

А. А. Харкевич

ЛИНЕЙНЫЕ
И НЕЛИНЕЙНЫЕ
СИСТЕМЫ

А. А. Харкевич
ИЗБРАННЫЕ ТРУДЫ

2

АКАДЕМИЯ НАУК СССР

**ИНСТИТУТ
ПРОБЛЕМ ПЕРЕДАЧИ
ИНФОРМАЦИИ**

А. А. Харкевич

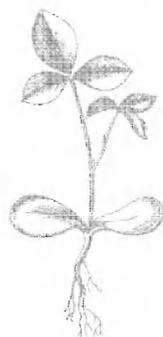
**ИЗБРАННЫЕ
ТРУДЫ**

в трех томах

2



ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА
1973



Scan AAW

А.А. Харкевич

ЛИНЕЙНЫЕ
И НЕЛИНЕЙНЫЕ
СИСТЕМЫ



ИЗДАТЕЛЬСТВО «НАУКА»

МОСКВА

1973

Линейные и нелинейные системы. Харкевич А. А. Избранные труды в трех томах. М., «Наука», 1973 г., стр. 566.

Во втором томе помещены работы, посвященные анализу линейных систем (определению их частотных характеристик, оценке искажений на основе квадратичного критерия и т. д.). Сюда же включены обширные труды в области нелинейных и параметрических явлений в радиотехнике и в области теории и практики спектрального анализа колебаний.

Таблиц 7. Иллюстраций 460. Библиогр. 126 назв.

**КОМИССИЯ ПО ИЗДАНИЮ ТРУДОВ
АКАДЕМИКА А. А. ХАРКЕВИЧА:**

Э. Л. БЛОХ, Е. М. ВЛАСОВА, В. А. ГАРМАШ,
А. Ю. ИШЛИНСКИЙ (председатель), В. М. НАЗАРОВ,
В. И. НЕЙМАН, И. А. ОВСЕВИЧ, Л. Д. РОЗЕНБЕРГ,
И. Г. РУСАКОВ, В. И. СИФОРОВ, Б. С. ЦЫБАКОВ

О ДЕТЕКТИРОВАНИИ БИЕНИЙ

1. Современный генератор звуковой частоты должен удовлетворять следующим основным требованиям: а) управление частотой должно быть возможно простым, б) мощность на выходе не должна зависеть от частоты и в) получаемый от генератора ток звуковой частоты должен быть практически свободен от гармоник. В наибольшей мере всем этим требованиям может удовлетворить генератор, построенный по принципу биений, получающихся как результат сложения двух колебаний высокой частоты. Но если выполнение первых двух условий вытекает непосредственно из принципиальных особенностей генератора, то для выполнения третьего необходимо принятие специальных мер. Гармоники звуковой частоты могут появиться либо за счет наличия гармоник в обоих колебаниях высокой частоты, либо благодаря перегрузке (т. е. переходу на нелинейные части характеристик) усилительных ламп и трансформаторов, либо, наконец, вследствие неправильного детектирования.

Детектирование должно быть таково, чтобы в результате его состав тока в цепи детектора определялся не более как тремя слагаемыми:

1) чисто синусоидальный ток звуковой частоты, 2) постоянная составляющая и 3) составляющие высокой частоты любого порядка и в любом количестве. Постоянная составляющая задерживается соответствующими переходными органами; составляющие высокой частоты устраняются ограничивающим сверху фильтром.

2. Иностранными авторами указывался следующий путь к разрешению задачи детектирования биений: детекторная характеристика берется прямолинейной, а амплитуды обоих колебаний высокой частоты неравными, причем отношение их доходит до 30.

Сущность этого метода выясняется из следующего его рассмотрения. Пусть даны напряжения $A_1 \sin \omega_1 t$ и $A_2 \sin \omega_2 t$, подводимые к детектору. В результате сложения имеем

$$V = \sqrt{A_1^2 + A_2^2 + 2A_1A_2 \cos(\omega_1 - \omega_2)t} \sin\left(\frac{\omega_1 + \omega_2}{2}t + \varphi\right), \quad (1)$$

т. е. переменное напряжение высокой частоты $(\omega_1 + \omega_2)/2$ с переменной фазой φ и с амплитудой, меняющейся со звуковой частотой.

той по закону, выражаемому радикалом. Если предположить, что детекторная характеристика задана уравнениями:

$$\begin{aligned} \text{при } -\infty < V < 0 \text{ имеем } I=0, \\ \text{при } 0 < V < +\infty \text{ имеем } I=sV, \end{aligned}$$

т. е. линейная зависимость между током и напряжением имеет силу только для положительных значений напряжения (отсчитываемых от некоторого напряжения, условно принимаемого за 0), то такое предположение равносильно отбрасыванию всех отрицательных значений функции $f = \sin\left(\frac{\omega_1 + \omega_2}{2}t + \varphi\right)$, а следовательно, нарушению непрерывности ее.

Поэтому в самом общем виде можно представить ток в цепи детектора формулой

$$I = s \sqrt{A_1^2 + A_2^2 + 2A_1A_2 \cos(\omega_1 - \omega_2)t} \times \left[C_0 + \sum_1^h C_k \sin k\left(\frac{\omega_1 + \omega_2}{2}t + \varphi_k\right) \right]. \quad (2)$$

Из этой формулы совершенно очевидно, что составляющая звуковой частоты будет содержаться только в произведении радикала на первый член в скобках, т. е. на постоянную составляющую периодической функции высокой частоты, все же остальные произведения дадут ряд составляющих высокой частоты.

Отсюда же хорошо уясняется вообще смысл и необходимость детектирования для извлечения тока низкой частоты из сложного модулированного тока: если модулированный ток представляет собою произведение некоторых двух периодических функций низкой и высокой частоты, то первая объективно существует (т. е. существует, как слагаемое, и может быть отделена соответствующими фильтрующими устройствами) только при наличии в последней постоянной составляющей, которая и появляется после детектирования.

Вводим для упрощения письма обозначения

$$\begin{aligned} A &= \sqrt{p} \sqrt{1 + m \cos x} = \sqrt{A_1^2 + A_2^2 + 2A_1A_2 \cos(\omega_1 - \omega_2)t}; \\ m &= \frac{2A_1A_2}{A_1^2 + A_2^2}; \quad p = A_1^2 + A_2^2; \quad x = (\omega_1 - \omega_2)t. \end{aligned}$$

Отбрасывая составляющие высокой частоты, перепишем выражение (2) для тока в виде

$$I = sC_0 p^{1/2} (1 + m \cos x)^{1/2}. \quad (3)$$

Для рассматриваемого случая функция высокой частоты может быть выражена уравнениями

$$0 < \left[\left(\frac{\omega_1 + \omega_2}{2} \right) t + \varphi \right] < \pi, \quad f = \sin \left[\left(\frac{\omega_1 + \omega_2}{2} \right) t + \varphi \right],$$

$$\pi < \left[\left(\frac{\omega_1 + \omega_2}{2} \right) t + \varphi \right] < 2\pi, \quad f = 0.$$

Постоянная составляющая такой функции, как показывает разложение, равна $1/\pi$.

Таким образом, окончательно

$$I = \frac{1}{\pi} S p^{1/2} (1 + m \cos x)^{1/2} = \frac{1}{\pi} SA. \quad (4)$$

Для суждения о составе тока найдем разложение в ряд Фурье функции A , для чего представим ее в виде ряда, пользуясь формулой бинома Ньютона

$$A = p^{1/2} (1 + m \cos x/2 - m^2 \cos^2 x/8 + m^3 \cos^3 x/16 - 5m^4 \cos^4 x/128). \quad (5)$$

Так как функция симметрична относительно средней ординаты, то $a_{k=2n-1} = b_{k=2n} = 0$, а потому амплитуды гармоник будут просто

$$C_1 = \frac{1}{\pi} \int_{-\pi}^{+\pi} A \sin x dx, \quad C_2 = \frac{1}{\pi} \int_{-\pi}^{+\pi} A \cos 2x dx,$$

$$C_3 = \frac{1}{\pi} \int_{-\pi}^{+\pi} A \sin 3x dx. \quad (6)$$

Подставляя в (6) значение A из (5) и интегрируя соответствующие ряды, получим следующие значения C_k :

$$C_1 = p^{1/2} \left(\frac{1}{2} m + \frac{3}{64} m^3 + \frac{35}{2048} m^5 + \dots \right),$$

$$C_2 = p^{1/2} \left(\frac{1}{16} m^2 + \frac{5}{256} m^4 + \frac{21}{2048} m^6 + \dots \right),$$

.

Для частного случая $A_1 = A_2 = 1$ можно найти разложение совершенно точно, воспользовавшись известным тригонометрическим соотношением

$$-\pi < x < +\pi, \quad +\sqrt{2} \sqrt{1 + \cos x} = 2 \cos \frac{1}{2} x.$$

Разложение дает для этой функции

$$C_1 = \frac{8}{3\pi}, \quad C_2 = \frac{3}{15\pi}, \quad C_k = \frac{2}{\pi} \cdot \frac{1}{k^2 - 1/4}.$$

Разложение представлено графически на рис. 1.

Пользуясь уравнениями (7), можно найти зависимость относительной величины, например, второй гармоники от отношения A_1/A_2

$$\frac{C_2}{C_1} = \frac{\frac{1}{16} m^2 + \frac{5}{256} m^4 + \frac{21}{2048} m^6 + \dots}{\frac{1}{2} m + \frac{3}{64} m^3 + \frac{35}{2048} m^5 + \dots}$$

Очевидно, что отношение это стремится к нулю, когда m стремится к нулю, так как

$$A = \sqrt{A_1^2 + A_2^2 + A_1 A_2 \cos x},$$

числитель дроби есть полином степени на единицу выше знаменателя.

Но

$$m = \frac{2A_1 A_2}{A_1^2 + A_2^2} = \frac{2}{A_1/A_2 + A_2/A_1}.$$

Это отношение стремится к нулю, когда любое из отношений A_1/A_2 или A_2/A_1 стремится к бесконечности. При этом обратное отношение обращается в нуль. Таким образом, действительно, при увеличении одной из амплитуд высокой частоты по отноше-

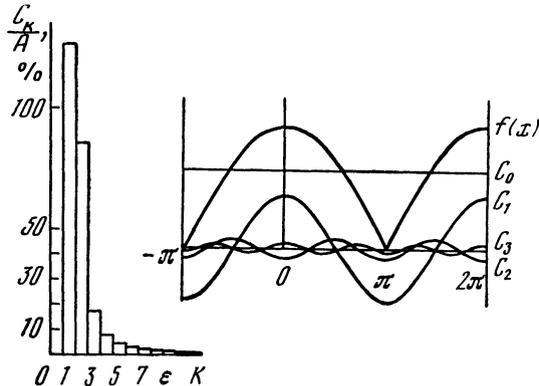
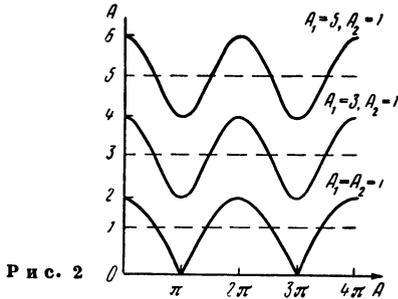


Рис. 1.

нию к другой относительная величина гармоник убывает, т. е. кривая приближается к синусоиде. Это иллюстрируется рис. 2, на котором представлены графики функции A при $A_1/A_2=1$, 3 и 5. По этим кривым легко проследить постепенное закругление $\text{minimum}'\text{ов}$ и заострение $\text{maximum}'\text{ов}$ по мере возрастания величины A_1/A_2 . Результаты вычисления относительной величины второй гармоники, т. е. C_2/C_1 в функции A_1/A_2 показывают, что отношение C_2/C_1 , равное 20% при $A_1=A_2$, падает довольно быстро, достигая величин $< 1\%$ при A_1/A_2 порядка 25—30.

3. Нетрудно усмотреть основной принципиальный недостаток описанного метода: при любых значениях A_1/A_2 все гармоники неизбежно будут содержаться в кривой тока звуковой частоты. Возможен метод детектирования, не дающий никаких гармоник звуковой частоты. Суть метода заключается в том, что детектирование совершается по квадратичному закону.



Р и с. 2

Пусть детекторная характеристика есть парабола вида

$$I = a + bV + kV^2. \quad (8)$$

Полагая по-прежнему (см. (1))

$$V = A_1 \sin \omega_1 t + A_2 \sin \omega_2 t \quad (9)$$

и подставляя (9) в (8), получим

$$\begin{aligned} I &= a + bA_1 \sin \omega_1 t + bA_2 \sin \omega_2 t + kA_1^2 \sin^2 \omega_1 t + \\ &+ kA_2^2 \sin^2 \omega_2 t + 2kA_1A_2 \sin \omega_1 t \sin \omega_2 t = \\ &= \left[a + k \frac{A_1^2 + A_2^2}{2} \right] + \left[b(A_1 \sin \omega_1 t + A_2 \sin \omega_2 t) - \right. \\ &- \left. \frac{1}{2} k(A_1^2 \cos 2\omega_1 t + A_2^2 \cos 2\omega_2 t + 2A_1A_2 \cos(\omega_1 + \omega_2)t) \right] + \\ &+ [kA_1A_2 \cos(\omega_1 - \omega_2)t]. \end{aligned}$$

В этом выражении первый член есть постоянная составляющая, второй содержит члены высокой частоты, а третий представляет собою чистую косинусоиду звуковой частоты, свободную от каких бы то ни было гармоник.

Таким образом, описанный метод детектирования является принципиально совершенным. Дальнейшей задачей является выяснение условий его практического применения.

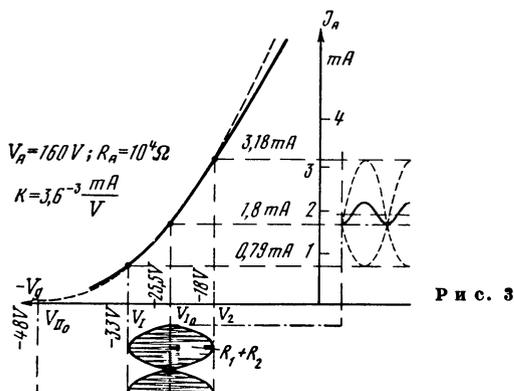
4. На практике осуществление метода квадратичного детектирования приводит к отысканию на реальной характеристике участка, с достаточной степенью приближения могущего быть представленным параболическим законом (8).

Уравнение (8) может быть представлено в виде

$$I = I_0 + k(V_0 + V)^2. \quad (10)$$

Связь с коэффициентами уравнения (9) дается соотношениями $a = I_0 + kV_0^2$, $b = 2kV_0$.

По графическому смыслу I_0V_0 суть координаты вершины параболы. Графический анализ реальной характеристики с целью отыскания на ней участка, приближающегося к общему закону (9) или (10), весьма затруднителен; задача значительно упрощается, если ограничить выбор семейством парабол, вершины которых



Р и с. 3

лежат на оси абсцисс, т. е. $I_0 = 0$. Тогда уравнение (10) перепишется в виде

$$I = k(V_0 + V)^2. \quad (11)$$

Заданная в виде $I = f(V)$ характеристика перестраивается в виде $\sqrt{I} = f_1(V)$, на полученной кривой выбирается прямолинейный участок, ограниченный значениями, например, V_1 и V_2 . Продолжая участок до пересечения с осью абсцисс в некоторой точке V_0 , можем полагать, что выбранный участок в принятом нами масштабе выражается уравнением

$$\sqrt{I} = \sqrt{k}(V_0 + V).$$

В обычном масштабе по возведению обеих частей уравнения в квадрат получим параболу (11), которой принадлежит выбранный участок.

Построение показано на рис. 3 для лампы УТ-1, для которой сняты динамические характеристики $I_a = f(V_g)$ при различных R_a . Там же указаны значения k , который представляет собою, так сказать, квадратичную крутизну и выражается в mA/V^2 .

5. Выбрав рабочий участок характеристики, следует задать режим работы детектора таким образом, чтобы мгновенные значения напряжения не выходили за пределы этого участка. Напряжение на детекторе может меняться (см. (9)) от $+(A_1+A_2)$ до $-(A_1+A_2)$, т. е. всего на $2(A_1+A_2)$. Таким образом, выбор амплитуд напряжения высокой частоты ограничен условием

$$2(A_1 + A_2) \leq V_2 - V_1. \quad (12)$$

С другой стороны, ранее было получено, что амплитуда тока звуковой частоты пропорциональна произведению A_1A_2 . Но про-

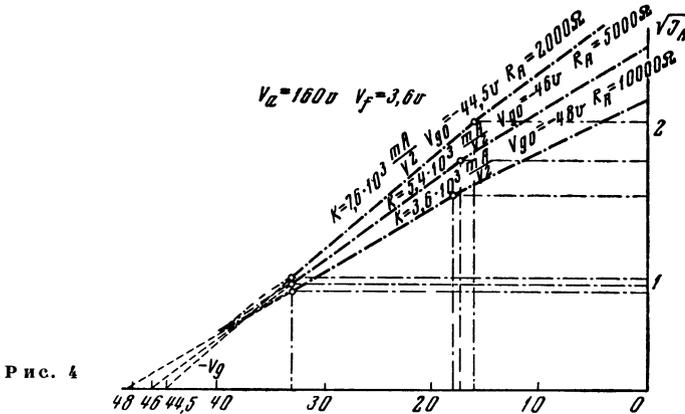


Рис. 4

изведение двух величин, сумма которых постоянна, имеет максимум при равенстве этих величин. Отсюда и из (12) может быть получено практическое правило для выбора наивыгоднейших (т. е. наибольших) амплитуд напряжения

$$A_1 = A_2 = \frac{1}{4}(V_2 - V_1). \quad (13)$$

6. Возвращаясь к выражению для тока в цепи квадратичного детектора, перепишем его (за вычетом членов высокой частоты) в виде

$$I = k \left[\left(V_0^2 + \frac{A_1^2 A_2^2}{2} \right) + A_1 A_2 \cos(\omega_1 - \omega_2) t \right]. \quad (14)$$

Пользуясь соотношением (13), можно выразить ток через параметры характеристики, т. е. через величины V_0 , V_1 и V_2 следующей формулой:

$$I = k \left[\left(V_0^2 + \frac{V_2 + V_1}{16} \right)^2 + \frac{(V_2 + V_1)^2}{16} \cos(\omega_1 - \omega_2) t \right]^*.$$

* Интересно заметить, что амплитуда тока звуковой частоты не зависит от V_0 , а следовательно, и от линейной крутизны характеристики на данном участке.

Пример* (см. рис. 4). Детектор: лампа УТ-1, $V_a = 160$ в, $R_a = 10^4$ ом.

Характеристика: $k = 3,6 \cdot 10^{-3}$ мА/В². $V_0 = -48$, $V_1 = -33$, $V_2 = -18$ в**. $A_1 = A_2 = 15/4 = 3,75$ в.

$$I = 3,6 \cdot 10^{-3} \left[\left(22,5^2 + \frac{15^2}{15} \right) + \frac{15^2}{16} \cos(\omega_1 - \omega_2)t \right] \cong 1,87 + 0,05 \cos(\omega_1 - \omega_2)t.$$

Амплитуда переменного напряжения, снимаемого с сопротивления в анодной цепи, равна $0,05 \cdot 10^{-3} \cdot 10^4 \cong 0,5$ в.

* На рис. 4 получение синусоидального тока в анодной цепи пояснено построением верхней и нижней огибающих (штриховые кривые) и синусоиды, получаемой в результате поординатного вычитания огибающих (сплошная кривая).

** Ось ординат предполагается проходящей через середину рабочего участка характеристики; в рассматриваемом примере абсцисса этой точки равна $-25,5$ в; эта величина должна быть вычтена из приведенных значений V_0 , V_1 и V_2 .

О СПЕЦИАЛЬНОЙ ЧАСТОТНОЙ ХАРАКТЕРИСТИКЕ ДЛЯ ОЦЕНКИ ИСКАЖЕННОЙ ФОРМЫ КРИВОЙ

При передаче тех или иных сигналов и вообще при воспроизведении различных функций времени важно знать, как искажает воспроизводящая система воздействующую на нее функцию. Для оценки этих искажений естественно обратиться непосредственно к рассмотрению самих функций времени. Однако в слаботочной технике укоренилось обыкновение представлять явления спектрами, а свойства передающих систем — частотными характеристиками их. Иначе говоря, функции времени на основе понятий ряда и интеграла Фурье заменяются функциями частоты.

Когда дело касается искажений формы кривой, то говорят, что для оценки этих искажений недостаточно знать частотную характеристику системы, но требуется знать еще и фазную характеристику. Если при этом подразумевается, что определенной частотной характеристике могут соответствовать различные фазные характеристики, то такое мнение ошибочно, так как частотная характеристика определяет систему полностью. Несомненно лишь одно, что обычная частотная характеристика неудобна для оценки искажений формы кривой, так как для нахождения последних требуются дополнительные математические операции. Поэтому может представить интерес попытка построения такой специальной частотной характеристики, которая, не выходя из круга привычных частотных представлений и соответствующей экспериментальной методики, позволила бы непосредственно судить об искажениях формы кривой в той системе, которая представлена этой характеристикой.

Теоретические соображения. Если поведение x данной системы под воздействием $F(t)$ описывается дифференциальным уравнением

$$\dot{H}(p)x = F(t), \quad (1)$$

где $H(p)$ — символическая форма дифференциального оператора n -го порядка, то частотная характеристика данной системы выражается как

$$f(\omega) = |1/H(j\omega)|. \quad (2)$$

Функция $x(t)$, являющаяся решением уравнения (1), вообще говоря, отличается от функции $F(t)$, стоящей в правой части, что и имеется в виду, когда говорится об искажении формы кривой. Будем в дальнейшем для краткости называть функцию $F(t)$

сигналом, а функцию $x(t)$ — записью. Кроме того, введем понятие неискаженной записи, определяемой как

$$x_0 = F/q, \quad (3)$$

где q — не зависящий от p член оператора $H(p)$. Если последний не содержит отрицательных степеней p , то в этом частном случае $q = H(0)$. Заметим попутно, что система, уравнение которой не содержит не зависящего от p члена, вообще непригодна для представления сигнала F записью x .

Далее воспользуемся понятием погрешности, определяемой как

$$\delta = x_0 - x^*. \quad (4)$$

Теперь нужно выбрать критерий для оценки искажения формы сигнала. Для этого естественно, как это было указано Н. Н. Андреевым, воспользоваться понятием квадратичной погрешности. Мы выберем в качестве мерил погрешности суммарную квадратичную погрешность, выражаемую интегралом

$$S^2 = \int_0^{\infty} [\delta(t)]^2 dt. \quad (5)$$

Само собою разумеется, что этот критерий пригоден только для случая, когда интеграл (5) сходится.

Пусть теперь функция $\delta(t)$ разложена в интеграл Фурье, т. е. пусть

$$\delta(t) = \frac{1}{\pi} \int_0^{\infty} [A(\omega) \cos \omega t + B(\omega) \sin \omega t] d\omega,$$

где, как обычно,

$$A(\omega) = \int_0^{\infty} \delta(t) \cos \omega t dt, \quad B(\omega) = \int_0^{\infty} \delta(t) \sin \omega t dt.$$

Воспользуемся теоремой Рэйли (Rayleigh), на основании которой [1]

$$S^2 = \int_0^{\infty} [\delta(t)]^2 dt = \frac{1}{\pi} \int_0^{\infty} \{[A(\omega)]^2 + [B(\omega)]^2\} d\omega. \quad (6)$$

Из этого соотношения мы и делаем вывод, положенный в основу всего дальнейшего: функция частоты, стоящая под знаком интеграла в правой части, может рассматриваться как особого вида частотная характеристика данной системы, позволяющая

* Предполагается, что функции F , x и δ тождественно равны нулю при $t < 0$.

непосредственно судить об искажении формы сигнала с точки зрения критерия (5)*.

Теперь нам нужно установить связь между уравнением системы и видом функции частоты в выражении (6). Здесь надо прежде всего заметить, что критерий (5) зависит от вида функции F и, следовательно, является неопределенным. Вопрос получает необходимую определенность, если выбрать раз навсегда некоторую функцию времени в качестве стандартного испытательного сигнала. Такой выбор неизбежно содержит некоторый элемент произвола. Не приводя здесь никаких соображений для обоснования нашего выбора, примем в качестве стандартного сигнала так называемую единичную функцию I , определяемую условиями

$$\begin{aligned} I &= 0 \text{ при } t < 0, \\ I &= 1 \text{ при } t > 0. \end{aligned}$$

При воздействии сигнала такой формы получается, как известно, запись в виде так называемой переходной функции $h(t)$, определяемой операционным уравнением

$$h = \frac{1}{H(p)} I. \quad (7)$$

Таким образом, в этом случае

$$\delta(t) = \frac{1}{q} I - h(t) \quad (8)$$

(для $t > 0$).

Мы должны теперь найти разложение правой части (8) в интеграл Фурье. С одной стороны, известно, что I может быть выражена через интегральный синус [2]

$$I = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\sin \omega t}{\omega} d\omega. \quad (9)$$

С другой стороны, функция $h(t)$ разлагается по формуле [3]

$$h(t) = \frac{1}{2} a(0) + \frac{1}{\pi} \left[\int_0^{\infty} \frac{a(\omega)}{\omega} \sin \omega t d\omega + \int_0^{\infty} \frac{b(\omega)}{\omega} \cos \omega t d\omega \right]. \quad (10)$$

Здесь

$$a + jb = 1/H(j\omega). \quad (11)$$

* Мы перешли от функций времени к функциям частоты. Так как последние рассматриваются в конечном интервале частот, то вопрос о сходимости интеграла в правой части (6) практически отпадает. Вопрос о том, как при рассмотрении функций времени поступать с интегралом (5) в том случае, когда он расходится, будет обсужден в другом месте.

Вычитая (10) из (9) и подставляя в (8), получаем

$$\delta(t) = \frac{1}{2} \left[\frac{1}{q} - a(0) \right] + \frac{1}{\pi} \left[\int_0^{\infty} \frac{1/q - a(\omega)}{\omega} \sin \omega t d\omega - \int_0^{\infty} \frac{b(\omega)}{\omega} \cos \omega t d\omega \right]. \quad (12)$$

Первый член, не зависящий от ω , часто бывает равен нулю; мы будем для простоты иметь в дальнейшем в виду именно этот случай.

К тому же результату мы пришли и другим путем; именно, если

$$h = \frac{1}{H(p)} I,$$

то

$$\frac{1}{q} - h = \left(\frac{1}{q} - \frac{1}{H(p)} \right) I = \frac{1}{G(p)} I.$$

Принимая во внимание (11), можем сразу записать разложения $1/G(j\omega)$ на вещественную и мнимую часть

$$\frac{1}{G(j\omega)} = \frac{1}{q} - \frac{1}{H(j\omega)} = \left[\frac{1}{q} - a \right] - j[b],$$

после чего, определив вспомогательную функцию g операционным уравнением

$$g = \frac{1}{G(p)} I$$

и подставляя g и ее компоненты в (10), получаем результат (12). Теперь мы можем записать (5) в явной форме

$$S^2 = \int_0^{\infty} [\delta(t)]^2 dt = \int_0^{\infty} \frac{(1-a)^2 + b^2}{\omega^2} d\omega = \int_0^{\infty} \frac{d\omega}{[\omega |G(j\omega)|]^2}. \quad (13)$$

Функция $1/\omega |G(j\omega)|$, единственным образом определяющая суммарную погрешность, может квалифицироваться как спектр абсолютной погрешности.

Множитель $1/\omega$ указывает на то, что независимо от вида функций H и G , спектр погрешности убывает с частотой по гиперболическому закону. Это происходит оттого, что и спектр выбранного нами сигнала имеет такую же форму. Введя отношение спектра абсолютной погрешности к спектру сигнала, мы получим функцию $1/|G(j\omega)|$, которую можно назвать относительным спектром погрешности. Эта функция и предлагается в качестве единой частотной характеристики.

Желательно иметь возможность суждения о суммарной абсолютной погрешности.

Если принять во внимание, что амплитуда каждого из частичных колебаний, суммируемых интегралом Фурье, представляется произведением спектра на дифференциал частоты, то можем написать

$$\frac{d\omega}{\omega |G(j\omega)|} = \frac{1}{|G(j\omega)|} d(\lg \omega). \quad (14)$$

Последняя форма весьма удобна для практического применения: на основании (14) получаем, что мерой абсолютной погрешности в указанном смысле может служить площадь, ограниченная осями и кривой $1/|G(j\omega)|$ при логарифмическом масштабе по оси абсцисс.

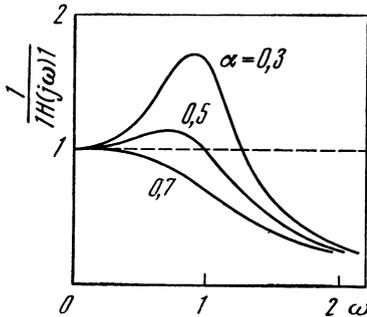


Рис. 1

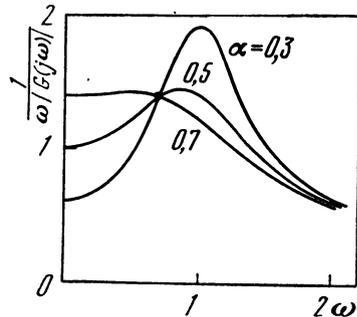


Рис. 2

Геометрический смысл преобразования (14) состоит в том, что убывание ординат заменяется соответствующим сжатием масштаба абсцисс.

Величина $1/|G(j\omega)|$ геометрически представляется модулем разности вектора неискаженной записи, направленного по вещественной оси — вектора сравнения и вектора записи $1/H(j\omega)$.

Рассмотрим пример. Пусть дано

$$H(p) = mp^2 + rp + s, \quad (15)$$

что соответствует, например, механической системе с одной степенью свободы.

Пользуясь обычными обозначениями $\omega_0 = \sqrt{s/m}$, $\alpha = r/2m$, имеем

$$\begin{aligned} \frac{1}{H(j\omega)} &= \frac{1}{s} \cdot \frac{1}{1 - \omega^2/\omega_0^2 + j2\alpha\omega}, \\ a &= \frac{1}{s} \cdot \frac{1 - \omega^2/\omega_0^2}{(1 - \omega^2/\omega_0^2)^2 + (2\alpha\omega)^2}, \\ b &= -\frac{1}{s} \cdot \frac{2\alpha\omega}{(1 - \omega^2/\omega_0^2)^2 + (2\alpha\omega)^2}. \end{aligned}$$

Обычная частотная характеристика

$$\frac{1}{|H(j\omega)|} = \frac{1}{s} \cdot \frac{1}{\sqrt{(1 - \omega^2/\omega_0^2)^2 + (2\alpha\omega)^2}} \quad (16)$$

Новая частотная характеристика

$$\frac{1}{|G(j\omega)|} = \frac{1}{s} \sqrt{\frac{\omega^4/\omega_0^4 + 4\alpha^2\omega^2}{(1 - \omega^2/\omega_0^2)^2 + (2\alpha\omega)^2}} \quad (17)$$

На рис. 1 построены графики величины $1/|H(j\omega)|$ по формуле (16), т. е. обычные резонансные кривые.

На рис. 2 построены спектры абсолютной погрешности, т. е. значения $1/\omega |G(j\omega)|$; на рис. 3 построены относительные спек-

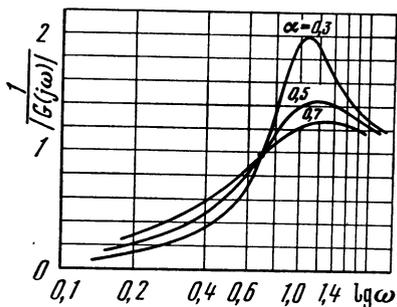


Рис. 3

тры погрешности, т. е. значения $1/|G(j\omega)|$ по формуле (17) при логарифмическом масштабе по оси абсцисс. На всех чертежах принято $s = \omega_0 = 1$; $\alpha = 0,3; 0,5; 0,7$.

Замечания о технике измерений. Снятие предлагаемой частотной характеристики сводится к регистрации зависимости от частоты модуля разности векторов интересующей нас величины на входе и на выходе четырехполюсника, как это показано, например, для напряжений на схеме рис. 4.

Возможно применение компенсационных методов измерения; предпочтительны, однако, прямые методы, обеспечивающие возможность простой автоматизации.

Действительная схема может быть устроена, как показано на рис. 5. Подобная схема применялась Гютцмахером [4] для измерения суммы, а не разности напряжений, что, впрочем, не меняет дела. Интересно отметить, что Гютцмахер измерял по разности оба вектора и их сумму для получения треугольника напряжений. Зависимость суммы векторов от частоты Гютцмахер условно называет фазной характеристикой (Phasenkurve). Точно такая же схема применена в ламповом фазометре, описанном Крейпером [5]. Съемка характеристик может быть автоматизирована при условии обеспечения постоянства режима.

Для демонстрации этой возможности автором применена весьма простая схема, подобная схеме лампового фазометра Траутвейна

[6]. Эта схема изображена на рис. 6. Действие ее состоит в том, что при $V_2 = \mu V_1$ действие анода компенсирует действие сетки, вследствие чего ток приближается к нулю (на этом основано общеизвестное применение этой схемы для измерения коэффициента усиления). При увеличении переменного напряжения на аноде постоянный ток в цепи катода возрастает за счет работы

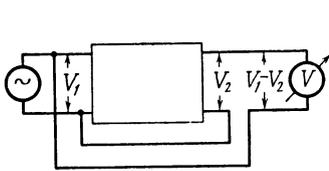


Рис. 4

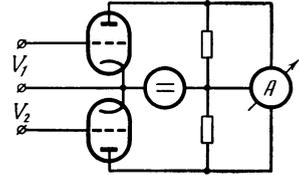


Рис. 5

диола катод—анод, а при увеличении напряжения на сетке — за счет работы диода катод—сетка. Таким образом можно использовать включенную по подобной схеме лампу для измерения вектора разности приведенных напряжений на аноде и на сетке. Градуировка лампы производилась по схеме рис. 7. Данные схемы таковы: лампа УО-104, частота 500 $\mu\text{ц}$, $R_a = 5000$ ом , напряжение на нем 40 V . Зависимость тока в цепи катода от сопротивления R_g представлена рис. 8 ($1 - R_0 = 0$; $2 - R_0 = 3000$; $3 - R_0 = 5000$ ом). При изменении знака разности напряжений на сетке и на аноде градуировка определяется двумя различными ветвями этой характеристики. Разница по крутизне и форме обеих ветвей хорошо видна, если перегнуть чертеж по ординате, соответствующей минимуму, т. е. нулевой разности приведенных напряжений. На рис. 8 результат такого переноса левых ветвей показан штрихом. Вариацией сопротивления в аноде (R_0) можно до некоторой степени уменьшить расхождение кривых в известных

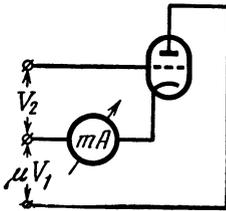


Рис. 6

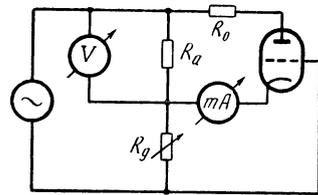


Рис. 7

пределах измеряемых напряжений, как это видно из сопоставления кривых, относящихся к различным R_0 , на рис. 8. Метрические качества схемы в общем довольно сомнительны, однако для демонстрационного опыта она вполне пригодна¹.

В качестве объекта испытания была выбрана схема рис. 9. Эта интересная схема обладает тем свойством, что если принять

¹ Надо заметить, что эта схема могла бы дать прекрасные результаты при применении очень простой специальной трехэлектродной лампы с двумя симметричными анодами по обе стороны катода.

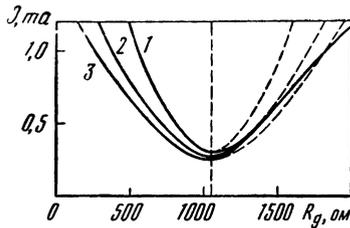


Рис. 8

за воздействие питающий схему ток, а поведение схемы определять напряжением на конденсаторе, то оператор схемы принимает форму (15), в которой нужно только, как обычно, заменить t , r и s соответственно на L , R и $1/C^*$ [7].

Таким образом свойства этой схемы описываются приведенными выше формулами (16) и (17). Полная схема опыта представлена на рис. 10, на котором означают: Γ — гетеродинный генератор, X — автоматический характеристограф, Π — переключатель. При изображенном на чертеже положении переключателя снимается характеристика $1/G |j\omega|$, при втором положении — характеристика $1/H(j\omega)$.

Данные схемы: $R_i = 15 \cdot 10^3$, $R = 1,1 \cdot 10^3$ ом, $L = 0,4$ H, $C = 1$ или $0,25$ мкф, $R_a = 5 \cdot 10^3$, $R_0 = 3 \cdot 10^3$ ом. Результат опыта показан на снимках рис. 11 и 12. На рис. 11 сняты характеристики $1/G(j\omega)$; кривая a снята при $C = 0,25$ мкф, кривая b — при $C = 1$ мкф. Кривая c представляет собою величину напряжений на R_a (в уменьшенном масштабе) и снята для проверки постоянного режима. Надо заметить, что кривая c снята при тех же данных, что и кривая a , т. е. при больших вариациях напряжения.

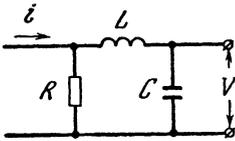


Рис. 9

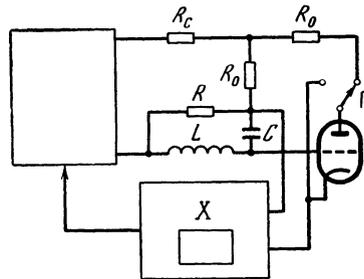
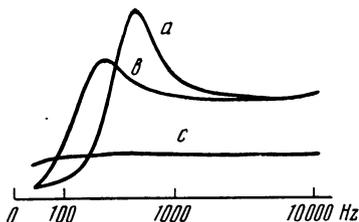


Рис. 10

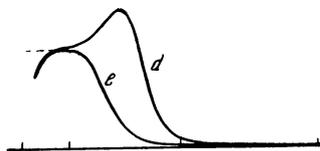
На рис. 12 сняты характеристики $1/H(j\omega)$; кривые d и e соответствуют значениям емкости $0,25$ и 1 мкф.

Провал около нуля обусловлен нижней и пограничной частотой усилителя генератора; резкое падение выходного напряжения наблюдается около 10 гц. Истинный ход кривых намечен на рис. 12 штрихом.

* Эта схема применяется для коррекции влияния сеточной емкости в усилителях.



Р и с. 11



Р и с. 12

Снимок рис. 11 соответствует теоретически характеристикам рис. 3, а снимок рис. 12 — характеристикам рис. 1.

Не представит, по-видимому, никаких затруднений видоизменить соответственным образом снятие частотных характеристик любого четырехполюсника, будь то прибор электрический, электромеханический, электроакустический или даже чисто механический. Желательна лишь постройка специальных вольтметров, приспособленных для измерения векторной разности, подобных описанным выше.

* * *

Предполагаемая новая характеристика в принципиальном отношении обладает тем преимуществом, что, будучи единственной, дает количественное суждение об искажении формы кривой. В практическом же отношении следует признать, что снятие этой характеристики не только вполне возможно, но сопровождается лишь весьма незначительным усложнением обычной схемы для снятия частотных характеристик.

Возможность съемки приведенных выше экспериментальных характеристик любезно предоставлена мне акустической лабораторией ВНИИМ'а, за что я искренне признателен руководителю лаборатории Ю. Н. Егорову.

Л и т е р а т у р а

1. Н. Н. Андреев. Решетка, призма, резонатор. ЖРФХО, 17, 5, 1915, 1—33.
2. В. И. Смирнов. Курс высшей математики, т. 2. Гостехиздат, 1933.
3. Дж. Р. Карсон. Электрические нестационарные явления и операционное исчисление. ГНТИУ, 1934, стр. 199.
4. M. Gützmaeher. Durchlassbereich, Phasenlaufzeit und Klirrfaktor von Fernkabeln. ENT, 6, 1929, 386—395.
5. В. Л. Крейцер. Об одном способе измерения угла сдвига фаз в большом диапазоне частот. ИЭСТ, 1, 1935, 40—47.
6. А. Хунд. Hochfrequenzmesstechnik. Berlin, J. Springer, 1928, 235—236.
7. Г. В. Брауде. О коррекции частотных и фазных характеристик усилительных устройств. Журн. техн. физ., 4, 1934, 1714—1730, 1818—1828.

О ПРИМЕНЕНИИ КРИТЕРИЯ КВАДРАТИЧНОЙ ПОГРЕШНОСТИ К ОЦЕНКЕ ЛИНЕЙНЫХ ИСКАЖЕНИЙ

Для оценки искажений, вносимых линейными системами, предлагается критерий суммарной квадратичной погрешности, определяемой как

$$S_T = \int_0^T \delta^2 dt = \int_0^T (x - x_0)^2 dt,$$

где $x_0 = F/q$ — неискаженная запись сигнала F , x — фактическая запись; q — не зависящий от p ($= d/dt$) член в общем уравнении системы

$$H(p)x = F.$$

В качестве испытательного сигнала предлагается $F=1$ (единичная функция). При наличии в операторе $H(p)$ только положительных степеней p интеграл может браться в пределах от 0 до ∞ ; при наличии отрицательных степеней p верхний предел должен определяться из условий практического применения данной системы.

Общие соображения иллюстрируются примерами, относящимися к системе с одной степенью свободы.

Вопрос об искажениях весьма актуален во многих отраслях техники. Постановка этого вопроса своеобразна для каждой из этих отраслей. Однако можно выделить два основных подхода к проблеме искажений: в одной группе технических применений существенны искажения формы кривой (осциллография, телеграфия, сейсмометрия и др.), в другой же играют роль искажения амплитудного спектра (телефония и вообще электроакустика). Оставляя пока в стороне вопрос о лишь кажущемся различии в этих подходах, надо отметить также два методических приема, применяющихся при рассмотрении вопросов, связанных с искажениями.

Один прием состоит в прямом исследовании функций времени, представляющих собою первичное явление (сигнал) и его искаженное воспроизведение (запись). Второй прием состоит в рассмотрении частотной характеристики воспроизводящей системы, основанном на замене сигнала и записи как функций времени, соответствующими функциями частоты на основе понятия ряда

и интеграла Фурье. Оба эти приема одинаково законны, хотя второй не всегда быстрее ведет к цели.

Замечательно, однако, что при рассмотрении проблемы искажений посредством того или иного методического приема в большой литературе по этому вопросу зачастую чрезвычайно неясно говорится о том критерии, с точки зрения которого оцениваются искажения. Очевидно, всякие рассуждения на тему об искажениях и, в частности, об их уменьшении в значительной степени обесцениваются, если при этом не указан критерий, посредством которого производится оценка искажений, или если этот критерий не может быть непосредственно выведен из условий применения данной аппаратуры, или сопоставлен с ними.

Примеры различной трактовки проблемы искажений приведены в приложении.

При рассмотрении сигнала и его искаженной записи наиболее естественно, казалось бы, оценивать искажения с точки зрения весьма универсального критерия квадратичной погрешности. В настоящей заметке приводятся результаты применения этого критерия, относящиеся к некоторым простейшим примерам. Эти результаты могут оказаться полезными применительно к тем областям, где искажение формы кривой играет решающую роль.

Погрешность при записи единичной функции (без запаздывания). Выберем в качестве примера простейшую систему с одной степенью свободы, уравнение которой пригодно для описания многих применяемых в технике приборов, например, обычного осциллографа. Уравнение это гласит

$$(mp^2 + rp + s)x = F,$$

где $p = d/dt$; F — сигнал; x — запись (для механической системы физический смысл коэффициентов и переменных таков: m , r и s — соответственно масса, активное сопротивление и упругость; x — смещение, F — сила).

Уравнение (1) может быть переписано в обычных обозначениях

$$(p^2 + 2\alpha p + \omega_0^2)x = \frac{1}{m}F^*. \quad (1)$$

Здесь α — показатель затухания; ω_0 — собственная частота при отсутствии затухания.

Зададимся вопросом о наивыгоднейшем относительном затухании $k = \alpha/\omega_0^{**}$, т. е. о том значении этого параметра, при котором получается наименьшее искажение. Искажение будем оце-

* Надо заметить, что точно такое уравнение описывает поведение одной электрической схемы, применяемой для коррекции усилителей на высокой частоте.

** Величина, в 2π раз большая, носит название логарифмического декремента.

нить посредством квадратичной погрешности. Погрешность определим как

$$\delta = \frac{F}{s} - x = x_0 - x, \quad (2)$$

где $x_0 = F/s$ определяется как неискаженная запись. Квадрат средней квадратичной погрешности выразится как

$$\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \delta^2 dt.$$

Само собою разумеется, что величина этого интеграла зависит от выбора как T_1 и T_2 , так и функции F . Выбор последней определяется условиями эксплуатации данной аппаратуры; так, например, в технике высокого напряжения применяются испытательные импульсы вида $e^{-at} - e^{-bt}$. Я предпочту испытательный сигнал в форме единичной функции

$$I = \begin{cases} 0 & \text{при } t < 0, \\ 1 & \text{при } t \geq 0 \end{cases}$$

по соображениям, отчасти обоснованным ниже. Очевидно, при таком выборе интервал (3) при бесконечном верхнем пределе сходится, поэтому удобно взять его именно в пределах от 0 до ∞ , оперируя не со средней, а с суммарной квадратичной погрешностью, выражаемой следующим образом:

$$S = \int_0^{\infty} \delta^2 dt. \quad (4)$$

Задача состоит в нахождении минимума этого интеграла по параметру k . Решение уравнения (1) при $F=1$ и нулевых начальных условиях дает, как известно, так называемую переходную функцию h рассматриваемой системы

$$x(t) = h(t) = \frac{1}{s} \left[1 - e^{-\alpha t} \left(\frac{\alpha}{\omega'} \sin \omega' t + \cos \omega' t \right) \right], \quad (5)$$

где $\omega' = \sqrt{\omega_0^2 - \alpha^2}$ — собственная частота.

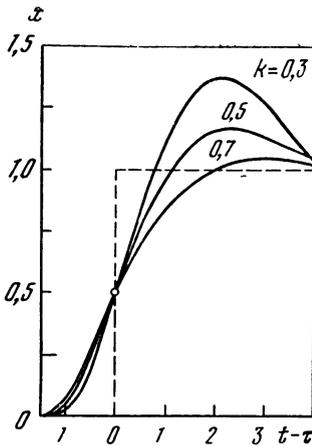
Графики функции (5) при различных значениях параметра k представлены на рис. 1, на котором, как и на всех последующих, приняты для упрощения численные значения $s=m=\omega_0=1$, что, понятно, не нарушает общности результатов.

Мгновенная погрешность равна

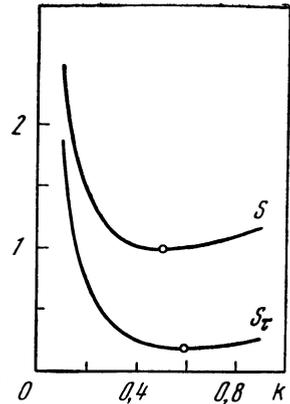
$$\delta = \frac{1}{s} e^{-\alpha t} \left(\frac{\alpha}{\omega'} \sin \omega' t + \cos \omega' t \right),$$

что, по подстановке в (4), дает

$$S = \frac{1}{s^2} \int_0^{\infty} e^{-2\alpha t} \left(\frac{\alpha}{\omega'} \sin \omega' t + \cos \omega' t \right)^2 dt. \quad (6)$$



Р и с. 1



Р и с. 2

Интеграл (6) легко берется; получаем

$$S = \frac{1}{s^2 \omega_0} \left(\frac{1}{4k} + k \right).$$

Эта функция имеет минимум при значении

$$k = 0,5. \quad (1)$$

Это и есть, таким образом, то значение относительного затухания $k = \alpha / \omega_0$, при котором получается наименьшая суммарная квадратичная погрешность при записи единичной функции. График функции (7) представлен на рис. 2.

Погрешность при записи единичной функции (с запаздыванием). Если запись отличается от сигнала лишь тем, что сдвинута по времени на определенный интервал, т. е. если

$$x(t) = \frac{1}{s} F(t - \tau),$$

то в известном смысле искажения отсутствуют, так как для совмещения приведенных к одинаковому масштабу кривых достаточно сдвинуть начало координат кривой x на величину τ влево. Примером такого положения вещей может служить сигнал, распространяющийся с конечной скоростью, но без изменения формы вдоль идеальной линии.

На основании этих соображений можно предъявить-воспроизводящей системе менее жесткие требования и оценивать вносимые ею искажения погрешностью, определяемой разностью соответственно сдвинутых во времени сигнала и записи, т. е. определить погрешность как

$$\delta_\tau = x_0(t - \tau) - x(t) = x_0(t') - x(t) \quad (8)$$

и соответственно суммарную квадратичную погрешность как

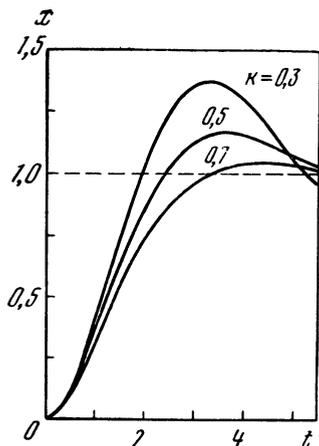
$$S_{\tau} = \int_0^{\infty} \delta_{\tau}^2 dt. \quad (9)$$

Прежде всего нужно выбрать величину τ . Естественно сделать этот выбор таким образом, чтобы искажения, определяемые интегралом (9), были наименьшими. Нужно, следовательно, найти минимум этого интеграла по τ . Разобьем интеграл на две части, имея в виду, что

$$x_0 = \frac{1}{s} I = \begin{cases} 0 & \text{при } t' < 0, \\ 1/s & \text{при } t' > 0, \end{cases} \quad (10)$$

$$S_{\tau} = \int_0^{\tau} x^2 dt + \int_{\tau}^{\infty} (1/s - x)^2 dt,$$

$$\frac{dS_{\tau}}{d\tau} = [x(\tau)]^2 - \left[\frac{1}{s} - x(\tau) \right]^2.$$



Р и с. 3

Приравнивая производную нулю, находим

$$x(\tau)_{\min} = 1/2s,$$

или, пользуясь (5),

$$e^{-\alpha\tau} \left(\frac{\alpha}{\omega'} \sin \omega'\tau + \cos \omega'\tau \right) = \frac{1}{2}. \quad (11)$$

Последнее соотношение представляет собою трансцендентное уравнение, из которого можно определить τ . Для решения пригоден, например, метод *regula falsi*. Применяя этот или какой-либо другой метод, находим следующие значения τ в зависимости от k (при $m=s=\omega_0=1$):

k	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
τ	1,094	1,135	1,181	1,234	1,292	1,358	1,430	1,503	1,586

На рис. 3 изображены графики функции x [по формуле (5)] те же, что и на рис. 1, но сдвинутые относительно сигнала на соответствующие τ .

Вернемся к интегралу (10). Представим его в виде

$$S_{\tau} = \int_0^{\tau} [x^2 - (1/s - x)^2] dt + \int_0^{\infty} (1/s - x)^2 dt.$$

Второй интеграл есть S (см. (7)). Таким образом,

$$S_\tau = \frac{1}{s^2\omega_0} \left(\frac{1}{4k} + k \right) + \frac{2}{s} \int_0^\tau x dt - \frac{1}{s^2} \tau.$$

Выполняя интегрирование и используя (11), находим окончательно

$$S_\tau = \frac{1}{s^2\omega_0} \left[\frac{1}{4k} - k + \omega_0\tau - \frac{2}{\sqrt{1-k^2}} e^{-k\omega_0\tau} \sin \sqrt{1-k^2} \omega_0\tau \right]. \quad (12)$$

График функции (12) изображен на рис. 2 совместно с графиком функции (7); из сопоставления кривых видно, что минимальные значения суммарной квадратичной погрешности без запаздывания и с запаздыванием различаются примерно в пять раз.

Для нахождения минимума функции (12) по k нужно составить и вычислить производную

$$\frac{dS_\tau}{dk} = \frac{\partial S_\tau}{\partial k} + \frac{\partial S_\tau}{\partial \tau} \cdot \frac{d\tau}{dk}.$$

Для вычисления производной $d\tau/dk$ послужит соотношение (11), из которого, пользуясь соответствующими правилами, получаем

$$\frac{d\tau}{dk} = \frac{1}{\omega_0 \sqrt{1-k^2}} (1 - \omega_0\tau \operatorname{ctg} \sqrt{1-k^2} \omega_0\tau).$$

Не приводя промежуточных выкладок, выпишем прямо выражение для искомой производной

$$\begin{aligned} \frac{dS_\tau}{dk} = \frac{1}{s^2\omega_0} \left\{ \frac{2}{\sqrt{1-k^2}} (1 - 2e^{-k\omega_0\tau} \cos \sqrt{1-k^2} \omega_0\tau) \times \right. \\ \times (1 - \omega_0\tau \operatorname{ctg} \sqrt{1-k^2} \omega_0\tau) - 1 - \frac{1}{4k^2} - \\ \left. - 2e^{-k\omega_0\tau} \left[\left(\frac{k}{(1-k^2)^{3/2}} - \frac{\omega_0\tau}{\sqrt{1-k^2}} \right) \sin \sqrt{1-k^2} \omega_0\tau - \right. \right. \\ \left. \left. - \frac{k\omega_0\tau}{1-k^2} \cos \sqrt{1-k^2} \omega_0\tau \right] \right\}. \end{aligned}$$

Приравнявая эту производную нулю, получаем трансцендентное уравнение, решая которое, находим

$$k = 0,587. \quad (II)$$

Это значение, будучи подставлено в (11), дает

$$\omega_0\tau = 1,349. \quad (III)$$

Погрешность при записи некоторых других функций (без запаздывания). Вычислим погрешности, получающиеся при записи некоторых функций, выбранных таким образом, что единичная функция является для них предельным случаем. Так как из уже проделанных вычислений видно, что случай, когда допускается запаздывание записи относительно сигнала, приводит

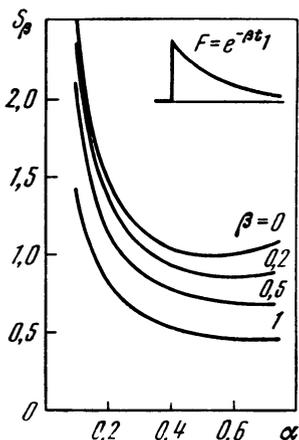


Рис. 4

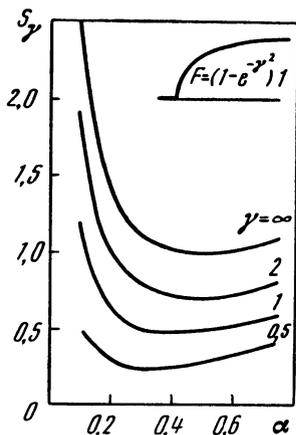


Рис. 5

к довольно громоздким вычислениям, ограничимся рассмотрением погрешности лишь при общем начале оси времен для сигнала и для записи, т. е. не допуская запаздывания последней. Кроме того, для упрощения формул будем сразу пользоваться выбранными для построения графиков численными значениями $s = m = \omega_0 = 1$.

В качестве первого примера возьмем функцию

$$F = e^{-\beta t} 1. \quad (13)$$

Запись такого сигнала системой с одной степенью свободы выражается как

$$x = \frac{1}{\beta^2 - 2\alpha\beta + 1} \left[e^{-\beta t} - e^{-\alpha t} \left(\frac{\alpha - \beta}{\omega'} \sin \omega' t + \cos \omega' t \right) \right]. \quad (14)$$

Составляя выражение для суммарной квадратичной погрешности по формуле (4) и выполняя действия, получаем

$$S_\beta = \frac{1}{(\beta^2 - 2\alpha\beta + 1)^2} \left[\frac{(2\alpha - \beta)^2 (2\alpha\beta + 1) + 1}{4\alpha} - \frac{4\alpha\beta (2\alpha - \beta)}{\beta^2 + 2\alpha\beta + 1} \right]. \quad (15)$$

При β , стремящемся к нулю, выражения (14) и (15) переходят в пределе соответственно в (5) и (7), что соответствует случаю сигнала в форме единичной функции. Графики функции (15) в зависимости от α при различных значениях β построены на рис. 4. Как видим, при увеличении β минимум функции смещается в сторону больших α .

Возьмем теперь сигнал в форме

$$F = (1 - e^{-\gamma t}) 1. \quad (16)$$

В этом случае

$$x = 1 - \frac{1}{\gamma^2 - 2\alpha\gamma + 1} \left\{ e^{-\gamma t} + \gamma e^{-\alpha t} \left[\frac{2\alpha^2 - \alpha\gamma - 1}{\omega'} \sin \omega' t + (2\alpha - \gamma) \cos \omega' t \right] \right\} \quad (17)$$

и суммарная квадратичная погрешность

$$S_{\gamma} = \frac{\gamma^2}{(\gamma^2 - 2\alpha\gamma + 1)^2} \left\{ \frac{(2\alpha^2 - \alpha\gamma - 1)^2 + (\gamma - 2\alpha)^2 \left(\alpha^2 + 2\frac{\alpha}{\gamma} + 1 \right)}{4\alpha} + \right. \\ \left. + (\gamma - 2\alpha) \left[2\frac{4\alpha^2 - \gamma^2 - 1}{\gamma^2 + 2\alpha\gamma + 1} - \frac{1}{2}(2\alpha^2 - \alpha\gamma - 1) \right] \right\}. \quad (17)$$

На рис. 5 видно, что в данном случае минимум погрешности при уменьшении γ смещается в сторону меньших α . Случай $\gamma = \infty$ соответствует сигналу в форме единичной функции.

Представляет интерес случай двухступенчатого сигнала вида

$$F = \frac{1}{2}(1 + I_{\delta}), \quad (18)$$

т. е. функции, которая при $t=0$ скачком достигает значения $1/2$, а при $t=\delta$ вторым скачком достигает значения 1. Относительно такого сигнала можно а priori утверждать, что при $\delta=0$ мы получим те же соотношения, что и в случае единичной функции, а при δ , стремящемся к бесконечности, суммарная квадратичная погрешность составит в пределе $1/2$ от погрешности в случае единичной функции. Вычисление дает

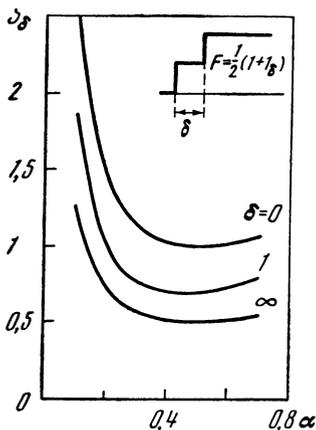
$$S_{\delta} = \frac{1}{2} S + \frac{1}{2} e^{-\alpha\delta} \left[S \cos \omega'\delta - \frac{1 - 4\alpha^2}{4\omega'} \sin \omega'\delta \right]. \quad (19)$$

Соответствующие графики показаны на рис. 6.

Замечательно, что во всех разобранных примерах погрешность получается меньше, чем в случае единичной функции. По-видимому, можно высказать следующее положение: суммарная квадратичная погрешность записи любой функции, удовлетворяющей следующим условиям: а) функция тождественно равна нулю до известного значения $t=t'$; б) функция монотонно изменяется в интервале $t' < t < \infty$ и в) функция ограничена в промежутке значений между 0 и 1, меньше такой же погрешности при записи единичной функции. Математическая формулировка этого положения, которое мне, к сожалению, не удалось пока доказать, в общем виде такова

$$\int_0^{\infty} \left[\frac{1}{q} 1 - h(t) \right]^2 dt \geq \int_0^{\infty} \left[\frac{1}{q} F(t) - x(t) \right]^2 dt,$$

где h — переходная функция системы; F — произвольная функция, удовлетворяющая вышеуказанным условиям; x — запись этой функции; $q = \dot{H}(0)$, где $H(p)$ — дифференциальный оператор системы, не содержащий отрицательных степеней p ($=d/dt$). Высказанное предположение служит одним из оснований для выбора единичной функции в качестве стандартного испытательного сигнала. Это же соображение может быть выражено с



Р и с. 6

несколько иной точки зрения посредством следующего чисто физического рассуждения: всякая система, уравнение которой содержит положительные степени (т. е. производные), дает тем большие погрешности, чем быстрее изменения сигнала; наоборот, отрицательные степени p (т. е. интегралы) вызывают погрешность при медленных изменениях сигнала; единичная же функция содержит в себе как раз сочетание бесконечно быстрого и бесконечно медленного изменения. Можно поэтому считать, что применение единичной функции в качестве стандартного испытательного сигнала дает и достаточно универсальный и наиболее суровый метод испытания.

В пользу выбора единичной функции в качестве стандартного испытательного сигнала говорят еще и следующие два соображения:

1. Существует разработанный и удобный для практических применений математический аппарат для вычисления переходных функций и т. п. расчетов.

2. Экспериментальная методика с применением сигнала в форме единичной функции весьма проста: например, в электрических цепях дело сводится к простому замыканию или размыканию цепи.

Случай, когда оператор системы содержит отрицательные степени p . Необходимо рассмотреть еще один вопрос, относящийся к тому случаю, когда уравнение системы, кроме положительных степеней p , содержит также и отрицательные, т. е. когда оператор $H(p)$ в уравнении

$$H(p)x = F$$

принимает следующую, более общую, форму .

$$H(p) = ap^n + bp^{n-1} + \dots + q + rp^{-1} + \dots + zp^{-m}.$$

В этом случае очевидно $H(0) = \infty$ и, следовательно, $x(\infty) = 0$ и интеграл (4) расходится. Совершенно естественно в этом случае суммировать квадратичную погрешность не от 0 до ∞ , а от 0 до некоторого времени T

$$S_T = \int_0^T \delta^2 dt.$$

Выбор величины T должен определяться условиями применения той аппаратуры, к которой относится данный конкретный

расчет; для аппаратуры, воспроизводящей слышимые звуковые явления, время T должно быть определенным образом связано с наибольшим периодом слышимого звукового колебания¹. Проще обстоит дело, например, в телеграфии постоянным током; время T в этом случае есть наибольшая длительность сигнала. Так или иначе, вопрос о выборе T есть вопрос о критерии для оценки искажений; как отмечалось в начале этой заметки, этот критерий должен быть непосредственно выведен из условий данной аппаратуры.

Обратимся к примеру. Пусть в той же системе, к которой относится уравнение (1), запись производится не смещением, а скоростью, так что уравнение это, написанное относительно скорости, принимает форму

$$\left(p + 2\alpha + \frac{\omega_0^2}{p}\right)v = \frac{1}{m}F. \quad (20)$$

При $F=1$ решение уравнения может быть получено простым дифференцированием (5)

$$v = \frac{1}{\omega'm} e^{-\alpha t} \sin \omega't. \quad (21)$$

По виду уравнения (20) можно непосредственно заключить о том, что для получения малых искажений нужно выбирать относительно большие значения α ; другими словами, система должна быть сверхпериодической. В этом случае решение принимает другую форму

$$v = \frac{1}{am} e^{-\alpha t} \operatorname{sh} at,$$

где $a = j\omega'$, или, в прежних обозначениях,

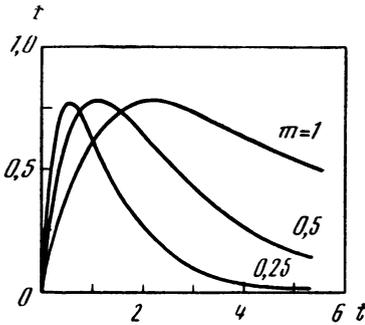
$$v = \frac{1}{r} \cdot \frac{2k}{\sqrt{k^2 - 1}} e^{-k\omega_0 t} \operatorname{sh} \sqrt{k^2 - 1} \omega_0 t.$$

Неискаженную запись определим в согласии с предыдущим как $v_0 = F/r$. Тогда

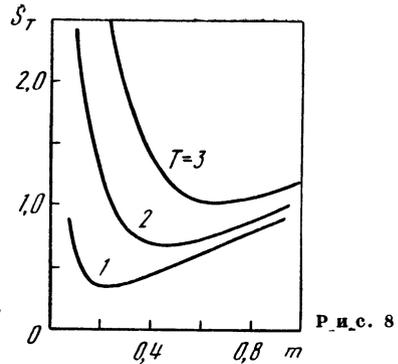
$$\delta = \frac{1}{r} \left(1 - \frac{2}{\sqrt{k^2 - 1}} e^{-k\omega_0 t} \operatorname{sh} \sqrt{k^2 - 1} \omega_0 t \right).$$

Не производя пока дальнейших вычислений, рассмотрим последнее соотношение. Как видим, в отличие от примеров предыдущих параграфов, погрешность не имеет минимума по перемен-

¹ Здесь приходится ссылаться на частотные зависимости ввиду того, что именно такими зависимостями описываются до последнего времени свойства слуха. Можно было бы ставить вопрос более определенно, если бы существовали физиологические исследования, непосредственно устанавливающие способность слуха реагировать, например, на определенную скорость изменения давления.



Р и с. 7



Р и с. 8

ным α и ω_0 , так как во всяком случае выгодно увеличивать α (при неизменном r) и уменьшать ω_0 (при неизменном m), т. е., иначе говоря, выгодно уменьшать m и s бесконечно. Можно найти оптимум по одной из переменных m или s , если связать их между собою условием $ms = \text{const}$ *. Тогда решение можно записать в виде

$$v = \frac{2}{r\varepsilon} e^{-t/2m} \text{sh} \frac{\varepsilon t}{2m}, \quad (22)$$

где

$$\sqrt{1 - 4ms/r^2} = \varepsilon.$$

На рис. 7 построены графики функции (22) при $r=1$, $\varepsilon=0,5$ и для различных значений m . Как видно из кривых, увеличение m (и связанное с ним уменьшение s) сопровождается увеличением погрешности при малых значениях t и уменьшением погрешности для больших значений t . Этот результат можно было бы, впрочем, предвидеть на том основании, что в (22) величина m входит лишь в комбинацию с t и, следовательно, изменение m можно рассматривать просто как изменение масштаба времени.

Рассматривая кривые рис. 7, легко убедиться в том, что наиболее выгодное значение параметра m зависит от верхнего предела интегрирования при вычислении суммарной квадратичной ошибки. Именно, чем больше T , тем больше наиболее выгодное значение m . После этих общих соображений можно довести вычисления до конца. Запишем выражение для суммарной квадратичной погрешности

$$S_T = \int_0^T \delta^2 dt = \frac{1}{r^2} \int_0^T \left[1 - \frac{2}{\varepsilon} e^{-\frac{t}{2m}} \text{sh} \frac{\varepsilon t}{2m} \right]^2 dt,$$

* Нельзя не признать, что такое условие носит несколько искусственный характер, но в данном примере, которым я воспользовался исключительно ради его простоты, я не сумел найти какого-либо соотношения, заимствованного из практики.

что, по выполнении интегрирования, дает

$$S_T = \frac{1}{r^2} \left[T + m \left\{ \frac{8e^{-T/2m}}{\varepsilon(1-\varepsilon^2)} \left(\operatorname{sh} \frac{\varepsilon T}{2m} + \varepsilon \operatorname{ch} \frac{\varepsilon T}{2m} \right) + \frac{8}{1-\varepsilon^2} - \frac{2e^{-T/m}}{\varepsilon^2(1-\varepsilon^2)} \left(\operatorname{ch} \frac{\varepsilon T}{m} + \varepsilon \operatorname{sh} \frac{\varepsilon T}{m} \right) + \frac{2}{\varepsilon^2(1-\varepsilon^2)} + \frac{2}{\varepsilon^2} (e^{-T/m} - 1) \right\} \right]. \quad (23)$$

Графики этой функции, построенные на рис. 8 при значениях $r=1$, $\varepsilon=0,5$ и $T=1, 2, 3$, иллюстрируют высказанные выше соображения.

* * *

Применение критерия квадратичной погрешности к оценке искажений представляется вполне целесообразным с принципиальной точки зрения.

К вопросу о методике экспериментального определения суммарной квадратичной погрешности я предполагаю вернуться.

Приложение к постановке вопроса об искажении (библиографическая справка). Нижеследующий краткий обзор, отнюдь не претендующий на полноту, имеет целью привести ряд примеров различного подхода к проблеме искажений и их устранении путем выбора наиболее выгодных соотношений величин. Эта проблема раньше всего возникла в связи с разработкой осциллографа; поэтому литературе, относящейся к осциллографу, уделено относительно большее внимание.

В исследованиях по теории осциллографа в качестве одной из наиболее ранних упоминается иногда работа Корню [2]. В этой работе речь идет о синхронизации, т. е. о приведении какой-либо колебательной системы в состояние вынужденных колебаний с заданной частотой. Рассматривая общее решение неоднородного уравнения, Корню приходит к заключению, что для получения установившегося режима необходимо, чтобы собственные колебания синхронизируемой системы были затухающими, причем стабильный режим устанавливается тем быстрее, чем больше показатель затухания. Таким образом, если принять, что отличие движения синхронизируемой системы в начальной фазе процесса от синхронизирующей силы есть в известном смысле искажение, то мерой этого искажения с точки зрения Корню является время установления процесса, т. е. время, в течение которого собственные колебания практически исчезают.

Блондель в работе, в которой он сообщает об изобретенном им осциллографе [3], устанавливает требования, которым должен удовлетворять осциллограф, но не дает критерия для оценки его погрешности. В следующей работе [4] ставится проблема интегральной синхронизации (в отличие от простой синхронизации Корню), которая есть не что иное, как проблема неискаженной записи.

Прежде всего Блондель отмечает, что всякое нарушение непрерывности регистрируемого явления сопровождается возникновением собственных колебаний; с этой точки зрения надлежит применять критическое затухание $\alpha/\omega_0=1$ (в наших обозначениях), т. е. применять аperiodический осциллограф. Далее, рассматривая решение уравнения движения с периодической правой частью, Блондель ставит условие, чтобы ни одна из гармоник записываемой периодической величины не была записана в масштабе большем, нежели основная частота: это условие приводит для затухания к соотношению

$$\alpha > \sqrt{(\omega_0^2 - \omega^2)/2},$$

где ω — основная частота¹.

¹ Полагая последнюю для предельного случая равной нулю, получаем $\alpha/\omega_0 > \sqrt{1/2}$, что по сопоставлению с первым условием дает $1 > \alpha/\omega_0 >$

Кроме того, отмечая зависящий от частоты фазный сдвиг между той или иной гармонической слагающей явления и ее записью, Блондель устанавливает условие малости этого сдвига.

Возвращаясь к тому же вопросу в более поздней работе [5], Блондель рассматривает его с точки зрения частотных характеристик. С этой целью он строит семейство характеристик при различных значениях α/ω_0 , а именно: 0; 0,4; 0,25; 0,50; 0,60; 0,667; 0,707; 0,75; 1,00 и 1,25. Представляя затем часть графика в области низких частот в более крупном масштабе, Блондель констатирует, что наиболее благоприятный вид характеристики получается при $\alpha/\omega_0=0,707$. Это значение относительного затухания Блондель рекомендует для записи непрерывных функций. Для функций же разрывных он по-прежнему предлагает применять $\alpha/\omega_0 \geq 1$, т. е. затухание выше критического.

Аналогичные рассуждения имеются в книге Орлиха [6], который указывает в приведенном численном примере, что при $\alpha/\omega_0=1$ и $\omega/\omega_0=0,02$ амплитуда основной частоты уменьшается на $4 \cdot 10^{-4}$ своего неискаженного значения, тогда как для девятой гармоники относительное уменьшение составляет 3,2%. Здесь же отмечается, что высшие гармонические имеют обычно малые по сравнению с основной частотой амплитуды, вследствие чего такое искажение мало скажется на форме записанной кривой. Относительно фазного сдвига Орлих указывает, что в первом приближении фазный угол пропорционален частоте, а следовательно, для всех гармоник, для которых это приближение имеет силу, получается некоторое одинаковое запаздывание, вследствие чего вся записанная кривая без искажения формы оказывается смещенной по оси времен вправо.

В статье А. Н. Крылова [7], вошедшей в главнейшей части в его известное сочинение по теории дифференциальных уравнений [8], вопрос об искажении записи поставлен совершенно определенно; мерой искажения является мгновенная погрешность в той же форме, как она была определена в данной заметке, т. е. $\delta=x_0-x$. Правая часть неоднородного уравнения, т. е. возбуждающая сила, вовсе не предположена периодической, она задана произвольной функцией времени и характеризуется скоростью нарастания и убывания. Зная последние величины, можно оценить и погрешность записи. Начальные условия выбраны нулевые, что соответствует реальным физическим условиям работы рассматриваемого рода приборов.

Вертгейм-Соломонсон [9] предлагает уменьшить затухание осциллографа, отмечая, что при критическом затухании правильное соотношение между гармониками соблюдается, например, с точностью до 1% лишь до частот $\omega=0,1\omega_0$. Он предлагает, в частности, полупериодическое затухание, но, не останавливаясь, однако, только на таком затухании, дает таблицы поправочных коэффициентов для следующих значений α/ω_0 : 0,707 (полупериодическое затухание); 0,614; 0,585 и 0,530. Последние три значения соответствуют наибольшим относительным погрешностям по амплитуде отдельных гармоник 3,5 и 10% в интервале частот от 0 до ω_0 .

В сравнительно недавней работе Хофман [10] исходит из аналогичных соображений. Допуская ошибку при записи различных частот до 5% и ссылаясь на то, что разницей в фазном сдвиге между значениями α/ω_0 0,6

$> \sqrt{1/2}$. Надо заметить, что Блондель допустил, по-видимому, ошибку в определениях. Он называет, ссылаясь на Кюри, $\alpha' = A/4\sqrt{2KC}$, где A , K и C — коэффициенты уравнения $K\theta + A\dot{\theta} + C\ddot{\theta} = F$.

При таком обозначении критическое затухание соответствует значению α' , равному не 1, как указывает Блондель, а 2. Значение 1 получается, если обозначить, как это и делают в последующем другие авторы, $\alpha' = 1/2\sqrt{KC}$. Определенная таким образом величина и есть то, что было вызвано в настоящей заметке относительным затуханием (degré d'amortissement, Dämpfungsgrad).

и 0,7 можно пренебречь, Хофман высказывается в пользу меньшего затухания, рекомендуя значение $\alpha/\omega_0=0,6$.

Большое влияние на последующие исследования оказала работа Буша [11]. Буш допускает запаздывание записи относительно сигнала и требует выполнения следующего соотношения между записью и сигналом

$$Ax_0(t + \tau) = F(t),$$

после чего раскладывает левую часть этого выражения по формуле Тейлора и подставляет в дифференциальное уравнение

$$m\dot{x} + r\dot{x} + sx = A \left(x_0 + \tau\dot{x}_0 + \frac{\tau^2}{2!}x_0 + \dots + \frac{\tau^n}{n!}x_0^{(n)} + \dots \right).$$

Затем Буш приравнивает коэффициенты при подобных членах и получает $A = s$, $\alpha/\omega_0 = \sqrt{1/2} = 0,707$ (полуаперриодическое затухание); $\omega_0\tau = \sqrt{2} = 1,41$ (ср. соотношения II и III настоящей заметки). Далее, определяя погрешность как

$$\delta = x_0(t + \tau) - x(t + \tau),$$

Буш получает

$$\begin{aligned} \delta &= \frac{1}{A} (m\dot{x} + r\dot{x} + sx) - \left(x + \frac{\tau}{1!}\dot{x} + \frac{\tau^2}{1!}\ddot{x} + \dots + \frac{\tau^n}{n!}x^{(n)} + \dots \right) = \\ &= \frac{\tau^3}{3!}x + \dots + \frac{\tau^n}{n!}x^{(n)} + \dots \end{aligned}$$

Последняя величина квалифицируется как присущая осциллографу погрешность.

Результаты этих простых выкладок не должны переоцениваться. Именно полученные соотношения вовсе не соответствуют какому-то универсальному оптимуму, так как наиболее выгодные соотношения (при условии выбора определенного критерия) могут быть найдены лишь для определенного вида функций $F(t)$. Об этом совершенно определенно говорит в конце своей работы и сам Буш, который рекомендует применение полуаперриодического затухания ($\alpha/\omega_0=0,707$) для всех гладких и непрерывных кривых, а в случае кривых с разрывами и углами допускает применение критического затухания ($\alpha/\omega_0=1$), хотя и полагает, что полуаперриодический осциллограф может дать более точные записи в том смысле, что истинный ход явления может быть легче восстановлен по искаженной записи.

Выкладки Буша почти буквально повторены в позднейших работах [12—14].

Цёллих [12] делает заключение в менее осторожной форме, нежели Буш; применительно к гладким кривым Цёллих утверждает, что «ошибки получаются наименьшие при относительном затухании $\alpha/\omega_0 = \sqrt{1/2}$ ». Помимо соображений, относящихся к частотным характеристикам, Цёллих пользуется (так же, как и Буш) в качестве критерия для оценки степени совершенства записи двумя параметрами записи единичной функции, а именно: быстротой нарастания отклонения и величиной первого выброса за установившееся значение отклонения, выражаемой в долях последнего. Отмечая, что чем больше затухание, тем медленнее нарастает отклонение, Цёллих склоняется с точки зрения этого соображения к меньшим значениям затухания; задаваясь же величиной первого выброса 1%, Цёллих приходит к значению $\alpha/\omega_0=0,83$ (для полуаперриодического затухания величина выброса составляет 4,3%).

Далее небезынтересно отметить, что, рассматривая вопрос о записи гармонической функции, Цёллих оперирует с временем запаздывания τ (так называемое время пробега) и ищет условия, при которых эта величина

в наименьшей мере зависит от частоты. Для этой цели он разворачивает в ряд выражение для τ

$$\tau = \frac{1}{\lambda} \operatorname{arctg} \operatorname{tg} \frac{2\alpha'\lambda}{1-\lambda^2} = 2\alpha' \left[1 + \lambda^2 \left(1 - \frac{4}{3} \alpha'^2 \right) + \lambda^4 \left(1 - 4\alpha'^2 + \frac{16}{5} \alpha'^4 \right) + \lambda^6 \left(1 - 8\alpha'^2 + 16\alpha'^4 - \frac{64}{7} \alpha'^6 \right) + \dots \right],$$

где $\alpha' = \alpha/\omega_0$; $\lambda = \omega/\omega_0$.

Основываясь на малости последней величины, Цёллих отбрасывает члены высших порядков и находит, что наиболее выгодное значение α/ω_0 с точки зрения наименьшей зависимости τ от частоты есть $\sqrt{3/4} = 0,866$.

В позднейших статьях Цёллиха [13, 15, 16] повторяются те же мысли. Эти статьи изобилуют графиками, из рассмотрения которых Цёллих слова заключает, что оптимальные значения параметров с точки зрения частотной и фазной характеристик не совпадают, хотя и незначительно различаются.

Крылов и Боголюбов [17], исходя из критерия квадратичной погрешности при записи периодической функции, составляют выражение для погрешности, которое существенно упрощается, если положить $\alpha/\omega_0 \sqrt{(1+\sqrt{2})}/2 = 0,777$. Приняв такое значение, авторы дают простые неравенства, посредством которых оцениваются средняя квадратичная и мгновенная погрешности записи.

В более поздней статье [18] Крылов и Боголюбов, отмечая, что применение точек зрения Корню и Бонделя приводят к противоречивым результатам, рассматривают проблему общей синхронизации (т. е. записи с наименьшими искажениями), пользуясь, как и прежде, критерием средней квадратичной ошибки.

Рассматривая случай периодической функции, представляемой рядом Фурье, и пользуясь обобщенным методом смещения диаграммы (т. е. допуская запаздывание записи относительно сигнала), авторы вычисляют таблицы, в одной из которых даны значения средней квадратичной ошибки для различных запаздываний. Выбирая наиболее выгодное значение запаздывания, авторы получают формулы, по которым может быть вычислена мгновенная, а затем и средняя квадратичная погрешность. Оптимальные значения параметров, а также и величина погрешности зависят от числа принятых в расчет членов ряда Фурье, выражающего сигнал, т. е., иначе говоря, от формы кривой.

Интересно проследить трактовку проблемы искажений в других областях, например, применительно к устройствам, применяемым в телевидении.

Робинсон [19], исследуя вопросы работы телевизионных усилителей, исходит из задачи воспроизведения прямоугольной периодической функции с низкой основной частью (частота кадров). Отмечая, что горизонтальные части такой кривой будут воспроизведены с искажением, выражающимся в совпадении этих частей со временем вследствие влияния цепи связи (конденсатор с утечкой), Робинсон оценивает искажение относительной погрешностью, которая не превышает $RC/2f_0$, где RC — постоянная времени цепи связи, f_0 — основная частота. Переходя к искажениям на высоких частотах и к вопросу о коррекции их, Робинсон отказывается от рассмотрения формы кривой и трактует вопрос с точки зрения фазных и частотных характеристик. Рассматривая схему корректированной цепи связи, Робинсон строит серию частотных характеристик, из рассмотрения которых выводит заключение о наиболее выгодном соотношении параметров. Далее, автор устанавливает соотношение, наиболее выгодное с точки зрения фазной характеристики. Нужно заметить, что Робинсон, в отличие от других авторов, считает фазную характеристику хорошей тогда, когда фазный сдвиг наименьшим образом отличается от нуля (а не когда он пропорционален частоте, т. е. когда время прогеба постоянно). Иначе говоря, Робинсон не допускает запаздывания записи относительно сигнала.

Брауде [1, 20] отвергает распространенный метод выбора параметров на основании рассмотрения частотных характеристик и применяет расчет-

ный метод, состоящий в разворачивании по формуле Тейлора и приравнявании нулю производных по частоте для частотной характеристики и времени пробега (или непосредственно самого фазного сдвига). Так, например, для электрической схемы, уравнение которой совпадает с обычным уравнением осциллографа (вследствие чего сопоставление результатов Брауде с вышеприведенными выводами, относящимися к осциллографу, приобретает особый интерес), Брауде приравняет только первую производную частотной характеристики нулю при $\omega=0$, что дает для α/ω_0 одно уравнение, из которого находится значение этой величины: $\alpha/\omega_0 = \sqrt{1/2} = 0,707$. Это значение признается оптимальным с точки зрения частотной характеристики.

Для фазной характеристики, приравнявая нулю третью производную фазного сдвига (или, что то же, вторую производную времени пробега), Брауде находит наиболее выгодное значение $\alpha/\omega_0 = \sqrt[3]{4} = 0,866$. Последний результат совпадает с данными Цёллиха, что и неудивительно, если принять во внимание сходность методов, применяемых обоими авторами. Отмечая, что в рассмотренной схеме нельзя одновременно получить оптимум как для частотной, так и для фазной характеристик, Брауде переходит к исследованию тем же методом более сложных схем.

Лурье [21] также рассматривает искажения в усилителе с точки зрения воспроизведения формы прямоугольной периодической кривой. Разделяя искажения при низких и при высоких частотах (что по сути дела соответствует погрешности при воспроизведении медленных и быстрых процессов), Лурье выбирает в качестве мерила последних быстроту нарастания функции, определяемую через время, потребное для достижения определенной доли предельного значения. Наряду с этим Лурье пользуется также и относительной погрешностью. Рассматривая случай скорректированного на высокой частоте каскада, он допускает превышение мгновенного значения над номинальным (первый выброс осциллографа) на 1%, что дает для соотношения параметров определенное значение, которое автор считает наиболее выгодным с точки зрения искажения формы кривой, и относительно которого он замечает, что оно лежит между значениями, соответствующими наилучшей частотной и наилучшей фазной характеристикам.

Рассматривая все вышепоименованные исследования, можно сделать следующее общее заключение. Все примененные в этих исследованиях методы делятся на две группы: одна группа методов оперирует непосредственно с сигналом и записью как функциями; в этой группе понятие искажения обычно определено достаточно отчетливо посредством мгновенной или средней квадратичной погрешности. Вторая группа методов оперирует с частотными и фазными характеристиками; при этом подразумевается, что сигнал и запись представлены соответственными интегралами Фурье.

Нужно отметить, что с течением времени понятие частотной характеристики заметно вырождается, отрываясь от своего физико-математического корня; частотная и фазная характеристики трактуются как некоторые самостоятельные сущности; при применении этих характеристик для оценки той или иной системы с точки зрения вносимых ею искажений зачастую совершенно выпадает связь между этими характеристиками и реальным поведением системы. Достаточно указать, что в связи с применением характеристик задается иногда вопрос о том, какая из двух характеристик — частотная или фазная — важнее с точки зрения искажений, т. е. по которой из двух характеристик следует выбирать параметры? Этот вопрос лишен реального смысла; форма здесь заслоняет содержание. На этот вопрос можно было бы отвечать так: для получения наилучшей фазной характеристики нужно исходить из фазной характеристики; для получения наилучшей частотной характеристики нужно исходить из частотной характеристики; для получения же наименьшего искажения формы кривой (с точки зрения какого-либо определенного критерия, например, с точки зрения мгновенной, средней или суммарной погрешности) нужно исходить не из частотной и фазной характеристик, а из соотношения между этими характеристиками и погрешностью, положенной в основу избранного критерия. Между тем это

соотношение как раз совершенно игнорируется некоторыми авторами исследований по вопросу об искажениях и их коррекции.

Правильный подход к проблеме рисуется мне в следующем виде. Следует прежде всего точно определить функции аппарата и установить, во-первых, предъявляемые к нему требования, во-вторых, вытекающие из этих требований свойства аппарата, и, в-третьих, критерий для их оценки. Далее, выбрав тот или иной способ описания свойств аппарата, надлежит установить соответствие между этим способом и выбранным критерием и не упускать из виду этого соответствия в течение всего последующего исследования.

Л и т е р а т у р а

1. Г. В. Брауде. О коррекции частотных и фазовых характеристик усилительных устройств. Журн. техн. физ., 4, 9, 1934, 1714—1739; 4, 10, 1934, 1818—1828.
2. А. Cornu. Sur la condition de stabilité du mouvement d'un système oscillant soumis à une liaison synchrone pendulaire. C. r. Acad. sci., 104, 22, 1887, 1463—1470.
3. А. Blondel. Oscillographes, nouveaux appareils pour l'étude des oscillations électriques lentes. C. r. Acad. sci., 116, 10, 1893, 502—506.
4. А. Blondel. Conditions générales que doivent remplir les instruments enregistreurs ou indicateurs: problème de la synchronisation intégrale. C. r. Acad. sci., 116, 15, 1893, 748—752.
5. А. Blondel. Théorie des oscillographes. L'éclairage électrique, 33, 43, 1902, 115—125.
6. E. Orlich. Aufnahme und Analyse von Wechselstromkurven. Braunschweig, F. Vieweg, 1906.
7. А. Н. Крылов. Некоторые замечания о крешерах и индикаторах. Изв. АН СССР, 6, 9, 1909, 623—654.
8. А. Н. Крылов. О некоторых дифференциальных уравнениях математической физики, имеющих приложения в технических вопросах. Л., Изд-во АН СССР, 1933.
9. J. K. A. Wertheim-Solomonson. Some points in the use and the Theory of the oscillograph. Electrician, 69, 9, 1912, 357—360.
10. W. Hoffman. Die richtige Dämpfung der Oszillographenschleifen. WWSK, 12, 2, 1933, 142—155.
11. H. Bush. Eine einfache Ableitung der vollständigen Theorie des Oszillographen. Phys. Z., 13, 1912, 615—623.
12. H. Zölllich. Prüfung von Messgeräten zur Aufzeichnung sich rasch verändernder Gressen. WWSK, 1, 1, 1920, 24—63.
13. H. Zölllich. Aufzeichnung schnell veränderlichen Vorgänge. Allgem. ATM, 365-1, Lief. 40, 1934.
14. W. Hort. Technische Schwingungslehre. 2 Aufl. Berlin, J. Springer, 1922.
15. H. Zölllich. Aufzeichnung schnell veränderlichen Vorgänge. Wiedergabe einfacher und zusammengesetzter harmonischer Schwingungen. ATM, 363-3, Lief. 41, 1934.
16. H. Zölllich. Aufzeichnung schnell veränderlicher Vorgänge. Weidergabe ungunstige Vorgänge. ATM, 365-4, Lief. 42, 1934.
17. Н. Крылов, Н. Боголюбов. Sur la théorie des appareils indicateur. J. phys., 7, 1, 3, 1930, 77—92.
18. Н. Крылов, Н. Боголюбов. Sur la théorie mathématique des oscillographes. C. r. Acad. sci., 187, 21, 1928, 938—940.
19. G. D. Robinson. Theoretical notes on certain features of television receiving circuits. Proc. IRE, 21, 6, 1933, 833—843.
20. Г. В. Брауде. О коррекции усилительных устройств на низкой частоте. Журн. техн. физ., 6, 8, 1936, 1334—1343.
21. О. Б. Лурье. Нестационарные явления в широкополосных усилителях, Журн. техн. физ., 6, 2, 1963, 319—343.

ОБ ЭКСПЕРИМЕНТАЛЬНОМ ПОДБОРЕ ПАРАМЕТРОВ СЛОЖНЫХ СИСТЕМ

В предыдущих статьях [1] рассматривался вопрос о применении критерия суммарной квадратичной погрешности к оценке линейных искажений и о технике измерений, связанных с применением этого критерия. Ниже будут рассмотрены некоторые дополнительные вопросы постановки измерений с точки зрения одного специального применения предлагаемого метода.

Подобные системы. Одно из первых практических затруднений, встречаемых при измерении суммарной квадратичной погрешности (СКП), состоит в том, что СКП может оказаться абсолютно малой, а потому ее измерение невозможным вследствие недостаточной чувствительности измерительного прибора. Если речь идет об измерении некоторых вполне определенных данных объектов, то единственная — вовсе не исключенная — возможность осуществления измерения состоит в повышении чувствительности измерительного прибора. Если же задача состоит в исследовании общих свойств некоторой системы, то может оказаться весьма полезным применение принципа подобия, который мы вкратце и рассмотрим.

Прежде всего нужно напомнить, что СКП выражается определенным интегралом по времени от квадрата мгновенной погрешности. Величина этого интеграла зависит от масштабов функции и независимой переменной. Функция, т. е. мгновенная погрешность, в случае электрических измерений (которые и имеются в виду) определяется значениями токов или напряжений. Очевидно, что масштабы здесь ограничены как общетехническими соображениями, так и, опять-таки, свойствами измерительного прибора. Что касается независимой переменной, то ее масштаб в данном случае, к сожалению, не зависит от нашего усмотрения. Но если мы не можем ускорить течение времени, то во всяком случае можем осуществить систему, в которой всякий раз процесс протекает в любое число раз медленнее, чем в данной. В этом смысле и понимается в дальнейшем принцип подобия. Системой, подобной данной, будет называться такая система, поведение которой отличается от поведения данной только масштабом времени. При построении подобной системы необходимо соблюсти некоторые простые правила. Во-первых, само собою разумеется, подобная система должна иметь уравнение того же вида, как и уравнение данной системы. Во-вторых, должны быть выпол-

нены известные соотношения между параметрами. Об этих соотношениях можно говорить по-разному: проще всего, пожалуй, поставить вопрос следующим образом: время входит в аналитическое выражение переходной функции линейной системы только в виде сложных аргументов, представляющих собою произведение времени на корни характеристического уравнения. Эти корни имеют размерность $[T^{-1}]$ и носят название собственных частот, показателей затухания или величин, обратных постоянным времени. Необходимо, очевидно, чтобы все корни уравнения изменились в одинаковое число раз; это соображение и дает правило для построения подобной системы¹.

Пусть, например, дана электрическая цепь, состоящая из последовательно соединенных R , L и C . Мы имеем $2\alpha = R/L$ (удвоенный показатель затухания), $\omega_0^2 = 1/LC$ (квадрат резонансной частоты), $\beta = 1/RC$ (постоянная времени), $\omega'^2 = 1/LC - R^2/4L^2$ (квадрат собственной частоты) и т. д.

Если измененное время

$$\bar{t} = At, \quad (1)$$

то, очевидно, должно быть

$$\bar{\alpha}\bar{t} = \alpha t, \quad \bar{\alpha} = \alpha/A, \quad \bar{\omega}_0\bar{t} = \omega_0 t, \quad \bar{\omega}_0 = \omega_0/A.$$

Отсюда

$$\frac{\bar{R}}{\bar{L}} = \frac{1}{A} \cdot \frac{R}{L}, \quad (2a)$$

$$LC = A^2 \bar{L}\bar{C}. \quad (2б)$$

Выбрав одну из трех переменных \bar{R} , \bar{L} или \bar{C} на основании тех или иных технических соображений, находим два остальных с помощью уравнений (2). Пусть, например, выбрана величина \bar{R} . Тогда для \bar{L} и \bar{C} имеем

$$\bar{L} = AL\bar{R}/R,$$

$$\bar{C} = ACR/\bar{R}.$$

При более сложной системе число коэффициентов, а следовательно, и число корней характеристического уравнения возрастает; при этом, однако, число корней равно порядку уравнения, а число независимых коэффициентов в приведенной форме уравнения — на единицу меньше, так что задача всегда разрешима.

Аналогичные системы. В целом ряде действительных схем, рассматриваемых в качестве объектов измерения, могут встретиться затруднения с измерением тех или иных величин. Так, например, может оказаться, что напряжение вообще удобнее

¹ Другой подход к вопросу основан на обращении непосредственно к оператору системы и на умножении p^n на $a-n$, где A — масштабный коэффициент времени.

измерять, нежели ток; измерения напряжения на самоиндукции практически неосуществимы вследствие наличия активного сопротивления во всякой реальной катушке; измерение тока в чисто емкостной ветви наталкивается на значительные затруднения вследствие влияния сопротивления самого прибора. Подобного рода затруднения заставляют искать схемы, аналогичные данным по своим свойствам, но отличные по устройству. Так, например, в слаботочной технике давно уже известны так называемые обратные схемы, проводимость которых выражается аналитически так же, как сопротивление прямых схем. Иначе говоря, уравнение напряжений прямых схем и уравнения токов обратных схем аналогичны [2].

Мы будем называть аналогичной схемой (или вообще системой) всякую схему, описываемую уравнением, подобным уравнению данной системы. Оказывается, что число аналогичных схем довольно велико — больше, чем обычно полагают. Это позволяет значительно обогатить метод аналогий вообще. Заметим, однако, что с точки зрения поставленной цели аналогии интересуют нас не как возможность изменения способа описания явления, а как возможность подобрать практически удобный объект измерения.

Рассмотрим пример. Пусть требуется составить схему, аналогичную механической системе с одной степенью свободы, уравнение которой гласит

$$H(p)x = (mp^2 + rp + s)x = F.$$

Формальное решение этого уравнения будет

$$x = \frac{1}{s} \cdot \frac{1}{\frac{m}{s}p^2 + \frac{r}{s}p + 1}.$$

Очевидно, что аналогичная схема может быть составлена из элементов R , L и C . Оказывается, однако, что число аналогичных схем равно шести, а число переменных, которыми замещается величина X , восьми. Все эти возможные комбинации даны на рис. 1*. Остается сделать выбор. Описанный ранее прибор приспособлен для из-

* Заметим, что если бы мы согласились включить в число переменных не только токи и напряжения, но и их производные, то число возможных комбинаций еще возросло бы. Кроме того, интересно указать, что соотношения рис. 1 дают основания для построения восьми различных систем электромеханических аналогий.

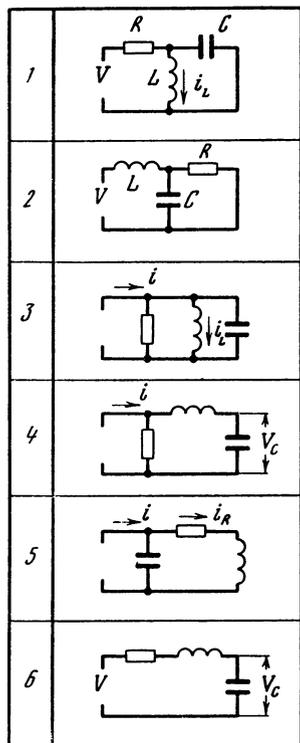


Рис. 1

Таблица 1

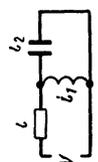
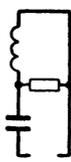
№	Схема	Функция					
		1	2	3	4	5	6
		$i(V)$	$i_1(V)$	$i_2(V)$	$i_1(i)$	$i_2(i)$	$V_R(V)$
1		$\frac{V}{R} \cdot \frac{LCp^2 + 1}{LCp^2 + \frac{L}{R}p + 1}$	$\frac{V}{R} \cdot \frac{1}{LCp^2 + \frac{L}{R}p + 1}$	$\frac{V}{R} \cdot \frac{LCp^2}{LCp^2 + \frac{L}{R}p + 1}$	$i \cdot \frac{1}{LCp^2 + 1}$	$i \cdot \frac{LCp^2}{LCp^2 + 1}$	$V \cdot \frac{LCp^2 + 1}{LCp^2 + \frac{L}{R}p + 1}$
2		$\frac{V}{R} \cdot \frac{RCp + 1}{LCp^2 + \frac{L}{R}p + 1}$	$\frac{V}{R} \cdot \frac{RCp}{LCp^2 + \frac{L}{R}p + 1}$	$\frac{V}{R} \cdot \frac{1}{LCp^2 + \frac{L}{R}p + 1}$	$i \cdot \frac{RCp}{RCp + 1}$	$i \cdot \frac{1}{RCp + 1}$	$V \cdot \frac{1}{LCp^2 + \frac{L}{R}p + 1}$
3		$\frac{V}{R} \cdot \frac{LCp^2 + RCp}{LCp^2 + \frac{L}{R}p + 1}$	$\frac{V}{R} \cdot \frac{LCp^2}{LCp^2 + \frac{L}{R}p + 1}$	$\frac{V}{R} \cdot \frac{RCp}{LCp^2 + \frac{L}{R}p + 1}$	$i \cdot \frac{\frac{L}{R}p}{\frac{L}{R}p + 1}$	$i \cdot \frac{1}{\frac{L}{R}p + 1}$	$V \cdot \frac{LCp^2}{LCp^2 + \frac{L}{R}p + 1}$
4		$\frac{V}{R} \cdot \frac{\frac{L}{R}p + 1}{\frac{L}{R}p}$	$\frac{V}{pL}$	VpC	$i \cdot \frac{1}{LCp^2 + \frac{L}{R}p + 1}$	$i \cdot \frac{LCp^2}{LCp^2 + \frac{L}{R}p + 1}$	V

Таблица 1 (продолжение)

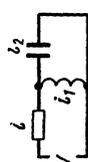
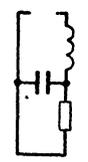
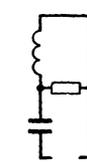
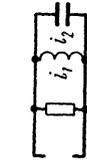
		Функция				
№	Схема	7	8	9	10	11
		$V_L(V)$	$V_C(V)$	$V_R(i)$	$V_L(i)$	$V_C(i)$
1		$V \cdot \frac{\frac{L}{R} p}{LCp^2 + \frac{L}{R} p + 1}$	$V \cdot \frac{\frac{L}{R} p}{LCp^2 + \frac{L}{R} p + 1}$	iR	$iR \cdot \frac{\frac{L}{R} p}{LCp^2 + 1}$	$iR \cdot \frac{\frac{L}{R} p}{LCp^2 + 1}$
2		$V \cdot \frac{LCp^2 + \frac{L}{R} p}{LCp^2 + \frac{L}{R} p + 1}$	$V \cdot \frac{1}{LCp^2 + \frac{L}{R} p + 1}$	$iR \cdot \frac{1}{RCp + 1}$	ipL	$iR \cdot \frac{1}{RCp + 1}$
3		$V \cdot \frac{LCp^2}{LCp^2 + \frac{L}{R} p + 1}$	$V \cdot \frac{\frac{L}{R} p + 1}{LCp^2 + \frac{L}{R} p + 1}$	$iR \cdot \frac{\frac{L}{R} p}{\frac{L}{R} p + 1}$	$iR \cdot \frac{\frac{L}{R} p}{\frac{L}{R} p + 1}$	$i \cdot \frac{1}{pC}$
4		V	V	$iR \cdot \frac{\frac{L}{R} p}{LCp^2 + \frac{L}{R} p + 1}$	$iR \cdot \frac{\frac{L}{R} p}{LCp^2 + \frac{L}{R} p + 1}$	$iR \cdot \frac{\frac{L}{R} p}{LCp^2 + \frac{L}{R} p + 1}$

Таблица 1 (продолжение)

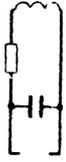
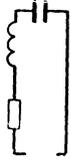
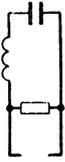
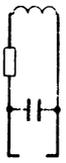
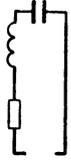
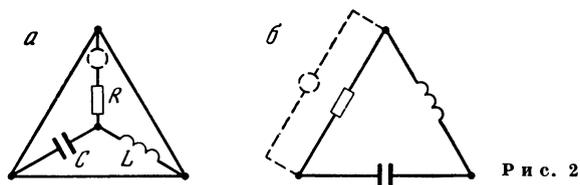
№	Схема	Функции					
		1	2	3	4	5	6
		$i (V)$	$i_1 (V)$	$i_2 (V)$	$i_1 (i)$	$i_2 (i)$	$V R (V)$
5		$\frac{V \cdot LCp^2 + RCp + 1}{R \cdot LCp^2 + 1}$	$\frac{V}{R}$	$\frac{V}{pL + \frac{1}{pC}}$	$i \cdot \frac{LCp^2 + 1}{LCp^2 + RCp + 1}$	$i \cdot \frac{RCp}{LCp^2 + RCp + 1}$	V
6		$\frac{V \cdot LCp^2 + RCp + 1}{R \cdot LCp^2 + \frac{L}{R} \cdot p}$	$\frac{V}{pL}$	$\frac{V}{R + \frac{1}{pC}}$	$i \cdot \frac{RCp + 1}{LCp^2 + RCp + 1}$	$i \cdot \frac{LCp^2}{LCp^2 + RCp + 1}$	$V \cdot \frac{RCp}{RCp + 1}$
7		$\frac{V \cdot LCp^2 + RCp + 1}{R \cdot \frac{L}{R} \cdot p + 1}$	VpC	$\frac{V}{pL + R}$	$i \cdot \frac{LCp^2 + RCp}{LCp^2 + RCp + 1}$	$i \cdot \frac{1}{LCp^2 + RCp + 1}$	$V \cdot \frac{1}{\frac{L}{R} \cdot p + 1}$
8		$\frac{V \cdot RCp}{R \cdot LCp^2 + RCp + 1}$	—	—	—	—	$V \cdot \frac{RCp}{LCp^2 + RCp + 1}$

Таблица 1 (окончание)

№	Схема	Функция				
		7	8	9	10	11
		$V_L(V)$	$V_C(V)$	$V_R(i)$	$V_L(i)$	$V_C(i)$
5		$V \cdot \frac{LCp^2}{LCp^2 + 1}$	$V \cdot \frac{1}{LCp^2 + 1}$	$iR \cdot \frac{LCp^2 + 1}{LCp^2 + RCp + 1}$	$iR \cdot \frac{LCp^2}{LCp^2 + RCp + 1}$	$iR \cdot \frac{1}{LCp^2 + RCp + 1}$
6		V	$V \cdot \frac{1}{RCp + 1}$	$iR \cdot \frac{LCp^2}{LCp^2 + RCp + 1}$	$iR \cdot \frac{LCp^2 + \frac{L}{R}p}{LCp^2 + RCp + 1}$	$iR \cdot \frac{\frac{L}{R}p}{LCp^2 + RCp + 1}$
7		$V \cdot \frac{\frac{1}{R}p}{\frac{L}{R}p + 1}$	V	$iR \cdot \frac{1}{LCp^2 + RCp + 1}$	$iR \cdot \frac{\frac{L}{R}p}{LCp^2 + RCp + 1}$	$iR \cdot \frac{\frac{L}{R}p + 1}{LCp^2 + RCp + 1}$
8		$V \cdot \frac{LCp^2}{LCp^2 + RCp + 1}$	$V \cdot \frac{1}{LCp^2 + RCp + 1}$	iR	ipL	$i \cdot \frac{1}{pC}$

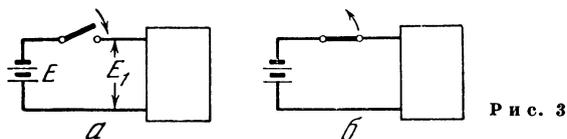
мерения напряжений; с этой точки зрения пригодны схемы 2, 4, 5 и 6. Если входное сопротивление прибора невелико, то его подключение параллельно емкости может повлиять на работу схем 2 и 5. Теперь вопрос идет о задании входного режима. Схемы 2 и 5 различаются между собою тем, что первая дает требуемое напряжение как функцию входного напряжения, а вторая — как функцию входного тока. О способах задания этих режимов говорится в следующем параграфе.



Р и с. 2

Для того чтобы исчерпать все возможные комбинации для трех элементов R , L и C , составлена табл. 1. В ней содержится 8 строк и 11 столбцов. Строки соответствуют восьми возможным схемам; столбцы дают все токи в функции входных напряжений, токи — в функции входных токов, напряжения — в функции входных напряжений и напряжения — в функции входных токов. Комбинации рис. 1 занимают в табл. 1 следующие места (первый индекс относится к строке, второй — к столбцу): 1.2; 2.3; 2.6; 4.4; 5.11; 7.5; 7.9 и 8.8.

Порядок составления схем табл. 1 соответствует рис. 2. Схемы 1, 2 и 3 получаются из схемы рис. 2, *a* путем введения источника в один из лучей звезды; схема 4 получается путем включения источника между центром звезды и внешним проводом. Схемы



Р и с. 3

5—7 получаются из схемы рис. 2, *б* путем подключения источника параллельно одной из сторон треугольника: схема 8 получается путем введения источника в контур треугольника.

Входной режим. Согласно выбранному определению измерение СКП производится при сигнале в форме единичной функции. Как видно из предыдущего, речь может идти об изменении по закону единичной функции либо входного напряжения, либо входного тока. Приложение напряжения в форме единичной функции осуществляется замыканием ключа в схеме рис. 3, *a*. После замыкания входной ток будет определяться переходной проводимостью подключенной схемы. Если мы хотим, чтобы по закону единичной функции применялся входной ток, то для этого, как известно,

необходимо, чтобы внутреннее сопротивление источника было значительно больше сопротивления питаемой схемы. В пределе мы должны были бы иметь бесконечно большие э. д. с. и внутреннее сопротивление источника. Это легче сказать, чем сделать. На практике желаемый результат очень просто достигается размыканием ключа по схеме рис. 3, б. Тогда протекавший по схеме постоянный ток внезапно прекращается, что соответствует его изменению по закону $1-I$, где I — символ единичной функции. По существу это, конечно, одно и то же; необходимо только устранить действие постоянной составляющей [3].

Надо заметить, что при размыкании ключа становятся излишними все те предосторожности, которые необходимо соблюдать при работе ключа на замыкание и о которых говорилось в одной из предыдущих статей. К сожалению, указанный простой и удобный прием не универсален: он применим только для схем с проводимостью для постоянного тока, т. е. иначе говоря — для схем, у которых $1/H(0) \neq 0$. Комбинируя схемы и входные режимы, можно получить разнообразие возможностей. В связи с этим интересно отметить, что поведение схем под действием сигнала в форме единичной функции можно трактовать как свободное колебание, т. е. как решение однородного уравнения $H(p)x=0$ при различных начальных условиях. Рассмотрим схемы 5—7 табл. 1 при питании их по схеме рис. 3, б. После того, как ключ разомкнут, схема изолирована от источника, ее поведение описывается однородным уравнением

$$(Lp + R + 1/Cp) = 0,$$

т. е., иначе говоря, все три схемы сводятся к универсальному треугольнику рис. 2, б. Однако начальные условия для всех трех схем различны. В схеме 5 ток через самоиндукцию в начальный момент равен нулю, а конденсатор имеет заряд $Q_0 = i_0 RC$. В схеме 6 начальный заряд равен нулю, а через самоиндукцию течет ток i_0 . Наконец, в схеме 7 начальные условия определяются и начальным током i_0 и начальным зарядом Q_0 . Решение уравнения должно в этом случае, очевидно, выражаться на основании принципа суперпозиции суммой решений для двух предыдущих случаев. Это подтверждается рассмотрением операторных выражений, которые, если взять за переменную напряжение на конденсаторе, можно найти в столбце 11 табл. 1. Решение для схем 5—7 в явной форме имеют следующий вид:

схема 5

$$V_C = i_0 R e^{-\alpha t} \left(\frac{\alpha}{\omega'} \sin \omega' t + \cos \omega' t \right);$$

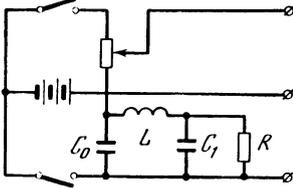
схема 6

$$V_C = -\frac{i_0}{\omega' C} e^{-\alpha t} \sin \omega' t = -i_0 R \frac{\omega_0^2}{2\alpha\omega'} e^{-\alpha t} \sin \omega' t;$$

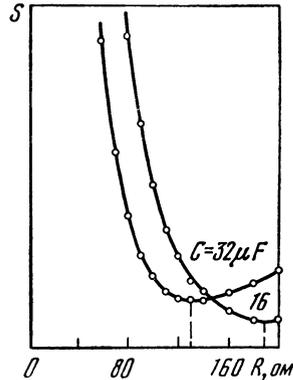
схема 7

$$V_C = i_0 R e^{-\alpha t} \left(\frac{2\alpha^2 - \omega_0^2}{2\alpha\omega'} \sin \omega' t + \cos \omega' t \right).$$

Интересно заметить, что в то время, как в первом случае V_C не может превзойти $i_0 R$, ни при каких обстоятельствах, в двух других случаях при неограниченном уменьшении α амплитуда V_C



Р и с. 4



Р и с. 5

стремится к бесконечности. Приведенные рассуждения показывают в общих чертах, как вариации начальных условий могут быть реализованы путем вариации схемы.

Пример исследования: схема коррекции. Рассмотрим одну схему коррекции на высоких частотах, применяемую для широкополосных усилителей [4—6]. Эта схема в точности соответствует схеме 7 в табл. 1, а условия ее использования описываются операторным выражением столбца 11. Задача состоит в нахождении наивыгоднейшего относительного затухания

$$k = \alpha/\omega_0 = R/2\sqrt{C/L}.$$

При таком выборе переменного параметра вопросы подобия не играют никакой роли, так как k — безразмерная величина. Мы можем, следовательно, выбрать постоянные испытываемой схемы полностью по своему усмотрению. При неизменных L и C k будут зависеть только от R . Схема измерения представлена на рис. 4. Ключи размыкаются одновременно; это достигается применением реле. Некоторая неточность работы реле не играет при этом роли, так как из-за неодновременного срабатывания обоих контактов отсчет по прибору изменяется на постоянную величину, что не меняет общего характера зависимости. Впрочем опыт показал, что разновременность работы контактов может быть сделана ничтожно малой и практически неощутимой; для проверки ключи приводятся в действие при отключенных L и C . Потенциометр R_0 имеет своим назначением сбалансировать некоторую несиммет-

рию измерительного прибора. Три зажима в правой части схемы принадлежат измерителю СКП, уже описанному ранее. Данные схемы: $E=20$ в, $L=0,5$ гн ($+14 \Omega$), $C=16$ и 32 мф, R изменялось от 60 до 200 ом. Результаты измерения показаны на рис. 5. Как видно из графиков, оптимальное значение относительного затухания оказывается между 0,55 и 0,6.

Для данной схемы возможно теоретическое исследование. Мы имеем

$$\begin{aligned} \frac{V_C}{i_0 R} &= \frac{\frac{L}{R} p + 1}{LCp^2 + RCp + 1}, \\ \frac{V_C}{i_0 R} &= 1 - e^{-\alpha t} \left(\frac{2a^2 - \omega_0^2}{2a\omega'} \sin \omega' t + \cos \omega' t \right), \\ \delta &= e^{-\alpha t} \left(\frac{2a^2 - \omega_0^2}{2a\omega'} \sin \omega' t + \cos \omega' t \right), \\ S &= \int_0^{\infty} \delta^2 dt = \frac{1}{\omega_0} \left(k - \frac{1}{4k} + \frac{1}{16k^3} \right). \end{aligned}$$

Эта функция имеет минимум при $k = \sqrt{(\sqrt{13} - 1)/8} = 0,571$.

Соответствующие значения приведены на рис. 5: как видим, определение наиболее выгодного значения параметра экспериментальным путем является достаточно надежной операцией.

Интересно сравнить между собою две схемы, применяемые для коррекции на высоких частотах. Эти схемы значатся под № 5 и 7 в табл. 1. О второй из этих схем только что говорилось, первая была исследована ранее.

Суждение о сравнительных свойствах обеих схем можно составить, сопоставляя нижеследующие данные (табл. 2).

Таблица 2

	Схема 5	Схема 7
$\frac{V_C}{i_0 R}$	$\frac{1}{LCp^2 + RCp + 1}$	$\frac{L}{R} p + 1$
$h(t)$	$1 - e^{-\alpha t} \left(\frac{\alpha}{\omega'} \sin \omega' t + \cos \omega' t \right)$	$1 - e^{-\alpha t} \left(\frac{2d^2 - \omega_0^2}{2a\omega'} \sin \omega' t + \right. \\ \left. + \cos \omega' t \right)$
S	$k + 1/4k$	$k - 1/4k + 1/16k^3$
k_{\min}	0,5	0,571
S_{\min}	1	0,47

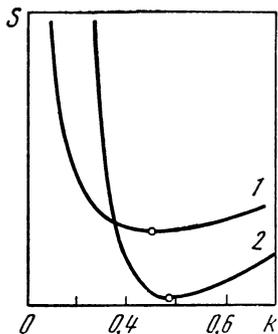


Рис. 6

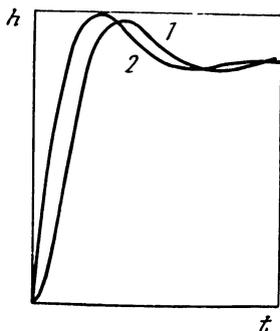


Рис. 7

Из этой таблицы сразу видно, что схема 7 более совершенна. Ее преимущество состоит главным образом в том, что переходная функция первой схемы в начальный момент нарастает на параболу, а второй схемы — по прямой. Иначе говоря, $h(0) = 0$ для первой схемы и постоянной величине для второй схемы.

Эту постоянную легко найти; для этого нужно найти предел умноженного на p операционного выражения для переходной проводимости при p , стремящемся к бесконечности, т. е.

$$h(0) = \lim_{p \rightarrow \infty} \frac{p \left(\frac{L}{R} p + 1 \right)}{LCp^2 + RCp + 1} = \frac{1}{RC}.$$

Иллюстрациями к этим рассуждениям и табл. 2 служит рис. 6, на котором изображены зависимости $S(k)$ для обеих схем ($1 - S = k + 1/4 k$; $2 - S = k - 1/4 k + 1/16 k^3$) и рис. 7, на котором построены переходные функции обеих схем при наивыгоднейших значениях относительного затухания ($1 -$ схема 5 при $k = 0,5$; $2 -$ схема 7 при $k = 0,571$).

Пример исследования: схема, аналогичная индукционному микрофону. Эквивалентная схема индукционного микрофона

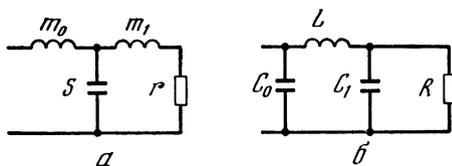
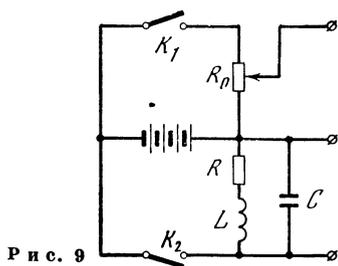


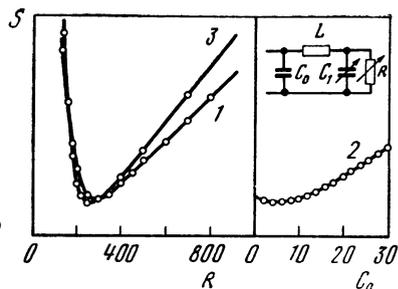
Рис. 8

Уэнте [7] для малых времен (высоких частот) имеет вид, показанный на рис. 8, а. Поставим себе задачей подобрать наивыгоднейшие с точки зрения минимума СКП значения параметров m_1 и r при заданных и неизменных m_0 и S . Погрешность в данном случае будет определяться как разность между единичной функцией и переходным сопротивлением (или проводимостью) исследуемой схемы. Здесь нужно заметить, что схема рис. 8, а импульсивна (пользуясь краткой американской терминологией) по со-



Р и с. 9

Р и с. 10



противлению, а поэтому непригодна для исследования путем изменения входного тока по закону единичной функции¹.

Можно построить обратную схему (рис. 8, б), проводимость которой определяется уравнением, аналогичным уравнению для сопротивления схемы 7а. Для обеих схем имеем следующее соответствие между параметрами:

схема 7а

$$m_0 m_1 S V,$$

схема 7б

$$C_0 C_1 1/L 1/R.$$

Задача состоит в нахождении минимума СКП по двум переменным параметрам: C_1 и R . Эта аналитическая задача решается с помощью системы двух уравнений

$$\partial S(R, C_1)/\partial R = 0, \quad \partial S(R, C_1)/\partial C_1 = 0.$$

Практически задача будет решаться следующим образом. Будем изменять один из двух параметров, например R , сохраняя неизменное произвольное значение $C_1 = A$, и найдем $S_1 = S(R, A)$. Найдя по графику значение $R = B$, соответствующее минимуму этой функции, будем строить функцию $S_2 = S(B, C_1)$. Найдём значение C_1 , соответствующее минимуму этой последней функции, и т. д. Минимум функции нескольких переменных находится, таким образом, рядом последовательных приближений. В рассматриваемом случае двух переменных такая последовательность действий имеет простой геометрический смысл. Функция $f(S_1 R_1 C_1) = 0$ представляет собою уравнение некоторой поверхности в системе координат R, C_1, S . Кривая S_1 представляет собою линию пересечения этой поверхности с плоскостью, параллельной SOR . Из наинизшей точки этой кривой мы движемся по линии пересечения поверхности с плоскостью, параллельной SOC_1 , и т. д. Полная схема опыта показана на рис. 9. Данные схемы $C_0 = 16$ мф, $L = 1$ гн, R изменялось от 100 до 1000 ом; C_1 изменялось от 0 до 30 мф. Результаты опыта иллюстрируются тремя

¹ Я просмотрел это обстоятельство. Эту ошибку обнаружил и указал мне А. М. Нанос, за что я ему благодарен.

кривыми рис. 10. 1 — $S(R)$ при $C_1=0$; 2 — $S(C_1)$ при $R=300$ ом; 3 — $S(R)$ при $C_1=4$ мкф. Прежде всего снята кривая 1 зависимости S от R при $C_1=0$. Значение R , соответствующее минимуму, оказалось равным 300 ом. Затем снята кривая 2 в зависимости от C_1 при $R=300$ ом. Весьма расплывчатый минимум обнаруживается при $C_1 \cong 4$ мкф. Снимая еще раз зависимость от R , находим мало отличающееся от предыдущего значение $R \cong 250$ ом. Продолжая такие последовательные действия, мы могли бы уточнить найденные значения. В данном случае это не представляется необходимым. Замечательно, что минимум по R получается весьма острый, тогда как минимум по C_1 выражен весьма слабо. Из этого непосредственно следует, что правильный подбор первой величины имеет гораздо большее значение, нежели второй.

В заключение интересно сопоставить полученные из описанного опыта цифровые значения постоянных с вычисленными Л. Я. Гутиным [8]. Исходя из требования наименьшей зависимости сопротивления от частоты, Л. Я. Гутин находит следующие значения: $m_0=1$ г, $m_1=0,5$ г, $S=15 \cdot 10^8$ дин/см, $p=35 \cdot 10^3$ ом.

Определяем масштаб времени $A = \sqrt{LC_0 s / m_0} = \sqrt{16 \cdot 10^{-6} \cdot 15 \cdot 10^8} \cong 155$. Далее

$$R = Am_0 / C_0 r = 155 / 16 \cdot 10^{-6} \cdot 35 \cdot 10^3 = 275 \text{ ом.}$$

Это совпадение с данными описанного опыта даже несколько удивительно, если иметь в виду существенное различие критериев. Что касается величины C_1 , то по Л. Я. Гутину мы должны были бы иметь $C_1=8$ мкф. Как видим из рис. 10, разница в значении СКП при $C_1=4$ и 8 мкф совершенно незначительна.

Более сложные объекты со многими переменными параметрами не исследовались, так как можно полагать, что идея и техника исследования достаточно ясны из изложенного. Нужно еще добавить, что если характер зависимостей СКП от различных параметров не представляет интереса и требуется лишь подобрать наиболее выгодные значения параметров, то описанное выше последовательное снятие ряда кривых является излишним. Достаточно, визуально наблюдая отсчеты по прибору и изменяя в произвольном порядке значения параметров, добиваться минимального отсчета, при достижении наименьшего значения СКП можно зафиксировать подобранные значения параметров.

Л и т е р а т у р а

1. А. А. Харкевич. Журн. техн. физ., 7, 10, 1937, 515—530, 1083—1095.
2. Т. Е. Шу. Четырехполосники и фильтры, § 29.
3. V. Bush. Operational circuit analysis. N. Y., J. Wiley and Sons, 1929.
4. Г. В. Брауде. Журн. техн. физ., 4, 9, 1934, 1714—1739.
5. О. В. Лурье. Журн. техн. физ., 6, 2, 1936, 319—343.
6. G. D. Robinson. Proc. IRE, 21, 6, 1933, 833—843.
7. E. C. Wente, A. L. Thuros. Mowing coil telephone receivers and microphones. JASA, 3, 1, 1931, 44—55.
8. Л. Я. Гутин. К теории электродинамического микрофона. Журн. техн. физ., 6, 11, 1936, 1885—1904.

О КОРРЕКЦИИ ИСКАЖЕНИЙ, ВНОСИМЫХ ОПТИЧЕСКОЙ ЩЕЛЬЮ

В систему записи и воспроизведения в звуковом кино, а также в тракт передачи и приема в телевидении, бильдтелеграфии и т. п. входит читающий орган, в дальнейшем называемый кратко щелью. Наиболее существенным признаком щели является ее размер, или, точнее, время пробега по ней. Время пробега определяет вносимые щелью искажения. Идеальная щель должна была бы обладать бесконечно малым временем пробега, т. е. бесконечно малой длиной (под длиной понимается размер в направлении развертки).

Естественно стремиться к коррекции искажений, вносимых щелью. Настоящая заметка посвящена рассмотрению различных возможностей коррекции.

Характеристики щели. Характеристикой щели можно назвать отношение светового потока к величине, характеризующей запись. Для прямоугольной щели это отношение имеет следующее символическое выражение:

$$h = \frac{\text{sh } D\tau}{D\tau} e^{-D\tau}. \quad (1)$$

Постоянные множители опущены; начало координат выбрано у первого края щели. Обозначены: D — оператор Хевисайда; $2\tau = l/v$ — время пробега щели; l — длина щели; v — скорость развертки. Частотную характеристику щели получим, подставив $j\omega$ вместо D в формулу (1) и взяв модуль полученного выражения

$$A(\omega) = \sin \omega\tau / \omega\tau. \quad (2)$$

Переходная функция щели проще всего может быть получена непосредственным рассуждением, если рассмотреть явление, состоящее в прохождении перед щелью границы между двумя различными плотностями.

Аналитически переходную функцию получим так

$$\frac{\text{sh } D\tau}{D\tau} e^{-L\tau} = \frac{1}{2D\tau} (1 - e^{-2D\tau}) \leftarrow \div \frac{t}{2\tau} 1 - \frac{t-2\tau}{2\tau} 1_{2\tau}, \quad (3)$$

$$\text{т. е. } h(t) = \begin{cases} \frac{t}{2\tau} [0 < t < 2\tau], \\ 1 [2\tau < t < \infty]. \end{cases} \quad (4)$$

Коррекция при помощи резонансной системы. Старый метод коррекции, основанный на рассмотрении частотной характеристики (2), состоял в том, что падение частотной характеристики при приближении к первому корню, т. е. к значению 1, компенсировалось подъемом характеристики корректирующего звена в виде резонансной системы. В звуковом кино в качестве такой системы применялись особые контуры в схеме усилителя, либо использовались резонансные свойства рекордера [1]. Установилась даже своего рода норма: задавался подъем характеристики корректирующего звена на 18 дб. Считая, что при этом можно скорректировать такой же завал, можно найти граничную частоту скорректированной характеристики. Для этого послужит уравнение

$$20 \lg \frac{\sin \omega_0 \tau}{\omega_0 \tau} = -18,$$

откуда $\omega_0 \approx 2,78 / \tau = 0,88 \omega_1$. Уравнение характеристики корректирующего звена имеет вид

$$A_1(\omega) = \frac{1}{\sqrt{(1 - \omega^2/\omega_0^2)^2 + 4k^2\omega^2/\omega_0^2}}. \quad (5)$$

Здесь обозначение $k = \alpha / \omega_0$, где α — показатель затухания. Так как затухание очень мало, то приблизительно $A_1(\omega)_{\max} \approx 1/2k = A_1(\omega_0)$. Отсюда находим значение k : $k = 0,063$.

Теперь перепишем выражение (2) для частотной характеристики щели в виде

$$A(\omega) = \frac{\sin \pi \frac{\omega}{\omega_1}}{\pi \omega / \omega_1} \quad (6)$$

и для частотной характеристики корректора

$$A_1(\omega) = \frac{1}{\sqrt{(1 - 1,29\omega^2/\omega_1^2)^2 + 2,06 \cdot 10^{-2}\omega^2/\omega_1^2}}. \quad (7)$$

На рис. 1 построены частотные характеристики щели по формуле (6) (кривая *a*), корректора по формуле (7) (кривая *b*) и результирующая характеристика (кривая *c*), полученная путем перемножения первых двух. Как видим, вполне удовлетворительная коррекция частотной характеристики таким образом достигнута. Рассмотрим переходные процессы в такого рода системе. В символической форме характеристика скорректированной системы выражается как $f = gh$, где g — характеристика корректора. Для нахождения переходной функции удобнее всего воспользоваться теоремой Бореля. Так как $h(0) = 0$, то

$$f(t) = \int_0^t h(t - \sigma) g(\sigma) d\sigma.$$

Далее

$$\begin{aligned}
 & 0 [t < 0], \\
 h(t) &= \frac{1}{2\tau} (1 - 1_{2\tau}) = \frac{1}{2\tau} [0 < t < 2\tau], \\
 & 0 [t > 2\tau]; \\
 & 0 [\sigma > t], \\
 h(t - \sigma) &= \frac{1}{2\tau} [t - 2\tau < \sigma < t], \\
 & 0 [\sigma < t - 2\tau].
 \end{aligned}$$

Таким образом,

$$\begin{aligned}
 f(t) &= \frac{1}{2\tau} \int_{t-2\tau}^t g(\sigma) d\sigma [t - 2\tau > 0], \\
 & \frac{1}{2\tau} \int_0^t g(\sigma) d\sigma [t - 2\tau \leq 0].
 \end{aligned} \tag{8}$$

Нам нужно найти переходную функцию корректора. Его характеристика в символической форме имеет вид

$$g = \frac{1}{D^2/\omega_0^2 + 2kD/\omega_0 + 1} = \frac{\omega_0^2}{(D - a_1)(D - a_2)},$$

где $a_1 = -\alpha + j\omega'$, $a_2 = -\alpha - j\omega'$.

По таблицам готовых формул находим

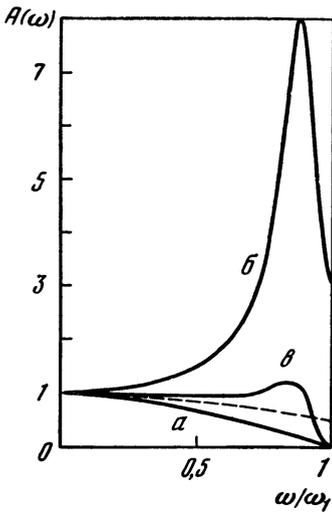
$$g(t) = 1 - e^{-\alpha t} \left(\frac{\alpha}{\omega'} \sin \omega' t + \cos \omega' t \right),$$

или, полагая (вследствие малости α) $\omega' \approx \omega_0$,

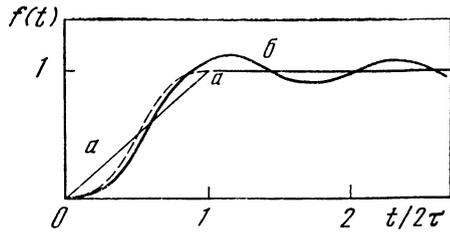
$$g(t) \approx 1 - e^{-\alpha t} (k \sin \omega_0 t + \cos \omega_0 t). \tag{9}$$

Подставляя (9) в (8) и выполняя интегрирование, получим

$$\begin{aligned}
 & \frac{t}{2\tau} - \frac{k}{\omega_0 \tau} - \frac{e^{-\alpha t}}{2\omega_0 \tau} [(1 - k^2) \sin \omega_0 t - 2k \cos \omega_0 t] [0 < t < 2\tau], \\
 f(t) &= 1 - \frac{e^{-\alpha t}}{2\omega_0 \tau} [(1 - k^2) \sin \omega_0 t - 2k \cos \omega_0 t] + \\
 & + \frac{e^{-(\alpha - 2\tau)}}{2\omega_0 \tau} [(1 - k^2) \sin \omega_0 (t - 2\tau) - 2k \cos \omega_0 (t - \\
 & - 2\tau)] [2\tau < t < \infty].
 \end{aligned} \tag{10}$$



Р и с . 1



Р и с . 2 .

Вычисленная по этим формулам переходная функция построена на рис. 2 (кривая б). Ломаная линия (а) представляет переходную функцию некорректированной щели. Очень любопытно отметить, что, несмотря на наличие корректора с очень малым затуханием, колебания с его частотой сравнительно слабо выражены. Это объясняется тем, что колебание, возникающее при переходе первого края щели, гасится другим колебанием, возникающим при переходе второго края и имеющим почти противоположную фазу. Сказанное можно усмотреть из формул (10). Еще лучше уясняется картина явлений, если представить себе, что щель корректируется звеном с нулевым затуханием. Оказывается, что при этом достигаются наивыгоднейшие соотношения.

В самом деле, в этом случае

$$g(t) = 1 - \cos \omega_0 t;$$

$$f(t) = \begin{cases} \frac{1}{2\pi}(\omega_0 t - \sin \omega_0 t) & [0 < t < 2\tau], \\ 1 - \frac{1}{2\pi}[\sin \omega_0 t - \sin(\omega_0 t - 2\pi)] = 1 & [2\tau < t < \infty]. \end{cases} \quad (11)$$

График функций (11) нанесен на рис. 2 штрихом. Интересно также и то, что частотная характеристика в этом случае сохраняет теоретически конечное значение, равное половине при $\omega = \omega_0$, в чем нетрудно убедиться, раскрыв неопределенность

$$\lim_{\omega \rightarrow \omega_0} \frac{\sin \pi \frac{\omega}{\omega_0}}{\pi \frac{\omega}{\omega_0} \left(1 - \frac{\omega^2}{\omega_0^2}\right)}.$$

График этой характеристики нанесен штрихом на рис. 1.

Полная коррекция методом умножения. Возможности описанного в предыдущем параграфе метода, как мы убедились, ограничены. Возникает естественный вопрос: можно ли и как осуществить коррекцию щели, применяя принципы полной коррекции [2]?

Для достижения полной коррекции методом умножения нужно выполнить условие

$$gh = 1, \quad (12)$$

где h — по-прежнему характеристика щели; g — характеристика корректора.

Из (12) имеем

$$g = 1/h = \text{sh } D\tau / D\tau. \quad (13)$$

Здесь предположено, что $h = \text{sh } D\tau / D\tau$, т. е. что начало координат выбрано посредине щели.

Спрашивается, какая физическая система может обладать характеристикой типа (13)? Для ответа на этот вопрос найдем переходную функцию корректора. Мы имеем

$$g = D\tau \frac{1}{\text{sh } D\tau} = 2D\tau e^{-D\tau} (1 - e^{-2D\tau})^{-1} = 2D\tau (e^{-D\tau} + e^{-3D\tau} + e^{-5D\tau} + \dots)$$

и, следовательно,

$$g(t) = 2\tau (I_\tau + I_{3\tau}^* + I_{5\tau}^* + \dots),$$

где I_τ^* означает импульсивную функцию Хевисайда—Дирака с разрывом в точке $t = \tau$. Итак, искомая система обладает тем свойством, что при возбуждении единичной функцией в ней возникает бесконечно длящийся процесс, характеризующийся сериями импульсов, повторяющихся через равные промежутки времени (рис. 3). Подобного рода явление может иметь место в ограниченной волновой системе при полном отражении от ее границ. Рассмотрим, например, соотношения для упругого стержня, по которому может распространяться без затухания волна сжатия, возбуждаемая на одном из концов. Для этого случая, полагая для стержня длины l граничные условия в виде

$$E \frac{dx}{dy} \Big|_{y=0} = -P; \quad \frac{dx}{dy} \Big|_{y=l} = 0,$$

можно получить решение волнового уравнения в следующей операционной форме:

$$\frac{x}{P} = \frac{c}{DE} \cdot \frac{\text{ch } \frac{D}{c} (y-l)}{\text{sh } \frac{D}{c} l},$$

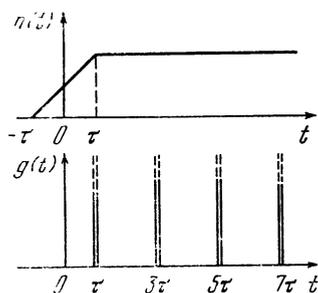


Рис. 3

ИЛИ

$$\frac{v}{p} = \frac{c}{E} \cdot \frac{\operatorname{ch} \frac{D}{c}(y-l)}{\operatorname{sh} \frac{D}{c}l}.$$

В этих формулах E — модуль упругости, c — скорость распространения возмущения по стержню, x — смещение данного

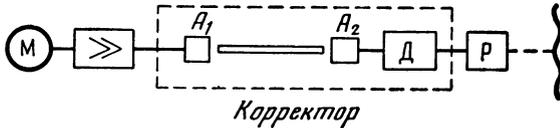


Рис. 4

сечения вдоль координаты y , направленной по оси стержня, dx/dy — деформация, $v=Dx$ — скорость смещения сечения. Физический смысл граничных условий состоит в том, что к концу стержня $y=0$ приложено давление p , а конец $y=l$ свободен. Поведение свободного конца получим, положив $y=l$,

$$\frac{v_l}{p} = \frac{c}{E} \cdot \frac{1}{\operatorname{sh} \frac{D}{c}l}.$$

Мы имеем уже почти то, что нам нужно. Остается еще ввести в тракт дифференцирующий орган, местоположение которого, вообще говоря, совершенно безразлично.

Принципиальная схема записи с коррекцией характеристики щели методом умножения принимает показанный на рис. 4 вид. На этой схеме означают М — микрофон, Р — рекордер, A_1 — прибор, возбуждающий корректирующий стержень, A_2 — прибор, воспринимающий колебание на втором конце стержня, Д — дифференцирующий орган. Возможны многочисленные варианты технического осуществления этой схемы. В качестве волновой системы можно применить не только стержень, но и структуру, что было бы удобно с точки зрения возможности настройки, а также и электрическую систему, что, вероятно, оказалось бы технически наиболее удачным решением. Необходимо только соблюсти равенство времен пробега по корректирующей системе по щели, т. е. выполнить условие $l/c = \tau$.

Это условие должно быть выполнено с большою точностью, что может быть и неплохо, так как позволило бы с соответственной точностью устанавливать ширину щели или скорости развертки, используя корректирующую систему как эталон времени. Нужно добавить, что смягчение границ щели позволило бы смягчить со-

ответственным образом и технические требования к элементам устройства.

Полная коррекция методом сложения. Для осуществления полной коррекции методом сложения нужно подобрать корректирующую систему так, чтобы выполнялось условие

$$h + g = 1, \quad (15)$$

откуда

$$g = 1 - h = 1 - \frac{\text{sh } D\tau}{D\tau} e^{-D\tau}$$

(начало координат у края щели).

Действия сложения и вычитания могут быть перенесены на функции времени, так что если

$$h(t) = \frac{t}{2\tau} I - \frac{t - 2\tau}{2\tau} I_{2\tau},$$

то

$$g(t) = 1 - h(t) = (1 - t/2\tau)(I - I_{2\tau}) \quad (16)$$

(рис. 5). Нам нужно теперь придумать систему, обладающую требуемой переходной функцией (16). Построение такой системы оказывается более трудным для коррекции методом умножения.

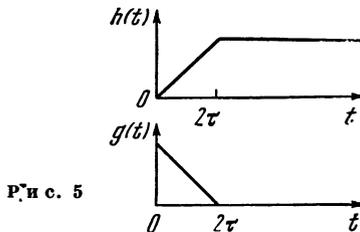


Рис. 5

Переходная функция $g(t)$ обращается в нуль через конечный промежуток времени. Такого рода соотношения могут наблюдаться в **ограниченной** волновой системе. Это общее соображение дает ключ к решению задачи.

В качестве одного из вариантов решения можно составить следующую схему: пусть дана неограниченная волновая система (стержень, струна), на которой в начальный момент создается заданное распределение возмущения. В последующие моменты возмущение побегит с соответствующей скоростью в обе стороны от места возбуждения. Относящиеся сюда соображения с большой подробностью разбираются в руководстве по уравнениям математической физики в связи с решением задачи о поведении струны методом характеристик.

Пусть в данной неограниченной волновой системе в момент времени $t=0$ задано неизвестное пока распределение смещения $x_0 = \varphi(y)$, где y — координата вдоль направления распространения. В фиксированной точке $y=2l$ будет наблюдаться смещение, выражаемое следующей функцией времени:

$$x_{2l}(t) = \frac{1}{2} \varphi(2l - ct),$$

где c — скорость распространения возмущения.

Пусть x_{2l} и есть переменная, определяющая действие устройства; тогда по условию должно быть

$$x_{2l}(t) = g(t) = (1 - t/2\tau)(1 - I_{2\tau}).$$

Отсюда сразу находим

$$\begin{aligned} & 0 \quad [y > 0], \\ \varphi(y) &= \frac{y}{l} \quad [0 < y < 2l], \\ & 0 \quad [2l < y]. \end{aligned}$$

Остается подтвердить правильность намеченной схемы путем составления операционного выражения для ее характеристики. Для этого будем рассматривать каждый элемент корректирующей системы как источник и просуммируем действие всех элементарных источников в точке $y=2l$.

Действие каждого элемента с координатой y можно представить в виде следующей функции:

$$dx = \frac{1}{2} \varphi(y) I_{\sigma}^* d\sigma,$$

где σ — время пробега от рассматриваемого элемента до точки $y=2l$. Таким образом,

$$\sigma = (2l - y)/c \quad [0 < \sigma < 2l/c].$$

В операционной форме

$$dx = \frac{1}{2} \varphi(y) D e^{-D\sigma} d\sigma$$

и соответственно

$$x = \frac{1}{2} \int_0^{2l/c} \varphi(y) D e^{-D\sigma} d\sigma,$$

Переходя к переменной y и подставляя значение $\varphi(y)$, получим

$$x = \frac{e^{-D \frac{2l}{c}}}{2l} \int_0^{2l} y d\left(e^{D \frac{y}{c}}\right) = 1 - e^{-D \frac{l}{c}} \frac{\operatorname{sh} D \frac{l}{c}}{D \frac{l}{c}},$$

что и требовалось доказать.

Принципиальная схема коррекции с использованием изложенных идей показана на рис. 6. Корректор состоит из отрезка

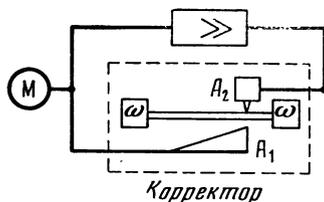


Рис. 6

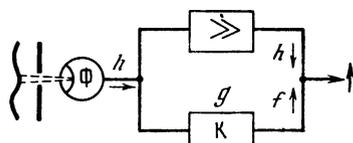


Рис. 7

волновой системы, включенной между двумя волновыми сопротивлениями ω , что обеспечивает в отрезке такой же режим, как и в бесконечной системе. Аппарат A_1 создает в отрезке заданное пространственное распределение начального возмущения; действие этого аппарата может быть основано, например, на создании соответствующего распределения магнитного или электрического поля. Аппарат A_2 воспринимает результирующее возмущение в надлежашей точке отрезка. Обратная связь между A_2 и A_1 через основную цепь предполагается отсутствующей.

Коррекция методом сложения при воспроизведении. Интересные соображения возникают при рассмотрении вопроса о принципиальной применимости коррекции методом сложения при воспроизведении. Дело в том, что при воспроизведении корректор стоит *п о с л е щ е л и*; таким образом, на систему, какова бы она ни была, действует уже искаженный сигнал. Вследствие этого условие полной коррекции придется формулировать совершенно иначе. Принципиальная схема показана на рис. 7. После фотозлемента мы имеем искаженный сигнал h . Корректирующая система k под действием этого сигнала должна дать такую функцию f , чтобы было $h+f=1$. Но $f=hg$. Итак,

$$h(1+g) = 1,$$

откуда

$$g(t) = 1/h - 1 = 2D\tau(1 - e^{-2D\tau})^{-1}.$$

В функции времени

$$g(t) = 2\tau(-I + I^* + I_{2\tau}^* + I_{4\tau}^* + \dots).$$

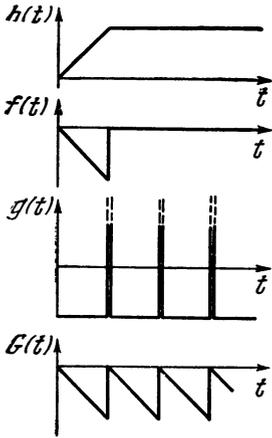


Рис. 8

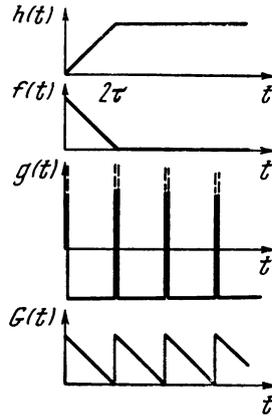


Рис. 9

Предполагая возможность применения дифференциатора, рассмотрим еще функцию

$$G = D^{-1}g = 2\pi(1 - e^{-2D\tau})^{-1},$$

$$G(t) = 2\tau(-1/2\tau - 1 + I_{2\tau} + I_{4\tau} + \dots).$$

Графики всех функций показаны на рис. 8. Как видим, для коррекции в данном случае может быть использована ограниченная волновая система, в которой возбуждается импульс треугольной формы, бегающий затем по системе взад и вперед.

Следует указать на еще один возможный вариант, состоящий в том, что скорректированный сигнал смещается без искажения по оси времен. Это можно записать так:

$$h + f = e^{-D\tau},$$

где T — время запаздывания. В рассматриваемом случае естественно положить $T = 2\tau$. Тогда будет

$$f = e^{-2D\tau}h - \leftarrow \div - \frac{t}{2\tau}(1 - I_{2\tau}),$$

$$g = 2D\tau e^{-2D\tau}(1 - e^{-2D\tau}) - 1 \leftarrow \div 2\tau \left(-\frac{1}{2\tau} + I_{2\tau}^* + I_{4\tau}^* + \dots \right),$$

$$G = 2\tau e^{-2D\tau}(1 - e^{-2D\tau}) - \frac{1}{D} \leftarrow \div 2\tau \left(-\frac{t}{2\tau} + I_{2\tau} + I_{4\tau} + \dots \right)$$

(см. рис. 9).

Применение щели непрямоугольной формы. Все предыдущие соображения относились к щели прямоугольной формы. Соблюдение этой формы вовсе необязательно. Более того, можно указать на крупный и, по-видимому, неустраняемый принципиальный недостаток прямоугольной щели: периодический сигнал, период которого (выраженный в единицах длины) укладывается целое число

раз в длине щели, не дает модуляции, и, следовательно, не может быть ни записан, ни воспроизведен. Этот недостаток может быть устранен путем перехода к другим формам щели. Понятно, что здесь возможны бесчисленные варианты. Рассмотрим один частный случай, приводящий к очень простым соотношениям. Пусть щель имеет экспоненциальную форму, т. е.

$$z = e^{-\alpha x} \quad [0 < x < \infty],$$

где z — ширина щели, или, в более общем виде, прозрачность щели, которая может быть выполнена, например, в форме оптического клина. Переходную функцию такой щели легко вычислить

$$h(t) = \int_0^{\alpha t} e^{-\sigma} d\sigma = 1 - e^{-\alpha t} = 1 - e^{-\beta t}.$$

В операционной форме $h = \beta / (D + \beta)$.

Для коррекции методом умножения нужно иметь корректор с характеристикой

$$g = 1/h = (D + \beta)/\beta.$$

Подобной характеристикой обладает, например, схема из последовательно соединенных R и L , для которой отношение напряжения к току, т. е. переходное сопротивление, выражается формулой

$$U/i = L(D + R/L).$$

Подобрав $R/L = \beta$, получим полную коррекцию

$$hg = \frac{\beta}{D + \beta} L \left(D + \frac{R}{L} \right) = R.$$

Интересно войти в некоторые подробности разбираемого примера. Прежде всего заметим, что мы предположили снятие со схемы напряжения при заданном, т. е. не зависящем от нагрузки, токе. В реальном случае имеется источник (например, лампа) с заданной э. д. с. E и хотя и большим, но конечным внутренним сопротивлением R_i . Таким образом, реально используется отношение не U/i , а U/E . Для последнего имеем:

$$\gamma = \frac{U}{E} = \frac{U}{i} \cdot \frac{i}{E} = \frac{LD + R}{LD + R + R_i} \approx \frac{D + R/L}{D + R_i/L}.$$

Таким образом, с учетом реальных соотношений получим для скорректированной характеристики

$$\gamma h = \frac{D + R/L}{D + R_i/L} \cdot \frac{\beta}{D + \beta} = \frac{\beta}{D + R_i/L}.$$

Выигрыш от применения коррекции сводится, стало быть, к уменьшению постоянной времени в отношении R/R_i . Здесь, как и обычно, качественные и количественные требования всту-

пают в противоречие, так как требуемое усиление должно, очевидно, возрасти в том же отношении.

Второй вопрос относится к форме щели. При выводе она предполагалась экспоненциальной, т. е. бесконечной. Практически, очевидно, возможно ограничить длину щели участком, на протяжении которого прозрачность убывает в определенное число раз, например на 40 дб. Тогда показатель α определится длиной щели, задаваемой, в свою очередь, конструктивными соображениями.

* * *

Не следует, конечно, относиться к изложенным соображениям как к сколько-нибудь технически разработанным предложениям. Цель этой статьи совершенно иная — она состоит в том, чтобы показать некоторые новые пути, открывающие довольно широкие теоретические возможности, те или иные из которых при дальнейшей разработке могут привести и к технически приемлемым решениям.

Изложенные здесь соображения были предварительно доложены на заседании кафедры телевидения ЛЭИС; я очень признателен участникам заседания за сделанные ими во время дискуссии очень полезные для меня замечания.

Л и т е р а т у р а

1. В. А. Бузров. Оптическая запись звука. «Искусство», 1937.
2. Журн. техн. физ., 7, 3, 1937, 250—256.

О ДИАГРАММАХ СИЛА—ПЕРЕМЕЩЕНИЕ В ПРИМЕНЕНИИ К ИССЛЕДОВАНИЮ КОЛЕБАНИЙ

1. Как известно, при исследовании колебательных явлений широко применяются диаграммы, на которых действительное движение представлено некоторой траекторией изображающей точки на так называемой фазовой плоскости, т. е. в координатах скорость—перемещение (координатная система (v, x)).

Возможен, конечно, и другой выбор системы координат. Из имеющихся здесь возможностей представляет интерес координатная система сила—перемещение (F, x) .

Дело в том, что при решении многих проблем, в частности проблем, относящихся к установлению автоколебаний, основным физическим критерием является критерий энергетический. Именно из рассмотрения энергетического баланса мы можем сделать непосредственные заключения об амплитуде установившихся колебаний, их устойчивости и т. п.

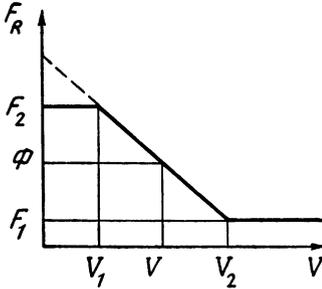
Диаграммы в системе (F, x) обладают как раз тем свойством, что энергетические соотношения интерпретируются ими особенно наглядно. В самом деле, площадь, замкнутая некоторым контуром на плоскости (F, x) , равна работе, причем обход контура по часовой стрелке соответствует энергии, затраченной системой, а обход его в противоположном направлении — энергии, полученной системой. Диаграммы (F, x) могут иметь разный смысл в зависимости от того, какую силу мы примем за координату. Именно, в качестве F можно выбрать как внешнюю силу, так и одну из реакций системы, или же сумму нескольких таких реакций. Выбор переменных определяется, понятно, особенностями задачи.

Замкнутый цикл на плоскости (F, x) в общем случае должен содержать не менее двух петель с противоположным направлением обхода. Таким образом, простейшая диаграмма (F, x) есть фигура типа восьмерки. Если диаграмма изображает поведение автоколебательной системы, то при изменении амплитуды обе петли, в общем случае, будут изменяться неодинаковым образом. Можно подобирать такие соотношения, при которых площади петель сравняются. При этом условии, очевидно, достигается баланс между вложенной и затраченной энергиями; это условие соответствует, таким образом, колебаниям со стационарной амплитудой.

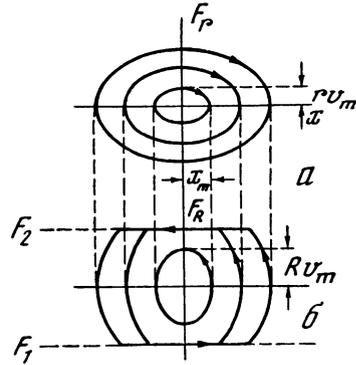
Для суждения об устойчивости следует увеличить (уменьшить) амплитуду колебания. Если при этом окажется, что потери энергии растут быстрее (убывают медленнее), чем полученная извне

энергия, то это значит, что колебания устойчивы. График таких колебаний на обычной фазовой плоскости представляет собою устойчивый предельный цикл.

2. В качестве примера рассмотрим, основываясь на диаграммах (F, x) , действие маятника Фруда, причем будем решать задачу в сильно схематизированной форме. А именно: предположим, во-первых, что движение маятника несущественно отличается от



Р и с. 1



Р и с. 2

синусоидального ¹, во-вторых, аппроксимируем характеристику трения ломаной линией, изображенной на рис. 1.

Аналитическое представление этой характеристики таково:

$$\bar{F}_R = \begin{cases} \bar{F}_2 & [0 < \bar{v} < \bar{v}_1], \\ \bar{F}_3 - \bar{v}R & [\bar{v}_1 < \bar{v} < \bar{v}_2], \\ \bar{F}_1 & [\bar{v}_2 < \bar{v} < \infty], \end{cases} \quad (1)$$

где $\bar{F}_3 = \bar{F}_1 - \bar{v}_2 R$.

Постоянная скорость V предположена равной полусумме скоростей \bar{v}_1 и \bar{v}_2 , т. е. рабочая точка лежит посередине прямолинейного падающего участка характеристики.

Черточками вверху обозначены относительные скорости и полные силы, т. е.

$$\bar{v} = v + V, \quad \bar{F} = F + \Phi,$$

где $\Phi = (\bar{F}_1 + \bar{F}_2)/2$; V — скорость вала; v — скорость маятника (отсчитываемая в противоположную сторону).

Если заменить переменные, взяв вместо \bar{v} и \bar{F} $v = \bar{v} - V$ и $F = \bar{F} - \Phi$, то характеристика трения сместится так, что начало

¹ Заметим, в частности, что при таком предположении диаграммы (v, x) становятся совершенно бесполезными.

координат перейдет в рабочую точку и формулы (1) примут вид

$$F_R = \begin{cases} F_2 & [-V < v < v_1], \\ -vR & [v_1 < v < v_2], \\ F_1 & [v_2 < v < \infty]. \end{cases} \quad (1a)$$

Мы уже сделали исходное предположение о характере движения. Поэтому диаграмма ничего нового в этом отношении нам дать не может. Для рассмотрения же вопросов энергетического баланса в данном случае удобнее разбить диаграмму на две, одна из которых будет изображать работу внешней силы в координатах (F_R, x) , вторая — затраченную на линейном трении системы работу в координатах (F_r, x) . Сила сопротивления, обусловленного линейным трением, равна

$$F_r = rv. \quad (2)$$

Так как движение синусоидально, то график (2) будет представлять собой семейство эллипсов с полуосями x_m и rv_m , где x_m и v_m — соответственно амплитуды перемещения и скорости (рис. 2, а).

Что же касается (1), то, пока $F_R < (F_2 - F_1)/2 = F_0$, график этой функции также будет представляться эллипсами с полуосями x_m и Rv_m . Вертикальная полуось этих эллипсов (рис. 2, б) больше, чем у эллипсов рис. 2, а, так как по смыслу задачи $R > r$ (в противном случае автоколебания вообще невозможны).

Возможные значения F_R ограничены, однако, пределами F_1 и F_2 , как это показано на рис. 2, б. При дальнейшем возрастании амплитуды график F_R представляется поэтому усеченным эллипсом.

Рассмотрим теперь вопрос об устойчивости колебаний. При малых амплитудах энергия, затраченная за цикл, равна

$$A_r = \pi r v_m x_m, \quad (3)$$

тогда как энергия, вложенная в систему за цикл,

$$A_R = \pi R v_m x_m. \quad (4)$$

Так как $R > r$, то $A_R > A_r$, и колебания нарастают.

При некоторой амплитуде (которую мы определим в дальнейшем) достигается баланс энергии: $A_R = A_r$. При весьма больших амплитудах можно положить приближенно

$$A_R \approx 4x_m F_0, \quad (5)$$

т. е. A_R растет пропорционально первой степени x_m , а A_r по-прежнему равно

$$A_r = \pi r v_m x_m = \pi \omega r x_m^2,$$

т. е. возрастает пропорционально квадрату амплитуды. В рассматриваемом случае стационарные колебания будут поэтому устойчивыми.

Для того чтобы найти амплитуду этих колебаний, нужно составить уравнение баланса энергии.

Площадь усеченного эллипса выражается формулой

$$S = 2ab \left[\frac{d}{b} \sqrt{1 - \frac{d^2}{b^2}} + \arcsin \frac{d}{b} \right] \quad (9)$$

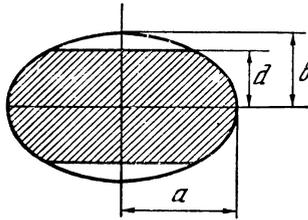
(обозначения приведены на рис. 3). В нашем случае

$$a = x_m, \quad b = Rv_m, \quad d = F_0.$$

Воспользовавшись этим и приравнивая (6) и (3), получим уравнение баланса

$$f(z) = \frac{2}{\pi} (z \sqrt{1 - z^2} + \arcsin z) = r/R, \quad (7)$$

где $z = d/b = F_0/Rv_m$.



Р и с. 3

Амплитуда скорости определяется из уравнения (7), если задана характеристика трения (F_1 , F_2 и R) и параметр r . Для облегчения решения задачи на рис. 4 приведен график функции $f(z)$, так что, зная r/R , можно сразу найти соответствующее значение z , а затем и значение

$$v_m = F_0/Rz. \quad (8)$$

Заметим, что при малых z (или при малых r/R) мы имеем приближенное равенство

$$\frac{4}{\pi} z \approx \frac{r}{R}, \quad (9)$$

откуда

$$v_m \approx \frac{4}{\pi} \cdot \frac{F_0}{r}. \quad (10)$$

Смысл последнего результата очень прост: мы перешли к предельным соотношениям, при которых характеристика трения вырождается, принимая разрывную форму (рис. 5, а). При этом, очевидно, всем значениям скорости $v > V$ будет соответствовать $F_R = F_1$, а всем значениям $v < V$ $F_R = F_2$ (рис. 5, б). Система колеблется с собственной частотой. Эти колебания можно трактовать как вынужденные колебания, находящиеся в резонансе с вынуждающей силой, амплитуда первой гармоники которой равна $F_m = 4F_0/\pi$. Это и есть первая гармоника разложения F_R (рис. 5, б) в ряд Фурье. Высшие гармоники не должны приниматься в расчет,

так как с самого начала мы предположили движение синусоидальным.

Остается коснуться вопроса о к. п. д. системы. Будем полагать, что сопротивление r есть полезная колебательная нагрузка. Тогда полезная мощность равна $P_r = v_m^2 r/2$. Мощность на валу двигателя примем равной произведению постоянных составляющих силы и скорости, т. е. положим

$$P_0 = V\Phi,$$

где $V = (v_1 + v_2)/2$, $\Phi = (\bar{F}_1 + \bar{F}_2)/2$.

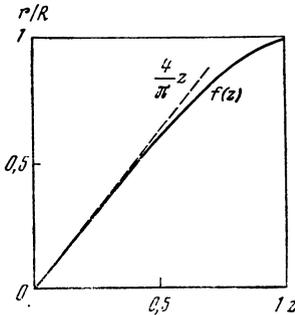


Рис. 4

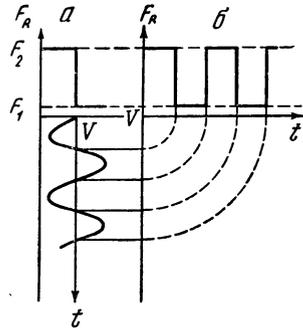


Рис. 5

При малых к. п. д. P_0 не зависит от колебательной нагрузки (т. е. реакции на двигатель малы); к. п. д. будет равен отношению $\eta = P_r/P_0 = v_m^2 r/2V\Phi$. Подставляя значение v_m из (8), найдем, $\eta = F_0^2 r/2R^2 z^2 V\Phi$, а воспользовавшись приближенным соотношением (9), получим

$$\eta \approx \frac{8}{\pi^2} \cdot \frac{E_0^2}{rV\Phi}.$$

В рамках сделанных допущений к. п. д. зависит, таким образом, не от крутизны падающего участка характеристики трения, а только от координат рабочей точки (V , Φ) и перепада силы трения ($F_0 = (F_2 - F_1)/2$).

3. Возвращаясь к общим положениям п. 1, построим теперь диаграмму движения в координатах (F , x), где $F = F_r + F_R$. Аналитическое выражение F имеет следующий вид:

$$\begin{aligned} F_2 + rv & \quad [-V < v < v_1], \\ F = (r - R)v & \quad [v_1 < v < v_2], \\ F_1 + rv & \quad [v_2 < v < \infty] \end{aligned}$$

(см. рис. 6)

Для построения диаграммы нужно, как известно, исключить параметр t . Это сделано графически. На рис. 7 приведено построение функции $F(t)$ по результирующей характеристике $\bar{F}(v)$ и заданной синусоидальной скорости $v(t)$. На том же графике под

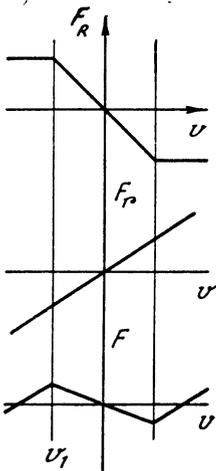


Рис. 6

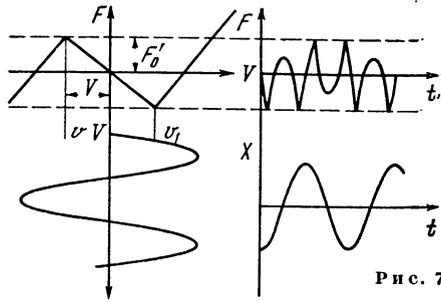


Рис. 7

кривой $F(t)$ нанесена синусоида $x(t)$. Беря последовательно соответствующие пары значений F и x , получаем диаграмму рис. 8, состоящую из трех замкнутых контуров. Средний изображает энергию, затрачиваемую системой, два крайних — энергию, получаемую от двигателя. Рис. 7 и 8 соответствуют установившемуся колебанию системы при следующих значениях параметров: $z=0,5$; $f(z)=0,61=r/R$.

4. Перейдем теперь к рассмотрению случая, когда экспериментальная характеристика трения аппроксимирована некоторой аналитической функцией.

Здесь выясняется существенное преимущество диаграмм (F, x) и связанных с ними энергетических представлений с чисто вычислительной точки зрения. Для построения диаграммы (F, x) так же, как и диаграммы (v, x) , нужно исключить время их параметрически заданных выражений $F(t)$ и $x(t)$ (или $v(t)$ и $x(t)$).

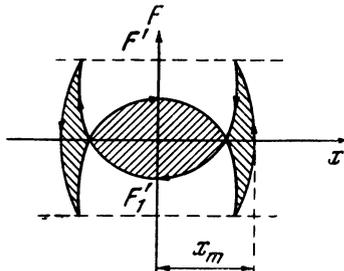


Рис. 8

Сделать это аналитически нелегко, так как задача сводится к решению системы трансцендентных уравнений.

Если же ограничиться вычислением значений энергий, то все затруднения отпадают, так как выражения для силы и скорости

могут быть непосредственно проинтегрированы в параметрической форме. Действительно, энергия равна

$$A = \oint F dx = \int_0^T F(t) \frac{dx}{dt} dt.$$

Пусть, к примеру, характеристика трения, отнесенная к абсолютной скорости и переменной составляющей силы, задана полиномом третьей степени, как это принято у Стрелкова¹,

$$F_R = av + bv^2 + cv^3. \quad (11)$$

Полагая, по-прежнему,

$$v = v_m \sin \omega t, \quad x = -x_m \cos \omega t,$$

получим

$$F_R(t) = av_m \sin \omega t + bv_m^2 \sin^2 \omega t + cv_m^3 \sin^3 \omega t.$$

Прибавляя

$$F_r(t) = rv = rv_m \sin \omega t,$$

получим полную силу

$$F(t) = F_R(t) + F_r(t).$$

Полная энергия за цикл должна равняться нулю, в чем и заключается условие баланса энергии

$$\int_0^{x_m} F dx = \int_0^{\pi/2} F_r(t) dx(\omega t) = 0.$$

Мы имеем

$$dx(\omega t) = x_m \sin \omega t d(\omega t)$$

и уравнение

$$\int_0^{\pi/2} [(a+r) \sin^2 \omega t + bv_m \sin^3 \omega t + cv_m^2 \sin^4 \omega t] d(\omega t) = 0,$$

которое дает

$$v_m^2 + \frac{32}{9\pi} \cdot \frac{b}{c} v_m + \frac{4}{3} \cdot \frac{a+r}{c} = 0.$$

Отсюда амплитуда скорости

$$v_m = \frac{16}{9\pi} \cdot \frac{b}{c} + \sqrt{\left(\frac{16}{9\pi} \cdot \frac{b}{c}\right)^2 - \frac{4}{3} \cdot \frac{a+r}{c}}.$$

У Стрелкова влияние члена, содержащего v^2 , выпадает, по-видимому, вследствие несовершенства примененного им метода. Мы придем к результату Стрелкова, положив $b = 0$

$$v_m = \sqrt{-\frac{4}{3} \cdot \frac{a+r}{c}}.$$

¹ С. П. Стрелков. Маятник Фроуда. Журн. техн. физ., 3, 1933, 563—573.

ВЫНУЖДЕННЫЕ КОЛЕБАНИЯ СИСТЕМЫ, УДАРЯЮЩЕЙСЯ ОБ ОГРАНИЧИТЕЛЬ¹

1. Если система с одной степенью свободы ограничена в своем движении таким образом, что движущаяся масса ударяется о неподвижное тело, то, как показывает непосредственный опыт, период свободных колебаний системы уменьшается.

Отсюда можно заключить, что в случае вынужденных колебаний такой системы резонанс будет иметь место при частоте, отличной от собственной частоты аналогичной системы без ограничителя. Введение ограничителя есть по существу введение некоторой связи. На основании общих теорем динамики следует поэтому ожидать повышения собственных частот системы. При этом нужно, конечно, иметь в виду, что после введения ограничителя рассматриваемая система становится нелинейной.

Вопрос о колебаниях системы с ограничителем имеет прикладное значение. Так, например, рабочие орудия типа резонансных электрических отбойных молотков по своей динамической схеме сводятся именно к такого рода системе.

2. Будем рассматривать модель, представленную схематически на рис. 1, при следующих обычных предположениях:

а) внешняя сила синусоидальна;

б) удар происходит практически мгновенно и характеризуется коэффициентом восстановления скорости $R < 1$;

в) при свободном движении (в интервале между ударами) потери отсутствуют;

г) возможно устойчивое периодическое движение с периодом, равным периоду вынуждающей силы.

Примерный характер получаемого движения иллюстрируется рис. 2, на котором сверху приведен график перемещения, внизу — график скорости.

3. Уравнение движения, описывающее поведение системы в интервале между ударами, гласит

$$m\ddot{x} + sx = F_m \cos(\omega t + \varphi). \quad (1)$$

Чтобы учесть влияние удара, введем условия для начала и конца этого интервала, длительность которого, по нашему предположению, равна периоду вынуждающей силы

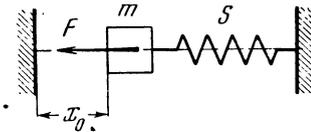
$$\begin{array}{ll} \text{при } t = 0 & \text{при } t = T \\ x = -x_0, & x = -x_0, \\ v = v_2 = Rv_1; & v = -v_1. \end{array} \quad (2)$$

¹ Совместно с И. Г. Русаковым.

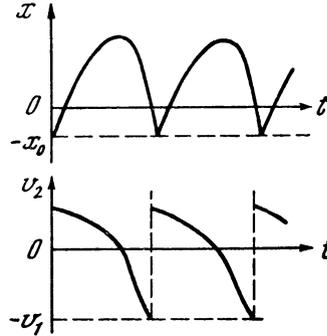
Решение уравнения (1) имеет обычный вид

$$x = x_{1m} \cos(\omega_0 t + \psi) + x_{2m} \cos(\omega t + \varphi). \quad (3)$$

Первый член описывает здесь собственное колебание, амплитуда которого зависит от начальных условий, второй член — вынужденное колебание, амплитуда которого зависит от внешней



Р и с. 1



Р и с. 2

силы. Выражение (3) совместно с условиями (2) дает полное решение поставленной задачи. Мы получаем

$$x_{1m} = \frac{v_1 + v_2}{2} \cdot \frac{1}{\omega_0 \sin \pi \zeta}, \quad x_{2m} = -\frac{F_m}{\omega^2 m} \cdot \frac{1}{1 - \zeta^2}, \quad (4)$$

$$\psi = -\pi \zeta, \quad \varphi = \arcsin \frac{v_1 - v_2}{2} \cdot \frac{1}{\omega x_{2m}},$$

где для краткости обозначено $\omega_0/\omega = \zeta$ ($\omega_0^2 = s/m$).

4. Найдем теперь выражение для скорости при ударе, обозначенной у нас через v_1 . Мы имеем из (3)

$$-x_0 = x_{1m} \cos \psi + x_{2m} \cos \varphi.$$

Подставляя сюда (4) и $v_2 = Rv_1$, получим квадратное уравнение

$$v_1^2 (1 + f^2) + v_1 4x_0 \frac{\omega}{1 - R} f + 4 \frac{\omega^2 (x_0^2 - x_{2m}^2)}{(1 - R)^2} = 0,$$

которое дает

$$\Phi = \frac{v_1}{\omega x_1} = \frac{2}{1 - R} \cdot \frac{\lambda f + \sqrt{\frac{1 + f^2}{(1 - \zeta^2)^2} - \lambda^2}}{1 + f^2}. \quad (5)$$

Здесь введены сокращенные обозначения

$$x_1 = \frac{F_m}{\omega^2 m}; \quad \lambda = -\frac{x_0}{x_1}; \quad f = f(\zeta, R) = \frac{1 + R}{1 - R} \cdot \frac{\text{ctg } \pi \zeta}{\zeta}.$$

5. Прежде всего нужно заметить, что устойчивые периодические колебания возможны не при любых значениях λ и ζ . Так как

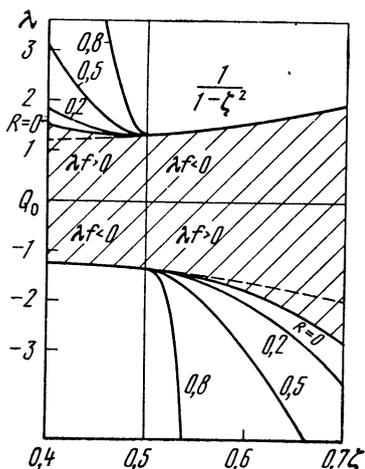


Рис. 3

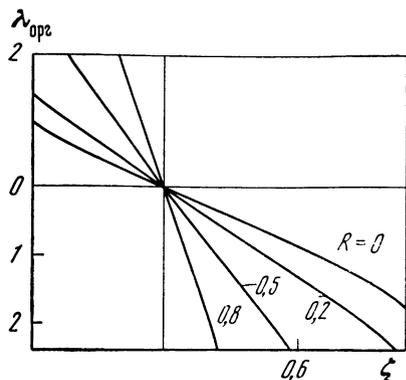


Рис. 4

v_1 и x_1 , а следовательно, и Φ — вещественные и существенно положительные величины, то рассмотрение (5) приводит к следующим условиям, ограничивающим значения λ :

$$\lambda \leq \frac{\sqrt{1+f^2}}{1-\zeta^2} \quad (\text{граница вещественности } \Phi \text{ при } \lambda f > 0) \quad (6)$$

$$\lambda \leq \frac{1}{1-\zeta^2} \quad (\text{граница положительных значений } \Phi \text{ при } \lambda f < 0).$$

На рис. 3 приведены графики функций (6) для различных значений R и интервала $0,4 < \zeta < 0,7$. Заштрихованная область соответствует значениям λ и ζ , для которых колебания возможны при любом R .

Следует, может быть, пояснить, что положительные значения λ отвечают тому случаю, когда между начальным положением массы и ограничителем существует зазор длиной x_0 , а отрицательные значения λ — случаю, когда зазор отсутствует и масса прижата к ограничителю упругой силой пружины, сжатой предварительно на величину x_0 . Таким образом, $\lambda=0$ означает, что масса касается ограничителя без нажима.

6. Займемся теперь рассмотрением экстремальных значений функции (5). Из уравнения $\partial\Phi/\partial\lambda=0$ получим

$$\lambda_{\text{opt}} = \frac{f}{1-\zeta^2}. \quad (7)$$

Подстановка (7) в (5) дает

$$\Phi_{\text{max}} = \frac{2}{1-R} \left| \frac{1}{1-\zeta^2} \right|. \quad (8)$$

Формула (7) позволяет подобрать наиболее выгодное значение λ при данном неизменном ζ . Графики $\lambda_{\text{opt}}(\zeta)$ приведены на рис. 4. Заметим, что λ зависит от F_m , так что оптимальные соотношения

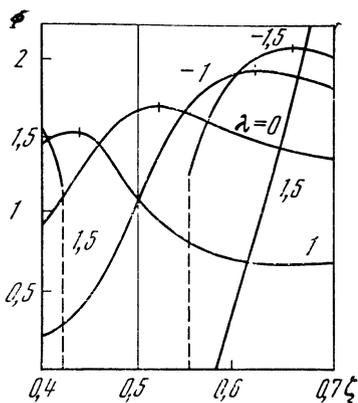


Рис. 5

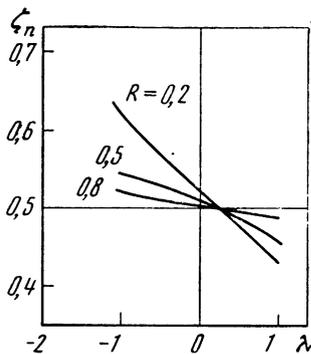


Рис. 6

достигаются как путем изменения x_0 , так и путем изменения F_m , что характерно для нелинейной системы.

Обратимся к изменениям ζ . При неизменном λ зависимость $\Phi(\zeta)$ дает резонансные кривые. Семейство таких кривых для $R=0,2$ изображено на рис. 5. На рис. 5 нанесены также ветви характеристики, соответствующие значениям $\lambda=-1,5$ и $\lambda=1,5$. Как это следует из рис. 3, колебания при последнем значении параметра λ невозможны в интервале $0,44 < \zeta < 0,58$.

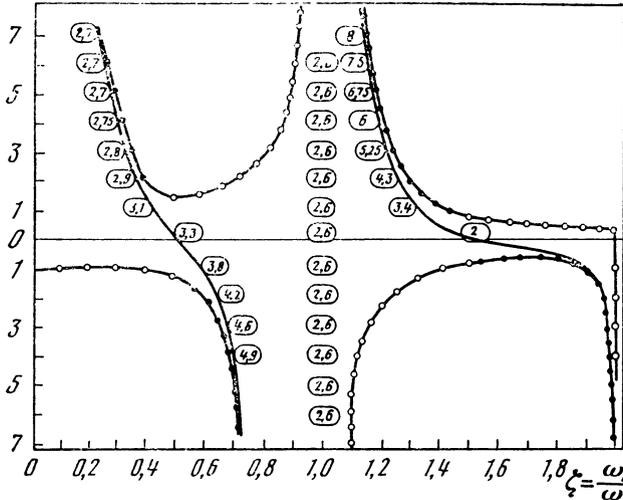
Уравнение $\partial\Phi/\partial\zeta=0$ дает значение ζ_{opt} . Графики функций $\zeta_{opt}(\lambda)$ построены на рис. 6. (Точность этих графиков невелика, так как они получены путем непосредственного вычисления резонансных кривых.)

Для рассмотрения полученных зависимостей в более широких пределах изменения ζ на рис. 7 нанесен график функций (6) для значения $R=0,2$ (— линия максимумов по λ , ○—○— граница нуля скорости, ●—●— граница вещественности скорости, цифры в кружочках — значения скорости Φ). На этом графике изображена область существования периодического решения, ограниченная вышеуказанными условиями. Кроме кривых, изображающих границы, на графике построены также кривые значений λ , соответствующих максимуму функции (5) по λ (численные значения этих максимумов проставлены на графике).

Следует заметить, что максимумы Φ по λ и по ζ , вообще говоря, не совпадают, но в пределах не особенно больших значений λ расхождение между ними незначительно. Полученные кривые оптимальных значений λ и ζ служат количественной иллюстрацией высказанного вначале общего положения о повышении резонансной частоты вследствие наличия ограничителя.

7. Выбор функции (5) предполагает, что нас интересует значение скорости удара v_1 при вариациях настройки (т. е. величины ζ)

$$\lambda = \frac{x_0 \omega^2 m}{F m}$$



Р и с. 7

путем изменения собственной частоты системы ω_0 . При этом частота вынуждающей силы ω предположена неизменной.

Если же настройка производится изменением ω , то вместо (5) удобнее взять функцию

$$\Psi = \frac{v_1}{\omega_0 x_s} \zeta \Phi = \frac{2}{1-R} \cdot \frac{1}{\zeta} \cdot \frac{\delta f + \sqrt{\left(\frac{\zeta}{1-\zeta^2}\right)^2 (1+f^2) - \delta^2}}{1+f^2}, \quad (9)$$

где $x_s = F_m / \omega_0^2 m = F/s$; $\delta = x_0/x_s$.

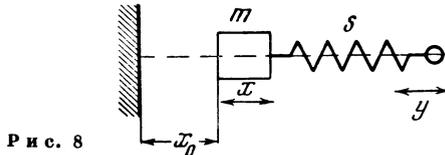
Очевидно, что зависимость $\lambda_{\text{опт}}$ от ζ получается несколько иной, чем та, которую мы имели ранее. Общий характер соотношений, однако, сохраняется.

8. Рассмотрим теперь другую динамическую модель. Пусть задана не сила, приложенная к массе, а перемещение второго конца пружины (рис. 8). И в этом случае общее решение уравнения движения выражается формулой (3); однако, если задано $v = y_m \cos(\omega t + \varphi)$, то $x_{2m} = -y_m \zeta^2 / (1 - \zeta^2)$, тогда как раньше мы имели $x_{2m} = -x_s \zeta^2 / (1 - \zeta^2)$.

Рассматриваемый случай приводится, таким образом, к предыдущему путем простой замены x_s на y_m . Так, например, в формуле (9) нужно положить $\delta = x_0/v_m$.

9. Остается обсудить найденные соотношения. В отличие от простой линейной системы настройка системы с ограничителем может быть осуществлена как непосредственным изменением отношения собственной частоты ($\omega_0 = \sqrt{s/m}$) системы к рабочей частоте, так и изменением расстояния ограничителя от положения равновесия системы, или, вернее, отношения этого расстояния

к амплитуде вынуждающей силы (x_0/F_m). Влияние последнего фактора настройки особенно значительно вблизи $\zeta = \omega_0/\omega = 1/2$ и $\zeta = 1^{1/2}$; при небольших свободных ходах ударяющей массы, а также при небольших начальных нажимах ограничителя на массу (в случае $\lambda < 0$) или при работе со значительными силами F_m здесь выгодно выбирать условия настройки. Условия вблизи $\zeta = 1/2$ рассмотрены подробнее выше. Следует отметить, что ре-



Р и с. 8

жимы $\zeta = 1/2$ и $\zeta = 1^{1/2}$ выгодны также и с точки зрения независимости настройки от величины R .

Сравнивая области вблизи $\zeta = 1/2$ и $\zeta = 1^{1/2}$, можно подметить, что при очень небольших значениях λ выгоднее работать вблизи $\zeta = 1/2$. Если конструктивными соображениями допускаются сравнительно большие (до $\lambda = 1^{1/2}$) отводы ограничителя, то можно получить некоторый выигрыш в скорости удара, настраиваясь на резонанс при $\zeta \approx 1,3 \div 1,5$. Однако большого выигрыша таким путем получить не удастся, и, вообще, надежнее работать без начального отодвигания ограничителя или даже при небольшом начальном нажиме ограничителя при $\zeta \approx 1/2$. Наконец, если особенно важно освободиться от рассматриваемого влияния величины силы F_m или положения ограничителя, может оказаться более удобным работать при настройке на резонанс в обычном линейном смысле ($\zeta = 1$), когда теоретически влияние λ на скорость удара не должно иметь места.

О РАСЧЕТЕ ОДНОГО РОДА КОРРЕКТИРУЮЩИХ СХЕМ

В технике применяются корректирующие схемы, имеющие назначение уменьшить искажения, вносимые тем или иным аппаратом. В частности, предлагались схемы для коррекции осциллографа [1].

Принципиальная возможность коррекции ясна, однако расчет корректирующей схемы, т. е. определение ее параметров по заданным допускам на искажения, возможен, вообще говоря, лишь тогда, когда задан характер записываемого процесса.

В некоторых случаях оказывается возможным свести проблему коррекции к исследованию изменения определенных свойств аппарата в результате коррекции, не затрагивая вопроса о характере регистрируемых явлений. В этих случаях вопрос о расчете может быть поставлен и доведен до окончательного разрешения, исходя только из свойств самого аппарата и корректирующей схемы.

Коррекция осциллографа путем увеличения упругости. В осциллографии принято считать, что шлейф осциллографа дает искажения тем меньшие, чем выше собственная частота шлейфа (предполагается, что шлейф надлежащим образом задемпфирован). Это положение легко вывести из рассмотрения уравнения шлейфа. Полагая, что шлейф может рассматриваться как система с одной степенью свободы, имеем

$$(mp^2 + rp + s)x = Ni_k, \quad (1)$$

где p — оператор Хевисайда; m , r и s — соответственно приведенные масса, трение и упругость; x — перемещение шлейфа (т. е. светового луча, если запись производится фотографически); N — коэффициент электромеханической связи; i_k — ток, протекающий через шлейф.

Уравнение (1) может быть переписано в виде

$$(p^2 + 2\alpha p + \omega_0^2)x = \frac{N}{m}i_k. \quad (2)$$

Здесь α — коэффициент затухания; ω_0 — собственная частота (без поправки на затухание). Если член, содержащий ω_0^2 , велик по сравнению с остальными, то приблизительно

$$\omega_0^2 x \cong \frac{N}{m}i_k \quad \text{или} \quad x \cong \frac{N}{m\omega_0^2}i_k = \frac{N}{s}i_k. \quad (3)$$

Соотношение (3) служит выражением неискаженной записи.

Таким образом, увеличение ω_0 , вообще говоря, уменьшает вносимые шлейфом искажения. Увеличение ω_0 возможно путем увеличения s и уменьшения m . Последнее имеет практический предел, и значительное повышение возможно главным образом за счет увеличения s .

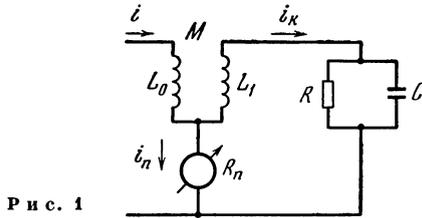
Здесь нужно, однако, иметь в виду, что при увеличении s падает статическая чувствительность шлейфа, которую мы определим как

$$\gamma_0 = \left. \frac{x}{i} \right|_{p=0}.$$

В нашем случае

$$\gamma_0 = N/s.$$

Посмотрим, как изменится чувствительность при изменении собственной частоты в заданное число раз. Пусть первоначальная



Р и с. 1

упругость есть s_1 , а измененная упругость s_2 . Остальные параметры остаются неизменными.

Мы имеем $\gamma_1 = N/s_1$; $\gamma_2 = N/s_2$; $\omega_1^2 = s_1/m$; $\omega_2^2 = s_2/m$. Вводя обозначения $\gamma_2/\gamma_1 = \varepsilon$; $\omega_2/\omega_1 = \mu$, получаем

$$\varepsilon = 1/\mu^2. \quad (4)$$

Таким образом, повышение собственной частоты путем увеличения упругости ведет к значительному уменьшению чувствительности.

Схема Ирвина. Одна из схем для коррекции осциллографа, предложенная Ирвином [2], изображена на рис. 1. Суть дела заключается в том, что форма тока i_k отличается от формы тока i , и притом как раз так, что запись x воспроизводит ток i с достаточной точностью, несмотря на несовершенство механической части осциллографа.

Составим уравнение схемы. Пользуясь законами Кирхгофа, получим, пренебрегая сопротивлениями R_k и pL_2 ,

$$i = i_1 + i_k; \quad \frac{1}{1/R + pC} i_1 + pMi = 0.$$

Из этих уравнений находим

$$i_k = \left(p^2 MC + p \frac{M}{R} + 1 \right) i. \quad (5)$$

С другой стороны, уравнение движения шлейфа есть

$$(p^2m + pr + s)x = Ni_k. \quad (I)$$

Сопоставляя (1) и (5), получаем

$$(p^2m + pr + s)x = N(p^2MC + pM/R + 1)i$$

или

$$\frac{x}{i} = N \frac{p^2MC + pM/R + 1}{p^2m + pr + s}. \quad (6)$$

Рассмотрим это выражение. Дробь в правой части имеет в числителе и знаменателе одинаковые по структуре трехчлены относительно p . Эта дробь может равняться постоянной величине, или когда все коэффициенты при p^2 и p равны нулю, или когда трехчлены подобны, т. е. когда между коэффициентами существует соотношение

$$\frac{MC}{m} = \frac{M/R}{r} = \frac{1}{s} = \beta.$$

В этом случае трехчлены делятся друг на друга без остатка и частное равно β , так что

$$x/i = N\beta = N/s.$$

Мы получаем, таким образом, тот же результат, что и полученный ранее ценою увеличения s . Очевидно, что коррекция может дать количественный выигрыш, т. е. что определенное улучшение свойств системы может быть достигнуто при меньшей потере чувствительности, чем если бы такое же улучшение достигалось путем увеличения упругости.

Предшествующий вывод привел нас к полной коррекции, т. е. к полному устранению искажений. Такой результат есть следствие сделанных при выводе упрощений. Мы повторим вывод заново, введя в расчет сопротивления R_k и pL_2 , которыми ранее пренебрегли. Теперь не будут учтены только активное сопротивление вторичной обмотки и рассеяние трансформатора; практически возможно сделать эти величины достаточно малыми. Кроме того, мы пренебрегаем для простоты реакцией шлейфа как преобразователя.

Законы Кирхгофа дают

$$i = i_1 + i_k; \quad \left(\frac{1}{1/R + pC} + pL_2\right)i_1 - R_k i_k + pMi = 0.$$

Из этих уравнений получаем для i_k

$$i_k = \frac{p^2(L_2 + M)C + p \frac{L_2 + M}{R} + 1}{p^2L_2C + p(L_2/R + R_kC) + R_k/R + 1}. \quad (7)$$

Сопоставление (7) и (1) дает

$$\frac{x}{i} = \frac{p^2(L_2 + M)C + p \frac{L_2 + M}{R} + 1}{p^2m + pr + s} \cdot \frac{N}{p^2L_2C + p(L_2/R + R_kC) + R_k/R + 1}. \quad (8)$$

В этом выражении первый сомножитель может быть обращен в постоянную величину, если положить

$$(L_2 + M)C/m = (L_2 + M)/Rr = 1/s = \beta. \quad (II)$$

Тогда

$$\frac{x}{i} = \frac{N\beta}{p^2L_2C + p(L_2/R + R_kC) + R_k/R + 1}. \quad (9)$$

Рассмотрим это выражение. Как видим, оно совершенно аналогично по структуре выражению для чувствительности некорректированной системы, которое на основании (I) имеет вид

$$x/i = N/(p^2m + pr + s). \quad (10)$$

Результат коррекции проявляется, очевидно, только в изменении коэффициентов трехчлена в знаменателе. В частности, если собственная частота некорректированной системы есть

$$\omega_0^2 = s/m,$$

то собственная частота корректированной системы определится отношением коэффициентов при нулевой и второй степени p в формуле (9)

$$\omega_k^2 = \frac{1 + R_k/R}{L_2C}.$$

Отсюда отношение собственных частот

$$\mu^2 = \frac{\omega_k^2}{\omega_0^2} = \frac{m}{s} \cdot \frac{1 + R_k/R}{L_2C}. \quad (11)$$

Обратимся к чувствительности. В некорректированной системе

$$\gamma_0 = N/s.$$

В корректированной системе на основании (9)

$$\gamma_k = \frac{N\beta}{1 + R_k/R} = \frac{N}{s(1 + R_k/R)}.$$

Отношение чувствительностей

$$\varepsilon = \frac{\gamma_k}{\gamma_0} = \frac{1}{1 + R_k/R}. \quad (12)$$

Подставляя (12) в (11), получим

$$\mu^2 = \frac{m}{s} \cdot \frac{1}{\varepsilon L_2C}.$$

Мы имели ранее

$$(L_2 + M)C/m = 1/s = \beta.$$

Отсюда

$$m/s = (L_2 + M) C$$

и, следовательно,

$$\mu^2 = \frac{L_2 + M}{\varepsilon L_2} = \frac{1 + M/L_2}{\varepsilon}.$$

Замечая, что

$$M/L_2 = \sqrt{L_1 L_2 / L_1} = \sqrt{L_1 / L_2} = n, \quad (13)$$

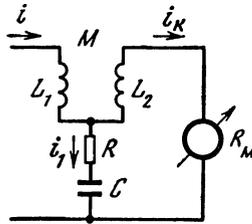
где n означает коэффициент трансформации, окончательно получаем

$$\mu^2 = (1 + n)/\varepsilon$$

или

$$\varepsilon = (1 + n)/\mu^2. \quad (14)$$

Эта формула устанавливает связь между ε и μ , т. е. показывает, как изменяется чувствительность с повышением собственной ча-



Р и с. 2

сты. Для суждения о преимуществах, доставляемых коррекцией, надлежит сравнить (14) и (4). Сравнение показывает, что применение коррекции позволяет получить при прочих равных условиях чувствительность, в $1 + n$ раз большую, чем при коррекции. Напомним, что чувствительность определяется по отношению к току i , предполагаемому заданным.

Схема Кертиса. Кертис [3] применял иную схему, которая в упрощенном и несколько измененном виде изображена на рис. 2.

Для этой схемы получаются более простые соотношения. Аналогично предыдущему

$$i = i_1 + i_k; \quad \left(R + \frac{1}{pC}\right) i_1 + pMi - (R_k + pL_2) i_k = 0,$$

откуда

$$i_k = i \frac{p^2 MC + pRC + 1}{p^2 L_2 C + p(R_k + R)C + 1} \quad (15)$$

и, следовательно,

$$\frac{x}{i} = \frac{p^2 MC + pRC + 1}{p^2 m + pr + s} \cdot \frac{N}{p^2 L_2 C + p(R_k + R)C + 1}. \quad (16)$$

Полагая

$$MC/m = RC/r = 1/s = \beta, \quad (III)$$

получим

$$\frac{x}{i} = \frac{N\beta}{p^2 L_2 C + p(R_k + R)C + 1} \cdot \quad (17)$$

Собственная частота корректированной системы $\omega_k^2 = 1/L_2 C$. Отношение собственных частот $\mu_k^2 = \omega_k^2 / \omega_0^2 = m/sL_2 C$. Но так как $m/s = MC$, то $\mu^2 = M/L_2$ или

$$\mu^2 = n. \quad (18)$$

Чувствительность корректированной системы на основании (17) равна

$$\gamma_k = N\beta = N/s,$$

но это есть не что иное, как чувствительность некорректированной системы. Итак

$$\varepsilon = 1. \quad (19)$$

Таким образом, чувствительность при коррекции по схеме Кертиса остается неизменной. В этом отношении схема Кертиса выгоднее схемы Ирвина, для которой уменьшение чувствительности выражается формулой (12).

Расчет. Мы можем теперь составить расчетные формулы для параметров корректирующих схем. Будем полагать, что параметры шлейфа заданы.

а) *Схема Ирвина.* По формуле (12)

$$R = \frac{R_k}{1/\varepsilon - 1}. \quad (20)$$

Далее, с помощью равенств (II) найдем

$$L_2 + M = \frac{r}{s} R, \quad (21)$$

$$C = \frac{m}{s(M + L_2)} = \frac{1}{\omega_0^2(M + L_2)} \quad (22)$$

и, наконец,

$$L_2 = \frac{M + L_2}{1 + n}. \quad (23)$$

Эти четыре формулы позволяют определить четыре параметра R , C , L_2 и M по заданным параметрам механической системы (m , r и s) и по значениям ε и n , одно из которых может произвольно выбираться.

В этой схеме расчета не учтен, однако, пока один важный фактор, а именно затухание корректированной системы. Дело в том, что необходимо, кроме повышения собственной частоты, позаботиться также и о надлежащем затухании системы.

Относительное затухание $k = \alpha/\omega_0$ берется в осциллографах обычно порядка 0,7—0,5. Часто полагают, что наиболее выгодное

значение относительного затухания есть $k = \sqrt{2}/2 = 0,707$. Введем в расчет это значение. Требование соблюдения определенной величины затухания дает нам еще одно уравнение и накладывает еще одну связь на соотношения величин n , ε и μ .

Для коэффициента затухания имеем из (9)

$$\alpha_k = \frac{L_2/R + R_k C}{2L_2 C},$$

собственная частота

$$\omega_k = \sqrt{\frac{1 + R_k/R}{L_2 C}}$$

или на основании (12)

$$\omega_k = \frac{1}{\sqrt{\varepsilon L_2 C}}.$$

Таким образом,

$$k = \frac{\alpha_k}{\omega_k} = \frac{L_2/R + R_k C}{2} \sqrt{\frac{\varepsilon}{L_2 C}}.$$

По принятому условию должно быть

$$k = \sqrt{2}/2,$$

это дает уравнение

$$\frac{L_2/R + R_k C}{2} \sqrt{\frac{\varepsilon}{L_2 C}} = \frac{\sqrt{2}}{2},$$

которое после преобразований приводится к виду

$$\left(\frac{L_2}{C}\right)^2 - 2R^2 \frac{L_2}{C} + R^4 \left(\frac{1}{\varepsilon} - 1\right)^2 = 0,$$

откуда

$$L_2/C = R^2 (1 - \sqrt{1 - (1/\varepsilon - 1)^2}).$$

Мы стремимся к тому, чтобы ε как можно меньше отличалось от единицы; если это условие выполнено, то $1/\varepsilon - 1$ есть малая дробь, и мы можем написать приближенно

$$\frac{L_2}{C} \cong R^2 \frac{1}{2} \left(\frac{1}{\varepsilon} - 1\right)^2.$$

После ряда алгебраических преобразований с применением формул (20)—(23) и (14) получаем следующее соотношение:

$$\mu^2 = 1 + n + 2k_0 \sqrt{2n}, \quad (24)$$

где $k_0 = r\omega_0/2s$ — относительное затухание некорректированной системы.

Таким образом, если задаться величиной n , то μ и ε определяются однозначно по формулам (24) и (14). Естественно предположить, что техническими условиями задается величина μ .

Тогда n , ε и все прочие параметры определяются однозначно. Формула для определения n получается из (24) в следующем виде:

$$n = \mu^2 - 1 + 4k_0^2(1 + \sqrt{(\mu^2 - 1)/2k_0^2 + 1}). \quad (25)$$

Итак, расчет производится в следующем порядке:

$$n = \mu^2 - 1 + 4k_0^2(1 + \sqrt{(\mu^2 - 1)/2k_0^2 + 1}), \quad (25)$$

$$\varepsilon = (1 + n)/\mu^2, \quad (14)$$

$$R = \frac{R_k}{1/\varepsilon - 1}, \quad (20)$$

$$L_2 + M = \frac{r}{s} R, \quad (21)$$

$$C = \frac{1}{\omega_0^2(L_2 + M)}, \quad (22)$$

$$L_2 = \frac{L_2 + M}{n + 1}. \quad (23)$$

б) *Схема Кертиса*. Для определения параметров служат, во-первых, равенства (III) и, во-вторых, уравнение для затухания, которое имеет вид

$$\frac{1}{2}(R + R_k) \sqrt{\frac{C}{L_2}} = \frac{\sqrt{2}}{2}.$$

Таким образом, имеем три уравнения, определяющие три параметра R , C и M

$$RC = r/s, \quad MC = 1/\omega_0^2, \quad R + R_k = \sqrt{2M/nC}.$$

Решение этих уравнений дает расчетные формулы

$$R = \frac{\sqrt{2}\mu k_0}{1 - \sqrt{2}\mu k_0}, \quad C = \frac{\sqrt{2}(1 - \sqrt{2}\mu k_0)}{\omega_0 \mu R_k}, \quad M = \frac{\mu R_k}{\sqrt{2}\omega_0(1 - \sqrt{2}\mu k_0)}.$$

При решении приняты во внимание соотношения (13) и (18).

Во всех полученных формулах фигурирует разность $1 - \sqrt{2}\mu k_0$. Эта разность, очевидно, должна быть положительной. Из этого следует, что большие значения μ , т. е. существенный эффект коррекции, может быть достигнут лишь при малых значениях относительного затухания механической системы. Предельное соотношение определяется неравенством

$$1 - \sqrt{2}\mu k_0 \geq 0,$$

т. е. если $k_0 \geq \sqrt{2}/2$, то μ не превосходит единицы, и коррекция вообще не может принести никакой пользы.

* * *

Изложенное представляет собою лишь предварительный набросок расчета, требующий, конечно, дальнейшего уточнения и в особенности экспериментальной проверки.

Вопрос заслуживает внимания, так как коррекция осциллографа позволяет в принципе разрешить три задачи: а) повышение чувствительности при заданной собственной частоте; б) сохранение собственной частоты при увеличении массы подвижной системы и в) повышение собственной частоты при сохранении чувствительности.

Относительно второй задачи следует, в частности, отметить, что она имеет значение в применении к быстроходным осциллографам с зеркальными шлейфами в связи с открывающейся возможностью увеличения размеров зеркала [4].

Л и т е р а т у р а

1. А. А. Харкевич. Журн. техн. физ., 7, 1937, 250; 6, 1936, 1071.
2. J. T. Irwin. Oscillographs. London, Is. Pitman, 1925.
3. A. M. Curtis. BSTJ, 12, 1933, 76.
4. H. Abraham. J. phys., 8, 1909, 265.

СПЕКТРЫ И АНАЛИЗ

Глава I

СПЕКТРЫ

§ 1. Введение

Когда Бернулли и Эйлер, а затем Фурье впервые применили разложение функций в тригонометрические ряды¹, то это разложение рассматривалось лишь как математическое средство для решения задач математической физики. Сам Фурье пользовался рядами, получившими его имя, для интегрирования уравнения теплопроводности. Метод Фурье стал классическим приемом решения волновых уравнений — уравнения струны и позднее телеграфного уравнения. Однако разложение Фурье долгое время не связывалось непосредственно с какими-либо физическими представлениями. Даже после открытия электрических колебаний и волн высказывалось сомнение в адекватности разложения Фурье происходящим физическим явлениям. Например, Герц (см. его переписку с Пуанкаре²) отрицательно относился к спектральным представлениям.

Долгое время спектральные представления применялись и развивались лишь сравнительно узким кругом физиков-теоретиков. Но, начиная с двадцатых годов, в связи с бурным развитием радиотехники, акустики, колебательной механики и вообще отраслей техники, опирающихся на теорию колебаний, спектральные представления необычайно широко распространились. Была установлена прямая связь между спектральным разложением и поведением реальных колебательных систем. Спектральный способ описания явлений получил всеобщее признание.

Более того, спектральный язык стал всеобщим языком, на котором объясняются между собой все, имеющие дело с техническими применениями разного рода колебаний. На спектральном (частотном) языке стали описывать не только явления, но и свойства аппаратуры.

¹ Любопытно отметить, что самая возможность такого разложения в те времена оспаривалась.

² Цитируется в работе Н. Н. Андреева [2].

Нет сомнения, что такое широкое развитие спектральных представлений сыграло огромную прогрессивную роль; благодаря им сложные колебательные явления стали доступны пониманию широких кругов техников и физиков.

Но история развития спектральных представлений показывает, что эти представления давали иногда осечку. Совершались и продолжают совершаться грубые ошибки. Возникали затяжные дискуссии по основным вопросам (например, дискуссия о боковых полосах при радиопередаче), происходили курьезные недоразумения (например, неправильное представление о ширине полосы при частотной модуляции). Обнаруживались разного рода парадоксы. А парадоксы, как замечательно сказал Л. И. Мандельштам, возможны лишь там, где нет полного понимания, «понимания второго рода», как он выражался (если не говорить о парадоксах, обусловленных несовершенством самой теории).

В действительности спектральный подход безупречен. Он никогда не приведет к ошибкам, если им разумно пользоваться. Вышеупомянутые ошибки и недоразумения — это не порок метода, а результат неумелого его применения.

Можно избежать многих ошибок, если не ограничивать свой кругозор спектральными рамками, а дополнять и углублять спектральный подход подходом временным.

Очень интересно проследить эволюцию спектральных представлений за последнее время. Первоначальное определение спектра основывается на преобразовании Фурье; интегрирование по времени выполняется в бесконечных пределах. Таким образом, преобразованию подвергается функция времени в целом; результат преобразования, т. е. спектр, зависит только от частоты. Однако учет реальных условий эксперимента заставляет ввести новое понятие — понятие текущего спектра. Текущий спектр определяется как результат преобразования Фурье, но с переменным верхним пределом интегрирования, в качестве которого фигурирует текущее время. Таким образом, появляется спектральная функция, зависящая не только от частоты, но и от времени — это уже некоторое промежуточное понятие, сближающее частотные и временные представления. Процесс сближения продолжается: вводится, несомненно, полезное понятие мгновенного спектра и связанное с ним понятие активной полосы спектра. От мгновенного спектра остается один лишь шаг до мгновенной частоты, после чего мы можем снова говорить о синусоиде с переменной частотой, т. е. восстановить в правах понятие, весьма решительно осужденное в свое время. Таким образом, спектральные представления, описав в своем развитии широкий круг, возвращаются почти что к исходным позициям, однако на значительно более высоком уровне; все основные и ряд промежуточных понятий ясно определены и образуют в совокупности мощное и гибкое орудие исследования.

В связи с развитием новых отраслей техники, главным образом радиотехники (например, импульсная техника, специальные виды модуляции и т. п.), от инженеров требуется более глубокое овладение спектральными представлениями. Поэтому одна из задач этой книги состоит в изложении основных вопросов теории спектров в объеме, по возможности отвечающем современным требованиям, и на уровне, несколько превышающем средний уровень знаний, даваемый высшим техническим учебным заведением.

Изложению этих вопросов посвящена первая глава «Спектры».

Практическое применение спектральных представлений неизбежно приводит к необходимости экспериментального осуществления разложения Фурье, т. е. к гармоническому анализу различных явлений. Хотя существует огромное число методов анализа и применяющих эти методы приборов-анализаторов, до сих пор многие основные вопросы анализа оставались недостаточно разработанными, а подчас и не вполне ясными. В частности, основные требования к анализу как к измерительному процессу и к анализатору как измерительному прибору зачастую вовсе не ставились и не обсуждались. Вероятно, это можно объяснить тем, что постановка таких вопросов связана с известными трудностями. Однако эти трудности преодолимы, и необходимо попытаться привести принципиальные вопросы анализа в некоторый порядок. Такая попытка сделана во второй главе «Анализ».

Вопросы, относящиеся к спектрам случайных процессов, выделены в отдельную — третью главу.

Ряд дополнительных вопросов, представляющих тот или иной интерес и имеющих прямое отношение к теме, рассмотрен в «Добавлениях».

§ 2. Ряд и интеграл Фурье

Понятие о разложении Фурье можно считать общеизвестным. Поэтому здесь напоминаются лишь основные соотношения и определения. Математические подробности читатель найдет в любом учебнике.

Мы начинаем с определения периодической функции

$$f(t) = f(t + nT). \quad (2.1)$$

Здесь T — постоянная величина, называемая периодом, n — любое целое число, положительное или отрицательное. Определение (2.1) выражает основное свойство периодической функции, состоящее в том, что ход явления периодически повторяется и что периодичность эта существует вечно, т. е. для всех времен от $-\infty$ до $+\infty$.

Из этого сразу можно заключить, что периодических явлений в строгом смысле определения (2.1) в действительности нет и быть не может. Периодическая функция есть полезная математи-

ческая абстракция; ее соотношение с действительными явлениями выяснится в дальнейшем.

Всякая — с несущественными для нас математическими ограничениями¹ — периодическая функция может быть представлена рядом по тригонометрическим функциям

$$f(t) = c_0 + \sum_{k=1}^{\infty} c_k \cos\left(2\pi k \frac{t}{T} - \varphi_k\right). \quad (2.2)$$

Периодическая функция $f(t)$ представлена, таким образом, суммой слагаемых вида

$$c_k \cos(2\pi kt/T - \varphi_k),$$

каждое из которых есть синусоидальное колебание с амплитудой c_k и начальной фазой φ_k . Значения c_k и φ_k должны быть надлежащим образом подобраны для того, чтобы равенство (2.2) выполнялось. Частоты колебаний, из которых составляется периодическая функция $f(t)$, образуют гармоническую последовательность; это значит, что частоты всех составляющих кратны основной частоте $1/T$. Отдельные составляющие носят название гармоник. Колебание с частотой $1/T$ называется первой гармоникой ($k=1$), с частотой $2/T$ — второй гармоникой ($k=2$) и т. д.

Выражение (2.2) может быть переписано в другой, очень употребительной форме

$$f(t) = c_0 + \sum_{k=1}^{\infty} (a_k \cos 2\pi kt/T + b_k \sin 2\pi kt/T), \quad (2.3)$$

где $a_k = c_k \cos \varphi_k$, $b_k = c_k \sin \varphi_k$, так что $c_k = \sqrt{a_k^2 + b_k^2}$, $\operatorname{tg} \varphi_k = b_k/a_k$. Коэффициенты и определяются по формулам

$$a_k = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos 2\pi k \frac{t}{T} dt, \quad (2.4)$$

$$b_k = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin 2\pi k \frac{t}{T} dt. \quad (2.5)$$

Величина c_0 выражает среднее значение функции за период, она называется часто постоянной составляющей и вычисляется по формуле

$$c_0 = \frac{1}{T} \int_{-T/2}^{T/2} f(t) dt. \quad (2.6)$$

¹ Именно: функция предполагается ограниченной кусочно-непрерывной и имеющей на протяжении периода конечное число экстремальных значений (условия Дирихле).

Если a_k , b_k и c_0 определены по формулам (2. 4)—(2. 6), то равенство (2. 3) есть тождество.

Замечательным свойством ряда Фурье является то, что если взять конечное число членов ряда, т. е. аппроксимировать периодическую функцию тригонометрическим полиномом, представив ее в виде

$$f(t) \simeq c_0 + \sum_{k=1}^N (a_k \cos 2\pi kt/T + b_k \sin 2\pi kt/T),$$

то для любого N получается наименьшее квадратичное отклонение от точного значения $f(t)$, если коэффициенты полинома определены по тем же формулам (2. 4)—(2. 6). С увеличением числа членов N приближение, разумеется, улучшается, и в пределе при $N \rightarrow \infty$ приближенное равенство переходит в точное.

Ряд Фурье может быть также записан в комплексной форме следующим образом:

$$f(t) = \sum_{k=-\infty}^{\infty} C_k e^{j2\pi kt/T}, \quad (2. 7)$$

где $2C_k = c_k e^{-j\varphi_k} = a_k - jb_k$; $c_k = 2|C_k|$; $c_0 = C_0$. Величина $2C_k$ есть комплексная амплитуда; C_k находится по формуле

$$C_k = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-j2\pi kt/T} dt. \quad (2. 8)$$

Суммирование в (2. 7) ведется по всем целым k как положительным, так и отрицательным, включая и нуль. Для того чтобы вернуться от (2. 7) к (2. 2) или (2. 3), нужно вспомнить, что вещественная часть каждого слагаемого под знаком сумм в (2. 7) четна по отношению к k , а мнимая часть — нечетна. В дальнейшем мы будем преимущественно пользоваться компактной формой (2. 7).

Ряд Фурье дает разложение периодической функции по тригонометрическим функциям. Это разложение может быть обобщено и на случай непериодической функции. Нестрогий, но наглядный путь к получению разложения Фурье непериодической функции состоит в применении предельного перехода при $T \rightarrow \infty$. Действительно, непериодическую функцию можно рассматривать как предельный случай периодической функции при неограниченно возрастающем периоде. Возьмем формулу (2. 7) и подставим в нее значение C_k из (2. 8)

$$f(t) = \frac{1}{T} \sum_{k=-\infty}^{\infty} e^{j2\pi kt/T} \int_{-T/2}^{T/2} f(t) e^{-j2\pi kt/T} dt.$$

Перейдем к пределу, устремляя T к бесконечности. Вместо $1/T$ введем круговую основную частоту $\omega_1 = 2\pi/T$. Эта величина есть

частотный интервал между соседними гармониками, частоты которых равны $2\pi k/T$. При предельном переходе сделаем замену по следующей схеме:

$$T \rightarrow \infty, \quad \omega_1 \rightarrow d\omega, \quad 2\pi \frac{k}{T} \rightarrow \omega,$$

где ω — текущая частота, изменяющаяся непрерывно; $d\omega$ — ее приращение. Сумма перейдет в интеграл, и мы получим

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{j\omega t} d\omega \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (2.9)$$

или

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega t} d\omega, \quad (2.10)$$

где

$$S(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt. \quad (2.11)$$

Формулы (2.10) и (2.11) являются основными формулами теории спектров. Они представляют собой пару преобразований Фурье, связывающих между собой две функции: вещественную функцию времени $f(t)$ и комплексную функцию частоты $S(\omega)$. Формула (2.10) представляет собой интеграл Фурье в комплексной форме. Смысл этой формулы состоит в том, что функция $f(t)$ представлена суммой синусоидальных составляющих. Но функция $f(t)$ предполагается непериодической; поэтому она может быть представлена только суммой бесконечно большого числа бесконечно малых колебаний, бесконечно близких по частоте. Комплексная амплитуда каждого отдельного колебания бесконечно мала; она равна

$$dC = \frac{1}{\pi} S(\omega) d\omega. \quad (2.12)$$

Частотный интервал между двумя соседними колебаниями также бесконечно мал; он равен $d\omega$.

Если ряд Фурье представляет периодическую функцию суммой хотя и бесконечного числа синусоид, но с частотами, имеющими определенные дискретные значения, то интеграл Фурье представляет непериодическую функцию суммой синусоид с непрерывной последовательностью частот. В составе непериодической функции имеются, как говорят, все частоты.

Одна из особенностей, отличающих интеграл Фурье от ряда Фурье, состоит в том, что ряд Фурье представляет периодическую функцию как сумму периодических же составляющих, тогда как

интеграл Фурье представляет непериодическую функцию суммой периодических составляющих. Таким образом, в случае интеграла Фурье сумма не обладает существенным свойством своих слагаемых, и это обстоятельство необходимо иметь в виду при рассуждениях общего характера о спектральном разложении по Фурье.

Отметим в заключение, что формулу (2.10) можно записать в вещественной форме; тогда интегрирование будет производиться только по положительным частотам. Введя обозначение $S(\omega) = A(\omega) + jB(\omega)$, получим (учитывая, что A — четная, а B — нечетная функция)

$$f(t) = \frac{1}{\pi} \int_0^{\infty} [A(\omega) \cos \omega t - B(\omega) \sin \omega t] d\omega. \quad (2.13)$$

Можно получить еще одну запись формулы (2.10), представив ее в виде

$$f(t) = \frac{1}{2\pi} \int_0^{\infty} [S(\omega) e^{j\omega t} + S(-\omega) e^{-j\omega t}] d\omega.$$

В квадратных скобках стоит сумма сопряженных величин, равная удвоенной вещественной части. Поэтому

$$f(t) = \frac{1}{\pi} \operatorname{Re} \int_0^{\infty} S(\omega) e^{j\omega t} d\omega. \quad (2.14)$$

§ 3. Спектры, определения и классификация

Вернемся к формуле (2.2) ряда Фурье и перепишем ее в виде

$$f(t) = c_0 + \sum_{k=1}^{\infty} c_k \cos(k\omega_1 t - \varphi_k).$$

Здесь $\omega_1 = 2\pi/T$ — основная частота. Как мы видим, сложная периодическая функция $f(t)$ вполне определяется совокупностью величин c_k и φ_k . Совокупность величин c_k носит название спектра амплитуд. Совокупность величин φ_k называется соответственно спектром фаз. Для многих применений достаточно знать спектр амплитуд; он применяется настолько часто, что когда говорят просто спектр, то подразумевают обычно именно амплитудный спектр. В остальных случаях делают соответственные оговорки. Мы будем поступать также.

Спектр периодической функции можно изобразить графически. Выберем для этого координаты c_k и $\omega = k\omega_1$. Спектр будет изображен в этой системе координат совокупностью дискретных точек, так как каждому значению $k\omega_1$ соответствует одно определенное c_k .

График, состоящий из отдельных точек, неудобен. Поэтому принято изображать амплитуды отдельных гармоник вертикальными отрезками соответствующей длины. В результате спектр периодической функции принимает вид, показанный на рис. 1. Это — дискретный спектр; его называют также линейчатым, заимствуя этот термин из оптики. Второе свойство спектра, изображенного на рис. 1, состоит в том, что этот спектр — гармонический. Это значит, что он состоит из равноотстоящих спектральных линий; частоты гармоник находятся в простых кратных соотношениях. Конечно, отдельные гармоники, иногда даже первая, могут отсутствовать, т. е. амплитуды их могут равняться нулю; это, однако, не нарушает гармоничности спектра.



Р и с. 1

Не следует думать, что только периодическая функция обладает дискретным спектром. Предположим, например, что сложное колебание есть результат сложения двух синусоидальных колебаний с несоизмеримыми частотами, скажем, ω_1 и $\sqrt{2} \omega_1$. Это колебание заведомо непериодическое, однако спектр его дискретен и состоит из двух спектральных линий.

Функция, обладающая дискретным спектром из произвольно расположенных на шкале частот спектральных линий, называется почти периодической и обладает многими интересными свойствами [3, 15].

Итак, дискретные, или линейчатые, спектры могут принадлежать как периодическим, так и непериодическим функциям. В первом случае линейчатый спектр обязательно гармонический.

Большое практическое значение имеет частный случай почти периодической функции, представляемой разложением вида

$$f(t) = \sum c_k \cos [(\omega_0 + k\omega_1)t - \varphi_k],$$

где k принимает как положительные, так и отрицательные значения. Спектр, отвечающий этому разложению, характеризуется тем, что линии его эквидистантны; поэтому мы будем называть такого рода линейчатый спектр квазигармоническим. Таковы, например, спектры периодически модулированных колебаний; ω_0 в этом случае есть не что иное, как несущая частота.

Обратимся теперь к спектрам непериодических функций. Мы уже знаем, что в результате предельного перехода от ряда к интегралу Фурье интервалы между отдельными спектральными линиями неограниченно сокращаются, линии сливаются, и вместо дискретных точек спектр должен изображаться непрерывной последовательностью точек, т. е. непрерывной кривой. Такого рода спектр называется сплошным.

Здесь нужно, однако, внести одно уточнение. Мы писали формулу интеграла Фурье в виде (2. 10)

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) e^{j\omega t} d\omega.$$

Подынтегральная функция выражает отдельное бесконечно малое слагаемое, т. е. колебание с бесконечно малой амплитудой dC

$$\frac{1}{\pi} S(\omega) e^{j\omega t} d\omega = dC e^{j\omega t}.$$

Отсюда находим $S(\omega) = dC/d\omega$.

Таким образом, величина $S(\omega)$ выражает не непосредственно амплитуду, а так называемую спектральную плотность. Однако обычно эту деталь опускают и называют $S(\omega)$ комплексным спектром непериодической функции, а абсолютное значение (модуль) этой величины $\Phi(\omega) = |S(\omega)|$ просто спектром. Это может повести к недоразумениям лишь в том случае, когда мы будем непосредственно сравнивать соотношения для периодических и непериодических функций.

Итак, мы имеем две основные разновидности спектров: линейчатые и сплошные. Гармонические линейчатые спектры принадлежат периодическим функциям, сплошные — непериодическим.

В заключение настоящего параграфа заметим, что теми или иными функциями времени могут выражаться изменения самых различных физических величин. Соответствующий смысл получают и спектры этих функций. В практике приходится, например, иметь дело со спектрами механических величин (смещения, скорости, ускорения, силы, давления и т. п.), электрических величин (тока, напряжения и т. п.). Кроме того, нас часто интересуют спектры квадратичных величин: мощности или энергии. Например, в оптике чаще всего имеют дело со спектрами этих величин.

§ 4. Некоторые теоремы о спектрах

Выведем теперь несколько общих теорем о спектрах, основанных на свойствах преобразования Фурье. Эти теоремы сходны с теоремами операционного исчисления и выводятся аналогичным путем: ведь преобразование Фурье и преобразование Лапласа,

составляющие основу операционного исчисления, находятся в близком родстве между собой.

Прежде всего отметим, что преобразование Фурье линейно. Из этого непосредственно следует, что к нему применим принцип наложения. Это обстоятельство можно выразить следующим соотношением:

$$S_1(\omega) + S_2(\omega) = \int_{-\infty}^{+\infty} e^{-j\omega t} [f_1(t) + f_2(t)] dt. \quad (4.1)$$

Смысл соотношения (4.1) может быть кратко выражен так: спектр суммы равен сумме спектров.

Теперь докажем теорему о спектре производной: если комплексный спектр функции $f(t)$ есть $S(\omega)$, то комплексный спектр $f'(t) = df/dt$ есть $j\omega S(\omega)$; комплексный спектр производной получается из комплексного спектра функции умножением на $j\omega$.

Для доказательства составим выражение

$$S_{(1)}(\omega) = \int_{-\infty}^{+\infty} e^{-j\omega t} f'(t) dt$$

и проинтегрируем его по частям. Получим

$$S_{(1)}(\omega) = f(t) e^{-j\omega t} \Big|_{-\infty}^{+\infty} + j\omega \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt.$$

Так как функция, представляемая интегралом Фурье, обращается в нуль при $t \rightarrow \pm\infty$, то имеем

$$S_{(1)}(\omega) = j\omega S(\omega). \quad (4.2)$$

Это доказательство может быть распространено на случай n -й производной. Прделав интегрирование по частям n раз, получим комплексный спектр n -й производной [при условии, что все производные функции до $(n-1)$ порядка включительно обращаются в нуль при $t \rightarrow \pm\infty$]

$$S_{(n)}(\omega) = (j\omega)^n S(\omega). \quad (4.3)$$

Подобным же образом может быть выведено выражение для комплексного спектра интеграла от данной функции.

Составляя выражение

$$S_{(-1)}(\omega) = \int_{-\infty}^{+\infty} e^{-j\omega t} \left(\int_{-\infty}^t f(\tau) d\tau \right) dt$$

и интегрируя его по частям, находим

$$S_{(-1)}(\omega) = \frac{1}{j\omega} S(\omega) \quad (4.4)$$

при условии, что

$$\int_{-\infty}^{\infty} f(t) dt = 0.$$

Это условие выполняется, например, для всякой нечетной функции, интегрируемой в бесконечных пределах.

Выведем теперь выражение для комплексного спектра функции, отличающейся от исходной запаздыванием на время τ . Мы можем записать

$$S_{\tau}(\omega) = \int_{-\infty}^{+\infty} e^{-j\omega t} f(t - \tau) dt.$$

Путем простой замены переменной по формуле $t_1 = t - \tau$ приходим к результату

$$S_{\tau}(\omega) = e^{-j\omega\tau} S(\omega). \quad (4.5)$$

Если в этом соотношении перейти от комплексных спектров к их модулям, то получим

$$|S_{\tau}(\omega)| = \Phi_{\tau}(\omega) = \Phi(\omega),$$

т. е. при запаздывании — или вообще при смещении функции по шкале времен — спектр ее остается неизменным. Иначе говоря, спектр не зависит от выбора начального момента для отсчета времени, чего и следовало ожидать.

Следующая теорема относится к транспозиции (переносу) спектров. Вопрос ставится следующим образом: какой функции соответствует спектр, смещенный по шкале частот на Ω ?

Так как

$$S(\omega + \Omega) = \int_{-\infty}^{+\infty} e^{-j(\omega + \Omega)t} f(t) dt,$$

то, следовательно, комплексным спектром искомого вида будет обладать функция

$$f_{\Omega}(t) = e^{-j\Omega t} f(t). \quad (4.6)$$

Выведем теперь некоторые более сложные соотношения. Возьмем выражение интеграла Фурье

$$f_1(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_1(\omega) e^{j\omega t} d\omega,$$

умножим обе части на $f_2(t)$ и проинтегрируем по t в пределах $\pm\infty$. Получим

$$\int_{-\infty}^{+\infty} f_1(t) f_2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f_2(t) dt \int_{-\infty}^{+\infty} S_1(\omega) e^{j\omega t} d\omega.$$

Изменим порядок интегрирования в правой части

$$\int_{-\infty}^{+\infty} f_2(t) dt \int_{-\infty}^{+\infty} S_1(\omega) e^{j\omega t} d\omega = \int_{-\infty}^{+\infty} S_1(\omega) d\omega \int_{-\infty}^{+\infty} f_2(t) e^{j\omega t} dt.$$

Таким образом,

$$\int_{-\infty}^{+\infty} f_1(t) f_2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_1(\omega) S_2(-\omega) d\omega. \quad (4.7)$$

Эта формула годится, например, для вычисления энергии, если известны спектры тока и напряжения, или любых двух других функций, произведение которых выражает мощность.

В вещественной форме формула (4.7) может быть представлена в виде (учитывая, что $S(-\omega) = S^*(\omega)$)

$$\int_{-\infty}^{+\infty} f_1(t) f_2(t) dt = \frac{1}{\pi} \int_0^{\infty} \Phi_1(\omega) \Phi_2(\omega) \cos(\varphi_1 - \varphi_2) d\omega; \quad (4.8)$$

для частного случая $f_1 = f_2$

$$\int_{-\infty}^{+\infty} f^2(t) dt = \frac{1}{\pi} \int_0^{\infty} \Phi^2(\omega) d\omega. \quad (4.9)$$

Последнее соотношение известно под названием теоремы Рэяли. Нам предстоит им пользоваться в дальнейшем. Формула (4.9) показывает, что по физическому смыслу функция

$$\Phi^2(\omega) = S(\omega) S^*(\omega)$$

представляет спектральную плотность энергии.

Перепишем формулу (4.7) в новых обозначениях

$$\int_{-\infty}^{+\infty} f_1(t) u_2(t) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_1(\nu) R_2(-\nu) d\nu$$

и определим функцию $u_2(t)$ соотношением

$$u_2(t) = f_2(t) e^{-j\omega t}.$$

Если обозначить через S_2 спектр функции f_2 , то на основании теоремы (4.6) будем иметь

$$R_2(\nu) = S_2(\omega + \nu)$$

и, следовательно,

$$\int_{-\infty}^{+\infty} f_1(t) f_2(t) e^{-j\omega t} dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_1(\nu) S_2(\omega - \nu) d\nu.$$

Таким образом, если S_1 и S_2 — соответственно спектры функций f_1 и f_2 , а S — спектр произведения $f = f_1 f_2$, то имеем

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_1(\nu) S_2(\omega - \nu) d\nu. \quad (4.10)$$

Интеграл в правой части носит название свертки функций S_1 и S_2 .

Соотношение (4.10) выражает спектр произведения двух функций через спектры каждой из них.

Выведем, наконец, еще одну формулу. Составим свертку двух функций времени t_1 и t_2

$$f(t) = \int_{-\infty}^{+\infty} f_1(\tau) f_2(t - \tau) d\tau$$

и вычислим спектр этой функции

$$\begin{aligned} S(\omega) &= \int_{-\infty}^{+\infty} e^{-j\omega t} dt \int_{-\infty}^{+\infty} f_1(\tau) f_2(t - \tau) d\tau = \\ &= \int_{-\infty}^{+\infty} f_1(\tau) d\tau \int_{-\infty}^{+\infty} e^{-j\omega t} f_2(t - \tau) dt = \\ &= \int_{-\infty}^{+\infty} e^{-j\omega t} \tau f_1(\tau) d\tau \int_{-\infty}^{+\infty} e^{-j\omega \mu} f_2(\mu) d\mu. \end{aligned}$$

Здесь после перемены порядка интегрирования сделана замена переменной по формуле $\mu = t - \tau$.

Итак, спектр функции $f(t)$ есть

$$S(\omega) = S_1(\omega) S_2(\omega). \quad (4.11)$$

Это соотношение дает возможность найти функцию времени, спектр которой известен и выражается произведением спектров двух функций. Полагая в (4.11) $S_2(\omega) = S_1^*(\omega)$, можно убедиться, что $\Phi^2(\omega)$ есть спектр функции

$$\int_{-\infty}^{\infty} f(\tau) f(t + \tau) d\tau = \int_{-\infty}^{+\infty} f(\tau) f(\tau - t) d\tau.$$

Применение соотношений (4. 1)—(4. 11) может в значительной степени облегчить вычисление спектров различных функций.

В заключение настоящего параграфа следует указать на одно интересное обстоятельство. Применяя разложение Фурье, мы имеем дело с парой преобразований Фурье

$$S(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt, \quad (4. 12)$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) e^{j\omega t} d\omega.$$

В этих формулах обращает на себя внимание то, что время t и круговая частота ω входят в них симметричным образом, на равных началах¹. Но из полной симметрии формул (4. 12) следует также, что всякая теорема теории спектров имеет парную теорему, не требующую особого доказательства и получаемую из данной теоремы простой формальной заменой переменной t , переменной ω и функций времени соответствующими спектральными плотностями.

Таблица 1

Исходные преобразования	
$S(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt$	$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) e^{j\omega t} d\omega$
$S_{\tau} = \int_{-\infty}^{+\infty} f(t - \tau) e^{-j\omega t} dt = e^{-j\omega\tau} S$ Теорема запаздывания	$f_{\Omega} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega + \Omega) e^{j\omega t} d\omega = e^{-j\Omega t} f$ Теорема смещения
$S = \int_{-\infty}^{+\infty} f_1 f_2 e^{-j\omega t} dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_1(\nu) S_2(\omega - \nu) d\nu$ Теорема о спектре произведения	$f = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_1 S_2 e^{j\omega t} d\omega = \int_{-\infty}^{+\infty} f_1(\tau) f_2(t - \tau) d\tau$ Теорема о спектре свертки

¹ Формулы могут быть сделаны совершенно симметричными, если, изменив определения, разнести множитель $1/2\pi$ на оба интеграла (т. е. ввести в обеих формулах множитель $1/\sqrt{2\pi}$), как часто и делают.

Пересматривая теоремы настоящего параграфа, мы убеждаемся в том, что они имеют парный характер. Для иллюстрации этого можно записать некоторые из них так (табл. 1).

§ 5. Текущий спектр

По основному определению (§ 2 и 3) спектральная плотность выражается формулой

$$S(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt. \quad (5.1)$$

Таким образом, для нахождения спектра необходимо выполнить интегрирование по времени в бесконечных пределах. Это возможно в принципе, если функция $f(t)$ задана и известна на всем бесконечном протяжении оси времен. Но если функция $f(t)$ есть отображение некоторого реального физического процесса, являющегося объектом нашего наблюдения, и если весь ход этого процесса не может быть в точности предсказан на основании теоретических соображений, то сведения о функции $f(t)$ мы получаем лишь в результате наших наблюдений. Поэтому мы можем выполнить интегрирование не в бесконечных пределах, как этого требует определение (5.1), а лишь до настоящего, текущего момента.

Все прошлое в принципе нам может быть известно, так что интегрирование может быть выполнено в пределах от $-\infty$ до текущего времени t . Измененное таким образом определение спектра принимает вид

$$S_t(\omega) = \int_{-\infty}^t f(t) e^{-j\omega t} dt. \quad (5.2)$$

Величина $S_t(\omega)$, являющаяся функцией не только частоты, но и времени, носит название текущего спектра.

В действительных условиях наблюдение процесса (или самый процесс) фактически может начинаться в некоторый момент t_0 , находящийся в прошлом на конечном удалении от текущего момента t . В этом случае момент t_0 может быть принят за начало отсчета времени, и мы можем определить текущий спектр следующим образом:

$$S_t(\omega) = \int_0^t f(t) e^{-j\omega t} dt \quad (5.3)$$

Мы будем в дальнейшем пользоваться обоими определениями текущего спектра.

Ясно, что связывание математического определения спектра с условиями реального эксперимента само по себе имеет большое

значение. Об этой стороне дела подробнее говорится в § 18. Но понятие текущего спектра является вообще весьма плодотворным.

Мы начали все изложение теории спектров со спектра периодической функции, определяемой соотношением (2. 1).

Периодическая функция есть математическая абстракция. Эта абстракция очень полезна. Но надо иметь в виду, что не может существовать никакого реального физического процесса, отвечающего определению (2. 1). Всякий действительный процесс имеет начало и конец и, следовательно, описывается выражением вида (2. 1) лишь на протяжении конечного промежутка времени. Мы называем действительный циклически повторяющийся процесс периодическим, если этот процесс длится достаточно долго. Мерилом длительности служит число периодов; длительность велика, если число периодов много больше единицы. Если взять короткий отрезок процесса, то он вовсе не будет иметь периодического характера. Периодичность процесса проявляется не сразу; лишь с течением времени обрисовываются характерные черты процесса. Текущий спектр как раз и выражает со спектральной точки зрения это развитие процесса.

Спектр короткого отрезка процесса — за небольшое время от его начала — однороден, так как короткий отрезок любого процесса есть просто короткий импульс. Если в дальнейшем происходит периодическое повторение некоторого цикла явления, то на текущем спектре начинают сформировываться максимумы на основной частоте и ее гармониках. Эти максимумы становятся все более острыми и высокими, а значение спектральной плотности в интервалах между максимумами все убывает и — лишь в пределе, при $t \rightarrow \infty$, — сплошной текущий спектр вырождается в линейчатый спектр периодического в строгом смысле процесса.

Конечно, при достаточно больших длительностях процесса максимумы делаются настолько узкими, что их можно уже трактовать практически как линии. Однако это не умаляет принципиального значения всего сказанного выше — периодический процесс есть лишь предел, к которому может стремиться с течением времени реальный повторяющийся процесс.

Для уяснения высказанных соображений построим текущий спектр синусоиды. Применяя определение (5. 3) и подставляя в него $f(t) = \sin \Omega t$, найдем

$$S_t = \int_0^t e^{-j\omega t} \sin \Omega t dt = \\ = \frac{\Omega_i}{\Omega^2 - \omega^2} \left[1 - e^{-j\omega t} \left(\cos \Omega t + j \frac{\omega}{\Omega} \sin \Omega t \right) \right]. \quad (5. 4)$$

Формулу (5. 4) можно существенно упростить, рассматривая значения спектральной плотности для дискретных моментов

$$t = t_n = nT/2 = n\pi/\Omega.$$

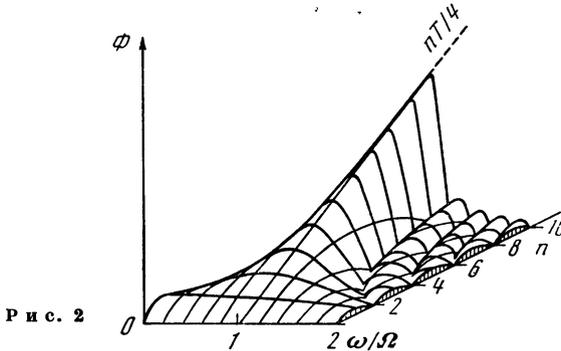
Подставив это значение в (5.4), получим

$$S_t = \frac{1}{\Omega} \cdot \frac{1}{1 - \left(\frac{\omega}{\Omega}\right)^2} [1 - (-1)^n e^{-jn\pi\omega\Omega}]$$

и спектр

$$\Phi_t = |S_t| = \frac{2}{\Omega} \cdot \frac{1}{1 - (\omega/\Omega)^2} \sin n \frac{\pi}{2} \cdot \frac{\omega}{\Omega}. \quad (5.5)$$

В этой формуле знак \sin относится к четному n , а знак \cos — к нечетному n . Величина n означает число полупериодов синусоиды с момента включения.



Неопределенность при $\omega = \Omega$ легко раскрывается

$$\Phi_t|_{\omega=\Omega} = t/2 = nT/4,$$

т. е. спектральная плотность на этой частоте нарастает со временем линейно.

Текущий спектр синусоиды, вычисленный по формуле (5.5), представлен на рис. 2 в виде рельефа. По горизонтальной оси, лежащей в плоскости чертежа, отложено отношение частот ω/Ω , по оси ординат — спектральная плотность; по горизонтальной оси, направленной от зрителя, — число полупериодов n . Это число, очевидно, пропорционально времени. Детали на левом склоне рельефа опущены, чтобы не усложнять чертежа.

Рис. 2 ясно показывает, что вначале спектр получается одно-родным; лишь постепенно сформировывается максимум на частоте Ω ; этот максимум с течением времени становится все более и более острым, но лишь в пределе при $t \rightarrow \infty$ фигура превратится в дискретную спектральную линию, которой мы изображаем периодическое, синусоидальное колебание. При этом спектральная плотность на частоте $\omega = \Omega$ будет бесконечно велика. Так оно и должно быть. По этому поводу нужно вспомнить то, что говорилось в § 4 о соотношении между спектром амплитуд составляющих периодической функции и спектральной плотностью непериодической функции.

§ 6. Мгновенный спектр

Введение понятия текущего спектра значительно расширяет круг спектральных представлений. Это понятие перекидывает мост между частотным и временным описанием явлений. Но ощущается потребность и в дальнейшем сближении частотной и временной точек зрения; в результате возникает еще одно понятие: понятие мгновенного спектра.

Обычное определение спектра (5. 1) дает функцию частоты: зависимость от времени выпадает, так как спектр отражает процесс в целом. Определение текущего спектра (5. 2) отражает всю предшествующую (вплоть до настоящего момента) историю процесса. Но имеет, очевидно, смысл и понятие о спектре, изменяющемся во времени и отражающем свойства процесса в данный момент. Представим себе, например, диалог между женщиной и женщиной, т. е. собеседников с низким и высоким голосом. Совершенно естественно представление, что, когда говорит мужчина, спектр звука располагается в области более низких частот, а когда говорит женщина, спектр перемещается в область более высоких частот. Остается дать этому интуитивно ясному понятию четкое математическое определение.

Простейшее определение мгновенного спектра может быть дано в следующем виде:

$$S_T(\omega, t) = \int_{t-T}^t f(\tau) e^{-j\omega\tau} d\tau, \quad (6. 1)$$

т. е. мгновенный спектр определен как спектр отрезка процесса длительностью T , непосредственно предшествующего данному моменту t . В этом определении мы имеем дело со скользящим интегрированием: интервал интегрирования имеет постоянную длину, но перемещается по оси времени; расположение интервала неизменно относительно текущего времени t .

Возможно и более общее определение мгновенного спектра, состоящее в том, что в подынтегральное выражение вводится скользящая (т. е. связанная с текущим временем) весовая функция. Определение принимает вид

$$S_r(\omega, t) = \int_{-\infty}^{+\infty} r(\tau - t) f(\tau) e^{-j\omega\tau} d\tau. \quad (6. 2)$$

Легко видеть, что определение (6. 1) есть частный случай (6. 2), если взять весовую функцию в виде

$$r(x) = \sigma(x + T) - \sigma(x),$$

где

$$\sigma(x) = \begin{cases} 0 & \text{при } x < 0, \\ 1 & \text{при } x > 0 \end{cases}$$

— единичная функция. Эти соотношения проясняются рис. 3. Фано [27] применяет определение (6.2), выбирая весовую функцию в виде¹

$$r(x) = e^{\alpha x} \sigma(-x)$$

(рис. 4). Эта функция учитывает все прошлое процесса, но с весом, экспоненциально убывающим по мере удаления от настоя-

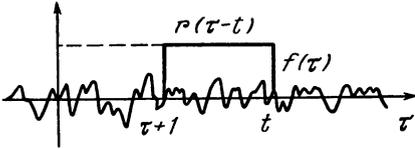


Рис. 3

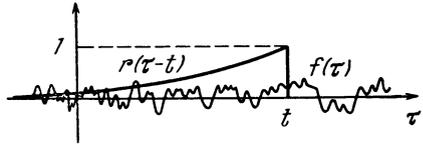


Рис. 4

щего момента. Основанием для такого выбора весовой функции является то, что она выражает фактический результат спектрального анализа при помощи реальных фильтров, характеризующихся постоянной времени $1/\alpha$. К этому вопросу нам еще придется вернуться в главе II.

Обратимся снова к определению (6.1) и перепишем его в виде

$$S_T(\omega, t) = \int_{-\infty}^t f(\tau) e^{-j\omega t} d\tau - \int_{-\infty}^{t-T} f(\tau) e^{-j\omega \tau} d\tau.$$

Таким образом, мгновенный спектр представлен разностью двух текущих спектров, или, иначе говоря, приращением, получаемым текущим спектром за промежуток времени T . При достаточно малом T это приращение могло бы быть выражено через производную текущего спектра по времени

$$\Delta S_t \approx \frac{\partial S_t}{\partial t} T.$$

Это рассуждение подводит нас к определению мгновенного спектра, предложенному Пэйджем [32]. Пэйдж определяет мгновенный спектр мощности следующим образом:

$$\rho(\omega, t) = \frac{\partial}{\partial t} |S_t(\omega)|^2, \quad (6.3)$$

где $S_t(\omega)$ — текущий спектр. Интеграл мгновенного спектра мощности по всей оси частот дает мгновенную мощность, т. е. квадрат функции $f(t)$

$$\frac{1}{\pi} \int_0^{\infty} \rho(\omega, t) d\omega = f^2(t).$$

¹ Здесь опущен фигурирующий у Фано постоянный множитель $(2\alpha)^{1/2}$, не имеющий для нас значения.

Интеграл от мгновенного спектра мощности по времени по всему прошлому дает квадрат модуля текущего спектра

$$\int_{-\infty}^t \rho(\omega, t) dt = |S_t(\omega)|^2.$$

Как видим, мгновенный спектр может определяться по-разному. Не следует этим смущаться: все определения произвольны. Нужно только выбрать целесообразное в данном случае (например, учитывающее способ действия измерительной аппаратуры) определение и четко сформулировать его с самого начала, чтобы ясно было, о чем идет речь. К сожалению, это элементарное требование далеко не всегда выполняется.

§ 7. Спектры модулированных колебаний

Перейдем теперь от общих соотношений к ряду приложений и займемся прежде всего вопросом о модуляции, играющим в современной технике очень важную роль.

Наибольшее значение имеет модуляция в технике связи. Всякий сигнал радиосвязи — будь то сигнал телеграфный, телефонный, телевизионный или любой другой — получается путем модуляции. Излучение радиостанции без модуляции подобно чистой странице, модулированное излучение подобно странице, на которой напечатаны те или иные буквы или знаки.

Большое значение имеет модуляция и в современной измерительной технике и в ряде специальных отраслей.

При передаче сигналов применяется некоторый физический агент, называемый переносчиком, и характеризующийся в отсутствие модуляции определенным числом постоянных параметров. Модуляция состоит в том, что тот или иной параметр переносчика изменяется во времени в соответствии с передаваемым сигналом. В простейшем случае, рассмотрением которого мы и ограничимся, в качестве переносчика применяется синусоидальное колебание¹. Запишем аналитическое выражение такого колебания

$$x = c_0 \sin(\omega_0 t + \varphi_0). \quad (7.1)$$

Здесь c_0 — амплитуда, ω_0 — частота, φ_0 — начальная фаза. В немодулированном колебании эти три параметра, полностью определяющих колебание, постоянны. В принципе возможно модулировать каждую из трех названных постоянных величин; мы будем иметь соответственно амплитудную модуляцию (АМ), частотную модуляцию (ЧМ) и фазовую модуляцию (ФМ). Разберем каждый вид модуляции подробно.

¹ При импульсном способе передачи переносчиком служит периодическая последовательность импульсов той или иной формы.

Воздействие, называемое модуляцией, мы можем характеризовать как умножение модулируемой величины на множитель $1 + mf(t)$, где $f(t)$ — модулирующая функция, определяемая так, чтобы $|f(t)| \leq 1$, а m — величина, характеризующая степень воздействия, могущая принимать значения от 0 до 1 и называемая глубиной модуляции.

При амплитудной модуляции модулированное колебание принимает вид

$$x = c_0 [1 + mf(t)] \sin(\omega_0 t + \varphi_0). \quad (7.2)$$

Начнем с простейшего случая синусоидальной модуляции, т. е. положим $f(t) = \sin \Omega t$. Подставляя в (7.2), получаем

$$\begin{aligned} x &= c_0 [\sin(\omega_0 t + \varphi_0) + m \sin(\omega_0 t + \varphi_0) \sin \Omega t] = \\ &= c_0 \left\{ \sin(\omega_0 t + \varphi_0) + \frac{m}{2} \cos[(\omega_0 - \Omega)t + \varphi_0] - \right. \\ &\quad \left. - \frac{m}{2} \cos[(\omega_0 + \Omega)t + \varphi_0] \right\}. \quad (7.3) \end{aligned}$$

Следовательно, модулированное колебание имеет дискретный спектр, состоящий из трех спектральных линий, как показано на рис. 5. Частота немодулированного колебания ω_0 носит название несущей частоты; возникшие в результате модуляции дополнительные частоты $\omega_0 - \Omega$ и $\omega_0 + \Omega$ называются боковыми частотами, или спутниками.

Итак, колебание с постоянной частотой, но с переменной амплитудой распадается на несколько синусоидальных колебаний. Напомним, что по определению синусоидальным является только колебание вида (7.1), т. е. колебание с постоянными амплитудой, частотой и фазой. Как угодно модулированная синусоида — это уже не синусоида. Эту истину можно выражать по-разному, но как бы мы ее ни формулировали, она является ключом к правильному пониманию вопросов модуляции.

Рассмотрим несколько более сложный случай, когда модулирующая функция периодична, т. е.

$$f(t) = \sum_{k=1}^{\infty} c_k \sin(k\Omega t + \varphi_k).$$

Тогда

$$\begin{aligned} x &= c_0 \left[1 + m \sum_{k=1}^{\infty} c_k \sin(k\Omega t + \varphi_k) \right] \sin(\omega_0 t + \varphi_0) = \\ &= c_0 \left\{ \sin(\omega_0 t + \varphi_0) + \frac{m}{2} \sum_{k=1}^{\infty} c_k \cos[(\omega_0 - k\Omega)t + \varphi_0 - \right. \\ &\quad \left. - \varphi_k] - \frac{m}{2} \sum_{k=1}^{\infty} c_k \cos[(\omega_0 + k\Omega)t + \varphi_0 + \varphi_k] \right\}. \end{aligned}$$

Модулированное колебание состоит из несущей частоты и двух групп, называемых боковыми полосами (обе суммы в фигурных скобках). Спектр модулированного колебания изображен на рис. 6.¹

Следует заметить, что правая боковая полоса воспроизводит спектр модулирующей функции, а левая представляет собой зеркальное отражение правой. Таким образом, при процессе модуляции осуществляется транспозиция (перенос) спектра модулирующей функции; спектр смещается на величину ω_0 по шкале

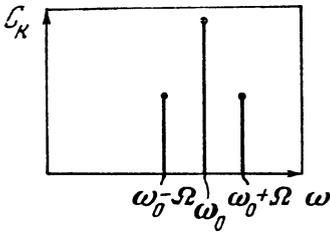


Рис. 5

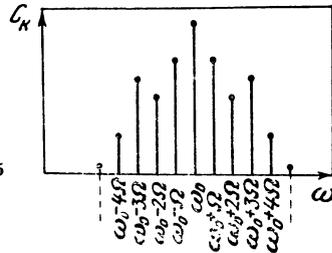


Рис. 6

частот. Общие заключения по этому поводу можно вывести из теоремы (4. 6). Нам еще представится случай вернуться к этому.

Очевидно, что если несущая частота и основная частота модулирующей функции несоизмеримы, то получаемое модулированное колебание непериодично и изображенный на рис. 6 спектр квазигармоничен.

Следует добавить, что при так называемой балансовой модуляции несущая частота отсутствует, и спектр состоит только из боковых полос. Математически это выражается тем, что амплитуда несущей частоты умножается не на $1 + mf(t)$, а просто на $f(t)$. Действительно, слагающая несущей частоты появляется в модуляционном спектре только за счет постоянной составляющей в выражении $1 + mf(t)$. При балансовой модуляции осуществляется простое перемножение модулирующей функции на колебание несущей частоты. Представим последнее в виде

$$x = c_0 \cos \omega_0 t = \frac{c_0}{2} (e^{j\omega_0 t} + e^{-j\omega_0 t}).$$

Для балансно-модулированного колебания получим

$$x = \frac{c_0}{2} (e^{j\omega_0 t} + e^{-j\omega_0 t}) f(t).$$

Воспользовавшись теоремой (4. 6), найдем, что спектр модулированного колебания есть

$$\frac{c_0}{2} S(\omega_0 - \omega) + S(\omega_0 + \omega),$$

где $S(\omega)$ — спектр модулирующей функции $f(t)$. Два члена в квадратных скобках выражают обе боковые полосы.

Представляет интерес вопрос о ширине модуляционного спектра; этот вопрос рассмотрен в Добавлении I.

Перейдем к несколько более трудному вопросу о частотной модуляции. Нужно сказать, что история развития теории модуляции дает наибольшее количество примеров путаницы в понятиях, связанной с неправильным применением спектрального воззрения на колебания. Даже сравнительно простая амплитудная модуляция послужила (и притом не так давно) поводом для совершенно ошибочных высказываний. Еще хуже обстояло дело с частотной модуляцией. Сущности дела долго не понимали инженеры и исследователи. Поучительно воспроизвести распространенное в свое время рассуждение о свойствах частотной модуляции: при частотной модуляции мы имеем колебание, частота которого непрерывно изменяется в пределах заданного нами интервала $\omega_0 \pm \Delta \omega$, от частоты модуляции Ω зависит частота изменения несущей частоты, но не величина этого изменения. Стало быть, спектр колебания должен быть сплошным, так как частота пробегает все значения в пределах интервала $\omega_0 \pm \Delta \omega$, а ширина спектра должна составлять $2\Delta \omega$. А так как эта ширина назначается нами произвольно, то можно сократить полосу, занимаемую на шкале частот передающей радиостанцией.

Все здесь неверно: и спектр получается не сплошной, а дискретный, и ширина его при узком интервале $2\Delta \omega$ (так называемая полоса качания) не зависит вовсе от величины этого интервала, а определяется, как и в случае АМ, шириной спектра модулирующей функции. Наконец, когда ЧМ получила практическое применение, то оказалось, что, имея значительные специфические преимущества, она требует полосы, раз в 15—20 более широкой, чем та, которая отпускается по международным правилам на АМ.

Выведем теперь основные соотношения. Предположим, что частота модулируется по косинусоидальному закону

$$\omega = \omega_0 + \Delta \omega \cos \Omega t = \omega_0 \left(1 + \frac{\Delta \omega}{\omega_0} \cos \Omega t \right),$$

где $\Delta \omega$ — частотное отклонение, а $\Delta \omega / \omega_0$ — относительное изменение, т. е. глубина модуляции частоты.

По своему определению круговая частота есть производная по времени от аргумента тригонометрической функции, представляющей колебание. Поэтому мы можем записать для частотно-модулированного колебания при синусоидальной модуляции

$$\begin{aligned} x = c_0 \sin \theta = c_0 \sin \left(\int_0^t \omega dt \right) &= c_0 \sin \left[\omega_0 t + \right. \\ &+ \left. \frac{\Delta \omega}{\Omega} \sin \Omega t \right] = c_0 [\sin \omega_0 t \cos (\beta \sin \Omega t) + \\ &+ \cos \omega_0 t \sin (\beta \sin \Omega t)], \end{aligned} \quad (7.4)$$

где $\beta = \Delta \omega / \Omega$ — так называемый индекс модуляции.

Рассмотрим сначала соотношения при малом индексе β . Замена в (7.4) \cos и \sin малого аргумента соответственно единицей и самим аргументом, получим

$$x \approx c_0 (\sin \omega_0 t + \beta \sin \Omega t \cos \omega_0 t),$$

$$\beta \ll 1$$

т. е. выражение, ничем не отличающееся от выражения для АМ колебания. Спектр ЧМ колебания при синусоидальной модуляции с малым индексом, так же как и спектр АМ колебания, состоит из несущей частоты и двух спутников с частотами $\omega_0 \pm \Omega$. Обратимся теперь к общему случаю, т. е. к случаю произвольного индекса β .

Используя известные формулы теории бesselевых функций¹ [11]

$$\cos(x \sin \varphi) = J_0(x) + 2 \sum_{n=1}^{\infty} J_{2n}(x) \cos 2n\varphi,$$

$$\sin(x \sin \varphi) = 2 \sum_{n=0}^{\infty} J_{2n+1}(x) \sin(2n+1)\varphi,$$

$$\cos(x \cos \psi) = J_0(x) + 2 \sum_{n=1}^{\infty} (-1)^n J_{2n}(x) \cos 2n\psi,$$

$$\sin(x \cos \psi) = 2 \sum_{n=1}^{\infty} (-1)^n J_{2n+1}(x) \cos(2n+1)\psi,$$

найдем

$$x = c_0 \left\{ \sin \omega_0 t \left[J_0(\beta) + 2 \sum_{n=1}^{\infty} J_{2n}(\beta) \cos 2n\Omega t \right] + \right. \\ \left. + \cos \omega_0 t \left[2 \sum_{n=0}^{\infty} J_{2n+1}(\beta) \sin(2n+1)\Omega t \right] \right\}.$$

Перемножая под знаками сумм, получим окончательно

$$x = c_0 \left\{ J_0(\beta) \sin \omega_0 t + \sum_{k=1}^{\infty} J_k(\beta) [\sin(\omega_0 + k\Omega)t + \right. \\ \left. + (-1)^k \sin(\omega_0 - k\Omega)t] \right\}. \quad (7.5)$$

Мы имеем, таким образом, колебание с линейчатым спектром. В отличие от АМ здесь при синусоидальной модуляции возникает бесконечный спектр. Однако практически он ограничен. Дело в том, что, как видим, амплитуды гармоник пропорциональны $J_k(\beta)$, а эти функции обладают своеобразным свойством: они сохраняют весьма малое значение до значений β тем больших, чем выше порядок k .

В практике пользуются понятием действительной ширины полосы, занимаемой спектром ЧМ колебания. Действительная ши-

¹ Последние две формулы можно получить, положив $\varphi = \pi/2 - \psi$.

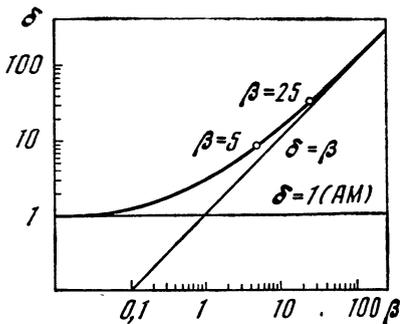


Рис. 7

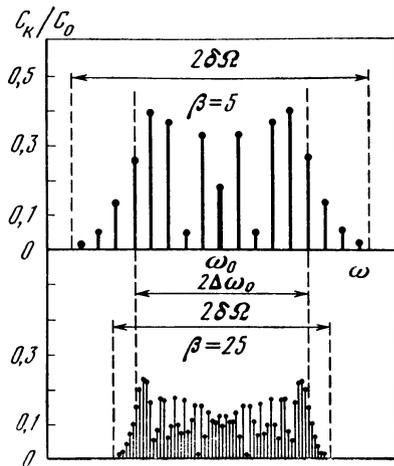


Рис. 8 ▶

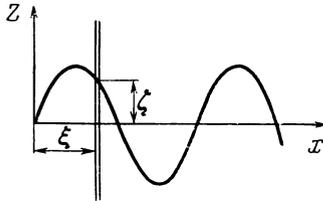
рина есть интервал на шкале частот, вне которого гармоники имеют относительную величину $\leq 0,01$. Пользуясь таблицами бесселевых функций, можно найти границы этого интервала, а также его зависимость от индекса модуляции β .

Эта зависимость имеет, очевидно, ступенчатый характер: ее сглаженный график показан на рис. 7. По оси ординат отложено обозначенное через δ отношение половины действительной ширины спектра модулированного колебания (т. е. ширины одной боковой полосы) к ширине спектра модулирующей функции. По оси абсцисс отложен индекс модуляции β . Нужно заметить, что представленная на рис. 7 зависимость с достаточной для практических надобностей точностью аппроксимируется простой формулой $\delta = 1 + \beta$.

Кривая зависимости δ от β вписывается в угол, образованный двумя прямыми. Горизонтальная прямая $\delta = 1$ отвечает условиям, существующим при амплитудной модуляции; как мы уже знаем, при АМ ширина полосы, занимаемой модуляционным спектром, всегда равна двойной ширине спектра модулирующей функции. Наклонная прямая $\delta = \beta$ является асимптотой нашей кривой при возрастающем β . Как видим, в пределе действительно ширина спектра делается равной полосе качания. Однако это соотношение достигается как раз при очень больших β , тогда как в вышеприведенном высказывании ожидалось, что это соотношение можно получить при очень малых β , на чем и основывалась предполагаемая возможность сужения полосы частот за счет применения ЧМ. В действительности же при ЧМ величина δ всегда больше единицы; другими словами, полоса частот, занимаемая спектром, при ЧМ всегда больше, чем полоса частот при АМ. Только при самых малых β полосы сравниваются.

При малых β спектр модулированного колебания при ЧМ получается бедный, практически не отличающийся от спектра при АМ, при синусоидальной модуляции он состоит практически из двух боковых линий (остальные очень малы). Такое положение сохраняется до тех пор, пока δ мало отличается от единицы. В дальнейшем число достигающих заметной величины гармоник растет. Величина δ непосредственно выражает число этих гармоник.

Интересно проследить вид спектра при больших значениях δ и β . На рис. 8 показаны спектры при синусоидальной ЧМ для $\beta=5$ и 25 .



Р и с. 9

Считая наибольшее значение $\Delta f = \Delta \omega / 2\pi = 75$ кГц, получим, что $\beta=5$ соответствует модулирующей частоте $F = \Omega / 2\pi = 15$ кГц, а $\beta=25$ соответствует частоте $F = 3$ кГц. На рис. 8 отмечена полоса качания $2\Delta\omega$. Как видим, действительная ширина спектра, равная $2\delta\Omega$, приближается к полосе качания. Величина $2\delta\Omega$ также отмечена на рис. 8. Два спектра на рис. 8 отвечают соотношениям между δ и β , отмеченным соответствующими двумя точками на кривой рис. 7.

Читатель, не привыкший иметь дело с колебаниями переменной частоты, может быть, испытывает затруднения с истолкованием выражения

$$\sin \theta = \sin \left(\int_0^t \omega dt \right).$$

Мы поясним этот пункт, рассмотрев простой пример. Предположим, что имеется правильно записанная фонограмма чистого тона, т. е. синусоидального колебания. Пусть теперь эта фонограмма воспроизводится на аппарате, в котором линейная скорость продвижения фонограммы непостоянна. Возникающие при этом искажения представляют собой, как известно, не что иное, как результат частотной модуляции.

Выведем соответствующие соотношения. Положим, что записано колебание $y = \sin \omega_0 t$. Фонограмма, полученная при постоянной линейной скорости v_0 , выражается формулой

$$z = \sin 2\pi \frac{x}{\lambda},$$

где $x = v_0 t$ — координата, отсчитываемая вдоль фонограммы; $\lambda = 2\pi v_0 / \omega_0$ — длина записанной волны. Пусть теперь при воспроизведении фонограмма движется с переменной скоростью $v(t)$. Читающий орган, например оптическая щель, воспринимает величину

$$\zeta = \sin 2\pi \frac{\zeta}{\lambda},$$

где ζ — ордината фонограммы в точке ξ ; ξ — путь, пройденный щелью относительно фонограммы (рис. 9). Но этот путь, очевидно, равен

$$\xi = \int_0^t v dt$$

и, следовательно,

$$\zeta = \sin \frac{2\pi}{\lambda} \left(\int_0^t v dt \right).$$

[Это совершенно аналогично выражению $\sin \theta = \sin \left(\int_0^t \omega dt \right)$. Ведь θ — это угол, на который поворачивается за время t вектор, вращающийся с переменной угловой скоростью ω .] Если, например,

$$v = v_0 \left(1 + \frac{\Delta v}{v_0} \cos \Omega t \right),$$

то

$$\begin{aligned} \zeta &= \sin \frac{2\pi}{\lambda} \left[\int_0^t (v_0 + \Delta v \cos \Omega t) dt \right] = \\ &= \sin \frac{2\pi}{\lambda} \left(v_0 t + \frac{\Delta v}{\Omega} \sin \Omega t \right) = \sin \left(\omega_0 t + \frac{\Delta \omega}{\Omega} \sin \Omega t \right), \end{aligned}$$

т. е. имеется типичное частотно-модулированное колебание (ср. (7.4)). В этом выражении

$$\Delta \omega = \frac{2\pi}{\lambda} \Delta v = \omega_0 \frac{\Delta v}{v_0}.$$

Обратимся теперь к фазовой модуляции. По общему определению мы должны, взяв немодулированное колебание

$$x = c_0 \sin (\omega_0 t + \varphi_0),$$

заменить в нем φ_0 величиной

$$\varphi = \varphi_0 + \Delta \varphi f(t).$$

Тогда мы получим

$$x = c_0 \sin [\omega_0 t + \varphi_0 + \Delta \varphi f(t)]. \quad (7.6)$$

Для частотной модуляции, вводя переменную частоту $\omega = \omega_0 + \Delta \omega f(t)$, мы получили бы

$$x = c_0 \sin \left[\omega_0 t + \varphi_0 + \Delta \omega \int_0^t f(t) dt \right] = \\ = c_0 \sin [\omega_0 t + \varphi_0 + \Delta \omega F(t)]. \quad (7.7)$$

Сравнивая (7.6) и (7.7), видим, что ФМ и ЧМ в сущности дают совершенно одинакового вида колебания. Разница заключается лишь в том, что при ФМ в аргумент синусоидальной функции входит модулирующая функция $f(t)$, а при ЧМ — ее интеграл.

Если модуляция синусоидальна, то различия в форме модулированных колебаний и их спектрах вообще усмотреть нельзя, так как интеграл от синусоиды есть косинусоида, т. е. опять-таки синусоида, лишь сдвинутая по фазе на $\pi/2$. Различия все же есть, но более тонкое. Дело в том, что если $f(t) = \sin \Omega t$, то

$$F(t) = -\frac{1}{\Omega} \cos \Omega t.$$

Появление множителя $1/\Omega$ означает необходимость введения соответствующих корректирующих цепей.

Если предположить, что дискриминаторы, т. е. элементы, превращающие ФМ и ЧМ в АМ, ведут себя одинаково, то в дальнейшей схеме приемников ФМ должно быть добавлено одно интегрирующее звено по сравнению с аналогичной схемой приемника ЧМ.

Дальнейшее сравнение ФМ и ЧМ завело бы нас очень далеко в специальные области современной радиотехники. Мы ограничимся высказанными общими соображениями и попытаемся лишь в заключение пояснить различие между всеми тремя видами модуляции по возможности наглядным образом.

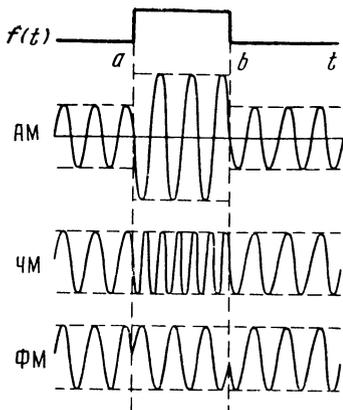
Как уже говорилось, при синусоидальной ЧМ и ФМ различия в форме модулированных колебаний усмотреть нельзя. Но различие это обнаруживается при более сложном законе модуляции, т. е. когда модулирующая функция обладает более или менее богатым спектром.

Предположим, что модулирующая функция представляет собой импульс следующего вида:

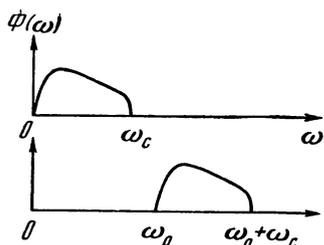
$$f(t) = \begin{cases} 0 & (t < a), \\ 1 & (a < t < b), \\ 0 & (t > b). \end{cases}$$

Тогда при $t=a$ скачком изменится амплитуда при АМ, частота при ЧМ и фаза при ФМ.

Обратное скачкообразное изменение до первоначальных значений произойдет при $t=b$. Это изображено на рис. 10. Как видим, имеется различие между модулированными колебаниями



Р и с. 10



Р и с. 11

в случае ЧМ и ФМ. Заметим кстати, что скачкообразное изменение фазы можно толковать как результат бесконечно быстрого изменения частоты на протяжении бесконечно малого промежутка времени. В конце концов все это сводится опять к основным определениям

$$\omega = d\theta/dt, \quad \theta = \int \omega dt.$$

§ 8. Перенос спектра

Для техники представляет большой интерес вопрос о возможности такого преобразования заданной функции времени, в результате которого спектр функции переместился бы по шкале частот. При обычной модуляции или преобразовании частоты получаются, как известно, две боковые полосы. Спрашивается, каким образом можно получить спектр, состоящий только из одной боковой полосы (ОБП)? Требуемое преобразование спектра показано на рис. 11. Мы предполагаем, что исходный спектр ограничен сверху частотой ω_c .

Нужно сразу пояснить, что теорема смещения (§ 4) не дает требуемого преобразования, так как она относится к смещению комплексного спектра $S(\omega)$, тогда как нам требуется сместить вещественный спектр амплитуд $\Phi(\omega) = |S(\omega)|$ [12].

Простейший способ получения смещенного спектра состоит в том, что сначала посредством балансной модуляции несущей частоты образуют двухполосный модуляционный спектр, а затем при помощи фильтров подавляют нижнюю боковую полосу. Представляют интерес и другие возможности, две из которых мы здесь рассмотрим.

Будем исходить из выражения (2. 14)

$$f(t) = \frac{1}{\pi} \operatorname{Re} \int_0^{\infty} S(\omega) e^{j\omega t} d\omega. \quad (8. 1)$$

Нам нужно построить функцию, которая имела бы спектральную плотность $S(\omega)$ на частоте $\omega_0 + \omega$. Для такой функции можем записать

$$f_1(t) = \frac{1}{\pi} \operatorname{Re} \int_0^{\infty} S(\omega) e^{j(\omega_0 + \omega)t} d\omega. \quad (8.2)$$

Это выражение можно преобразовать следующим образом:

$$\begin{aligned} f_1(t) &= \frac{1}{\pi} \operatorname{Re} e^{j\omega_0 t} \int_0^{\infty} S(\omega) e^{j\omega t} d\omega = \\ &= \frac{1}{\pi} \left(\cos \omega_0 t \operatorname{Re} \int_0^{\infty} S(\omega) e^{j\omega t} d\omega - \sin \omega_0 t \operatorname{Im} \int_0^{\infty} S(\omega) e^{j\omega t} d\omega \right) \end{aligned}$$

или

$$f_1(t) = f(t) \cos \omega_0 t + f^V(t) \sin \omega_0 t. \quad (8.3)$$

Определим функцию

$$f^V(t) = -\frac{1}{\pi} \operatorname{Im} \int_0^{\infty} S(\omega) e^{j\omega t} d\omega. \quad (8.4)$$

Для этого представим комплексный спектр в виде

$$S(\omega) = A(\omega) + jB(\omega).$$

Подставляя в (8.4), получим

$$\begin{aligned} f^V(t) &= -\frac{1}{\pi} \int_0^{\infty} (A \sin \omega t + B \cos \omega t) d\omega = \\ &= \frac{1}{\pi} \int_0^{\infty} \left[A \cos \left(\omega t + \frac{\pi}{2} \right) - B \sin \left(\omega t + \frac{\pi}{2} \right) \right] d\omega, \end{aligned} \quad (8.5)$$

тогда как для функции $f(t)$ аналогичным путем получаем

$$f(t) = \frac{1}{\pi} \int_0^{\infty} (A \cos \omega t - B \sin \omega t) d\omega. \quad (8.6)$$

Сопоставляя (8.5) и (8.6), видим, что функция $f^V(t)$ отличается от $f(t)$ тем, что все составляющие повернуты по фазе на $\pi/2$. Следовательно, для образования функции $f^V(t)$ нужно распола-

* Комплексная функция $F(t) = \frac{1}{\pi} \int_0^{\infty} S(\omega) e^{j\omega t} d\omega$, вещественная и мнимая составляющие которой здесь фигурируют, носит название «аналитического сигнала» и оказывается удобной в некоторых теоретических исследованиях.

гать фазовращателем, поворачивающим фазу на $\pi/2$ для всех частот, т. е. четырехполюсником, комплексный коэффициент передачи которого равен

$$K = j = e^{j\pi/2}. \quad (8.7)$$

Четырехполюсник с такими свойствами физически неосуществим; можно, однако, построить реальный четырехполюсник, в котором характеристика (8.8) осуществляется с тем или иным приближением в конечном интервале частот (исключая $\omega=0$ и ближайшую окрестность этого значения).

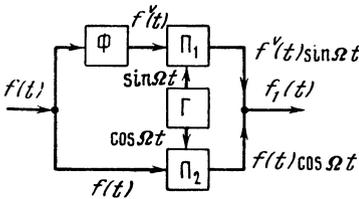


Рис. 12

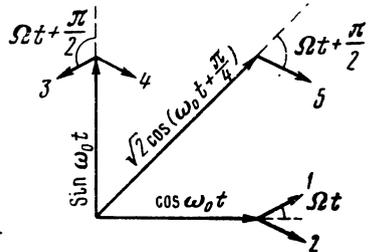


Рис. 13 ▶

Требуемое преобразование (8.3) выполняется скелетной схемой рис. 12, из которой: Φ — фазовращатель, поворачивающий фазу на $\pi/2$; Γ — двухфазный генератор, дающий два сдвинутых на $\pi/2$ колебания; $\sin \omega_0 t$ и $\cos \omega_0 t$; Π_1 и Π_2 — множители (балансные модуляторы).

Сущность преобразования (8.3) поясняется векторной диаграммой рис. 13, на которой представлен случай

$$f(t) = \cos \Omega t, \quad f^V(t) = -\sin \Omega t = \cos(\Omega t + \pi/2).$$

Для наглядности принята не балансная, а обычная АМ, так что составляющая несущей частоты ω_0 не уничтожается. По горизонтали отложен вектор несущего колебания $\cos \omega_0 t$; в результате модуляции функцией $f(t)$ появляются два спутника, представленные векторами 1 и 2, расположенными под углом $\pm \Omega t$ по отношению к основному вектору. По вертикали отложен вектор $\sin \omega_0 t$; модуляция функцией $f^V(t)$ дает два спутника, векторы которых 3 и 4 расположены под углом $\pm(\Omega t + \pi/2)$ к основному. Направления векторов 2 и 4 совпадают, а направления векторов 1 и 3 — противоположны. В результате сложения получается вектор несущей частоты и один-единственный спутник — вектор 5. Второй спутник устраняется, и это и есть однополосная синусоидальная модуляция¹.

¹ Векторная диаграмма показывает, что при наличии несущей однополосную модуляцию можно рассматривать как сложную амплитудно-фазовую модуляцию, на чем и основан один из способов получения ОБП, который мы не рассматриваем.

Можно показать, что рассмотренное преобразование есть частный случай многофазной модуляции, так что описанный механизм можно назвать двухфазной однополосной модуляцией. С технической точки зрения больший интерес представляет трехфазная однополосная модуляция [22].

Рассмотрим теперь другую возможность переноса спектра. Спектр, ограниченный сверху частотой ω_c , может быть представлен рядом Фурье (см. § 14)

$$S(\omega) = \Delta t \sum_{-\infty}^{\infty} f_k e^{-jk\omega\Delta t}, \quad (8.8)$$

где $\Delta t = \pi/\omega_c$, $f_k = f(k\Delta t)$. Спектр $S_1(\omega)$, сдвинутый на ω_0 по шкале частот, связан с исходным спектром $S(\omega)$ соотношением

$$S_1(\omega_0 + \omega) = S(\omega),$$

так что

$$S_1(\omega) = \Delta t \sum_{-\infty}^{\infty} f_k e^{jk(\omega_0 - \omega)\Delta t}. \quad (8.9)$$

Соответствующая этому спектру функция времени выражается соотношением

$$f_1(t) = \frac{1}{\pi} \operatorname{Re} \int_{\omega_0}^{\omega_0 + \omega_c} S_1(\omega) e^{j\omega t} d\omega. \quad (8.10)$$

Подставляя сюда выражение (8.8) для спектра, находим [8, 9]

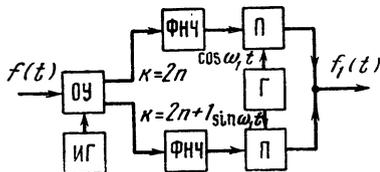
$$f_1(t) = \sum_{-\infty}^{\infty} f_k \frac{\sin \frac{\omega_c}{2}(t - k\Delta t)}{\frac{\omega_c}{2}(t - k\Delta t)} \cos\left(\omega_1 t - k \frac{\pi}{2}\right), \quad (8.11)$$

где $\omega_1 = \omega_0 + \omega_c/2$ — центральная частота перенесенного спектра. Для исходной функции $f(t)$, спектр которой выражается формулой (8.8), имеем

$$f(t) = \sum_{-\infty}^{\infty} f_k \frac{\sin \omega_c(t - k\Delta t)}{\omega_c(t - k\Delta t)}. \quad (8.12)$$

Рассмотрение формулы (8.11) указывает характер преобразования, которому нужно подвергнуть исходную функцию $f(t)$, чтобы получить функцию $f_1(t)$ со смещенным спектром. Такое преобразование может быть выполнено скелетной схемой рис. 14. На этой схеме: ИГ — импульсный генератор с периодом следования Δt ; ОУ — отсчетное устройство, назначение которого состоит в том, чтобы в моменты, задаваемые импульсами, поступающими от ИГ, брать отсчеты мгновенных значений функции $f(t)$, т. е. определять величины f_k и выдавать короткие импульсы с площадью, пропорциональной f_k . Кроме того, требуется, чтобы эти

импульсы, во-первых, получали требуемый знак, а во-вторых, чтобы они поочередно поступали на два выхода устройства, т. е. чтобы, например, импульсы четных номеров поступали на верхний выход, а нечетных номеров — на нижний. Далее следуют фильтры нижних частот ФНЧ с верхней граничной частотой $\omega_c/2$. Затем идут два перемножителя П, на вторые входы которых поступают сдвинутые по фазе на $\pi/2$ напряжения частоты $\omega_1 = \omega_0 + \omega_c/2$ от двухфазного генератора Г. Напряжения, снимаемые с обоих перемножителей, складываются.



Р и с. 14

Описанная схема интересна тем, что в ней отсутствует широкополосный фазовращатель. Однако для получения хороших результатов нужно, по-видимому, предъявить строгие требования к фильтрам нижних частот.

§ 9. Преобразование спектров при детектировании

Если некоторое периодическое колебание x подвергнуть нелинейной операции $y = \varphi(x)$, то полученное в результате этой операции колебание y будет обладать спектром, отличным от спектра x и, как правило, более богатым. Так, например, если первоначальное колебание представляет собой сумму двух синусоид

$$x = c_1 \sin \omega_1 t + c_2 \sin \omega_2 t$$

и, следовательно, имеет спектр, состоящий из двух спектральных линий, то после нелинейной операции φ мы получим в составе колебания y спектральные составляющие с частотами

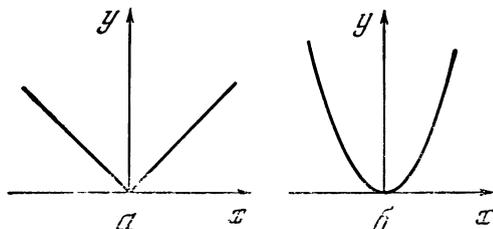
$$\omega_{mn} = m\omega_1 \pm n\omega_2,$$

где m и n — в общем случае любые положительные целые числа. Такого рода спектр носит название комбинационного, а частоты ω_{mn} называются комбинационными частотами.

Таким изменением спектра пользуются для измерения степени отклонения данной системы от линейности. На вход системы подается колебание x , изучается спектр получаемого на выходе колебания y . Комбинационные частоты не возникают в том единственном случае, когда φ (выражающая в данном случае характеристику исследуемой системы) — линейная функция.

Специальный вид нелинейной операции, преобразовывающей спектр, называется детектированием. В составе модулированного

колебания составляющей с частотой модуляции нет. Но эта составляющая нам нужна, так как она-то представляет собой передаваемый сигнал. Для того чтобы она вновь появилась, надо подвергнуть модулированное колебание некоторой нелинейной операции. Эта операция, имеющая целью образование составляющей с частотой модуляции, и носит название детектирования («детектирование» — обнаружение). В результате детектирования получается сложное колебание, в состав которого входит в качестве одной из составляющих интересующее нас колебание с частотой модуляции. Дальнейшее разделение слагаемых не представляет уже никаких затруднений.



Р и с. 15

Рассмотрение вопросов детектирования в общем виде было бы очень громоздким; мы ограничимся несколькими простейшими примерами.

Рассмотрим простое АМ колебание при синусоидальной модуляции. Первоначально мы имели колебание несущей частоты $\sin \omega_0 t$ и модулирующее колебание $1 + m \sin \Omega t$. В результате некоторой операции, которую мы называем модуляцией, оба эти колебания оказываются перемноженными, и мы имеем

$$x = (1 + m \sin \Omega t) \sin \omega_0 t.$$

В составе модулированного колебания, как мы знаем, уже нет составляющей с частотой Ω ; спектр x состоит из трех линий с частотами ω_0 , $\omega_0 + \Omega$ и $\omega_0 - \Omega$.

Если мы желаем теперь снова получить колебание с частотой Ω , то должны соответствующим образом продетектировать x . Операция детектирования в данном случае производит действие, обратное операции модуляции, поэтому в применении к модулированным сигналам детектирование называют иногда демодуляцией. Мы будем рассматривать только два основных вида детекторов: «линейный» детектор $y = |x|$ (рис. 15, а) и квадратичный детектор $y = x^2$ (рис. 15, б). Слово «линейный» поставлено для первого раза в кавычки, чтобы подчеркнуть, что на самом деле линейный детектор нелинеен и что линейный в подлинном смысле детектор невозможен (т. е. линейная система не детектируется).

Для детектирования модулированного колебания пригодно линейное детектирование. Воспользовавшись тем, что абсолют-

ная величина произведения равна произведению абсолютных величин сомножителей, можем записать

$$y = |x| = (1 + m \sin \Omega t) |\sin \omega_0 t|.$$

Но функция $|\sin \omega_0 t|$ может быть представлена следующим рядом Фурье:

$$|\sin \omega_0 t| = \frac{2}{\pi} \left(1 - \sum_{k=1}^{\infty} \frac{1}{4k^2 - 1} \cos 2k\omega_0 t \right),$$

откуда

$$y = |x| = \frac{2}{\pi} \left\{ (1 + m \sin \Omega t) - \sum_{k=1}^{\infty} \frac{1}{4k^2 - 1} \left[\cos 2k\omega_0 t + \right. \right. \\ \left. \left. + \frac{m}{2} \sin (2k\omega_0 + \Omega) t - \frac{m}{2} \sin (2k\omega_0 - \Omega) t \right] \right\}.$$

В этом выражении первый член (в круглых скобках) — модулирующая функция, которую мы и стремились получить; второй член объединяет под знаком суммы составляющие высоких частот $2k\omega_0$, $2k\omega_0 + \Omega$ и $2k\omega_0 - \Omega$, которые нетрудно отделить.

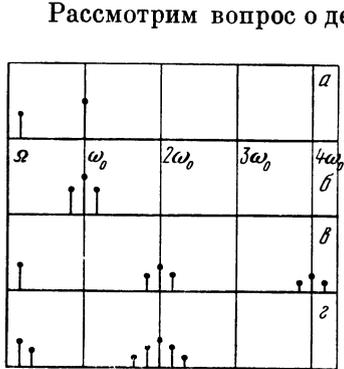
Если бы мы подвергли модулированное колебание квадратичному детектированию, то получили бы

$$y = x^2 = (1 + m \sin \Omega t)^2 \sin^2 \omega_0 t = \frac{1}{2} \left\{ 1 + \frac{m^2}{2} + \right. \\ \left. + 2m \sin \Omega t - \frac{m^2}{2} \cos 2\Omega t - \left(1 + \frac{m^2}{2} \right) \cos 2\omega_0 t - \right. \\ \left. - m \sin (2\omega_0 - \Omega) t + m \sin (2\omega_0 + \Omega) t + \right. \\ \left. + m^2 \cos 2(\omega_0 - \Omega) t + m^2 \cos 2(\omega_0 + \Omega) t \right\}.$$

Таким образом, в этом случае кроме постоянной составляющей $1 + m^2/2$ и пяти спектральных линий с высокими частотами: $2\omega_0$, $2\omega_0 - \Omega$, $2\omega_0 + \Omega$, $2(\omega_0 - \Omega)$, $2(\omega_0 + \Omega)$, мы получаем две спектральные линии с низкими частотами Ω и 2Ω . Следовательно, спектр модулирующего колебания, состоявший первоначально из одной линии с частотой Ω , оказывается искаженным, и данный вид детектирования может применяться только при очень малой глубине модуляции (так как отношение амплитуд второй и первой гармоник равно $m/4$).

На рис. 16 изображены разобранные случаи. На рис. 16, а представлены два исходных колебания — несущая частота ω_0 и модулирующая частота Ω ; на рис. 16, б изображен спектр модулированного колебания — несущая частота и два спутника; на

рис. 16, а представлен спектр колебания, получаемого в результате линейного детектирования модулированного колебания (нужно отметить появление линии с частотой Ω). Наконец, на рис. 16, з представлен спектр, получаемый в результате квадратичного детектирования (спектр ограничен, но имеются две линии низкой частоты: Ω и 2Ω).



Р и с. 16

Рассмотрим вопрос о детектировании биений. Биениями называют интерференционное явление, состоящее в периодическом изменении амплитуды результирующего колебания, составленного из двух простых синусоидальных колебаний с неравными частотами. Говорят, что частота биений равна разности частот образующих колебаний¹.

Положим, что мы ставим своей задачей получение в результате детектирования синусоидального колебания с разностной частотой. Мы имеем

$$x = \sin \omega_1 t + \sin \omega_2 t.$$

В данном случае следует применить квадратичное детектирование, которое дает

$$\begin{aligned} y = x^2 &= \sin^2 \omega_1 t + \sin^2 \omega_2 t + 2 \sin \omega_1 t \sin \omega_2 t = \\ &= 1 - \frac{1}{2} \cos 2\omega_1 t - \frac{1}{2} \cos 2\omega_2 t - \frac{1}{2} \cos (\omega_1 + \omega_2) t + \\ &\quad + \frac{1}{2} \cos (\omega_1 - \omega_2) t. \end{aligned}$$

Как видим, мы получили кроме постоянной составляющей и высоких частот $2\omega_1$, $2\omega_2$ и $\omega_1 + \omega_2$ требуемое колебание с частотой $\omega_1 - \omega_2$. Линейное детектирование этого не дает. Применяв его, мы получили бы

$$\begin{aligned} y = |x| &= |\sin \omega_1 t + \sin \omega_2 t| = 2 \left| \sin \frac{\omega_1 + \omega_2}{2} t \right| \left| \cos \frac{\omega_1 - \omega_2}{2} t \right| = \\ &= \frac{8}{\pi^2} \left[1 - \sum_{k=1}^{\infty} \frac{1}{4k^2 - 1} \cos k (\omega_1 + \omega_2) t \right] \times \\ &\quad \times \left[1 - \sum_{k=1}^{\infty} \frac{(-1)^k}{4k^2 - 1} \cos k (\omega_1 - \omega_2) t \right], \end{aligned}$$

¹ Это верно, если под частотой биений понимать частоту повторения максимумов или минимумов огибающей результирующего колебания. И это неверно, если под частотой биений понимать основную частоту наблюдаемого процесса. При несоизмеримых частотах образующих колебаний результирующее колебание вообще непериодично.

т. е., кроме постоянной составляющей и высоких частот, мы имели бы бесконечный спектр низкочастотных слагаемых с частотами $k(\omega_1 - \omega_2)$.

Квадратичное детектирование дает при детектировании биений нужный результат и в том более сложном случае, когда амплитуды образующих колебаний не равны, т. е. когда

$$x = \sin \omega_1 t + \varepsilon \sin \omega_2 t$$

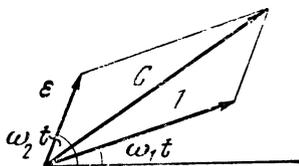
($\varepsilon \geq 1$). Выражение для огибающей биений при равенстве амплитуд есть

$$c(t) = \left| \cos \frac{\omega_1 - \omega_2}{2} t \right|.$$

Если амплитуды не равны, то для огибающей получается

$$c(t) = \sqrt{1 + \varepsilon^2 + 2\varepsilon \cos(\omega_1 - \omega_2)t}.$$

Это выражение проще всего получить из построения (рис. 17).



Р и с. 17

Только в пределе при очень малых или очень больших ε огибающая биений приближается к синусоиде. При $\varepsilon \ll 1$

$$\begin{aligned} c(t) &= \sqrt{1 + \varepsilon^2 + 2\varepsilon \cos(\omega_1 - \omega_2)t} \approx \sqrt{1 + 2\varepsilon \cos(\omega_1 - \omega_2)t} \approx \\ &\approx 1 + \varepsilon \cos(\omega_1 - \omega_2)t, \end{aligned}$$

и, следовательно, при таких условиях можно применять и линейное детектирование. Кривая биений при этом не отличается от кривой синусоидально-модулированного колебания при малой глубине модуляции. Совершенно так же обстоит дело и при $\varepsilon \gg 1$.

§ 10. Спектр суммы периодических функций

В § 4 уже говорилось о том, что преобразование Фурье линейно, и к нему применим поэтому принцип наложения. В случае периодических функций это можно записать так:

$$C_k = \frac{1}{T} \int_{-T/2}^{+T/2} dt \sum_i f_i(t) e^{-j2\pi kt/T} = \sum_i C_{ik},$$

т. е. комплексная амплитуда k -й гармоники спектра суммы функций равна сумме k -х гармоник спектров каждой отдельно взятой

функции. Это бесспорно, но нас интересуют обычно действительные амплитуды. Для них можно записать

$$c_k = 2 |C_k| = 2 \left| \sum_i C_{ik} \right|.$$

С геометрической точки зрения эта величина представляет собой замыкающую ломаной, стороны которой равны c_{ik} и отложены под соответствующими углами φ_{ik} . Положим, что даны два синусоидальных колебания с комплексными амплитудами

$$2C_1 = c_1 e^{-j\varphi_1}, \quad 2C_2 = c_2 e^{-j\varphi_2}.$$

Тогда

$$\begin{aligned} 2C &= 2(C_1 + C_2) = c_1 e^{-j\varphi_1} + c_2 e^{-j\varphi_2}, \\ c &= |c_1 e^{-j\varphi_1} + c_2 e^{-j\varphi_2}| = |c_1 \cos \varphi_1 + c_2 \cos \varphi_2 - \\ &\quad - j(c_1 \sin \varphi_1 + c_2 \sin \varphi_2)| = \\ &= \sqrt{(c_1 \cos \varphi_1 + c_2 \cos \varphi_2)^2 + (c_1 \sin \varphi_1 + c_2 \sin \varphi_2)^2} = \\ &= \sqrt{c_1^2 + c_2^2 + 2c_1 c_2 \cos(\varphi_1 - \varphi_2)}. \end{aligned} \quad (10.1)$$

Эта формула может применяться и в том случае, когда вместо постоянных фазовых углов φ мы подставляем как угодно зависящие от времени угловые аргументы θ . Тогда и амплитуда оказывается функцией времени, и выражение (10.1) получает смысл огибающей некоторого сложного колебания. Например, если вместо φ_1 и φ_2 подставить $\omega_1 t$ и $\omega_2 t$, то мы получим выражение для огибающей биений, возникающих при сложении двух синусоидальных колебаний с амплитудами c_1 и c_2 и частотами ω_1 и ω_2 .

Рассмотрим вопрос о спектре функции, получаемой в результате сложения двух одинаковых, но сдвинутых по времени периодических функций. Для некоторой периодической функции $f(t)$

$$C_{1k} = \frac{1}{T} \int_{-T/2}^{+T/2} f(t) e^{-j2\pi kt/T} dt.$$

Для такой же функции, но запаздывающей на время τ , имеем

$$C_{2k} = \frac{1}{T} \int_{-T/2}^{+T/2} f(t - \tau) e^{-j2\pi kt/T} dt$$

или, заменяя $t - \tau$ на t_1 ,

$$C_{2k} = \frac{1}{T} \int_{-T/2}^{+T/2} f(t_1) e^{-j\pi k \frac{t_1 + \tau}{T}} dt_1 = e^{-jk\Omega\tau} C_{1k}.$$

Если теперь сложить функции $f(t)$ и $f(t - \tau)$, то комплексная амплитуда k -й гармоники их суммы будет равна

$$C_k = C_{1k} + C_{2k} = C_{1k}(1 + e^{-jk\Omega\tau}),$$

а действительная амплитуда равна ¹

$$c_k = 2 |C_k| = c_k |1 + e^{-jk\Omega\tau}| = 2c_{1k} \left| \cos \frac{k\Omega\tau}{2} \right|. \quad (10.2)$$

Итак, для того чтобы получить спектр суммы двух одинаковых функций, сдвинутых на время τ (например, сумму сигнала и его отражения), достаточно умножить амплитуду каждой гармоники на $2|\cos(k\Omega\tau/2)|$.

Рассмотрим пример. Пусть дана периодическая последовательность коротких импульсов и пусть $\tau = T/2$. Тогда множитель в формуле (10. 2) принимает вид

$$2 \left| \cos k \frac{\pi}{2} \right| = \begin{cases} 0 & \text{при } k \text{ нечетном,} \\ 2 & \text{при } k \text{ четном.} \end{cases}$$

Таким образом, все нечетные гармоники выпадают. Так оно и должно быть: ведь если $\tau = T/2$, то это значит, что импульсы второй серии попадают в середину промежутков первой серии, т. е. получается вдвое более частое следование импульсов, и, стало быть, основная частота, а с нею и частоты всех гармоник увеличиваются вдвое.

Легко сообразить, что то же самое получится при $\tau = \frac{2n+1}{2} T$, т. е. когда τ равно любому нечетному числу полу-периодов. Если $\tau = \frac{2n+1}{4} T$, то из спектра выпадают вторая, шестая, десятая и т. д. гармоники.

Формула (10. 2) дает значение амплитуды k -й гармоники спектра суммы функций $f(t)$ и $f(t - \tau)$. Если мы составим не сумму, а разность этих двух функций, то, действуя аналогично предыдущему, найдем

$$c_k = c_{1k} |1 - e^{-jk\Omega\tau}| = 2c_{1k} \left| \sin \frac{k\Omega\tau}{2} \right|. \quad (10.3)$$

Предположим теперь, что τ — настолько малая величина, что справедливо приближенное равенство

$$f(t) - f(t - \tau) \doteq \Delta f(t) \approx \tau f'(t) = \dot{f}(t).$$

¹ Эту формулу можно было бы получить из (10. 1), положив $c_1 = c_2 = c_{1k}$, $\varphi_1 = k\Omega\tau$, $\varphi_2 = 0$.

Таким образом, мы выразили разность функций через производную. Найдем спектр \dot{f}

$$\begin{aligned} \dot{C}_k &= \frac{1}{T} \int_{-T/2}^{+T/2} \dot{f}(t) e^{-jk\Omega t} dt = \frac{\tau}{T} \int_{-T/2}^{+T/2} f'(\tau) e^{-jk\Omega \tau} d\tau = \\ &= \tau \left[\frac{f(\tau)}{T} e^{-jk\Omega \tau} \Big|_{-T/2}^{+T/2} + jk\Omega C_k \right]. \end{aligned}$$

Но так как функция $f(t)$ периодическая, то

$$\dot{C}_k = jk\Omega \tau C_k$$

и, следовательно,

$$\dot{c}_k = k\Omega \tau c_k.$$

Это соотношение могло бы быть получено из (10. 3) путем замены синуса его аргументом.

Все приведенные выше соотношения без труда распространяются и на случай почти-периодической функции; в этом случае во все формулы входит Ω_k вместо $k\Omega$.

§ 11. Спектры некоторых импульсов

Рассмотрим сначала спектры некоторых специальных разрывных функций, в частности так называемой единичной функции $\sigma(t)$, определяемой следующим образом:

$$\begin{aligned} 0 & \quad t < 0, \\ \sigma(t) &= 1/2 \quad t = 0, \\ 1 & \quad t > 0. \end{aligned} \quad (11.1)$$

Более общим является определение запаздывающей на время τ единичной функции:

$$\begin{aligned} 0 & \quad t < \tau, \\ \sigma(t - \tau) &= 1/2 \quad t = \tau, \\ 1 & \quad t > \tau. \end{aligned}$$

Далее вводится определение единичного импульса в виде

$$\int_{-\infty}^t \delta(u) du = \sigma(t), \quad (11.2)$$

причем функция $\delta(t)$ равна нулю повсюду, кроме точки $t=0$, где она равна бесконечности. Итак, $\delta(t)$ представляет бесконечно короткий импульс с площадью, равной единице. Функция $\delta(t)$ впервые введена в математический аппарат теоретической физики Дираком; она называется также функцией Дирака, или дельта-функцией.

Дифференцируя (11. 2) по t , получаем формально¹

$$\delta(t) = \frac{d}{dt} \sigma(t). \quad (11. 3)$$

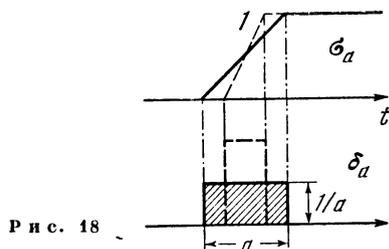
Качественное представление о свойствах функций σ и δ и некоторое обоснование соотношения (11. 3) можно получить путем предельных переходов. Рассмотрим вспомогательную функцию

$$\sigma_a(t) = \begin{cases} 0 & t < -a/2, \\ 1/2 + t/a & -a/2 < t < a/2, \\ 1 & t > a/2, \end{cases}$$

и ее производную

$$\delta_a(t) = \begin{cases} 0 & t < -a/2, \\ 1/a & -a/2 < t < a/2, \\ 0 & t > a/2. \end{cases}$$

Будем переходить к пределу при $a \rightarrow 0$. Заметим, что площадь прямоугольного импульса $\delta_a(t)$ остается равной единице незави-



симо от значения a (рис. 18). В пределе имеем

$$\begin{aligned} \sigma_a(t) &\rightarrow \sigma(t) & \delta_a(t) &\rightarrow \delta(t) \\ a \rightarrow 0 & & a \rightarrow 0 & \end{aligned}$$

Условившись об определениях, перейдем к спектрам. Для спектра функции δ получим

$$S_1 = \int_{-\infty}^{+\infty} e^{-j\omega t} \delta(t - \tau) dt = e^{-j\omega \tau}.$$

Таким образом, модуль спектра функции δ равен единице. Это значит, что функция δ обладает сплошным спектром, простирающимся до бесконечно больших значений частоты с неизменной спектральной плотностью.

¹ Операция дифференцирования, совершаемая здесь над разрывной функцией σ , в новейшее время узаконена в математике в связи с введением так называемых обобщенных функций.

При попытке вычислить спектр функции σ возникает затруднение, связанное с тем, что эта функция не абсолютно интегрируема и имеет конечное значение на бесконечности. Можно обойти это затруднение, умножив функцию σ на $e^{-\alpha t}$ *.

Тогда

$$S_0 = \int_{-\infty}^{+\infty} e^{-(\alpha+j\omega)t} \sigma(t-\tau) dt = \int_{\tau}^{\infty} e^{-(\alpha+j\omega)t} dt = \\ = \frac{1}{\alpha+j\omega} e^{-(\alpha+j\omega)\tau} \xrightarrow{\alpha \rightarrow 0} \frac{1}{j\omega} e^{-j\omega\tau}.$$

Модуль спектра функции σ есть $1/\omega$. К этому же результату можно прийти и другим путем. Воспользуемся известным соотношением [19]

$$\sigma(t) = \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\sin \omega t}{\omega} d\omega.$$

Интеграл в правой части можно рассматривать как вещественную форму интеграла Фурье для нечетной функции $\sigma(t) = 1/2$. В таком случае спектр этой функции (т. е. множитель при синусе) есть $1/\omega$.

Рассмотрим теперь спектр весьма короткого импульса произвольной формы. Импульс характеризуется тем, что функция, представляющая его, равна нулю вне некоторого очень малого интервала τ около $t=0$ **.

* Иногда пишут просто

$$\int_0^{\infty} e^{-j\omega t} dt = \frac{1}{j\omega},$$

т. е. полагают $e^{-j\infty} = 0$, что требует оговорок. Когда мы пишем $f(\infty)$, то понимаем под этим следующее:

$$f(\infty) = \lim_{x \rightarrow \infty} f(x).$$

Функция e^{-jx} при x , стремящемся к бесконечности, не стремится ни к нулю, ни вообще к какому бы то ни было определенному пределу. Поэтому выражение $e^{-j\infty}$ лишено всякого смысла. Функция e^{-jx} (как, впрочем, и e^{+jx}) имеет модуль, равный единице при любых значениях x , аргумент же изменяется циклически с возрастанием x . Другое дело, если показатель есть комплексная величина, т. е. если, например, функция записана в виде

$$f(x) = e^{(-\alpha+j\beta)x}.$$

Тогда при сколь угодно малом, но конечном вещественном и положительном α мы можем записать в указанном выше смысле

$$f(\infty) = 0.$$

** Совершенно необязательно располагать импульс в начале шкалы времен: как мы знаем, спектр не зависит от смещения по времени.

Мы можем записать для спектра такого импульса

$$S = \int_{-\infty}^{+\infty} e^{-j\omega t} f(t) dt = \int_{-\tau/2}^{+\tau/2} e^{-j\omega t} f(t) dt.$$

Но если τ мало, то $e^{\pm j\omega\tau/2}$ мало отличается от единицы, и

$$S \approx \int_{-\tau/2}^{+\tau/2} f(t) dt = q,$$

т. е. спектр равен постоянной величине, определяемой площадью импульса. Такое соотношение мы уже встретили, разбирая вопрос о спектре единичного импульса $\delta(t)$. Однако тогда речь шла о некоторой определенной функции, здесь же в интервале, в котором функция существует, она может иметь произвольный вид, т. е. импульс может иметь какую угодно форму, лишь бы длительность его была малой.

Этот вопрос играет здесь основную роль, и мы на нем немного задержимся. Мы требуем, чтобы τ было мало. Но требование малости лишено смысла до тех пор, пока не указано, с чем та или иная величина сравнивается. В рассматриваемом случае положение определяется тем, что вследствие малости τ функция $e^{\pm j\omega\tau/2}$ мало отличается от единицы. Но это выполняется при условии

$$\omega\tau/2 \ll 1 \quad \text{или} \quad \tau \ll T,$$

где T — период, соответствующий частоте ω ($T = 2\pi/\omega$).

Таким образом, мы приходим к простому, но очень важному как с принципиальной, так и с практической точки зрения выводу: одиночный импульс произвольной формы имеет сплошной спектр, который может быть приближенно выражен постоянной величиной, пропорциональной площади импульса, в пределах того интервала частот, в котором период остается большим по сравнению с длительностью импульса¹.

Мы будем еще неоднократно возвращаться к этому положению.

С повышением частоты, когда T делается уже сравнимым с τ , функция S начинает убывать. Прежде чем изучать общие закономерности, мы рассмотрим ряд примеров.

Вычислим спектр прямоугольного импульса высотой h и длительностью τ . Мы получаем

$$S = h \int_{-\tau/2}^{+\tau/2} e^{-j\omega t} dt = -\frac{h}{j\omega} e^{-j\omega t} \Big|_{-\tau/2}^{+\tau/2} = q \frac{\sin \omega\tau/2}{\omega\tau/2}, \quad (11.4)$$

где площадь импульса $q = h\tau$. График $\Phi = |S|$ для этого случая изображен на рис. 19.

¹ Здесь есть одна тонкость: сформулированное положение справедливо для незнакопеременного импульса. Относящиеся сюда подробности разбираются в Добавлении IV в конце работы.

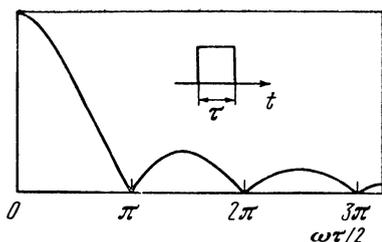


Рис. 19

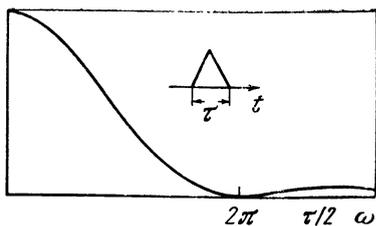


Рис. 20

В качестве иллюстрации предыдущего заметим, что при $\omega\tau/2 \ll 1$

$$\sin \omega \frac{\tau}{2} \approx \omega \frac{\tau}{2} \quad \text{и} \quad S \approx q.$$

Для импульса в форме треугольника с основанием τ и высотой h имеем

$$f(t) = \begin{cases} 0 & t < -\tau/2, \\ h & -\tau/2 < t < 0, \\ h & 0 < t < \tau/2, \\ 0 & \tau/2 < t. \end{cases}$$

Следовательно,

$$S = h \int_{-\tau/2}^0 (1 + 2t/\tau) e^{-j\omega t} dt + h \int_0^{\tau/2} (1 - 2t/\tau) e^{-j\omega t} dt = q \frac{1 - \cos \omega \frac{\tau}{2}}{\frac{1}{2} \left(\omega \frac{\tau}{2} \right)^2}, \quad (11.5)$$

где $q = h\tau/2$ (рис. 20).

Разлагая числитель в степенной ряд, можно убедиться, что отношение стремится к единице при $\omega \tau/2$, стремящемся к нулю. Для косинусоидального импульса, вырезанного из косинусоиды с периодом 2τ и амплитудой h , имеем

$$S = h \int_{-\tau/2}^{+\tau/2} e^{-j\omega t} \cos \pi \frac{t}{\tau} dt = q \frac{\cos \omega \frac{\tau}{2}}{1 - \left(\frac{2}{\pi} \omega \frac{\tau}{2} \right)^2}, \quad (11.6)$$

где $q = 2h\tau/\pi$ (рис. 21).

За последнее время привлек к себе внимание импульс, называемый в специальной литературе «колокольным». Этот импульс выражается функцией $f(t) = e^{-\beta^2 t^2}$.

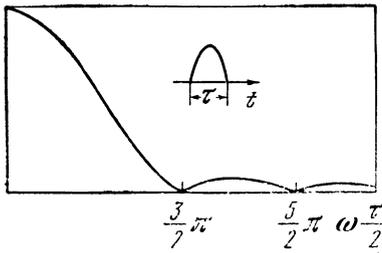


Рис. 21

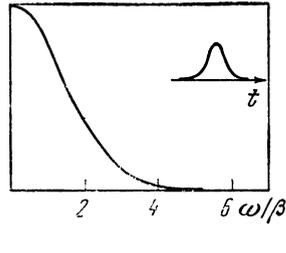


Рис. 22

Функция эта замечательна тем, что она двойственна себе, т. е. что ее преобразование Фурье дает такую же функцию. Действительно, для спектра колокольного импульса получаем

$$S = \int_{-\infty}^{+\infty} e^{-\beta^2 t^2} e^{-j\omega t} dt = e^{-\omega^2/4\beta^2} \int_{-\infty}^{+\infty} e^{-(\beta t + j\omega/2\beta)^2} dt = \frac{2}{\beta} e^{-\omega^2/4\beta^2} \int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{\beta} e^{-\omega^2/4\beta^2}$$

(рис. 22).

Рассмотрим еще несколько импульсов, начинающихся при $t=0$, но стремящихся к нулю лишь при $t \rightarrow \infty$. Пусть, например, дан экспоненциальный импульс $f(t) = e^{-\beta t} \sigma_0(t)$. Его спектр есть

$$S = \int_0^{\infty} e^{(-\beta + j\omega)t} dt = \frac{e^{-(\beta + j\omega)t}}{-\beta + j\omega} \Big|_0^{\infty} = \frac{1}{\beta + j\omega} \quad (11.7)$$

(рис. 23).

Спектр импульса в форме затухающей синусоиды

$$f(t) = e^{-\alpha t} \sin \omega_1 t \sigma_0(t)$$

имеет вид

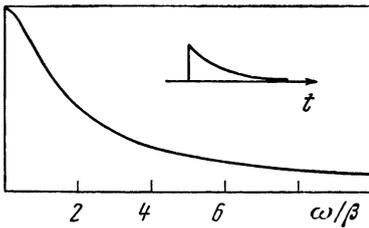
$$S = \int_0^{\infty} e^{-(\alpha + j\omega)t} \sin \omega_1 t dt = \frac{\omega_1}{\alpha^2 - \omega^2 + \omega_1^2 + 2j\alpha\omega}$$

или, вводя обозначения $\alpha^2 + \omega_1^2 = \omega_0^2$, $d = 2\alpha/\omega_0$ (при $d = 0,5$),

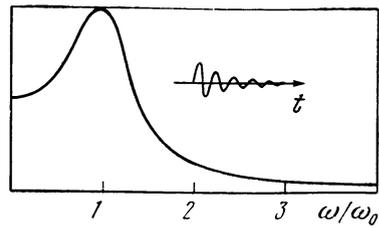
$$S = \frac{\omega_1}{\omega_0^2} \cdot \frac{1}{1 - \left(\frac{\omega}{\omega_0}\right)^2 + jd \frac{\omega}{\omega_0}} \quad (11.8)$$

(рис. 24). Вычислим еще спектр импульса в форме отрезка синусоиды, состоящего из целого числа периодов n

$$S = \int_{-nT/2}^{+nT/2} e^{-j\omega t} \sin \omega_0 t dt = \frac{2j\omega_0}{\omega_0^2 - \omega^2} (-1)^n \sin n\pi \frac{\omega}{\omega_0} \quad (11.9)$$



Р и с. 23



Р и с. 24

(рис. 25). Это есть, в сущности, не что иное, как текущий спектр синусоиды (см. § 5).

Очень часто импульсы определенной формы периодически повторяются. Установим связь между спектрами одиночного импульса и периодической последовательности таких же импульсов. Наперед ясно одно: спектр одиночного импульса есть спектр сплошной, так как импульс есть непериодическая функция. Если же импульс какой угодно формы периодически повторять, то мы получим периодическую функцию, обладающую дискретным гармоническим спектром.

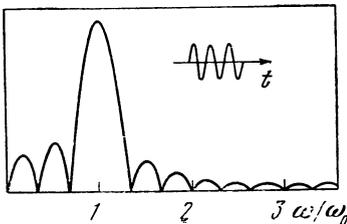
Пусть спектр одиночного импульса есть

$$S_0 = \int_{-\infty}^{+\infty} e^{-j\omega t} f(t) dt. \quad (11.10)$$

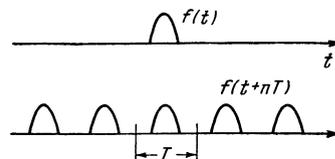
Если такой импульс повторять через промежутки времени T , то получится периодическая функция с периодом T (рис. 26). Спектр этой функции может быть получен по формуле

$$C_k = \frac{1}{T} \int_{-T/2}^{+T/2} e^{-j2\pi kt/T} f(t) dt. \quad (11.11)$$

Сопоставляя (11.10) и (11.11), мы видим, что значения непрерывной функции S_0 совпадают со значениями C_k (с точностью до



Р и с. 25



Р и с. 26

постоянного множителя $1/T$) при определенных значениях аргумента, а именно при

$$\omega = 2\pi k/T = k\omega_1,$$

где $\omega_1 = 2\pi/T$ — круговая частота повторения.

Таким образом, совокупность точек TC_k , определяющих дискретный спектр периодической последовательности импульсов, лежит на кривой S_0 , определяющей спектр одиночного импульса.

На этом примере легко проследить предельный переход от ряда к интегралу Фурье: если период повторения возрастает, т. е. если импульсы повторяются все реже и реже, то точки, изображающие линейчатый спектр, оставаясь на кривой S_0 , располагаются на ней все теснее, пока не образуют непрерывную последовательность, т. е. кривую, совпадающую с S_0 .

§ 12. Связь между длительностью импульса и шириной его спектра

Мы уже установили, что чем короче импульс, тем шире его спектр, в частности, бесконечно короткий импульс имеет бесконечно протяженный спектр с равномерной плотностью. В этом проявляется одно весьма общее и имеющее очень большое значение соотношение, к установлению которого мы подойдем постепенно.

Прежде всего заметим, что общее представление о связи между протяженностями по времени и по частоте вытекает непосредственно из общего свойства преобразования Фурье. Положим, что функция $f(t)$ имеет спектр $S(\omega)$. Изменим масштаб времени в a раз и найдем спектр функции $af(at)$ (множитель a перед функцией добавлен для сохранения площади, т. е. размеры графика функции увеличены по оси ординат во столько же раз, во сколько уменьшены в связи с изменением масштаба размеры по оси абсцисс)

$$S_a = a \int_{-\infty}^{+\infty} f(at) e^{-j\omega t} dt = S(\omega/a).$$

Таким образом, если длительность функции $f(t)$ уменьшена в a раз, то ровно во столько же раз возрастает ширина спектра функции. При этом предполагается, что определения длительности функции и ширины ее спектра остаются неизменными.

Теперь рассмотрим несколько примеров.

Возьмем снова прямоугольный импульс (см. § 11, формулу (11.4) и рис. 18). Для спектра такого импульса мы получили

$$S = q \frac{\sin \omega \frac{\tau}{2}}{\omega \frac{\tau}{2}}. \quad (12.1)$$

Сопоставим теперь длительность импульса и ширину его спектра. Под длительностью импульса Δt естественно в данном случае понимать величину τ . Что касается ширины спектра, то здесь нужно выбрать какое-либо определение, так как спектр импульса безграничен. Но так как спектр убывает с частотой, то можно, например, условиться считать за ширину спектра полосу частот Δf между нулем и тем значением частоты, когда спектр первый раз обращается в нуль. Из формулы (12. 1) видно, что это случится, когда аргумент синуса будет равен π . Отсюда следует равенство

$$2\pi\Delta f\Delta t/2 = \pi \quad \text{или} \quad \Delta f\Delta t = 1,$$

т. е. произведение длительности данного импульса Δt на ширину его спектра Δf (в выбранном нами произвольном определении) равно единице.

Для треугольного импульса мы получили

$$S = q \frac{1 - \cos \omega \frac{\tau}{2}}{\frac{1}{2} \left(\omega \frac{\tau}{2} \right)^2}. \quad (12. 2)$$

Первый нуль спектра будет при $\omega\Delta t/2 = 2\pi$, откуда $\Delta f\Delta t = 2$.

Мы сохраняем определения для Δf и Δt : Δf — полоса частот до первого перехода спектра через нуль, Δt — интервал, вне которого функция, представляющая импульс, равна нулю.

Для косинусоидального импульса

$$S = q \frac{\cos \omega \frac{\tau}{2}}{1 - \left(\frac{2}{\pi} \omega \frac{\tau}{2} \right)^2}, \quad (12. 3)$$

и искомое соотношение принимает вид $\Delta f\Delta t = 1,5$. Для всех рассмотренных примеров получается, что

$$\Delta f\Delta t \approx 1. \quad (12. 4)$$

Однако если бы мы захотели увеличить число примеров, то скоро натолкнулись бы на затруднение при выборе определений Δf и Δt для данного конкретного случая. Положим, например, что мы желаем установить соотношение типа (12. 4) для экспоненциального импульса

$$f(t) = e^{-\beta t} \sigma(t).$$

Предыдущее определение Δt непригодно, так как очевидно, что эта функция не равна нулю в интервале от 0 до ∞ . Спрашивается, как же определить длительность некоторым универсальным образом?

Прежде всего заметим, что вопрос о соотношении между Δf и Δt имеет не только теоретическое, но и очень большое практическое значение. В современной импульсной технике необходимо создавать очень короткие, но в то же время очень мощные им-

пульсы. Иначе говоря, в импульсе большая энергия должна быть сосредоточена в малом интервале времени.

С другой стороны, требуется, чтобы спектр импульса был как можно менее размыт, поскольку широкий спектр вызывает целый ряд серьезных трудностей в устройстве импульсной аппаратуры.

Таким образом, с одной стороны, мы требуем малого Δt , а с другой — малого Δf . Эти требования, как мы видели, вообще говоря, противоречивы. Однако можно искать вид импульса, для которого произведение $\Delta f \Delta t$ имеет наименьшее значение.

Если подходить к вопросу об определении величин Δf и Δt с практической точки зрения, то можно предложить следующее определение длительности; под длительностью импульса понимается промежуток времени, в котором сосредоточена подавляющая часть энергии импульса. Аналитически это определение можно сформулировать так:

$$\int_{t_0 - \Delta t/2}^{t_0 + \Delta t/2} f^2(t) dt = \eta \int_{-\infty}^{+\infty} f^2(t) dt = \epsilon A_t. \quad (12.5)$$

Здесь $A_t = \int_{-\infty}^{+\infty} f^2(t) dt$ — величина, пропорциональная полной энергии импульса; η — выражаемая правильной дробью относительная доля полной энергии импульса, приходящаяся на промежуток времени Δt . Уравнение (12.5) может с удобством решаться при помощи планиметрирования.

Аналогичным образом можно определить и ширину спектра

$$\int_0^{\Delta \omega} \Phi^2(\omega) d\omega = \eta \int_0^{\infty} \Phi^2(\omega) d\omega = \epsilon A_\omega. \quad (12.6)$$

Заметим, что по теореме Рэйли

$$A_\omega = \pi A_t. \quad (12.7)$$

Что касается величины t_0 , входящей в пределы интеграла в левой части (12.5), то в ряде случаев выбор этой величины не оставляет сомнений. Для симметричных импульсов, выражаемых четными функциями, $t_0 = 0$. Для импульсов, начинающихся при $t = 0$, формулу (12.5) следует переписать в виде

$$\int_0^{\Delta t} f^2(t) dt = \eta \int_0^{\infty} f^2(t) dt.$$

Вернемся к ранее рассмотренным импульсам и вычислим их длительности и ширины их спектров, основываясь на предложенном определении. Выбирая для η значение 0,9, получим следующую сводную табл. 2 интересующих нас величин (детали вычислений вынесены в Добавление VI в конце работы).

Мы ограничимся лишь одним замечанием по поводу этой таблицы: как видим, $\Delta f \Delta t$ оказывается наибольшим для импульсов, характеризующихся разрывом функции $f(t)$ (прямоугольный и экспоненциальный); меньшее значение $\Delta f \Delta t$ получается для импульсов с разрывом в первой производной (треугольный и косинусоидальный) и, наконец, наименьшее $\Delta f \Delta t$ оказывается у колокольного импульса, отличающегося тем, что выражающая его функция непрерывна со всеми своими производными.

Таблица 2

Импульс	Рис.	Δt	$\Delta \omega$	Δf	$\Delta f \Delta t$
Прямоугольный	19	$0,90\tau$	$5,1 \frac{1}{\tau}$	$0,81 \frac{1}{\tau}$	0,73
Треугольный	20	$0,541\tau$	$5,3 \frac{1}{\tau}$	$0,84 \frac{1}{\tau}$	0,46
Косинусоидальный	21	$0,596\tau$	$4,57 \frac{1}{\tau}$	$0,73 \frac{1}{\tau}$	0,43
Колокольный	22	$0,825 \frac{1}{\beta}$	$1,64\beta$	$0,26\beta$	0,22
Экспоненциальный	23	$1,155 \frac{1}{\alpha}$	$6,16\alpha$	$0,98\alpha$	1,13

Из всего изложенного можно заключить, что связь между Δf и Δt удовлетворяет в общем случае неравенству

$$\Delta f \Delta t \geq \mu, \quad (12.8)$$

где μ — некоторая постоянная, зависящая, конечно, от выбора определений Δf и Δt .

Определения, которыми мы только что воспользовались, при всей их простоте и практическом удобстве не позволяют, к сожалению, поставить и разрешить вопрос в общем виде. Мы введем новые определения для Δf и Δt , основанные на применении понятия о моментах функции.

Для большей наглядности поясним эти определения ссылкой на общеизвестные понятия и определения технической механики. Предположим, что нам дана произвольная плоская фигура в плоскости XU и требуется некоторым универсальным образом определить размер этой фигуры в направлении оси X (рис. 27). Таким универсальным измерителем может служить радиус инерции данной фигуры относительно главной оси Y_0 , параллельной U .

Напомним определения. Моментом инерции плоской фигуры относительно оси U называется интеграл $I = \int x^2 dA$. Главной

осью называется ось, проходящая через центр тяжести площади. Для определения координаты центра тяжести, т. е. положения главной оси, нужно найти так называемый статистический момент площади $M = \int x dA$, после чего координата центра тяжести определяется по формуле $x_0 = \int x dA / \int dA = M/A$. Главный момент инерции, т. е. момент относительно главной оси Y_0 будет

$$\begin{aligned} I_0 &= \int (x - x_0)^2 dA = \int x^2 dA - 2x_0 \int x dA + x_0^2 \int dA = \\ &= I + x_0^2 A - 2x_0 M = I - x_0^2 A = I - M^2/A. \end{aligned}$$

Если положить $I_0 = r^2 A$, то в этом определении величина r и называется радиусом инерции.

Итак,

$$r^2 = I/A - M^2/A^2, \quad (12.9)$$

и определенная таким образом величина может служить универсальной мерой ширины данной фигуры в направлении оси X .

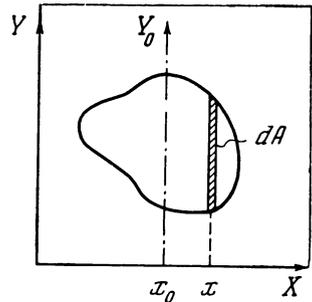
Обратимся теперь к импульсам и их спектрам. Графики как тех, так и других являются плоскими фигурами, к которым в полной мере применимо все сказанное выше. Для импульсов роль оси X выполняет ось времени, для спектров — ось частот. Элемент площади фигуры выразится произведением значения функции на дифференциал независимой переменной. Однако ввиду того что нам встречаются и знакопеременные функции, удобнее оперировать не самой функцией, а ее квадратом.

Таким образом, можно ввести следующие обозначения и определения:

$$\begin{aligned} A_t &= \int_{-\infty}^{+\infty} f^2(t) dt, & A_\omega &= \int_0^{\infty} \Phi^2(\omega) d\omega, \\ M_t &= \int_{-\infty}^{+\infty} t f^2(t) dt, & M_\omega &= \int_0^{\infty} \omega \Phi^2(\omega) d\omega, \\ I_t &= \int_{-\infty}^{+\infty} t^2 f^2(t) dt, & I_\omega &= \int_0^{\infty} \omega^2 \Phi^2(\omega) d\omega. \end{aligned}$$

Далее, согласно (12.9)

$$\Delta t^2 = I_t/A_t - M_t^2/A_t^2; \quad \Delta \omega^2 = I_\omega/A_\omega - M_\omega^2/A_\omega^2.$$



Р и с. 27

Отсюда

$$(\Delta\omega\Delta t)^2 = (I_\omega/A_\omega - M_\omega^2/A_\omega^2)(I_t/A_t - M_t^2/A_t^2)$$

или, используя (12.7) и вводя $\Delta f = \Delta\omega/2\pi$,

$$(\Delta f\Delta t)^2 = \frac{1}{4\pi^2 A_t^2} (I_\omega A_\omega - M_\omega^2)(I_t A_t - M_t^2). \quad (12.10)$$

Теперь ограничим и упростим задачу, предположив, что речь идет о четных функциях времени. Тогда

$$\Phi(\omega) = a(\omega) = \int_{-\infty}^{+\infty} f(t) \cos \omega t dt$$

и, кроме того, $t_0 = 0$, $M_t = 0$. Условимся, далее, что функция $f(t)$ нормирована, т. е. что

$$A_t = \int_{-\infty}^{+\infty} f^2(t) dt = 1$$

и соответственно

$$A_\omega = \int_{-\infty}^{+\infty} \Phi^2(\omega) d\omega = \pi.$$

Займемся теперь интегралом

$$I_t = \int_{-\infty}^{+\infty} t^2 f^2(t) dt.$$

Мы имеем

$$a(\omega) = \int_{-\infty}^{+\infty} f(t) \cos \omega t dt.$$

Продифференцируем обе части по ω

$$a'(\omega) = - \int_{-\infty}^{+\infty} t f(t) \sin \omega t dt = -b_1(\omega),$$

где $b_1(\omega)$ — спектр нечетной функции $f_1(t) = t f(t)$. Воспользовавшись теоремой Рэяли, получим

$$I_t = \int_{-\infty}^{+\infty} t^2 f^2(t) dt = \frac{1}{\pi} \int_0^{\infty} a'^2(\omega) d\omega = \frac{1}{\pi} N.$$

Введя все эти соотношения в (12.10), находим (индекс ω здесь и далее опущен)

$$(\Delta f\Delta t)^2 = \frac{1}{4\pi^5} N (\pi I - M^2). \quad (12.11)$$

Поставим себе теперь задачу найти наименьшее значение $\Delta f \Delta t$, т. е. подобрать такую функцию, которая дает минимальное значение интегральному выражению

$$K = N (\pi I - M^2). \quad (12.12)$$

Воспользуемся для этой цели, следуя Майеру и Леонтовичу [14], вариационными методами.

Составим первую вариацию величины (12.12)

$$\delta K = \delta N (\pi I - M^2) + N (\pi \delta I - 2M \delta M).$$

Для нахождения минимума мы должны приравнять эту вариацию нулю. Но при этом интегралы I , M и N примут вполне определенные постоянные значения, которые мы обозначим соответственно через i , m и n . После этого мы можем объединить варьируемые интегралы и записать

$$\delta \int_0^{\infty} [(\pi i - m)(a')^2 + (\pi n \omega^2 - 2mn\omega)a^2] d\omega = \delta \int_0^{\infty} F(\omega, a, a') d\omega = 0.$$

Теперь мы можем записать уравнение Эйлера

$$\partial E / \partial a - \partial / \partial \omega (\partial E / \partial a') = 0$$

в следующем виде

$$(\pi i - m^2) a'' - (\pi n \omega^2 - 2mn\omega) a = 0. \quad (12.13)$$

Умножим это уравнение на a и проинтегрируем от нуля до бесконечности. Интеграл в первом члене возьмем по частям

$$\int_0^{\infty} a a' d\omega = a a' \Big|_0^{\infty} - \int_0^{\infty} (a')^2 d\omega = -n$$

(так как $a(\infty) = a'(0) = 0$). Интегрируя второй член, заметим, что

$$\int_0^{\infty} \omega^2 a^2 d\omega = i, \quad \int_0^{\infty} \omega a^2 d\omega = m,$$

и таким образом мы получим соотношение $2\pi i = 3m^2$. Замечая, что, согласно (12.12) $K_{\min} = n(\pi i - m^2)$, мы можем теперь записать

$$K_{\min} = \frac{\pi}{3} n i = \frac{\pi}{3} \int_0^{\infty} (a')^2 d\omega \int_0^{\infty} \omega^2 a^2 d\omega.$$

Для оценки этого выражения обратимся к неравенству Буяковского

$$\int_a^b f_1^2(x) dx \int_a^b f_2^2(x) dx \geq \left[\int_a^b f_1(x) f_2(x) dx \right]^2.$$

В нашем случае

$$\int_0^{\infty} (a')^2 d\omega \int_0^{\infty} \omega^2 a^2 d\omega \geq \left[\int_0^{\infty} \omega a a' d\omega \right]^2.$$

Вычисляя интеграл в правой части неравенства по частям, найдем, что он равен $-\pi/2$. Таким образом, $K_{\min} \geq \pi^3/12$, $(\Delta f \Delta t)^2 \geq 1/48 \pi^2$, откуда, наконец, $\Delta f \Delta t \geq 1/4 \sqrt{3} \pi = 0,046$. Мы получили, таким образом, оценку $\Delta f \Delta t$ для наивыгоднейшего случая. Интересно сопоставить полученную величину со значением $\Delta f \Delta t$ для колокольного импульса, для которого, как мы видели, из всех рассмотренных нами импульсов произведение $\Delta f \Delta t$ получается наименьшим. Нельзя, понятно, воспользоваться ранее вычисленными данными, так как сейчас мы пользуемся совершенно другими критериями. Мы должны вычислить интегралы, входящие в формулу (12. 11).

Колокольный импульс и его спектр были определены выше (§ 11). $f(t) = e^{-\beta^2 t^2}$, $a(\omega) = \sqrt{\pi} e^{-\omega^2/4\beta^2}/\beta$.

Из условия нормировки

$$A_f = \int_{-\infty}^{+\infty} f^2(t) dt = 1$$

определяем $\beta = \sqrt{\pi/2}$. После этого находим

$$\begin{aligned} I &= \int_0^{\infty} \omega^2 a^2 d\omega = 2 \int_0^{\infty} e^{-\omega^2/\pi} \omega^2 d\omega = \pi^2/2, \\ N &= \int_0^{\infty} a'^2 d\omega = \frac{2}{\pi^2} \int_0^{\infty} e^{-\omega^2/\pi} \omega^2 d\omega = \frac{1}{2}, \\ M &= \int_0^{\infty} \omega a^2 d\omega = 2 \int_0^{\infty} e^{-\omega^2/\pi} \omega d\omega = -\pi. \end{aligned}$$

Подставляя эти значения в формулу (12. 12), получим

$$\Delta f \Delta t = \frac{1}{4\pi} \sqrt{\frac{\pi-2}{\pi}} = 0,048.$$

Эта величина, как видим, очень близка к теоретическому минимуму.

§ 13. Связь между спектрами и характеристиками линейной системы

Пусть дана некоторая линейная система, описываемая обыкновенным дифференциальным уравнением n -го порядка с постоянными коэффициентами

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = x(t). \quad (13. 1)$$

Применим к обеим частям уравнения преобразование Фурье

$$\int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt.$$

В правой части получится спектр функции $x(t)$, при интегрировании левой части применим формулу (4.3).

Тогда

$$[a_n(j\omega)^n + a_{n-1}(j\omega)^{n-1} + \dots + a_1j\omega + a_0] S_y = S_x \quad (13.2)$$

или, вводя сокращенное обозначение для многочлена в скобках,

$$Z(\omega) S_y = S_x,$$

здесь S_x и S_y — соответственно спектры функций $x(t)$ и $y(t)$. Таким образом,

$$S_y = S_x/Z, \quad (13.3)$$

и мы можем вычислить y по формуле

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{S_x}{Z} e^{j\omega t} d\omega. \quad (13.4)$$

Эта формула дает решение уравнения (13.1) по методу интеграла Фурье.

Введем еще обозначение

$$K(\omega) = 1/Z(\omega).$$

Эта величина, выражающая отношение комплексных амплитуд на выходе и на входе системы при синусоидальном режиме, называется комплексной частотной характеристикой.

Мы можем переписать (13.3) в виде

$$S_y = K S_x. \quad (13.5)$$

Теперь рассмотрим импульсное возбуждение системы. Пусть на вход подается единичный импульс

$$x(t) = \delta(t).$$

При этом на выходе будет наблюдаться отклик системы на такого рода воздействие, который мы обозначим

$$y(t) = g(t).$$

Функцию $g(t)$ назовем временной характеристикой системы¹.

¹ Возможны другие термины: $g(t)$ называется также импульсной реакцией;

$h(t) = \int_0^t g(\tau) d\tau$ называется переходной функцией, или переходной характеристикой.

Для единичного импульса $S=1$, следовательно, в этом случае согласно (12. 5)

$$S_y = K,$$

так что

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} K(\omega) e^{j\omega t} d\omega$$

и

$$K(\omega) = \int_{-\infty}^{+\infty} g(t) e^{-j\omega t} dt, \quad (13. 6)$$

т. е. комплексная частотная характеристика линейной системы есть спектр временной характеристики системы.

Приведем пример. Пусть дана электрическая цепь, составленная из R , L и C в последовательном соединении. Дифференциальное уравнение такой цепи имеет вид

$$L \frac{di}{dt} + Ri + \frac{1}{C} \int idt = e(t).$$

Обозначим

$$U = \frac{1}{C} \int idt,$$

тогда уравнение примет вид

$$\frac{d^2U}{dt^2} + 2\alpha \frac{dU}{dt} + \omega_0^2 U = \omega_0^2 e(t),$$

где, как обычно, $2\alpha = R/L$, $\omega_0^2 = 1/LC$.

В данном случае

$$K(\omega) = \frac{\omega_0^2}{\omega_0^2 - \omega^2 + 2\alpha j\omega}. \quad (13. 7)$$

Временную характеристику получим как решение уравнения

$$\frac{d^2g}{dt^2} + 2\alpha \frac{dg}{dt} + \omega_0^2 g = \omega_0^2 \delta(t)$$

или в операционной форме

$$g = pK(p),$$

откуда

$$g(t) = \frac{\omega_0^2}{\omega_1} e^{-\alpha t} \sin \omega_1 t, \quad (13. 8)$$

где $\omega_1^2 = \omega_0^2 - \alpha^2$.

Мы видим, что, действительно, формула (13. 7) дает спектр функции (13. 8) (ср. с формулой (11. 8)).

Из (13. 6) вытекает возможность экспериментального получения частотной характеристики данной системы путем анализа. Действительно, вместо того чтобы снимать частотную характеристику при синусоидальном возбуждении исследуемой системы, можно возбудить систему коротким импульсом и проанализировать получаемую при таком возбуждении функцию времени на выходе системы. Римский-Корсаков и Шумова [17] использовали эту возможность для акустических измерений, они возбуждали объект исследования периодически повторяемыми короткими импульсами и получали искомую частотную характеристику в форме линейчатого спектра при помощи автоматического анализатора.

Мы начали наше рассуждение со случая системы с сосредоточенными постоянными, описываемой уравнением в полных производных. Однако соотношение (13. 3) (или (13. 5)) справедливо и в том случае, когда система описывается уравнением в частных производных. При этом обычно Z и K — трансцендентные функции ω .

В заключение этого параграфа заметим еще, что в силу (13. 6) к характеристикам линейной системы применимо соотношение $\Delta f \Delta t \geq \mu$. Оно читается в данном применении следующим образом: произведение длительности Δt временной характеристики на ширину Δf частотной характеристики для любой линейной системы больше некоторой постоянной. При этом подразумевается, что для Δf и Δt выбраны соответствующие определения (см. § 12).

§ 14. Функции с ограниченным спектром

В технике связи при передаче различных сигналов мы имеем обычно дело с функциями времени, спектр которых ограничен, т. е. в спектре которых не содержатся частоты выше некоторой граничной. Такие функции обладают замечательным свойством, установленным впервые в 1933 г. Котельниковым [10] и выраженным им в теореме, играющей фундаментальную роль в теории связи и, в частности, в импульсной связи.

Свойство это состоит в том, что тогда как в общем случае функция времени вполне определяется бессчетным множеством своих значений (т. е. бесконечным числом значений на протяжении конечного интервала), функция с ограниченным спектром вполне определяется счетным множеством своих значений (т. е. конечным числом значений на протяжении конечного интервала). С геометрической точки зрения это означает, что если задать на протяжении конечного интервала вполне определенное количество точек, изображающих мгновенные значения функции с ограниченным спектром, то непрерывная кривая, представляющая график функции, может быть проведена через эти точки единственным образом. Это положение объясняется тем, что отсутствие высоких частот в составе функции накладывает серьезное ограничение на способы, которыми могут быть соединены каждые две соседние

точки, и смысл теоремы Котельникова состоит именно в утверждении, что при достаточно частом расположении точек эти ограничения приводят к тому, что кривая определяется этими точками полностью.

Эту мысль можно пояснить еще следующим примером. Положим, что относительно непрерывной линии известно, что она ломаная (в этом состоит в данном примере ограничение свойств линии), и положим, что нам даны точки изломов этой линии. Тогда, зная по свойству ломаной, что точки изломов соединены отрезками прямых, мы легко восстанавливаем всю ломаную линию.

Перейдем к доказательству теоремы, которая в формулировке ее автора гласит: «Любую функцию $f(t)$, состоящую из частот от 0 до f_c , можно передавать с любой точностью при помощи чисел, следующих друг за другом через $1/2f_c$ сек».

Доказательство состоит в разложении функции в особого вида ряд.

В общем случае

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) e^{j\omega t} d\omega, \quad (14.1)$$

где

$$S(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt. \quad (14.2)$$

Но в рассматриваемом частном случае функции с ограниченным спектром мы имеем вместо (14.1)

$$f(t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} S(\omega) e^{j\omega t} d\omega, \quad (14.3)$$

так как $S(\omega) = 0$ при $\omega > \omega_c$. Функция же $S(\omega)$ на конечном промежутке $(-\omega_c, \omega_c)$ может быть разложена в ряд Фурье по частотам следующим образом:

$$S(\omega) = \sum_{-\infty}^{+\infty} D_k e^{j\pi k \omega / \omega_c} \quad (14.4)$$

где $2\omega_c$ играет роль периода по частоте. Коэффициенты D_k ряда (14.4) определяются по обычной формуле

$$D_k = \frac{1}{2\omega_c} \int_{-\omega_c}^{\omega_c} S(\omega) e^{-j\pi k \omega / \omega_c} d\omega. \quad (14.5)$$

Подставляя (14. 4) в (14. 3), получим

$$f(t) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} \left(\sum_{-\infty}^{+\infty} D_k e^{j\pi k\omega/\omega_c} \right) e^{j\omega t} d\omega$$

или, изменяя порядок действий,

$$f(t) = \frac{1}{2\pi} \sum_{-\infty}^{+\infty} D_k \int_{-\omega_c}^{\omega_c} e^{j\omega(t+k\pi/\omega_c)} d\omega,$$

откуда после интегрирования находим

$$f(t) = \frac{1}{\pi} \sum_{-\infty}^{+\infty} D_k \frac{\sin \omega_c (t + k\pi/\omega_c)}{t + k\pi/\omega_c}. \quad (14. 6)$$

Теперь сравним (14. 5) и (14. 3). Как видим,

$$D_k = \frac{\pi}{\omega_c} f\left(-k \frac{\pi}{\omega_c}\right) = \Delta t f(-k\Delta t).$$

Подставляя в (14. 6) найденное значение D_k и меняя знаки под знаком суммы (так как суммирование производится по всем k от $-\infty$ до $+\infty$), получаем окончательно

$$f(t) = \sum_{-\infty}^{+\infty} f(k\Delta t) \frac{\sin \omega_c (t - k\Delta t)}{\omega_c (t - k\Delta t)}. \quad (14. 7)$$

Итак, функция с ограниченным спектром может быть представлена рядом (14. 7), коэффициенты которого являются отсчетами значений функции, взятыми через

$$\Delta t = \pi/\omega_c = 1/2f_c.$$

Этим теорема доказана.

В качестве чисел, определяющих функцию с ограниченным спектром, в предшествующем доказательстве фигурировали отсчеты мгновенных значений функции. Но это могут быть и другие независимые числа, как покажет нижеследующее рассуждение.

Всякая (с несущественными ограничениями) функция может быть представлена на произвольном конечном промежутке $(-\tau/2, \tau/2)$ тригонометрическим рядом

$$f(t) = \sum_{-\infty}^{+\infty} C_k e^{j2\pi kt/T}. \quad (14. 8)$$

Но если спектр функции $f(t)$ ограничен, т. е. если в ее составе нет частот выше граничной частоты ω_c , то и разложение (14. 8) ограничивается конечным числом слагаемых, и мы имеем в этом случае

$$f(t) = \sum_{-\pi}^{\pi} C_k e^{j2\pi kt/T}, \quad (14. 9)$$

где номер наивысшей гармоники n определяется из соотношения $2\pi n/\tau = \omega_c$ или $n = f_c \tau$.

Таким образом, функция полностью определяется на интервале τ посредством n комплексных амплитуд C_k и постоянной составляющей C_0 . Но так как каждая из амплитуд C_k в свою очередь определяется двумя числами (или, что то же, каждое слагаемое многочлена (14. 9) определяется амплитудой и фазой), то всего для полного определения функции на интервале τ нужно

$$m = 2n + 1 = 2f_c \tau + 1 \quad (14. 10)$$

чисел. Теперь нужно устремить τ к бесконечности, чтобы определить функцию $f(t)$ на протяжении всей оси времени. При этом, конечно, и n будет стремиться к бесконечности, так что вместо (14. 10) будем иметь асимптотическое равенство $m \sim 2n = 2f_c \tau$. Если передавать числа равномерно, то в единицу времени нужно передать $m/\tau = 2f_c$ чисел, а временной интервал между двумя числами будет

$$\Delta t = \tau/m = 1/2f_c.$$

Таким образом показано, что числами, определяющими функцию с ограниченным спектром, могут быть не только отсчеты мгновенных значений функции, но и коэффициенты ее разложения в ряд Фурье.

Мы имеем на практике дело с процессами, имеющими начало и конец, т. е. с функциями ограниченной длительности T . Ограниченные во времени функции не могут иметь ограниченный спектр (т. е. спектральную плотность, равную нулю вне конечного интервала) — эти условия противоречивы. Однако можно построить рассуждение на более общей основе, определив ширину спектра F как интервал частот, вне которого спектральная плотность меньше некоторой заданной величины. Оказывается, что общий смысл теоремы Котельникова, состоящий в том, что функция определяется на интервале T посредством $n = 2FT$ «координат», при такого рода обобщении сохраняется ¹.

§ 15. Интеграл Фурье и дискретные спектры

Спектр функции $f(t)$ вычисляется по формуле

$$S = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt. \quad (15. 1)$$

Это вычисление выполнимо при соблюдении обычных ограничений, накладываемых на функцию $f(t)$. Одно из этих ограничений состоит, как известно, в том, что функция $f(t)$ должна быть

¹ Мы не приводим здесь этого более широкого рассуждения, так как оно имеет теоретико-информационный характер и выходит за рамки нашего изложения. См. по этому поводу [9].

абсолютно интегрируема в бесконечных пределах. Очевидно, что это условие не выполняется в случае, когда $f(t)$ есть периодическая функция. Поэтому формула (15. 1) непосредственно неприменима для вычисления дискретных (линейчатых) спектров периодических функций.

Можно, однако, построить некоторое обобщение, хотя и не строгое математически, но во многих отношениях интересное и полезное.

Выясним прежде всего, что представляет собою с математической точки зрения дискретный спектр. Будем рассматривать для простоты чисто синусоидальное колебание с частотой ω_0 . Спектр такого колебания представляется одной-единственной спектральной линией с частотой $\omega = \omega_0$. Спектральная плотность на этой частоте равна бесконечности; при всех же остальных значениях частоты она равна нулю. Вместе с тем из определения спектральной плотности (см. § 3).

$$S = \pi dC/d\omega,$$

где C — амплитуда, следует, что интеграл от спектральной плотности по частоте, взятой на любом интервале (ω_1, ω_2) , включающем в себя $\omega = \omega_0$, будет равен

$$\frac{1}{\pi} \int_{\omega_1}^{\omega_2} S d\omega = \int_{\omega_1}^{\omega_2} \frac{dC}{d\omega} d\omega = C,$$

т. е. «площадь» контура, очерченного графиком спектральной плотности на спектральной диаграмме, есть постоянная конечная величина, равная амплитуде анализируемого синусоидального колебания.

Но описанные свойства функции, представляющей спектральную плотность, в точности совпадают со свойствами единичного импульса δ (см. § 11). Поэтому спектральная плотность синусоидального колебания с частотой ω_0 может быть представлена в виде

$$S = \pi C \delta(\omega - \omega_0). \quad (15. 2)$$

Это представление применимо и к постоянной составляющей, для которой надо положить $\omega_0 = 0$.

Таким образом, спектральная плотность любого периодического процесса может быть представлена формулой

$$S = \pi \sum_{k=0}^{\infty} C_k \delta(\omega - k\omega_1),$$

где C_k — амплитуда k -й гармоники; ω_1 — основная частота.

Обратимся теперь к преобразованиям Фурье. Если применить

преобразование (15. 1) к единичному импульсу, рассматриваемому как функция времени, то для спектра его получим

$$S = \int_{-\infty}^{+\infty} \delta(t) e^{-j\omega t} dt = 1. \quad (15. 3)$$

Подставляя этот результат в формулу обратного преобразования

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) e^{j\omega t} d\omega, \quad (15. 4)$$

мы должны были бы получить

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{j\omega t} d\omega = \delta(t). \quad (15. 5)$$

Это соотношение, справедливость которого совершенно неочевидна, может быть, однако, принято в качестве условного интегрального определения единичного импульса $\delta(t)$.

Если мы согласимся принять это определение, то, видоизменяя его, сразу же можем записать еще ряд соотношений

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{j\omega(t-\tau)} d\omega = \delta(t - \tau), \quad (15. 6)$$

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \bar{e}^{j\omega t} dt = \delta(\omega), \quad (15. 7)$$

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{j(\omega-\omega_0)t} dt = \delta(\omega - \omega_0). \quad (15. 8)$$

Кроме того, знаки показателей можно менять на обратные на том основании, что δ — четная функция.

Ниже следующее рассуждение послужит если не для доказательства справедливости, то для пояснения смысла соотношения (15. 5). Рассуждение это основано на том, что мы берем некоторую непрерывную функцию, к которой преобразование Фурье применимо, и делаем от этой функции предельный переход к единичному импульсу. Множество функций годится для этой цели. Мы выберем колокольный импульс, т. е.

$$f(t) = e^{-\beta^2 t^2} \quad (15. 9)$$

и запишем

$$S(\omega) = \int_{-\infty}^{+\infty} e^{-\beta^2 t^2} e^{-j\omega t} dt = \frac{\sqrt{\pi}}{\beta} e^{-\omega^2/4\beta^2}. \quad (15. 10)$$

Будем теперь стремиться β к нулю. При этом функция $e^{-\beta a/2}$ будет приближаться к единице, а функция $S/2\pi = e^{-\omega^2/4\beta}/2\sqrt{\pi\beta}$ — к единичному импульсу. В самом деле, с уменьшением β эта функция стремится к нулю для всех значений $\omega \neq 0$; при $\omega = 0$ она стремится к бесконечности. Интеграл же от этой функции сохраняет постоянное значение, равное единице

$$\frac{1}{2\sqrt{\pi\beta}} \int_{-\infty}^{+\infty} e^{-\frac{\omega^2}{4\beta}} d\omega = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-x^2} dx = 1.$$

Таким образом, все свойства единичного импульса налицо, и показано, что при $\beta \rightarrow 0$ выражение (15. 10) переходит в пределе в (15. 7), которое эквивалентно (15. 5).

В заключение отметим, что гармоническое колебание и единичный импульс составляют пару сопряженных функций, соотношения между которыми указаны ниже (здесь еще раз проявляется уже отмечавшаяся двойственность частотно-временных представлений):

$$\begin{array}{lll} f(t) & e^{j\omega_0 t} & \delta(t - \tau) \\ S(\omega) & 2\pi\delta(\omega - \omega_0) & e^{-j\omega\tau} \end{array}$$

Г л а в а II

АНАЛИЗ

§ 16. Постановка вопроса

Приступая к изложению вопросов анализа, постараемся прежде всего определить предмет исследования. Мы будем понимать под анализом получение спектра. Однако необходимо сразу же подчеркнуть, что речь пойдет не об анализе функций; а об анализе физических процессов. Остановимся на этом подробнее.

Если задача состоит в анализе функции, т. е. в нахождении спектра функции, то задача эта будет решаться по-разному в зависимости от того, как нами задана функция. Она может быть задана тройким способом: либо своим аналитическим выражением, либо графиком, либо таблицей.

В первом случае спектр вычисляется аналитически; в двух других случаях применяются либо графо-аналитические численные методы вычисления спектра, либо специальные приборы — анализаторы (механические, оптико-механические или оптико-электрические). Все эти методы хорошо разработаны и детально описаны; читатель найдет достаточное количество сведений по этим вопросам в энциклопедического характера книге Серебренникова [18].

Вся эта область нас совершенно не будет интересовать. Мы будем заниматься вопросами физического анализа. Этот вид анализа характеризуется тем, что спектр некоторого процесса получается во время течения процесса в результате его воздействия на определенный физический прибор, называемый анализатором. Таким образом, проблема анализа ставится как чисто физическая проблема, и, как мы увидим, ее разрешение связано с целым рядом чисто физических особенностей.

Как уже говорилось в § 3, может потребоваться анализ самых различных физических явлений; нас могут интересовать спектры механических величин — сил, скоростей, ускорений, смещений, моментов и т. д.; электрических величин — токов, напряжений, зарядов, индукций и т. д.; тепловых, акустических и многих других величин. Было бы крайне неудобно строить анализаторы для каждого рода анализируемой величины. В этом в наше время нет и необходимости. Дело в том, что современная тенденция в области техники измерений состоит в том, что все виды измерений сводятся по возможности к электрическим измерениям. Эта тенденция оправдана, во-первых, наличием громадного ассортимента первоклассных по точности и чрезвычайно чувствительных электроизмерительных приборов, а во-вторых, специфической гибкостью электрических измерений. Не вдаваясь в технические подробности, отметим лишь возможность отнесения измерительного прибора на любое расстояние от объекта измерения, а также то, что электрические измерения позволяют выполнять измерения быстро изменяющихся величин. Для этого служат электромеханические осциллографы, а для особо быстрых процессов — электронные.

Необычайно распространенные в современной технике электрические методы измерения неэлектрических величин основаны на применении приборов, преобразующих измеряемую величину в ту или иную электрическую величину. Эти приборы называются обычно датчиками. Всякое устройство для электрического измерения неэлектрической величины состоит в основном из двух элементов: датчика и подходящего электроизмерительного прибора. Так, например, измерение высоких температур производится электрическим пирометром, состоящим из термопары и милливольтметра; измерение звукового давления — комбинацией высококачественного микрофона и лампового вольтметра с соответствующим усилением. В этих двух примерах термопара и микрофон являются преобразователями и выполняют функции датчиков. Существует громадное количество самых разнообразных датчиков, позволяющих преобразовать какую угодно подлежащую измерению величину в электрический ток или напряжение с выполнением наиболее высоких метрологических требований.

При таком положении вещей, очевидно, что и анализ любого физического процесса, состоящего в изменении во времени той или иной физической величины, может быть сведен к анализу

электрического процесса, т. е. соответствующим образом изменяющегося тока или напряжения. Поэтому все современные технические анализаторы — это электрические приборы.

Мы определили анализ как операцию нахождения спектра. Спектр представляется совокупностью амплитуд составляющих различной частоты. Следовательно, анализатор есть прибор, позволяющий измерить амплитуду и частоту каждого из синусоидальных колебаний, входящего в состав сложного анализируемого колебания.

Всякий анализатор есть измерительный прибор. Поэтому мы должны уделить достаточное внимание метрологическим характеристикам анализатора и в первую очередь — его точности. Как мы увидим, погрешности анализатора обусловлены в значительной мере очень специальными обстоятельствами, лежащими в природе процесса анализа. В дальнейшем эти обстоятельства обсуждены со всей необходимой подробностью.

§ 17. Спектральные приборы

Для целей анализа может служить любой прибор, поведение которого так или иначе зависит от частоты воздействия на него. Такого рода приборы можно назвать спектральными. В основе действия спектральных приборов лежит одно из следующих явлений: интерференция, преломление при наличии дисперсии, резонанс ¹.

Первые два явления используются для построения целого ряда спектральных приборов в оптике. Интерференционным путем получают оптические спектры в интерференционных спектроколах, а также в дифракционной решетке. Спектр может быть получен также при помощи призмы, если показатель преломления заметно зависит от частоты, или, другими словами, если имеет место дисперсия фазовой скорости ².

Мы не будем останавливаться на теории этих приборов; нас интересует только общая их характеристика. Наиболее существенно, что названные спектральные приборы являются анализаторами, но анализаторами волновыми; они специально приспособлены для анализа волн, падающих на анализатор. Поэтому они как раз весьма удобны в оптике.

¹ Кроме того, для целей анализа может быть использована селективность, возникающая вследствие ортогональности тригонометрических функций. Рассмотрению этой возможности посвящен специальный § 20.

² Не следует думать, что получение спектра на основе дисперсии составляет монополию оптики. Электрический кабель также обладает дисперсией; колебания более высокой частоты распространяются по нему с большей фазовой скоростью. В результате этого на достаточно длинных (например, трансатлантических) кабелях можно наблюдать спектральное разложение сигнала во времени, прекрасно известное инженерам дальней связи. Недавно сделана попытка использовать это явление для анализа.

Резонатор — прибор, использующий для анализа явление резонанса, — более универсален, так как может применяться как для анализа волн, для чего его следует поместить в волновое поле, так и для анализа сосредоточенных воздействий. В оптике анализ посредством резонатора нельзя осуществить только потому, что мы еще не в состоянии построить электрический резонатор на частоты порядка 10^{14} *гц*, с которыми мы имеем дело в оптике¹. Для радиочастот резонатор уже вполне осуществим, не говоря уже о более низких ультразвуковых и звуковых частотах. А так как анализ различных явлений (кроме оптических) сводится, как уже говорилось, к анализу электрического тока главным образом в цепях с сосредоточенными постоянными, то возможные решения задачи анализа размежевываются в настоящее время так: в оптике применяются исключительно волновые анализаторы, т. е. интерференционные спектроскопы, дифракционные решетки и призматические спектроскопы; для всех остальных явлений, происходящих с частотами от радиочастот и ниже и допускающих преобразование анализируемой величины в электрическую (ток или напряжение), анализ осуществляется при помощи резонаторов.

Простейшим электрическим резонатором является колебательный контур, состоящий из сосредоточенных индуктивности, емкости и активного сопротивления. Однако эта простейшая форма в ряде случаев оказывается непригодной.

Во-первых, затухание обычного контура довольно велико, а вследствие этого разрешающая способность анализатора мала (об этом подробнее говорится в следующих параграфах). Обычный способ преодоления этого затруднения состоит в замене электрического резонатора механическим, обладающим, как правило, значительно меньшим затуханием. Схема анализатора с применением механического резонатора усложняется: механический резонатор включается между двумя преобразователями. Первый преобразователь превращает ток в механическую силу и возбуждает резонатор. Второй преобразователь воспринимает механическое колебание резонатора и снова преобразовывает это механическое колебание в электрическое. Очень удобно совместить преобразователи и собственно резонатор в одном элементе; так возникают пьезоэлектрические и магнитострикционные резонаторы. Кварцевая пластинка пьезоэлектрического резонатора является одновременно и преобразователем и колебательной системой. То же относится и к магнитострикционному стержню.

Во-вторых, зачастую практически невозможно построить резонатор на заданную частоту в виде системы с сосредоточенными постоянными. В этих случаях используют один из резонансов системы с распределенными постоянными. Но здесь интересно

¹ Можно, однако, заметить, что существуют естественные резонаторы на оптические частоты — атомы и молекулы. Действием этих резонаторов обусловлены такие, например, явления, как резонансное поглощение, резонансная флуоресценция и т. п.

отметить, что резонанс в такой системе — например, в отрезке линии или в стержне — есть волновое явление, и здесь мы снова возвращаемся к принципам, используемым в оптике. Ведь волновой резонанс в отрезке линии представляет собой не что иное, как результат интерференции; отрезок линии как спектральный прибор решительно ничем не отличается по существу происходящих в нем явлений от интерференционного спектроскопа. Таким образом, намеченная выше граница со спектральными приборами оптики не так уж резка. Для полноты картины следует отметить, что дифракционная решетка применялась для анализа явлений ультразвуковой частоты. Для этой цели изучаемое явление превращалось в ультразвуковое излучение в воздухе (эта операция необходима, так как решетка есть волновой анализатор). Волна попадала на соответствующих размеров дифракционную решетку. Дифракционный спектр наблюдался при помощи ультразвукового микрофона. Едва ли такая сложная схема целесообразна. В дальнейшем мы уделим основное внимание анализу посредством резонаторов.

§ 18. Одновременный и последовательный анализ

Анализ сложного колебания посредством резонаторов может быть осуществлен двумя различными способами. Первый способ состоит в применении набора резонаторов, настроенных на различные частоты и подвергающихся одновременно воздействию исследуемого колебания. Такой способ будем называть одновременным анализом. По второму способу применяется один-единственный резонатор с переменной настройкой. Этот вид анализа будем называть последовательным.

Очевидно, что одновременный анализ имеет перед последовательным преимущество в скорости выполнения анализа. Но, кроме того, оба метода имеют и более существенное различие. Дело в том, что процесс перестройки резонатора при последовательном анализе совершается во времени, и наперед ясно, — мы будем еще заниматься этим вопросом подробно, — что перестройка не может происходить очень быстро, так как в противном случае результаты анализа будут искажены переходными явлениями. Следовательно, последовательный способ анализа пригоден для периодических явлений или явлений, характер которых изменяется медленно (т. е. мало изменяется за время анализа). Для анализа же быстро изменяющихся свой ход явлений, в частности, для анализа одиночных импульсов, последовательный анализ совершенно непригоден. Правда, очень часто мы располагаем возможностью периодически повторять импульс в целях исследования. В этом случае последовательный способ пригоден для анализа импульсов; к этой возможности постоянно прибегают.

Сущность последовательного анализа состоит в том, что частота резонатора, плавно изменяясь, совпадает поочередно с ча-

стотами гармонических составляющих анализируемого колебания. Положение органа настройки позволяет отсчитать частоту, амплитуду гармоники показывает подключенный к резонатору индикатор. Примером подобного анализатора может служить самый обыкновенный волномер. Градуировка волномера при помощи мультивибратора есть в сущности операция анализа. Правда, анализ здесь качественный, так как амплитудами гармоник при градуировке не интересуются, но очевидно, что они могут быть измерены. При помощи того же волномера можно с успехом проанализировать, например, модулированное колебание. Однако такой простейший анализатор широкого применения иметь не может по следующим причинам.

От современного анализатора, будь то анализатор технических, звуковых, ультразвуковых или радиочастот, требуется обычно очень широкий диапазон частот. Например, звуковой диапазон охватывает около девяти октав. Построить достаточно селективный резонатор с плавным изменением настройки в таком диапазоне практически невозможно. Поэтому прибегают к видоизменению способа последовательного анализа, имеющему в технике анализа большое значение. Видоизменение это состоит в том, что вместо того чтобы передвигать резонансную частоту по шкале частот относительно неподвижных спектральных линий, заставляют передвигаться по шкале частот весь спектр относительно фиксированной резонансной частоты. И в том и в другом случае наблюдается последовательное совпадение отдельных спектральных линий с частотой резонатора вследствие относительного перемещения их на шкале частот.

Для получения смещающегося по шкале частот спектра необходимо выполнить соответствующее преобразование исходного спектра. Требуемое преобразование получается очень просто: достаточно умножить анализируемое колебание на синусоидальное напряжение переменной частоты. Такое перемножение осуществляется посредством того или иного модулятора.

Пусть анализируемое колебание есть

$$x = \sum c_k \cos (k\omega_1 + \varphi_k),$$

а вспомогательная частота (по немецкой терминологии «Suchton» — буквально «ищущий тон») есть

$$y = y_m \sin \Omega t,$$

где Ω — изменяемая по нашему произволу частота. Составим произведение

$$\begin{aligned} xy &= y_m \sin \Omega t \sum c_k \cos (k\omega_1 t + \varphi_k) = \\ &= \frac{1}{2} y_m \sum c_k \{ \cos [(\Omega - k\omega_1) t - \varphi_k] + \cos [(\Omega + k\omega_1) t + \varphi_k] \}. \end{aligned}$$

Мы получаем модуляционный спектр, воспроизводящий спектр анализируемого колебания в виде двух боковых полос, распо-

женных симметрично по отношению к несущей частоте Ω . Но так как эта частота может изменяться, то весь модуляционный спектр перемещается по шкале частот, следуя за Ω . Достаточно теперь выбрать пределы изменения вспомогательной частоты Ω так, чтобы через частоту резонатора прошла полностью одна из боковых полос преобразованного спектра. Такой вариант последовательного анализа получил очень большое распространение, вполне оправданное его значительными преимуществами. В технические детали мы здесь входить не будем.

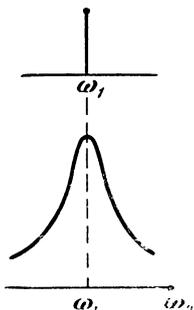
§ 19. Статическая разрешающая способность и погрешность анализатора

Разрешающая способность — важнейшая метрологическая характеристика анализатора. Под разрешающей способностью анализатора понимается вообще способность его разрешить (или разделить) две соседние спектральные линии. Количественной мерой разрешающей способности является наименьший интервал по частоте между двумя спектральными линиями, при котором они еще разделяются анализатором. Нужно определенным образом условиться, при каких обстоятельствах мы будем считать линии разделенными. Для этого нужно рассмотреть процесс анализа подробнее.

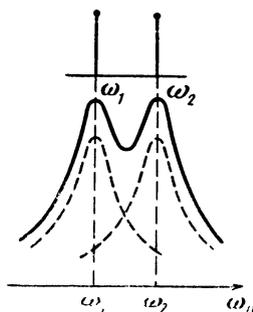
Начнем с последовательного анализа. Предположим, что имеется одна-единственная спектральная линия с частотой ω_1 . Пусть резонатор постепенно перестраивается, так что его резонансная частота, возрастая, проходит через значение ω_1 . При этом индикатор дает наибольший отсчет. Если же записать показания индикатора в функции резонансной частоты, то эта зависимость даст не что иное, как резонансную кривую (рис. 28).

Пусть теперь анализируемый спектр состоит из двух спектральных линий одинаковой интенсивности. В этом случае в резонаторе будут существовать одновременно колебания двух частот, и возникнут биения с разностной частотой. «Амплитуда» сложного колебания при этом не будет уже постоянной величиной. Однако можно условиться понимать под амплитудой ее наибольшее значение, равное сумме амплитуд обоих составляющих колебаний. Каждая из этих амплитуд зависит от настройки резонатора по отношению к соответствующей частоте, и, следовательно, повторяя предыдущее рассуждение применительно к каждой из двух спектральных линий, мы приходим к выводу, что показания индикатора будут представлены графически двугорбой кривой, получаемой при сложении двух смещенных резонансных кривых (рис. 29). Мы видим, что резонатор с затуханием изображает отдельную спектральную линию резонансной кривой. (Вообще невозможно построить анализатор, который давал бы спектр в виде линии.) Спектр же, состоящий из многих линий, будет изображен анализатором в виде плавной кривой с рядом максимумов, отвечающих

отдельным спектральным линиям. Задача анализа тем не менее разрешается, так как положение максимумов на шкале частот определяет положение спектральных линий, а высота максимумов определяет интенсивность линий. Нам остается определить условия разделения.



Р и с. 28



Р и с. 29

Для двух резонансных кривых рис. 29 можно записать

$$y_1 = \frac{1}{2} \frac{1}{\sqrt{\left(\frac{\omega - \omega_1}{\omega_1}\right)^2 + \frac{d^2}{4}}}, \quad y_2 = \frac{1}{2} \frac{1}{\sqrt{\left(\frac{\omega - \omega_2}{\omega_2}\right)^2 + \frac{d^2}{4}}},$$

где d — затухание резонатора. Введем $\omega_0 = (\omega_1 + \omega_2)/2$, $b = (\omega_2 - \omega_1)/2$. Тогда

$$\omega - \omega_1 = \omega - \omega_0 + b = \Delta\omega + b, \quad \omega - \omega_2 = \omega - \omega_0 - b = \Delta\omega - b,$$

и уравнение двугорбой кривой есть

$$y = y_1 + y_2 \approx \frac{1}{2} \left[\frac{1}{\sqrt{\left(\frac{\Delta\omega + b}{\omega_0}\right)^2 + \frac{d^2}{4}}} + \frac{1}{\sqrt{\left(\frac{\Delta\omega - b}{\omega_0}\right)^2 + \frac{d^2}{4}}} \right].$$

При небольшом затухании максимумы кривой будут расположены на частотах ω_1 и ω_2 , т. е. при $\Delta\omega = \pm b$; величина максимума приближенно равна

$$y_{\max} \approx \frac{1}{d} + \frac{1}{2} \cdot \frac{1}{\sqrt{(2b/\omega_0)^2 + d^2/4}} \approx \frac{1}{d}.$$

Отношение ординаты седла к величине максимума равно

$$\frac{y_0}{y_{\max}} = \frac{1}{\sqrt{(b/\omega_0)^2 + d^2/4}} = \frac{2}{\sqrt{(2b/\omega_0 d)^2 + 1}}.$$

Мы видим, что отношение это зависит как от относительного расстояния между линиями b/ω_0 , так и от затухания d , причем увеличение расстояния равносильно уменьшению затухания, и обратно.

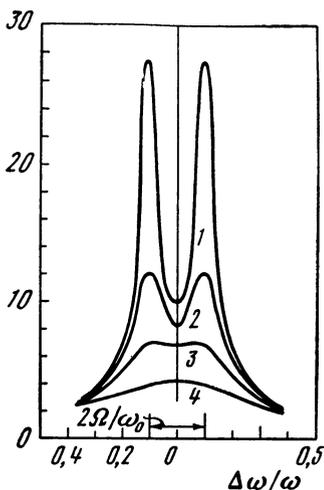


Рис. 30

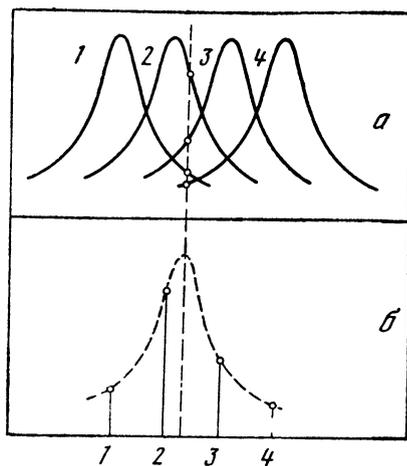


Рис. 31

Условимся считать, что две спектральные линии одинаковой интенсивности разделены, если отношение y_0/y_{\max} не превышает некоторой определенной величины, например $1/2$.

Тогда из неравенства

$$\frac{2}{\sqrt{1 + (2b/\omega_0 d)^2}} \leq \frac{1}{2}$$

найдем

$$2b/\omega_0 d \geq \sqrt{15} \approx 4 \quad \text{или} \quad d \leq \frac{1}{4} \cdot \frac{2b}{\omega_0}.$$

Итак, в рассмотренном случае разрешающая способность зависит только от затухания резонатора: чем затухание меньше, тем разрешающая способность больше; задавшись разрешающей способностью, можно определить наибольшее допустимое затухание.

Выведенные соотношения иллюстрируются графиком на рис. 30 ($1-d=0,04$; $2-d=0,10$; $3-d=0,20$; $4-d=0,40$; $\Omega/\omega_0=0,1$), на котором двугорбая кривая построена для различных значений затухания. Как видим, при определенном затухании седло пропадает, и кривая делается одногорбой. При этом разделение линий, очевидно, уже совершенно невозможно. Величину затухания, соответствующую такому изменению формы кривой, можно определить, положив $y_0/y_{\max}=1$; это дает

$$d = \frac{1}{\sqrt{3}} \cdot \frac{2b}{\omega_0}.$$

Обратимся теперь к одновременному анализу. Анализатор состоит из набора резонаторов, настроенных на различные ча-

стоты. Все резонаторы одновременно подвергаются действию анализируемого колебания и так или иначе реагируют на него. Пусть анализатор состоит из резонаторов 1, 2, 3, 4, . . . , резонансные кривые которых изображены на рис. 31, а. Положим далее, что анализируемое колебание состоит из одной-единственной спектральной линии, положение которой отмечено вертикальным штрихом. Каждый из резонаторов откликнется на это колебание по-своему; на рис. 31, а точками отмечены ординаты каждой из резонансных кривых, дающие величину показания индикатора, подключенного к данному резонатору. Совокупность этих показаний, которую мы в дальнейшем будем кратко называть показанием анализатора, представлена на рис. 31, б.

Итак, вместо единственной линии, имеющейся в действительности, показание анализатора представляется целым спектром, частоты которого вдобавок зависят от настройки резонаторов, а не от частоты анализируемого колебания. Однако заключение о полной несостоятельности одновременного анализа было бы слишком поспешным. Дело в том, — и это весьма существенно, — что показание анализатора неоднозначно, т. е. между спектром анализируемого колебания и показанием анализатора существует вполне определенная связь. Рассмотрим этот вопрос несколько подробнее.

Пусть анализатор составлен из некоторого числа резонаторов, различающихся по настройке, так что уравнение резонансной кривой k -го резонатора записывается в виде

$$y_k = \frac{1}{2} \cdot \frac{1}{\sqrt{\left(\frac{\omega - \omega_k}{\omega_k}\right)^2 + \frac{d^2}{4}}},$$

где ω_k — резонансная частота. Под воздействием синусоидального колебания с частотой ω_0 показание анализатора представится совокупностью величин

$$y_{k0} = \frac{1}{2} \cdot \frac{1}{\sqrt{\left(\frac{\omega_0 - \omega_k}{\omega_k}\right)^2 + \frac{d^2}{4}}}.$$

Но в силу тех же самых приближений, при помощи которых выводится выражение для резонансной кривой¹, последнее соотно-

¹ Точное выражение

$$y = \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_k^2}\right)^2 + \frac{\omega^2}{\omega_k^2} d^2}}$$

для случая малой расстройки, т. е. для селективного резонатора, упрощается, если положить

$$\omega \approx \omega_k, \quad 1 - \frac{\omega^2}{\omega_k^2} = \frac{(\omega + \omega_k)(\omega - \omega_k)}{\omega_k^2} \approx 2 \frac{\omega - \omega_k}{\omega_k},$$

что и приводит к приближенной формуле, которой мы пользуемся,

шение можно переписать в виде

$$y_{k0} = \frac{1}{2} \cdot \frac{1}{\sqrt{\left(\frac{\omega_k - \omega_0}{\omega_0}\right)^2 + \frac{d^2}{4}}},$$

а это выражение означает, что показание анализатора представляется совокупностью линий с частотами ω_k , вписанных в такую же резонансную кривую, но с резонансом при ω_0 . Последняя резонансная кривая на рис. 31, б намечена штрихом.

Из этого следует, что если мы составим анализатор из большого числа резонаторов с часто и равномерно расположенными по шкале частот резонансными частотами, то показание такого анализатора даст столь же частый спектр, так что определить по нему положение максимума станет уже возможным. При таком положении можно говорить о разрешающей способности анализатора при одновременном анализе совершенно так же, как и применительно к последовательному анализу. Можно также рассматривать кривую, изображающую показание последовательного анализатора, как предельный случай показания одновременного анализатора при неограниченном увеличении числа резонаторов. Выражая это же соотношение еще иначе, можно сказать, что показания при последовательном и одновременном анализе соотносятся так же, как спектры одиночного и периодически повторяемого импульса: первый спектр — сплошной, второй — линейчатый, вписывающийся в первый.

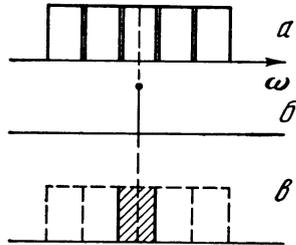
Возможность хорошего разрешения при помощи одновременного анализа можно иллюстрировать ссылкой на общеизвестный технический язычковый частотомер. Этот прибор состоит из набора механических резонаторов, обладающих большой селективностью. Если настройка выполнена, как обычно, с шагом 1 *гц*, то частота может быть уверенно отсчитана с точностью до 0,5 *гц*, так как, когда частота попадает в середину интервала между частотами двух соседних язычков, они колеблются одновременно с равными амплитудами.

Возможна принципиально иная постановка вопроса об анализе, особенно уместная в применении к одновременному анализу. Представим себе, что анализатор составлен из набора не обычных резонаторов, а идеальных полосовых фильтров с тесно примыкающими границами (рис. 32, а). Тогда при наличии одной спектральной линии (рис. 32, б) показание анализатора будет однозначным: на колебание, отвечающее этой линии, откликнется только тот фильтр, в пределах полосы пропускания которого лежит возбуждающая частота. Показание анализатора изображено условно на рис. 32, в. Однако ясно, что значение частоты не определяется точно; мы можем лишь утверждать, что она находится в пределах той или иной полосы.

Так же ясно, что мерой разрешающей способности такого полосового анализатора является ширина полосы. Если в пределах

одной и той же полосы оказывается более чем одна линия исследуемого спектра, то анализатор их не разделяет; он дает лишь некоторое суммарное значение для всех составляющих, частоты которых лежат в пределах той или иной полосы. При таких обстоятельствах естественно заставить полосовой анализатор измерять мощность, так как мощность в полосе выразится простой арифметической суммой мощностей отдельных составляющих.

В заключение настоящего параграфа рассмотрим вопрос о погрешности анализатора. Если исследуемое колебание синусоидально, то амплитуда и частота колебаний могут быть в принципе



Р и с. 32

определены при помощи анализатора точно. Если же анализируется сложное колебание, то возникает неустранимая погрешность. Погрешность эта обусловлена тем, что на резонатор воздействуют одновременно все составляющие сложного колебания, и мы не в состоянии разделить их влияние, а следовательно, и определить амплитуду и частоту каждой составляющей в отдельности.

Пусть комплексная амплитуда k -го резонатора под действием i -й гармоники спектра есть

$$y_{ki} = \frac{x_i}{1 - (\omega_i/\omega_k)^2 + jd_k\omega_i/\omega_k},$$

где ω_i — частота гармоники; x_i — ее амплитуда; ω_k — резонансная частота данного резонатора. Под действием всего спектра амплитуда резонатора будет

$$y_k = \sum_i \frac{x_i}{1 - (\omega_i/\omega_k)^2 + jd_k \frac{\omega_i}{\omega_k}}.$$

Предположим, что резонатор настроен точно на частоту одной из гармоник. Выделив слагаемое, относящееся к этой гармонике, получим

$$y_k = \frac{x_k}{jd_k} + \sum_{i \neq k} \frac{1}{1 - \left(\frac{\omega_i}{\omega_k}\right)^2 + jd_k \frac{\omega_i}{\omega_k}}.$$

В остающейся здесь сумме суммирование производится по всем остальным гармоникам, т. е. по всем i , не равным k . Положим теперь, что спектр гармоничен, т. е. что

$$\omega_i = \Omega i,$$

где Ω — основная частота. Кроме того, примем, что относительное затухание для всех резонаторов одинаково, т. е. положим $d_k = d$. Тогда

$$y_k = \frac{x_k}{jd} + \sum_{i \neq k} \frac{x_i}{1 - (i\Omega/\omega_k)^2 + jd i\Omega/\omega_k}.$$

Последнее выражение показывает, что отсчет по резонатору пропорционален амплитуде соответствующей гармоники, т. е.

$$|y_k| = \frac{x_k}{d}$$

только при условии, что суммой в правой части можно пренебречь. Таким образом, эта сумма и представляет собой погрешность, совершаемую при анализе.

Рассматривая формулу, мы видим, что погрешность тем меньше, чем реже спектр, т. е. чем больше основная частота Ω . Кроме того, относительная погрешность убывает с уменьшением затухания (т. е. с увеличением разрешающей способности), потому что при этом возрастает первый член в правой части по сравнению с суммой.

Все эти соображения применимы в равной мере как к одновременному, так и к последовательному анализу.

§ 20. Об анализе без резонаторов

До сих пор предполагалось, что анализ выполняется при помощи резонатора, т. е. некоторого устройства, обладающего определенной частотной избирательностью. Простейшим резонатором является колебательный контур; более совершенным избирательным органом может служить полосовой фильтр.

Существует, однако, и другая возможность, которую мы рассмотрим в общих чертах. Составим схему, показанную в скелетной форме на рис. 33. На этой схеме Π — перемножитель (идеальный модулятор), т. е. устройство, на выходе которого получается напряжение, пропорциональное произведению напряжений, поданных на два входа, Γ — генератор синусоидального напряжения с изменяемой частотой и начальной фазой, И — интегрирующая схема. Если генератор дает напряжение

$$U = \cos \omega t,$$

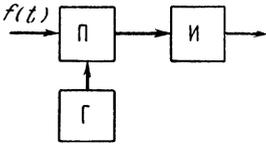
то на выходе мы получим

$$A_t(\omega) = \int_0^t f(\tau) \cos \omega \tau d\tau$$

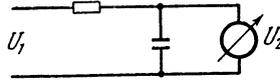
(предполагая, что схема включена в момент $t=0$). Если изменить начальную фазу напряжения генератора на $\pi/2$, то на выходе будем иметь

$$B_t(\omega) = \int_0^t f(\tau) \sin \omega \tau d\tau.$$

Величины A_t и B_t представляют собой косинусную и синусную составляющие текущего спектра. Соединив вместе две схемы рис. 33 с двумя генераторами, дающими сдвинутые на $\pi/2$ синусоидальные напряжения, и добавив звено, в котором A_t и B_t возводятся в квадрат и складываются с последующим извлечением



Р и с. 33



Р и с. 34

корня, получим анализатор, на выходе которого получается текущий амплитудный спектр функции $f(t)$

$$\Phi_t(\omega) = \sqrt{A_t^2(\omega) + B_t^2(\omega)}.$$

Можно видоизменить описанный метод получения спектра так, чтобы можно было обойтись одной схемой рис. 33. Для этого нужно при данной частоте ω изменять фазу φ напряжения генератора. Тогда на выходе схемы рис. 33 будем иметь

$$x(t, \omega, \varphi) = \int_0^t f(\tau) \sin(\omega\tau + \varphi) d\tau.$$

Найдем наибольшее значение x , считая фазу φ переменной. Для этого приравняем нулю производную $\partial x / \partial \varphi$

$$\begin{aligned} \frac{\partial x}{\partial \varphi} &= \int_0^t f(\tau) \cos(\omega\tau + \varphi) d\tau = \cos \varphi \int_0^t f(\tau) \cos \omega\tau d\tau - \\ &- \sin \varphi \int_0^t f(\tau) \sin \omega\tau d\tau = A \cos \varphi - B \sin \varphi = 0. \end{aligned}$$

Отсюда

$$\cos \varphi = B / \sqrt{A^2 + B^2}, \quad \sin \varphi = A / \sqrt{A^2 + B^2}.$$

Подставляя эти значения в выражение для x , получим

$$x_{\max} = A^2 / \sqrt{A^2 + B^2} + B^2 / \sqrt{A^2 + B^2} = \sqrt{A^2 + B^2} = \Phi_t(\omega)$$

Итак, если подбором фазы добиться максимума (при данной частоте) выходного напряжения схемы рис. 33, то получим непосредственно амплитудный спектр (модуль комплексного спектра).

Вышеописанная схема действует в полном соответствии с математическим определением спектра, выполняя все операции, входящие в это определение; короче говоря, схема рис. 33 выполняет преобразование Фурье. Следовательно, эта схема является идеальным анализатором, обладающим бесконечной разрешающей способностью.

Но теперь нужно заметить, что схема обладает свойствами идеального анализатора, лишь поскольку ее составные части идеальным образом выполняют свое назначение; мы предполагали, что перемножитель идеально перемножает, интегратор идеально интегрирует, а генератор дает безукоризненное синусоидальное напряжение. В действительной схеме все эти условия, разумеется, в той или иной мере нарушены.

Мы рассмотрим теперь схему, в которой одно из идеальных звеньев заменено реальным, а именно: вместо идеального интегратора И включим в схему рис. 33 простейшую интегрирующую схему в виде RC -звена (рис. 34). Положим, для простоты, что анализируемое напряжение чисто синусоидально, и что мы рассматриваем установившийся режим анализатора. Для установившегося режима коэффициент передачи интегрирующей схемы рис. 34 равен

$$K = \frac{U_2}{U_1} = \frac{1}{1 + j\omega RC}.$$

Напряжение, подаваемое на вход интегрирующей схемы, равно произведению анализируемого напряжения и напряжения генератора (пусть начальные фазы обоих напряжений совпадают)

$$U_1 = \cos \Omega t \cos \omega t = \frac{1}{2} \cos (\omega + \Omega) t + \frac{1}{2} \cos (\omega - \Omega) t.$$

Для комплексных амплитуд напряжения на выходе можем записать

$$\frac{1}{2} \cdot \frac{1}{1 + j(\omega + \Omega) RC}, \quad \frac{1}{2} \cdot \frac{1}{1 + j(\omega - \Omega) RC}.$$

При ω , мало отличающейся от Ω , первой величиной можно пренебречь, и мы получим

$$|K| \simeq \frac{1}{2} \cdot \frac{1}{\sqrt{1 + (\omega - \Omega)^2 \tau^2}} \quad \text{или} \quad |K| \simeq \frac{1}{2} \frac{1/\Omega \tau}{\sqrt{(1 - \omega/\Omega)^2 + 1/(\Omega \tau)^2}},$$

где $\tau = RC$ — постоянная времени интегрирующей схемы.

Полученное выражение только постоянным множителем $1/\Omega\tau$ отличается от выражения для коэффициента передачи простого колебательного контура

$$|K| = \frac{1}{\sqrt{(1 - \omega^2/\omega_0^2)^2 + \omega^2 d^2/\omega_0^2}} \approx \frac{1}{2} \frac{1}{\sqrt{(1 - \omega/\omega_0)^2 + d^2/4}}.$$

Таким образом, оказывается, что рассматриваемая нами схема анализатора обладает совершенно такими же свойствами в отношении избирательности, как и обычный резонатор, и, следовательно, к ней применимо все то, что говорилось ранее о разрешающей способности и погрешности анализатора с резонаторами.

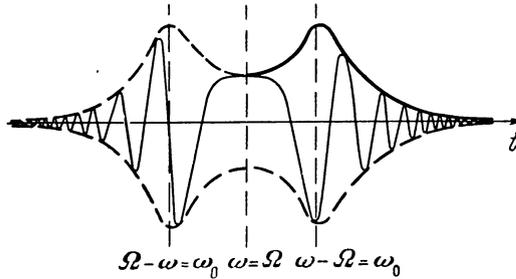
Интересно, что в рассматриваемом анализаторе роль затухания играет величина $1/\Omega\tau$. Так как для увеличения постоянной времени $\tau = RC$ нет практических ограничений, то можно получить анализатор с очень высокой разрешающей способностью. С другой стороны, разрешающую способность можно легко и в широких пределах изменять путем изменения R или C . Используя это, можно, например, исследовать тонкие детали какого-либо участка спектра.

Техническое выполнение описанной системы возможно в самых различных формах. В качестве поучительного примера приведем предложенное в свое время применение в качестве анализатора обычного ваттметра. В частности, если взять электродинамический прибор, то для использования его в качестве анализатора достаточно подать анализируемое напряжение и напряжение от генератора соответственно на две обмотки прибора (статора и ротора). Вращающий момент пропорционален произведению токов в обмотках; таким образом осуществляется перемножение. Интегрирование выполняется в механической части прибора за счет баллистических свойств ротора. Если желательно перемножать непосредственно напряжения (а не пропорциональные им токи), то вместо электродинамического ваттметра можно взять квадратный электрометр. Отсчет в такого рода устройствах берется обычно не по постоянному отклонению (при нулевой разностной частоте), а по амплитуде колебаний подвижной части прибора на низкой частоте ω_0 ее механического резонанса. Опишем происходящие при этом явления.

Если вспомогательная частота ω медленно изменяется, проходя через значение Ω , то на прибор воздействует переменная разностная частота $|\Omega - \omega|$, которая сначала убывает до нуля, а затем снова возрастает. При этом подвижная система прибора колеблется, амплитуда колебания имеет два максимума, когда разностная частота сравнивается с собственной частотой подвижной системы. Процесс, происходящий в измерительном приборе при медленном изменении ω , показан на рис. 35. Амплитуда сначала мала, затем нарастает до максимума, затем убывает до некоторого постоянного отклонения; затем все происходит в обратном порядке. Огибающая процесса есть не что иное, как резонансная кривая

подвижной системы прибора. С точки зрения основного принципа анализатора мы должны были бы взять отсчет при $\omega = \Omega$, когда стрелка остановится. Но практически гораздо удобнее взять отсчет в точках, где подвижная система резонирует.

В заключение нужно дать ответ на естественно возникающий вопрос: существует ли принципиальная разница между анализом при помощи резонаторов и анализом по вышеописанному методу?



Р и с. 35

Разница, конечно, существует и имеет принципиальный характер. Это та же разница, которая имеется между фильтрацией и гетеродинированием. Описанный метод есть в сущности метод преобразования частоты, или, в частности, метод синхронного детектирования по отношению к той или иной составляющей спектра. Принципиальный характер различия состоит в том, что фильтрация осуществляется посредством пассивной линейной системы с постоянными параметрами, тогда как рассмотренный гетеродинный анализатор с физико-математической точки зрения сводится к некоторой линейной системе с переменными параметрами.

§ 21. Работа резонатора

Рассматривая действие анализатора с резонаторами, мы считали до сих пор само собою разумеющимся, что если резонатор обладает частотной избирательностью, то он способен выделить те или иные спектральные составляющие, т. е. произвести анализ. Однако в свете соображений предыдущего параграфа возникает естественный вопрос: действительно ли резонатор способен анализировать, или, иначе, какими свойствами должен обладать резонатор, чтобы наблюдаемая на выходе величина действительно отображала спектр анализируемой функции? Для ответа на этот вопрос мы должны выяснить, при каких условиях величина на выходе резонатора связана со входной величиной посредством преобразования Фурье. Если такая связь существует, то это значит, что резонатор должным образом выполняет свое назначение.

В общем случае отклик резонатора, на вход которого в момент $t=0$ включено произвольное воздействие $f(t)$, может быть выражен интегралом Дюамеля

$$x(t) = \int_0^t f(\tau) g(t-\tau) d\tau, \quad (21.1)$$

где $g(t)$ — временная характеристика (импульсная реакция) резонатора. Предположим, что

$$g(t) = \sin \omega_0 t. \quad (21.2)$$

В этом случае

$$\begin{aligned} x(t) &= \int_0^t f(\tau) \sin \omega_0 (t-\tau) d\tau = \\ &= \sin \omega_0 t \int_0^t f(\tau) \cos \omega_0 \tau d\tau - \cos \omega_0 t \int_0^t f(\tau) \sin \omega_0 \tau d\tau = \\ &= A \sin \omega_0 t - B \cos \omega_0 t. \end{aligned}$$

Здесь A и B — по-прежнему косинусная и синусная составляющие текущего спектра. Они являются функциями времени и могут рассматриваться как медленно меняющиеся составляющие амплитуды колебания $x(t)$. Огибающая (т. е. зависящая от времени амплитуда) этого колебания равна, очевидно, $\sqrt{A^2 + B^2} = \Phi_t(\omega_0)$, т. е. текущему амплитудному спектру. Итак, если для анализа применен резонатор с временной характеристикой (21.2), то анализ производится в соответствии с математическим определением спектра, т. е. мы имеем идеальный анализатор. Но временной характеристикой (21.2) обладает контур без потерь, т. е. цепь, составленная из L и C ($R=0$). Поэтому следующий вопрос — это вопрос о том, что дает в качестве анализатора реальный резонатор. Для контура с потерями

$$g(t) = e^{-\alpha t} \sin \omega_0 t$$

и

$$\begin{aligned} x(t) &= \int_0^t f(\tau) e^{\alpha(\tau-t)} \cos \omega_0 \tau d\tau = \\ &= \sin \omega_0 t \int_0^t f(\tau) e^{\alpha(\tau-t)} \cos \omega_0 \tau d\tau - \cos \omega_0 t \int_0^t f(\tau) e^{\alpha(\tau-t)} \sin \omega_0 \tau d\tau = \\ &= A_\alpha \sin \omega_0 t - B_\alpha \cos \omega_0 t. \end{aligned}$$

В этом выражении A_α и B_α — соответственно косинусная и синусная составляющие текущего спектра, но не функции $f(\tau)$, а взве-

шенной функции

$$f_{\alpha}(\tau) = e^{\alpha(\tau-t)} f(\tau).$$

Появившаяся здесь скользящая весовая функция — это та самая функция, которой пользуется Фано (см. § 6); вышеприведенные соображения обосновывают, таким образом, предложенные Фано определения.

В более общем случае для узкополосного резонатора произвольного вида временная характеристика может быть представлена в виде

$$g(t) = p(t) \sin \omega_0 t + q(t) \cos \omega_0 t = \operatorname{Re} s(t) e^{j\omega_0 t},$$

где ω_0 — центральная частота пропускаемой резонатором полосы. В этом случае

$$\begin{aligned} x(t) &= \int_0^t f(\tau) g(t-\tau) d\tau = \int_0^t f(\tau) \operatorname{Re} s(t-\tau) e^{j\omega_0(t-\tau)} d\tau = \\ &= \operatorname{Re} e^{j\omega_0 t} \int_0^t f(\tau) r(\tau-t) e^{-j\omega_0 \tau} d\tau \end{aligned}$$

и интеграл представляет комплексный текущий спектр с весовой функцией

$$r(-t) = s(t) = p(t) - jq(t).$$

Вывод, вытекающий из всего сказанного, заключается в том, что всякий реальный резонатор дает нам не истинный спектр анализируемой функции, а спектр взвешенной функции, причем функция веса зависит от характеристики резонатора. Мы получаем истинный спектр лишь в пределе, когда временная характеристика резонатора приближается к $g(t) = \sin \omega_0 t$, а весовая функция, следовательно, к единице. На основании этих соображений может быть вычислена погрешность реального анализатора, чем мы, однако, сейчас заниматься не будем (см. § 26).

Из изложенного ясно, что метрологическая характеристика анализатора должна была бы содержать данные о характеристике резонатора или, по меньшей мере, о постоянной времени. В противном случае нельзя установить, что именно измеряет анализатор. К сожалению, в большинстве случаев указывается лишь ширина полосы пропускания резонатора, что недостаточно.

В качестве иллюстрации упомянем об интересном применении взвешенных спектров, рассматриваемых в качестве мгновенных, а именно о так называемой «видимой» речи [33]. Суть дела заключается в том, что мгновенный спектр (взвешенный при помощи полосовых фильтров) изображается непрерывно на движущемся экране. Так как мгновенный спектр есть функция двух переменных — частоты и времени, то спектр изображается на плоскости экрана в прямоугольных координатах; по оси абсцисс отклады-

вается время, по оси ординат — частота. Что же касается самой функции этих двух переменных, т. е. спектральной плотности, то она отображается яркостью свечения экрана в данной точке (а при записи — степенью почернения специальной бумаги, как в фототелеграфе). Речь анализируется в диапазоне примерно до 4 кгц, разделенном на полосы по 300 гц каждая. Получаемая картина дает достаточно полное отображение звуков речи; можно научиться читать «видимую речь» непосредственно с экрана. Первоначально устройство «видимой речи» предназначалось для глухих — чтобы заменить слуховое восприятие зрительным и, действительно, с успехом применялось для этой цели. Однако такого рода техника может быть использована также для исправления дефектов речи, для фонетических исследований, а также для построения автоматов, фиксирующих звуки речи или выполняющих поданные голосом команды и т. п.

§ 22. Действительные условия работы анализатора

Выводы § 19 требуют серьезных коррективов с точки зрения действительных условий работы анализатора, которые мы и постараемся сейчас разобрать.

В самом начале отмечалось, что периодическая функция вида $f(t) = f(t + nT)$ есть математическая абстракция. Но если бы даже существовало явление, отвечающее этому определению, то оно не существовало бы для анализатора. Ведь анализатор в действительности включается на конечное время, и, следовательно, анализируемое явление для анализатора представляется отрезком на протяжении от момента включения до момента выключения.

Анализатор может дать истинный спектр явления в том единственном случае, когда анализируемое явление полностью укладывается в этот отрезок времени, т. е. когда анализируется заведомо непериодическое явление типа импульса, равного нулю вне конечного промежутка времени, и при условии, что этот промежуток времени, т. е. длительность импульса, меньше времени включения анализатора. При анализе же длительных явлений, периодических или непериодических, анализатор принципиально не в состоянии дать истинный спектр этих явлений. По определению спектр выражается интегралом

$$S = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt. \quad (22.1)$$

Если анализатор выполняет требуемое умножение и интегрирование, то он во всяком случае не может произвести интегрирование в бесконечных пределах. Анализатор «не знает», что происходило до его включения, и тем более «не знает», что будет происходить после его выключения. Он может лишь произвести ин-

тегрирование на промежутке $t_1 < t < t_2$, где t_1 и t_2 — соответственно моменты включения и выключения.

В лучшем случае можно ожидать от анализатора, что он, действуя непрерывно, выполняет интегрирование на промежутке от момента включения до текущего момента. Если принять момент включения за начало отсчета времени, то анализатор может измерить величину

$$S_t = \int_0^t f(t) e^{-j\omega t} dt, \quad (22.2)$$

т. е. текущий спектр функции $f(t)$ (см. § 5). Чем больше время включения, тем больше текущий спектр приближается к истинному спектру длительного явления. Если время включения очень мало, то каков бы ни был характер явления, анализатор получает лишь короткий импульс, сплошной спектр которого имеет однородную плотность вплоть до частот, определяемых временем включения, и тем более высоких, чем время включения меньше, т. е. чем импульс короче.

Эти соображения легко проверить экспериментально. Синусоидальное колебание воспринимается слухом как чистый музыкальный тон. Если включать этот тон на все более короткие промежутки времени, то ясное ощущение высоты тона будет постепенно утрачиваться, пока, наконец, достаточно кратковременное включение не будет восприниматься просто как щелчок — это и соответствует короткому импульсу с широким сплошным спектром.

Из всего сказанного следует, что результат анализа принципиально зависит от времени. Учесть эту зависимость можно, опираясь на две различные (но приводящие, конечно, к одним и тем же выводам) точки зрения: временную и спектральную (частотную).

Приведем подробное рассуждение о процессе анализа с обеих точек зрения.

а. Временная точка зрения. Мы исходили в предыдущем при определении разрешающей способности резонатора из его резонансной кривой (частотной характеристики). Но резонансная кривая представляет поведение резонатора в установившемся режиме. Установившимся режимом мы называем режим, достигаемый по окончании переходных явлений, возникающих при включении, т. е. теоретически через бесконечно большое время. Практически переходными явлениями можно пренебречь уже по прошествии конечного времени. Однако время установления тем больше, чем больше избирательность резонатора, т. е. чем меньше его затухание. А в применении к анализу нас интересуют именно резонаторы с высокой избирательностью. Стало быть, исследуя работу анализатора, мы не можем, вообще говоря, игнорировать переходные явления. Составляя и решая соответствующие диффе-

ренциальные уравнения, мы получим амплитуду колебаний резонатора как функцию не только частоты (как в установившемся режиме), но и времени. Следовательно, резонансная кривая резонатора будет различной для различных моментов времени; ее уместно назвать динамической резонансной кривой в противопоставление обычной резонансной кривой, которую мы будем называть статической и которая является предельной формой динамической резонансной кривой при $t \rightarrow \infty$.

Таким образом, все ранее выведенные соотношения сохраняют смысл лишь в качестве предельных. Истинная разрешающая способность анализатора определяется его динамической резонансной кривой. Ее максимум тем тупее, чем меньше время включения; следовательно, анализ выполняется тем точнее, чем больше время анализа.

б. Спектральная точка зрения. Мы исходили в предыдущем при определении разрешающей способности анализатора из представления, что на анализатор воздействует периодическое, в частности, синусоидальное колебание. В действительности этого нет: анализатор включается лишь на конечное время. Следовательно, для каждого данного момента воздействие на резонатор может быть представлено непериодическим явлением, а именно отрезком синусоиды, длящимся от момента включения до текущего момента. Анализатор будет анализировать это явление с присущей ему статической разрешающей способностью. Но так как спектр самого воздействия с течением времени изменяется («текущий» спектр), то будет изменяться и показание анализатора. В частности, при анализе синусоиды в самые первые моменты, когда периодический характер явления еще не определился и текущий спектр практически однороден, анализатор даст почти одинаковые показания на всех частотах. Лишь в пределе, когда сплошной текущий спектр вырождается в дискретную спектральную линию, анализатор даст спектр, определяемый на основе стационарных представлений, изложенных выше. Эти представления сохраняют, таким образом, силу при выводе предельных соотношений. В действительности же показание анализатора оказывается тем более размытым, т. е. разрешающая способность его тем меньше, чем меньше время включения.

Обе точки зрения — временная и спектральная — теоретически равноценны. В дальнейшем использована временная точка зрения и вот по каким методическим соображениям:

1. Временная точка зрения представляется легче усвояемой и менее формальной.

2. Стоя на временной точке зрения, мы можем характеризовать анализатор как таковой независимо от режима его включения и характера анализируемого явления¹.

¹ Речь идет о том, что при анализе периодического явления мы можем учитывать конечность времени включения анализатора двумя способами:

1) характеризовать явления обычным спектром, а анализатор — динами-

3. Некоторые вопросы вообще затруднительно (хотя и возможно в принципе) излагать на спектральном языке. В частности, здесь имеется в виду вопрос о последовательном анализе. Все предыдущее относилось к одновременному анализу, когда речь идет об отклике на то или иное воздействие резонатора с постоянной настройкой. Но если обратиться к анализатору, содержащему резонатор с переменной настройкой, то мы имеем дело с явлениями более сложными, и оказывается более удобным говорить о них на временном языке.

И в случае последовательного анализа очевидно, что свойства резонатора могут описываться статической резонансной кривой лишь при бесконечно медленной перестройке. В действительности анализ происходит с конечной скоростью. Легко понять, что чем быстрее настройка резонатора проходит через частоту данной спектральной составляющей, тем в меньшей степени эта составляющая успеет раскачать резонатор. То же относится и к варианту с применением вспомогательной частоты, когда настройка резонатора неизменна, а спектр исследуемого колебания передвигается с той или иной скоростью по шкале частот.

В случае последовательного анализа свойства анализатора также могут с удобством описываться посредством динамической резонансной кривой. Теория и опыт показывают, что ее максимум тупее, чем у статической резонансной кривой, причем тем тупее, чем больше скорость относительного движения по шкале частот настройки резонатора и исследуемого спектра, т. е. чем больше скорость анализа. Таким образом, и в этом случае разрешающая способность тем меньше, чем больше скорость анализа.

Из всего сказанного следует, что понятие разрешающей способности анализатора лишено практического смысла, если оно не связано так или иначе со скоростью (или продолжительностью) анализа. Очевидно, что для полной характеристики свойств анализатора мы должны ввести понятие динамической разрешающей способности анализатора, зависящей как от параметров анализатора, так и от скорости анализа. Последующие параграфы посвящены подробному исследованию этого вопроса, причем одновременный и последовательный анализы рассмотрены порознь, так как постановка вопроса в этих двух случаях довольно существенно различается.

ческой частотной характеристикой; 2) характеризовать явление текущим спектром, а анализатор — обычной статической частотной характеристикой. Первый способ, не затрагивающий описания анализируемого явления, но зато охватывающий посредством динамической характеристики как свойства анализатора, так и условия его работы, представляется более естественным и удобным.

§ 23. Связь между разрешающей способностью анализатора и временем анализа

При анализе периодических явлений имеет силу очень общее положение, гласящее, что чем больше разрешающая способность анализатора, тем больше необходимое для анализа время.

Это положение легко уяснить себе на примере резонатора. С одной стороны, разрешающая способность резонатора тем больше, чем меньше затухание, т. е. чем острее резонансная кривая резонатора. С другой же стороны, чем меньше затухание резонатора, тем медленнее затухают собственные колебания, возникающие при включении, и тем дольше, следовательно, нужно выждать, пока режим резонатора можно будет считать уже установившимся. Очевидно, что отсчет по резонатору при анализе периодического явления следует брать при установившемся режиме. В противном случае нельзя пользоваться при определении разрешающей способности резонатора его статической резонансной кривой (об этом подробно говорилось в § 22.)

Пример с резонатором показывает, что высказанное вначале положение имеет в виду влияние переходных (неустановившихся) явлений в анализаторе и в приведенной формулировке относится к анализу периодических явлений.

Это положение имеет, как сказано, очень общий характер. Можно привести еще один пример, в котором справедливость этого положения на первый взгляд не очевидна. Речь идет о дифракционной решетке. Разрешающая способность дифракционной решетки как анализатора зависит, как известно, не от периода решетки, а от общего ее размера (от числа штрихов). Но при увеличении размера решетки возрастают пути пробега волн, а следовательно, и соответствующие времена. Понятно, что в оптике это не имеет никакого практического значения, но нас интересует принципиальная сторона вопроса.

Хотя суть дела в общем ясна, мы не можем, конечно, удовлетвориться пояснением положения о связи между разрешающей способностью и временем анализа на отдельных, хотя бы и убедительных, примерах. Задача наша состоит теперь в том, чтобы доказать это положение в общем виде.

Представим себе избирательный элемент анализатора в виде четырехполюсника, на вход которого подается величина x , на выходе получается величина y , причем избирательность проявляется в том, что отношение $K=y/x$, т. е. коэффициент передачи, зависит определенным образом от частоты, если x и y — колебательные величины.

С другой стороны, если возбудить четырехполюсник толчком на входе, то на выходе возникает переходное явление той или иной длительности. Нам нужно найти связь между шириной частотной характеристики (K) и длительностью переходного явления. Первая величина определяет разрешающую способность, вторая —

необходимое для анализа время. Интересующая нас зависимость была уже установлена ранее, в § 12. Именно, было показано, что

$$K(\omega) = \int_{-\infty}^{+\infty} g(t) e^{-j\omega t} dt,$$

т. е. что комплексный коэффициент передачи $K(\omega)$ есть не что иное, как спектр функции $g(t)$, которая представляет собой явление на выходе при подаче на вход возбуждения вида $\delta(t)$ и которую мы назвали временной характеристикой.

Заметим попутно, что нам удобнее пользоваться для определения длительности переходных явлений именно функцией $g(t) = h'(t)$, а не переходной функцией $h(t)$, соответствующей возбуждению вида $\sigma(t)$, так как последняя функция часто может стремиться с возрастанием t к постоянному пределу.

Итак, дело сводится теперь к тому, чтобы установить связь между длительностью Δt функции $g(t)$ и шириной Δf ее спектра $[K(\omega)]$. Эта задача рассмотрена подробно в § 11, в котором установлено, что $\Delta f \Delta t \geq \mu$. Применительно к рассматриваемому вопросу это соотношение удобно записать в виде

$$\Delta f \Delta t = A \quad (23.1)$$

и истолковать следующим образом: для данной системы необходимое время анализа обратно пропорционально разрешающей способности. Это положение, таким образом, доказано. Остается еще указать на существенную деталь — как следует понимать выражение «данная система».

Соотношение (23.1) не представляло бы для нас никакого интереса, если бы оно не допускало никаких изменений свойств системы. Вопрос ставится так: если дана функция

$$f(a, b, c, \dots, t),$$

где a, b, c, \dots — параметры, и если, естественно, спектр этой функции будет зависеть от тех же параметров, т. е.

$$\Phi = \Phi(a, b, c, \dots, \omega),$$

то можно ли утверждать, что постоянная A в выражении (23.1) сохранит свое значение при любых изменениях параметров? В таком общем виде это утверждение было бы неверным. Но заметим, что мы имеем дело, в частности, с такими функциями, в выражения которых параметры входят в качестве сомножителей при независимой переменной. Легко показать, что если умножить аргумент на постоянную величину, то будем иметь

$$f(\alpha t) = f(t_1), \quad \Phi(\omega/\alpha) = \Phi(\omega_1),$$

и, следовательно, в этом случае

$$\Delta f \Delta t = \Delta f_1 \Delta t_1 = A,$$

т. е. от изменения параметров произведение $\Delta f \Delta t$ не меняет своей величины.

Следовательно, под данной системой в вышеприведенной формулировке следует понимать систему, отвечающую данной схеме или, говоря более общим языком, описываемую данного вида дифференциальным уравнением, но без закрепления частных значений коэффициентов этого уравнения.

Так, например, если уменьшить затухание резонатора, то время анализа возрастет во столько же раз, во сколько увеличится разрешающая способность.

Теперь заметим, что к соотношению (23.1) можно подойти и с другой точки зрения. Если A есть величина, постоянная для данной системы, то как выбрать систему для получения наименьшего A ? Ведь у нас нет никаких оснований для того, чтобы считать обычный резонатор наилучшим видом избирательного элемента для целей анализа. Нельзя ли построить систему, обладающую лучшими свойствами, чем обычный резонатор, например, обладающую меньшим временем установления при той же разрешающей способности или большей разрешающей способностью при том же времени установления?

Из сказанного в § 12 по поводу колокольного импульса можно заключить, что этот вид импульса обладает особо выгодными свойствами в смысле малости произведения $\Delta f \Delta t$. Для того чтобы можно было использовать эти свойства для целей анализа, нужно подобрать такую функцию, которая, сохраняя колоколообразный характер, имела бы максимум на любой заданной частоте ω_0 . Легко сообразить, что если спектр обычного колокольного импульса

$$f(t) = e^{-\pi t^2/2}$$

имеет вид

$$\Phi(\omega) = \sqrt{2} e^{-\omega^2/2\pi},$$

то требуемым свойством будет обладать функция

$$\Phi_1(\omega) = \sqrt{2} e^{-(\omega-\omega_0)^2/2\pi}.$$

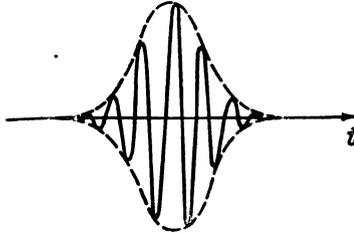
Такому спектру на основании теоремы (4.6) отвечает функция времени

$$f_1(t) = e^{j\omega_0 t} e^{-\pi t^2/2}.$$

Это колебание с колоколообразной огибающей показано на рис. 36.

Подобного рода импульс привлекает к себе внимание в теории связи, где он рассматривается как наивыгоднейший элементарный сигнал. Для нас эта функция имеет другой смысл: она представляет собой временную характеристику некоторой системы, которую мы предполагаем применить в качестве избирательного элемента анализатора. Иначе говоря, эта функция представляет явление на выходе системы при возбуждении на входе вида $\delta(t)$.

Можно высказать некоторые общие соображения о свойствах такой системы. Рассматривая рис. 36, можно заметить, что искомая система обладает тем существенным свойством, что колебания в ней возникают не сразу, в момент толчка, а постепенно нарастают до максимума, а затем уже начинают убывать. Следовательно, первым приближением к искомой системе могла бы послужить комбинация двух слабо связанных резонаторов, из которых первый воспринимает толчок, а затем постепенно передает энергию второму резонатору, на выходе которого можно таким образом наблюдать картину, до известной степени сходную



Р и с. 36

с рис. 36. Можно полагать, что для такой комбинации произведение $\Delta f \Delta t$ окажется меньшим, чем для обычного резонатора. Можно также рассчитывать получить лучшее приближение к требуемым свойствам путем последовательного усложнения системы. И действительно, требуемой характеристикой обладает система из достаточно большого числа слабо связанных контуров, осуществляемая практически в виде многокаскадного резонансного усилителя.

§ 24. Динамическая разрешающая способность резонатора

Пусть резонатор с неизменной настройкой включается при $t=0$ под синусоидальное воздействие с частотой ω . Найдем динамическую характеристику резонатора. Положим, что в качестве резонатора взята электрическая цепь из R , C и L в последовательном включении. Уравнение резонатора имеет вид

$$L \frac{di}{dt} + Ri + \frac{1}{C} \int idt = \sin \omega t \quad (t > 0)$$

или, дифференцируя по t и полагая $L=1$,

$$\frac{d^2i}{dt^2} + 2\alpha \frac{di}{dt} + \omega_0^2 i = \omega \cos \omega t \quad (t > 0). \quad (24.1)$$

Нам нужно не полное решение этого уравнения, а лишь значение амплитуды, зависящей от времени, т. е. выражение для огибающей. Воспользуемся для ее нахождения методом, состоящим в том, что мы ищем решение в виде

$$i = a(t) \sin \omega t + b(t) \cos \omega t, \quad (24.2)$$

где a и b — медленные функции времени, т. е. функции, медленно изменяющиеся по сравнению с $\sin \omega t$ или $\cos \omega t$. Медленность определяется тем, что производные этих функций являются малыми величинами первого порядка по сравнению с производной $\sin \omega t$ или $\cos \omega t$, т. е. по сравнению с ω .

Будем, кроме того, полагать, что и α представляет собой малую величину первого порядка, т. е. что мы имеем дело с цепью высокой избирательности. Дифференцируя (24. 2), подставляя в (24. 1) и приравнявая порознь коэффициенты при $\sin \omega t$ и $\cos \omega t$, получим после отбрасывания членов второго порядка малости (\ddot{a} , \ddot{b} , $\dot{\alpha}a$, $\dot{\alpha}b$) так называемые укороченные уравнения

$$\begin{aligned}(\omega_0^2 - \omega^2) a - 2\omega (\dot{b} + \alpha b) &= 0, \\(\omega_0^2 - \omega^2) b + 2\omega (\dot{a} + \alpha a) &= \omega.\end{aligned}\tag{24. 3}$$

Эти уравнения имеют уже первый порядок, в чем и состоит преимущество метода. Для решения уравнений (24. 3) воспользуемся операционным методом. Запишем операционное изображение этих уравнений

$$\begin{aligned}(\omega_0^2 - \omega^2) \bar{a} - 2\omega (p + \alpha) \bar{b} &= 0, \\(\omega_0^2 - \omega^2) \bar{b} + 2\omega (p + \alpha) \bar{a} &= \omega,\end{aligned}$$

откуда

$$\bar{a} = \frac{1}{2} \cdot \frac{p + \alpha}{(p + \alpha)^2 + \beta^2}, \quad \bar{b} = \frac{1}{2} \cdot \frac{\beta}{(p + \alpha)^2 + \beta^2},$$

где $\beta = (\omega_0^2 - \omega^2)/2\omega$.

Применяя общеизвестные формулы операционного исчисления, находим

$$\begin{aligned}a(t) &= \frac{1}{2} \cdot \frac{\alpha}{\alpha^2 + \beta^2} \left[1 - e^{-\alpha t} \left(\cos \beta t - \frac{\beta}{\alpha} \sin \beta t \right) \right], \\b(t) &= \frac{1}{2} \cdot \frac{\beta}{\alpha^2 + \beta^2} \left[1 - e^{-\alpha t} \left(\cos \beta t + \frac{\alpha}{\beta} \sin \beta t \right) \right].\end{aligned}$$

Нас интересует огибающая, выражаемая через a и b следующим образом:

$$c(t) = \sqrt{a^2(t) + b^2(t)}.$$

Выполняя вычисления, находим

$$c = \frac{1}{2} \sqrt{\frac{(1 - e^{-\alpha t} \cos \beta t)^2 + (e^{-\alpha t} \sin \beta t)^2}{\alpha^2 + \beta^2}},\tag{24. 4}$$

или

$$c = \sqrt{\frac{e^{-\alpha t} (\operatorname{ch} \alpha t - \cos \beta t)}{2(\alpha^2 + \beta^2)}}.$$

Эта формула и представляет собой искомое выражение для динамической характеристики резонатора. Величина c зависит,

как мы видим, от параметров резонатора, т. е. от α и ω_0 , а также от двух независимых переменных ω и t . Легко убедиться, что при $t \rightarrow \infty$ выражение (24. 4) стремится в пределе к

$$c \xrightarrow{t \rightarrow \infty} \frac{1}{2} \cdot \frac{1}{\sqrt{\alpha^2 + \beta^2}} = y. \quad (24. 5)$$

Это и есть выражение для статической характеристики, т. е. обычной резонансной кривой.

Мы получили выражение динамической характеристики резонатора, опираясь на временную точку зрения, т. е. рассматривая устанавливающийся режим резонатора. Тот же вопрос мог бы быть изложен на чисто спектральном языке следующим образом.

Будем решать задачу о поведении резонатора как стационарную, т. е. положим, что приложенное напряжение представлено суммой синусоидальных напряжений. Воспользуемся процедурой, описанной в § 13, т. е. применим к уравнению резонатора

$$L \frac{di}{dt} + Ri + \frac{1}{C} \int idt = E(t)$$

преобразование Фурье. Мы получим

$$(j\omega L + R + 1/j\omega C) S_i = Z S_i = S_E,$$

где S_E и S_i — соответственно спектры напряжения и тока. Далее,

$$i = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_i e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{S_E}{Z} e^{j\omega t} d\omega.$$

Но в рассматриваемом случае спектральная плотность является функцией не только частоты, но и времени. В самом деле, приложенное напряжение $E(t)$ предполагается синусоидальным, но включается на анализатор в момент $t=0$. Следовательно, для каждого данного момента времени t напряжение, воздействующее на анализатор, представляется следующей функцией:

$$E = \begin{cases} 0 & \text{при } u < 0, \\ \sin \Omega u & \text{при } 0 < u < t \end{cases}$$

и спектр этой функции в данный момент, т. е. текущий спектр, выражается формулой

$$S_E = \int_0^t \sin \Omega u e^{-j\omega u} du = \frac{\Omega}{\Omega^2 - \omega^2} \left[1 - e^{-j\omega t} \left(\cos \Omega t + j \frac{\omega}{\Omega} \sin \Omega t \right) \right],$$

где ω — текущая частота спектра; t — текущее время; u — переменная интегрирования. Таким образом, получаем для тока

$$i = \frac{\Omega}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{j\omega t} - \left(\cos \Omega t + j \frac{\omega}{\Omega} \sin \omega t \right)}{(\Omega^2 - \omega^2) \left[R + j \left(\omega L - \frac{1}{\omega C} \right) \right]} d\omega.$$

Подынтегральная функция имеет два простых сопряженных полюса, лежащих над вещественной осью

$$a = \pm \omega_1 + ja.$$

(Точки $a = \pm \Omega$ не являются полюсами, в чем легко убедиться, раскрыв получающуюся неопределенность.) Вычеты подынтегральной функции равны

$$\frac{\omega_1 \pm ja}{2j\omega_1 L} \cdot \frac{e^{j(\pm \omega_1 + ja)t} - \left(\cos \Omega t + j \frac{\pm \omega_1 + ja}{\Omega} \sin \Omega t \right)}{\Omega^2 - (\pm \omega_1 + ja)^2}.$$

Выполнив вычисление, находим (положив $L=1$)

$$i = \frac{\Omega}{(\Omega^2 - \omega_0^2)^2 + 4a^2\Omega^2} \left\{ e^{-at} \left[(\Omega^2 - \omega_0^2) \cos \omega_1 t - \frac{a}{\omega_1} (\Omega^2 + \omega_0^2) \sin \omega_1 t \right] - (\Omega^2 - \omega_0^2) \cos \Omega t + 2a\Omega \sin \Omega t \right\}.$$

Эта формула представляет собой точное решение задачи, совпадающее, разумеется, с решением, полученным любым другим способом. Вычислив огибающую, найдем

$$c = \Omega \sqrt{\frac{1 + \left(\frac{\omega_0}{\omega_1} \right)^2 e^{-2at} - 2 \frac{\omega_0}{\omega_1} e^{-at} \cos (\Omega - \omega_1) t}{(\Omega^2 - \omega_0^2)^2 + 4a^2\Omega^2}}.$$

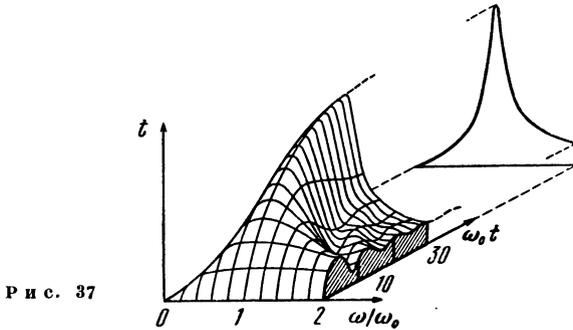
Это выражение несколько отличается от (24. 4) вследствие того, что формула (24. 4) не является точной; она получена из укороченных уравнений (24. 3). Интересно заметить, что в точном решении аргумент косинуса имеет более простой вид.

Перейдем теперь к исследованию динамической характеристики. Для того, чтобы можно было составить себе ясное представление о величине c как функции двух переменных ω и t , на рис. 37 изображен ее рельеф. Кривые, лежащие в плоскостях, параллельных $cO\omega/\omega_0$, представляют собой динамические резонансные кривые. Статическая резонансная кривая также показана на рисунке как предел, к которому стремятся динамические кривые.

Кривые, лежащие в плоскостях, параллельных $cO\omega_0 t$, представляют собой кривые установления амплитуды колебаний в

резонаторе при различных расстройках. Эти кривые имеют колебательный характер вследствие биений между вынуждающей частотой и частотой собственных колебаний резонатора, возбуждаемых в момент включения. Частота биений, очевидно, тем больше, чем больше расстройка. При $\omega = \omega_0$, как известно, процесс установления совершается по экспоненциальному закону.

Рассматривая рис. 37, мы видим, что острота резонансной кривой резонатора зависит от времени включения резонатора;



Р и с. 37

острота резонанса, а следовательно, и разрешающая способность тем больше, чем больше время включения. Так возникает понятие о динамической разрешающей способности анализатора. Динамическая разрешающая способность в отличие от статической зависит не только от параметров резонатора, но и от времени.

Выведем некоторые соотношения. Положим, что мы условились выражать остроту резонанса шириной $\Delta\omega$ резонансной кривой на высоте, равной $1/2$ максимальной ординаты. Пользуясь (24.4), заметим, что максимум амплитуды получается при $\beta=0$, так что $c_{\max} = (1 - e^{-\alpha t})/2\alpha$. Составляя выражение $c^2/c_{\max}^2 = 1/2$, получим уравнение

$$1 - \cos \beta t - \frac{1}{2} (\beta^2/\alpha^2 - 1) (\text{ch} \alpha t - 1) = 0.$$

При $t > 0$ это уравнение принимает вид

$$1 - \cos \beta t - \frac{1}{4} \beta^2 t^2 = 0.$$

Решая эти уравнения методом ложных корней, получаем для $\alpha=0,1$ решение, приведенное ниже

t	0	10	20	30
βt	2,78	2,87	3,13	3,57

Теперь нам нужно перейти от β к ω . Предполагая, как обычно делается в радиотехнике, расстройку малой и резонансную кривую симметричной¹, получим

$$\beta = \frac{\omega_0^2 - \omega^2}{2\omega} = \frac{(\omega_0 + \omega)(\omega_0 - \omega)}{2\omega} \approx \omega_0 - \omega = \frac{\Delta\omega}{2},$$

т. е.

$$2\beta \approx \Delta\omega.$$

Теперь мы можем составить следующее:

t	0	10	20	30	∞
$\Delta\omega/\omega_0$		0,575	0,313	0,238	0,200
$\Delta\omega t$	5,56	5,75	6,26	7,14	∞

Первая строка (рис. 38, *a*) дает относительную ширину резонансной кривой и характеризует, таким образом, динамическую разрешающую способность резонатора. С течением времени $\Delta\omega/\omega_0$ стремится к постоянному пределу d — эта величина выражает относительную ширину статической резонансной кривой.

Вторая строка (рис. 38, *b*) дает произведение $\Delta\omega t$. При малых t это есть постоянная величина и притом, что весьма замечательно, не зависящая от затухания резонатора. С течением времени острота резонансной кривой стремится к постоянному пределу, а время неограниченно нарастает. Следовательно, кривая зависимости $\Delta\omega t$ от t имеет асимптотой прямую, проходящую через начало координат. Наклон этой прямой равен, очевидно, d .

Для уяснения понятия динамической разрешающей способности анализатора на рис. 39 представлена картина развития показания анализатора с течением времени². Предполагается, что

¹ Указанные предположения необязательны. Можно получить тот же результат более общим путем. У нас обозначено

$$\beta = \frac{|\omega_0^2 - \omega^2|}{2\omega}, \quad \frac{\beta}{\omega_0} = \frac{1}{2} \left(\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0} \right).$$

Составляя и решая два квадратных уравнения относительно $\epsilon = \omega/\omega_0$, мы получаем, сохраняя лишь имеющие смысл корни

$$\epsilon_1 = \sqrt{(\beta/\omega_0)^2 + 1} - \beta/\omega_0, \quad \epsilon_2 = \sqrt{(\beta/\omega_0)^2 + 1} + \beta/\omega_0.$$

Таким образом,

$$\epsilon_2 - \epsilon_1 = (\omega_2 - \omega_1)/\omega_0 = \Delta\omega/\omega_0 = 2\beta/\omega_0,$$

откуда по-прежнему

$$\Delta\omega = 2\beta.$$

При этом выводе мы не предполагали ни малости расстройки, ни симметрии резонансной кривой.

² Рис. 39 построен по тем же данным, что и рис. 37. Однако детали опущены и весь рельеф несколько сглажен, так что рисунок воспроизводит явление лишь в общих чертах.

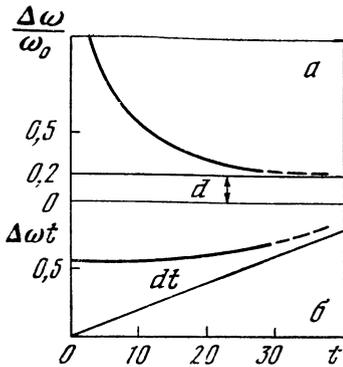


Рис. 38

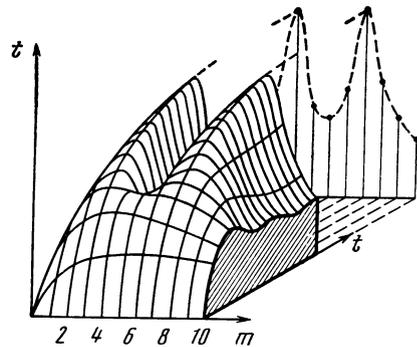


Рис. 39

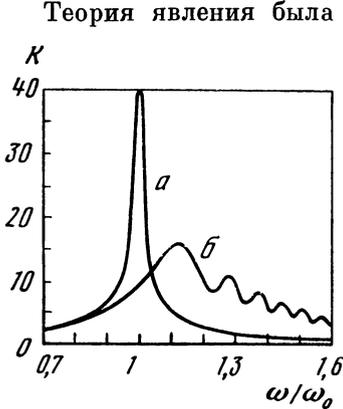
анализатор представляет собой набор резонаторов с одинаковым затуханием и с равноотстоящими резонансными частотами. На анализатор подается при $t=0$ два синусоидальных колебания. Как видно из рис. 39, в начальной стадии явления анализатор не разделяет этих колебаний. Лишь по происшествии некоторого времени начинает сформировываться седловина, постепенно углубляющаяся. В пределе получается установившееся показание анализатора, огибающая которого имеет вид двугорбой кривой (см. § 19).

§ 25. Динамическая характеристика резонатора при воздействии изменяющейся частоты

В предыдущем параграфе был рассмотрен вопрос о динамической разрешающей способности резонатора в системе одновременного анализа. Теперь мы разберем соотношения при последовательном анализе, т. е. в случае, когда изменяется возбуждающая резонатор частота.

При воздействии на резонатор изменяющейся во времени частоты наблюдаются резонансные явления. Однако если изменение частоты происходит достаточно быстро, то наиболее интенсивные колебания происходят не в тот момент, когда возбуждающая частота совпадает с резонансной, а несколько позже, так как резонатор не успевает раскачаться; по этой же причине и наибольшая амплитуда колебаний оказывается меньше, чем при резонансе в установившемся режиме. Если построить график зависимости амплитуды колебаний от мгновенной частоты, то получится динамическая резонансная кривая. Эта кривая при быстром изменении частоты может существенно отличаться от статической резонансной кривой, а именно: 1) положение максимума смещается по шкале частот (что соответствует смещению по шкале времен

в сторону запаздывания); 2) высота максимума убывает; 3) кривая становится асимметричной — левый скат делается более пологим; 4) полоса пропускания возрастает (т. е. разрешающая способность убывает); 5) явление осложняется еще убыстряющимися колебаниями амплитуды, обусловленными биениями между вынужденными колебаниями и собственными колебаниями, возникшими при прохождении через резонанс. Все эти особенности видны на рис. 40, на котором представлены статическая (а) и динамическая (б) резонансные кривые.



Р и с. 40

Теория явления была предметом многочисленных исследований [5, 8, 24, 28, 31]. Нижеследующее изложение основывается на работах Турбовича [20, 21], дающих достаточно общий подход к проблеме, и в то же время позволяющих получить ответ на вопросы, вытекающие из потребностей практики.

Пусть частота изменяется со временем по линейному закону $\omega = \omega_1 + \lambda t$ ($t > 0$). Напряжение, действующее на резонатор, можно записать в комплексной форме $U_1 = e^{j\theta(t)}$, где

$$\theta(t) = \int_0^t \omega dt = \omega_1 t + \frac{1}{2} \lambda t^2.$$

Выходное напряжение представим интегралом Дюамеля

$$U_2 = \int_0^t g(\tau) U_1(t - \tau) d\tau = \int_0^t g(\tau) e^{j\theta(t-\tau)} d\tau,$$

где $g(\tau)$ — импульсная реакция резонатора. Подставляя значение θ , находим

$$U_2 = e^{j(\omega_1 t + \frac{1}{2} \lambda t^2)} \int_0^t g(\tau) e^{-j(\omega_1 \tau + \lambda t \tau - \frac{1}{2} \lambda \tau^2)} d\tau.$$

Но

$$e^{j(\omega_1 t + \lambda t^2 / 2)} = U_1; \quad \omega_1 + \lambda t = \omega,$$

и, таким образом,

$$U_2 = U_1 \int_0^t g(\tau) e^{-j(\omega \tau - \lambda \tau^2 / 2)} d\tau.$$

Динамический коэффициент передачи определим как отношение комплексных напряжений U_2 и U_1

$$K_D = \frac{U_2}{U_1} = \int_0^t g(\tau) e^{-j\omega\tau} e^{j\lambda\tau^2/2} d\tau. \quad (25.1)$$

Это — точная формула, имеющая достаточно универсальный характер, так как она годится для любого резонатора.

Перейдем теперь к приближениям. Прежде всего заметим, что $g(t)$ есть функция, содержащая экспоненциально убывающие множители. Поэтому с незначительной погрешностью (которая оценивается ниже) можно заменить верхний предел на бесконечность

$$K_D \simeq \int_0^{\infty} g(\tau) e^{-j\omega\tau} e^{j\lambda\tau^2/2} d\tau. \quad (25.2)$$

При $\lambda \rightarrow 0$, т. е. при переходе к статическому режиму, получим известное выражение для статического коэффициента передачи как преобразования Фурье импульсной реакции

$$K = \int_0^{\infty} g(\tau) e^{-j\omega\tau} d\tau. \quad (25.3)$$

Таким образом, разность между динамическими коэффициентами передачи равна

$$\Delta K = K_D - K = \int_0^{\infty} g(\tau) e^{-j\omega\tau} (e^{j\lambda\tau^2/2} - 1) d\tau. \quad (25.4)$$

Далее отметим, что нас интересует случай, когда динамическая резонансная кривая мало отличается от статической, так как только при этом условии погрешность анализатора, рассчитанного по статическим параметрам, будет мала. Полагая аргумент $\lambda\tau^2$ малым для всех значений τ , для которых убывающая функция $g(\tau)$ имеет еще заметную величину, можем разложить выражение в круглых скобках под знаком интеграла в (25.4) в степенной ряд и ограничиться для первого приближения всего двумя членами разложения. Это даст

$$\Delta K \simeq \int_0^{\infty} g(\tau) e^{-j\omega\tau} \left(j \frac{1}{2} \lambda \tau^2 - \frac{1}{8} \lambda^2 \tau^4 \right) d\tau.$$

Теперь заметим, что появление множителей τ^n можно представить как результат n -кратного дифференцирования по параметру ω . Следовательно,

$$\Delta K \simeq -j \frac{1}{2} \lambda \frac{d^2}{d\omega^2} \int_0^{\infty} g(\tau) e^{-j\omega\tau} d\tau - \frac{1}{8} \lambda^2 \frac{d^4}{d\omega^4} \int_0^{\infty} g(\tau) e^{-j\omega\tau} d\tau,$$

т. е.

$$\Delta K \simeq -j \frac{1}{2} \lambda K'' - \frac{1}{8} \lambda^2 K^{IV}.$$

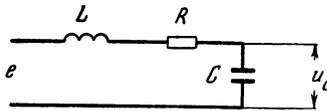
Итак, для динамического коэффициента передачи имеем

$$K_{\text{д}} = K + \Delta K = K - j \frac{1}{2} \lambda K'' - \frac{1}{8} \lambda K^{IV}. \quad (25.5)$$

Полагая $K = A + jB$, $C = |K|^2 = A^2 + B^2$, можем составить общее выражение для квадрата модуля динамического коэффициента передачи

$$C_{\text{д}} = |K_{\text{д}}|^2 = C + \lambda (AB'' - BA'') + \\ + \frac{1}{4} \lambda^2 [(A'')^2 + (B'')^2 - AA^{IV} - BB^{IV}]. \quad (25.6)$$

Приближенные формулы (25.5) и (25.6) имеют общий характер.



Р и с. 41

В дальнейшем мы будем рассматривать только простейший резонатор по схеме рис. 41, для которого

$$K \simeq \frac{1}{1 - \frac{\omega^2}{\omega_0^2} + j \frac{\omega}{\omega_0} d} \simeq \frac{1}{-\varepsilon + jd} = \frac{1}{jd} \cdot \frac{1}{1 + jx},$$

где $x = \varepsilon/d = 2(\omega - \omega_0)/\omega_0 d$ — обобщенная расстройка. С точностью до постоянного множителя

$$K = 1/(1 + jx), \quad (25.7)$$

этим выражением мы и воспользуемся в дальнейшем.

Поставим задачу дальнейшего исследования. Имея в виду практические вопросы расчета и построения анализаторов (а также характеристографов и других приборов, использующих модулированное по частоте испытательное напряжение), мы должны, очевидно, выяснить, как зависят те или иные параметры, определяющие резонансную кривую, от скорости изменения частоты. В частности, нас интересует: 1) смещение максимума s , 2) уменьшение высоты максимума ΔC , 3) смещение полосы пропускания $\frac{1}{2}(s_1 + s_2)$, 4) изменение ширины полосы пропускания $s_2 - s_1$. Смысл введенных здесь обозначений поясняется рис. 42, на котором по-прежнему a — статическая, b — динамическая резонансные кривые.

Для определения перечисленных величин будем исходить из выражения (25. 5) для динамического коэффициента передачи, переписав его в новых переменных

$$K_D = K - j \frac{1}{2} \lambda \frac{d^2 K}{d\omega^2} - \frac{1}{8} \lambda^2 \frac{d^4 K}{d\omega^4} =$$

$$= K - j \frac{1}{2} \lambda \left(\frac{dx}{d\omega} \right)^2 \frac{d^2 K}{dx^2} - \frac{1}{8} \lambda^2 \left(\frac{dx}{d\omega} \right)^4 \frac{d^4 K}{dx^4}$$

или

$$K_D = K - j\mu K_D'' - \frac{1}{8} \mu^2 K_D^{IV}, \quad (25. 8)$$

где $\mu = \lambda (dx/d\omega)^2 = 4\lambda/\omega_0^2 d^2$.

В формуле (25. 8) $K = K(x)$ и производные берутся по аргументу x . Для модуля динамического коэффициента передачи находим

$$|K_D^2| = C_D = C - 4\mu x C^3 -$$

$$- \mu^2 (5C^3 - 48x^2 C^5). \quad (25. 9)$$

Максимум статического коэффициента передачи достигается при $x=0$

$$C_{\max} = C(0) = 1, \quad C'(0) = 0.$$

Максимум динамического коэффициента передачи имеет место при некотором значении $x=s$, которое и является искомым. Разложим C_D' в ряд Маклорена, ограничившись членом первого порядка,

$$C_D'(s) = C_D'(0) + s C_D''(0) = 0.$$

Подставляя сюда значение C_D из (25. 9) и сохраняя лишь члены до μ , получим уравнение относительно s ; решение этого уравнения есть

$$s = 2\mu, \quad (25. 10)$$

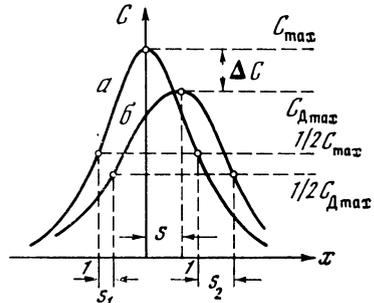
т. е. смещение максимума при динамическом резонансе имеет тот же порядок малости, что и параметр μ , пропорциональный скорости изменения частоты.

Для нахождения относительного изменения высоты максимума составим выражение

$$\Delta C = C_D(s) - C(0) = C_D(0) + s C_D'(0) - C(0).$$

Подставляя сюда C_D и s из (25. 9) и (25. 10), найдем с точностью до μ^2

$$\Delta C = -\mu^2. \quad (25. 11)$$



Р и с. 42

Таким образом, изменение высоты максимума имеет второй порядок малости относительно μ .

Определим теперь значения обобщенной расстройки x , при которых квадрат модуля коэффициента передачи равен половине максимального значения (в соответствии с обычным определением ширины полосы пропускания). Для статического коэффициента передачи

$$C(\pm 1) = \frac{1}{2} C(0) = \frac{1}{2}.$$

Для динамического коэффициента передачи

$$C_D(-1 + s_1) = C_D(1 + s_2) = \frac{1}{2} C_D(s) = \frac{1}{2} (1 - \mu)^2.$$

Разложив это выражение в ряд по искомым малым параметрам s_1 и s_2 (с точностью до μ^2), получим уравнения для определения этих величин. Разность $s_2 - s_1$ дает приращение относительной ширины полосы пропускания; вычисление дает

$$s_2 - s_1 = \frac{5}{2} \mu^2. \quad (25.12)$$

Расширение полосы пропускания имеет, как видим, второй порядок малости относительно μ . Центр полосы пропускания для динамической резонансной кривой смещен относительно $x=0$ на величину полусуммы s_1 и s_2 . Вычисление дает

$$\frac{1}{2} (s_1 + s_2) = \mu. \quad (25.13)$$

Сравнивая (25.13) и (25.10), замечаем, что смещение центра полосы вдвое меньше смещения максимума. Это обусловлено уже упоминавшейся асимметрией динамической резонансной кривой.

Итак, одночленные выражения первого приближения для рассмотренных параметров динамической резонансной кривой пропорциональны μ или μ^2 , т. е. имеют первый или второй порядок малости, если μ — величина первого порядка малости. Грубое искажение резонансной кривой, показанное на рис. 42, едва ли встретится в практике построения радиоизмерительных устройств; динамическая резонансная кривая рис. 42 относится к значению $\mu \approx 30$ ($d=0,025$).

Вернемся от обобщенных переменных к исходным. Обозначим через $\Delta\omega$ ширину полосы пропускания; знаком δ обозначим приращение, получаемые теми или иными параметрами в динамическом режиме. Тогда вместо формул (25.10)—(25.12) будем иметь

$$\delta\omega_0/\omega_0 = \mu d = 4\lambda/\omega_0^2 d, \quad (25.10')$$

$$\delta C_{\max}/C_{\max} = -\mu^2 = -16\lambda^2/\omega_0^4 d^4, \quad (25.11')$$

$$\delta(\Delta\omega)/\Delta\omega = 5\mu^2/4 = 20\lambda^2/\omega_0^4 d^4. \quad (25.12')$$

Необходимое время анализа (или период частотной модуляции при пилообразном законе модуляции) может быть определено по формуле

$$T = \frac{2}{\pi} \cdot \frac{F}{\mu (\Delta f)^2},$$

где F — размах модуляции, т. е. ширина обследуемого диапазона частот; Δf — разрешающая способность, т. е. ширина полосы пропускания (при выводе этой формулы положено $\Delta f/f_0 = d$).

Так, например, для гетеродинного анализатора звуковой частоты ($F = 10$ кГц) с электромеханическим резонатором, настроенным на частоту 50 кГц, при разрешающей способности 25 Гц (что соответствует $d = 5 \cdot 10^{-4}$), задаваясь $\mu = 0,1$ (т. е. допуская расширение полосы на величину порядка 1%), получим

$$T = \frac{2}{\pi} \cdot \frac{10^4}{0,1 (25)^2} \approx 100 \text{ сек.}$$

На практике считают достаточным ставить менее жесткие требования: полагают $\mu \leq 1$. При μ порядка единицы вышеприведенные приближенные формулы, разумеется, уже не применимы.

Вернемся к исходному выражению (25. 1) для динамического коэффициента передачи. Мы заменили в этом выражении верхний предел на бесконечность (см. (25. 2)), полагая, что вносимая этой заменой погрешность невелика. Нам нужно теперь оценить величину погрешности, равной

$$R = \int_0^t g(\tau) e^{-j\omega\tau} e^{j\lambda\tau^2/2} d\tau.$$

Модуль интеграла меньше интеграла от модуля; поэтому

$$|R| < \int_0^t |g(\tau)| d\tau.$$

Подставляя

$$g(t) = \omega_0 e^{-\alpha t} \sin \omega_0 t,$$

запишем

$$|R| < \omega_0 \int_0^t e^{-\alpha\tau} d\tau = \frac{2}{\alpha} e^{-\alpha t}.$$

Но

$$\alpha t = \alpha \frac{\omega - \omega_1}{\lambda} = \frac{x - x_1}{\mu}$$

и, стало быть,

$$|R| < \frac{2}{d} e^{-(x-x_1)\mu}.$$

Относительную погрешность получим, отнеся R к $K_{\max} = 1/jd$

$$\left| \frac{R}{K_{\max}} \right| = 2e^{-(x-x_1)\mu}.$$

Все расчеты велись с точностью до μ^2 . Для того чтобы относительной погрешностью можно было пренебречь, достаточно, чтобы она имела следующий порядок малости, т. е. порядок μ^3 . Отбрасывая не влияющий на порядок множитель 2, будем иметь

$$e^{-(x-x_1)\mu} \leq \mu^3$$

или после логарифмирования

$$x - x_1 \geq 3\mu \ln \frac{1}{\mu}.$$

Смысл этого неравенства состоит в том, что для получения достаточно точных результатов в полосе частот, начиная с обобщенной расстройки x , нужно, чтобы начальная частота, соответствующая значению x , была достаточно удалена от границы исследуемой полосы. Пусть, например, $\mu = 0,1$. Тогда $x - x_1 \geq 3 \cdot 0,1 \cdot \ln 10 = 0,69$.

Напомним, что ширина полосы пропускания простого резонатора, выраженная через обобщенную расстройку, равна 2. Таким образом, нужно отступить влево от края исследуемой полосы всего лишь на долю полосы пропускания резонатора, для того чтобы иметь право пренебречь погрешностью R при переходе от (25. 1) к (25. 2). Это условие, конечно, на практике всегда выполняется, и все соотношения, выведенные на основании формулы (25. 2), таким образом, подтверждены.

Предыдущее изложение основывалось на предположении малости отличия динамической частотной характеристики от статической; считалось само собою разумеющимся, что мы должны стремиться к такому положению. Это справедливо, если мы рассчитываем анализатор на основе параметров статической характеристики резонатора. В этом случае различие между статической и динамической характеристиками рассматривается как источник погрешности, и желательно, разумеется, чтобы эта погрешность была мала. Но возможен и иной подход к вопросу. Именно, если скорость изменения частоты постоянна, то анализатор может рассчитываться на основе параметров динамической характеристики. В этом случае скорость изменения частоты может быть как угодно велика, если только значения динамических параметров (например, ширина полосы пропускания) нас устраивают.

В заключение интересно заметить, что вопрос о явлениях, возникающих при воздействии изменяющейся частоты на резо-

нансную систему представляет интерес и для других отраслей техники. Так, при пуске машин с неуравновешенными вращающимися массами на гибком валу, могут возникать опасные колебания при прохождении через резонанс (так называемая «критическая скорость»). Математическая формулировка задачи в точности совпадает с нашей [8, 31]. Однако цель расчета противоположна: нужно подобрать безопасный режим пуска, т. е. обеспечить настолько быстрое изменение скорости вращения, чтобы, так сказать, «проскочить» через резонанс, не возбудив на критической скорости колебаний чрезмерной амплитуды. В наших обозначениях это значит, что нужно, задаваясь допустимым наибольшим значением $K_{Д \max}$, выбрать наименьшее значение $\mu \gg 1$.

§ 26. Анализ одиночных импульсов

Анализ непериодических явлений, в частности импульсов, и по постановке задачи и по картине явлений во многом отличается от анализа периодических явлений.

Для уяснения этих особенностей мы рассмотрим одну из возможностей анализа импульсов. Напомним, что для анализа импульсов метод последовательного анализа совершенно непригоден. Для анализа импульсов может применяться только метод одновременного анализа (в этом уже проявляется различие анализа импульсов и анализа периодических явлений).

Положим, что мы желаем применить для анализа импульсов анализатор, представляющий собой набор настроенных на различные частоты резонаторов. Такой анализатор может служить для анализа периодических явлений; его пригодность для анализа импульсов нужно еще установить.

Покажем прежде всего, что набор резонаторов без затухания может дать точные дискретные значения спектральной плотности импульса. Дело в том, что спектр импульса есть непрерывная кривая, представляющая спектральную плотность как функцию частоты. Ясно, что набор, состоящий из конечного числа резонаторов, может дать показание, представляемое графически лишь совокупностью отдельных точек. Можно лишь желать, чтобы эти точки лежали на кривой сплошного спектра. Кривая будет определена тем лучше, чем больше число точек, т. е. чем больше число резонаторов. Это соображение дает представление о том, как ставится в данном случае вопрос о разрешающей способности анализатора.

Пусть уравнение резонатора без затухания записано в виде

$$U'' + \omega_k^2 U = \omega_k^2 a_k f(t), \quad (26.1)$$

где $f(t)$ — анализируемый импульс, который мы будем полагать равным нулю вне промежутка $0 < t < t$; k — порядковый номер резонатора; U — выходное напряжение; a_k — коэффициент пропорциональности, задаваемый входным делителем.

Решение (26. 1) можно представить в виде

$$U(t) = a_k \int_0^t g(t-u) f(u) du, \quad (26. 2)$$

где $g(t)$ — временная характеристика, т. е. решение (26. 1) при равной части вида

$$f(t) = \delta(t).$$

Для рассматриваемого случая мы имеем

$$g(t) = a_k \omega_k \sin \omega_k t,$$

где $\omega_k = 1/\sqrt{L_k C_k}$ — собственная частота k -го резонатора. Таким образом

$$\begin{aligned} U(t) &= a_k \omega_k \int_0^t \sin \omega_k (t-u) f(u) du = \\ &= a_k \omega_k \left[\sin \omega_k t \int_0^t f(u) \cos \omega_k u du - \cos \omega_k t \int_0^t f(u) \sin \omega_k u du \right]. \end{aligned}$$

Найдем огибающую этого колебания; она будет выражена через сумму квадратов коэффициентов при $\sin \omega_k t$ и $\cos \omega_k t$

$$c = a_k \omega_k \left[\left(\int_0^t f(u) \cos \omega_k u du \right)^2 + \left(\int_0^t f(u) \sin \omega_k u du \right)^2 \right]^{1/2}. \quad (26. 3)$$

Эта величина (представляющая собой переменную «амплитуду» колебания) есть функция времени; зависимость от времени определяется тем, что t входит в верхние пределы интегралов.

Но после $t = \tau$ импульс $f(t)$, а с ним и подынтегральные выражения обращаются в нуль, и, следовательно, в дальнейшем, т. е. при $t > \tau$, величина c сохраняет постоянное значение, которое она приобрела к моменту времени $t = \tau$; соответствующее значение обозначим через c_τ . Физически это означает, что после окончания внешнего воздействия на резонатор амплитуда его колебания останется в дальнейшем неизменной (напомним, что мы пока рассматриваем резонатор без затухания). Но в таком случае мы можем записать выражение c_τ с любыми постоянными пределами интегрирования, включающими в себя промежуток $0 - \tau$, в частности, и с бесконечными пределами

$$c_\tau = a_k \omega_k \left[\left(\int_{-\infty}^{+\infty} f(u) \cos \omega_k u du \right)^2 + \left(\int_{-\infty}^{+\infty} f(u) \sin \omega_k u du \right)^2 \right]^{1/2}. \quad (26. 4)$$

В интегралах в правой части легко узнать выражения для синусной и косинусной компонент спектра, так что

$$c_\tau = a_k \omega_k \sqrt{a^2 + b^2} = a_k \omega_k \Phi(\omega). \quad (26.5)$$

Если теперь выбрать множитель $a_k \omega_k$ одинаковым для всех резонаторов, то получается, что наблюдаемые по окончании действия импульса амплитуды колебания резонаторов (с точностью до постоянного множителя) дают точные дискретные значения спектра для частот.

Что же касается промежутка $0 < t < \tau$, то в течение этого времени амплитуды колебания резонаторов дают текущий спектр импульса, как это видно из формулы (26. 3).

Нужно заметить, что рассмотренный анализатор, составленный из резонаторов без затухания и дающий, как мы убедились, идеальный анализ импульса, принципиально непригоден для анализа периодических явлений. Дело в том, что для резонатора без затухания установившийся режим вообще не существует, а анализ периодических явлений мыслится именно в установившемся режиме. В этом проявляется еще одно существенное различие анализа периодических явлений и импульсов.

Следует еще заметить, что время анализа импульса определяется длительностью импульса и никак не зависит от свойств анализатора. Следовательно, рассмотренное в [23] соотношение между временем анализа и разрешающей способностью анализатора к анализу импульсов неприменимо; здесь опять-таки проявляется различие между анализом периодических явлений и импульсов.

Посмотрим теперь, как изменятся соотношения, если мы составим анализатор импульсов из реальных резонаторов с затуханием. Уравнение резонатора в этом случае следующее:

$$\dot{U} + 2\alpha\dot{U} + \omega_k^2 U = a_k \omega_k^2 f(t). \quad (26.6)$$

Временная характеристика имеет вид

$$g(t) = \omega_k^{-\alpha k t} \sin \omega_k t.$$

Решение уравнения (26. 6) есть

$$\begin{aligned} U(t) &= a_k \omega_k \int_0^t e^{-\alpha_k(t-u)} \sin \omega_k(t-u) f(u) du = \\ &= a_k \omega_k e^{-\alpha k t} \left[\sin \omega_k t \int_0^t e^{\alpha_k u} f(u) \cos \omega_k u du - \right. \\ &\quad \left. - \cos \omega_k t \int_0^t e^{\alpha_k u} f(u) \sin \omega_k u du \right]. \end{aligned}$$

Для огибающей при $t > \tau$ можем записать

$$c = a_k \omega_k e^{-\alpha_k t} \left[\left(\int_{-\infty}^{+\infty} e^{\alpha_k u} f(u) \cos \omega_k u du \right)^2 + \left(\int_{-\infty}^{+\infty} e^{\alpha_k u} f(u) \sin \omega_k u du \right)^2 \right]^{1/2}. \quad (26.7)$$

Сравнивая (26.7) и (26.4), замечаем следующее. Во-первых, теперь амплитуда при $t > \tau$ не сохраняет постоянного значения, а убывает за счет множителя $e^{-\alpha_k t}$, так как свободное колебание затухает. Это значит, что показание анализатора зависит от того, в какой момент (конечно, после $t = \tau$) оно будет отсчитано. Однако если сделать показатель затухания α_k одинаковым для всех резонаторов, то с течением времени спектр не будет искажаться; будет изменяться только масштабный коэффициент, что несущественно при том условии, что отсчет по всем резонаторам берется одновременно. Во-вторых, наличие затухания проявляется в том, что в подынтегральных выражениях появился весовой множитель $e^{\alpha u}$. Его наибольшее значение равно $e^{\alpha \tau}$; это значение тем меньше отличается от единицы, чем меньше произведение $\alpha \tau$, т. е. чем меньше затухание, и чем короче импульс.

В ряде случаев может оказаться возможным положить приближенно $e^{\alpha \tau} \approx 1$, т. е. пренебречь влиянием затухания, и свести, таким образом, рассматриваемый случай к предыдущему. Если же этого сделать нельзя, то показание анализатора не будет отвечать истинному спектру импульса, и задача состоит в том, чтобы оценить совершаемую погрешность.

Из формулы (26.7) видно, что анализатор дает нам спектр не функции $f(t)$, а измененной (взвешенной) функции $f_\alpha(t) = e^{\alpha t} f(t)$. Но если спектр функции $f(t)$ есть

$$S(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt,$$

то спектр измененной функции $f_\alpha(t)$ будет

$$S_\alpha = \int_{-\infty}^{+\infty} e^{\alpha t} f(t) e^{-j\omega t} dt = S(\omega + j\alpha).$$

Таким образом, вместо истинных значений $\Phi(\omega_k)$ мы отсчитаем по анализатору значения

$$\Phi_\alpha = |S(\omega_k + j\alpha)|. \quad (26.8)$$

По этой формуле можно подсчитать обусловленное наличием затухания искажение спектра для данного конкретного импульса.

В качестве примера разберем случай анализа прямоугольного импульса длительностью τ . Аналитическое выражение такого импульса может быть записано в виде

$$f(t) = \sigma(t) - \sigma(t - \tau).$$

Таким образом,

$$\begin{aligned} \frac{1}{a_k} U(t) &= \int_0^t g(t-u) f(u) du = \int_0^t g(t-u) [\sigma(u) - \sigma(u - \tau)] du = \\ &= \int_0^t g(t-u) du - \int_0^t g(t-u) \sigma(u - \tau) du = \int_0^t g(u) du \quad 0 < t < \tau, \\ &= \int_0^\tau g(t-u) du - \int_{t-\tau}^t g(u) du \quad \tau < t < \infty. \end{aligned}$$

Подставляя сюда значение

$$g(t) = \omega_k e^{-\alpha_k t} \sin \omega_k t,$$

выполняя интегрирование и полагая $\alpha_k \ll \omega_k$, получим приближенные решения

$$\frac{1}{a_k} U(t) = \begin{cases} 1 - e^{-\alpha_k t} \cos \omega_k t & 0 < t < \tau, \\ e^{-\alpha_k t} [e^{\alpha_k \tau} \cos \omega_k (t - \tau) - \cos \omega_k t] & \tau < t < \infty. \end{cases} \quad (26.9)$$

Полагая $a_k \omega_k = 1$, $\alpha_k = \alpha$, после несложных вычислений получим для огибающей

$$c(t) = \begin{cases} \frac{1}{\omega_k} (1 + e^{-\alpha t}) & 0 < t < \tau, \\ \frac{1}{\omega_k} M e^{-\alpha t} & \tau < t < \infty, \end{cases} \quad (26.10)$$

где

$$M = \sqrt{1 - 2e^{\alpha\tau} \cos \omega_k \tau + e^{2\alpha\tau}}. \quad (26.11)$$

Физическая картина явлений такова: при $t=0$ $f(t)$ изменяется скачком от 0 до 1, в результате чего возникает свободное колебание резонатора; при $t = \tau$ $f(t)$ изменяется скачком в обратном направлении, т. е. от 1 до 0. При этом возникает точно такое же колебание, но с обратным знаком и с запаздыванием на τ (рис. 43, а).

Амплитуда результирующего колебания, устанавливающегося после $t = \tau$ (рис. 43, б) зависит, понятно, от относительной фазы обоих свободных колебаний; фазовый сдвиг зависит в свою очередь как от частоты ω_k , так и от интервала τ . Вот почему величина M есть функция аргумента $\omega_k \tau$. Эта зависимость и определяет способность анализатора разобраться в спектральных соотношениях.

Заметим, что колебание x непрерывно, но огибающая c претерпевает разрыв при $t = \tau$. Первая строка формулы (26. 10) включает

постоянную составляющую; огибающая по формуле (26. 10) отмечена на рис. 43,б.

Для спектра прямоугольного импульса длительностью τ , существующего между $-\tau/2$ и $+\tau/2$, мы имели (см. § 11)

$$S(\omega) = \frac{\sin \omega\tau/2}{\omega\tau/2}.$$

В нашем случае импульс задан на промежутке $0 < t < \tau$; значит мы имеем для его спектра

$$S(\omega) = e^{-j\omega\tau/2} \frac{\sin \omega\tau/2}{\omega\tau/2}.$$

Применим формулу (26. 8) для получения искаженного спектра

$$\Phi_\alpha = |S(\omega + j\alpha)| = \left| e^{\frac{\tau}{2}(-j\omega + \alpha)} \frac{\sin \frac{\tau}{2}(\omega + j\alpha)}{\frac{\tau}{2}(\omega + j\alpha)} \right|,$$

что дает после выполнения вычислений

$$\Phi_\alpha = \frac{1}{\omega\tau} \sqrt{1 - 2e^{\alpha\tau} \cos \omega\tau + e^{2\alpha\tau}}.$$

Мы видим, таким образом, что, как оно и должно быть, величина M (см. формулы (26. 10) и (26. 11)) непосредственно выражает наблюдаемый спектр. Что же касается искажения спектра, то оно проявляется наиболее заметным образом в том, что вместо нулей истинного спектра при $\omega\tau = 2n\pi$ искаженный спектр Φ_α имеет при тех же аргументах конечные значения

Очень интересно отметить, что до $t = \tau$ анализатор еще, так сказать, «не знает», с чем он имеет дело. Он получает в начале импульса толчок, возбуждающий колебания во всех резонаторах. В дальнейшем никакие дополнительные возмущения на резонаторы не действуют вплоть до $t = \tau$. Значит анализатор воспринимает прямоугольный импульс как единичную функцию, и его показание дает спектр этой функции, как видно из первой строки формулы (26. 10). При $t = \tau$ положение внезапно определяется; импульс сформировался и закончился, и с этого момента анализатор дает спектр импульса¹.

$$\Phi_{\alpha_{\min}} = \frac{1 - e^{\alpha\tau}}{2n\pi}.$$

Очень интересно отметить, что до $t = \tau$ анализатор еще, так сказать, «не знает», с чем он имеет дело. Он получает в начале импульса толчок, возбуждающий колебания во всех резонаторах. В дальнейшем никакие дополнительные возмущения на резонаторы не действуют вплоть до $t = \tau$. Значит анализатор воспринимает прямоугольный импульс как единичную функцию, и его показание дает спектр этой функции, как видно из первой строки формулы (26. 10). При $t = \tau$ положение внезапно определяется; импульс сформировался и закончился, и с этого момента анализатор дает спектр импульса¹.

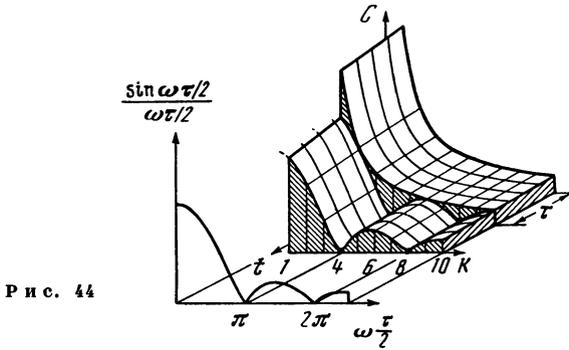
Рис. 43

¹ Общая картина явления ясна, но остается еще показать, почему при $t < \tau$ анализатор не дает текущего спектра прямоугольного импульса, чего сле-

Возьмем численный пример. Пусть анализатор состоит из 10 резонаторов со следующими данными: $\omega_k = k\omega_0$, $\omega_0 = 1$, $\alpha_k = \alpha = 0,1$, длительность прямоугольного импульса пусть будет $\tau = \pi/2$. Тогда нули спектра придутся на значения $k=4$ и $k=8$. Множитель M имеет следующие значения:

$$\begin{aligned} \sqrt{1 + e^{2\alpha\tau}} &= 1,54 \quad \text{для } k=1, 3, 5, 7, 9, \\ M = 1 + e^{\alpha\tau} &= 2,17 \quad \text{для } k=2, 6, 10, \\ |1 - e^{\alpha\tau}| &= 0,17 \quad \text{для } k=4, 8. \end{aligned}$$

По формулам (26. 10) на рис. 44 построен рельеф $c(k, t)$, где амплитуда представлена как функция номера резонатора и времени.



Р и с. 44

При $t < \tau$ анализатор дает гиперболический спектр единичной функции. При $t = \tau$ амплитуда изменяется скачком; в этот момент сформировывается спектр прямоугольного импульса. Как видим, амплитуды резонаторов $k=4$ и $k=8$ не равны нулю (искажение спектра). Истинный спектр прямоугольного импульса изображен в подходящем масштабе на переднем плане. В общем анализатор воспроизводит спектр вполне удовлетворительно.

Поскольку для всех резонаторов выбрано одинаковое значение показателя затухания α , то любое сечение рельефа плоскостью, параллельной sOk (т. е. снятие показания анализатора в любой момент времени), дает фигуры, различающиеся только вертикальным масштабом.

Итак, мы убедились, что анализатор в форме набора резонаторов может с успехом применяться для анализа импульсов¹.

довало бы ожидать на основании формулы (26. 3). Дело в том, что, вычислив огибающую по этой формуле, мы действительно получили бы текущий спектр прямоугольного импульса, при этом оказалось бы, что «амплитуда» c изменяется с частотой $\omega_k/2$, т. е. с частотой того же порядка, что и частота колебания резонатора. Наблюдать изменение амплитуды при таких условиях нельзя, и, следовательно, такое рассмотрение вопроса в случае нашего примера было бы формальным.

¹ К этому выводу приходит и Римский-Корсаков [16]. Надо заметить, что в своей работе он применяет интересный, так сказать, синтетический под-

Разобрав в деталях картину происходящих при этом явлений, мы смогли подметить ряд специфических черт анализа импульсов.

К анализу импульсов посредством резонаторов можно подойти несколько иначе. Прежде всего заметим, что установившаяся ко времени окончания импульса амплитуда резонатора является мерой накопленной резонатором колебательной энергии; энергия пропорциональна квадрату амплитуды. Из этого вытекает возможность видоизменения способа отсчета по анализатору, а именно: можно замкнуть колеблющийся резонатор на активное сопротивление, в котором выделится запасенная в резонаторе энергия, и тем или иным способом измерить эту энергию.

Но можно поступить и иначе — с самого начала измерять потребляемую в резонаторе энергию. Оказывается, что по окончании процесса эта энергия с известным приближением пропорциональна квадрату спектральной плотности импульса на частоте резонатора. Докажем это.

Возьмем сразу резонатор с затуханием, для которого мгновенное значение наблюдаемой величины (см. выше) есть

$$\begin{aligned} x(t) &= a_k \omega_k \int_0^t e^{-\alpha_k(t-u)} \sin \omega_k(t-u) f(u) du = \\ &= a_k \omega_k e^{-\alpha_k t} \left[\sin \omega_k t \int_0^t e^{\alpha_k u} f(u) \cos \omega_k u du - \right. \\ &\quad \left. - \cos \omega_k t \int_0^t e^{\alpha_k u} f(u) \sin \omega_k u du \right]. \end{aligned}$$

Составим выражение для энергии, затраченной ко времени

$$E(t) = \int_0^t x^2 dt.$$

Полная энергия, выделившаяся в каждом резонаторе по окончании как импульса, так и процессов в резонаторе, может быть записана в виде

$$\begin{aligned} E_k &= a_k^2 \omega_k^2 \int_0^{\infty} e^{-2\alpha_k t} \left[\sin \omega_k t \int_{-\infty}^{+\infty} f_\alpha(u) \cos \omega_k u du - \right. \\ &\quad \left. - \cos \omega_k t \int_{-\infty}^{+\infty} f_\alpha(u) \sin \omega_k u du \right]^2 dt. \end{aligned}$$

ход к проблеме, ставя сперва требования к идеальному анализатору и находя затем его физическую характеристику. Для рассматриваемого применения таким идеальным анализатором оказывается именно резонатор без затухания.

Здесь обозначено

$$f_{\alpha}(u) = e^{\alpha u} f(u).$$

Внутренние интегралы представляют собой спектральные компоненты A и B функции f_{α} . Так как эти величины не зависят от t , то имеем

$$\begin{aligned} E_k &= a_k^2 \omega_k^2 \left[a^2 \int_0^{\infty} e^{-2\alpha_k t} \sin^2 \omega_k t dt + b^2 \int_0^{\infty} e^{-2\alpha_k t} \cos^2 \omega_k t dt - \right. \\ &\quad \left. - ab \int_0^{\infty} e^{-2\alpha_k t} \sin 2\omega_k t dt \right] = \\ &= \frac{a_k^2 \omega_k^2}{4\alpha_k} \left[a^2 + b^2 + \frac{d^2/4}{1 + d^2/4} (b^2 - a^2) - \frac{d}{1 + d^2/4} ab \right]. \end{aligned}$$

При малом затухании, т. е. при $\alpha_k \ll \omega_k$ $\alpha \ll 1$, получаем приближенно (пренебрегая двумя последними членами в скобках)

$$E_k \approx \frac{a_k^2 \omega_k^2}{4\alpha_k} (a^2 + b^2) = b_k \Phi_{\alpha}^2 \approx b_k \Phi^2,$$

так как Φ_{α} при $d \ll 1$ мало отличается от Φ . Если теперь выбрать параметры резонаторов так, чтобы величина b_k не зависела от k , т. е. была бы для всех резонаторов одинакова, то получаем, что выделенная в каждом резонаторе энергия пропорциональна (с точностью, определяемой сделанными при выводе приближениями) квадрату спектральной плотности, соответствующей частоте данного резонатора. Это и требовалось доказать.

Можно придумать различные способы отсчета величины E_k . Можно, например, подключить к каждому резонатору квадратичный электродинамический баллистический прибор. Можно интегрировать не в механической части (т. е. пользуясь баллистическим отсчетом), а при помощи интегрирующей электрической цепочки. Можно интегрировать, пользуясь тепловой инерцией теплового прибора. Можно, наконец, представить себе индикатор следующего устройства: активное сопротивление каждого резонатора образует нагревательную обмотку, наложенную на шарик термометра (или введенную внутрь шарика). Отсчет производится по высоте столбиков в наборе термометров, которые, будучи расположены в ряд, дают наглядную картину спектра. Постоянной времени такого термометрического индикатора очень легко управлять путем изменения теплоемкости термометров и их термоизоляции.

Описанный способ анализа сводится по сути дела к измерению энергии сложного колебания, приходящейся на полосу частот, вырезаемую данным резонатором. Очевидно, можно произвести более грубое усреднение, применив полосный анализ энергии, для чего резонаторы следует заменить соответственными полосовыми фильтрами.

§ 27. Спектральное представление случайных процессов

До сих пор мы применяли спектральное представление к тем или иным определенным функциям времени. Теперь необходимо распространить спектральные представления на объекты существенно иной природы — на случайные процессы, играющие в технике и математике большую и все возрастающую роль.

Случайным процессом мы называем функцию непрерывно изменяющегося аргумента t , значения которой представляют собою случайные величины.

Из этого определения сразу видно, что если вычислить спектральную плотность случайного процесса $\xi(t)$ по обычной формуле

$$S(\omega) = \int_{-\infty}^{\infty} \xi(t) e^{-j\omega t} dt, \quad (27.1)$$

то полученная функция $S(\omega)$ будет случайной функцией. Мы получим то, что можно назвать спектром одной из возможных реализаций случайного процесса. В условиях же действительного наблюдения за течением некоторого случайного процесса мы можем получить лишь текущий спектр данной реализации, т. е. случайную функцию

$$S_T(\omega) = \int_0^T \xi(t) e^{-j\omega t} dt. \quad (27.2)$$

Желательно ввести такое определение спектра случайного процесса, которое привело бы нас к неслучайной функции частоты.

Случайный процесс определяется распределением вероятностей значений величины ξ для любого момента t . Для последующей теории имеют большое значение не только самые функции распределения вероятностей, но и первые два момента

$$M(\xi) = \int x\varphi(x) dx = \xi_0, \quad (27.3)$$

$$M[\xi(t_1)\xi(t_2)] = \iint x_1x_2\varphi(x_1, x_2) dx_1dx_2 = B(t_1, t_2), \quad (27.4)$$

где $\varphi(x)$ и $\varphi(x_1, x_2)$ — соответственно одно- и двумерные плотности вероятностей, в общем случае зависящие от времени. Однако в этом параграфе мы будем рассматривать только стационарные (в широком смысле) процессы, для которых обе величины (27.3) и (27.4) не зависят от времени; первая величина есть просто постоянная, называемая средним значением случайного процесса,

вторая величина зависит только от разности $|t_2 - t_1| = \tau$ и носит название функции корреляции.

Рассматриваемые здесь стационарные случайные процессы обладают эргодическим свойством, состоящим в том, что средние по множеству с вероятностью единицы равны средним по времени, взятым для любой реализации. Обозначая средние по времени чертой сверху, можем записать условие эргодичности в виде

$$\xi_0 = M(\xi) = \overline{\xi(t)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) dt, \quad (27.5)$$

$$\begin{aligned} B(\tau) &= M[\xi(t)\xi(t-\tau)] = \overline{\xi(t)\xi(t-\tau)} = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t)\xi(t-\tau) dt. \end{aligned} \quad (27.6)$$

Заметим сразу, что

$$B(0) = M[\xi^2(t)] = \overline{\xi^2(t)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi^2(t) dt. \quad (27.7)$$

Таким образом, дисперсия случайного процесса

$$D(\xi) = M(\xi)^2 - M^2(\xi) = \overline{\xi^2(t)} - \overline{\xi(t)}^2 = B(0) - \xi_0^2.$$

По физическому смыслу ξ_0 есть постоянная составляющая, $B(0)$ — полная мощность процесса, ξ_0^2 — мощность постоянной составляющей, дисперсия выражает мощность переменной составляющей. Впрочем чаще всего мы имеем дело с процессами, для которых $\xi_0 = 0$, так что $B(0)$ непосредственно выражает мощность процесса.

Дальнейшее основывается на теореме Хинчина, согласно которой функция корреляции $B(\tau)$ может быть представлена в виде

$$B(\tau) = \int_{-\infty}^{\infty} e^{j\omega\tau} dF(\omega), \quad (27.8)$$

где $F(\omega)$ — неубывающая ограниченная функция. Если функция $F(\omega)$ дифференцируема, то, обозначая

$$dF/d\omega = G(\omega)/2,$$

получим вместо (27.8)

$$B(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} G(\omega) e^{j\omega\tau} d\omega, \quad (27.9)$$

так что $G(\omega)$ есть обычное преобразование Фурье для $B(\tau)$

$$G(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} B(\tau) e^{-j\omega\tau} d\tau. \quad (27.10)$$

Функция $G(\omega)$ и есть то, что называется статистическим спектром случайного процесса, и формула (27.10) является основным определением этой функции. Следует заметить, что как $B(\tau)$, так и $G(\omega)$ — четные функции своих аргументов. Поэтому (27.9) и (27.10) могут быть записаны в вещественной форме двух косинус-трансформаций Фурье

$$B(\tau) = \int_0^{\infty} G(\omega) \cos \omega\tau d\omega, \quad (27.11)$$

$$G(\omega) = \frac{2}{\pi} \int_0^{\infty} B(\tau) \cos \omega\tau d\tau. \quad (27.12)$$

Физический смысл функции $G(\omega)$ легко выяснить, положив в (27.11) $\tau=0$. Мы получим

$$B(0) = P = \int_0^{\infty} G(\omega) d\omega, \quad (27.13)$$

где через P обозначена полная мощность процесса. Формула (27.13) показывает, что функция $G(\omega)$ выражает мощность процесса $\xi(t)$, приходящуюся на полосу частот $d\omega$. Иначе говоря, $G(\omega)$ представляет собою спектральную плотность мощности. На этом основании $G(\omega)$ называют часто энергетическим спектром процесса $\xi(t)$. Мощность, заключенную в конечной полосе между ω_1 и ω_2 , можно определить интегрированием $G(\omega)$ в соответствующих пределах

$$P_{12} = \int_{\omega_1}^{\omega_2} G(\omega) d\omega.$$

Статистический спектр можно выразить через текущий спектр реализации

$$S_T(\omega) = \int_0^T \xi(t) e^{-j\omega t} dt.$$

Составим выражение для энергии процесса $\xi(t)$, выделившейся за время T

$$E_T = \int_0^T \xi^2(t) dt = \frac{1}{\pi} \int_0^{\infty} |S_T(\omega)|^2 d\omega.$$

Это соотношение выражает теорему Рейли для конечного промежутка времени T . Среднюю за время T мощность получим, деля на T ,

$$P_T = \frac{E_T}{T} = \frac{1}{\pi T} \int_0^{\infty} |S_T(\omega)|^2 d\omega. \quad (27.14)$$

Эта величина зависит, вообще говоря, от T , но для стационарного процесса стремится с увеличением T к постоянному пределу, который и выражает мощность процесса

$$P = \lim_{T \rightarrow \infty} P_T = \frac{1}{\pi} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^{\infty} |S_T(\omega)|^2 d\omega. \quad (27.15)$$

Сопоставляя (27.13) и (27.15), мы видим, что статистический спектр связан с текущим спектром соотношением

$$G(\omega) = \frac{1}{\pi} \lim_{T \rightarrow \infty} \frac{|S_T(\omega)|^2}{T}. \quad (27.16)$$

Выведем еще некоторые формулы для вычисления $G(\omega)$. Возьмем формулу (27.10) и подставим в нее значение функции корреляции в соответствии с ее определением (см. (27.4)):

$$B(\tau) = M[\xi(t)\xi(t-\tau)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \varphi(x_1, x_2, \tau) dx_1 dx_2.$$

В полученном выражении

$$G(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{-j\omega\tau} d\tau \int_{-\infty}^{\infty} x_1 x_2 \varphi(x_1, x_2, \tau) dx_1 dx_2$$

выполним сперва интегрирование по τ . Обозначив

$$\psi(x_1, x_2, \omega) = \int_{-\infty}^{\infty} \varphi(x_1, x_2, \tau) e^{-j\omega\tau} d\tau,$$

можем записать

$$G(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \psi(x_1, x_2, \omega) dx_1 dx_2. \quad (27.17)$$

Эта формула представляет $G(\omega)$ как второй момент распределения некоторого случайного процесса; это распределение $\psi(x_1, x_2, \omega)$ есть преобразование Фурье для распределения $\varphi(x_1, x_2, \tau)$ исходного процесса.

Еще одно соотношение мы получим, исходя из определения текущего спектра случайного процесса

$$S_t(\omega) = \int_{-\infty}^t \xi(x) e^{-j\omega x} dx. \quad (27.18)$$

Введем случайную функцию $\xi_t(x)$, определяемую как

$$\xi_t(x) = \begin{cases} \xi(x) & \text{при } x < t, \\ 0 & \text{при } x > t. \end{cases}$$

С помощью этого определения можно переписать (27.18) в виде

$$S_t(\omega) = \int_{-\infty}^{\infty} \xi_t(x) e^{-j\omega x} dx. \quad (27.19)$$

Обратное преобразование для (27.19) будет

$$\xi_t(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_t(\omega) e^{j\omega x} d\omega. \quad (27.20)$$

Умножим обе части этого выражения на $\xi_t(x - \tau) dx$ и проинтегрируем в бесконечных пределах

$$\int_{-\infty}^{\infty} \xi_t(x) \xi_t(x - \tau) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \xi_t(x - \tau) dx \int_{-\infty}^{\infty} S_t(\omega) e^{j\omega x} d\omega. \quad (27.21)$$

Выполним в правой части сначала интегрирование по x . Интеграл

$$\int_{-\infty}^{\infty} \xi_t(x - \tau) e^{j\omega x} dx = e^{j\omega\tau} \int_{-\infty}^{\infty} \xi_t(x) e^{j\omega x} dx = e^{j\omega\tau} S_t^*(\omega)$$

выражается через комплексно-сопряженный текущий спектр. Если теперь вернуться к функции $\xi(x)$ в левой части (27.21) на основе определения $\xi_t(x)$, т. е. заменить соответственно верхний предел интегрирования, то получим

$$\int_{-\infty}^t \xi(x) \xi(x - \tau) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |S_t(\omega)|^2 e^{j\omega t} d\omega.$$

Продифференцируем обе части по t

$$\xi(t) \xi(t - \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\partial}{\partial t} |S_t(\omega)|^2 e^{j\omega t} d\omega.$$

Беря средние по множеству от обеих частей этого выражения, получаем

$$M [\xi(t), \xi(t - \tau)] = B(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} M \frac{\partial}{\partial t} |S_t(\omega)|^2 e^{j\omega\tau} d\omega. \quad (27.22)$$

Сличая (27.22) с (27.9), мы видим, что

$$G(\omega) = \frac{1}{\pi} M \left(\frac{\partial}{\partial t} |S_t(\omega)|^2 \right). \quad (27.23)$$

Таким образом, оказывается, что спектр $G(\omega)$ случайного процесса есть среднее по множеству от мгновенного спектра мощности в определении Пэйджа (см. (6.3)); аналогичное определение применяется Лемпердом [29].

Мы располагаем теперь целым рядом формул. Для того чтобы лучше в них ориентироваться, заметим, что $B(\tau)$ и $G(\omega)$ связаны между собой парой преобразований Фурье (27.9) и (27.10). Эти преобразования симметричны по своей структуре, из чего можно заключить, что всякой формуле для $B(\tau)$ должна соответствовать парная ей формула для $G(\omega)$. Эта симметрия становится очевидной, если собрать все приведенные выше формулы и сгруппировать их

$$B(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} G(\omega) e^{j\omega\tau} d\omega \quad (27.9)$$

$$= M [\xi(t) \xi(t - \tau)] \quad (27.4)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \varphi(x_1, x_2, \tau) dx_1 dx_2 \quad (27.4)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) \xi(t - \tau) dt \quad (27.6)$$

$$\pi G(\omega) = \int_{-\infty}^{\infty} B(\tau) e^{-j\omega\tau} d\tau \quad (27.10)$$

$$= M \left(\frac{\partial}{\partial t} |S_t(\omega)|^2 \right) \quad (27.23)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \psi(x_1, x_2, \omega) dx_1 dx_2 \quad (27.17)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} |S_T(\omega)|^2 \quad (27.16)$$

По поводу этих формул можно еще заметить следующее: последняя строка этой группировки содержит соотношения (27.6) и (27.16), получаемые на основании эргодичности стационарного процесса и представляющие собою средние по времени. Эти формулы могут послужить основой для построения аппаратуры для измерения $B(\tau)$ и $G(\omega)$ — коррелометров и анализаторов. Но эти формулы не могут непосредственно служить для вычисления $B(\tau)$ и $G(\omega)$, так как они представляют средние по времени для некоторой реализации. Следовательно, для выполнения вычислений по этим формулам нужно располагать аналитическим выра-

жением реализации в виде некоторой (неслучайной) функций, определенной на всем протяжении оси времени. В действительности случайный процесс задается математически не одной из своих реализаций, а тем или иным распределением; поэтому как функция корреляции $B(\varphi)$, так и спектр $G(\omega)$ вычисляются с неизбежным усреднением по множеству. Напротив, при эксперименте мы имеем чаще всего дело именно с одной реализацией случайного процесса. Техника вычисления спектров показана на примерах следующего параграфа.

Для спектров случайных процессов можно вывести ряд соотношений, аналогичных приведенным в § 4. Некоторые из этих соотношений собраны в табл. 3.

Таблица 3

$\eta(t)$	$B_\eta(\tau)$	$G_\eta(\omega)$
$\xi_1(t) + \xi_2(t)$	$B_1(\tau) + B_2(\tau)$	$G_1(\omega) + G_2(\omega)$
$\xi(at)$	$B(a\tau)$	$G\left(\frac{\omega}{a}\right)$
$\xi(t - t_0)$	$B(\tau)$	$G(\omega)$
$\xi(t) e^{-j\Omega t}$	$B(\tau) e^{j\Omega\tau}$	$G(\omega + \Omega)$
$\xi^{(n)}(t)$	$(-1)^n B^{(2n)}(\tau)$	$\omega^{2n} G(\omega)$
$\xi_1(t) \xi_2(t)$	$B_1(\tau) B_2(\tau)$	$\frac{1}{\pi} \int_{-\infty}^{\infty} G_1(\nu) G_2(\omega - \nu) d\nu$

Предполагается, что $M\xi=0$, что ξ_1 и ξ_2 — независимы, и что условие дифференцируемости выполняется там, где это требуется.

Для спектров случайных процессов и их функций корреляции имеет силу общего характера связь между шириной спектра и величиной, называемой интервалом корреляции, а именно

$$\Delta f \Delta \tau \geq \mu, \quad (27.24)$$

где μ — постоянная порядка единицы. Интервал корреляции $\Delta \tau$ выражает эффективную протяженность функции корреляции и может определяться по-разному. Общие соображения по поводу соотношения (27.24), справедливые для любых функций, связанных парой преобразований Фурье, были изложены в § 12. Мы укажем здесь на простые определения [30], приводящие (27.24) к тождеству

$$\Delta f \Delta \tau \equiv 1. \quad (27.25)$$

Запишем выражения для спектра и функции корреляции

$$G(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} B(\tau) \cos \omega \tau d\tau,$$

$$B(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} G(\omega) \cos \omega \tau d\omega.$$

Положим в первом выражении $\omega = 0$, а во втором $\tau = 0$; получим

$$G(0) = \frac{1}{\pi} \int_{-\infty}^{\infty} B(\tau) d\tau, \quad B(0) = \frac{1}{2} \int_{-\infty}^{\infty} G(\omega) d\omega. \quad (27.26)$$

Введем определения

$$\Delta\omega = 2\pi\Delta f = \frac{1}{G(0)} \int_{-\infty}^{\infty} G(\omega) d\omega, \quad \Delta\tau = \frac{1}{B(0)} \int_{-\infty}^{\infty} B(\tau) d\tau. \quad (27.27)$$

Отсюда на основании (27.26) сразу получаем (27.25). Смысл определений (27.27) состоит в том, что площади под кривыми $G(\omega)$ и $B(\tau)$ приравниваются площадям прямоугольников с основаниями $\Delta\omega$ и $\Delta\tau$ и высотами $G(0)$ и $B(0)$ соответственно.

Пусть, например,

$$B(\tau) = e^{-\alpha|\tau|}, \quad G(\omega) = \frac{2}{\pi} \cdot \frac{\alpha}{\alpha^2 + \omega^2}.$$

Имеем

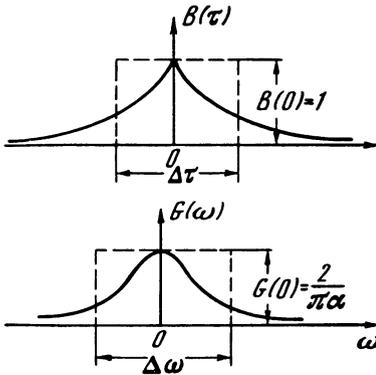
$$\int_{-\infty}^{\infty} B(\tau) d\tau = 2/\alpha, \quad B(0) = 1, \quad \Delta\tau = \alpha/2;$$

$$\int_{-\infty}^{\infty} G(\omega) d\omega = 2, \quad G(0) = 2/\pi\alpha, \quad \Delta\omega = 1/\pi\alpha, \quad \Delta f = 2/\alpha$$

(см. рис. 45).

Определения (27.27) имеют и недостатки. Первый из них состоит в том, что эти определения не применимы непосредственно в случае $G(0) = 0$ ($B(0)$ никогда не равно нулю). Между тем этот случай часто встречается. В частности, нас может интересовать процесс, спектр которого заключен в полосе частот от ω_1 до ω_2 . Можно, соответственно изменив определения, и в этом случае представить график спектра эквивалентным прямоугольником. Что касается функции корреляции, то она имеет в рассматриваемом случае колебательный характер, и здесь выясняется второй недостаток определений (27.27). Он заключается в том, что в то

время как функция $G(\tau)$ всегда неотрицательна, функция $B(\tau)$ может менять знак, а потому интеграл от функции в этом случае дает плохой критерий для оценки интервала корреляции. Эти недостатки исключаются в определениях, содержащих те или иные моменты квадратов функций (см. § 12) или их абсолютных значений, и т. п. Но, совершенствуя определения в этом направлении, мы теряем их простоту и лишаемся возможности применять тождество (27. 25).



Р и с. 45

§ 28. Спектры некоторых стационарных процессов

В качестве первого примера рассмотрим стационарный случайный процесс, представляемый случайной величиной, принимающей с равной вероятностью одно из двух значений $\pm a$, с переменной знака, происходящей в случайные моменты. На рис. 46 изо-

бражена одна из реализаций такого процесса. Для полной характеристики процесса нужно задать еще распределение для точек перехода через нуль, или, как коротко говорят, распределение нулей. Пусть задано такое распределение в виде вероятности $p(n, \tau)$ того, что на интервале τ окажется ровно n нулей.

Найдем сначала функцию корреляции, пользуясь основным ее определением (27. 4). Для этого нужно предварительно составить выражение для двумерной плотности вероятностей, т. е. плотности совместной вероятности случайным величинам $\xi = \xi(t)$ и $\xi_\tau = \xi(t + \tau)$ находиться соответственно в интервалах $(x_1, x_1 + dx_1)$ и $(x_2, x_2 + dx_2)$. Распределение в нашем случае дискретно и обладает следующими свойствами. Если на интервале τ имеется четное число нулей, то $\xi(t)$ и $\xi(t + \tau)$ имеют одинаковый знак, т. е. равны с одинаковой вероятностью либо $+a$, $+a$, либо $-a$, $-a$. Если же на интервале τ число нулей нечетно, то $\xi(t)$ и $\xi(t + \tau)$ имеют разные знаки и равны либо $+a$, $-a$, либо $-a$, $+a$. Все это можно записать аналитически при помощи дельта-функций следующим образом:

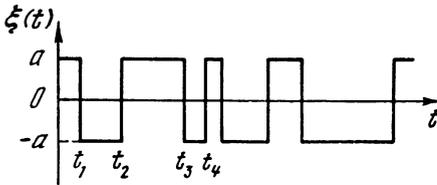
$$\begin{aligned} \varphi(x_1, x_2, \tau) = & \frac{1}{2} [\delta(x_1 - a, x_2 - a) + \\ & + \delta(x_1 + a, x_2 + a)] \sum_{k=0}^{\infty} p(2k, \tau) + \frac{1}{2} [\delta(x_1 - a, x_2 + a) + \\ & + \delta(x_1 + a, x_2 - a)] \sum_{k=0}^{\infty} p(2k + 1, \tau). \end{aligned}$$

В этом выражении

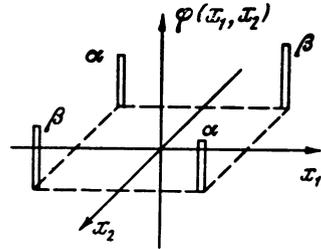
$$\sum_{k=0}^{\infty} p(2k, \tau) = p_{\text{Ч}}, \quad \sum_{k=0}^{\infty} p(2k+1, \tau) = p_{\text{Н}}$$

— соответственно вероятности появления на интервале τ четного и нечетного числа нулей.

Для наглядности распределение для рассматриваемого процесса изображено на рис. 47. Двумерные дельта-функции представлены условно столбиками конечной высоты (и объема $\alpha = p_{\text{Ч}}/2$ или $\beta = p_{\text{Н}}/2$).



Р и с. 46



Р и с. 47

Подставим найденное распределение в общее выражение для функции корреляции

$$B(\tau) = \frac{1}{2} p_{\text{Ч}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 [\delta(x_1 - a, x_2 - a) + \delta(x_1 + a, x_2 + a)] dx_1 dx_2 + \frac{1}{2} p_{\text{Н}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 [\delta(x_1 - a, x_2 + a) + \delta(x_1 + a, x_2 - a)] dx_1 dx_2.$$

При вычислении интегралов с дельта-функциями следует иметь в виду, что по определению этой функции

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = f(x_0).$$

Учитывая это соотношение, после двукратного интегрирования получим

$$B(\tau) = a^2(p_{\text{Ч}} - p_{\text{Н}}).$$

Рассмотрим частный случай пуассонова распределения

$$p(n, \tau) = \frac{(\mu\tau)^n}{n!} e^{-\mu\tau},$$

где μ — средняя частота нулей. В этом частном случае

$$B(\tau) = a^2 e^{-\mu\tau} \left(\sum_{k=0}^{\infty} \frac{(\mu\tau)^{2k}}{2k!} - \sum_{k=0}^{\infty} \frac{(\mu\tau)^{2k+1}}{(2k+1)!} \right) = a^2 e^{-\mu\tau} \sum_{k=0}^{\infty} (-1)^k \frac{(\mu\tau)^k}{k!}$$

или

$$B(\tau) = a^2 e^{-2\mu\tau}. \quad (28.1)$$

Отсюда находим и спектр (по (27. 12))

$$G(\omega) = \frac{2}{\pi} a^2 \int_0^{\infty} e^{-2\mu\tau} \cos \omega\tau d\tau = \frac{2}{\pi} a^2 \frac{2\mu}{4\mu^2 + \omega^2}. \quad (28.2)$$

Мы получили требуемый результат, опираясь на основные определения $\beta(\tau)$ и $G(\omega)$. Однако этот же результат можно получить гораздо быстрее, минуя составление и интегрирование двумерного распределения, точнее говоря, заменяя соответствующую часть вычисления устным рассуждением. Вот, к примеру, рассуждение Райса [35, стр. 118].

Рассмотрим произведение $\xi\xi_\tau$. Оно равно либо a^2 , либо $-a^2$ в зависимости от того, четно или нечетно число нулей на интервале τ . Среднее значение произведения $\xi\xi_\tau$ будет поэтому равно $a^2 p_{\text{ч}}$ плюс $(-a^2) p_{\text{н}}$, где $p_{\text{ч}}$ и $p_{\text{н}}$ — вероятности появления четного и нечетного числа нулей на интервале τ . Итак,

$$B(\tau) = a^2 (p_{\text{ч}} - p_{\text{н}})$$

(в подлиннике рассуждение лишь немного длиннее).

В качестве следующего примера возьмем стационарный случайный процесс, характеризующийся тем, что на интервале T_k между двумя случайно расположенными точками процесс имеет случайное, но неизменное значение ξ_k [7, стр. 372—376]. График возможной реализации такого процесса показан на рис. 48. Для нахождения функции корреляции потребуется рассуждение, отличающееся от предыдущего.

Произведение двух значений $\xi(t)$ и $\xi(t+\tau)$ будет различно в зависимости от того, лежат ли моменты t и $t+\tau$ в одном и том же интервале T_k или в различных. В первом случае

$$\xi(t) \xi(t+\tau) = \xi_k^2.$$

Во втором случае

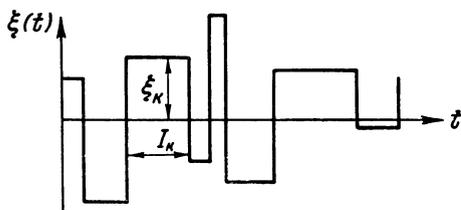
$$\xi(t) \xi(t+\tau) = \xi_t \xi_k$$

и для получения функции корреляции нужно сложить усредненные по множеству значения этих произведений, умноженные на соответствующие вероятности. Итак, мы должны прежде всего

найти вероятность того, что отрезок длиной τ укладывается полностью на интервале T_k . Эта вероятность равна вероятности того, что на протяжении отрезка длиной τ , случайным образом брошенного на ось времени, не окажется ни одной границы между интервалами T_k , т. е. ни одного нуля. Эта вероятность сразу определяется, если дано распределение $p(n, \tau)$: искомая вероятность есть $p(0, \tau)$. Так, в случае пуассонова распределения

$$p(0, \tau) = e^{-\mu\tau} = p(\tau).$$

Вероятность того, что отрезок τ не вмещается в интервал T_k (т. е. что t и $t + \tau$ находятся в разных интервалах, например,



Р и с. 48

T_k и T_i) равна, очевидно, $1 - p(\tau)$. Таким образом, для функции корреляции имеем

$$B(\tau) = M(\xi_k^2) p(\tau) + M(\xi_i \xi_k) [1 - p(\tau)].$$

Считая, что ξ_i и ξ_k независимы и одинаково распределены, что $M(\xi) = 0$ и что, следовательно, $M(\xi_i \xi_k) = 0$, получим

$$B(\tau) = M(\xi^2) p(\tau),$$

а для случая пуассонова распределения

$$B(\tau) = M(\xi^2) e^{-\mu\tau}. \quad (28.3)$$

Отсюда сейчас же находится и спектр

$$G(\omega) = \frac{2}{\pi} M(\xi^2) \frac{\mu}{\mu^2 + \omega^2}. \quad (28.4)$$

Интересно отметить различие формул (28.1) и (28.3). Казалось бы, ранее рассмотренный процесс рис. 46 является частным случаем процесса рис. 48 и что можно получить выражение для функции корреляции для процесса рис. 46 из формулы (28.3), положив в ней $M(\xi^2) = a^2$. Как видим, это не так. Разница объясняется тем, что для процесса рис. 46, принимающего попеременно значения $\pm a$, нельзя просто положить $M(\xi_i \xi_k) = 0$, как это сделано для более общего процесса рис. 48. Произведение $\xi_i \xi_k$ принимает попеременно значения $\pm a^2$, так что имеется зависимость от расстояния между интервалами (т. е. от разности $k-i$), т. е. в конечном счете от τ . Эту зависимость мы и учитываем, когда бе-

рем вероятности четных и нечетных чисел нулей, что и приводит к убыванию функции корреляции по закону $e^{-2\mu\tau}$, а не $e^{-\mu\tau}$.

В предыдущих примерах спектр вычислялся как преобразование Фурье для функции корреляции. Покажем теперь ход вычисления, когда за основу берется формула

$$G(\omega) = \frac{1}{\pi} \lim_{T \rightarrow \infty} \frac{1}{T} |S_T(\omega)|^2.$$

В качестве примера возьмем процесс $\pm a$ с переменной знака в случайные моменты (рис. 45) [13, стр. 48—49; 23, стр. 194—196]. Выражение для реализации можно записать в виде

$$\xi(t) = a \sum_{k=-\infty}^{\infty} (-1)^k \sigma(t - t_k).$$

Для текущего спектра будем иметь

$$S_T(\omega) = a \int_0^T \left[\sum_{k=0}^{\infty} (-1)^k \sigma(t - t_k) \right] e^{-j\omega t} dt.$$

Поменяем порядок операций суммирования и интегрирования и заодно заменим пределы, положив $T = nT_{cp}$, где $T_{cp} = 1/\rho$ — средний интервал между соседними нулями. Тогда получим

$$S_T(\omega) = a \sum_{k=1}^n (-1)^k \int_{-\infty}^{\infty} \sigma(t - t_k) e^{-j\omega t} dt.$$

Интеграл выражает спектр запаздывающей единичной функции (см. § 11), так что

$$S_T(\omega) = \frac{a^2}{j\omega} \sum_{k=1}^n (-1)^k e^{-j\omega t_k}.$$

Для получения квадрата модуля умножим на сопряженную величину

$$|S_T(\omega)|^2 = S_T(\omega) S_T^*(\omega) = \frac{a^2}{\omega^2} \sum_{k=1}^n \sum_{l=1}^n (-1)^{k+l} e^{-j\omega(t_k - t_l)}.$$

Итак, выражение для спектра случайного процесса имеет вид

$$G(\omega) = \frac{a^2 \mu}{\pi \omega^2} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \sum_{l=1}^n (-1)^{k+l} e^{-j\omega(t_k - t_l)}.$$

Введем $m = k - l$ и перепишем двойную сумму в виде

$$\frac{1}{n} \sum_{k=1}^n \sum_{l=1}^n (-1)^{k+l} e^{-j\omega(t_k - t_l)} = \frac{1}{n} \sum_{m=-(n-1)}^{n-1} (n - m) (-1)^m e^{-j\omega(t_{l+m} - t_l)}.$$

(Дело в том, что всего имеется n^2 слагаемых. Из них для n слагаемых $m=0$; для $n-1$ слагаемых $m=1$ и для такого же числа $m=-1$; для $n-2$ слагаемых $m=2$ и для $n-2$ слагаемых $m=-2$ и так далее.)

Рассмотрим отдельно экспоненциальный множитель. Его можно преобразовать следующим образом:

$$e^{-j\omega(t_{l+m} - t_l)} = e^{-j\omega(t_{l+m} - t_{l+m-1})} e^{-j\omega(t_{l+m-1} - t_{l+m-2})} \dots \\ \dots e^{-j\omega(t_{l+1} - t_l)} = e^{-j\omega T_1} e^{-j\omega T_2} \dots e^{-j\omega T_m} = \prod_{i=1}^m e^{-j\omega T_i},$$

где T_i — последовательность интервалов между нулями на промежутке $t_k - t_l = t_{l+m} - t_l$.

Внесем теперь множитель $1/n$ под знак суммы, перейдем к пределу и запишем отдельно суммы для положительных и отрицательных m . Получим

$$G(\omega) = \frac{a^2 \mu}{\pi \omega^2} \left[1 + \sum_{m=1}^{\infty} (-1)^m \prod_{i=1}^m e^{-j\omega T_i} + \sum_{m=1}^{\infty} (-1)^m \prod_{i=1}^m e^{+j\omega T_i} \right].$$

Для завершения вычисления мы должны были бы знать последовательность T_i , определяющую данную реализацию. Эта последовательность нам не дана; дана плотность распределения $\varphi(T)$ интервалов T_i . Поэтому на данном этапе мы прибегаем к усреднению по множеству.

Задача состоит в нахождении среднего значения суммы

$$\sum_{m=1}^{\infty} (-1)^m \prod_{i=1}^m e^{-j\omega T_i}.$$

Предположим, что интервалы T_i статистически независимы. Тогда независимы и величины $e^{-j\omega T_i}$ и среднее произведение равно произведению средних. Мы можем записать

$$M \left[\sum_{m=1}^{\infty} (-1)^m \prod_{i=1}^m e^{-j\omega T_i} \right] = \sum_{m=1}^{\infty} (-1)^m \prod_{i=1}^m M(e^{-j\omega T_i}).$$

Предполагается, что все интервалы T_i имеют одинаковое распределение, выражаемое плотностью $\varphi(T)$. Тогда

$$M(e^{-j\omega T_i}) = \int_0^{\infty} \varphi(T) e^{-j\omega T} dT = \chi(\omega),$$

т. е. среднее значение $e^{-j\omega T}$ есть преобразование Фурье для плотности вероятностей $\varphi(T)$. Функция $\chi(\omega)$ называется характеристической функцией. Теперь мы можем записать

$$\sum_{m=1}^{\infty} (-1)^m \prod_{i=1}^m M(e^{-j\omega T_i}) = \sum_{m=1}^{\infty} (-1)^m \chi^m(\omega) = -\frac{\chi(\omega)}{1 + \chi(\omega)}.$$

Вторая сумма дает сопряженную величину. Складывая комплексно-сопряженные величины, получаем удвоенную действительную часть, так что окончательно

$$G(\omega) = \frac{a^{2\mu}}{\pi\omega^2} \left[1 - 2 \operatorname{Re} \left(\frac{\chi(\omega)}{1 + \chi(\omega)} \right) \right].$$

Определим спектр для частного случая пуассонова распределения нулей

$$\varphi(T) = \mu e^{-\mu T}.$$

Найдем характеристическую функцию

$$\chi(\omega) = \int_0^{\infty} \varphi(T) e^{-j\omega T} dT = \mu \int_0^{\infty} e^{-(\mu + j\omega)T} dt = \frac{\mu}{\mu + j\omega}.$$

Далее

$$\frac{\chi(\omega)}{1 + \chi(\omega)} = \frac{\mu}{2\mu + j\omega},$$

и в результате

$$G(\omega) = \frac{2}{\pi} a^2 \frac{2\mu}{4\mu^2 + \omega^2},$$

что совпадает с (28.2).

§ 29. Понятие спектра в применении к нестационарным процессам

В предыдущем мы занимались только эргодическими стационарными процессами. Но на практике приходится иметь дело и с процессами нестационарными, причем существенно, что такие процессы встречаются достаточно часто и имеют большое значение. К числу нестационарных процессов относятся многие виды сигналов связи. Так, например, обычное модулированное колебание

$$\eta(t) = \xi \cos \omega t,$$

где $\xi(t)$ — стационарный случайный процесс, является процессом нестационарным, что сразу видно хотя бы из того, что одномерное

распределение пропорционально $\cos \omega t$, т. е. зависит от времени. При таком положении необходимо поставить вопрос о применимости спектральных представлений к некоторому интересующему нас классу нестационарных процессов. Дело сводится, конечно, к некоторому пересмотру и расширению определений, которыми мы до сих пор пользовались.

Как мы видели (см. § 27), вычисление спектра случайного процесса непременно связано с операцией усреднения. Для эргодических стационарных процессов средние по множеству и по времени с вероятностью единицы равны друг другу. Поэтому безразлично, каким способом усреднения мы воспользуемся. Выбор усреднения определяется тем, как нам задан случайный процесс. В условиях эксперимента мы имеем дело обычно с отдельной реализацией случайного процесса; в этих условиях мы прибегаем к усреднению по времени. При теоретическом же исследовании случайный процесс задается распределением вероятностей, и в этом случае мы пользуемся усреднением по множеству.

Представим случайный процесс как функцию $\xi(\vartheta, t)$ двух переменных: случайного параметра ϑ и непрерывно текущего времени t . Зафиксировав некоторый момент $t=t_0$, мы получим случайную величину $\xi(\vartheta, t_0)$, распределение которой в общем случае зависит от t_0 . Закрепив же некоторое значение случайного параметра ϑ , мы получаем одну из возможных реализаций случайного процесса, представляющую собой уже неслучайную функцию времени.

Операция усреднения есть в сущности операция интегрирования с тем или иным весом. Так, среднее по времени находится из выражения

$$\overline{\xi(t)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \xi(t) dt.$$

Среднее по множеству вычисляется, как

$$M(\xi) = \int_{-\infty}^{\infty} x \varphi(x) dx.$$

Но если мы имеем дело с функцией двух переменных, то, в общем случае, усреднение по одной переменной даст величину, зависящую от второй переменной, и наоборот. Применительно к случайным процессам это значит, что среднее по множеству, в общем случае, зависит от времени, а средние по времени образуют случайное множество. Эргодический стационарный процесс тем и замечателен, что для него средние по времени и по множеству равны друг другу, а из этого непосредственно следует, что они не зависят от второй переменной (так как функции разных

аргументов могут быть равны друг другу только в том случае, когда эти функции представляют собой постоянные, т. е. величины, не зависящие от аргументов).

Таким образом, среднее для нестационарного процесса, рассматриваемого как функция двух переменных, может определяться как результат двукратного усреднения — по множеству и по времени. Порядок усреднения безразличен, так как дело сводится всего лишь к изменению порядка интегрирования. Так, произведя для нестационарного процесса усреднение по множеству, мы получим величину, зависящую от времени; для устранения этой зависимости следует произвести повторное усреднение, на этот раз по времени. Усредняя сперва по времени, получим случайное множество; требуется дополнительная операция усреднения по этому множеству.

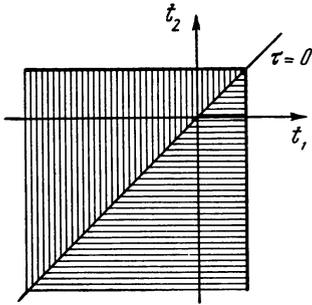


Рис. 49

Для полноты картины следовало бы разобрать вопрос о том, для какого класса нестационарных процессов среднее в указанном смысле, во-первых, существует, а во-вторых, не лишается

того физического смысла, который мы обычно усредняемой величине приписываем. Не предпринимая здесь такого исследования, заметим лишь, что в качестве примеров в дальнейшем рассматриваются только такие нестационарные процессы, для которых функция корреляции и спектр, вычисленные путем двойного усреднения, сохраняют свой обычный смысл¹. Во избежание недоразумений мы будем называть функцию корреляции и спектр нестационарных процессов, полученные двукратным усреднением, средней функцией корреляции и средним спектром.

Покажем теперь, что средние функция корреляции и спектр по-прежнему связаны между собой парой преобразования Фурье. В качестве исходных возьмем следующие два определения:

$$B(\tau, t) = M |\xi(t) \xi(t - \tau)|, \quad G(\omega) = \frac{1}{\pi} M \left(\frac{\partial}{\partial t} |S_t(\omega)|^2 \right).$$

Запишем выражение для текущего спектра

$$S_t(\omega) = \int_{-\infty}^t \xi(t_1) e^{-j\omega t_1} dt_1$$

¹ Бунимович [4] применяет двойное усреднение, имея в виду нестационарные процессы вида $\eta(t) = \psi[\xi(t), f(t)]$, где ψ — знак некоторой функции, $\xi(t)$ — случайный процесс, $f(t)$ — периодическая или почти периодическая функция. Все примеры следующего параграфа принадлежат этому классу.

и умножим на сопряженную величину

$$S_t(\omega) S_t^*(\omega) = |S_t(\omega)|^2 = \int_{-\infty}^t \int_{-\infty}^t \xi(t_1) \xi(t_2) e^{-j\omega(t_1-t_2)} dt_1 dt_2.$$

Введем $\tau = t_1 - t_2$ и разобьем область интегрирования на две, как показано на рис. 49. Это даст

$$|S_t(\omega)|^2 = \int_{-\infty}^t dt_1 \int_{-\infty}^0 \xi(t_1) \xi(t_1 - \tau) e^{-j\omega\tau} d\tau + \\ + \int_{-\infty}^t dt_2 \int_0^{\infty} \xi(t_2) \xi(t_2 + \tau) e^{-j\omega\tau} d\tau.$$

Усредним обе части этого равенства по множеству, замечая, что

$$M[\xi(t_1) \xi(t_1 - \tau)] = B(t_1, t_1 - \tau) = B(\tau, t_1) = M[\xi(t_1) \xi(t_2)] = \\ = M[\xi(t_2) \xi(t_2 + \tau)].$$

Мы получим

$$M|S_t(\omega)|^2 = 2 \int_{-\infty}^t dt_1 \int_0^{\infty} B(\tau, t_1) \cos \omega\tau d\tau.$$

Дифференцируя по t и подставляя результат в исходное определение спектра, находим

$$G(\omega, t) = \frac{2}{\pi} \int_0^{\infty} B(\tau, t) \cos \omega\tau d\tau.$$

Отсюда усреднением по времени получаем связь между средним спектром и средней функцией корреляции

$$\overline{G(\omega, t)} = \frac{2}{\pi} \int_0^{\infty} \overline{B(\tau, t)} \cos \omega\tau d\tau.$$

Справедливо и обратное преобразование

$$\overline{B(\tau, t)} = \int_0^{\infty} \overline{G(\omega, t)} \cos \omega\tau d\omega.$$

(сравним с формулами (27. 11) и (27. 12)).

Вывод из всего сказанного состоит в том, что мы можем воспользоваться для вычисления функции корреляции и спектров нестационарных процессов формулами, приведенными в § 27, добавив в них второе усреднение. С точки зрения вычислений существенно, что это второе усреднение может производиться на любом этапе вычисления, как показано на примерах следующего параграфа.

§ 30. Спектры некоторых нестационарных процессов

Отличительным признаком нестационарного (в широком смысле) процесса является то, что его одно- и двумерные распределения или соответственно первый и второй моменты зависят от времени. Иначе говоря, нестационарность выражается в том, что среднее значение, либо дисперсия, либо функция корреляции являются функциями времени. Мы будем применять этот критерий в рассматриваемых ниже примерах.

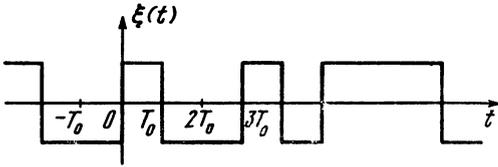


Рис. 50

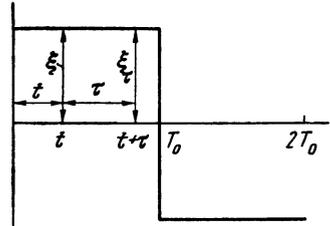


Рис. 51 ▶

В качестве первого примера рассмотрим случайный процесс, характеризующийся тем, что случайная величина принимает одно из двух равновероятных значений $\pm a$; перемена знака может происходить только в фиксированные моменты

$$t = t_k = kT_0.$$

К описанному случайному процессу сводится, например, телеграфный сигнал при передаче равномерным кодом. На рис. 50 изображена одна из реализаций.

Составим выражение для распределения вероятностей. Одномерное распределение состоит из двух вероятностей $p(a) = p(-a) = 1/2$. Выражение для одномерной плотности вероятностей можно записать с помощью дельта-функций

$$\varphi(x) = \frac{1}{2} [\delta(x - a) + \delta(x + a)].$$

Одномерное распределение, как видим, от времени не зависит. Теперь рассмотрим распределение для двумерной величины ξ , ξ_τ , представляющей собой совокупность двух значений случайного процесса, отстоящих друг от друга на τ . Двумерное распределение, оказывается, зависит от времени. Именно, если $t + \tau < T_0$, то моменты t и $t + \tau$ находятся внутри одного и того же интервала T_0 и, следовательно, ξ и ξ_τ имеют одинаковый знак (рис. 51). Таким образом,

$$\varphi_1(x_1, x_2) = \frac{1}{2} [\delta(x_1 - a, x_2 - a) + \delta(x_1 + a, x_2 + a)] \quad t + \tau < T_0.$$

Если же $t + \tau > T_0$, то ξ и ξ_τ с равной вероятностью могут иметь любые знаки и, стало быть,

$$\varphi_2(x_1, x_2) = \frac{1}{4} [\delta(x_1 - a, x_2 - a) + \delta(x_1 + a, x_2 + a) + \delta(x_1 - a, x_2 + a) + \delta(x_1 + a, x_2 - a)] \quad t + \tau > T_0.$$

Таким образом, двумерная плотность вероятностей имеет на разных интервалах разные выражения, а это и значит, что она зависит не только от τ , но и от t . Итак, рассматриваемый процесс нестационарен.

Как указывалось выше, усреднение может производиться на любом этапе вычисления. Воспользовавшись этим, усредним сразу по времени самое двумерное распределение. Мы имеем

$$\varphi(x_1, x_2, \tau, t) = \begin{cases} \varphi_1(x_1, x_2) & \text{при } t + \tau < T_0, \\ \varphi_2(x_1, x_2) & \text{при } t + \tau > T_0. \end{cases}$$

Перепишем неравенство $t + \tau < T_0$ в виде $0 < t < T_0 - \tau$. Усредняя распределение по интервалу T_0 , получим

$$\overline{\varphi(x_1, x_2, \tau, t)} = \frac{1}{T_0} \int_0^{T_0 - \tau} \varphi_1(x_1, x_2) dt = (1 - \tau/T_0) \varphi_1(x_1, x_2)$$

для τ , находящегося в интервале $0, T_0$. При $t + \tau > T_0$, т. е. при $T_0 - \tau < t < \infty$ получим, усредняя по интервалу $T \rightarrow \infty$,

$$\overline{\varphi(x_1, x_2, \tau, t)} = \varphi_2(x_1, x_2)$$

для τ в интервале T_0, ∞ . Итак, для усредненного распределения можем записать

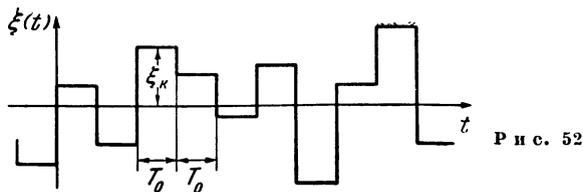
$$\begin{aligned} \overline{\varphi(x_1, x_2, \tau)} = \overline{\varphi(x_1, x_2, \tau, t)} = & \frac{1}{2} [\delta(x_1 - a, x_2 - a) + \delta(x_1 + a, x_2 + a)] \times \\ & \times (1 - \tau/T_0) \quad \text{при } 0 < \tau < T_0, \\ & \frac{1}{4} [\delta(x_1 - a, x_2 - a) + \delta(x_1 + a, x_2 + a) + \\ & + \delta(x_1 - a, x_2 + a) + \delta(x_1 + a, x_2 - a)] \\ & \text{при } T_0 < \tau < \infty. \end{aligned}$$

Подставляя усредненное распределение в формулу

$$B(\tau, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \overline{\varphi(x_1, x_2, \tau, t)} dx_1 dx_2,$$

найдем среднюю функцию корреляции

$$B(\tau, t) = \begin{cases} a^2 (1 - \tau/T_0) & 0 < \tau < T_0, \\ 0 & T_0 < \tau < \infty, \end{cases}$$



Р и с. 52

а затем и средний спектр

$$\begin{aligned} \overline{G(\omega, t)} &= \frac{2}{\pi} \int_0^{\infty} \overline{B(\tau, t)} \cos \omega \tau d\tau = \frac{2}{\pi} a^2 \int_0^{T_0} \left(1 - \frac{\tau}{T_0}\right) \sin \omega \tau d\tau = \\ &= \frac{1}{\pi} a^2 T_0 \left(\frac{\sin \omega T_0/2}{\omega T_0/2}\right)^2. \end{aligned} \quad (30.1)$$

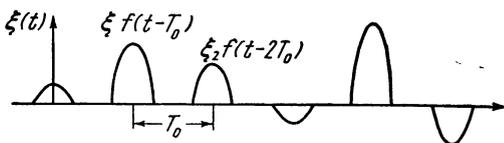
Для второго примера возьмем процесс несколько более общего вида: ось времени по-прежнему разбита на равные интервалы T_0 ; в каждом интервале случайная величина принимает независимые значения ξ_k , определяемые плотностью вероятностей $\varphi(x)$ (рис. 52). Для двумерной плотности вероятностей имеем следующие выражения:

$$\varphi(x_1, x_2, \tau, t) = \begin{cases} \varphi(x_1) \delta(x_1 - x_2) & t + \tau < T_0, \\ \varphi(x_1) \varphi(x_2) & t + \tau > T_0. \end{cases}$$

Если $M(\xi) = 0$, то после усреднения находим для функции корреляции и спектра

$$\begin{aligned} \overline{B(\tau, t)} &= \begin{cases} M(\xi^2) \left(1 - \frac{\tau}{T_0}\right) & 0 < \tau < T_0, \\ 0 & T_0 < \tau < \infty; \end{cases} \\ \overline{G(\omega, t)} &= \frac{1}{\pi} T_0 M(\xi^2) \left(\frac{\sin \omega T_0/2}{\omega T_0/2}\right)^2. \end{aligned} \quad (30.2)$$

Рассмотренный процесс можно представить, как последовательность прямоугольных импульсов со скважностью единица,



Р и с. 53

модулированных по амплитуде. Нужно заметить, что спектры различных видов импульсной модуляции удобнее находить, исходя из текущего спектра, подобно тому, как это было показано в § 28.

Рассмотрим общий случай амплитудно-импульсной модуляции. Процесс представляет собой последовательность импульсов произвольной формы $f(t)$, модулированных по амплитуде, т. е. умно-

женных на случайный множитель ξ_k (рис. 53). Такой процесс может быть записан в виде

$$\eta(t) = \sum_{k=-\infty}^{\infty} \xi_k f(t - kT_0).$$

Составим выражение для текущего спектра

$$S_T(\omega) = \int_0^T \left[\sum_{k=-\infty}^{\infty} \xi_k f(t - kT_0) \right] e^{-j\omega t} dt.$$

Переменим порядок действий с одновременной заменой пределов

$$S_T(\omega) = \sum_{k=1}^n \xi_k \int_{-\infty}^{\infty} f(t - kT_0) e^{-j\omega t} dt = A(\omega) \sum_{k=1}^n \xi_k e^{-jk\omega T_0},$$

где $A(\omega) = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt$ — спектр одиночного импульса, представляемого функцией $f(t)$.

Умножим $S_T(\omega)$ на сопряженную величину

$$S_T(\omega) S_T^*(\omega) = |S_T(\omega)|^2 = |A(\omega)|^2 \sum_{k=1}^n \sum_{l=1}^n \xi_k \xi_l e^{-j(k-l)\omega T_0}.$$

Остается разделить на $T = nT_0$, перейти к пределу при $T \rightarrow \infty$ и взять среднее по множеству

$$\overline{G(\omega, t)} = \frac{1}{\pi T_0} M \left\{ |A(\omega)|^2 \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \sum_{l=1}^n \xi_k \xi_l e^{-j(k-l)\omega T_0} \right\}.$$

Мы будем иметь n членов, для которых $k=l$; после суммирования и деления на n эти члены дадут в среднем $M(\xi^2)$. Среднее значение произведения $\xi_k \xi_l$ (в предположении, что ξ_k и ξ_l независимы) равно произведению средних, т. е. $(M\xi)^2$; эта величина выходит из-под знака суммы. Вводя $m=k-l$, после преобразований, аналогичных проделанным в § 28, получим

$$\overline{G(\omega, t)} = \frac{1}{\pi T_0} |A(\omega)|^2 \left\{ M(\xi^2) - (M\xi)^2 - (M\xi)^2 \sum_{-\infty}^{\infty} e^{-jm\omega T_0} \right\}$$

(здесь слагаемое $-(M\xi)^2$ добавлено для того, чтобы сумма включала это слагаемое, соответствующее значению $m=0$). Сумма в правой части представляет в виде ряда Фурье некоторую периодическую функцию частоты. Легко видеть, что эта функция есть не что иное, как последовательность дельта-функций вида

$$F(\omega) = \omega_0 \sum_{m=1}^{\infty} \delta(\omega - m\omega_0)$$

в чем можно убедиться, разложив функцию $F(\omega)$ в ряд Фурье. Функция $F(\omega)$ представляет собой линейчатый гармонический спектр; $\omega_0 = 2\pi/T_0$ — основная частота спектра, равная частоте следования импульсов. Таким образом, $G(\omega)$ содержит сплошную и дискретную части; последняя отсутствует при $M\xi = 0$. Окончательное выражение для спектра записывается в виде

$$\overline{G(\omega, t)} = \frac{1}{\pi T_0} |A(\omega)|^2 [M(\xi^2) - (M\xi)^2 + (M\xi)^2 F(\omega)]. \quad (30.3)$$

Применим эту формулу к процессу рис. 52, представляющему собой последовательность прямоугольных импульсов длительностью T_0 . В этом случае мы имеем

$$A(\omega) = T_0 \frac{\sin \omega T_0/2}{\omega T_0/2}, \quad M(\xi) = 0$$

и, следовательно,

$$\overline{G(\omega, t)} = \frac{1}{\pi} T_0 M(\xi^2) \left(\frac{\sin \omega T_0/2}{\omega T_0/2} \right)^2$$

— результат, полученный ранее другим способом.

Разберем теперь случай импульсов, модулированных по фазе. Процесс представляет собой последовательность импульсов одинаковой формы и величины, имеющих случайные смещения ϵ_k во времени относительно фиксированных равноотстоящих моментов kT_0 ; иначе говоря, период следования импульсов T_0 получает случайные приращения ϵ_k . Величины ϵ_k предполагаются независимыми; процесс задается одномерным распределением $\varphi(\epsilon)$. На рис. 54 к примеру изображены прямоугольные импульсы.

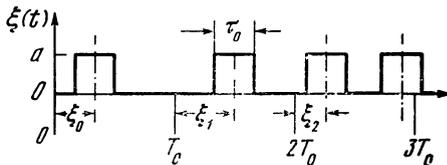


Рис. 54

Мы не будем составлять выражения для двумерного распределения; заметим лишь, что произведение $\xi \xi_\tau$ равно a^2 при одновременном выполнении двух условий

$$kT_0 + \epsilon_k - \frac{1}{2} \tau_0 < t < kT_0 + \epsilon_k + \frac{1}{2} \tau_0,$$

$$lT_0 + \epsilon_l - \frac{1}{2} \tau_0 < t + \tau < lT_0 + \epsilon_l + \frac{1}{2} \tau_0.$$

Если хотя бы одно из этих неравенств не соблюдается, то $\xi \xi_\tau = 0$. Отсюда видно, что функция корреляции зависит не только от τ , но и от t , так что рассматриваемый процесс нестационарен.

Мы будем вычислять спектр непосредственно для импульсов произвольной формы. Процесс можно записать в виде

$$\xi(t) = \sum_{k=-\infty}^{\infty} f(t - kT_0 - \varepsilon_k),$$

где $f(t)$ — по-прежнему функция, представляющая отдельный импульс. Текущий спектр выражается как

$$S_T(\omega) = \int_0^T \left[\sum_{k=-\infty}^{\infty} f(t - kT_0 - \varepsilon_k) \right] e^{-j\omega t} dt$$

или после перемены порядка действий и замены пределов

$$S_T(\omega) = \sum_{k=1}^n \int_{-\infty}^{\infty} f(t - kT_0 - \varepsilon_k) e^{-j\omega t} dt = A(\omega) \sum_{k=1}^n e^{-j\omega(kT_0 + \varepsilon_k)},$$

где $A(\omega)$ — спектр одиночного импульса. Далее,

$$|S_T(\omega)|^2 = |A(\omega)|^2 \sum_{k=1}^n \sum_{l=1}^n e^{-j\omega(k-l)T_0} e^{-j\omega\varepsilon_k} e^{j\omega\varepsilon_l}.$$

Усреднение по множеству дает

$$M |S_T(\omega)|^2 = |A(\omega)|^2 |\chi(\omega)|^2 \sum_{k=1}^n \sum_{l=1}^n e^{-j\omega(k-l)T_0},$$

где

$$\chi(\omega) = M(e^{-j\omega\varepsilon}) = \int_{-\infty}^{\infty} \varphi(\varepsilon) e^{-j\omega\varepsilon} d\varepsilon$$

— характеристическая функция распределения ε . Преобразуя двойную сумму точно так же, как это делалось раньше, получим

$$\overline{G(\omega, t)} = \frac{1}{\pi T_0} |A(\omega)|^2 \left(1 - |\chi(\omega)|^2 + |\chi(\omega)|^2 \sum_{m=-\infty}^{\infty} e^{-j\omega m T_0} \right)$$

или

$$\overline{G(\omega, t)} = \frac{1}{\pi T_0} |A(\omega)|^2 [1 - |\chi(\omega)|^2 + |\chi(\omega)|^2 F(\omega)], \quad (30.4)$$

где

$$F(\omega) = \omega_0 \sum_{m=1}^{\infty} \delta(\omega - m\omega_0).$$

Перейдем теперь к исследованию синусоидальных колебаний, модулированных случайным процессом по амплитуде или по фазе.

Начнем с амплитудной модуляции, как более простой. Случайный процесс, спектр которого мы должны найти, записывается в виде

$$\eta(t) = \xi(t) \cos \omega_0 t,$$

где $\xi(t)$ — стационарный процесс. Составим произведение

$$\eta(t)\eta(t+\tau) = \frac{1}{2}\xi(t)\xi(t+\tau)[\cos\omega_0(2t+\tau) + \cos\omega_0\tau].$$

Усреднив это произведение по множеству, получим

$$B_\eta(\tau, t) = \frac{1}{2}B_\xi(\tau)[\cos\omega_0(2t+\tau) + \cos\omega_0\tau].$$

Сделаем теперь усреднение по времени; первый член в квадратных скобках выпадет, и мы будем иметь

$$\overline{B_\eta(\tau, t)} = \frac{1}{2}B_\xi(\tau)\cos\omega_0\tau.$$

Отсюда находим средний спектр

$$\begin{aligned} \overline{G_\eta(\omega, t)} &= \frac{2}{\pi} \int_0^\infty \overline{B_\eta(\tau, t)} \cos\omega\tau d\tau = \frac{1}{\pi} \int_0^\infty B_\xi(\tau) \cos\omega_0\tau \cos\omega\tau d\tau = \\ &= \frac{1}{2\pi} \int_0^\infty B_\xi(\tau) \cos(\omega_0 + \omega)\tau d\tau + \frac{1}{2\pi} \int_0^\infty B_\xi(\tau) \cos(\omega_0 - \omega)\tau d\tau. \end{aligned}$$

Итак,

$$\overline{G_\eta(\omega, t)} = \frac{1}{4}[G_\xi(\omega_0 + \omega) + G_\xi(\omega_0 - \omega)], \quad (30.5)$$

т. е. спектр модулированного колебания состоит из двух боковых полос, повторяющих спектр модулирующего процесса и расположенных симметрично относительно несущей частоты ω_0 . Таким образом, обычное строение спектра амплитудно-модулированного колебания сохраняется и в том случае, когда модулирующая функция есть случайный процесс.

Обратимся к фазовой модуляции. Здесь дело обстоит значительно сложнее [6].

Задача состоит в нахождении спектра случайного процесса вида

$$\xi(t) = \cos\theta_1(t),$$

где $\theta_1(t)$ — случайный процесс. Но так как по смыслу фаза (аргумент косинуса) должна в среднем расти пропорционально времени, то мы будем рассматривать частный вид фазовой модуляции, выражаемый соотношением

$$\xi(t) = \cos[\omega_0 t + \theta(t)]. \quad (30.6)$$

Здесь $\theta(t)$ уже может быть стационарным процессом, что мы и будем предполагать.

Прежде всего нужно показать, что процесс (30.6) нестационарен. Для этого составим выражение для плотности вероятностей двумерной случайной величины ξ , ξ_τ . Мы имеем

$$\begin{aligned}\xi &= \xi(t) = \cos [\omega_0 t + \theta(t)], \\ \xi_\tau &= \xi(t + \tau) = \cos [\omega_0(t + \tau) + \theta(t + \tau)].\end{aligned}\quad (30.7)$$

Плотность вероятностей для двумерной величины θ , θ_τ задана и равна $\varphi_\theta(\vartheta_1, \vartheta_2, \tau)$. Искомая плотность $\varphi_\xi(x_1, x_2, \tau)$ должна быть выражена через φ_θ . Дело сводится к замене переменных по известной формуле

$$\varphi_\xi(x_1, x_2, \tau) = \varphi_\theta(\vartheta_1, \vartheta_2, \tau) \frac{1}{\partial(x_1, x_2)/\partial(\vartheta_1, \vartheta_2)},$$

где

$$\frac{\partial(x_1, x_2)}{\partial(\vartheta_1, \vartheta_2)} = \begin{vmatrix} \frac{\partial x_1}{\partial \vartheta_1} & \frac{\partial x_1}{\partial \vartheta_2} \\ \frac{\partial x_2}{\partial \vartheta_1} & \frac{\partial x_2}{\partial \vartheta_2} \end{vmatrix}$$

— якобиан преобразования. В нашем случае x_1 зависит только от ϑ_1 , а x_2 — только от ϑ_2 . Поэтому

$$\varphi_\xi = \varphi_\theta \frac{1}{\frac{dx_1}{d\vartheta_1} \cdot \frac{dx_2}{d\vartheta_2}}$$

и мы имеем

$$\varphi_\xi(x_1, x_2, \tau, t) = \frac{\varphi_\theta(\vartheta_1, \vartheta_2, \tau)}{\sin[\omega_0 t + \vartheta_1] \sin[\omega_0(t + \tau) + \vartheta_2]}.$$

Итак, плотность вероятностей φ_ξ зависит от времени, и процесс (30.6), следовательно, нестационарен.

Перейдем к вычислению спектра, для чего найдем предварительно функцию корреляции. Мы имеем

$$B_\xi(\tau, t) = M(\xi\xi_\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \varphi_\xi(x_1, x_2, \tau, t) dx_1 dx_2.$$

В нашем случае

$$\varphi_\xi(x_1, x_2, \tau, t) dx_1 dx_2 = \varphi_\theta(\vartheta_1, \vartheta_2, \tau) d\vartheta_1 d\vartheta_2.$$

Таким образом,

$$B_\xi(\tau, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \varphi_\theta(\vartheta_1, \vartheta_2, \tau) d\vartheta_1 d\vartheta_2$$

или, подставляя выражения x_1 и x_2 через ϑ_1 и ϑ_2 из (30.7),

$$B_\xi(\tau, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cos(\omega_0 t + \vartheta_1) \cos[\omega_0(t + \tau) + \vartheta_2] \varphi_\theta(\vartheta_1, \vartheta_2, \tau) d\vartheta_1 d\vartheta_2.$$

Преобразуя произведение косинусов, получим два члена, один из которых при усреднении по времени выпадает. Таким образом,

$$\begin{aligned}\overline{B_{\xi}(\tau, t)} &= \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cos(\omega_0\tau + \vartheta_2 - \vartheta_1) \varphi_{\theta}(\vartheta_1, \vartheta_2, \tau) d\vartheta_1 d\vartheta_2 = \\ &= \frac{1}{2} M [\cos(\omega_0\tau + \theta_{\tau} - \theta)] = \frac{1}{2} \cos \omega_0\tau M [\cos(\theta - \theta_{\tau})] - \\ &\quad - \frac{1}{2} \sin \omega_0\tau M [\sin(\theta - \theta_{\tau})].\end{aligned}$$

Мы ограничимся случаем симметричного распределения разности $\zeta(\tau, t) = \theta(t) - \theta(t + \tau)$.

В этом случае второй член выпадает, и мы получаем

$$\overline{B_{\xi}(\tau, t)} = \frac{1}{2} \cos \omega_0\tau M [\cos \zeta(\tau, t)].$$

Отсюда находим спектр процесса $\xi(t)$

$$\begin{aligned}\overline{G_{\xi}(\omega, t)} &= \frac{1}{\pi} \int_0^{\infty} \cos \omega_0\tau \cos \omega\tau M(\cos \zeta) d\tau = \\ &= \frac{1}{2\pi} \int_0^{\infty} M(\cos \zeta) \cos(\omega_0 + \omega)\tau d\tau + \frac{1}{2\pi} \int_0^{\infty} M(\cos \zeta) \cos(\omega_0 - \omega)\tau d\tau\end{aligned}$$

или

$$\overline{G_{\xi}(\omega, t)} = \frac{1}{4} [G(\omega_0 + \omega) + G(\omega_0 - \omega)]. \quad (30.8)$$

Итак, и в случае фазовой модуляции спектр состоит, как и следовало ожидать, из двух боковых полос. Спектр каждой из них выражается формулой

$$G(\omega) = \frac{2}{\pi} \int_0^{\infty} M(\cos \zeta) \cos \omega\tau d\tau. \quad (30.9)$$

Найдем среднее значение $\cos \zeta$. Пусть θ имеет нормальное распределение, в таком случае и разность $\zeta = \theta - \theta_{\tau}$ также имеет нормальное распределение, полностью определяемое дисперсией

$$\begin{aligned}D(\zeta) &= \sigma^2 = M(\theta - \theta_{\tau})^2 = M\theta^2 + M\theta_{\tau}^2 - 2M\theta\theta_{\tau} = \\ &= 2[D(\theta) - B_{\theta}(\tau)] = 2[B_{\theta}(0) - B_{\theta}(\tau)],\end{aligned}$$

и мы можем теперь найти

$$\begin{aligned}M(\cos \zeta) &= \int_{-\infty}^{\infty} \cos z\varphi(z) dz = \frac{2}{\sqrt{2\pi\sigma}} \int_0^{\infty} \cos ze^{-z^2/2\sigma^2} dz = e^{\sigma^2/2} = \\ &= e^{-[B_{\theta}(0) - B_{\theta}(\tau)]}\end{aligned}$$

и выражение для спектра

$$G(\omega) = \frac{2}{\pi} \int_0^{\infty} \cos \omega \tau e^{-[B_{\theta}(0) - B_{\theta}(\tau)]} d\tau. \quad (30.10)$$

Эта формула выражает спектр фазово-модулированного колебания через функцию корреляции модулирующего процесса, распределение которого предполагается нормальным¹. Чтобы довести вычисления до конца, мы рассмотрим частный случай: будем иметь в виду частотную модуляцию с большим индексом.

В случае частотной модуляции задается не непосредственно процесс $\theta(t)$, а процесс $\nu(t) = d\theta/dt$, выражающий мгновенное отклонение частоты от среднего значения ω_0 . Таким образом, мы можем записать для частотно-модулированного колебания

$$\xi(t) = \cos [\omega_0 t + \int \nu(t) dt].$$

Спектр процесса $\xi(t)$ нужно выразить через характеристики процесса $\nu(t)$. Удобно воспользоваться преобразованием Фурье. Если

$$B_{\nu}(\tau) = \int_0^{\infty} G_{\nu}(\omega) \cos \omega \tau d\omega,$$

то

$$B_{\theta}(\tau) = \int_0^{\infty} \frac{1}{\omega^2} G_{\nu}(\omega) \cos \omega \tau d\omega$$

и входящая в предыдущие формулы величина σ может быть выражена через спектр мгновенного частотного отклонения $G_{\nu}(\omega)$ следующим образом:

$$\frac{1}{2} \sigma^2 = B_{\theta}(0) - B_{\theta}(\tau) = \int_0^{\infty} \frac{1}{\omega^2} G_{\nu}(\omega) (1 - \cos \omega \tau) d\omega.$$

Пусть спектр G_{ν} равномерен и ограничен сверху частотой Ω , т. е.

$$G_{\nu}(\omega) = \begin{cases} g = \text{const} & 0 < \omega < \Omega, \\ 0 & \Omega < \omega < \infty. \end{cases}$$

Тогда

$$\frac{1}{2} \sigma^2 = g \int_0^{\Omega} \frac{1 - \cos \omega \tau}{\omega^2} d\omega.$$

Для наших целей удобно представить этот интеграл в следующем виде:

$$\frac{1}{2} \sigma^2 = g \Omega \frac{1}{2} \tau^2 f(\Omega \tau),$$

¹ Аналогичная формула иным путем получена Задэ [37].

где

$$f(\Omega\tau) = 1 - \frac{2}{3 \cdot 4!} (\Omega\tau)^2 + \frac{2}{5 \cdot 6!} (\Omega\tau)^4 - \dots$$

Произведение

$$g\Omega = \int_0^{\infty} G_v(\omega) d\omega$$

есть средний квадрат процесса $v(t)$, т. е. средний квадрат отклонения частоты от среднего значения. Обозначим эту величину через $\Delta\omega^2$. Введя обозначение индекса модуляции $\beta = \Delta\omega/\Omega$, можем переписать выражение для σ^2 в виде

$$\frac{1}{2}\sigma^2 = \frac{1}{2}\beta^2(\Omega\tau)^2 f(\Omega\tau).$$

Теперь перейдем к приближению для случая [6] большого индекса, т. е. для $\beta \gg 1$. Случай большого индекса характеризуется тем, что в выражении для искомого спектра

$$G(\omega) = \frac{2}{\pi} \int_0^{\infty} \cos \omega\tau e^{-\frac{1}{2}\beta^2(\Omega\tau)^2 f(\Omega\tau)} d\tau$$

экспоненциальный множитель убывает тем быстрее, чем больше β . Это значит, что подынтегральное выражение становится пренебрежимо малым при тех меньших значениях τ , чем больше β . Следовательно, при большом β можно ограничиться малыми значениями τ . На этом и основано приближение. В частности, можно положить

$$f(\Omega\tau) \simeq 1$$

с погрешностью, не превосходящей 0,01, если $\Omega\tau \leq 0,6$. При этом условии будем иметь

$$e^{-\frac{1}{2}\beta^2(\Omega\tau)^2 f(\Omega\tau)} \simeq e^{-\frac{1}{2}\beta^2(\Omega\tau)^2} < \varepsilon,$$

если

$$\beta^2 > \frac{2 \ln \frac{1}{\varepsilon}}{(\Omega\tau)^2},$$

откуда, подставляя наибольшее значение $\Omega\tau = 0,6$ и выбирая $\varepsilon = 0,01$, найдем $\beta > 3$.

Соображения, приводящие к описанному приближению, легко обобщаются на случай произвольного спектра G , модулирующего процесса $v(t)$ при условии, что ширина Ω спектра G , достаточно мала по сравнению со среднеквадратичным частотным отклонением $\Delta\omega$. Для получения этого более общего результата доста-

точно в выражении для σ^2 разложить $1 - \cos \omega \tau$ в ряд и сохранить первый член ряда. Это даст

$$\begin{aligned} \frac{1}{2} \sigma^2 &= \int_0^{\infty} \frac{1}{\omega^2} G_v(\omega) (1 - \cos \omega \tau) d\tau \simeq \\ &\simeq \frac{1}{2} \tau^2 \int_0^{\infty} G_v(\omega) d\omega = \frac{1}{2} \tau^2 \Delta \omega^2 = \frac{1}{2} \beta^2 (\Omega \tau)^2, \end{aligned}$$

т. е. то же, что и раньше.

Итак, при большом индексе модуляции, или, что то же, при большом частотном отклонении, мы имеем для спектра приближенное выражение

$$G(\omega) = \frac{2}{\pi} \int_0^{\infty} \cos \omega \tau e^{-\frac{1}{2} \Delta \omega^2 \tau^2} d\tau = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\Delta \omega} e^{-\frac{1}{2} \frac{\omega^2}{\Delta \omega^2}}. \quad (30.11)$$

Здесь следует отметить, что при большом индексе спектр частотно-модулированного колебания не зависит от формы и ширины спектра модулирующего процесса; спектр определяется лишь частотным отклонением, причем ширина спектра частотно-модулированного колебания пропорциональна частотному отклонению.

§ 31. Замечания об анализе случайных процессов

В этом параграфе обсуждаются некоторые вопросы, относящиеся к технике спектрального анализа случайных процессов и к погрешностям, возникающим при анализе, т. е. при экспериментальном определении спектров. Для простоты будем иметь в виду только стационарные эргодические процессы.

Прежде всего заметим, что, как указывалось выше, в условиях эксперимента мы имеем дело обычно с одной из реализаций случайного процесса. Поэтому в основе действия анализаторов спектра случайных процессов лежит определение спектра, содержащее усреднение по времени, например

$$G(\omega) = \frac{1}{\pi} \lim_{T \rightarrow \infty} \frac{1}{T} |S_T(\omega)|^2, \quad (31.1)$$

где

$$S_T(\omega) = \int_0^T \xi(t) e^{-j\omega t} dt$$

— текущий спектр реализации, а величина

$$\frac{1}{\pi} |S_T(\omega)|^2$$

выражает текущий спектр энергии, выделяемой процессом $\xi(t)$ за время T . Величина же

$$\frac{1}{\pi} \cdot \frac{1}{T} |S_T(\omega)|^2 = G_T(\omega)$$

есть спектральная плотность средней за время T мощности процесса $\xi(t)$. Как $|S_T(\omega)|$, так и $G_T(\omega)$ — случайные величины. Лишь в пределе G_T сходится к G , т. е.

$$G(\omega) = \lim_{T \rightarrow \infty} G_T(\omega). \quad (31.2)$$

Совершенно ясно, что в условиях реального эксперимента мы располагаем только конечным временем T . Поэтому в результате однократного опыта мы получаем случайную величину $G_T(\omega)$, которая, вообще говоря, отличается от истинного значения $G(\omega)$. Если же опыт (одной и той же длительности T) повторяется многократно, то получается множество значений $G_T(\omega)$ и в пределе мы могли бы найти истинный спектр, как

$$G(\omega) = M G_T(\omega). \quad (31.3)$$

Перейдем теперь к определению погрешности измерения, т. е. уклонения измеренного спектра от истинного его значения. Сначала определим энергию, выделяемую процессом $\xi(t)$ за время T . Эта энергия равна

$$E_T = \int_0^T \xi^2(t) dt = \frac{1}{\pi} \int_0^\infty |S_T(\omega)|^2 d\omega$$

или

$$E_T = T \int_0^\infty G_T(\omega) d\omega.$$

Очевидно, что и E_T — случайная величина. Ее среднее значение есть

$$E = M E_T = T M \int_0^\infty G_T(\omega) d\omega = T \int_0^\infty G(\omega) d\omega.$$

Средний квадрат уклонения E_T от E — это и есть средний квадрат ошибки измерения. Таким образом, средний квадрат ошибки выражается непосредственно дисперсией величины

$$\sigma^2 = D E_T = M(E_T^2) - M^2(E_T) = M(E_T^2) - E^2. \quad (31.4)$$

Мы имеем

$$E_T^2 = \left(T \int_0^\infty G_T(\omega) d\omega \right)^2 = T^2 \int_0^\infty \int_0^\infty G_T(\nu_1) G_T(\nu_2) d\nu_1 d\nu_2,$$

Для среднего значения этой величины Райс [34] выводит следующее выражение:

$$M(E_T^2) = T^2 \int_0^\infty \int_0^\infty G(\nu_1) G(\nu_2) \left[1 + \left(\frac{\sin(\nu_1 + \nu_2) T/2}{(\nu_1 + \nu_2) T/2} \right)^2 + \right. \\ \left. + \left(\frac{\sin(\nu_1 - \nu_2) T/2}{(\nu_1 - \nu_2) T/2} \right)^2 \right] d\nu_1 d\nu_2, \quad (31.5)$$

откуда

$$\sigma^2 = T^2 \int_0^\infty \int_0^\infty G(\nu_1) G(\nu_2) \left[\left(\frac{\sin(\nu_1 + \nu_2) T/2}{(\nu_1 + \nu_2) T/2} \right)^2 + \right. \\ \left. + \left(\frac{\sin(\nu_1 - \nu_2) T/2}{(\nu_1 - \nu_2) T/2} \right)^2 \right] d\nu_1 d\nu_2. \quad (31.6)$$

Удобнее нормировать ошибку, т. е. ввести среднеквадратичную относительную ошибку, определив ее как $\varepsilon = \sigma/E$.

Теперь представим себе анализатор в виде идеального фильтра с коэффициентом передачи, равным единице в полосе частот от ω_1 до ω_2 и равным нулю вне этого интервала. Процесс анализа состоит в измерении энергии, выделяемой на выходе фильтра за время T . Если анализатор имеет высокую разрешающую способность, то полоса $\Omega = \omega_2 - \omega_1$ узка, и мы можем считать, что в пределах этой полосы истинная спектральная плотность постоянна и равна g_0 . В таком случае $E = Tg_0\Omega$, а для среднего квадрата погрешности имеем

$$\sigma^2 = T^2 g_0^2 \int_{\omega_1}^{\omega_2} \int_{\omega_1}^{\omega_2} \left[\left(\frac{\sin(\nu_1 + \nu_2) T/2}{(\nu_1 + \nu_2) T/2} \right)^2 + \left(\frac{\sin(\nu_1 - \nu_2) T/2}{(\nu_1 - \nu_2) T/2} \right)^2 \right] d\nu_1 d\nu_2.$$

Для очень малых T имеем (полагая $\sin x/x \approx 1$) $\sigma^2 \approx 2T^2 g_0^2 \Omega^2$ и $\sigma/E \approx \sqrt{2}$; для очень больших T получается асимптотическое выражение $\sigma/E \sim \sqrt{2\pi/\Omega T} = 1/\sqrt{FT}$. Следует заметить, что асимптотическое выражение вида $\sigma/E \sim A/\sqrt{\Omega T}$ имеет силу не только для идеального фильтра, который мы рассматриваем, но и для других типов фильтров; постоянная A зависит от свойства фильтра [26].

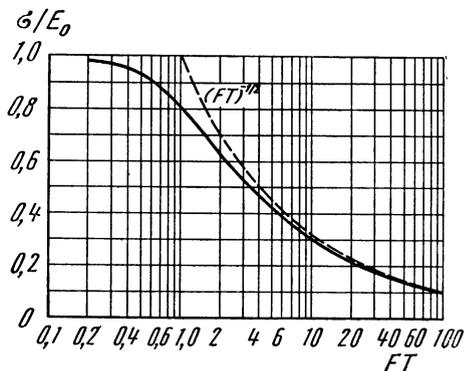
Для случая полосы, ширина которой мала по сравнению со средней частотой, т. е. при условии

$$\Omega = \omega_2 - \omega_1 \ll \omega_0 = \frac{1}{2}(\omega_1 + \omega_2)$$

подсчитаны [34] значения относительной погрешности как функции аргумента FT ($F = f_2 - f_1 = \Omega/2\pi$) (выкладки Райса очень громоздки, и мы их не приводим); график этой зависимости представлен на рис. 55. Как видим, ошибка убывает с возрастанием T очень медленно. Ошибка тем больше, чем меньше F . Так, например, для

того чтобы ошибка была меньше 10%, нужно взять $FT=100$; при разрешающей способности 100 *гц* измерение должно продолжаться не менее 1 *сек.*

Заметим, что формула (31.6) выражает погрешность через истинный спектр $G(\omega)$, который неизвестен, так как он-то и должен быть определен в результате измерений. Возникающая здесь принципиальная трудность обойдена путем предположения, что спектральная плотность постоянна в узкой полосе Ω . Вычислить погрешность только на основании результатов измерения спектра



Р и с. 55

нельзя принципиально. Для вычисления погрешности нужно знать нечто большее, а именно корреляцию спектра, который является моментом второго порядка; нужно, следовательно, измерять моменты четвертого порядка. Подробности можно найти в специальных работах по этому вопросу [25, 26].

Обратимся теперь к вопросам методики и техники спектрального анализа случайных процессов. Прежде всего отметим, что из основного определения спектра случайного процесса

$$G(\omega) = \frac{2}{\pi} \int_0^{\infty} B(\tau) \cos \omega \tau d\tau$$

вытекает возможность косвенного измерения спектра, состоящая в том, что из опыта находится функция корреляции $B(\tau)$, а затем спектр получается путем преобразования Фурье (которое может выполняться либо путем численной обработки экспериментальных данных, либо при помощи того или иного прибора, выполняющего преобразование Фурье). При таком способе нахождения спектра в качестве основного измерительного прибора применяется не анализатор, а коррелометр. По вопросам измерения функций корреляции имеется обширная литература, обзор которой выходит за рамки нашей темы. Заметим лишь, что существует мнение [7], что определение спектра через функцию корреляции более эко-

номно с точки зрения суммы требуемых измерительных и вычислительных операций, нежели прямое измерение спектра.

Что же касается методов непосредственного анализа, то простейший анализатор представляется в виде набора полосовых фильтров; в качестве выходных индикаторов должны применяться интегрирующие квадратичные приборы, измеряющие энергию. Это — наиболее чистый способ измерения, так как применение ваттметров сводится по существу к интегрированию с весом; весовая функция зависит от свойств как самого прибора, так и схемы его включения, что вносит элемент произвола в результат измерения. Точнее говоря, результат измерения представляет собой величину, определение которой вытекает из свойств измерительного устройства. Пример такого специализированного определения можно видеть в уже цитированной работе Фано [27].

Другой возможный способ анализа состоит в том, что измеряется энергия на выходе фильтра нижних частот с подвижной верхней границей, т. е. энергия в полосе от 0 до ω . Эта энергия равна

$$E_T(\omega) = \int_0^{\omega} G_T(\nu) d(\nu).$$

Если найденную функцию продифференцировать по частоте, то получим

$$G_T(\omega) = \frac{\partial}{\partial \omega} E_T(\omega).$$

Возможно применение приема, вообще очень употребительного в технике анализа, а именно периодическое повторение образца анализируемого процесса. Практически это осуществляется путем записи, например, на магнитной ленте, с последующим склеиванием достаточно длинного отрезка записи в кольцо. В месте склейки образуется, разумеется, разрыв если не самой функции, то ее производных. Это обстоятельство оказывает на результат анализа специфическое влияние, с которым иногда приходится считаться [36].

§ 32. О возможностях сжатия спектра

Со времени зарождения техники электрической связи одна проблема не переставала волновать специалистов — проблема уменьшения ширины спектра сигнала, или, короче, проблема сжатия спектра. Многочисленные попытки не привели пока к радикальному решению проблемы в целом. Вместе с тем имеется уже ряд работ, представляющих определенные достижения в частных аспектах проблемы. Не предпринимая здесь анализа этих работ, ограничимся общей постановкой задачи, для чего нам потребуется привлечь основные понятия теории информации.

Совершенно очевидно, что сжатие спектра некоторого сигнала, рассматриваемого как случайный процесс $\xi(t)$, возможно только в результате некоторого преобразования, которому подвергается исходный процесс. Мы образуем новый процесс

$$\eta(t) = \Psi[\xi(t)], \quad (32.1)$$

где Ψ — символ некоторой функциональной операции, и нужно подобрать операцию Ψ так, чтобы процесс $\eta(t)$ обладал более узким спектром, чем исходный процесс $\xi(t)$. Если преобразование обратимо, т. е. если существует однозначная обратная операция Ψ^{-1} , посредством которой исходный процесс может быть восстановлен, т. е.

$$\xi(t) = \Psi^{-1}[\eta(t)] = \Psi^{-1}\{\Psi[\xi(t)]\}, \quad (32.2)$$

то, очевидно, задача полностью решена. Нужно сразу указать, что требование обратимости выполняется далеко не для всех преобразований. Так, например, сокращающая спектр операция, выполняемая идеальным фильтром и состоящая в полном устранении части спектра процесса, необратима.

Требование полной обратимости (32.2) является очень жестким, и может быть разумным образом смягчено, если опираться на информационное представление об исходном процессе. Именно, можно потребовать, чтобы операция Ψ была инвариантна по отношению к содержащейся в исходном процессе информации. Иными словами, при преобразовании информация не должна теряться (возрасти она, очевидно, не может, так как в преобразовании не участвует какой-либо источник информации). При такой постановке вопроса можно представить себе, что исходный процесс зависит от двух функций, одна из которых содержит всю информацию, а вторая не содержит информации вовсе. Функция, не несущая информации, может быть добавлена на приемной стороне системы связи; в передаче этой функции надобности нет. Итак, требования к операции Ψ сводятся теперь к следующим двум: 1) преобразованный процесс $\eta(t)$ должен обладать более узким спектром и 2) преобразованный процесс должен сохранять информацию исходного процесса.

Напомним теперь некоторые определения, применяемые в теории информации. Мы будем иметь в виду в качестве исходных квантованные процессы, т. е. процессы с дискретным распределением вероятностей, так как в этом случае определения просты и недвусмысленны.

Пусть исходный процесс представляет собой случайную функцию с ограниченным спектром, простирающимся от 0 до F^* . По теореме Котельникова такой процесс может быть заменен последовательностью мгновенных значений, отсчитанных через ин-

* Мы опускаем здесь тонкости, относящиеся к информационной трактовке теоремы Котельникова (см. сноску на стр. 146).

тервалы $\Delta t = 1/2F$. На промежутке T таких отсчетов будет $n = T/\Delta t = 2FT$. Пусть, далее, мгновенное значение может принимать одно из m равновероятных значений. Всего возможно

$$N = m^n \quad (32.3)$$

реализаций. В таком случае информация, содержащаяся в процессе $\xi(t)$, определяется, как

$$I = \log_2 N = n \log_2 m \quad (32.4)$$

или

$$I = 2FT \log m. \quad (32.5)$$

Из этой формулы сразу видно, что при неизменной информации I для сокращения ширины спектра E существуют только две возможности: либо за счет увеличения T , либо за счет увеличения m . Такого рода преобразования можно осуществить по-разному, ниже приведены два простейших примера.

Пример 1. Пусть сигнал $\xi(t)$ записан, а затем воспроизведен с пониженной скоростью движения носителя. При этом спектр сожмется во столько же раз, во сколько увеличится длительность. Восстановление исходного сигнала производится путем повторной записи с последующим воспроизведением с повышенной скоростью. Происходящие преобразования могут быть записаны в следующем виде:

$$\begin{aligned} \eta(t) &= \Psi[\xi(t)] = \xi(at), \\ \xi(t) &= \Psi^{-1}[\eta(t)] = \eta\left(\frac{1}{a}t\right). \end{aligned}$$

Пример 2. Группа из M знаков исходной последовательности может быть представлена M -значным числом, записанным по системе счисления с основанием m . Но та же группа может быть записана однозначным числом по системе с основанием $l > m$, причем, очевидно, должно быть $l \geq m^M$. При таком преобразовании за одно и то же время нужно передать один знак вместо M знаков. Это значит, что ширина спектра сокращается в $M = \log l / \log m$ раз.

Возможности сокращения ширины спектра в рамках формулы (32.5), как видим, очень ограничены. Гораздо больший интерес представляют преобразования, затрагивающие статистику сигнала.

Выражение для информации, содержащейся в последовательности из n символов, может быть в общем случае записано в виде

$$I = nI',$$

где I' — информация на один символ (энтропия). В уже рассмотренном случае m равновероятных значений

$$I' = \log m = I'_0. \quad (32.6)$$

Если значения, которые может принимать процесс, неравновероятны, и вероятность i -го значения обозначена через p_i , то информация на символ определяется как

$$I' = -\sum_{i=1}^m p_i \log p_i = I'_1. \quad (32.7)$$

Легко видеть, что при $p_i = 1/m$ (32.7) переходит в (32.6). Формула (32.7) предполагает, что символы неравновероятны, но статистически независимы.

В наиболее же общем случае, когда между символами существуют статистические связи, можно характеризовать статистическую структуру сигнала многомерной вероятностью данной реализации

$$p_k = P_k(\xi_1, \xi_2, \dots, \xi_n)$$

и в этом случае

$$I' = -\frac{1}{n} \sum_{k=1}^{N=m^n} p_k \log p_k = I'_2. \quad (32.8)$$

Все дальнейшее основано на том, что $I'_0 > I'_1 > I'_2$. Поэтому, если записать выражение для информации в общем виде

$$I = nI' = 2FTI',$$

то станет ясным, что, не изменяя длительности T и числа уровней m , можно сократить ширину спектра F за счет преобразования статистики процесса, т. е. за счет изменения величины I' . Это преобразование должно заключаться в разрушении статистических взаимосвязей (переход от I'_2 к I'_1) с последующим выравниванием вероятностей независимых символов (переход от I'_1 к I'_0). Информация, содержащаяся в последовательности равновероятных и независимых символов, т. е.

$$I_0 = nI'_0 = 2FTI'_0$$

является наибольшей; иначе говоря, сжатие спектра на статистической основе при $I' = I'_0$ невозможно.

Величина

$$R = 1 - I'/I'_0 \quad (32.9)$$

носит название избыточности. Таким образом, сказанное выше может быть сформулировано следующим образом: статистическое сжатие спектра возможно за счет устранения избыточности. Для неизменного числа уровней выражение для избыточности можно записать еще в виде

$$R = 1 - F_0 T_0 / FT, \quad (32.10)$$

где F и T — фактические ширина спектра и длительность сигнала, а F_0 и T_0 — наименьшие значения тех же величин для идеально

преобразованного сигнала (т. е. для сигнала, для которого $I' = I'_0$). Из (32.10) следует, что при неизменной длительности сигнала ширина спектра может быть уменьшена в $F/F_0 = 1/(1-R)$ раз.

Итак, для выяснения возможности статистического сжатия спектра нужно уметь оценить величину избыточности, а для осуществления этой возможности нужно найти преобразование, устраняющее или, по крайней мере, уменьшающее избыточность. Эта обширная проблема с точки зрения теории информации есть не что иное, как проблема статистически-оптимального кодирования. Обсуждение этой проблемы и путей к ее решению завело бы нас слишком далеко. Мы ограничимся лишь одним замечанием, поясняющим принципиальную сторону дела.

Воспользуемся данными, относящимися к модулированным колебаниям (§ 30). Известно, что ширина спектра амплитудно-модулированного колебания ровно в два раза больше ширины спектра модулирующего процесса. Так как в результате модуляции количество информации не меняется, то модулированное колебание имеет избыточность

$$R = 1 - F/2F = 0,5$$

(предполагается, что модулирующий процесс лишен избыточности). Путем детектирования восстанавливается исходный (модулирующий) процесс. Таким образом, можно рассматривать модуляцию как преобразование, увеличивающее избыточность (и расширяющее спектр), а детектирование — как операцию, устраняющую избыточность (и сжимающую спектр). Следовательно, если статистика исходного процесса подобна статистике амплитудно-модулированного колебания, то операция типа детектирования даст желаемое сокращение спектра. Другая возможность устранения избыточности, приводящая также к сокращению спектра вдвое, состоит в устранении одной из боковых полос модуляционного спектра.

Те же соображения относятся и к колебанию, модулированному по фазе (или по частоте). Дело несколько усложняется в данном случае тем, что строение боковых полос не повторяет строения модулирующего спектра, как при АМ. Поэтому ширину спектра ЧМ-колебания нужно предварительно определить. Возьмем случай большого индекса и нормально распределенного модулирующего процесса, когда спектр боковой полосы записывается в виде (формула (30.11))

$$G(\omega) = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\Delta\omega} e^{-\omega^2/2\Delta\omega_0^2}.$$

Для определения ширины спектра воспользуемся моментным критерием (§ 12), т. е. положим

$$\Omega_0^2 = 1/A - M^2/A^2,$$

где $A = \int_0^{\infty} G d\omega$, $M = \int_0^{\infty} \omega G d\omega$, $I = \int_0^{\infty} \omega^2 G d\omega$. Выполнив интегрирование, получим $A = 1$, $M = (\sqrt{2/\pi}) \Delta\omega$, $I = \Delta\omega^2$, откуда $\Omega_0 = (\sqrt{1 - 2/\pi}) \times \Delta\omega \simeq 0,8\Delta\omega$, а полная ширина спектра $2\Omega \simeq 1,6\Delta\omega \simeq 2\Delta\omega$.

Таким образом, можно считать, что по порядку величины ширина спектра ЧМ-колебания определяется удвоенным частотным отклонением (т. е. полосой качания). Если принять $\Delta f = \Delta\omega/2\pi = 75$ кГц, то ширина спектра составит $F \simeq 2\Delta f = 150$ кГц. При ширине спектра модулирующего процесса 5 кГц $F/F_0 = 30$, а избыточность $R = 1 - F/F_0 \simeq 0,97$. Итак, для процесса, статистика которого близка к статистике ЧМ-колебания с большим индексом, операция типа частотного детектирования может привести к существенному сокращению спектра.

ДОБАВЛЕНИЯ

I. О ширине спектра произведения функций

В § 4 доказана теорема о спектре произведения двух функций времени: если даны две функции времени $f_1(t)$ и $f_2(t)$ и их спектры $S_1(\omega)$ и $S_2(\omega)$, то спектр произведения функций $f_1(t)$ и $f_2(t)$ выражается формулой

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S_1(\nu) S_2(\omega - \nu) d\nu.$$

Аналогичное соотношение имеется и для спектров мощности (см. § 27)

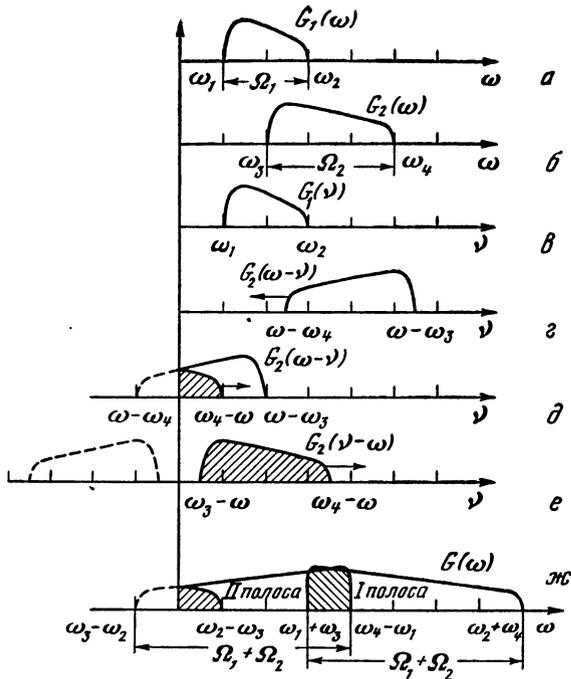
$$G(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} G_1(\nu) G_2(\omega - \nu) d\nu. \quad (\text{I. 1})$$

В качестве одного из применений этой формулы рассмотрим вопрос о ширине спектра произведения двух функций.

Вопрос ставится так: даны две функции, относительно которых известно, что спектры их ограничены, т. е. имеют ширину $\Omega_1 = \omega_2 - \omega_1$ для первой функции и $\Omega_2 = \omega_4 - \omega_3$ для второй. Это свойство заданных нам спектров может быть записано в виде

$$\begin{aligned} G_1(\omega) &\neq 0 \quad \text{при} \quad \omega_1 < \omega < \omega_2, \\ G_2(\omega) &\neq 0 \quad \text{при} \quad \omega_3 < \omega < \omega_4. \end{aligned}$$

Нам нужно определить ширину спектра произведения функций. Спектр произведения не равен нулю, если не равна нулю подынтегральная функция в выражении (I. 1). Это условие выполняется, когда спектры $G_1(\nu)$ и $G_2(\omega - \nu)$ перекрываются, т. е. в той области значений ω , для которой на некотором интервале



Р и с. 56

значений ν оба спектра $G_1(\nu)$ и $G_2(\omega - \nu)$ не равны нулю. Из этих соображений можно вывести систему неравенств, которая даст нам ответ на поставленный вопрос. Составление этой системы сильно облегчается при помощи графиков рис. 56.

На рис. 56, *a* и *б* изображены заданные спектры $G_1(\omega)$ и $G_2(\omega)$. На рис. 56, *в* представлен спектр $G_1(\nu)$; этот график ничем не отличается от графика рис. 56, *a*. На рис. 56, *г* построен спектр $G_2(\omega - \nu)$. Он отличается от рис. 56, *б* тем, что начало отсчета перенесено в точку $\nu = \omega$ и спектр повернут справа налево. ω — это текущая частота результирующего спектра. При ее изменении в сторону уменьшения спектр передвигается влево, как показано стрелкой. Условие перекрытия спектров рис. 56, *в* и *г* может быть записано в виде

$$\omega_2 > \omega - \omega_4$$

ИЛИ

$$\omega < \omega_2 + \omega_4. \quad (I. 2)$$

Перекрытие продолжается, пока соблюдено условие

$$\omega - \omega_3 > \omega_1$$

ИЛИ

$$\omega > \omega_1 + \omega_3. \quad (I. 3)$$

Объединяя неравенства (I. 2) и (I. 3), можем записать для первой полосы перекрытия

$$\omega_1 + \omega_3 < \omega < \omega_2 + \omega_4, \quad (\text{I. 4})$$

откуда ширина первой полосы

$$\omega_2 + \omega_4 - (\omega_1 + \omega_3) = \Omega_1 + \Omega_2.$$

При дальнейшем уменьшении ω спектр $G_2(\omega - \nu)$ заходит в область отрицательных частот, как показано штрихом на рис. 56, *д*. Действительный же спектр располагается в области положительных значений частоты; его можно построить путем перегибания спектра по оси ординат. Перегнутая часть спектра отмечена штриховкой на рис. 56, *д*. При дальнейшем уменьшении частоты ω полностью перевернутый спектр движется вдоль оси ν вправо, как показано на рис. 56, *е*. Последние два рисунка показывают, что спектры вступают во вторую полосу перекрытия, границы которой определяются неравенствами

$$\omega_1 < \omega_4 - \omega, \quad \omega_2 > \omega_3 - \omega,$$

т. е.

$$\omega_3 - \omega_2 < \omega < \omega_4 - \omega_1. \quad (\text{I. 5})$$

Ширина второй полосы

$$\omega_4 - \omega_1 - (\omega_3 - \omega_2) = \Omega_1 + \Omega_2.$$

Таким образом, мы получили следующее правило:

спектр произведения двух функций состоит из двух полос, ширина каждой из которых равна сумме ширин спектров обоих множителей. Границы полос определяются неравенствами (I. 4) и (I. 5).

Это еще не означает, что общая ширина спектра произведения равна просто удвоенной сумме ширин спектров множителей. Дело в том, что одна из полос может захватывать область отрицательных частот; при этом она перегибается, как описано выше, и ее действительная ширина сокращается. Кроме того, обе полосы могут частично перекрываться между собой.

И то и другое показано на рис. 56, *ж*. Все графики рис. 56 относятся к следующим числовым значениям: $\omega_1=1$, $\omega_2=3$, $\omega_3=2$, $\omega_4=5$, $\Omega_1=2$, $\Omega_2=3$, $\omega_2 + \omega_4=8$, $\omega_1 + \omega_3=3$, $\omega_4 - \omega_1=4$, $\omega_3 - \omega_2=-1$. Спектр произведения в этом примере занимает непрерывную полосу $0 < \omega < 8$. Таким образом, ширину спектра произведения можно определить с помощью следующего правила:

общая ширина спектра произведения двух функций не менее суммы, но не более удвоенной суммы ширин спектров обоих множителей.

Вышеприведенные правила имеют достаточно общий характер. Их можно, например, применить и к дискретным (линейчатым) спектрам. Можно привести еще два правила, относящиеся только к дискретным спектрам:

число линий в каждой из двух полос спектра произведения двух периодических функций равно произведению чисел линий в спектрах множителей.

Но отдельные линии могут сливаться. Поэтому

общее число линий в спектре произведения двух периодических функций не меньше произведения, но не больше удвоенного произведения чисел линий в спектрах множителей.

Для применения этих правил следует иметь в виду, что наличие постоянной составляющей выражается спектральной линией с частотой $\omega=0$.

В качестве примера рассмотрим спектр сложномодулированного колебания. Пусть несущая частота ω_0 промодулирована по частоте частотой ω_1 , а затем по амплитуде частотой ω_2 . Аналитически такое колебание выражается формулой

$$x = c_0 (1 + m \sin \omega_2 t) \sin (\omega_0 t + \beta \sin \omega_1 t),$$

где c_0 — амплитуда несущей; m — глубина АМ; β — индекс ЧМ.

Если обозначить

$$\sin (\omega_0 t + \beta \sin \omega_1 t) = f_1(t),$$

$$1 + m \sin \omega_2 t = f_2(t),$$

то вопрос о спектре сложномодулированного колебания сводится к вопросу о спектре произведения. Спектр $G_1(\omega)$ функции $f_1(t)$ при малом индексе состоит из трех линий с частотами ω_0 , $\omega_0 \pm \omega_1$. Спектр $G_2(\omega)$ функции $f_2(t)$ состоит из двух линий с частотами 0 и ω_2 . Для получения результирующего спектра надо передвигать спектр $G_2(\omega - \nu)$, имеющий две линии с частотами ω и $\omega - \omega_2$, вдоль оси ν . В результирующем спектре линии будут возникать всякий раз, когда совпадают любые две линии спектров $G_1(\nu)$ и $G_2(\omega - \nu)$. Отметив все возможные совпадения, мы найдем, что спектр рассматриваемого сложномодулированного колебания будет состоять из девяти линий с частотами

$$\omega_0, \omega_0 \pm \omega_1, \omega_0 \pm \omega_2, \omega_0 \pm \omega_1 \pm \omega_2.$$

Крайние частоты спектра равны $\omega_0 \pm (\omega_1 + \omega_2)$.

II. Спектры некоторых частотно-модулированных колебаний

В § 7 вычислен спектр для простейшего случая синусоидальной ЧМ

$$\omega = \omega_0 \left(1 + \frac{\Delta\omega}{\omega_0} \cos \Omega t \right).$$

Общий случай ЧМ может быть записан в виде

$$\omega = \omega_0 [1 + mf(t)],$$

откуда

$$\vartheta = \omega_0 t + mF(t),$$

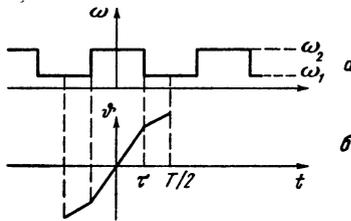
где

$$F(t) = \int_0^t f(\tau) d\tau.$$

Так как модулированное колебание имеет вид

$$x = \sin \vartheta = \sin [\omega_0 t + mF(t)],$$

то, очевидно, что представление функции $F(t)$ каким бы то ни было (например, степенным или тригонометрическим) родом ни к чему хорошему не приведет. Необходимо стараться предста-



Р и с. 57

вить $F(t)$ в конечной форме и притом таким образом, чтобы интеграл, которым выражается спектр, мог быть вычислен. Мы приведем здесь два примера такого рода вычислений.

а. Спектр колебания с частотой, модулированной по закону прямоугольной ломаной линии. Закон изменения частоты следующий (рис. 57, а):

$$\begin{aligned} \omega &= \omega_2 & \text{при} & \quad -\tau < t < \tau, \\ \omega &= \omega_1 & \text{при} & \quad -T/2 < t < -\tau, \quad \tau < t < T/2 \end{aligned}$$

и т. д. На рис. 57, б дан график ϑ , имеющий следующее аналитическое выражение на промежутке $(-T/2, T/2)$:

$$\begin{aligned} \vartheta &= \omega_2 t & \text{при} & \quad -\tau < t < \tau, \\ \vartheta &= \omega_1 t + b\tau & \text{при} & \quad \tau < t < T/2, \\ \vartheta &= \omega_1 t - b\tau & \text{при} & \quad -T/2 < t < -\tau, \end{aligned}$$

где $b = \omega_2 - \omega_1$. Так как ϑ — нечетная функция, то спектр функции $x = \sin \vartheta$ выразится через синусные коэффициенты

$$c_k = b_k = \frac{1}{T} \left[\int_{-\tau}^{\tau} \sin \omega_2 t \sin k\Omega t dt + \int_{-T/2}^{-\tau} \sin (\omega_1 t - b\tau) \sin k\Omega t dt + \int_{\tau}^{T/2} \sin (\omega_1 t + b\tau) \sin k\Omega t dt \right].$$

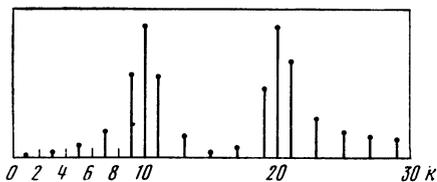
Выполняя простые преобразования и интегрируя, находим

$$c_k = \frac{1}{T} \left\{ \frac{1}{\omega_2 - k\Omega} \sin(\omega_2 - k\Omega)\tau - \frac{1}{\omega_2 + k\Omega} \sin(\omega_2 + k\Omega)\tau + \right. \\ \left. + \frac{1}{\omega_1 - k\Omega} [\sin(\omega_1 T/2 + b\tau - k\tau) - \sin(\omega_2 - k\Omega)\tau] - \right. \\ \left. - \frac{1}{\omega_1 + k\Omega} [\sin(\omega_1 T/2 + b\tau + k\tau) - \sin(\omega_2 + k\Omega)\tau] \right\}.$$

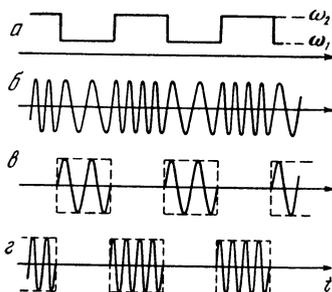
В частном случае $\tau = T/4$, т. е. когда зубцы и впадины ломаной (рис. 57, а) имеют одинаковую ширину, получаем

$$c_k = \frac{1}{4} \left\{ \frac{\sin \frac{\pi}{2} \left(\frac{\omega_2}{\Omega} - k \right)}{\frac{\pi}{2} \left(\frac{\omega_2}{\Omega} - k \right)} - \frac{\sin \frac{\pi}{2} \left(\frac{\omega_2}{\Omega} + k \right)}{\frac{\pi}{2} \left(\frac{\omega_2}{\Omega} + k \right)} + \right. \\ \left. + \frac{\sin \frac{\pi}{4} \left(\frac{\omega_1}{\Omega} - k \right)}{\frac{\pi}{4} \left(\frac{\omega_1}{\Omega} - k \right)} \cos \frac{\pi}{4} \left(\frac{\omega_1}{\Omega} + 2 \frac{\omega_2}{\Omega} - 3k \right) + \right. \\ \left. + \frac{\sin \frac{\pi}{4} \left(\frac{\omega_1}{\Omega} + k \right)}{\frac{\pi}{4} \left(\frac{\omega_1}{\Omega} + k \right)} \cos \frac{\pi}{4} \left(\frac{\omega_1}{\Omega} + 2 \frac{\omega_2}{\Omega} + 3k \right) \right\}.$$

На рис. 58 представлен спектр, вычисленный по этой формуле ($\omega_1/\Omega = 10$, $\omega_2/\Omega = 20$). Спектр этот имеет максимумы при $k\Omega = \omega_1$ и $k\Omega = \omega_2$, что и следовало ожидать.



Р и с. 58



Р и с. 59

Спектр рассматриваемого ЧМ-колебания может быть вычислен на основе других соображений. На рис. 59, а представлен тот же закон изменения частоты, что и на рис. 57, а, а на рис. 59, б — соответствующее ЧМ-колебание. Это колебание может быть разложено на два, показанных на рис. 59, в и г. Но каждая из этих составляющих представляет собою не что иное, как колебание соответствующей несущей частоты, модулированное по амплитуде прямоугольными импульсами. Спектр такого рода колебания находится без всяких затруднений, после чего спектр ЧМ-колебания находится как сумма спектров обоих АМ-колебаний. Этот

прием может быть применен и в том случае, когда частота принимает не два, а любое число фиксированных значений. Таким способом можно, следовательно, вычислять спектры сигналов не только обычного частотного телеграфа, но и спектры ДЧТ — двухканального частотного телеграфа, работающего на четырех фиксированных частотах. Техника вычисления не меняется, если

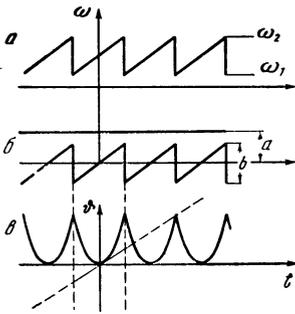


Рис. 60

имеется скачок фазы при переходе с частоты на частоту; в этом случае несколько усложняются, конечно, выкладки.

б. Спектр частотно-модулированного колебания с частотой, изменяющейся по пилообразному закону. Этот случай, так называемый «свип-сигнал», несколько сложнее предыдущего.

Пусть $\omega = a + bt/T$ при $-T/2 < t < T/2$ (рис. 60, а), где $a = (\omega_1 + \omega_2)/2$, $b = \omega_2 - \omega_1$. На рис. 60, б обе составляющие частоты изображены раздельно.

Для аргумента ϑ получаем $\vartheta = at + bt^2/2T$ при $-T/2 < t < T/2$ (рис. 60, в), и модулированное колебание записывается в виде

$$x = \sin(at + bt^2/2T).$$

Вычислим компоненты спектра

$$a_k = \frac{2}{T} \int_{-T/2}^{T/2} \sin\left(at + \frac{b}{2T} t^2\right) \cos k\Omega t dt,$$

$$b_k = \frac{2}{T} \int_{-T/2}^{T/2} \sin\left(at + \frac{b}{2T} t^2\right) \sin k\Omega t dt.$$

Для вычисления этих интегралов надо сначала сделать тригонометрические преобразования. Например, для a_k получим

$$a_k = \frac{1}{T} \int_{-T/2}^{T/2} [\sin(\alpha t + \gamma^2 t^2) + \sin(\beta t + \gamma^2 t^2)] dt,$$

где обозначено для сокращения $\alpha = a + k\Omega$, $\beta = a - k\Omega$, $\gamma^2 = b/2T$.

В этой форме нужно дополнить аргумент до полного квадрата, прибавляя и вычитая соответственно $\alpha^2/4\gamma^2$ и $\beta^2/4\gamma^2$. Затем, разлагая снова синусы разности, найдем

$$a_k = \frac{1}{T\gamma} \left(\cos \frac{\beta^2}{4\gamma^2} C_y + \sin \frac{\beta^2}{4\gamma^2} S_y - \cos \frac{\alpha^2}{4\gamma^2} C_x - \sin \frac{\alpha^2}{4\gamma^2} S_x \right).$$

Здесь применены следующие обозначения:

$$C_x = \int_{x_1}^{x_2} \cos x^2 dx, \quad C_y = \int_{y_1}^{y_2} \cos y^2 dy, \quad S_x = \int_{x_1}^{x_2} \sin x^2 dx,$$

$$S_y = \int_{y_1}^{y_2} \sin y^2 dy,$$

$$x_1 = \frac{1}{2} \left(\frac{\alpha}{\gamma} - \gamma T \right), \quad y_1 = \frac{1}{2} \left(\frac{\beta}{\gamma} - \gamma T \right), \quad x_2 = \frac{1}{2} \left(\frac{\alpha}{\gamma} + \gamma T \right),$$

$$y_2 = \frac{1}{2} \left(\frac{\beta}{\gamma} + \gamma T \right).$$

Аналогичная формула получается и для b_k

$$b_k = \frac{1}{T\gamma} \left(\cos \frac{\beta^2}{4\gamma^2} S_y - \sin \frac{\beta^2}{4\gamma^2} C_y + \cos \frac{\alpha^2}{4\gamma^2} S_x - \sin \frac{\alpha^2}{4\gamma^2} C_x \right).$$

Нас интересует амплитудный спектр. Амплитуда k -й гармоники равна

$$c_k = \sqrt{a_k^2 + b_k^2}.$$

Проделав вычисления, получим

$$c_k = \frac{1}{T\gamma} \left[C_y^2 + S_y^2 + C_x^2 + S_x^2 - 2 \cos \frac{\beta^2 + \alpha^2}{4\gamma^2} (C_y C_x - S_y S_x) - \right. \\ \left. - 2 \sin \frac{\beta^2 + \alpha^2}{4\gamma^2} (C_y S_x + S_y C_x) \right]^{1/2}.$$

В заключение заметим, что (в отличие от АМ)-спектр ЧМ-колебания в общем случае несимметричен относительно центральной частоты, как бы мы ее ни определили; характер асимметрии спектра зависит от закона изменения частоты. Симметричный спектр получается только в случае симметричного (относительно некоторого центрального значения) закона изменения частоты. Примером может служить рассмотренный в § 7 случай синусоидального изменения частоты.

III. Активная полоса спектра

Представления, связанные с мгновенным спектром, развивает предложенное Д. В. Агеевым понятие активной полосы спектра [1].

Будем исходить из определения мгновенного спектра мощности по Пейджи (см. § 6)

$$\rho(\omega, t) = \frac{\partial}{\partial t} |S_t(\omega)|^2,$$

где

$$S_t(\omega) = \int_{-\infty}^t f(u) e^{-j\omega u} du$$

— текущий спектр. По определению мгновенная мощность

$$f^2(t) = \frac{1}{\pi} \int_0^{\infty} \rho(\omega, t) d\omega.$$

Но можно определить некоторый интервал частот, в пределах которого сосредоточена подавляющая доля мощности. Это условие может быть записано в виде

$$\frac{1}{\pi} \int_{\omega_1}^{\omega_2} \rho(\omega, t) d\omega = \eta f^2(t),$$

где η — правильная дробь, немного меньшая единицы. Интервал (ω_1, ω_2) может быть назван активной полосой. Аналогичным критерием мы пользовались в § 12, но теперь речь идет о мгновенных спектрах, а следовательно, ω_1 и ω_2 являются функциями времени. Поэтому не только ширина интервала (ω_1, ω_2) , но и положение его на шкале частот зависит от времени.

К определению активной полосы можно подойти и по-другому. По определению текущего спектра

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_t(\omega) e^{j\omega t} d\omega = \frac{1}{\pi} \int_0^{\infty} (A_t \cos \omega t - B \sin \omega t) d\omega,$$

где $A_t + jB_t = S_t(\omega)$.

Введем теперь функцию $f_1(t)$, выражаемую аналогичным интегралом, но в конечных пределах

$$f_1(t) = \frac{1}{\pi} \int_{\omega_1}^{\omega_2} (A_t \cos \omega t + B_t \sin \omega t) d\omega$$

и потребуем, чтобы разность $\Delta f = f(t) - f_1(t)$ была (с точки зрения некоторого определенного критерия) достаточно мала, т. е. чтобы функция $f_1(t)$ могла служить удовлетворительным приближением к $f(t)$. Интервал (ω_1, ω_2) будет при этом определен, как активная полоса спектра.

В цитированной работе [1] рассмотрена в качестве примера функция вида

$$f(t) = \sin \frac{1}{2} \lambda t^2,$$

т. е. ЧМ-колебание, частота которого меняется по линейному закону

$$\omega = \lambda t.$$

Показано, что ширина активной полосы выражается соотношением

$$\omega_2 - \omega_1 = \Delta\omega = b \sqrt{d\omega/dt} = b \sqrt{\lambda},$$

где b — коэффициент тем больший, чем лучше требуется приближение. Показано также, что центральная частота активной полосы совпадает с мгновенной частотой, так что в рассматриваемом случае активная полоса, сохраняя постоянную ширину, скользит по шкале частот, следуя за измерениями мгновенной частоты ω ; этот результат вполне согласуется с нашими интуитивными представлениями.

IV. Разложение спектров по спектрам составляющих функций

Линейность преобразования Фурье, выражаемая соотношением (4.1), позволяет применить общий прием приближенного вычисления спектра, основанный на следующих соображениях.

Пусть данная функция $f(t)$ аппроксимирована конечной суммой некоторых произвольно выбранных функций f_k , т. е.

$$f(t) \approx \sum_{k=1}^N f_k(t).$$

Если спектры функций f_k известны, то, обозначая их через S_k , имеем

$$S = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt \approx \int_{-\infty}^{+\infty} e^{-j\omega t} \left(\sum_{k=1}^N f_k(t) \right) dt = \sum_{k=1}^N \int_{-\infty}^{+\infty} f_k e^{-j\omega t} dt = \sum_{k=1}^N S_k.$$

Это соотношение может служить основой для многочисленных вариантов формул и таблиц для вычисления спектров заданных функций.

Мы ограничимся здесь двумя примерами, относящимися к непериодическим функциям импульсивного характера, в частности к функции, не равной нулю на протяжении времени τ .

Пусть, например, функция существует на промежутке $-\tau/2 < t < \tau/2$. Выберем разложение по тригонометрическим функциям, т. е. разложим $f(t)$ в ряд Фурье на указанном промежутке. Мы получим

$$f(t) \approx \sum_{-N}^N C_k e^{j2\pi k t / \tau},$$

где

$$C_k = \frac{1}{\tau} \int_{-\tau/2}^{\tau/2} f(t) e^{-j2\pi k t / \tau} dt.$$

Таким образом, в этом примере элементарная функция есть

$$f_k = C_k e^{j2\pi k t / \tau}.$$

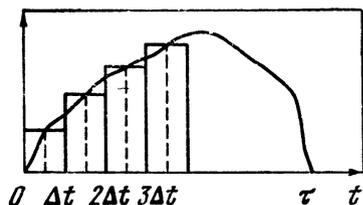
Спектр этой функции имеет вид

$$S_k = C_k \int_{-\tau/2}^{\tau/2} e^{-j(\omega - 2\pi k/\tau)t} dt = \frac{(-1)^k 2C_k}{\omega - 2\pi k/\tau} \sin \omega\tau/2,$$

и, следовательно, искомый спектр есть

$$S \approx 2 \sin \omega \frac{\tau}{2} \sum_{-N}^N \frac{(-1)^k C_k}{\omega - 2\pi k/\tau}.$$

Приближение будет, очевидно, тем лучшим, чем больше число слагаемых суммы.



Р и с. 61

В качестве второго примера применим аппроксимацию заданной функции ступенчатой ломаной линией, разбив интервал $0 < t < \tau$ на n частей Δt , т. е. положив $\tau = n\Delta t$ и выбрав элементарную функцию в форме прямоугольного импульса. Тогда можем записать (рис. 61)

$$f(t) \approx \sum_{k=0}^{n-1} \{\sigma_0 [t - k\Delta t] - \sigma_0 [t - (k+1)\Delta t]\} f[(k + 1/2)\Delta t].$$

Элементарный импульс равен

$$f_k(t) = f\left[\left(k + \frac{1}{2}\right)\Delta t\right] \{\sigma_0 [t - k\Delta t] - \sigma_0 [t - (k+1)\Delta t]\}.$$

Его спектр имеет вид

$$S_k = f\left[\left(k + \frac{1}{2}\right)\Delta t\right] \int_{k\Delta t}^{(k+1)\Delta t} e^{-j\omega t} dt = \frac{2}{\omega} e^{-j\omega(k+1/2)\Delta t} \sin \omega\Delta t/2,$$

следовательно, искомый спектр имеет форму

$$S = \Delta t \frac{\sin \omega\Delta t/2}{\omega\Delta t/2} \sum_{k=0}^{n-1} f\left[\left(k + \frac{1}{2}\right)\Delta t\right] e^{-j\omega(k+1/2)\Delta t}.$$

Приближение улучшается с возрастанием n . В пределе при $n \rightarrow \infty$, т. е. при $\Delta t \rightarrow 0$, множитель $\frac{\sin \omega\Delta t/2}{\omega\Delta t/2}$ обращается в еди-

ницу, а сумма переходит в интеграл Фурье, дающий точное значение спектра функции $f(t)$, т. е.

$$S = \int_0^{\tau} f(t) e^{-j\omega t} dt.$$

V. Спектр короткого знакопеременного импульса

Высказанное в § 11 положение о спектре короткого импульса, состоящее в том, что спектр такого импульса однороден вплоть до частот, период которых сравним с длительностью импульса, требует уточнения. Дело в том, что это положение справедливо, если на протяжении длительности импульса функция, которой он выражается, не меняет знака, т. е., например, если на промежутке $(-\tau/2, \tau/2)$ импульс определен четной функцией. Если же это условие не выполняется, т. е. если, например, на указанном промежутке импульс определен нечетной функцией, то положение радикально меняется.

Мы имеем для спектра импульса

$$S = \int_{-\tau/2}^{\tau/2} f(t) e^{-j\omega t} dt.$$

Разобьем интеграл на два, и пусть функция $f(t)$ нечетна. Тогда

$$\begin{aligned} S &= \int_{-\tau/2}^0 f(t) e^{-j\omega t} dt + \int_0^{\tau/2} f(t) e^{-j\omega t} dt = \int_0^{\tau/2} f(t) (e^{j\omega t} - e^{-j\omega t}) dt = \\ &= 2j \int_0^{\tau/2} f(t) \sin \omega t dt \approx 2j\omega \int_0^{\tau/2} t f(t) dt = j\omega M, \end{aligned}$$

где M — момент первого порядка функции $f(t)$ относительно начала. Приближение законно, если $\omega\tau/2 \ll 1$. Итак, до частот, определяемых сравнимостью периода с длительностью, спектр знакопеременного импульса выражается линейной функцией частоты.

В качестве примера сравним спектры отрезков косинусоиды и синусоиды. Для косинусоидального импульса имеем

$$S = 2 \int_0^{\tau/2} \cos \Omega t \cos \omega t dt = \frac{\sin(\Omega + \omega)\tau/2}{\Omega + \omega} + \frac{\sin(\Omega - \omega)\tau/2}{\Omega - \omega} \approx \tau.$$

Для синусоидального импульса

$$S = -2j \int_0^{\tau/2} \sin \Omega t \sin \omega t dt = 2j \left(\frac{\sin(\Omega + \omega)\tau/2}{\Omega + \omega} - \frac{\sin(\Omega - \omega)\tau/2}{\Omega - \omega} \right).$$

Здесь мы должны для перехода к пределу при убывающем τ воспользоваться уже двумя членами разложения синуса в степенной ряд. Это дает

$$2j \left(\frac{\sin(\Omega + \omega)\tau/2}{\Omega + \omega} - \frac{\sin(\Omega - \omega)\tau/2}{\Omega - \omega} \right) \approx \\ \approx j \frac{\tau^3}{4 \cdot 3!} [(\Omega + \omega)^2 - (\Omega - \omega)^2] = j\omega \frac{\tau^3}{3!} \Omega.$$

Для пояснения полученных соотношений можно еще сослаться на то, что всякая нечетная функция может рассматриваться как производная некоторой четной функции. Но спектр производной получается путем умножения спектра функции на $j\omega$ (см. (4. 2)). Следовательно, если спектр короткого четного импульса однороден, т. е. выражается постоянной величиной, то спектр нечетного импульса должен содержать множитель $j\omega$.

Наконец, следует указать еще на спектры разрывных функций (§11). Спектр единичного импульса $\delta(t)$, являющегося четной функцией, равен единице. Спектр же импульса второго порядка $\frac{d}{dt}\delta(t)$, являющегося нечетной функцией и характеризующегося двумя последовательными скачками в $+\infty$ и $-\infty$ около $t=0$, равен $j\omega$.

Если короткий импульс совершает несколько перемен знака, то его спектр может выражаться и более высокими степенями частоты.

VI. Подробности вычисления Δf и Δt

Для читателя, который пожелал бы проверить выкладки § 12 или проделать аналогичные, ниже приведены некоторые детали этих выкладок.

а. Прямоугольный импульс. $\Delta t = 0,9\tau$ — это очевидно. Далее,

$$\Phi = \tau \frac{\sin \omega\tau/2}{\omega\tau/2},$$

$$\int_0^{\Delta\omega} \Phi^2 d\omega = \tau^2 \int_0^{\Delta\omega} \frac{\sin^2 \omega\tau/2}{(\omega\tau/2)^2} d\omega = 2\tau \int_0^{\Delta\omega\tau/2} \frac{\sin^2 z}{z^2} dz.$$

Здесь и ниже встречаются интегралы вида

$$\int \frac{\sin \cos z}{z^n} dz,$$

которые вычисляются по частям, пока не закончатся на слагаемом $\text{Si}(z)$ или $\text{Ci}(z)$. В результате мы получаем для $\Delta\omega\tau$ уравнение

$$\frac{2}{\pi} \left[\text{Si}(\Delta\omega\tau) - \frac{\sin^2 \Delta\omega\tau/2}{\Delta\omega\tau/2} \right] = \eta = 0,9,$$

решение которого, легко получаемое методом ложных корней, есть $\Delta \omega \tau = 5,1$.

б. *Треугольный импульс.* Для треугольного импульса из условия

$$2 \int_0^{\Delta t/2} (1 - 2t/\tau)^2 dt = \eta \tau/3$$

получаем уравнение

$$(\Delta t/\tau)^3 + 3(\Delta t/\tau)^2 - 3\Delta t/\tau = \eta = 0,9,$$

решая которое, находим

$$\Delta t/\tau = 0,541.$$

Спектр треугольного импульса есть

$$\Phi = \tau \frac{1 - \cos \omega \tau/2}{(\omega \tau/2)^2}.$$

Составляя выражение

$$\tau^2 \int_0^{\Delta \omega} \frac{(1 - \cos \omega \tau/2)^2}{(\omega \tau/2)^4} d\omega = \eta \frac{2}{\pi} \tau$$

и выполняя интегрирование, получим уравнение

$$\frac{4}{\pi} \left\{ \text{Si}(x) - \frac{1}{2} \text{Si}\left(\frac{x}{2}\right) - \frac{2}{x^3} \left[3 + \left(\cos x - 4 \cos \frac{x}{2} \right) - \right. \right. \\ \left. \left. - \frac{x}{2} \left(\sin x - 2 \sin \frac{x}{2} \right) - \frac{x^2}{2} \left(\cos x - \cos \frac{x}{2} \right) \right] \right\} = 0,9,$$

решение которого есть

$$x = \Delta \omega \tau = 5,3.$$

в. *Косинусоидальный импульс.* Для определения Δt имеем соотношение

$$2 \int_0^{\Delta t/2} \cos^2 \pi \frac{t}{\tau} dt = \eta \tau/2,$$

что приводит к уравнению

$$\frac{\Delta t}{\tau} + \frac{1}{\pi} \sin \pi \frac{\Delta t}{\tau} = 0,9.$$

Отсюда находим

$$\Delta t/\tau = 0,596.$$

Спектр косинусоидального импульса есть

$$\Phi = \frac{\pi}{2} \tau \frac{\cos \omega \tau/2}{(\pi/2)^2 - (\omega \tau/2)^2},$$

и выражение для определения $\Delta\omega$ принимает вид

$$\frac{\pi^2}{4} \tau^2 \int_0^{\Delta\omega} \frac{\cos^2 \omega\tau/2}{[(\pi/2)^2 - (\omega\tau/2)^2]^2} d\omega = \eta \frac{\pi}{2} \tau.$$

Входящий сюда интеграл вычисляется путем разложения подынтегрального выражения на простые дроби. В результате интегрирования получаем уравнение

$$\frac{1}{\pi} \left[-\frac{1 - \cos x}{x} + \text{Si}(x) + \frac{1}{\pi} \ln x - \frac{1}{\pi} \text{Ci}(x) \right] \Big|_{x_1}^{x_2} = 0,9,$$

где $x_1 = \pi - \Delta\omega\tau$, $x_2 = \pi + \Delta\omega\tau$.

Здесь возникает затруднение: x_1 , несомненно, отрицательно, и мы не можем подставлять его в качестве аргумента ни под знак \ln , ни под знак Ci . Это затруднение, однако, очень легко преодолеть. Рассмотрим функцию

$$\varphi(x) = \ln x - \text{Ci}(x).$$

Вспоминая известное разложение $\text{Ci}(x)$ в степенной ряд, получим

$$\varphi(x) = -E + \frac{x^2}{2 \cdot 2!} - \frac{x^4}{4 \cdot 4!} + \frac{x^6}{6 \cdot 6!} - \dots,$$

где E — эйлерова постоянная. Таким образом, оказывается, что $\varphi(x)$ — четная функция, и, следовательно, мы имеем право при выполнении вычислений подставлять под знаки \ln и Ci вместо отрицательных аргументов положительные. Принимая, кроме того, во внимание, что $\text{Si}(x)$ есть функция нечетная, и проделывая все вычисления, находим $\Delta\omega\tau = 4,57$.

VII. По поводу общего критерия для оценки Δf и Δt

Использованный в §12 общий критерий для оценки Δf и Δt , основанный на понятии о радиусе инерции плоской фигуры, оказывается очень острым, и применение его не всегда возможно.

Вычислим на основе этого критерия произведение $\Delta f \Delta t$ для треугольного импульса, для которого

$$f(t) = 1 - 2|t|/\tau,$$

$$\Phi(\omega) = \frac{\tau}{2} \cdot \frac{1 - \cos \omega\tau/2}{1/2(\omega\tau/2)^2}.$$

Находим

$$A_t = 2 \int_0^{\tau/2} \left(1 - 2\frac{t}{\tau}\right)^2 dt = \frac{13}{12} \tau, \quad \Delta t^2 = I_t/A_t = 9,43 \cdot 10^{-2} \tau^2,$$

$$M_t = 0, \quad \Delta t = 0,307 \tau,$$

$$I_t = 2 \int_0^{\tau/2} t^2 (1 - 2t/\tau)^2 dt = \frac{49}{480} \tau^3.$$

Теперь найдем ширину спектра

$$A_{\omega} = \pi A_t = 3,4\tau,$$

$$M_{\omega} = \frac{\tau^2}{4} \int_0^{\infty} \frac{(1 - \cos \omega\tau/2)^2}{1/4 (\omega\tau/2)^4} d\omega = 4 \int_0^{\infty} \frac{\sin^4 x}{x^3} dx = \frac{8}{3},$$

$$I_{\omega} = \frac{\tau^2}{4} \int_0^{\infty} \omega^2 \frac{(1 - \cos \omega\tau/2)^2}{1/4 (\omega\tau/2)^4} d\omega = \frac{16}{\tau} \int_0^{\infty} \frac{\sin^2 x}{x^2} dx = \frac{8\pi}{\tau}.$$

Таким образом,

$$\Delta\omega^2 = I_{\omega}/A_{\omega} - (M_{\omega}/A_{\omega})^2 = 6,8/\tau^2,$$

$$\Delta\omega = 2,6/\tau, \quad \Delta f = 0,415/\tau$$

и, наконец, $\Delta f \Delta t = 0,415 \cdot 0,307 = 0,127$, т. е. результат примерно втрое больше теоретического минимума.

Если мы теперь попробуем проделать аналогичные вычисления для прямоугольного импульса, то для Δt мы легко найдем значение, равное $\tau/\sqrt{12} = 0,288\tau$. Что же касается Δf , то интегралы для M_{ω} и I_{ω} оказываются расходящимися.

Это объясняется, вообще говоря, тем, что прямоугольный импульс характеризуется разрывом самой функции, и его спектр убывает лишь как $1/\omega$. Рассмотренный же в начале треугольный импульс характеризуется разрывом первой производной, и, следовательно, его спектр убывает как $1/\omega^2$, благодаря чему интегралы оказываются сходящимися.

Л и т е р а т у р а

1. Агеев Д. В. Активная полоса частотного спектра функции времени. Тр. Горьковск. политехн. ин-та, 11, 1, 1955, 5—10.
2. Андреев Н. Н. Электрические колебания и их спектры. 1917.
3. Бор Г. Почти-периодические функции. Гостехиздат, 1934.
4. Бунимович В. И. Флюктуационные процессы в радиоприемных устройствах. «Советское радио», 1951.
5. Быкова Н. О. Воздействие напряжения меняющейся частоты на резонансные системы. Тр. МАП СССР, 1956, № 28.
6. Гоноровский И. С. Радиосигналы и переходные явления в радиоцепях. Связьиздат, 1954.
7. Джеймс Х., Никольс Н., Филлипс Р. Теория следящих систем. ИЛ, 1953.
8. Кац А. К. Вынужденные колебания при прохождении через резонанс. Инж. сб. Ин-та механики АН СССР, 3, 2, 1947, 100—125.
9. Колмогоров А. Н. Теория передачи информации. Изд-во АН СССР, 1965.
10. Котельников В. А. О пропускной способности «эфира» и проволок в электросвязи. Материалы к I Всесоюзному съезду по вопросам технической реконструкции дела связи и развития слаботочной промышленности. 1933. Всесоюзный энергетический комитет.
11. Кузьмин Р. О. Бесселевы функции. ОНТИ, 1935.
12. Лев А. Ю., Яхinson Б. И. О смещении спектра сигналов. ЭС, 4, 1956, 63—74.
13. Лоусон Дж., Уленбек Дж. Пороговые сигналы. «Советское радио», 1952.

14. Майер А. Г., Леонтович Е. А. Об одном неравенстве, связанном с интегралом Фурье. Докл. АН СССР, 4, 7, 1934, 353—360.
15. Привалов И. И. Ряды Фурье. Гостехиздат, 1934.
16. Римский-Корсаков А. В. Об анализе колебаний. Труды Комиссии по акустике АН СССР, 6, 1951, 66—81.
17. Римский-Корсаков А. В., Шумова Н. Д. Об одном методе снятия частотных характеристик. Журн. техн. физ., 8, 1938.
18. Серебренников М. Г. Гармонический анализ. Гостехиздат, 1948.
19. Смирнов В. И. Курс высшей математики, т. 2. Гостехиздат, 1948.
20. Турбович И. Т. О погрешности измерения частотных характеристик методом частотной модуляции. Радиотехника, 9, 12, 1954, 31—35.
21. Турбович И. Т. Динамические частотные характеристики избирательных систем. Радиотехника, 12, 11, 1957, 39—49.
22. Штейн Б. Б., Варганов Г. Г. Выделение одной боковой полосы с помощью многофазной модуляции. Научн.-техн. сб. МЭИС, 1950, 13—24.
23. Цянь-Сюэ-Сэнь. Техническая кибернетика. ИЛ, 1956.
24. Barber N. F., Ursell F. The response of a resonant system to a gliding tone. Philos. Mag., 39, 1948, 345—361.
25. Blassel P. Erreur due à une durée d'intégration finie dans la détermination des fonctions d'autocorrélation, Ann. Télécommun., 8, 12, 1953, 406—414.
26. Davenport W. B., Johnson R. A., Middleton D. Statistical errors in measurements on random time function. J. Appl. Phys., 23, 4, 1952, 377—388.
27. Fano R. M. Short time autocorrelation functions and power spectra. JASA, 22, 5, 1950, 546—550.
28. Hok G. Response of linear systems to excitation of a frequency varying linearly with time. J. Appl. Phys., 19, 3, 1948, 242—250.
29. Lampard D. G. Definitions of «bandwidth» and «time duration» of signals which are connected by an identity. Trans. IRE, CT-3, 4, 1956, 286—288.
30. Lampard D. G. Generalization of the Wiener-Khintchine theorem to non-stationary processes. J. Appl. Phys., 25, 6, 1954, 802—803.
31. Lewis E. M. Vibration during acceleration through a critical speed. Trans. Amer. Soc. Mech. Engrs., 54, 1932, 24, 253—261.
32. Page C. H. Instantaneous power spectra. J. Appl. Phys., 23, 1, 1952, 103—106.
33. Plotter R. K., Kopp G. A., Green H. C. Visible speech. N. Y., Van Nostrand, 1947.
34. Rice S. O. Filtered thermal noise fluctuation of energy as a function of interval length. JASA, 14, N 4, 1943, 216—227.
35. Rice S. O. Mathematical analysis of random noise. BSTJ, 23, 1944, N 3, 282—332, 24, 1, 1945, 46—156. Русский перевод см. в сборнике «Теория передачи электрических сигналов при наличии помех». ИЛ, 1953.
36. Spletner L. M. Errors in power spectra due to finite sample. J. Appl. Phys., 25, 5, 1954, 653—659.
37. Zadeh L. A. Correlation function and spectra of phase- and delay-modulated signals. Proc. IRE, 39, 4, 1951, 425—427.

АВТОКОЛЕБАНИЯ

ПРЕДИСЛОВИЕ

Общеизвестна ведущая роль советской науки в области нелинейных колебаний и, в частности, автоколебаний.

Мы имеем превосходные, единственные в своем роде книги: «Теория колебаний» А. А. Андропова и С. Э. Хайкина, «Автоколебательные системы» К. Ф. Теодорчика и другие.

Но и предлагаемая небольшая книга может оказаться нелишней, так как она отличается от названных сочинений как целевой установкой, так и способом изложения.

Цель этой книги — по возможности просто и наглядно изложить основы автоколебаний, освободив предмет от трудностей, связанных с количественным его исследованием. Другая цель — описать с единой точки зрения различные виды автоколебаний, имеющих значение и интерес для физики и техники.

В книге приведены некоторые подробности автоколебательных явлений, иногда малоизвестных. Разумеется, невозможно выдерживать высокий профессиональный уровень в каждой из затрагиваемых весьма различных областей — на это книга и не претендует.

Книга представляет собою попытку вывести общие качественные представления об автоколебаниях на основе рассмотрения многочисленных примеров. Качественное исследование может обойтись без аналитического аппарата. Специальный аппарат нелинейной теории колебаний иногда отпугивает читателя. Между тем, основные понятия сравнительно просты. Поэтому автор решился попробовать написать книгу «без математики» — в основном тексте нет ни одной формулы.

Если эта попытка будет способствовать распространению знаний об автоколебаниях, цель автора будет достигнута,

ВВЕДЕНИЕ

Если присмотреться к окружающему, то оказывается, что автоколебательные системы самого различного свойства встречаются нам повсюду.

Ламповый генератор, электромагнитный прерыватель, духовые музыкальные инструменты — все это автоколебательные системы; колебания, возникающие в них, — это автоколебания.

Часы — также автоколебательная система. К автоколебательным системам следует также отнести поршневые двигатели (внутреннего сгорания и паровой).

Механические автоколебания, обусловленные трением, необычайно распространены в виде разного рода скрипов. Сюда же относятся и автоколебания режущего инструмента (например, резцов на токарных станках).

Огромное значение имеют в современной технике колебания разного рода регуляторов, которые при известных условиях приобретают свойства автоколебательных систем.

Целый ряд явлений природы имеет автоколебательный характер. Многие процессы в живом организме представляют собою автоколебания. Например, дыхание и работа сердца — автоколебания.

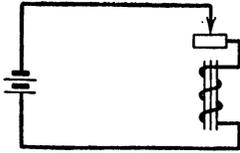
Из этого беглого перечня уже можно составить себе представление о широком значении автоколебаний не только для техники, но и вообще для естествознания. Если можно столь несходные явления объединить термином «автоколебания», то это означает, что эти явления обладают определенными общими чертами и что они могут рассматриваться с некоторой единой точки зрения, и к их изучению может применяться единая методика. Установление общих точек зрения и методики исследования и составляет содержание этой книги. Многие из перечисленных вначале видов автоколебаний будут рассмотрены в качестве примеров.

§ 1. Автоколебательная система

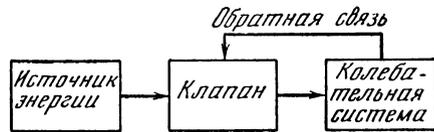
Мы называем автоколебательной системой устройство, преобразовывающее движение постоянного направления в колебательное движение. Термин «движение» применен здесь не в механическом, а в общем смысле; так, например, электрическая автоколебательная система, — скажем, ламповый генератор — преобразует постоянный ток в переменный.

Не всякое устройство, выполняющее такого рода преобразование, является автоколебательной системой. Возьмем к примеру кривошипный механизм. Он может преобразовать равномерное вращение кривошипа в колебательное движение шатуна. Мы не назовем, однако, это устройство автоколебательной системой. Следовательно, первоначальное определение требует уточнения.

Рассматривая внимательно кривошипный механизм, мы отмечаем прежде всего, что оба вида движения — равномерное вращательное и колебательное — связаны в нем жесткой кинематической связью, т. е. каждому положению кривошипа соответствует единственное вполне определенное положение шатуна. Из этого, между прочим, следует, что кривошипный механизм как преобразователь обратим — он может с равным успехом преобразовывать колебательное движение во вращательное, для чего он широко и применяется в технике, например, в поршневых двигателях.



Р и с. 1



Р и с. 2

Аналогичные соотношения имеют место и в альтернаторе — электрической машине, вырабатывающей переменный ток в результате вращения ротора. И альтернатор может быть обращен: если питать его переменным током, то он превращается в синхронный двигатель.

Автоколебательные системы, вообще говоря, необратимы. Этого обстоятельства, однако, еще недостаточно для исчерпывающего определения автоколебательной системы. В поисках такого определения мы рассмотрим действие автоколебательной системы.

Возьмем в качестве примера простейший электромагнитный прерыватель (о нем еще много будет говориться в дальнейшем). Электрическая цепь прерывателя состоит из батареи и обмотки электромагнита (рис. 1). Электромагнит воздействует на якорь, который, колеблясь, размыкает и замыкает контакт. Наличие контакта, управляемого движением якоря, — существеннейшая особенность устройства. Без размыкаемого контакта не могло бы, очевидно, быть никаких автоколебаний.

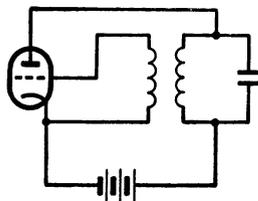
Схематизируя устройство прерывателя, мы отмечаем следующие составные части его:

1. Источники энергии — в данном случае батарея.
2. Клапан, т. е. орган, регулирующий поступление энергии в колебательную систему, — в данном случае контакт.
3. Колебательная система — в данном случае якорь.

Эти части находятся между собою во взаимодействии, состоящем в том, что, с одной стороны, клапан управляет движением колебательной системы, а с другой стороны, движение системы управляет работой клапана.

Такого рода обратное воздействие, чрезвычайно характерное для всякой автоколебательной системы, носит общее название обратной связи. При помощи обратной связи осуществляется то, так сказать, самоуправление, в результате которого могут возникнуть устойчивые незатухающие колебания.

На основании этих соображений можно представить общую схему автоколебательной системы, как показано на рис. 2. Что касается источника энергии, то он входит в состав автоколебательной системы, если мы определим ее как источник колебаний, и не входит в состав автоколебательной системы,



Р и с. 3

если мы определим ее как преобразователь постоянного движения в колебательное. Выбор того или иного определения — дело нашего усмотрения. Остановимся на первом. Тогда определение автоколебательной системы, устанавливающее как ее основную функцию, так и способ действия, будет выглядеть следующим образом.

Автоколебательной системой называется устройство, способное создавать незатухающие колебания и характеризующееся наличием источника энергии, клапана, регулирующего поступление энергии в колебательную систему, и обратной связи с колебательной системой на клапан.

В качестве простого примера для закрепления введенных представлений рассмотрим обычный ламповый генератор. Легко распознать на его схеме (рис. 3) вышеперечисленные основные части автоколебательной системы. Источник энергии — анодная батарея. Колебательная система — контур в анодной цепи. Роль клапана выполняет сетка; она управляет анодным током в соответствии с потенциалом на ней относительно катода. Обратная связь с анодным контуром на сетку осуществляется при помощи катушки связи, индуктивно связанной с контуром. В данном случае все устройство автоколебательной системы видно совершенно отчетливо. Нужно, однако, заранее предупредить, что механизм обратной связи далеко не всегда так очевиден, как в рассмотренных примерах электромагнитного прерывателя и лампового генератора. Иногда обратная связь действует по скрытым каналам; надо уметь ее найти, так как без этого нельзя понять действия автоколебательной системы.

§ 2. Энергетика автоколебаний

Одна из наиболее общих точек зрения, с которой можно рассматривать любое физическое явление, — это энергетическая точка зрения. Стоя на этой точке зрения, мы интересуемся прежде всего энергетическим балансом явления. Конечно, знание энергетического баланса не определяет специфических черт данного частного явления, но зато энергетическое рассмотрение целой группы родственных явлений позволяет установить наиболее общие черты, их объединяющие. А так как это и является нашей ближайшей целью, то энергетическое рассмотрение автоколебательных систем вполне целесообразно.

С энергетической точки зрения действие автоколебательной системы может быть описано следующим образом. Пусть колебания уже происходят. Энергия этих колебаний должна была бы убывать вследствие неизбежных потерь (или вследствие передачи колебательной энергии потребителю). Но убыль энергии пополняется за счет источника, от которого известная порция энергии за каждый период (или вообще периодически) пропускается клапаном в колебательную систему. Совершенно ясно, что установившиеся (т. е. незатухающие и не возрастающие по амплитуде) колебания возможны в том единственном случае, когда поступление энергии от источника за период (или в единицу времени) в точности равно потере энергии за то же время. Это условие баланса энергии и есть условие существования незатухающих колебаний. Также ясно, что если баланс будет нарушен в том смысле, что пополнение потерь будет недостаточным, то колебания будут затухать. Если же, наоборот, энергия будет поступать от источника в избытке, то колебания будут нарастать.

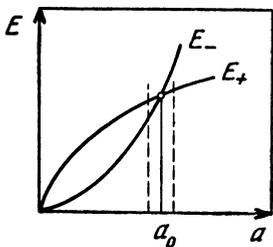
Теперь можно составить себе представление и о механизме самовозбуждения. Если вначале, т. е. при самых малых амплитудах, приход энергии больше расхода, то система будет раскачиваться, т. е. амплитуда колебаний будет возрастать. Если бы это соотношение между получаемой и теряемой энергиями всегда сохранялось, то амплитуда росла бы неограниченно. Таким образом, для получения установившегося режима необходимо, чтобы, начиная с некоторого значения амплитуды, потери росли с дальнейшим увеличением амплитуды быстрее, чем пополнение, получаемое от источника. При этом условии возможно достижение энергетического баланса, о котором говорилось раньше.

Все это очень наглядно поясняется графиком рис. 4, на котором энергия, получаемая от источника и обозначенная через E_+ , и энергия, теряемая колебательной системой, обозначенная через E_- , изображены в зависимости от амплитуды a ¹.

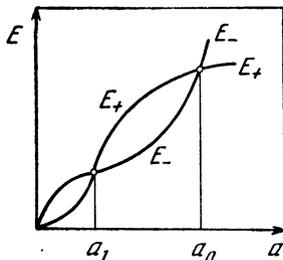
¹ Этот график не относится к какому-либо конкретному устройству. Предполагается только, что E_+ и E_- нарастают с увеличением амплитуды; это является в общем вполне правдоподобным предположением.

Точка пересечения кривых E_+ и E_- отвечает балансу энергий; абсцисса этой точки соответствует значению установившейся амплитуды a_0 . Диаграмма рис. 4 дает также сразу простейшее понятие об устойчивости колебательного движения.

Напомним, что устойчивым равновесием системы называется такое ее состояние, в которое она стремится вернуться, будучи из него выведена, а затем предоставлена самой себе. Неустойчивым равновесием, напротив, называется состояние, в котором система не может удержаться, так как при малейшем отклонении от этого состояния она стремится еще дальше уйти от него ¹. Эти



Р и с. 4



Р и с. 5

понятия и определения статики могут быть перенесены и на состояние движения, в частности колебательного. Мы будем называть движение устойчивым, если система, колебательный режим которой нарушен каким-либо внешним воздействием, возвращается к первоначальному режиму после прекращения этого воздействия. Наоборот, движение будет неустойчивым, если сколь угодно малое внешнее воздействие выбивает систему из колебательного режима, от которого система в дальнейшем все более отходит. Подчеркивая аналогию между устойчивостью равновесия и устойчивостью движения, мы будем называть первую — статической устойчивостью, вторую — динамической устойчивостью ².

Посмотрим теперь, устойчиво ли движение в условиях рис. 4.

Для проверки устойчивости возьмем два значения амплитуды меньше и больше a_0 . Легко видеть, что в первом случае (см. штриховые вертикальные линии на рис. 4) $E_+ > E_-$ и, следовательно, амплитуда колебаний будет нарастать, пока не достигнет значения a_0 . Во втором случае $E_+ < E_-$ и, следовательно, амплитуда

¹ Существует еще один вид равновесия: безразличное равновесие. Оно характеризуется тем, что все положения системы являются равновесными и что система, следовательно, может оставаться в любом положении, будучи предоставлена самой себе. Простейшими моделями, поясняющими различие трех видов равновесия, могут служить: тяжелый шарик на вогнутой поверхности, тот же шарик на выпуклой поверхности и тот же шарик на горизонтальной плоскости.

² Строгие определения устойчивости, данные Ляпуновым, нам пока не необходимы.

будет убывать, пока опять-таки не достигнет значения a_0 . Таким образом, точка пересечения с абсциссой a_0 есть точка устойчивого динамического равновесия.

Заметим, что начало координат, где пересекаются кривые E_+ и E_- , есть точка статического равновесия.

Однако это равновесие неустойчиво, в чем легко убедиться, рассмотрев соотношения при сколь угодно малой, но не равной нулю амплитуде. Таким образом, условие самовозбуждения автоколебательной системы есть не что иное, как условие неустойчивости ее в состоянии покоя.

Кривые E_+ и E_- могут пересекаться и не в одной точке. Рассмотрим график рис. 5. Здесь имеются три точки пересечения — три состояния равновесия. Рассуждая аналогично предыдущему, мы легко установим, что два из них устойчивы (абсциссы 0 и a_0), а третье — неустойчиво (абсцисса a_1). Из того, что начало координат есть точка устойчивого равновесия, следует, что система, предоставленная самой себе, будет сохранять состояние покоя, т. е. не будет самовозбуждаться. Однако это не значит, что в такой системе устойчивые автоколебания невозможны. Действительно, представим себе, что каким-либо внешним толчком в системе возбуждены колебания с амплитудой, превосходящей a_1 . Если это случится, то в дальнейшем система уже раскачается самостоятельно до амплитуды a_0 .

Этот случай, т. е. случай, когда требуется определенный конечной величины начальный толчок для того, чтобы вызвать самораскачивание автоколебательной системы, носит название ж е с т к о г о с а м о в о з б у ж д е н и я. В противоположность этому случай, соответствующий рис. 4, когда для раскачивания системы достаточно сколь угодно малого начального возмущения, всегда достаемого теми или иными флуктуациями, — называется м я г к и м с а м о в о з б у ж д е н и е м¹.

Кривые E_+ и E_- могут пересекаться и в большем числе точек. Все точки устойчивого равновесия (за исключением начала) означают возможные устойчивые режимы системы, в которых она может генерировать незатухающие колебания.

На основании приведенных рассуждений можно сделать еще одно заключение принципиального характера. Если бы система была л и н е й н а, то как потери, так и прибыль энергии росли бы по одинаковому закону, а именно, пропорционально квадрату амплитуды. Графически E_+ и E_- были бы представлены двумя параболой², проходящими через начало координат, но больше

¹ Пользуясь моделью, упомянутой в сноске на стр. 258, можно уподобить систему с несколькими чередующимися точками устойчивого и неустойчивого равновесия тяжелому шарик, лежащему на волнистой поверхности. Чтобы шарик мог попасть из одной впадины в другую, нужно помочь ему перебраться через разделяющий эти впадины гребень.

² Всякое отклонение от линейности проявляется в отклонении графика энергии от параболического закона. Можно предложить видоизменение графика энергии, состоящее в том, что по оси абсцисс откладывается не ам-

нигде не пересекающимися. Одна из этих парабол, следовательно, всеми своими точками лежит выше другой. Либо система не возбуждается вовсе, либо она мягко самовозбуждается, но амплитуда колебаний неограниченно растет. Отсюда следует, что режим установившихся колебаний конечной амплитуды возможен только в **нелинейной** системе.

Таким образом, определение автоколебательной системы, данное в § 1, следует дополнить в том смысле, что в состав всякой реальной автоколебательной системы, способной генерировать колебания с установившейся амплитудой, непременно входит нелинейный элемент. При этом нужно заметить, что нелинейность может быть присуща как колебательной системе или нагрузке, так и механизму клапана, или цепи обратной связи, — одним словом, нелинейными свойствами может обладать любое из звеньев, входящих в состав автоколебательной системы.

Итак, мы видим, что даже самый поверхностный энергетический анализ автоколебаний снабжает нас сразу целым ассортиментом важных понятий и определений.

Нам нужно теперь ввести некоторые приемы более детального описания энергетических соотношений при автоколебаниях.

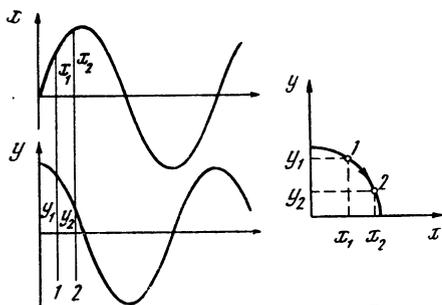
§ 3. Диаграммы работы

Колебательный процесс можно описать, изобразив периодическую зависимость от времени одной или нескольких переменных величин, в изменении которых и состоит данный процесс. Так, например, механический колебательный процесс может быть определен через периодическое изменение смещения, или скорости, или ускорения, или силы реакции и т. п. Можно снять экспериментально **осциллограммы** названных величин. Они будут представлять собой периодические кривые в прямоугольных координатах; по оси ординат откладывается мгновенное значение измеряемой величины, по оси абсцисс — время. Осциллограммы могут дать нам достаточно полное представление о колебательном процессе. Но процесс этот может быть представлен и иначе.

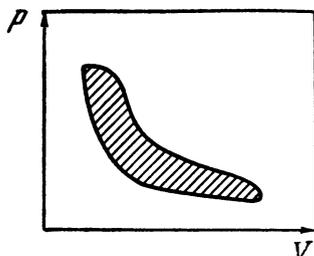
Положим, что нам известно, как изменяются со временем смещение и скорость, т. е. пусть обе эти величины выражены как функции времени. Исключим время и выразим одну колебательную величину функцией другой величины. Иначе говоря, каждому мгновенному значению смещения сопоставим мгновенное значение скорости для одного и того же момента времени. Затем построим график движения в прямоугольных координатах, в которых по одной оси отложена скорость, а по другой — смещение.

плитуда, а ее квадрат. Тогда для линейной системы график энергии будет представляться **прямой** линией и всякое нарушение линейности будет проявляться в отклонении графика от прямой, что и удобно, и привычно.

Полученное таким образом изображение отличается тем, что явная зависимость от времени на нем отсутствует, хотя каждая точка графика отвечает определенному моменту времени на протяжении периода колебания. Если процесс периодический, т. е. если по прошествии периода явление в точности повторяется, — а только такие процессы нас пока и интересуют, — то вышеописанный график будет представлять собою замкнутую кривую (или ломаную) линию. Лежащая на этой кривой точка изображает состояние системы в некоторый определенный момент — она



Р и с. 6



Р и с. 7

так и называется: «изображающая точка». С течением времени изображающая точка будет двигаться по замкнутому контуру, представляющему колебание, и по истечении периода совершает один полный оборот. Координатная плоскость в такого рода координатах, — когда по осям отложены две какие-либо колебательные величины, — носит название фазовой плоскости, а графики движения в этих координатах называются диаграммами на фазовой плоскости, или фазовыми диаграммами.

Исключительно приспособленным для непосредственного наблюдения фазовых диаграмм прибором является обычный электронный осциллоскоп. Для получения на его экране фазовой диаграммы колебательного процесса достаточно подвести к каждой из двух пар отклоняющих пластин напряжение, пропорциональное той или другой колебательной величине.

Уравнение фазовой диаграммы может быть получено аналитически путем исключения времени из двух выражений, дающих зависимость двух каких-либо колебательных величин от времени.

Эта же операция с удобством выполняется графически при наличии двух осциллограмм. Проводя ряд вертикальных линий, мы отмечаем значения двух переменных, соответствующие одному и тому же моменту времени, а затем переносим измеренные отрезки в качестве прямоугольных координат точки на фазовую диаграмму. Это построение показано на рис. 6.

Для получения фазовых диаграмм часто выбирают пару переменных так, чтобы одна являлась производной по времени

от другой (например, смещение и скорость, заряд и ток и т. п.).

Для наших целей удобнее выбрать переменные по другому принципу, а именно так, чтобы энергетические соотношения получили наглядное графическое представление. Мы достигнем этого, если выберем пару переменных, произведение которых выражает энергию, или работу. Этому требованию удовлетворяют в механическом случае пара сила—смещение, в электрическом случае пара напряжение—заряд. Откладывая по осям значения такой пары величин, мы получим, что энергия, или работа, выражается некоторой площадью на диаграмме. На этом основании мы будем в дальнейшем называть такого рода диаграммы *д и а г р а м м а м и р а б о т ы*. Они не представляют собою ничего нового. Обычные так называемые индикаторные диаграммы, снимаемые при помощи весьма простых приборов при испытании поршневых двигателей, представляют собою не что иное, как диаграммы работы. Индикатор, применяемый при таких испытаниях, представляет собой пишущий прибор, барабан которого вращается вместе с валом двигателя, а движение пера вдоль образующей барабана осуществляется манометрическим устройством, так что смещение пера пропорционально давлению в цилиндре. Таким образом, на накрутой на барабан бумаге наносится график в прямоугольных координатах: по одной оси (вдоль окружности барабана) отсчитывается изменение объема V (пропорциональное смещению поршня), а по другой оси — давление p . Типичная индикаторная диаграмма дизеля показана на рис. 7. Площадь внутри замкнутого контура дает непосредственно работу, совершаемую двигателем за один цикл. Координаты p и V — давление и объем — обладают требуемым свойством: их произведение имеет размерность энергии.

На нескольких простейших механических примерах мы выясним теперь основные свойства диаграмм работы. В качестве переменных выберем силу F и перемещение x . Под F будем понимать внешнюю силу, действующую на систему. Работа этой силы будет представлять собою вкладываемую в систему энергию. Условимся присваивать силе и перемещению одинаковые знаки, если сила действует в направлении вызываемого ею движения. Тогда работа внешней силы будет иметь положительный знак. Изобразим на диаграмме работы процесс, состоящий в периодическом возвратно-поступательном движении некоторого тела при наличии силы трения, величина которой постоянна (не зависит от скорости), но направлена, конечно, в сторону, противоположную движению. Сила трения тормозит движение; внешняя сила должна ее преодолеть. При этих условиях диаграмма работы выглядит, как показано на рис. 8. При движении вправо мы должны приложить силу $+F$, при перемене направления движения мы должны изменить и направление силы; величина силы по нашему предположению остается неизменной. Площадь прямоуголь-

ника, описываемого изображающей точкой, выражает работу, совершаемую внешней силой за цикл движения.

Рассмотрим теперь процесс, состоящий в периодическом поднятии и опускании некоторого груза. В этом случае внешняя сила постоянна как по величине, так и по направлению: она преодолевает неизменную силу тяжести. На рис. 9 изображена соответствующая диаграмма. Как видим, она вырождается в отрезок прямой, пробегаемый изображающей точкой туда и обратно. Замкнутый контур выродился в линию; площадь, охватываемая таким вырожденным контуром, равна нулю; нулю равна и совершаемая за цикл работа. Действительно, работа, затрачиваемая при подъеме груза, полностью возвращается источнику энергии при опускании груза.

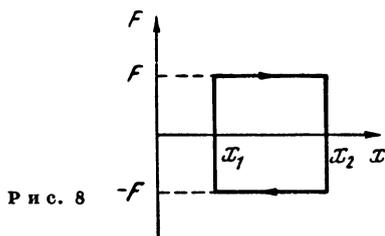


Рис. 8

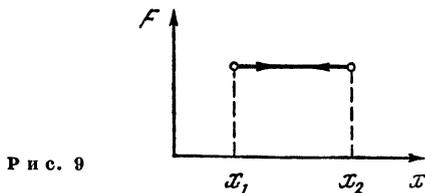


Рис. 9

В качестве третьего примера рассмотрим периодическое растяжение пружины. Соответствующая диаграмма работы показана на рис. 10. И в этом случае работа за цикл равна нулю, так как работа, затраченная на растяжение пружины и запасенная в ней в форме потенциальной энергии деформации, возвращается обратно при сокращении пружины. Заметим, что это обстоятельство никак не связано с линейностью пружины; зависимость силы от удлинения пружины может выражаться какой угодно кривой. Суть дела заключается в том, что если пружина и нелинейна, но идеально упруга в том смысле, что при ее деформации не происходит ведущих к потерям необратимых процессов, то растяжение и сокращение такой пружины будут изображаться одной и той же линией, пробегаемой изображающей точкой в противоположных направлениях.

Сопоставляя рис. 9 и 10 с рис. 8, мы можем сделать очень важное заключение, а именно: при однозначной зависимости силы от смещения работа совершена быть не может. Для того чтобы работа могла быть произведена, необходима не меньше чем двузначная зависимость силы от смещения. Однозначной зависимостью мы называем такую, при которой каждому значению смещения, т. е. каждому положению системы, отвечает вполне определенное и притом единственное значение силы, независимо от направления движения (как на рис. 9 и 10). При двузначной зависимости некоторому положению системы отвечают два зна-

чения силы: одно для движения вперед, другое для движения назад (рис. 8). Понятно, что можно говорить как о двузначной зависимости силы от смещения, так и, наоборот, о двузначной зависимости смещения от силы. При двузначной зависимости изображающая точка описывает замкнутый контур, охватывающий площадь, не равную нулю, и, следовательно, совершается определенная работа.

В качестве примера упомянем еще кривую намагничивания ферромагнетика. Если гистерезис отсутствует, то мы имеем кривую линию. Если же имеется гистерезис, то при циклическом перемагничивании образца получается двузначная зависимость индукции B от намагничивающего поля H ; эта зависимость выражается общеизвестной гистерезисной петлей, площадь которой выражает потери на гистерезис за один цикл перемагничивания.

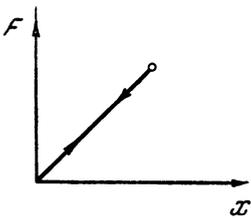


Рис. 10

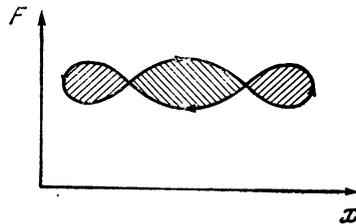


Рис. 11

Еще два замечания. Во-первых, от направления обхода зависит знак работы. При выбранных нами переменных и с соблюдением правила знаков получается, что обход контура по часовой стрелке соответствует работе, затрачиваемой источником и вкладываемой в систему. Противоположное направление обхода означало бы, что энергия поступает из системы в источник, т. е. возвращается источнику. В сущности рис. 9 и 10 можно толковать как наложение двух одинаковых контуров с противоположными направлениями обхода.

Во-вторых, зависимость F от x может быть и многозначной. Тогда диаграмма работы может принять вид, как на рис. 11, образуя несколько замкнутых контуров. В этом случае работа выражается разностью площадей всех контуров, обгаемых по часовой стрелке, и всех контуров, обгаемых в противоположном направлении.

Мы применим теперь диаграммы работы к исследованию различных автоколебательных систем, рассуждая так: потери колебательной энергии или передача ее тому или иному потребителю существуют всегда. Следовательно, для того чтобы автоколебания были возможны, необходима передача энергии от источника в систему. Диаграммы работы покажут нам, происходит ли это и при каких условиях.

§ 4. Статическая двузначность

Как мы увидим дальше, целесообразно различать два механизма автоколебаний. Один из них характеризуется тем, что требуемая двузначность зависимости F от x содержится в статической характеристике автоколебательной системы, т. е. может быть обнаружена при сколь угодно медленном изменении положения системы. Двузначность такого рода мы назовем статической двузначностью. О другом возможном механизме речь будет дальше.

Рассмотрим в качестве примера автоколебательной системы со статической двузначностью электрические часы старинного, но исправно действующего и применяемого поныне устройства. Устройство это состоит в следующем (рис. 12).

К маятнику 1 часов приделан стальной якорь 2, притягиваемый электромагнитом 3. На стержне маятника укреплен палец 4. Электрическая цепь часов состоит из электромагнита 3, контактных пружин 6 и батареи 7. На нижней из контактных пружин 6 подвешена свободно собачка 5.

Действие этого механизма поясняется рис. 13, на котором изображен ряд последовательных расположений деталей, отмеченных на рис. 12 штриховым кружком.

1. Маятник, а с ним и палец движутся влево. Палец еще не коснулся собачки.

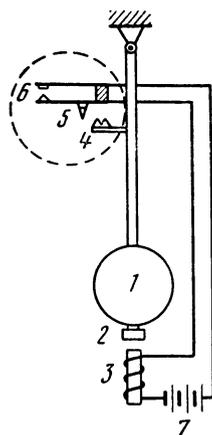
2. Палец коснулся собачки и отклоняет ее своей левой боковой частью.

3. Крайнее левое положение пальца. Собачка заскочила в зубец на пальце.

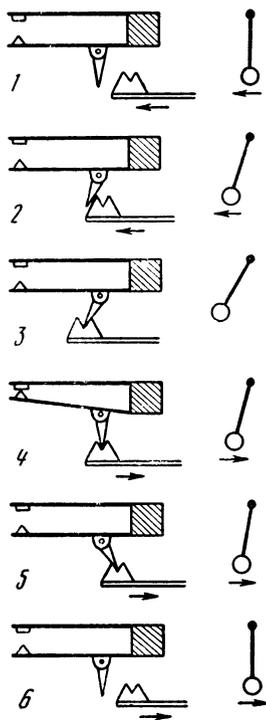
4. Палец идет вправо. Заскочившая в зубец собачка поворачивается принудительно, давит снизу на контактную пружину, замыкая контакт.

5. Палец уходит вправо. Собачка принимает наклонное положение, и контакт размыкается.

6. Палец и собачка занимают положение 1. Палец движется вправо.



Р и с. 12



Р и с. 13

Как видим, смысл всего этого устройства состоит именно в том, чтобы получить желаемую двузначность (разумеется, изобретатель этих часов едва ли рассуждал о них в таких терминах!). Контакт не замыкается при ходе маятника влево, но он замыкается при ходе маятника вправо. Электромагнит включается, и маятник получает ускоряющий импульс, который и необходим для поддержания незатухающих колебаний. Заметим, что если амплитуда колебаний больше чем нужно, то контакт не замыкается; при этих условиях палец продвигается влево дальше, чем показано на рис. 13 (положение 3), и собачка соскальзывает по заты-

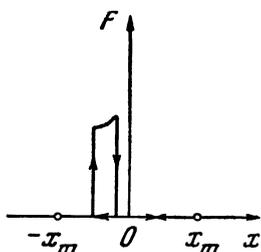


Рис. 14

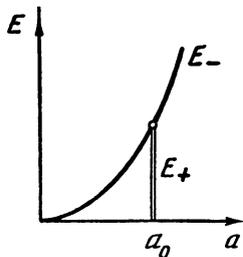


Рис. 15

лочной части пальца, не заскакивая в зубец. Таким образом, устройство поддерживает вполне определенную амплитуду, при которой положение 3 является крайним.

Нетрудно усмотреть в описанном устройстве основные части автоколебательной системы: источник энергии — батарея, колебательная система — маятник, клапан — замыкающийся контакт. Обратная связь с колебательной системы на клапан состоит в управлении контактом со стороны маятника; механизм этого управления только что описан. Остается на основе диаграммы работы убедиться в том, что в колебательную систему вкладывается энергия, т. е. что автоколебания в этой системе действительно могут происходить.

Построим диаграмму в координатах F , x , выбрав начало координат в среднем положении маятника (рис. 14). Пока маятник идет вправо и возвращается в среднее положение, внешней силы нет; эта часть диаграммы совпадает с осью x . Далее маятник идет влево от среднего положения, сила по-прежнему на него не действует. Лишь при возвращении маятника из крайнего левого положения на короткое время замыкается контакт и маятник получает короткий импульс силы. Работа этого импульса покрывает потери энергии за промежуток времени между импульсами.

Своеобразный вид имеют графики зависимости энергии от амплитуды. Положим, что механизм потерь близок к линейному, т. е. что теряемая энергия растет с увеличением амплитуды по закону, близкому к квадратичному, и изображается графически кривой, сходной с параболой (рис. 15, кривая E_-). Что касается

получаемой маятником энергии, то она равна нулю до тех пор, пока маятник еще не дотрагивается своим пальцем до контактного устройства. При определенном значении амплитуды контактное устройство начинает работать, как выше было описано, а при дальнейшем увеличении амплитуды бездействует, и энергия в колебательную систему не поступает. Таким образом, график зависимости энергии, поступающей от источника, представляется линией, совпадающей с осью абсцисс во всех точках, кроме точки a_0 , т. е. при установившейся амплитуде ¹.

Как видим, рассмотренные часы представляют собою систему с очень жестким самовозбуждением. Для того чтобы пустить их в ход, нужно дать маятнику отклонение, большее установившейся амплитуды. Впрочем все часы представляют собою системы с жестким самовозбуждением ².

Мы не будем задерживаться на других возможных устройствах подобного же рода, полагая, что одного примера достаточно для уяснения смысла статической двузначности.

§ 5. Динамическая двузначность

Рассмотрим теперь работу электромагнитного прерывателя — широко распространенного устройства, известного всем хотя бы в виде обычного электрического звонка. Схема такого прерывателя была приведена на рис. 1.

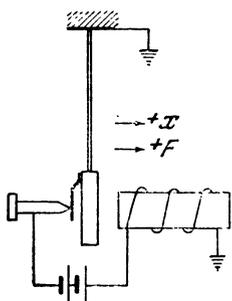
Несмотря на кажущуюся простоту устройства, его действие долго понимали неправильно. Особенность работы прерывателя сразу обнаруживается, если мы попытаемся построить диаграмму работы на основании статической характеристики.

Условившись относительно знаков силы и смещения, как показано на рис. 16, и обозначая через x_0 положение якоря, в котором замыкается контакт, мы получим диаграмму работы (рис. 17). Правее x_0 контакт разомкнут и внешняя сила равна нулю. Левее x_0 контакт замкнут и на якорь действует сила, тем большая, чем ближе якорь к электромагниту.

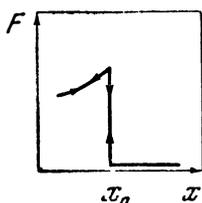
В этой картине нужно обратить внимание на то, что контакт замыкается (при ходе якоря влево) и размыкается (при ходе вправо) в одном и том же положении якоря. Стало быть, сила есть о д н о з н а ч н а я функция положения, и диаграмма

¹ Все вышеприведенное описание относится к случаю, когда работа одного импульса в точности равна потерям за период. В действительности в часах описанной конструкции работа импульса в несколько десятков раз больше. Поэтому после каждого импульса маятник совершает соответствующее число постепенно затухающих колебаний, расходуя полученный им запас энергии, до тех пор, пока амплитуда не достигнет того значения, при котором срабатывает контактное устройство.

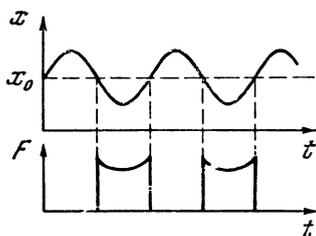
² Может показаться на первый взгляд, что это не относится к ручным и карманным часам. Но дело в том, что необходимые для их пуска начальные толчки эти часы в избытке получают в процессе завода.



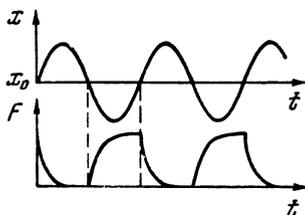
Р и с. 16



Р и с. 17



Р и с. 18



Р и с. 19

работы должна выродиться в линию, как и показано на рис. 17. Следовательно, энергия в колебательную систему прерывателя не поступает, и он принципиально не может работать. Но, однако, он работает!

Для того чтобы разобраться в сути дела, начнем с построения осциллограмм силы и движения. Предположим для простоты, что зависимость движения якоря от времени синусоидальна. Хотя в действительности это и не так, но истинная форма движения несущественна для качественного рассуждения. На рис. 18 вверху изображена синусоида, представляющая изменение положения якоря со временем. Ниже изображена зависимость силы от времени. В точках $x = x_0$ контакт включает и выключает ток, и, следовательно, возникает и исчезает сила, которую мы считаем пропорциональной току.

Если мы на основании этих двух осциллограмм построим диаграмму работы, как описано было в § 3 (см. рис. 6), то получим рис. 17.

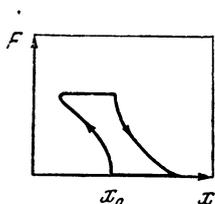
Все дело заключается в том, что рис. 18 не соответствует действительности и что на нем не показано влияние решающего в данном случае обстоятельства. Именно, ток не нарастает мгновенно до полного своего значения из-за задерживающего действия индуктивности обмотки электромагнита. (Эта обмотка обязательно обладает заметной индуктивностью; если бы она ею не обладала, то это означало бы, что она не способна создавать магнитный поток, т. е. выполнять свое прямое назначение.)

Ток и не убывает мгновенно до нуля при переходе якоря через положение x_0 в обратном направлении благодаря возникновению искры. Если принять все это во внимание, то наши осциллограммы примут вид,

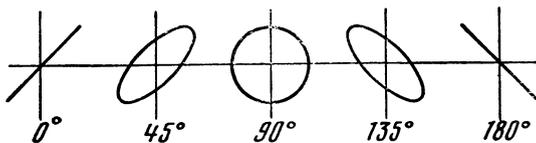
показанный примерно на рис. 19. Здесь характерно общее запаздывание кривой силы (по сравнению с рис. 18).

Если теперь построить на основании осциллограмм рис. 19 диаграмму работы, то получится картина, представленная на рис. 20. Двузначность налицо, энергия в систему вкладывается и возможность автоколебаний, таким образом, показана.

Однако двузначность в рассмотренном случае имеет совсем другой характер, нежели статическая двузначность, рассмотренная в предыдущем параграфе. В случае прерывателя статическая двузначность отсутствует (рис. 17); двузначность возникает только при колебаниях в результате запаздывания силы, или, говоря другим языком, вследствие сдвига по фазе. Ни о запаздывании,



Р и с. 20



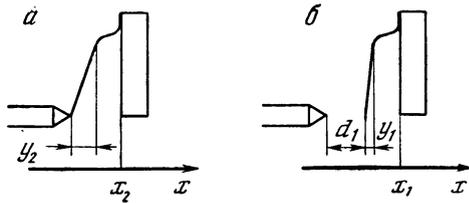
Р и с. 21

ни о фазном сдвиге применительно к статическим закономерностям речи быть не может; эти понятия неразрывно связаны со временем, с динамикой процесса. Поэтому естественно назвать такого рода двузначность *динамической*.

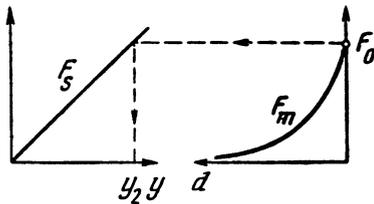
Радиотехник легко освоится с этим понятием, вспомнив динамическую характеристику триода с индуктивной нагрузкой в цепи анода. Эта характеристика, которую легко можно наблюдать на экране электронного осциллоскопа, имеет вид наклонной прямой при низких частотах или при очень малой индуктивной компоненте анодной нагрузки. При увеличении же индуктивной компоненты динамическая характеристика превращается в эллипс. Это обусловлено тем, что анодный ток уже заметно сдвигается по фазе (запаздывает) относительно напряжения на сетке.

Впрочем вопрос имеет значительно более общий характер. Речь идет, в сущности, о сложении двух взаимно перпендикулярных колебаний. Результат этого сложения представляется обычно графически в форме общеизвестных фигур Лиссажу, которые мы рассматриваем здесь для частного случая двух колебаний одинаковой частоты. В любом элементарном учебнике можно найти серию картинок, изображающих фигуры Лиссажу для этого случая при различном сдвиге фаз. Такая серия воспроизведена на рис. 21 для двух синусоидальных колебаний одинаковой частоты и амплитуды.

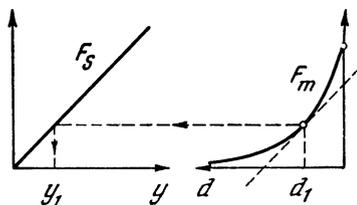
Подобного рода построения позволяют весьма простым образом изложить вопрос об ортогональности функций, имеющий



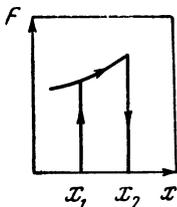
Р и с. 22



Р и с. 23



Р и с. 24



Р и с. 25

прямое отношение к нашей теме. Попытка такого изложения вынесена в добавление (см. стр. 366).

Вернемся, однако, к прерывателю. Выяснив способ его действия, мы можем теперь заметить, что обычный прерыватель неприемлем в качестве возбуждителя для электрических часов. Дело в том, что обычный прерыватель работает на частотах порядка десятков и даже сотен герц, тогда как частота маятника часов всего 1 или 0,5 герца. При такой частоте относительное запаздывание (или сдвиг фаз) будет слишком малым и недостаточное количество энергии не будет передано маятнику. Вот почему в электрических часах применяются возбуждители со статической двузначностью.

Интересно в качестве примера разобрать одно очень простое видоизменение устройства обычного прерывателя, сообщаемое ему свойство статической двузначности и, следовательно, способность работать на любых, сколь угодно низких частотах.

В принципе для этого необходимо, чтобы замыкание и размыкание контакта происходили в разных положениях якоря. Практически этого можно достигнуть, намагнитив контактный винт так, чтобы стальная контактная пружина к нему прилипла. Разберем подробно, что при этом получится.

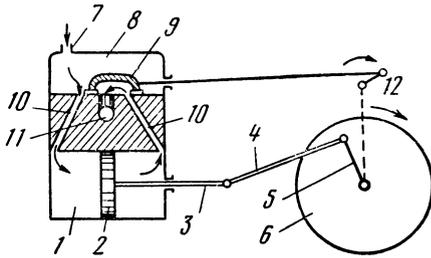
Прежде всего рассмотрим механизм замыкания и размыкания контакта. Для этого введем две переменные: координату якоря x и деформацию контактной пружины y . Когда якорь приближается к электромагниту (движение вправо, рис. 22, *a*), контактная пружина остается притянутой к намагниченному контактному винту до тех пор, пока сила упругости деформированной пружины не преодолет силы притяжения, после чего пружина оторвется от винта и контакт разомкнется. Если отсчитывать смещение якоря от положения, в котором не деформированная пружина только касается винта, то в условиях рис. 22, *a* смещение якоря и деформация пружины равны. Следовательно, для того чтобы найти положение якоря x_2 , в котором происходит размыкание контакта, достаточно найти соответствующее значение деформации пружины y_2 . Это легко сделать при помощи построения рис. 23, где показаны графики сил, действующих на пружину. Слева изображена упругая сила F_s , прямо пропорциональная деформации y . Справа кривой линией представлена зависимость силы притяжения между пружинной и винтом в функции расстояния d между ними. До отрыва это расстояние равно нулю, и сила притяжения имеет наибольшее значение F_0 . Снося это значение, как показано пунктиром, на прямую F_s , находим значение деформации y_2 , а следовательно, и равное ей смещение якоря x_2 в момент отрыва.

Замыкание контакта происходит следующим образом. Когда якорь удаляется от электромагнита, двигаясь влево (рис. 22, *b*), то пружина испытывает все возрастающую силу притяжения, которая, наконец, оказывается достаточной, чтобы притянуть пружину к винту. В этот момент пружина переходит через положение неустойчивого равновесия. Значения деформации y_1 и положения якоря x_1 в этот момент также находятся графически при помощи тех же кривых, что и на рис. 23, но несколько иным построением. Прежде всего мы находим расстояние d_1 , на котором находится точка неустойчивого равновесия. Для этого проводим касательную к F_m параллельно F_s (рис. 24). Сносим точку касания на прямую F_s и определяем, таким образом, деформацию y_1 . Координата якоря x_1 в момент замыкания равна, как нетрудно сообразить, сумме y_1 и d_1 .

Таким образом, координаты точек замыкания и размыкания контакта определены, и мы можем построить диаграмму работы. Она изображена на рис. 25, который полезно сравнить с рис. 17.

§ 6. Поршневые двигатели

Рассмотрим в качестве примеров автоколебательных систем поршневые двигатели и начнем с обыкновенной паровой машины. Само собою разумеется, что мы не будем здесь касаться новейших усовершенствований в этой отрасли техники; мы воспользуемся лишь простейшей схемой паровой машины.

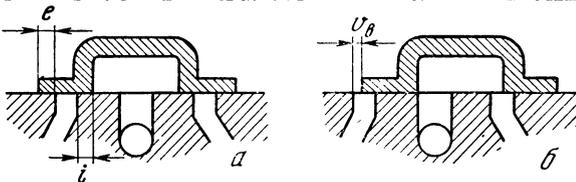


Р и с. 26

Эта схема изображена на рис. 26. В цилиндре 1 движется поршень 2, который через посредство штока 3 и шатуна 4 вращает кривошип 5, а с ним и маховик 6. Пар из котла через трубу 7 попадает в коробку 8. Здесь происходит парораспределение, которым заведует золотник 9. Его действие состоит в том, что он либо направляет пар из коробки в цилиндр по впускным трубам 10, либо сообщает цилиндр с выпускной трубой 11, откуда пар выходит наружу (в атмосферу или в холодильник). Движением золотника управляет эксцентрик 12, сидящий на валу машины.

Рассматривая систему как автоколебательную, мы видим ее составные части: источник энергии — паровой котел, колебательная система — поршень¹, клапан — это, очевидно, золотник. Обратная связь состоит в управлении золотником со стороны вала, т. е. со стороны колебательной системы.

Данная система обладает статической двузначностью. Это характеризуется тем, что положения поршня и золотника определенным образом согласованы через посредство углового расположения кривошипа и золотникового эксцентрика. Согласование



Р и с. 27

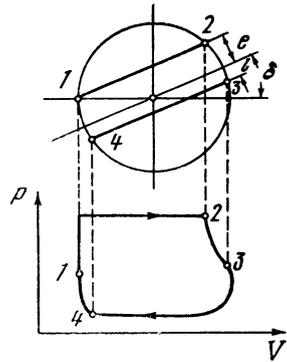
состоит в том, что при движении поршня вправо золотник открывает левую впускную трубу и одновременно дает возможность отработавшему пару из правой части цилиндра выйти в выпускную

¹ К колебательной системе можно отнести, кроме поршня, также и кинематически связанные с ним шатун и кривошип и даже маховик с валом. Вообще же и здесь и ниже мы будем понимать под колебательной системой ту часть автоколебательной системы, которая колеблется, т. е. в которой наблюдается периодическое изменение интересующей нас физической величины. Необязательно, чтобы колебательная система в изолированном виде обладала способностью совершать колебания с определенным собственным периодом. Пример вырожденной колебательной системы нам еще встретится в § 15.

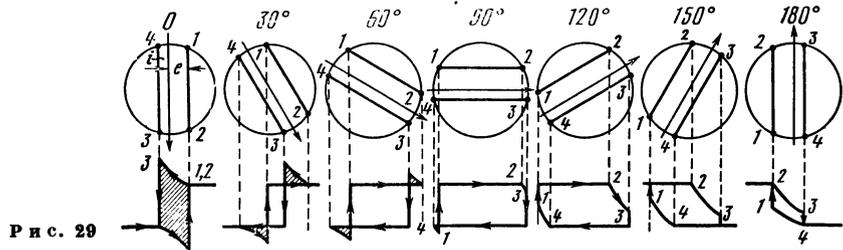
трубу; при движении же поршня влево золотник, переместившись, создает обратное положение.

Чтобы разобраться в работе машины, нужно детально рассмотреть парораспределение. На рис. 27, а золотник изображен в своем среднем положении. Подошва золотника перекрывает отверстие впускной трубы. Внешний перекрыш обозначен буквой *e*, внутренний — буквой *i*. На рис. 27, б изображено положение золотника, соответствующее крайнему левому положению поршня. В этом положении поршня золотник уже должен прикрыть впуск в левую часть цилиндра. Это достигается путем расположения эксцентрика золотника под углом, несколько большим 90° по отношению к кривошипу в сторону опережения. Дополнительный (сверх 90°) угол обозначается δ и называется углом опережения золотника.

Влияние всех этих величин на порядок парораспределения поясняется обычно при помощи так называемой золотниковой диаграммы. Было предложено много построений такого рода диаграмм. На рис. 28 показано одно из них. За основу диаграммы берется окружность, радиус которой изображает в известном масштабе радиус кривошипа, и, стало быть, диаметр



Р и с. 28



Р и с. 29

окружности в том же масштабе изображает полный ход поршня между крайними положениями. Проекция точки, равномерно движущейся по окружности, дает мгновенное положение поршня. Но проекция другой точки, сдвинутой относительно первой в сторону опережения на угол $90^\circ + \delta$, дает также мгновенное положение золотника. Если же повернуть диаграмму для золотника на этот же угол, то одна и та же точка даст нам положение как поршня, так и золотника.

Пересечение горизонтального диаметра с окружностью отмечает крайние левое и правое положения поршня. Вертикальный диаметр отмечает среднее положение поршня. Для построения

линии среднего положения золотника проводим наклонную прямую под углом δ к горизонтальному диаметру (или под углом $90^\circ + \delta$ к вертикальному). На расстояниях от этой прямой, равных внешнему и внутреннему перекрышам e и i , проводим еще две параллельные прямые. Если теперь обходить окружность по часовой стрелке, начиная от точки I , то мы получим последовательно все фазы парораспределения. А так как нам известно в каждый момент и положение поршня, то мы можем сразу строить индикаторную диаграмму. Все это изображено на рис. 28; сверху — золотниковая диаграмма, внизу — индикаторная диаграмма, т. е. диаграмма работы в координатах p (давление) и V (объем, пропорциональный смещению поршня).

Последовательность явлений, происходящих по одну сторону поршня, такова:

1. Начало парораспределения. Золотник открывает впуск. Давление сразу становится равным рабочему давлению пара и остается неизменным при движении поршня вправо до тех пор, пока золотник не перекроет впуск.

2. Начало расширения. Впуск прекращен. При дальнейшем движении поршня происходит расширение уже поступившего в цилиндр пара. Давление его соответственно падает.

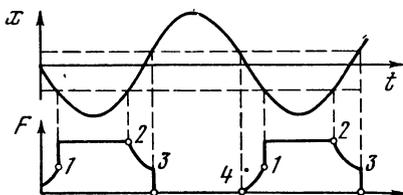
3. Начало выпуска. Золотник соединяет цилиндр с выпускной трубой. Давление быстро падает до атмосферного. Вскоре же начинается обратное движение поршня; давление остается пока неизменным.

4. Начало сжатия. Впуск прекращается. Поршень, продолжая двигаться влево, сжимает оставшийся пар, давление которого соответственно повышается.

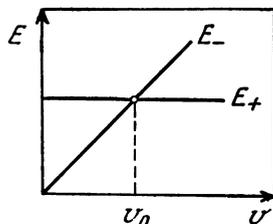
1. Начало парораспределения. Золотник открывает впуск... и далее цикл повторяется в описанном выше порядке.

То же самое, но со сдвигом на 180° происходит по другую сторону поршня. Работа, совершаемая машиной за один цикл (т. е. за один оборот вала), выражается удвоенной площадью индикаторной диаграммы рис. 28. Площадь диаграммы существенным образом зависит от выбранного сдвига между кривошипом и эксцентриком. Ясно, что имеется наиболее выгодное значение этого сдвига. Его можно найти, проследив, как видоизменяется диаграмма работы в зависимости от фазы распределения. На рис. 29 представлена серия золотниковых диаграмм и соответствующих индикаторных диаграмм для фазы распределения, выраженной через полный угол опережения золотника $90^\circ + \delta$ в пределах от 0 до 180° (т. е. для δ от -90 до $+90^\circ$). Диаграммы упрощены в том отношении, что предположены мгновенное нарастание давления до рабочего в момент впуска и мгновенное спадание давления до атмосферного в момент выпуска. Как видим, при фазе 0 работа отрицательна. При фазе 30° на диаграмме работы наряду с отрицательными (заштрихованными) областями имеется положительная. При дальнейшем увеличении опережения соотношения изме-

няются в пользу положительных областей. При фазе 120° ($\delta=30^\circ$) отрицательных областей уже нет; это — наивыгоднейшее опережение для данной серии. При еще большем увеличении опережения площадь диаграммы сокращается. Интересно отметить, что при фазе 180° ($\delta=90^\circ$) работа, совершаемая машиной, не равна нулю, как можно было бы ожидать, основываясь на том, что открытие и закрытие впуска и выпуска происходят при одних и тех же положениях поршня. Возможность совершения работы связана с процессами сжатия и разрежения, благодаря которым первая гармоника действующей на поршень силы сдвинута отно-



Р и с. 30



Р и с. 31

сительно положения поршня на угол, меньший 180° . Другими словами, имеется опережающая составляющая силы, хотя и относительно небольшая. Это видно из рис. 30, на котором изображены осциллограммы смещения поршня x и действующей на него (с одной стороны) силы F при фазе 180° .

В качестве второго примера рассмотрим обычный автомобильный бензиновый двигатель. Его устройство и работа в настоящее время настолько широко известны, что описывать их здесь нет надобности. Отметим лишь интересные для нас подробности.

В автомобильном двигателе имеется двойное распределение, причем оба рода распределения управляются валом двигателя. Во-первых, имеется распределение впуска и выпуска. Для этого служат пружинные клапаны — по одному впускному и выпускному на каждый цилиндр. Впускной клапан открывается только во время всасывания, выпускной — только во время выхлопа. Во время сжатия и в течение рабочего такта оба клапана закрыты. Работой клапанов управляет специальный распределительный вал, снабженный кулачками. Так как полный цикл работы цилиндра совершается за два оборота главного вала (т. е. за два полных хода поршня), то распределительный вал, связанный с главным (коленчатым) валом зубчатой передачей, делает вдвое меньшее число оборотов.

Во-вторых, имеется распределение зажигания. В автомобильном двигателе зажигание горючей смеси производится электрической искрой, проскакивающей в свече. Момент зажигания имеет очень большое значение. Как известно, двигатель развивает

наибольшую мощность, если вспышка происходит не в верхней мертвой точке (по окончании такта сжатия и перед началом рабочего такта), а несколько раньше. Это называется опережением зажигания. Установка зажигания чрезвычайно критична, для ее регулировки предусмотрены специальные приспособления. В старинных автомобилях манетка опережения зажигания была выведена на рулевое колесо и водитель все время регулировал опережение вручную. В современных автомобилях управление опережением автоматизировано. Дело в том, что наилучшее опережение зависит от скорости двигателя; оно возрастает с увеличением числа оборотов. Современные двигатели довольно быстроходны (2000—3000 оборотов в минуту). Поэтому вариации опережения довольно значительны. Для автоматического увеличения опережения с возрастанием скорости предусмотрены два автомата: центробежный и вакуумный. Действие первого в общем понятно, второй управляется разрежением во всасывающей трубе, которое возрастает со скоростью.

Из этого беглого описания видно, какое большое значение имеет фаза распределения, т. е. сдвиг цикла обратной связи относительно рабочего цикла системы. Усилия направлены к тому, чтобы правильным выбором этого сдвига получить при прочих равных условиях наибольшую площадь диаграммы работы, т. е. наибольшую мощность при данном расходе топлива, а значит, и наибольший коэффициент полезного действия¹.

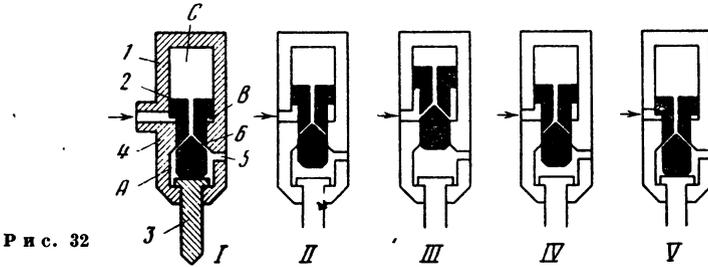
Одно заключительное замечание: дело в том, что в поршневых двигателях амплитуда колебания есть постоянная величина, вполне определяемая конструктивными параметрами двигателя, в частности радиусом кривошипа. Поэтому результатом увеличения подводимой энергии (повышение давления пара в паровой машине или обогащение смеси в бензиновом двигателе) будет неизбежно повышение частоты колебаний, т. е. скорости двигателя. Устойчивое значение скорости будет достигнуто тогда, когда подводимая энергия сравняется с энергией, теряемой на валу (нагрузка плюс собственные потери двигателя). Увеличение нагрузки вызовет соответствующее понижение скорости. Чтобы найти установившееся значение скорости из условия энергетического баланса, мы можем поступить так: обозначив через E_+ и E_- по-прежнему энергии, соответственно получаемую и теряемую системой за цикл, построить эти величины в функции с к о р о с т и (а не амплитуды, как мы делали раньше). E_+ от скорости не зависит (или почти не зависит: мы не учитываем газодинамических эффектов), E_- растет в первом приближении пропорционально скорости (при постоянной нагрузке). Получается график (рис. 31), на котором v_0 означает установившуюся скорость (или частоту).

¹ Здесь сознательно не затрагиваются многие другие вопросы, как, например, роль степени сжатия, влияние качества топлива и явления детонации и т. д.

§ 7. Пневматические ударные инструменты

К автоколебательным системам относятся разнообразные и широко применяющиеся в промышленности ударные пневматические инструменты, как, например, отбойный молоток, клепальный молоток, пневматические лопаты, ломы, трамбовки, шпальные подбойки, колотушки для очистки труб и многие другие.

По устройству ударного механизма все эти инструменты можно разделить на золотниковые и беззолотниковые. Обе эти группы



представляют для нас интерес. Мы начнем с беззолотникового механизма, как более простого в конструктивном отношении.

Отличительной особенностью этого механизма является то, что он имеет только одну подвижную часть — поршень, который одновременно служит бойком, ударяя по рабочему инструменту, и, кроме того, сам же управляет распределением воздуха, т. е. является золотником.

Таким образом, в этом устройстве опережение золотника принципиально исключено. Как мы сейчас увидим, условия работы беззолотникового механизма сходны с условиями работы паровой машины при фазе распределения 180° . Рассмотрим устройство и действие беззолотникового механизма.

На рис. 32 представлен схематически один из самых ранних беззолотниковых ударных механизмов в различных положениях на протяжении одного цикла автоколебаний. В цилиндре 1 ходит поршень 2, ударяющий по рабочему инструменту 3. Сжатый воздух подается в цилиндр через патрубок 4. Нижняя полость А всегда сообщена с атмосферой через отверстие 5, так что давление в этой полости равно атмосферному. В кольцевой полости В под заплечиком поршня давление всегда равно давлению подаваемого сжатого воздуха. Распределение осуществляется при помощи отверстий б, высверленных в теле поршня и соединенных осевым каналом поршня.

Действие системы мы рассмотрим, начиная с момента удара.

I. Удар только что нанесен. Давление в полости С равно атмосферному, так как эта полость сообщается с полостью А через осевой канал и отверстия б. На поршень действует направ-

хода через крайнее положение (точка 4) поршень устремляется вниз и наносит с большой скоростью удар (точка 1). Удар предполагается неупругим, так что поршень теряет скорость и возобновляет движение вверх лишь под действием силы F_0 .

Не останавливаясь на различных конструктивных вариантах этого устройства, перейдем к золотниковым механизмам, имеющим наибольшее значение для техники. Различают два конструктивных варианта: с цилиндрическим золотником и с плоским клапаном. Мы рассмотрим лишь первый вариант.

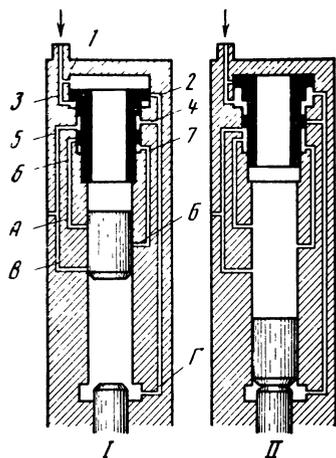
По поводу золотниковых ударных пневматических механизмов следует сделать одно замечание общего характера. Эти механизмы можно уже уподобить паровой машине, так как два главнейших элемента устройства паровой машины — поршень и золотник — имеются и в рассматриваемых механизмах. Однако здесь есть и существенная разница.

Дело в том, что между поршнем и золотником в паровой машине существует жесткая кинематическая связь. В пневматических же ударных механизмах «поршень служит распределительным органом, управляющим движением золотника, а золотник — распределительным органом, управляющим движением поршня» [6, стр. 160].

Обратная связь, таким образом, налицо, однако она не является жесткой. Механизм этого взаимного управления выясняется при рассмотрении типичной схемы устройства пневматического молотка, показанной на рис. 35, основные части которого — цилиндр, поршень-ударник и золотник — показаны в двух различных положениях. Нижеследующее обстоятельное описание работы молотка заимствовано из практического руководства [7, стр. 10—12].

«1. Золотник находится в нижнем положении. Сжатый воздух по радиальному каналу 1 поступает в верхнюю камеру, т. е. во внутреннюю полость молотка, расположенную над ударником, и здесь своим давлением стремительно гонит ударник вниз по направлению к вставному инструменту, т. е. обжимке, зубилу и т. п. Происходит рабочий ход ударника.

Одновременно сжатый воздух давит сверху на золотник и фиксирует его в нижнем положении. Через радиальный канал 3 сжатый воздух поступает в кольцевую полость между золотником и золотниковой коробкой и здесь давит на первый уступ золотника, частично разгружая золотник от давления на него сверху. Кольцевая площадь первого уступа меньше, чем кольцевая площадь золотника, на которую давит воздух сверху. В тот момент, когда верхний торец ударника проходит радиальный канал ствола Б,



Р и с. 35

сжатый воздух из верхней камеры через канал B , 7 — B и 7 поступает в кольцевую полость второго уступа и здесь давит на второй уступ золотника. Сумма давлений снизу вверх на первый и второй уступы почти равна давлению на золотник сверху, и поэтому золотник находится почти в равновесии.

При движении ударника вниз воздух, находящийся в нижней камере, т. е. во внутренней полости молотка, расположенной под ударником, через каналы $Г$, $2-4-Г$ и 4 и далее через кольцевую выточку золотника через каналы 5 и $5-B$ и выхлопное отверстие канала $5-B$ выходит наружу, не препятствуя движению ударника вниз.

Когда верхний торец ударника минует радиальный канал B , сжатый воздух из верхней камеры через каналы C и $5-B$ будет выходить наружу, и давление в верхней камере упадет. При этом упадет также давление на золотник, который под действием усилия, действующего на первый и второй уступы, переместится в верхнее положение.

Ударник, двигаясь по инерции дальше, нанесет удар по хвостовику вставного инструмента.

2. Золотник находится в верхнем положении. При этом радиальный канал закрыт... и сжатый воздух в верхнюю камеру не поступает. Также закрыт радиальный канал 4 и поэтому нижняя камера не имеет сообщения с атмосферой.

Сжатый воздух через радиальные каналы 3 и затем 2 и далее через продольный канал $2-4-Г$ и канал $Г$ поступает в нижнюю камеру, давит на ударник снизу и гонит его вверх. Происходит холостой ход ударника. Одновременно сжатый воздух давит на первый уступ золотника и фиксирует его в верхнем положении.

При движении ударника вверх воздух, находящийся в верхней камере, через каналы $7-B$ и $6-A$ и далее через верхнюю часть канала $5-B$ выходит наружу, не препятствуя движению ударника вверх. Когда верхний торец ударника пройдет канал A , в верхней камере начнется сжатие воздуха.

Когда нижний торец ударника пройдет канал B , нижняя камера окажется сообщенной с атмосферой через канал $5-B$ и отверстие в нем. Давление в нижней камере упадет. Однако по инерции ударник будет продолжать движение вверх, сжимая находящийся над ним в верхней камере воздух. В верхней камере образуется воздушная подушка, препятствующая удару ударника по внутреннему торцу рукоятки, а сжатый в верхней камере воздух давит на верхнюю кольцевую поверхность золотника, и в некоторый момент это давление преодолет давление сжатого воздуха на первый и второй уступы золотника и золотник переместится вниз. При этом открывается радиальный канал 1 , через который сжатый воздух поступает в верхнюю камеру».

После этого весь цикл повторяется. Основываясь на вышеприведенном описании, мы могли бы построить диаграмму работы молотка данного типа, но мы и так уделили ему много внимания.

Ясно, что в золотниковом механизме может быть достигнута большая мощность и большая эффективность. Эти механизмы применяются для тяжелых работ. Частота ударов — порядка 1000—2000 в минуту.

§ 8. Фрикционный маятник [8, 9]

Фрикционный маятник (или маятник Фроуда) представляет собою механическую автоколебательную систему, в которой механизм возбуждения колебаний связан с зависимостью силы трения от относительной скорости трущихся тел.

Устройство фрикционного маятника показано на рис. 36. Оно состоит в том, что маятник прикреплен к муфте, насаженной на вращающийся с постоянной скоростью вал.

При известных обстоятельствах возникают автоколебания маятника. Мы исследуем условия возникновения этих автоколебаний и их энергетический баланс.

Типичная характеристика трения, т. е. зависимость силы трения F_R от скорости скольжения (относительной скорости двух трущихся тел), показана на рис. 37, а. Характеристика такого вида типична для случая, когда трущиеся поверхности смочены

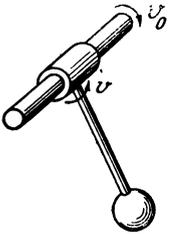


Рис. 36

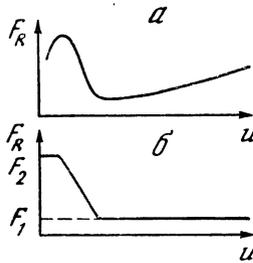


Рис. 37

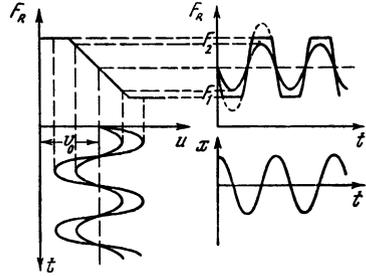


Рис. 38

той или иной жидкостью. Для упрощения исследования аппроксимируем характеристику трения ломаной линией (рис. 37, б).

Положим, как оно и есть в действительности, что движение маятника близко к синусоидальному. Значит, и скорость изменяется синусоидально, и мы можем, опираясь на характеристику трения, рис. 37, б, построить осциллограмму силы. Построение этой осциллограммы показано на рис. 38 для двух значений амплитуды. В одном случае изменения скорости лежат в пределах прямолинейного падающего участка характеристики трения, в другом — заходят за изломы характеристики и выходят на ее горизонтальные участки. По поводу этого построения нужно заметить следующее. Во-первых, по оси абсцисс отложена относительная линейная скорость u . Она равна сумме линейной скорости вала v_0 и линейной скорости муфты маятника v , если положительное направление v выбрано против v_0 . Во-вторых, сила трения, рассматриваемая как внешняя по отношению к маятнику сила, есть не движущая, а тормозящая сила, направленная против движения. Поэтому в качестве внешней силы F нужно брать силу трения F_R с обратным знаком. Если мы не примем это обстоятельство во внимание, то не получим правильного фазного сдвига и надлежащего направления обхода диаграммы работы. На рис. 38 под осциллограммой силы построена осциллограмма перемещения. Из двух этих осциллограмм может быть получена диаграмма работы (рис. 39). Она представляется в общем случае эллипсом, усеченным двумя парал-

тельными горизонтальными прямыми, отвечающими значениям силы F_1 и F_2 . При малых амплитудах эллипс не усекается; он вписывается полностью между указанными прямыми.

Рис. 39 дает нам представление о том, как изменяется вкладываемая в систему энергия в зависимости от амплитуды колебания. Получается, что для малых амплитуд, т. е. для неусеченного эллипса, площадь растет пропорционально квадрату амплитуды (так как площади подобных фигур пропорциональны квадратам линейных размеров). Для больших же амплитуд закон возрастания площади приближается к линейному, так как один из размеров фигуры (усеченного эллипса) остается неизменным. Что же касается потерь, то можно считать, что они растут пропорционально квадрату амплитуды.

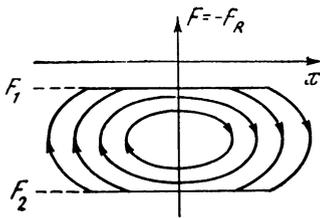


Рис. 39

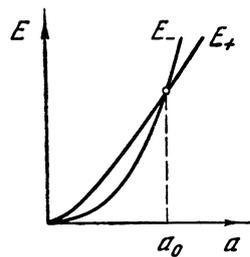


Рис. 40

На основании этих соображений можем построить графики зависимости E_+ и E_- от амплитуды (рис. 40). Эти графики позволяют найти установившуюся амплитуду колебаний. Все изложенное представляет собою уже достаточное приближение к действительным соотношениям в системе типа фрикционного маятника.

Из сказанного вытекает, между прочим, что для автоколебаний описанной системы существенную роль играет наличие падающего участка на характеристике трения. Действительно, если бы сила трения не убывала со скоростью, а нарастала, то при помощи рис. 38 нетрудно сообразить, что в этом случае фаза силы изменилась бы на 180° и, следовательно, направление обхода на диаграмме работы рис. 39 также изменилось бы на обратное. Это означало бы, что энергия не только не вкладывается в систему, но, наоборот, отнимается от нее.

Понятие падающей характеристики оказывается весьма общим, так как оно тесным образом связано с понятием устойчивости. Поэтому данное понятие может с успехом применяться при рассмотрении различных автоколебательных систем.

По поводу характеристик трения при смачивании (так называемое граничное трение) следует еще заметить, что форма рис. 37, а является типичной для самых различных условий. Оказывается, что даже при смазывании трущихся поверхностей маслом форма характеристик трения сохраняется, и на ней имеется явно выраженный падающий участок. Если так, то возможны автоколебания,

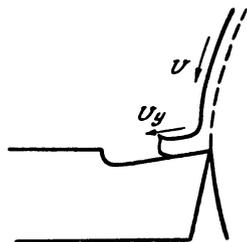
и становится непонятным, почему смазывание устраняет автоколебания (например, скрип).

Дело заключается в том, что в случае масла в качестве смазывающей жидкости падающий участок характеристики трения перемещается в область очень малых относительных скоростей. Практически встречающиеся скорости лежат уже на возрастающем участке характеристики.

§ 9. Автоколебания резца [10]

При обработке металла на токарном станке возникают иногда вибрации, представляющие собою автоколебательное явление. Явление это вредно, так как при вибрации резца резко ухудшается качество обработанной поверхности — она становится волнистой. Возрастает износ резца. И, наконец, вибрация может затруднить увеличение скорости резания, т. е. повышение производительности станка.

Явления, происходящие при резании, весьма сложны. Необходимо прежде всего, рассмотрев эти явления хотя бы в самых общих чертах, выделить те из них, которые могут иметь прямое отношение к возникновению автоколебаний¹.



Р и с. 41

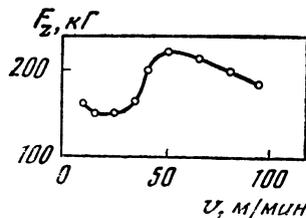
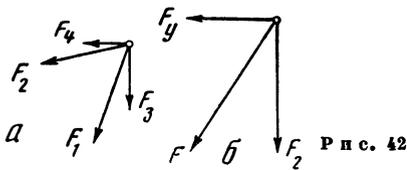
Взаимное расположение резца и обрабатываемой детали показано на рис. 41. Со стороны изделия на резец действуют следующие усилия: 1) усилие резания F_1 , т. е. усилие, требуемое для пластической деформации металла и отделения стружки; 2) сила F_2 трения стружки, сбегающей по передней грани резца; 3) сила F_3 трения изделия о заднюю грань резца; 4) сила F_4 давления упругодеформированного изделия на заднюю грань резца. Примерные направления всех этих сил показаны на рис. 42, а (предполагается, что все они приложены к лезвию резца). Однако измерить эти силы порознь затруднительно. Поэтому при экспериментальных исследованиях условий резания ограничиваются измерением двух проекций результирующей силы F : радиальной силы F_y и касательной силы F_z (рис. 42, б).

Если зависимость касательной силы F_z от скорости резания v имеет падающий участок, то возможны автоколебания, механизм которых ничем по существу не отличается от механизма фрикционных автоколебаний, разобранных в § 8. Все, что гово-

¹ Мы будем говорить в дальнейшем о колебаниях резца. Однако ясно, что речь идет всегда об относительном движении резца и обрабатываемого изделия. Нижеописываемые явления могут происходить при периодических деформациях как резца, так и изделия или всех деталей станка, определяющих взаимное расположение резца и изделия.

рилось в этом параграфе по поводу падающей характеристики, применимо и в данном случае.

Опыт показывает, что для чугуна и высокоуглеродистой стали касательная сила (определяемая главным образом усилием резания) практически от скорости не зависит. Значит, необходимые для возбуждения автоколебаний условия отсутствуют. И действительно, вибрация резца при обработке этих металлов обычно не наблюдается. Вибрация возникает при обработке вязких малоуглеродистых сталей. На рис. 43 дана зависимость F_z от v для стали «15» при таком режиме: глубина резания $a_p = 4$ мм, подача 0,24 мм. Как видим, F_z быстро нарастает, начиная от ско-



рости около 20 м/мин, и достигает максимума при скорости около 50 м/мин. После этого F_z начинает спадать. Наблюдение за вибрацией показывает, что она возникает как раз при скоростях, приходящихся на падающий участок характеристики, т. е. начиная от 50 м/мин и выше.

Всякий, наблюдавший практически явление вибрации резца, знает, что колебания не срываются в широких пределах изменения глубины резания и подачи. И это находит себе объяснение в кривых зависимости F_z от v для различных режимов: величина усилия растет с увеличением сечения стружки, но общий характер зависимости сохраняется, в том числе и расположение падающего участка.

Если не вдаваться в анализ физических причин, обуславливающих представленный на рис. 43 ход зависимости F_z от v , что для наших целей излишне, то вопрос о механизме возбуждения вертикальных колебаний резца можно было бы считать в общем ясным. Однако нужно упомянуть и о других факторах, играющих существенную роль.

Прежде всего заметим, что возможно возбуждение колебаний в радиальном направлении за счет трения сбегающей по передней грани резца стружки. Сила этого трения в значительной мере определяет радиальную компоненту силы F_y . Радиальная сила весьма значительна, она достигает примерно половины касательной силы F_x . Это и неудивительно, если принять во внимание громадное вертикальное давление стружки на резец. О давлении, развиваемом стружкой на переднюю грань резца, можно судить

хотя бы по характерному износу передней грани, придающему ей форму «лунки».

Теперь заметим, что стружка скользит по передней грани резца в радиальном направлении с постоянной скоростью, хотя и не равной скорости резания, но ей пропорциональной. Дело в том, что при выходе на переднюю грань резца стружка усаживается. Коэффициент усадки по данным многочисленных измерений можно принять равным примерно 0,5. Таким образом, постоянная скорость скольжения стружки по резцу в радиальном направлении — обозначим ее через v_y — равна примерно половине v . Если теперь предположить, что зависимость силы трения стружки по резцу от скорости ее скольжения имеет падающий участок, то, очевидно, возможны радиальные автоколебания. Мы не располагаем опытными данными относительно зависимости силы трения F_2 от скорости скольжения v_y . Но v_y пропорциональна v , а F_2 является существенной слагающей радиальной силы F_y . Следовательно, о возможности возникновения радиальных автоколебаний можно судить по зависимости F_y от v , которая экспериментально определена. Эта зависимость (для тех же условий, что и рис. 43) показана на рис. 44. Как видим, она имеет такой же характер, как и зависимость F_x от v . Итак, радиальные колебания возможны, и притом в тех же условиях, т. е. при тех же скоростях резания, что и касательные колебания.

Это не является, конечно, совпадением. Ведь мы имеем, в сущности, дело с единым процессом, и рассмотрение соотношений в проекциях на две произвольно выбранные оси едва ли способствует уяснению явлений.

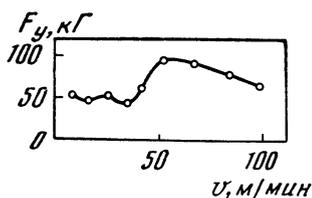
Помимо сказанного нужно еще учитывать следующее.

1. При колебаниях резца изменяется сечение стружки, что создает дополнительные переменные компоненты действующих на резец усилий.

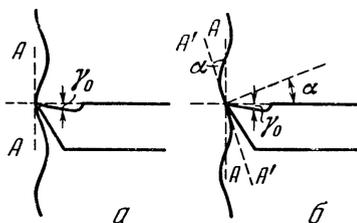
2. При колебаниях резца могут также изменяться рабочие углы резца, что влечет за собою те же последствия.

Можно показать, что последнее обстоятельство само по себе может составить механизм радиальных автоколебаний. Этот механизм поясняется на рис 45.

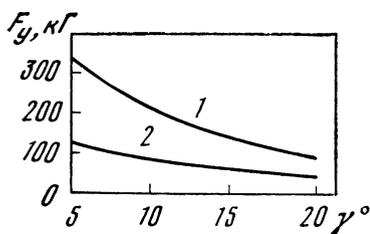
Благодаря радиальным колебаниям резец вырезает на поверхности изделия волнообразную линию. При этом плоскость резания наклоняется (положение $A'A'$ на рис. 45, б). Рабочий передний угол γ при этом изменяется; его величина определяется суммой постоянного угла γ_0 — начального переднего угла и переменного угла α , представляющего собою наклон волнистой поверхности изделия в данной точке. Угол α принимает как положительные, так и отрицательные значения. Заметим, что этот угол пропорционален радиальной скорости резца. На рис 45, а резец изображен в крайнем левом положении. В этом положении его скорость, а с нею и угол α равны нулю. На рис. 45, б резец изобра-



Р и с. 44



Р и с. 45



Р и с. 46



Р и с. 47

жен в среднем положении. В этот момент его скорость и угол α имеют наибольшие значения.

С другой стороны, известно, что и усилие резания, и его горизонтальная составляющая F_y зависят от величины переднего угла. Именно, с возрастанием γ сила F_y убывает, как показано на рис. 46 (на этом графике, относящемся к стали «1035», кривая 1 относится к сечению стружки $5 \times 0,75 \text{ мм}^2$, кривая 2— $3 \times 0,48 \text{ мм}^2$). Но угол γ , точнее его переменная составляющая α , пропорционален радиальной скорости v_r . Значит, сила F_y есть убывающая функция радиальной скорости v_r . А раз имеется падающая характеристика зависимости силы от скорости, значит, возможны автоколебания.

Есть основание полагать, что этот механизм возбуждения автоколебаний имеет практическое значение. Одним из таких оснований является тот факт, что специальная заточка резца, показанная на рис. 47, в значительной мере устраняет вибрации. Очевидно, что при такой форме резца условия скольжения стружки по передней грани существенно изменяются.

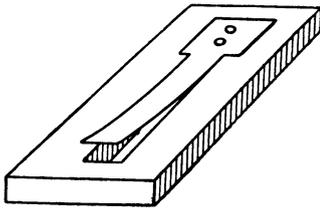
Мы не будем вдаваться в дальнейшие подробности. Приходится, к сожалению, признать, что сложная проблема автоколебаний резца изучена еще далеко не полностью. Неизвестно еще, какой из возможных механизмов автоколебаний играет здесь решающую роль. Поэтому нет еще и радикальных мер борьбы с вибрацией резца.

Однако нет сомнений в том, что постановка этой проблемы как автоколебательной и применение

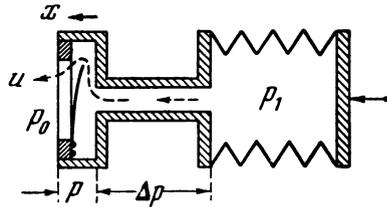
к ее решению общих принципов теории автоколебаний, а также использование современной экспериментальной техники приведут нас в ближайшем времени к полному овладению проблемой вибрации резца.

§ 10. Гармонный язычок

В этом параграфе речь пойдет об автоколебаниях в системе с воздушным возбуждением — об автоколебаниях в таких музыкальных инструментах, как различные гармонии, в том числе баян, аккордеон, концертино, а также в фисгармонии и в язычковых регистрах органа. Теория автоколебаний гармонного язычка впервые была дана не так давно Б. П. Константиновым [11]. Вообще можно подчас удивляться, как поздно приходит правильное понимание давно наблюдаемых и используемых явлений.



Р и с. 48



Р и с. 49

Гармонный язычок представляет собою гонкую металлическую пластинку, прикрепленную одним концом к более толстой пластинке — планке. Планка снабжена сквозным отверстием — проемом, в который язычок может входить с очень небольшим зазором. Свободный конец язычка в начальном положении несколько отогнут вверх (рис. 48).

В гармонии имеется целый набор планок с настроенными на различные частоты язычками. Каждая планка имеет сообщение с мехом гармонии через специальный канал, называемый входной камерой, или резонатором. Планки располагаются так, что избыточное давление, развиваемое мехом, действует на каждую планку сверху, пригибая язычок к проему. При изменении знака избыточного давления язычок не возбуждается. Это обстоятельство и позволяет применить два набора язычков, из которых один возбуждается при сжатии меха, а второй — при растяжении.

Взаимное расположение планки с язычком входной камеры и меха показано схематически на рис. 49. Нужно отметить следующие величины: атмосферное давление p_0 , давление в мехе p_1 , перепад давления в камере при колебаниях Δp , рабочую разность давлений, действующую на язычок p . Эта последняя величина меньше постоянной разности давлений между мехом и атмосферой на величину перепада давления в камере. Далее действие системы

определяется объемной скоростью воздуха, устремляющегося через проем наружу (как показано на рис. 49 стрелкой) под действием разности давлений p .

Прежде всего нужно установить зависимость объемной скорости u от положения конца язычка. Смещение язычка из начального положения мы обозначим через x . Легко сообразить, что, по мере того как язычок пригибается к планке, живое сечение, через которое проходит воздух, уменьшается, а следовательно, при постоянной разности давлений уменьшается и объемная скорость. Она достигнет наименьшего значения, когда язычок полностью войдет в проем. Когда же при дальнейшем прогибании язычка он выйдет из проема по другую сторону планки, то объемная скорость будет нарастать. Зависимость объемной скорости от положения язычка, снятая в статическом режиме (т. е. при неподвижном язычке — без колебаний), показана на рис. 50. Это — основная характеристика системы. Очень важно заметить, что на этой характеристике имеется падающий участок и что начальное положение язычка выбирается именно так, чтобы рабочая точка попала на падающий участок.

При помощи статической характеристики объемной скорости можно определить закон изменения скорости во времени, если известна форма колебаний язычка. Опыт показывает, что язычок колеблется практически синусоидально. Зная это, получаем осциллограмму объемной скорости воздуха простым построением, показанным на рис. 51 (для упрощения характеристика объемной скорости представлена двумя отрезками прямой; такая аппроксимация пригодна в случае очень тонкой планки).

Таким образом, при синусоидальном движении язычка изменение объемной скорости несинусоидально благодаря резкому излому характеристики скорости. Этим объясняется то, что звук гармонии, будучи очень богат обертонами, обладает специфическим тембром. Нетрудно предвидеть, — и опыт подтверждает это, — что при больших давлениях и, следовательно, при больших амплитудах спектр звука гармонии будет еще богаче.

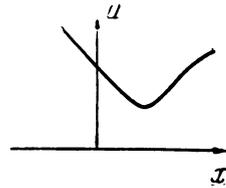
Обратная связь в данной автоколебательной системе состоит в том, что движение язычка оказывает влияние на рабочую разность давлений. Это влияние осуществляется путем изменений объемной скорости в зависимости от положения язычка; изменения скорости в свою очередь вызывают колебания давления. Дело в том, что столб воздуха в камере обладает известной инерцией. При протекании воздуха по камере с неизменной скоростью эта инерция себя не проявляет. Если же скорость изменяется, т. е. если имеется ускорение, то возникает динамический перепад давления, пропорциональный массе воздуха в камере и ускорению. При положительном ускорении этот перепад уменьшает рабочую разность давлений, при отрицательном ускорении, т. е. при замедлении, — увеличивает. Учитывая это обстоятельство, мы можем построить осциллограмму силы, действующей на язычок (эта сила

пропорциональна, очевидно, разности давлений p). Сопоставив эту осциллограмму с синусоидой, изображающей движение язычка, получим диаграмму работы.

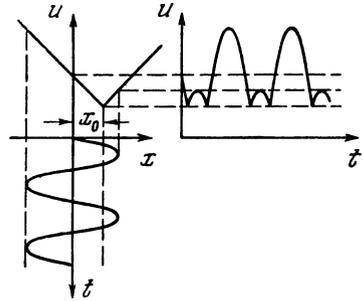
На рис. 52 изображены осциллограммы следующих величин (сверху вниз): смещения язычка x , объемной скорости u , динамического перепада Δp , пропорционального ускорению воздуха, рабочей разности давлений p . При помощи первой и четвертой осциллограмм построена диаграмма работы рис. 53. Как видим, она имеет форму эллипса, от которого отсекается сегмент, обходимый в противоположном направлении.

Теперь можно составить себе представление об энергетическом балансе. При очень малых амплитудах язычка объемная скорость меняется почти по синусоидальному закону. Когда же амплитуда превзойдет величину, соответствующую точке излома характеристики скорости (или, проще говоря, когда язычок при колебаниях начнет уже проходить сквозь планку), то скорость примет резко несинусоидальный характер, показанный на рис. 51 и 52.

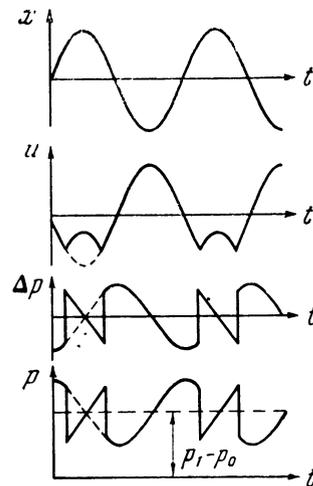
На диаграмме работы, имеющей для малых амплитуд вид эллипса, появится срез, отнимающий с дальнейшим увеличением амплитуды все большую долю площади эллипса. Поэтому если при малых амплитудах вкладываемая в язычок энергия растет пропорционально квадрату амплитуды, то в дальнейшем рост E_+ замедляется. В то же время теряемая энергия E_- растет неизменно пропорционально квадрату амплитуды (в рассматриваемом случае E_- зависит главным образом от потерь на звукоизлучение).



Р и с. 50



Р и с. 51

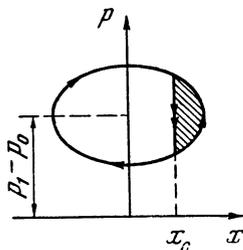


Р и с. 52

Таким образом, мягко самовозбуждающийся язычок достигает предельной амплитуды.

Эти соображения очень сходны с высказанными в § 8 по поводу фрикционного маятника. В частности, графики для E_+ и E_- имеют вид рис. 40.

Сказанное требует еще уточнения вот в каком отношении: при увеличении постоянного давления в мехе возрастает и постоянная разность давлений $p_1 - p_0$. Под действием этой разности язычок пригибается к проему (речь идет о статическом прогибе). Вследствие этого смещается рабочая точка на характеристике объемной скорости. Если давление настолько велико, что рабочая точка



Р и с. 53

попадает на излом характеристики, то колебания вообще не возникают. В этом легко убедиться практически: при чрезмерном давлении колебания срываются. Кроме того, можно заметить, что срыв колебаний наступает раньше для слабо отогнутых язычков.

Из предыдущего видно, какую большую, решающую роль в механизме самовозбуждения язычка играет входная камера. Без камеры язычки вообще не возбуждаются в обычных режимах. Вместе с тем можно отметить, что рассмотренная система представляет собою пример динамической двузначности: фазовый сдвиг силы относительно движения получается только «на ходу» в результате запаздывания, обусловленного инерцией воздуха в камере. В этом смысле данная система аналогична электрическому звонку.

§ 11. Часы

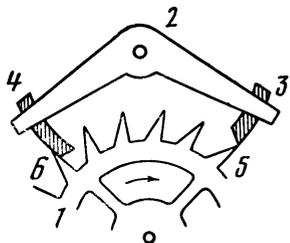
Вернемся еще раз к часам, чтобы разъяснить один важный специальный вопрос.

Об электрических часах уже говорилось в § 4. Теперь мы рассмотрим обычные механические часы.

Существует большое число конструкций механизмов, управляющих движением маятника; эти конструкции называются на профессиональном языке х о д а м и.

Не перечисляя существующих разновидностей ходов, остановимся на одном из вариантов анкерного механизма — так называемом ходе Грэхема. Его устройство и действие поясняются рис. 54. Зубчатое колесо I получает вращающий момент (на ри-

сунке вправо) от гиревого или пружинного завода. Якорь (анкер) 2 качается на оси, будучи жестко связан с маятником. Якорь снабжен двумя палеттами 3 и 4. Через посредство палетт импульс сообщается маятнику два раза за период при прохождении его через среднее положение. Это происходит следующим образом: когда зубец колеса попадает на скошенный торец палетты — так называемую «плоскость импульса», — то, проскальзывая по ней под действием вращающего колеса момента, зубец подталкивает якорь, а с ним и маятник. В это время вторая палетта опускается и удерживает очередной зубец, пока якорь снова не вернется в среднее положение. На рис. 54 изображено положение, когда

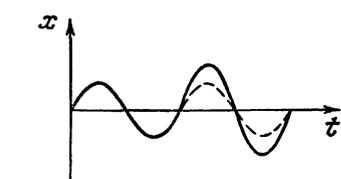
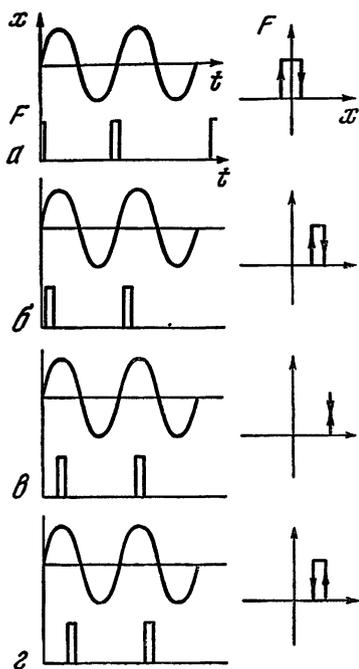


Р и с. 54

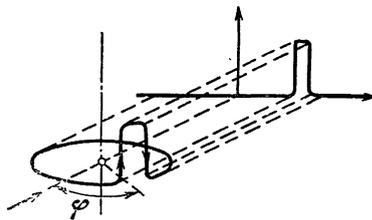
маятник, двигаясь вправо, только что прошел среднее положение и получил импульс в том же направлении, так как зубец 5 проскользнул по скосу палетты 3. Палетта 4 уже опустилась и удерживает очередной зубец 6. Этот зубец сообщит импульс противоположного направления через посредство палетты 4 маятнику при проходе его через среднее положение справа налево.

Таким образом, дело сводится к тому, что маятник получает более или менее короткий импульс два раза за период. Так обстоит дело во всех часах, в том числе и в электрических часах, рассмотренных в § 4. Правда, в тех часах маятник получает не два, а один импульс за период, но это не имеет никакого принципиального значения. В последующих рассуждениях мы будем иметь в виду именно тот случай, когда маятнику за период сообщается один импульс.

Вопрос, который имелось в виду разобрать, — это вопрос о фазе импульса. То, что во всех хороших часах импульс сообщается в момент прохождения маятника через среднее положение, т. е. когда скорость его наибольшая, — общеизвестно. Но надо выяснить, что произойдет, если фаза импульса изменится. Оказывается, — и это для часов чрезвычайно существенно, — что действие импульса может проявляться не только в увеличении энергии маятника, что необходимо для поддержания его незатухающих колебаний, но и в изменении его периода. Само собою разумеется, что при конструировании часов нужно позаботиться о таком выборе соотношений, чтобы влияние воздействия импульса на период колебаний маятника было минимальным.



Р и с. 57



Р и с. 56

◀ Р и с. 55

Займемся энергетической стороной вопроса. На рис. 55 даны осциллограммы движения и силы при разных фазах импульса.

а. Импульс сообщается в среднем положении маятника (фаза импульса — 0).

б. Импульс сообщается примерно на середине размаха (фаза импульса около 45°).

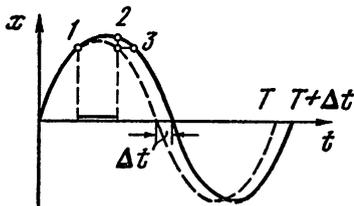
в. Импульс сообщается в крайнем положении маятника, когда скорость его меняет знак (фаза импульса — 90°).

г. Импульс того же знака сообщается при обратном движении маятника примерно на середине размаха (фаза импульса около 135°).

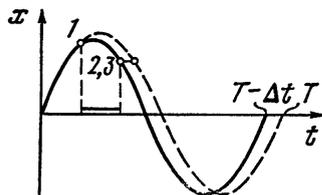
Здесь же построены и диаграммы работы. Как видим, площадь диаграммы наибольшая при фазе нуль. При фазе 90° контур диаграммы вырождается в линию. Это значит, что энергия в маятник не вкладывается. При фазе свыше 90° направление обхода диаграммы меняется. Это значит, что импульс не только не подталкивает маятник, но, напротив, тормозит его. Изменения диаграммы работы в зависимости от фазы импульса легко проследить, если представить себе проекцию проволочной модели рис. 56 на плоскость, параллельную оси, при постепенном повороте модели вокруг оси.

Теперь выясним влияние фазы импульса на период. Если импульс действует с нулевой фазой, т. е. при проходе маятника

через среднее положение, то действие импульса состоит в увеличении скорости в этот момент. Имея большую начальную скорость, маятник совершит колебание с большей амплитудой, но продолжительность цикла его движения от этого не изменится. Это иллюстрируется рис. 57, на котором показано действие одиночного импульса. Происходящее изменение скорости на этом и последующих рисунках грубо преувеличено для наглядности, а влияние



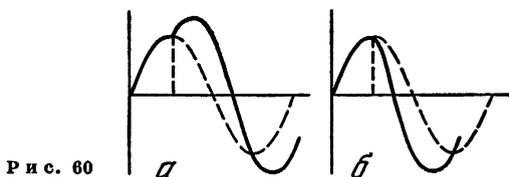
Р и с. 58



Р и с. 59

затухания не показано. На рис. 57 штрихом намечено движение маятника, которое он совершал бы, если бы не получил импульса.

Посмотрим теперь, что получится, если мы приложим импульс в крайнем положении маятника. Условимся называть импульс положительным, если он направлен в ту же сторону, в которую отклонен маятник. На рис. 58 показано действие положительного импульса с фазой 90° . Штрихом намечена синусоида невозмущенного движения маятника. Продолжительность импульса отмечена черточкой на оси. Импульс начинает действовать в точке 1. Скорость маятника нарастает в направлении импульса в течение всего времени его действия. Импульс прекращается в точке 2.



Р и с. 60

В точке 3 маятник имеет точно такую же скорость, какую он имел бы при невозмущенном движении в том же положении. Однако в это положение маятник приходит с запозданием, которое обусловлено действием импульса. Энергия в маятник не вложена, работа импульса равна нулю. Поэтому амплитуда последующего колебания маятника остается такой же, как и при невозмущенном движении. Стало быть, единственный результат действия импульса — это увеличения периода колебания на время Δt , т. е. соответствующее замедление хода часов.

Если на маятник с той же фазой 90° действует отрицательный импульс (или, что то же самое, положительный импульс с фазой 270°), то и в этом случае изменения амплитуды колебания произойти не должно, так как работа импульса равна нулю. Что же касается влияния такого импульса на период, то оно иллюстрируется рис. 59. Действие импульса, уменьшающего скорость, проявляется в том, что в момент прекращения импульса маятник имеет такую же скорость, какую он имел бы при невозмущенном движении в том же положении на Δt позже. Иными словами, действие импульса состоит в сокращении периода колебаний маятника, т. е. в ускорении хода часов.

Заметим, что действие отрицательного импульса с фазой 90° есть не что иное, как эффект удара маятника об ограничитель. А, как известно, всякое ограничение свободных колебаний укорачивает их период¹.

Рассмотренные соотношения тесно связаны с теми, которые можно вывести, рассматривая комплексную обратную связь (т. е. обратную связь с произвольной фазой) в радиотехнических устройствах. Об этом говорится в следующем параграфе.

В заключение, рискуя несколько запутать вопрос, следует все же предостеречь читателя от рассуждений, опирающихся на представление о бесконечно коротком импульсе. Эта абстракция в ряде случаев весьма полезна, и пользование ею к недоразумениям не ведет, однако данный случай составляет одно из исключений.

Мы понимаем под бесконечно коротким импульсом (речь идет о физической его трактовке) такое воздействие на механическую систему, которое вызывает мгновенное скачкообразное (точнее, совершающееся на протяжении бесконечно малого времени) изменение скорости системы. Предполагается, что система обладает конечной массой, т. е. что для изменения ее скорости нужен конечный импульс. Это в свою очередь означает, что бесконечно короткий импульс должен характеризоваться действием бесконечно большой силы на протяжении бесконечно малого промежутка времени.

Так вот, если бы мы попробовали провести предшествующие рассуждения о влиянии фазы импульса на колебания маятника для бесконечно короткого импульса, то мы пришли бы к частично неверным выводам.

В самом деле, пусть бесконечно короткий импульс с фазой 90° действует на маятник. На рис. 60 показаны осциллограммы: *a* — действие положительного импульса, *b* — действие отрицательного импульса. Если сравнить этот рисунок с рис. 58 и 59, то мы увидим, что суждение о влиянии на период, которое мы составили бы на основании рис. 60, правильно. Но как при поло-

¹ Автор наблюдал случай, когда стенные часы с открытым маятником внезапно пошли раза в полтора быстрее. Оказалось, что причиной этой аномалии было то, что маятник при каждом взмахе ударялся о корешок тетради, положенной на столик около часов.

жительном, так и при отрицательном бесконечно коротком импульсе получается увеличение амплитуды последующего колебания. Так оно и должно быть: всякое увеличение абсолютной величины скорости маятника (т. е. независимо от знака) при данном его положении должно вызвать увеличение амплитуды колебания.

Итак, использование представления о бесконечно коротком импульсе требует известной осторожности.

Окончательный же вывод, к которому мы приходим, таков: как с точки зрения наибольшей работы импульса, так и с точки зрения наименьшего его влияния на величину периода наиболее выгоднейшей фазой импульса является фаза нуль, т. е. когда импульс сообщается маятнику при проходе его через среднее положение.

§ 12. Роль фазы обратной связи

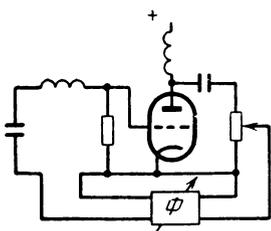
Мы уже не один раз видели, что фаза обратной связи играет существенную роль в вопросах автоколебаний. Вопрос о роли фазы будет теперь рассмотрен еще с несколько иной и в некотором отношении более общей точки зрения.

Рассмотрим схему рис. 61. Она состоит из колебательной системы в виде контура LCR . Напряжение U_R подается на сетку триода. С сопротивления в анодной цепи снимается переменная составляющая усиленного напряжения и через фазовращатель вводится снова в контур. Таким образом, мы имеем обычную систему с обратной связью, в которой могут при известных условиях возникнуть автоколебания. Проследим в подробностях, что будет происходить при изменении фазы напряжения обратной связи U' . Для этого воспользуемся обычной векторной диаграммой, применяемой в теории переменных токов. Предполагая все токи и напряжения синусоидальными, можем составить диаграмму напряжений в контуре в виде рис. 62.

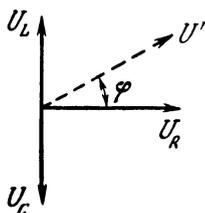
Здесь U_L , U_C и U_R — соответственно напряжения на индуктивности, емкости и сопротивлении. U' — напряжение обратной связи, величину и фазу которого мы можем по желанию изменять. Фазный сдвиг напряжения U' обозначим через φ и будем отсчитывать его от вектора V_R .

Заметим, что диаграмма рис. 62 представляет также сопротивление в контуре. Для перехода от напряжений к сопротивлениям достаточно разделить напряжения на ток в контуре. При такой постановке вопроса влияние напряжения U' можно представить как результат введения в контур некоторого сопротивления Z' , равного частному от деления U' на ток в контуре. Тогда, отбрасывая всю цепь обратной связи, можно представить контур эквивалентной схемой рис. 63.

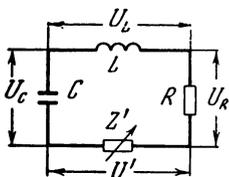
Теперь рассмотрим различные фазные соотношения. При $\varphi = 0$ U' совпадает по фазе с U_R , так что Z' есть активное сопротивление. При такой фазировке обратная связь просто увеличивает



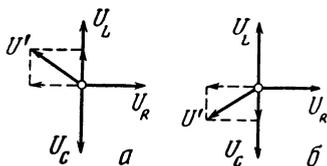
Р и с. 61



Р и с. 62



Р и с. 63



Р и с. 64

активное сопротивление контура. Это есть так называемая отрицательная обратная связь.

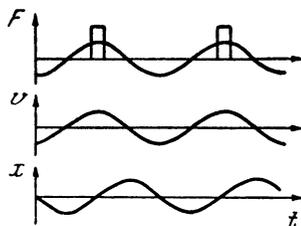
При $\varphi = 180^\circ$ U' имеет обратную фазу по отношению к U_R . Если оба напряжения равны друг другу по абсолютной величине (т. е. по амплитуде), то действие активного сопротивления уничтожается действием обратной связи. Это эквивалентно тому, что Z' представляет собой отрицательное активное сопротивление, численно равное R . При этом теряемая в контуре энергия восполняется обратной связью. Такой режим есть режим самовозбуждения, а фазные соотношения при этом характеризуются как положительная обратная связь.

Если $\varphi = +90^\circ$, то напряжение U' противоположно по фазе напряжению U_C и совпадает по фазе с U_L . Сопротивление Z' имеет при этом, очевидно, индуктивный характер. Но с равным правом можно говорить, что обратная связь вводит в контур отрицательную емкость.

При $\varphi = -90^\circ$ аналогичным образом возрастает напряжение U_C . Сопротивление Z' имеет емкостный характер; можно говорить также о введении отрицательной индуктивности.

Как видим, изменяя фазу обратной связи, можно полностью или частично компенсировать влияние каждого из параметров R , L или C . Заметим, что фаза обратной связи $\varphi = 180^\circ$ дает нам самовозбуждение без изменения собственной частоты колебательной системы. Если φ меньше 180° (рис. 64, а), то кроме компенсации U_R происходит еще частичная компенсация U_R (или, что то же, увеличение U_L). Это ведет к уменьшению собственной частоты контура. Если же φ больше 180° (рис. 64, б), то самовозбуждение происходит на частоте выше номинальной.

Основываясь на этих представлениях, мы можем очень просто рассматривать действие автоколебательной системы. Возьмем, к примеру, часы. В этой механической системе роли переменных напряжений U_R , U_L и U_C будут играть соответственно переменные силы: сила трения, сила инерции и возвращающая сила (зависящая от веса маятника). Сила, соответствующая напряжению обратной связи U' , доставляется анкерным механизмом; она имеет характер периодических импульсов. Разложим эту силу на гармонические составляющие и заметим, что работу производит только первая гармоника; высшие гармоники (при чисто синусоидальном движении маятника) работы не совершают. Это вытекает



Р и с. 65

из ортогональности тригонометрических функций (см. добавление в конце книги).

Сила, действующая на маятник, и ее первая гармоника изображены на рис. 65. Наивыгоднейшие соотношения получатся, если первая гармоника силы будет совпадать по фазе со скоростью, т. е. будет противоположна по фазе силе трения. При этом сила будет опережать по фазе на 90° движение маятника. Скорость и смещение также представлены на рис. 65 для указанного наивыгоднейшего случая.

Теперь нам ясно, что изменение фазы неминуемо повлияет на период маятника, и мы, таким образом, подтверждаем заключение, сделанное в предыдущем параграфе на основании других соображений.

Из сказанного, между прочим, следует, что закон изменения силы, действующей на маятник, не играет никакой роли. Не требуется вовсе, чтобы импульсы были короткими. Достаточно, чтобы сила была периодическая (с периодом, равным периоду маятника; это условие обеспечивается автоматически работой механизма обратной связи) и чтобы первая гармоника этой силы опережала смещение маятника на 90° . При этом условии будет, во-первых, совершаться наибольшая работа, а во-вторых, будет отсутствовать влияние обратной связи на период маятника.

§ 13. Падающая характеристика и отрицательное сопротивление

Мы уже встречались с падающей характеристикой в случае зависимости силы от скорости. Падающая характеристика обнаруживается и в электрических автоколебательных системах.

Понятие падающей характеристики является довольно общим, равно как и тесно связанное с ним понятие отрицательного сопротивления. Эти понятия и их взаимную связь легче всего уяснить, рассмотрев режим стационарных синусоидальных колебаний.

Мы начнем несколько издалека, установив сначала фазовые соотношения и их влияние на энергетический баланс. При этом нам придется отчасти повторить сказанное в предыдущих параграфах.

Ранее было установлено, что если в колебательном режиме сила F и перемещение x совпадают по фазе, то совершаемая работа равна нулю. Наибольшая работа получается для сдвига фаз между силой и смещением на 90° . Однако от знака этого сдвига, т. е. от того, опережает сила смещение или запаздывает относительно него, зависит направление обхода диаграммы работы, а следовательно, и знак работы.

На рис. 66, *а* изображены осциллограммы смещения и силы для случая силы, опережающей смещение на 90° , а на рис. 66, *б* — соответствующая диаграмма работы. На рис. 67, *а* изображены осциллограммы и диаграмма работы для случая запаздывающей на 90° силы. Как видим, в первом случае сила совершает работу, вкладываемую в систему, а во втором случае энергия извлекается из системы, на что указывает направление обхода диаграммы рис. 67, *б*.

Посмотрим теперь, как обстоит дело с фазовыми соотношениями между силой F и скоростью v . При синусоидальном движении скорость (т. е. производная от смещения по времени) всегда опережает смещение по фазе на 90° . Следовательно, в случае опережающей смещение силы, т. е. при соотношениях, соответствующих рис. 66, мы имеем совпадение силы и скорости по фазе. В случае же, когда сила отстает от смещения на 90° , она отстает от скорости на 180° , и, следовательно, в этом случае сила и скорость находятся в противоположных фазах. Изобразив зависимость силы от скорости для этих двух случаев, мы получим диаграммы рис. 68, *а* и *б*. Первая соответствует соотношениям рис. 66, вторая — рис. 67.

Характеристика рис. 68, *а* имеет положительный наклон, характеристика рис. 68, *б* — отрицательный. Характеристика с отрицательным наклоном — это есть падающая характеристика.

Мы видим, таким образом, что система с падающей характеристикой является источником колебательной энергии. Наличие в системе элемента с падающей характеристикой означает возможность возникновения автоколебаний. Иначе говоря, введение в систему элемента с падающей характеристикой может превратить эту систему в автоколебательную.

Отношение приращения силы к приращению скорости называется механическим сопротивлением (речь идет о сопротивле-

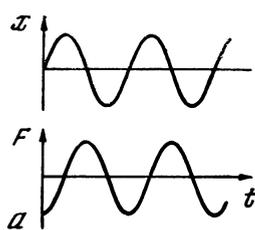


Рис. 66

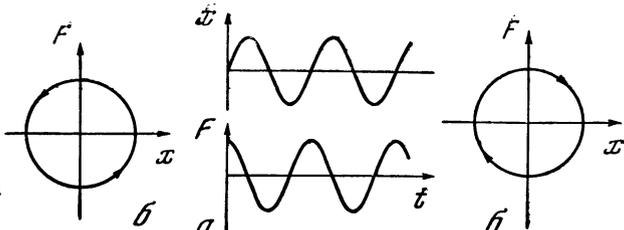


Рис. 67

нии при колебательном (движении)¹. Это отношение представляется наклоном графика зависимости силы от скорости. Наклон графика положителен на рис. 68, а и отрицателен на рис. 68, б. На этом основании, характеризуя систему с падающей характеристикой, говорят, что она обладает отрицательным сопротивлением. Это просто другой способ описания свойств системы с падающей характеристикой. Итак, можно сказать, что отрицательное сопротивление является источником энергии, тогда как положительное сопротивление — ее потребителем.

Если колебания возбуждены в системе с положительным со-

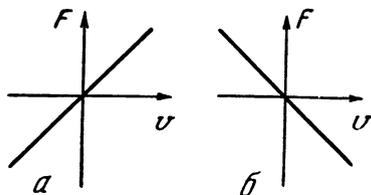


Рис. 68

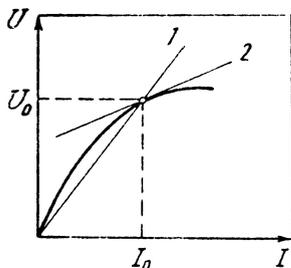


Рис. 69

¹ Напомним, что если в электрической цепи вольтамперная характеристика нелинейна, то можно по-разному определять сопротивление цепи. Сопротивление постоянному току определяется для данного значения тока (или напряжения) наклоном секущей линии 1 (рис. 69). Сопротивление же переменному току при данной постоянной составляющей тока (или напряжения) определяется наклоном касательной 2 к вольтамперной характеристике в данной точке. Это сопротивление, выражаемое отношением бесконечно малого приращения напряжения к бесконечно малому приращению тока, называется дифференциальным сопротивлением. Именно это определение сопротивления и имеется здесь в виду.

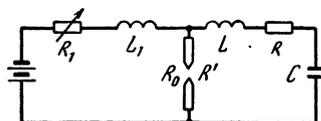


Рис. 70

противлением и если система затем предоставлена самой себе, колебания будут затухать. Если сопротивление равно нулю, мы будем иметь незатухающие колебания. На этом основании можно говорить о всякой автоколебательной системе, совершающей установившиеся колебания, что в ней наряду с неизбежным положительным сопротивлением имеется также равное по абсолютной величине отрицательное сопротивление, так что сумма обоих сопротивлений — результирующее сопротивление — равна нулю.

Действие обратной связи в автоколебательной системе можно трактовать как внесение в систему отрицательного сопротивления. Действительно, недовозбужденный генератор, т. е. генератор, в котором обратная связь не доведена еще до самовозбуждения, может применяться в качестве колебательной системы с очень малым затуханием. Это опять иной аспект тех же энергетических соотношений: действие обратной связи состоит в подведении к колебательной системе энергии; в недовозбужденном режиме эта энергия частично восполняет потери в системе. Стоя на этой точке зрения, можно сформулировать условия самовозбуждения следующим образом: автоколебательная система может самовозбуждаться, если отрицательное сопротивление сравнивается с положительным, т. е. если результирующее сопротивление обращается в нуль. Случай, когда результирующее сопротивление меньше нуля, т. е. когда оно отрицательно, не отвечает установившемуся режиму, так как при отрицательном результирующем сопротивлении приход энергии больше расхода, и система колеблется с возрастающей амплитудой.

В качестве примера и для выяснения еще одной подробности рассмотрим дуговой генератор, схема которого показана на рис. 70. Дуга питается от источника постоянного тока через реостат R_1 и дроссель L_1 ; последний защищает цепь источника от колебательного тока. Параллельно дуге подключена цепь из R , L и C . По отношению к колебательному току в этой цепи дуга включена последовательно.

Сопротивление дуги переменному току обозначено через R' . Таким образом, возможность возникновения автоколебаний определяется соотношением между положительным сопротивлением R и сопротивлением дуги R' , которое может быть и отрицательным.

В схеме рис. 70 возможны колебания трех родов: 1) малые синусоидальные колебания на падающем участке характеристики; 2) разрывные колебания, связанные с периодическим гашением дуги, и 3) так называемые колебания с обратным зажиганием. Мы рассматриваем здесь только колебания первого рода. Что касается колебаний второго рода, имеющих для техники наибольшее значение, то их механизм ничем не отличается по существу от механизма колебаний в генераторе с неоновой лампой, который будет рассмотрен позже.

Вольтамперная характеристика дуги показана на рис. 71, а. Если выбрать рабочую точку на падающем участке характеристики (отмечена кружком), то сопротивление дуги переменному току R' (определяемое наклоном вольтамперной характеристики в данной точке) будет отрицательно. Не следует, однако, смешивать это сопротивление с сопротивлением постоянному току R_0 , которое всегда положительно; оно равно отношению постоянных составляющих напряжения и тока (U_0 и I_0 на рис. 71, а). Таким образом, дуга обладает положительным сопротивлением для постоянного тока и отрицательным сопротивлением для переменного. Иными словами, дуга поглощает энергию постоянного тока и генерирует энергию переменного тока.

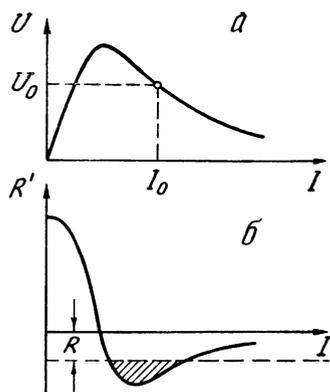


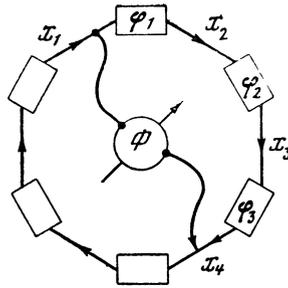
Рис. 71

Условия самовозбуждения иллюстрируются рис. 71, б, на котором построен график R' , т. е. наклона вольтамперной характеристики (производной от U по I). Сопротивление R' , положительное вначале, убывает и, перейдя через нуль, становится отрицательным. Если теперь перенести ось абсцисс на отрезок R вниз, как показано штрихом на рис. 71, б, то относительно этой новой оси мы получим график результирующего сопротивления в контуре, включающем в себя дугу. Область самовозбуждения характеризуется участком, на котором результирующее сопротивление отрицательно.

В заключение заметим, что условие неположительности результирующего сопротивления есть удобная в ряде случаев формулировка условий самовозбуждения. Однако изучение баланса энергий и исследование устойчивых амплитуд автоколебаний удобнее вести, опираясь на зависимости от амплитуды величин E_+ и E_- , как это мы и делали раньше.

§ 14. Анализ фаз

Исследование автоколебаний имеет практическую ценность, если в результате этого исследования мы получаем возможность отвечать на ряд основных вопросов, а именно: возможны ли в данной системе автоколебания; при каких условиях они возможны; если невозможны, то что нужно сделать, чтобы система приобрела свойства автоколебательной; если колебания возникают, то как можно видоизменить систему, чтобы исключить возможность самовозбуждения, и т. п. По крайней мере на первый из этих вопросов мы можем уже частично дать ответ.

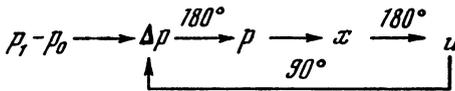


Р и с. 72

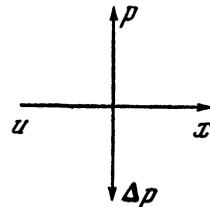
В предыдущих параграфах установлено и показано на различных примерах, что автоколебания возможны в том случае, когда в результате действия обратной связи возникает сила, совпадающая по фазе со скоростью (или, что то же, опережающая смещение на 90°). Это положение позволяет нам анализировать различные системы и выяснять их способность к автоколебаниям.

Если в системе имеется отрицательное сопротивление, т. е. падающая характеристика зависимости силы от скорости, то вопрос ясен, так как требуемое соотношение фаз налицо.

Но очень часто в системе более или менее сложным образом связан между собою целый ряд физических величин; существует целая цепь, в каждом звене которой имеется свой фазный сдвиг. В этих случаях нужно проанализировать всю цепь и найти в конце концов интересующее нас соотношение фаз между силой и скоростью (или между напряжением и током в случае электрической системы).



Р и с. 73



Р и с. 74

Сделаем несколько предварительных замечаний. Во-первых, заметим, что во всякой замкнутой цепи преобразования колебательных величин суммарный фазный сдвиг после обхода всей цепи равен нулю, а на любом участке этой цепи, включающем в себя несколько звеньев, фазный сдвиг равен сумме сдвигов в каждом отдельном звене. Такого рода замкнутая цепь представлена на рис. 72, где x_1, x_2, \dots, x_k означают те или иные колебательные величины. Первое звено преобразует величину x_1 в величину x_2 ; мы обозначим это звено и его действие символом $x_1 \rightarrow x_2$. Колебательная величина x_2 на выходе звена сдвинута относительно x_1 на угол φ_1 . Во втором звене $x_2 \rightarrow x_3$ величина на выходе x_3 сдвинута относительно x_2 на φ_2 и так далее. Если подклю-

чить фазометр, как показано на рис. 72, т. е. так, чтобы измерялся фазовый сдвиг между x_4 и x_1 , — при этом измеряется фазный сдвиг на цепочке из трех звеньев $x_1 \rightarrow x_2$, $x_2 \rightarrow x_3$ и $x_3 \rightarrow x_4$, — то фазометр покажет, очевидно, фазный сдвиг, равный сумме сдвигов φ_1 , φ_2 и φ_3 . Если теперь обойти в с ю замкнутую цепь, то результирующий фазный сдвиг будет, очевидно, равен нулю, так как в этом случае речь идет о фазном сдвиге между какой-либо из величин x_k и ею самой.

Во-вторых, заметим, что во всякой автоколебательной системе образуется замкнутая цепь взаимосвязанных колебательных величин. Такая замкнутая цепь возникает благодаря наличию обратной связи, соединяющей выход системы с ее входом. Выделив эту замкнутую цепь, мы можем применить к ней сказанное выше о соотношениях фаз в замкнутой цепи.

Теперь заметим, что вопрос о том, возможны ли в данной системе автоколебания, сводится к вопросу о том, какова фаза возникающей в результате действия обратной связи колебательной силы по отношению к колебательной скорости. Если сила совпадает со скоростью по фазе, то она будет раскачивать систему — это значит, что автоколебания возможны. Если же сила противоположна скорости по фазе, то она будет тормозить колебательное движение, которое становится затухающим. В этом случае автоколебания невозможны. Итак, мы должны представить себе, что колебания в системе тем или иным способом возбуждены, затем исследовать фазовые соотношения и установить, возможны ли нарастающие колебания.

Для выполнения этого анализа нужно:

1. Выделить замкнутую цепь колебательных величин, включая обратную связь.

2. Выделить то основное звено, фазные соотношения в котором играют решающую роль (например, звено, связывающее силу со скоростью или со смещением, или напряжение с током и т. п.).

3. Определить тем или иным способом фазные сдвиги во всех остальных звеньях.

4. Просуммировать все эти сдвиги и вычесть сумму из нуля; результат даст искомый сдвиг в основном звене.

Для того чтобы облегчить себе эту операцию, полезно составить функциональную схему системы в виде той самой цепи взаимосвязанных величин, о которой только что говорилось. Все эти соображения мы поясним на примерах.

В качестве первого примера выберем гармонный язычок. В этой системе мы имеем дело со следующими величинами (см. § 10): $p_1 - p_0$ — номинальная разность давлений, Δp — перепад давления в камере, обусловленный ускорением воздуха в камере, p — рабочая разность давлений, меньшая номинальной на величину перепада, x — смещение язычка, u — объемная скорость воздуха. Взаимосвязь между всеми этими величинами может быть представлена функциональной схемой рис. 73.

На этой схеме $p_1 - p_0$ — источник энергии; переменные величины Δp , p , x , u образуют замкнутую цепь взаимодействия; цепь замкнута обратной связью с u на Δp .

Искомой величиной является фазный сдвиг в звене $p \rightarrow x$. Нужно установить фазные сдвиги во всех остальных звеньях.

1. Звено $\Delta p \rightarrow p$. Рабочая разность давлений p равна постоянной номинальной разности $p_1 - p_0$ за вычетом перепада. Значит, переменная составляющая p сдвинута по фазе относительно Δp на 180° . Этот сдвиг отмечен на схеме.

2. Звено $x \rightarrow u$. Связь между смещением язычка и изменением объемной скорости известна: она выражается падающей характеристикой. Значит, между x и u также имеется сдвиг на 180° .

3. Звено $u \rightarrow \Delta p$ (обратная связь). Перепад пропорционален ускорению воздуха, т. е. производной от скорости по времени. Значит, в звене $u \rightarrow \Delta p$ имеется сдвиг $+90^\circ$.

Результирующий сдвиг составляет $+90^\circ$. Следовательно, в остающемся звене сдвиг фазы между p и x составляет -90° , т. е. сила опережает смещение на 90° , что и требуется для автоколебаний.

Фазные соотношения очень наглядно представляются векторными диаграммами теории переменных токов. Эти диаграммы можно применять и в случае несинусоидальных колебаний, относя их к первым гармоникам рассматриваемых переменных величин. Для системы гармонного язычка векторная диаграмма изображена на рис. 74.

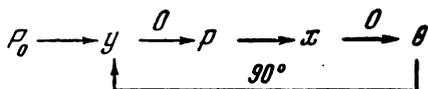
В рассмотренном примере интересно вот что: имеется падающая характеристика зависимости u от x . Это, однако, не та падающая характеристика, которая непосредственно обуславливает появление в системе отрицательного сопротивления. Но всякая падающая характеристика означает сдвиг фаз на 180° между переменными составляющими величин, взаимозависимость которых данная падающая характеристика выражает. Таким образом, мы видим, что падающая характеристика зависимости u от x играет в данном случае решающую роль в образовании надлежащих фазных соотношений. Если бы не было этой падающей характеристики, то сила не совпала бы по фазе со скоростью, а была бы с нею в противофазе, или, что то же, не опережала бы смещение язычка на 90° , а, наоборот, отставала бы на 90° и автоколебания были бы невозможны.

В качестве второго примера рассмотрим паровую машину. Здесь мы введем ряд обозначений, которые не фигурировали в нашем описании действия паровой машины в § 6. Именно, обозначим: x — смещение поршня, y — смещение золотника, p — давление пара в цилиндре, θ — угол поворота кривошипа и вала. Величины x и θ будем отсчитывать от одного и того же положения; тогда они будут совпадать по фазе. Смещение золотника y и давление пара p будем также считать совпадающими по фазе. Тогда схема и векторная диаграмма для рассматриваемого случая

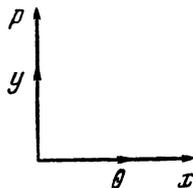
будут иметь вид, показанный на рис. 75 и 76. В звеньях $y \rightarrow p$ и $x \rightarrow \theta$ фазный сдвиг равен нулю. В звене $\theta \rightarrow y$ (обратная связь) фазный сдвиг равен $+90^\circ$, так как движение золотника происходит с опережением (точнее, сдвиг составляет $90^\circ + \delta$, см. § 6). Таким образом, результирующий сдвиг составляет $+90^\circ$, а искомый сдвиг в звене $p \rightarrow x$ составляет -90° , как оно и должно быть.

Само собою разумеется, что фазный сдвиг между силой и смещением не должен составлять ровно 90° . Достаточно хотя бы небольшого опережения, т. е. наличия составляющей по фазе со скоростью; возможность самовозбуждения устанавливается затем исследованием энергетического баланса.

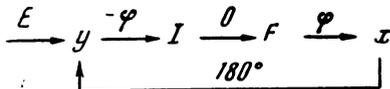
Рассмотрим еще один пример — электрический звонок. Здесь для целей анализа удобно ввести еще одну условную переменную: функцию коммутации, которая выражает проводимость цепи для постоянного тока. Обозначим ее через y ; тогда, если I — ток, F — сила, x — смещение якоря, схема будет иметь вид рис. 77. В звене $I \rightarrow F$ фазного сдвига нет. Что же касается звена обратной связи $x \rightarrow y$, то из зависимости между этими величинами, представленной на рис. 78, следует, что фазный сдвиг между ними составляет 180° (речь идет, разумеется, о первых гармониках). В звене $y \rightarrow I$ имеется запаздывание на некоторый угол φ , обусловленное влиянием индуктивности и искрообразованием. Стало быть, в зве-



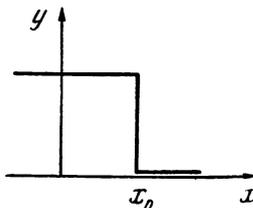
Р и с. 75



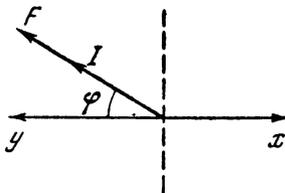
Р и с. 76



Р и с. 77



Р и с. 78



Р и с. 79

не $F \rightarrow x$ величина x сдвинута по фазе относительно F на угол $180^\circ + \varphi$, т. е. имеется требуемое опережение на угол $180^\circ - \varphi$, как показано на векторной диаграмме рис. 79.

В заключение этого параграфа нужно пояснить, что наличие благоприятных фазных соотношений еще не означает возможности существования устойчивых автоколебаний. Необходимое для этого условие состоит в том, что доставляемая в систему энергия должна покрывать потери. И энергетические соотношения можно, впрочем, исследовать при помощи векторных диаграмм. Для этого нужно строить диаграммы с соблюдением некоторого определенного масштаба и наряду с движущимися силами, обусловленными действием обратной связи, строить на диаграмме также тормозящие силы, находящиеся в противофазе со скоростью и зависящие от потерь и от реакции потребителя. Условие существования устойчивых автоколебаний будет тогда с точки зрения векторной диаграммы состоять в том, что сумма проекций всех сил на вектор скорости равна нулю.

Все сказанное в этом параграфе представляет собою изложенные в несколько необычном виде соображения, относящиеся к известному в теории автоколебаний условию «баланса фаз и амплитуд».

Эти соображения приводят к известным критериям устойчивости Найквиста, Цыпкина, Михайлова. Однако эти критерии здесь не разбираются и вот почему: во-первых, их применение требует знания количественных соотношений — мы же ограничиваемся рассмотрением только качественной стороны дела, во-вторых, эти критерии приспособлены для описания свойств системы в целом, тогда как для нас представляет интерес выяснение роли каждого звена сложной системы в отдельности.

§ 15. Релаксационные колебания

Пересматривая различные виды автоколебаний, мы различаем два типа колебаний, называемых обычно почти гармоническими¹ и релаксационными.

Внешнее различие этих двух типов колебаний есть различие в форме. Форма почти гармонических колебаний близка к синусоидальной; их частота определяется обычно собственной частотой задающей колебательной системы.

Релаксационные колебания резко несинусоидальны. Они зачастую принимают форму разрывных колебаний, когда наблюдаемая величина изменяется скачком.

Существенное различие почти гармонических и релаксационных колебаний определяется различием в порождающих эти коле-

¹ Почти гармонические колебания называют иногда томсоновскими; это название связано с формулой Томсона, выражающей собственную частоту системы через ее параметры (например, индуктивность и емкость).

бания механизмах и, в частности, в свойствах колебательных систем.

Если колебательная система состоит из двух элементов, в каждом из которых может запасаться энергия, то процесс колебания состоит в периодическом обмене запасом энергии между обоими элементами. Например, для случая механической колебательной системы, составленной из массы и упругости, запас энергии системы то сосредоточивается в виде кинетической энергии движущейся массы, то переходит в потенциальную энергию деформированного упругого элемента. Аналогично обстоит дело для электрического колебательного контура, составленного из индуктивности и емкости. Свободные колебания таких систем мало отличаются от синусоидальных. Посредством обратной связи можно превратить такие системы в почти гармонические автоколебательные.

Иначе обстоит дело в релаксационной системе. Одним из ее признаков является то, что в ней имеется не два накопителя энергии (например, индуктивность и емкость), а лишь один (например, емкость).

«Периодическое движение в системе становится по существу обменом энергии между внешней средой и системой, в которой имеется лишь один накопитель энергии.

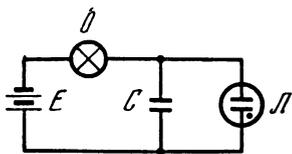
Обмен энергией происходит следующим образом: имеется система, из которой энергия притекает к накопителю, и, вообще говоря, другая система, куда энергия уходит из накопителя. Колебание складывается в этом случае из двух частей:

- а) приток энергии из первой системы к накопителю,
- б) отток энергии от накопителя во вторую систему.

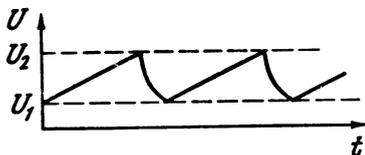
Накопитель энергии является как бы ее транзитным складом, попеременно наполняющимся и пустеющим. Конечно, вся физическая картина колебаний этого рода резко отличается от картины томсоновских колебаний» [5, стр. 163].

Основываясь на этих соображениях, можно составить представление о том, каково должно быть действие клапана в релаксационной системе. Клапан должен иметь два положения: «открыт» и «закрыт». В положении «закрыт» клапан отключает от накопителя сток (потребляющую систему); в этом положении происходит накопление энергии. В положении «открыт» клапан общедает накопитель со стоком, в результате чего накопитель отдает свой запас энергии. Но при этом совершенно очевидно, что клапан должен открываться при одном значении энергии накопителя, а закрываться при другом, а именно, меньшем значении. Таким образом, характеристика клапана должна характеризоваться перепадом между указанными двумя значениями.

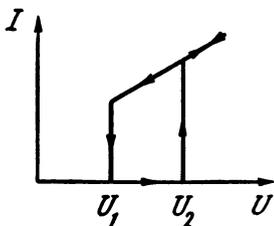
Из предыдущего следует, что релаксационные колебания должны иметь вполне определенный размах, зависящий от перепада характеристики клапана. Что же касается частоты релакса-



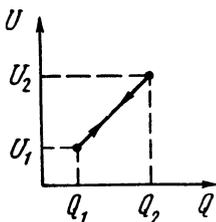
Р и с. 80



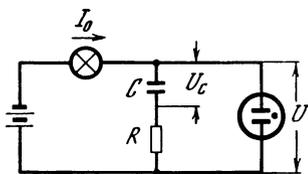
Р и с. 81



Р и с. 82



Р и с. 83



Р и с. 84

ционных колебаний, то она зависит от скорости процессов накопления и расходования запасов энергии накопителя, т. е., в конечном счете, от параметров системы.

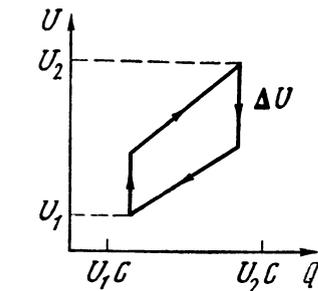
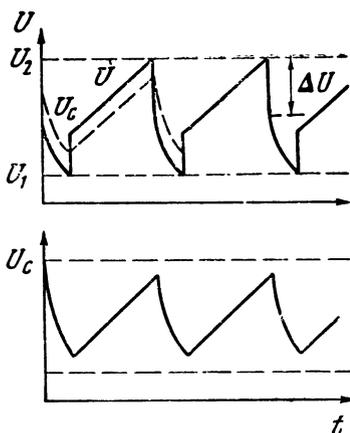
Все эти представления можно уяснить себе, рассмотрев какую-либо конкретную релаксационную систему. В предыдущем мы имели пример релаксационной системы: это — пневматический молоток. Единственный накопитель энергии в его устройстве — масса поршня, способная накапливать кинетическую энергию и отдавать ее при ударе.

В качестве другого примера мы рассмотрим генератор с неоновой лампой. Схема генератора изображена на рис. 80. Она состоит из батареи E , ограничителя тока O (например, диода в режиме насыщения), конденсатора C и неоновой лампы L (лампы тлеющего разряда). Действие этой схемы описывается обычно следующим образом: с момента включения батареи в конденсатор течет зарядный ток, постоянство которого поддерживается ограничителем. Вследствие этого заряд конденсатора, а следовательно, и напряжение на нем, нарастает линейно со временем. Когда напряжение достигнет напряжения зажигания неоновой лампы L , лампа вспыхивает (т. е. происходит ионизация заключенного в ней разреженного газа и возникает большая ионная проводимость). Конденсатор быстро разряжается через проводимость лампы, пока напряжение не упадет до значения U_1 — напряжения потухания. Лампа гаснет (т. е. ионизация, а с нею и проводимость пропадают), после чего возобновляется процесс

зарядки конденсатора, и так далее. В результате получается периодическое изменение напряжения между значениями U_1 и U_2 , как показано на осциллограмме рис. 81. Характеристика неоновой лампы как клапана показана на рис. 82.

Это описание само по себе не встречает возражений, однако попытка построить диаграмму работы сразу наталкивается на затруднения. В каких переменных строить диаграмму? В данной чисто электрической системе мы должны выбрать переменные U — напряжение и Q — заряд. При этом понимается, что U — внешнее напряжение, доставляемое источником. Что же касается Q , то эта величина должна выражать заряд в колебательной системе. Что является в данной схеме колебательной системой? Эту роль играет, очевидно, конденсатор, периодически заряжающийся и разряжающийся. Клапаном является неоновая лампа, управляющая чередованием процессов заряжения и разряжения конденсатора. Обратная связь с колебательной системы на клапан состоит в том, что неоновая лампа зажигается и гаснет в зависимости от напряжения на конденсаторе. Таким образом, полный состав автоколебательной системы налицо. Однако напряжение на конденсаторе всегда пропорционально его заряду, и наоборот. Никакой двузначности, ни статической, ни динамической, в этой зависимости нет и быть не может. Значит, диаграмма работы имеет вид, показанный на рис. 83, т. е. представляет собою отрезок прямой, независимо от вида осциллограммы, и, следовательно, автоколебания невозможны. Между тем схема работает. Дело усложняется еще и тем, что источник все время расходует энергию на покрытие потерь в неоновой лампе, через которую протекает пульсирующий ток, так что мощность, теряемая в лампе, имеет периодическую составляющую. Однако энергетические взаимоотношения источника энергии (т. е. батареи) с лампой не имеют прямого отношения к автоколебаниям; в колебательную систему, как мы убедились, энергия не поступает.

Нам не удастся свести концы с концами вот по какой причине, связанной с особенностями данной схемы. Дело в том, что в этой схеме нужно учесть наличие какого-либо потребителя колебательной энергии. В ранее рассмотренных случаях этого не нужно было делать: мы просто предполагали, что вложенная в колебательную систему энергия так или иначе расходуется. При этом механизм, пополняющий убыль энергии в колебательной системе, действует независимо от наличия потребителя. В данной же схеме энергия в колебательную систему вкладывается лишь в меру ее расходования. Следовательно, в отличие от ранее рассмотренных случаев в схеме с неоновой лампой амплитуда установившихся колебаний не зависит от энергетического баланса. И действительно, как мы видели, в схеме без потерь (рис. 80) устанавливается колебательный режим с вполне определенной амплитудой, определяемой разницей между напряжениями зажигания и потухания. Таким обра-



Р и с. 86

Р и с. 85

зом, для данной схемы график зависимости E_+ и E_- от амплитуды нечего строить, так как E_+ и E_- всегда равны друг другу.

Введя в схему потребитель колебательной энергии, мы сможем построить диаграмму работы. Пусть последовательно с конденсатором включено омическое сопротивление R (рис. 84). По этому сопротивлению будет проходить переменный ток, и, следовательно, в нем будет выделяться некоторая мощность. Благодаря наличию сопротивления общее напряжение U уже не будет равняться напряжению U_C на конденсаторе. Во время зарядки конденсатора U_C меньше U на величину падения напряжения на сопротивлении R . Эта разность остается во время зарядки постоянной, так как величина тока поддерживается ограничителем. После зажигания лампы напряжение на ней меньше напряжения на конденсаторе, так как ток, идущий через сопротивление R , меняет свое направление и знак падения напряжения меняется на обратный. Таким образом, напряжение U в момент зажигания лампы изменяется скачком. Величина этого скачка обозначена через ΔU на последующих рисунках. На осциллограммах рис. 85 изображено изменение во времени напряжений U и U_C . По этим осциллограммам, принимая во внимание, что заряд Q прямо пропорционален U_C , можно построить диаграмму работы, которая и изображена на рис. 86. При увеличении сопротивления R энергия, поглощаемая за цикл, а следовательно, и доставляемая источником, будет возрастать. При этом площадь диаграммы работы будет возрастать за счет роста вертикального отрезка ΔU .

На этом мы закончим исследование генератора с неоновой лампой. В следующих параграфах нам еще встретятся системы с аналогичными свойствами.

§ 16. Фрикционные автоколебания

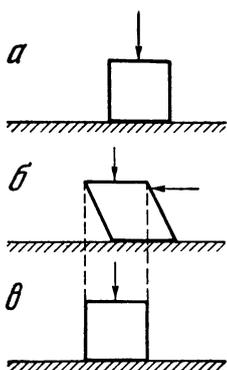
Фрикционные (т. е. обусловленные трением) релаксационные автоколебания — пожалуй, наиболее распространенный вид автоколебаний. Фрикционные колебания в форме различных скрипов воспринимаются нами постоянно. Скрипит паркет под нашими ногами, скрипит кожаная обувь, скрипят вилка по тарелке и палец по стеклу, скрипят двери и ворота, скрипит пробка флакона, скрипит (скрежещет) трамвай на закруглении пути, скрипит гвоздь, вытягиваемый из доски. Совершенно аналогичные явления происходят и в скрипке, на что прямо указывает краткое и точное название этого музыкального инструмента.

Наблюдая перечисленные и многие другие явления, мы замечаем следующие три обстоятельства: во-первых, скрип наблюдается повсеместно, откуда следует, что условия самовозбуждения этого вида автоколебаний легко осуществляются. Во-вторых, эти колебания достигают иногда очень большой мощности, о чем можно судить хотя бы по слуховым ощущениям. И, в-третьих, этот вид колебаний обладает поразительной устойчивостью. Какая-нибудь калитка может на протяжении многих лет скрипеть своим особым «голосом», не подвергаясь никаким изменениям.

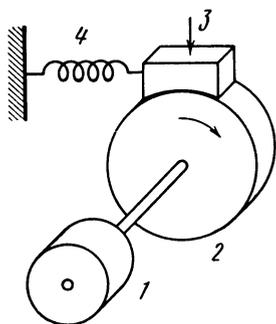
Скрип возникает при скольжении одного тела по другому. В общих чертах картина явления такова: когда на упругое тело, прижатое к другому, действует касательная сила, то сперва относительного движения по поверхности соприкосновения обоих тел нет, так как они сцеплены между собою силой трения покоя. Поэтому приложенная сила вызывает только деформацию тела. Когда же при увеличении приложенной силы и соответственном возрастании деформации упругая сила превзойдет, наконец, силу трения покоя, произойдет срыв. Тело скачком перейдет в новое положение, потеряв упругую деформацию, после чего все будет происходить в том же порядке. На рис. 87 все это показано схематически: *a* — упругое тело прижато к другому вертикальной силой; *b* — приложена горизонтальная (касательная) сила; накапливается упругая деформация; *в* — произошел срыв. Упругое тело заняло новое положение.

То, что колебания описанного типа зависят именно от трения, всем хорошо известно. Мы знаем из повседневного опыта, что стоит смазать трущиеся поверхности (маслом — металлические, мылом — деревянные), чтобы скрип исчез, хотя бы на некоторое время.

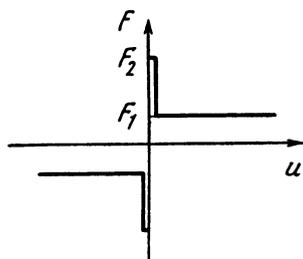
Закономерности фрикционных автоколебаний удобно изучить на какой-либо простой модели, показанной на рис. 88, где изображен двигатель *1*, равномерно вращающий маховик *2*. К ободу маховика прижата колодка *3*, удерживаемая пружиной *4*. Колодка увлекается маховиком при его вращении. Этому препятствует пружина; при известном растяжении она отрывает колодку и



Р и с. 87



Р и с. 88



Р и с. 89

возвращает ее на прежнее место. Все это возможно благодаря резкому различию между силами трения, покоя и скольжения. Поэтому для фрикционных автоколебательных систем основное значение имеет характеристика трения, т. е. зависимость силы трения от относительной скорости трущихся тел.

Предположим для начала (упрощая действительные соотношения), что характеристика трения имеет вид, показанный на рис. 89, т. е. что сила трения покоя (при относительной скорости, т. е. скорости скольжения, равной нулю) имеет значение F_2 большее, чем при скольжении. При этом предполагается, что сила трения скольжения F_1 не зависит от относительной скорости.

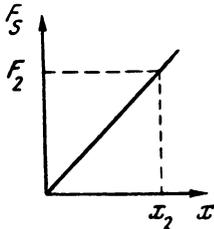
Зная характеристику трения и упругость пружины, можем найти то положение x_2 колодки, в котором происходит срыв. Это положение определяется по характеристике пружины, т. е. по зависимости упругой силы F_s от растяжения x , как показано на рис. 90. Обычно пренебрегают для упрощения влиянием массы колодки. Тогда график ее движения представится ломаной линией, образованной на осциллограмме рис. 91.

Однако, хотя на первый взгляд это и не очевидно, при сделанных предположениях (т. е. при характеристике трения вида рис. 89 и в пренебрежении массой колодки) автоколебания невозможны. Действительно, для возобновления цикла, т. е. для того чтобы после срыва колодка вновь увлеклась маховиком, необходимо, чтобы относительная скорость колодки и маховика хотя бы на мгновение стала равной нулю; тогда в действие вступит трение покоя. Но это невозможно, так как после срыва колодка либо движется навстречу маховику, либо стоит на месте. И в последнем случае относительная скорость не равна нулю —

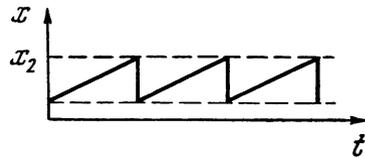
она равна скорости маховика. Таким образом, сила трения есть сила трения скольжения; она лишь дает некоторое постоянное натяжение пружине.

Следовательно, для того чтобы получить описание действительно возможного явления, мы должны отказаться от одного из двух упрощений, которые мы вначале приняли. Отбросим второе упрощение, т. е. примем во внимание влияние массы колодки.

Тогда с момента срыва скорость колодки будет изменяться по почти синусоидальному закону. Частота этой синусоиды —

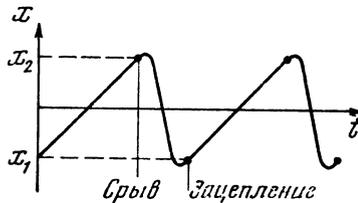


Р и с. 90



Р и с. 91

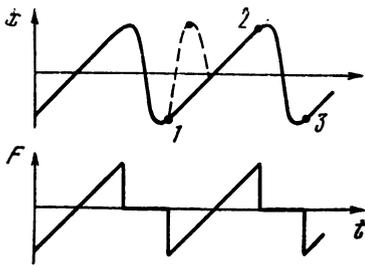
это собственная частота колебательной системы, составленной из упругости пружины и массы колодки. При возвратном движении колодки (на рис. 88 влево) она по инерции проходит то положение, в котором пружина не напряжена; она продолжает двигаться в том же направлении, она сжимает пружину, а затем, достигнув крайнего левого положения, под действием сжатой пружины начинает двигаться с ускорением вправо, т. е. в ту же сторону, что и маховик. При этом в некоторый момент скорости колодки и маховика сравниваются, т. е. относительная скорость становится равной нулю. В этот момент колодка сцепляется с маховиком и начинает двигаться вместе с ним до следующего срыва. Описанное движение изображено на рис. 92, на котором $x=0$ означает положение колодки при ненапряженной пружине.



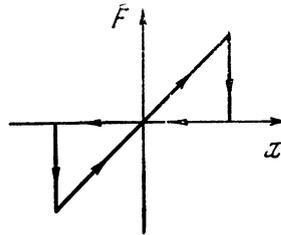
Р и с. 92

Итак, при разрывной характеристике трения вида рис. 89 наличие незатухающих автоколебаний может быть объяснено только с учетом массы колебательной системы.

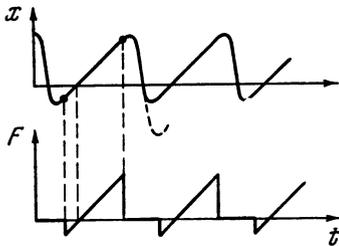
Анализируя состав автоколебательной системы рис. 88, мы разыскиваем ее главнейшие части: источник энергии — двигатель, колебательную систему — пружину (или пружина плюс масса колодки). Роль клапана играет механизм сцепления, который либо связывает колодку с маховиком трением покоя, либо



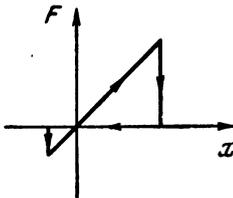
Р и с. 93



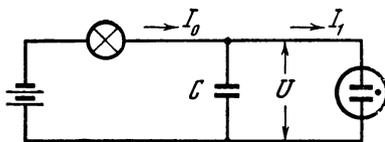
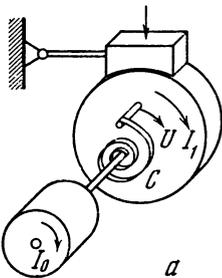
Р и с. 94



Р и с. 95



Р и с. 96



Р и с. 97

отпускает колодку в момент срыва. Обратная связь с колебательной системой на клапан состоит в том, что упругая сила пружины обуславливает срыв колодки в определенном ее положении.

Мы видим, что рассмотренная система во многом аналогична генератору с неоновой лампой. Эта система также обладает тем свойством, что амплитуда установившихся колебаний определяется не балансом энергий, а фиксированными предельными значениями смещения. Точно так же и диаграмма работы может быть построена лишь после введения в схему потребителя энергии.

На рис. 93 изображены осциллограммы смещения и силы. Сцепление осуществляется на прямолинейном участке 1—2. На этом участке движение равномерное, т. е. без ускорения. Поэтому наличие массы не оказывает влияния, и внешняя сила работает только против силы упругости пружины. Сперва пружина сжата и внешняя сила отрицательна. Лишь на второй половине участка 1—2 внешняя сила растягивает пружину.

Диаграмма работы показана на рис. 94. Она состоит из двух треугольников, обходимых в противоположных направлениях. В силу полной симметрии осциллограммы оба треугольника оди-

наковы и, следовательно, работа внешней силы равна нулю. Так оно и должно быть — ведь мы не учитывали потерь энергии колебательной системой. Если же эти потери учесть, то положение изменится. Энергия будет теряться на протяжении свободного хода системы, т. е. на участке 2—3. Движение будет, следовательно, представляться на этом участке *з а т у х а ю щ е й* синусоидой; размах в сторону отрицательных x будет меньше; симметрия осциллограммы нарушится. Осциллограммы получают вид рис. 95, а соответствующая диаграмма работы — рис. 96. Как видим, треугольник, обходимый по часовой стрелке, больше; это означает, что источник силы совершает работу, которая, очевидно, идет на покрытие потерь.

Интересно заметить, что небольшим видоизменением устройства фрикционной автоколебательной системы можно привести ее к полному физическому подобию схеме генератора с неоновой лампой. Такое видоизмененное устройство изображено на рис. 97, а.

Тут же, на рис. 97, б повторена схема генератора с неоновой лампой. Механическое устройство состоит по-прежнему из двигателя, маховика, пружины и колодки, но теперь пружина включена между двигателем и маховиком, а колодка неподвижна. Для того чтобы подчеркнуть аналогию, на рис. 97, а и б аналогичные величины отмечены одинаковыми буквами. Аналогия такова:

Механическая система

Постоянная скорость вращения, обеспечиваемая двигателем.
Момент закручивания пружины.
Пульсирующая скорость маховика.

Электрическая система

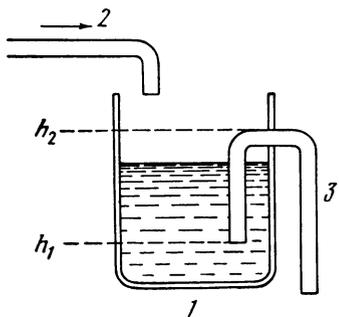
Постоянный ток, создаваемый батареей с ограничителем.
Напряжение на конденсаторе.
Пульсирующий ток через неоновую лампу

Эта аналогия настолько полна, что почти все рассуждения, относящиеся к генератору с неоновой лампой, могут быть (после «перевода на механический язык») повторены применительно к фрикционной автоколебательной системе, данной на рис. 97, а. Заметим мимоходом, что построить электрическую схему, аналогичную системе рис. 88, не удается.

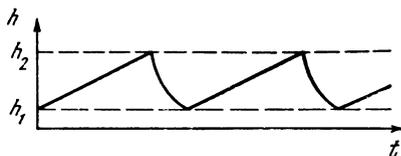
§ 17. Примеры гидравлических систем

Рассмотрим примеры автоколебательных систем с движением жидкости. Одна из наиболее простых автоколебательных систем этого рода — так называемый сосуд Гантала. Устройство его показано на рис. 98. Сосуд 1 наполняется водой, вытекающей с постоянной скоростью из питающей трубы 2. Сосуд снабжен сифоном 3. Когда уровень воды в сосуде достигнет высоты h_2 , сифон вступает в действие и быстро опорожняет сосуд до уровня h_1 . Дальнейшее действие сифона прекращается, так как вода из него

уходит. Уровень воды снова начинает повышаться, и так далее. Таким образом, возникает периодическое изменение уровня, зависимость которого от времени показана рис. 99. Описанное устройство постоянно приводится в качестве модели генератора с неоновой лампой (см. § 15). И действительно, аналогия получается довольно полная. Бак уподобляется конденсатору, уровень



Р и с. 98



Р и с. 99

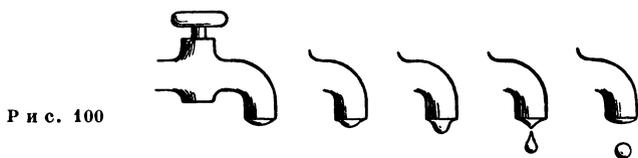
воды в баке — напряжению на конденсаторе, количество воды в баке — заряду конденсатора. Наполнение бака соответствует заряданию конденсатора, опорожнение — разряданию. Сифон выполняет функции, аналогичные функциям неоновой лампы. Момент, когда уровень достигает верхнего колена сифона, вследствие чего сифон начинает действовать, соответствует моменту, когда напряжение на конденсаторе достигает напряжения зажигания. Понижение уровня воды до h_1 соответствует падению напряжения на конденсаторе до напряжения потухания неоновой лампы. В обеих системах происходит, таким образом, срыв при определенном значении колебательной величины, управляющей работой клапана (т. е. уровня воды в случае сосуда Тантала и напряжения в случае генератора с неоновой лампой).

Нужно, впрочем, заметить, что вполне аналогичные явления происходят в еще более простой водяной модели. Речь идет об образовании капель.

Пусть из неплотно завернутого крана понемногу сочится вода. Она будет образовывать на срезе крана выпуклую поверхность (рис. 100), охватывающую постепенно возрастающий объем воды. Здесь находятся в равновесии следующие силы: сила поверхностного натяжения, сила сцепления частиц воды и сила тяжести. Последняя все время растет, так как вода непрерывно прибывает. Наконец, сила тяжести преодолевает противодействующие силы, и происходит срыв: капля срывается. Процесс, происходящий, как все это наблюдали, со строгой периодичностью, также представляет собою полнейшую аналогию релаксационным процессам в сосуде Тантала и в генераторе с неоновой лампой.

Рассмотрим более сложную по устройству и действию систему, известную под названием гидравлического тарана. Этот прибор, изобретенный в 1798 г., служит в качестве насоса, приводимого в действие водой и поднимающего воду на более высокий уровень, чем первоначальный (разумеется, ценою уменьшения количества поднимаемой воды).

Устройство гидравлического тарана показано схематически на рис. 101. Вода из сосуда 1 по трубе 2 попадает в коробку 3, снабженную двумя клапанами — выпускным K_1 , открываемым внутрь, и рабочим K_2 , открываемым наружу в воздушный колпак 4. Сюда нагнетается под давлением вода, поднимающаяся затем по трубе 5 на более высокий уровень.



Действие прибора состоит в следующем. Вначале клапаны закрыты: клапан K_1 гидростатическим давлением столба h_1 , клапан K_2 — столба H . Для запуска системы нужно нажать клапан K_1 (жесткое самовозбуждение!). Вода начнет вытекать через K_1 ; под действием напора h_1 ее скорость постепенно будет увеличиваться. При этом на клапане K_1 образуется возрастающий перепад давления, который, наконец, закроет клапан, преодолев его вес. Истечение внезапно прекратится, и произойдет явление гидравлического удара, состоящее в том, что давление в камере резко повысится. Механизм этого явления, разобранный в свое время Н. Е. Жуковским, таков: если движущийся столб воды внезапно останавливается, то его кинетическая энергия превращается в потенциальную, т. е. давление в столбе воды резко повышается. Образовавшееся сжатие распространяется со скоростью звука в направлении, противоположном первоначальному движению воды. Вследствие малой сжимаемости воды давление могло бы достигнуть колоссальных значений. Однако явление смягчается тем, что всякая труба не вполне жестка и может упруго растягиваться внутренним давлением. Тем не менее в обычной водопроводной трубе вследствие гидравлического удара возможно повышение давления в 10—12 раз. Этого совершенно достаточно для того, чтобы разорвать трубу. Вот почему все водопроводные краны делаются винтовыми: они могут перекрыть струю лишь постепенно.

Вернемся, однако, к тарану. Повысившееся вследствие гидравлического удара давление в коробке 3 открывает клапан K_2 и вода поступает в колпак 4. Заключенный в нем воздух служит в качестве буфера, смягчающего толчки и обеспечивающего более

равномерное течение воды по трубе 5. Когда волна сжатия распространится по трубе 2 до сосуда 1, она отразится оттуда в виде волны разрежения. В момент прихода этой волны в коробку клапан K_2 закроется, а клапан K_1 откроется — на этот раз автоматически, и в дальнейшем весь процесс повторится ¹.

В реальном устройстве частота колебаний достигает порядка 100 в минуту. Высота подъема H может превышать h_1 в 20 раз. На рис. 102 показаны осциллограммы скорости воды в трубе 2 и давления в камере 3. Последняя величина выражена через высоту водяного столба. Ход явления таков: в момент a клапан K_1

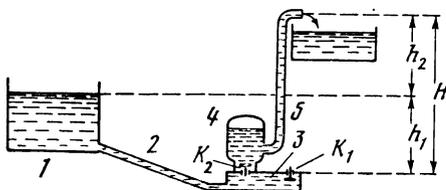


Рис. 101

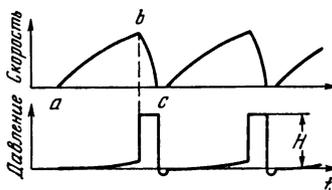


Рис. 102

открывается, и вода течет в коробку со все возрастающей скоростью. При этом давление в коробке равно атмосферному, так как клапан K_1 открыт. Однако когда скорость достигает больших значений, давление в коробке несколько повышается, и в момент b клапан закрывается. С этого момента скорость быстро убывает, а давление в коробке подскакивает вследствие гидравлического удара. После остановки воды (момент c) приходит волна разрежения, открывающая клапан K_1 .

§ 18. Релаксационные и почти гармонические колебания

Хотя различие между обоими названными типами колебаний и велико, но, как это обычно бывает, граница между ними нерезка и возможен непрерывный переход одного типа колебаний в другой. Изучение деталей этого перехода как раз и позволяет наилучшим образом усвоить различие между почти гармоническими и релаксационными колебаниями и понять, что оба эти типа колебаний являются в определенном смысле предельными случаями некоторой более общей картины колебаний. Выше уже говорилось, что внешнее различие между почти гармоническими и релаксационными колебаниями состоит в том, что первые синусоидальны,

¹ Следует заметить, что объяснение явлений в гидравлическом таране без учета разрежения невозможно. Между тем в некоторых описаниях волновые явления игнорируются, и картина процесса остается вследствие этого совершенно непонятной.

вторые же имеют резко несинусоидальный, а иногда и разрывной характер. Вторая внешняя отличительная черта состоит в том, что амплитуда релаксационных колебаний не зависит от нагрузки.

Внутреннее различие обоих типов колебаний определяется прежде всего различием в свойствах колебательных систем. В то время как почти гармоническая колебательная система состоит из двух накопителей, обменивающихся между собою энергией, в релаксационной системе имеется только один накопитель, периодически запасаящий энергию, а затем отдающий ее.

Второе внутреннее различие состоит в том, что в релаксационной системе работа клапана представляет собою скачкообразный периодический переход из одного положения в другое («открыт — закрыт») при различных значениях энергии накопителя, тогда как механизм действия почти гармонической колебательной системы описывается обычно при помощи понятия падающей характеристики или отрицательного сопротивления. Из этого последнего противопоставления следует, между прочим, что для того, чтобы плавный переход релаксационных колебаний в почти гармонические был возможен, необходимо, чтобы характеристика клапана обладала обоими признаками: и перепадом и падающим участком.

О возможности колебаний различного рода в одной и той же системе уже упоминалось в связи с дуговым генератором (§ 13), который может генерировать как малые почти синусоидальные колебания на падающем участке характеристики, так и релаксационные колебания, связанные с периодическим гашением дуги.

Мы рассмотрим более подробно другой пример.

Вернемся к фрикционной автоколебательной системе (рис. 88). Если масса колодки пренебрежимо мала, то в системе происходят чисто релаксационные колебания по пилообразному закону, изображенному на рис. 91. Если увеличить массу, то форма колебаний несколько изменится и будет соответствовать рис. 92. Если бы мы увеличивали массу и далее, то пришли бы постепенно к синусоидальным колебаниям. Но экспериментально осуществить постепенное увеличение массы затруднительно. Мы можем получить тот же переход не путем изменения массы, а путем изменения скорости вращения маховика.

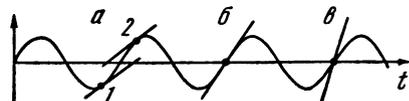
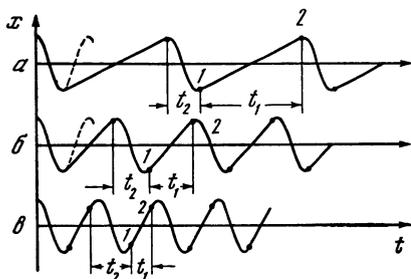
Рассмотрим рис. 103, на котором изображены колебания при различных скоростях маховика.

a — скорость маховика, представляемая графически наклоном прямолинейного участка кривой колебания, мала. Промежуток времени t_1 , на протяжении которого колодка равномерно движется вместе с маховиком, значительно больше промежутка t_2 , на протяжении которого совершается обратное движение; график движения имеет почти пилообразную форму, т. е. мы имеем почти предельный случай чисто разрывных колебаний.

b — скорость значительно увеличена. Разница между промежутками t_1 и t_2 заметно сократилась.

v — скорость еще увеличена. Теперь уже t_2 больше, чем t_1 . Форма колебаний почти синусоидальна.

При дальнейшем увеличении скорости t_1 уменьшается, стремясь к нулю, а t_2 увеличивается, стремясь к периоду свободных колебаний системы. Напомним, что при малой скорости t_2 немного более половины периода свободных синусоидальных колебаний системы. Предельные соотношения поясняются еще вспомогательным рис. 104, где представлено синусоидальное свободное колебание. Здесь же изображены прямые, наклон которых пред-



Р и с. 104

Р и с. 103

ставляет скорость маховика. Сцепление и срыв происходят в точках, в которых наклон этих прямых и наклон касательной к синусоиде, дающий мгновенное значение колебательной скорости, совпадают. Цифрами 1 и 2 на рис. 103 и 104 обозначены соответственно точки сцепления и срыва. Рис. 104, а соответствует примерно соотношениям рис. 103, в. Буквой б на рис. 104 обозначен предельный случай, когда наибольшая мгновенная скорость сравнивается со скоростью маховика. Наконец, буквой в обозначено такое положение, когда колодка вообще не может быть подхвачена маховиком, так как скорость последнего всегда больше скорости колодки. Следовательно, при очень большой скорости автоколебания уже невозможны. Однако здесь есть еще одна тонкость. Наблюдая все описанные явления экспериментально, мы отмечаем одно обстоятельство, резко бросающееся в глаза, а именно: пока колебания имеют пилообразный характер, с увеличением скорости вращения частота колебаний растет, а амплитуда остается неизменной (см. рис. 103, а и б). Когда же колебания приобрели уже синусоидальный характер, то при увеличении скорости вращения частота остается неизменной — ведь она равна собственной частоте системы, — зато амплитуда растет до поры до времени, пока колебания не сорвутся вовсе.

Но если амплитуда растет, то растет и колебательная скорость (при неизменной частоте). Следовательно, возможность сцепления сохраняется до больших значений скорости, чем это следовало бы из рис. 104 в предположении постоянства амплитуды. Наблюдается следующее явление: колебания прекращаются во время увеличения скорости при гораздо большем значении скорости,

чем то, при котором они возникают, если постепенно уменьшать скорость.

Описанные явления легко объясняются при помощи элементарных энергетических соображений. При увеличении скорости возрастает развиваемая двигателем мощность. Естественно предположить, что при этом возрастает и мощность колебаний. Но мощность колебаний растет с колебательной скоростью, которая равна произведению частоты колебаний на их амплитуду. Следовательно, при постоянной амплитуде мощность может возрастать за счет увеличения частоты, а при постоянной частоте — за счет увеличения амплитуды. Эти закономерности мы и наблюдаем в вышеописанном опыте.

Итак, во фрикционной автоколебательной системе возможен плавный переход от релаксационных колебаний к почти гармоническим. Такого рода переход может наблюдаться и в других релаксационных системах по мере того, как начинает играть все более заметную роль второй накопитель энергии, влиянием которого в чисто релаксационной системе пренебрегают. Таким образом, становится ясным, что релаксационная система есть вырожденный предельный случай системы более общего вида.

§ 19. Термопрерыватель

Биметаллическая пластинка, изгибающаяся при нагреве и размыкающая электрический контакт, составляет основу целого ряда устройств, применяемых в различных отраслях техники.

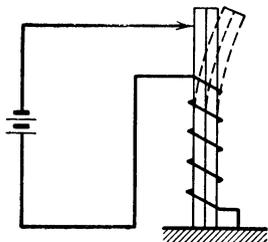
Укажем прежде всего на терморегулятор — устройство для поддержания постоянства температуры, применяемое в термостатах, в которых не требуется высокой степени постоянства температуры¹. Действие регулятора состоит в том, что биметаллическая пластинка управляет цепью электрического нагревателя непосредственно или при помощи реле. Когда температура дойдет до установленного верхнего предела, пластинка изогнется и разомкнет контакт. Когда же температура в термостате понизится и пластинка охладится, она выпрямится и снова замкнет контакт. Температура начнет повышаться, и так далее.

Процесс регулирования имеет, таким образом, колебательный характер; описанный терморегулятор является автоколебательной системой. Ее составные части: источник энергии — источник тока, питающего нагреватель, колебательная система — нагреватель с термостатом. Колебания имеют тепловую природу; они состоят в периодическом изменении температуры под влиянием периодического подвода и отвода тепла. Клапаном служит биме-

¹ Для точных термостатов применяются более «деликатные» регуляторы типа газовых термометров с контактом. Впрочем механизм автоколебаний системы термостат—регулятор не зависит от конструкции регулятора; все существенные черты явлений остаются неизменными.

таллическая пластинка, включающая и выключающая нагреватель. Обратная связь состоит в том, что окружающая среда нагревает или охлаждает биметаллическую пластинку, заставляя ее соответственным образом деформироваться.

Сходным образом устроены многие системы передачи показаний, например электрические дистанционные термометры, применяемые в автомобилях для передачи на приборный щиток сведений о температуре воды и масла. В этой последней системе как датчик, так и приемник представляют собою биметаллические пластинки, снабженные нагревательной обмоткой. В датчике



Р и с. 105

имеется контакт, так что пластинка датчика совершает автоколебания. Но кроме нагрева током, протекающим по нагревательной обмотке, пластинка датчика получает еще дополнительный нагрев от окружающей среды. Повышение температуры среды изменяет среднее положение пластинки; режим датчика изменяется в том смысле, что общее время, в течение которого контакт разомкнут, возрастает. Это вызывает уменьшение среднего тока, проходящего через приемник. Последний представляет собой по существу тепловой амперметр с большой постоянной времени. Его шкала проградуирована непосредственно на температуру.

Если в терморегуляторе автоколебания представляются хотя неизбежным, но в общем нежелательным явлением, то в целом ряде устройств создание автоколебаний является основным их назначением. Таковы, например, всевозможные «мигалки» — устройства для периодического (с большим периодом) включения и выключения огней — рекламных и сигнальных (например, освещение пожарных сигналов, автомобильных указателей поворота и т. п.). Схема такого терморерывателя крайне проста; она показана на рис. 105. Биметаллическая пластинка состоит из двух слоев с различными коэффициентами теплового расширения. На рис. 105 предполагается, что левый слой имеет больший коэффициент расширения. Если так, то при нагревании пластинка изогнется вправо, как намечено на рисунке штрихом.

Рассмотрим действие терморерывателя, заметив, что ток есть однозначная функция положения пластинки ¹.

¹ Фазный сдвиг, игравший решающую роль в случае электромагнитного

Когда по нагревателю начинает проходить ток, то выделяется тепло, нагревающее пластинку. Скорость нагревания зависит от теплоемкости пластинки. Чем теплоемкость больше, тем медленнее повышается температура. Пренебрегая потерями тепла, можно полагать, что температура пластинки нарастает со временем линейно (если не меняется ток, питающий нагреватель). Когда же ток прекратился, пластинка начинает отдавать тепло более холодной окружающей среде и сама остывает. Скорость остывания зависит от условий теплоотдачи, т. е. от внешней теплопроводности пластинки.

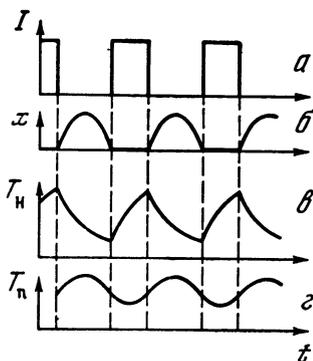
Однако такое описание явлений никак не объясняет возникновения автоколебаний. Действительно, стоит только контакту замкнуться, как дополнительное количество тепла, выделившееся в нагревателе, выгнет пластинку и ослабит контакт. А так как характеристика реального контакта может и не иметь разрыва, то возможно равновесное положение без всяких колебаний. Это равновесие характеризуется тем, что выделяется ровно столько тепла, сколько отдается окружающей среде (при неизменном токе).

Следовательно, в нашем описании пропущена какая-то существенная деталь. И действительно, мы упустили из вида теплоемкость самого нагревателя, а главное — условия теплообмена между нагревателем и пластинкой. В действительности благодаря конечной теплопроводности между нагревателем и пластинкой рост температуры пластинки отстает по отношению к нагревателю. Поэтому, когда контакт уже разомкнулся и выделение тепла прекратилось, дальнейший нагрев пластинки еще продолжается за счет передачи тепла от более горячего нагревателя; пластинка далеко отойдет от контакта, прежде чем ее температура сравняется с температурой нагревателя, после чего она начнет вместе с ним охлаждаться. Когда же в результате охлаждения контакт снова замкнется, пластинка будет еще в течение некоторого времени охлаждаться, пока вследствие повышающейся температуры нагревателя не образуется возрастающий поток тепла в пластинку. Этот поток сперва только скомпенсирует теряемое пластинкой тепло (в этот момент прекратится ее охлаждение), и лишь после этого она начнет нагреваться.

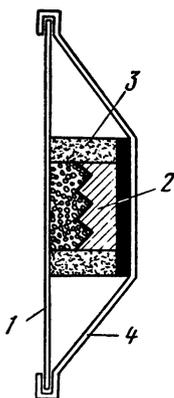
Все эти явления показаны на осциллограммах рис. 106 (*a* — изменение тока в цепи нагревателя, *b* — отклонение пластинки, *c* — температура нагревателя, *z* — температура пластинки).

Интересно заметить, что мы не встретились бы с затруднением, если бы начали с рассмотрения автоколебаний терморегулятора:

прерывателя, здесь не принимается во внимание, во-первых, потому, что колебания терморегулятора происходят с малой частотой (порядка десятых долей герца), а во-вторых, потому, что в цепи принципиально может не быть индуктивности — нагревательная обмотка вовсе не обязана сю обдавать; например, она может быть бифилярной.



Р и с. 106



Р и с. 107

в этом случае ясно видно, что температуры нагревателя и пластинки различны. Однако именно благодаря очевидности этого обстоятельства (в случае регулятора) мы рисковали бы не оценить должным образом его решающее значение.

Данная система довольно сложна в том отношении, что мы имеем, в сущности, колебания двоякой природы — механические и тепловые. Мы можем по нашему желанию рассматривать колебательную систему как механическую, а все тепловые явления отнести к механизму обратной связи. Можем поступить и иначе — рассматривать тепловые колебания, а механические явления отнести к механизму управления. Мы не будем развивать эти соображения.

В заключение отметим еще интересные тепловые автоколебания, наблюдаемые иногда в угольных микрофонах. Конструкция такого микрофона показана на рис. 107. Между угольной диафрагмой 1 и угольной колодкой 2 засыпан угольный же порошок, удерживаемый войлочным кольцом 3, не препятствующим колебаниям диафрагмы. Диафрагма завальцована по краю в корпус 4. При включении постоянного тока диафрагма может совершать медленные (с периодом порядка минут) автоколебания.

Механизм этих автоколебаний таков. При прохождении тока порошок нагревается и нагревает диафрагму. Будучи закреплена по краю, она выпучивается наружу, ослабляя давление на порошок. Сопротивление порошка возрастает, ток резко убывает, количество выделяемого тепла также убывает, так что система начинает охлаждаться. Диафрагма возвращается назад, надавливает на порошок, его сопротивление падает, ток возрастает, интенсивный нагрев возобновляется, и так далее.

Бывают случаи, когда диафрагма имеет тенденцию прогибаться внутрь. При этом никаких автоколебаний не возникает; происходит монотонное увеличение тока, приводящее к гибели

микрофона. Микрофоны с такой тенденцией представляют собою производственный брак. В некоторых конструкциях ставятся даже специальные пружинки, отжимающие диафрагму наружу.

Легко усмотреть полнейшую аналогию между тепловыми автоколебаниями микрофона и автоколебаниями терморезервуара

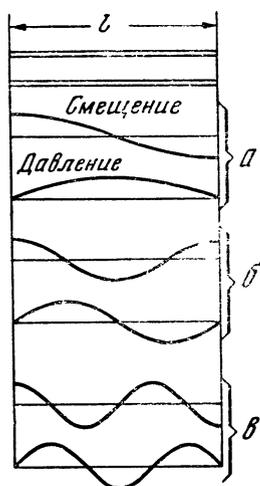
§ 20. Автоколебания в трубах

В трубах могут быть возбуждены мощные автоколебания воздушного столба. Хотя практическое использование этого вида автоколебаний ограничивается созданием музыкальных, а также сигнальных звуков, мы уделим трубам довольно большое внимание, так как в качестве примеров автоколебательных систем они дают нам много поучительного.

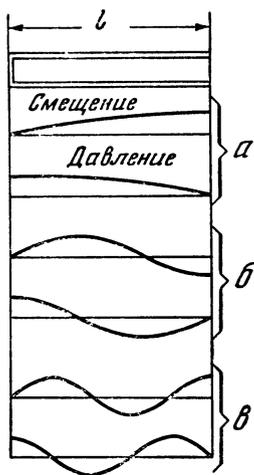
Автоколебания в трубах характерны тем, что задающей колебательной системой является труба, т. е. система с распределенными постоянными. Поэтому перед тем, как рассматривать отдельные частные случаи, необходимо в двух словах напомнить свойства труб и основные относящиеся к ним понятия.

Самое существенное — это то, что в трубе распространяются с конечной скоростью звуковые волны, чем и определяются интересующие нас свойства трубы. Эти свойства могут рассматриваться с двух точек зрения: с точки зрения установившегося синусоидального режима и с точки зрения неустановившегося режима. Нам понадобятся и те и другие представления.

1. У с т а н о в и в ш и й с я р е ж и м. В установившемся режиме труба рассматривается как отрезок волновой системы с определенными условиями, заданными на концах трубы. Если труба возбуждается на одном конце, то вследствие отражения от другого конца в трубе образуется стоячая волна. При определенных соотношениях возникают резонансы трубы, причем в отличие от системы с сосредоточенными постоянными труба обладает бесчисленным множеством резонансов, частоты которых для трубы неизменного сечения образуют гармоническую, или почти гармоническую, последовательность. Говорят также, что труба может возбуждаться как в основной частоте, так и в обертонах. Различные обертоны характеризуются возникновением в трубе того или иного числа узлов. Например, основная частота, или основной тон, открытой с обоих концов трубы соответствуют показанному на рис. 108 распределению вдоль трубы давления и смещения частиц воздуха. На открытых концах трубы всегда образуются пучности смещения (и скорости) частиц и узлы давления. Посреди трубы образуются для основного тона узел смещения и пучность давления. Как видим, при возбуждении в основном тоне на длине трубы укладывается половина длины звуковой волны данной частоты. Следующая возможная форма колебаний показана на рис. 108, б. Это — первый обертон, характеризующийся уже двумя узлами смещения. При колебаниях



Р и с. 108

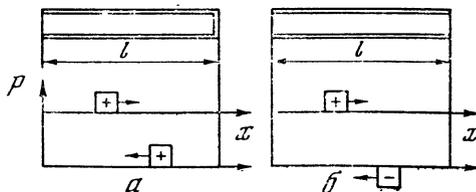


Р и с. 109

в первом обертоном на длине трубы укладывается целая длина волны. Для второго обертона на длине трубы укладывается три полуволны (рис. 108, в) и так далее. Таким образом, труба может быть возбуждена на частотах, для которых длина трубы составляет целое число полуволн, или четное число четвертей волн.

Для трубы с одним закрытым концом на этом конце получается узел смещения и пучность давления, а на открытом конце — пучность смещения и узел давления. На длине трубы при основном тоне укладывается четверть волны (рис. 109, а), при первом обертоном — три четверти волны (рис. 109, б), при втором обертоном — пять четвертей волны и так далее. Таким образом, закрытая на одном конце труба имеет собственные частоты, определяемые тем, что на длине трубы укладывается нечетное число четвертей длины волны. Основной тон трубы, закрытой с одного конца, вдвое ниже, чем у трубы такой же длины, открытой с обоих концов.

2. Н е у с т а н о в и в ш и й с я р е ж и м. Рассмотрение трубы в неустановившемся режиме состоит в том, что мы наблюдаем за распространением по трубе импульсов. Под импульсом понимается малая по сравнению с длиной трубы область, в которой тем или иным способом задано измененное значение какой-либо из величин, характеризующих звуковое поле. Так, например, если мы говорим о распространении по трубе импульса сжатия, то по-



Р и с. 110

нимаем под этим, что по трубе передвигается со скоростью звука область, внутри которой давление выше атмосферного. Наибольшее значение для последующего имеют закономерности отражения импульсов. Если импульс сжатия отражается от закрытого конца, то после отражения он возвращается обратно также в виде импульса сжатия. Если же отражение происходит от открытого конца, то импульс сжатия превращается в результате отражения в импульс разрежения.

Можно условиться описывать явления при помощи понятия избыточного давления; тогда при сжатии избыточное давление положительно, при разрежении — отрицательно.

Картина отражения импульсов представлена на рис. 110. Рис. 110, *а* изображает отражение импульса давления от закрытого конца трубы, рис. 110, *б* — от открытого конца трубы.

Для импульсов скорости или смещения получаются обратные соотношения: импульс скорости сохраняет знак при отражении от открытого конца и меняет знак на обратный при отражении от закрытого конца.

Переходя к вопросу о роли трубы в составе автоколебательной системы, заметим, что во всех рассмотренных ниже примерах независимо от механизма возбуждения принято, что труба является задающей системой. Иными словами, частота автоколебаний почти совпадает с одной из собственных частот трубы. Мы не можем получить произвольную частоту автоколебаний; изменяя параметры возбудителя, мы можем лишь перескочить с одного обертона на другой. Хроматическая последовательность, получаемая в любом духовом инструменте, обеспечивается изменением рабочей длины трубы посредством клапанов, включающих дополнительные колена. В тромбоне возможно и плавное изменение длины (а следовательно, и частоты) благодаря наличию выдвижной кулисы. Число клапанов в духовом инструменте обычно невелико. Те или иные комбинации клапанов (так называемая аппликатура) позволяют обычно получить хроматический звукоряд в пределах октавы. Весь диапазон инструмента перекрывается совместным действием клапанов и «передувания», т. е. перескока на обертон трубы.

Духовые музыкальные инструменты делятся на язычковые (кларнет, гобой, саксофон и др.) и мундштуковые (труба, тромбон и др.). В последних возбудителем являются губы музыканта, вдавливаемые в чашечку мундштука и работающие на манер голосовых связок.

В органе в высоких регистрах также применяются трубы с язычковым возбуждением. Рассмотрим сперва этот способ возбуждения.

В параграфе, посвященном гармонному язычку, язычок рассматривался как задающая система, т. е. автоколебания имели частоту язычка. Собственные же частоты примкнутой акустической системы не принимались во внимание (на этом основании

к входной камере и не применялся ремесленный термин «резонатор»).

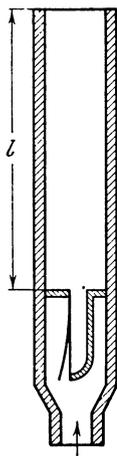
В случае же трубы, возбуждаемой язычком, мы имеем обратное положение вещей: труба является задающей системой, а собственная частота язычка игнорируется. Конечно, строго говоря, мы имеем дело с так называемой связанной системой. Однако язычок, применяемый для возбуждения трубы, делается обычно мягким и с частотой более низкой, чем собственная частота трубы. Движение язычка управляется давлением, действующим изнутри трубы. Самое большое, что мы можем сделать, это, воздействуя на язычок, получить небольшие поправки на частоту трубы. Надо сказать, что эта возможность постоянно используется, так как получить в точности весь необходимый звукоряд в духовых инструментах практически невозможно. Труба всегда немного фальшивит; чтобы избежать этого, музыкант напрягает губы и с силой вжимает их в мундштук. Таким образом ему удается немного изменить частоту автоколебаний¹. Этим же путем достигается перескок на обертоны. При игре на кларнете музыкант короче и крепче прихватывает язычок губой.

Труба возбуждаемая язычком, изображена схематически на рис. 111. Механизм самовозбуждения удобно рассматривать, наблюдая за распространением и действием отдельных импульсов, т. е. кратковременных сгущений и разрежений. Пусть в первый момент язычок отогнут, как показано на рис. 111. Сжатый воздух врывается в трубу и создает в ее начале сгущение — импульс сжатия. Сразу же давление сжатого воздуха прижимает язычок, который, действуя как клапан, замыкает вход в трубу и прекращает дальнейшее поступление в нее сжатого воздуха. Импульс сжатия бежит вдоль трубы. Достигнув открытого конца, он отражается от него уже в виде импульса разрежения, который распространяется в обратном направлении. Претерпев отражение в начале трубы, которое является закрытым концом, импульс разрежения бежит снова к открытому концу, откуда возвращается в виде импульса сжатия. В момент его прибытия к началу трубы давления внутри и снаружи трубы почти сравниваются, и язычок снова открывает вход в трубу. Продолжительность цикла равна учетверенному времени пробега импульсом длины трубы. Отсюда определяется частота основного тона, которая, конечно, получается такой же, как была определена выше на основании рассмотрения стоячих волн.

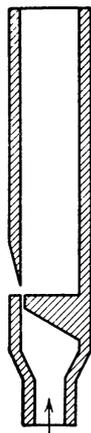
¹ Интересно отметить, что новые необыгранные трубы «жестче»: на них трудно играть. Со временем труба обыгрывается. Дело заключается в том, что с течением времени на внутренних стенках трубы откладывается слой окислов. Затухание трубы возрастает и резонанс становится менее острым. Следовательно, легче вызвать небольшие изменения частоты. Понятно, что обыгрывание можно заменить предварительной обработкой внутренних стенок трубы. Эти подробности могут показаться излишними. Однако они малоизвестны, а между тем представляют несомненный физический интерес.

Из приведенного описания процесса видно, что язычок работает как клапан, управляемый со стороны колебательной системы, т. е. со стороны трубы. В этом и состоит механизм обратной связи в рассматриваемой системе. Можно отметить еще следующее обстоятельство. Если бы колебания воздуха в трубе происходили без потерь, то давление в начале трубы по окончании вышеописанного цикла было бы в точности равно первоначальному давлению, т. е. давлению, развиваемому источником сжатого воздуха. В таком случае даже при открытом входе в трубу сжатый воздух в нее поступать не будет, так как внутреннее давление равно внешнему. В трубе происходили бы при таких условиях незатухающие свободные колебания. Но в действительности имеются потери, и весьма значительные, в частности потери на звукоизлучение. Поэтому импульсы ослабевают, и получаемое к концу цикла внутреннее давление меньше внешнего. В результате этого в трубу поступает при открытии язычка свежая порция сжатого воздуха, доводящая давление в импульсе сжатия до первоначального значения. Таким образом происходит автоматическое поддержание энергетического баланса. При повышении внешнего давления возрастает, естественно, мощность колебаний.

В более тонкие детали явлений мы вдаваться не будем, так как нам предстоит рассмотреть еще ряд механизмов самовозбуждения.



Р и с. 111



Р и с. 112

Заканчивая описание язычкового механизма, упомянем еще о том, что точно такое же устройство применяется в пневматических сигнальных приспособлениях, например в автомобильных гудках (в настоящее время вытесненных электрическими) и в мощных сигнальных аппаратах, устанавливаемых на побережье в качестве звуковых маяков.

Обратимся теперь к классической безязычковой органной трубе. Ее устройство изображено на рис. 112. Сжатый воздух поступает в камеру, выходное отверстие которой имеет форму сравнительно узкой щели. Струя воздуха, выходящая из щели, попадает на клиновидно скошенную часть стенки трубы.

Органная труба являлась издавна — за недостатком других объектов — предметом многочисленных исследований. В новейшее же время, когда как раз и теория и экспериментальная техника достигли высокой степени развития, интерес к органной трубе почти совсем угас. Поэтому наши сведения о явлениях в органной трубе основываются на сравнительно старых исследованиях. Господствующие представления о механизме звуко-

образования следует считать сомнительными. Сущность этих представлений состоит в том, что звук возбуждается посредством так называемых клиновых тонов, возникающих при попадании струи воздуха на клин. Механизм возникновения клиновых тонов связан с вихреобразованием (о клиновых тонах и вихрях подробнее говорится в § 21). Экспериментальным подтверждением изложенной точки зрения является то, что если вдуть в трубу дым, то при стробоскопическом наблюдении ясно видны образующиеся вихри. Картина вихреобразования запечатлена на многочисленных фотографиях; их можно найти в любом более старом учебнике акустики, в котором еще рассматриваются органичные трубы.

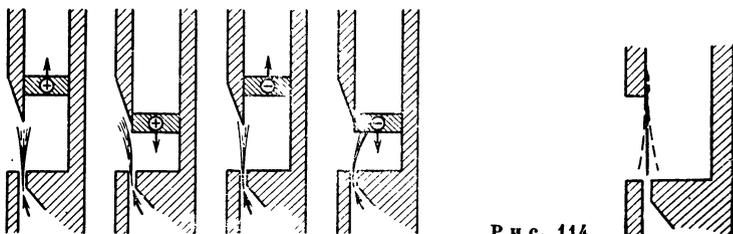
В последующем изложении принята другая точка зрения. Она состоит в том, что хотя вихри и возникают — отрицать это не приходится, — но их образование есть побочное явление, не определяющее механизм возбуждения трубы. Основной же механизм представляется в следующем виде.

Предположим, что в начальный момент в результате вдувания воздуха в начале трубы возникает импульс сжатия. Этот импульс побежит к открытому концу трубы и, отразившись от него, вернется в виде импульса разрежения. Когда импульс разрежения достигнет начала трубы, то давление в начале трубы станет меньше внешнего атмосферного давления. Под действием разности давлений струя воздуха, вытекающая из щели, отклонится внутрь трубы и создаст в ней повышенное давление. В результате этого разность давлений изменит знак и струя отклонится наружу. В то же время новый импульс сжатия побежит по трубе. Один цикл совершается за время пробега импульса вдоль трубы в обоих направлениях. Так оно и должно быть, так как в данном случае мы имеем дело с трубой, открытой с двух концов. Если труба закрыта на удаленном конце (такие закрытые трубы также применяются в органах), то описание хода явлений соответственно изменяется. Именно, первый импульс сжатия возвращается после отражения от закрытого конца трубы также в виде импульса сжатия. Возникающая разность давлений отклоняет струю наружу. Отражившись от открытого конца, импульс снова пускается в путь в виде импульса разрежения. После нового отражения от закрытого конца импульс разрежения возвращается опять к началу трубы. В этот момент происходит втягивание струи внутрь трубы и кладется начало следующему циклу. Таким образом, в случае закрытой трубы продолжительность цикла равна учетверенному времени пробега импульсом длины трубы в один конец. Последовательность явлений иллюстрируется рис. 113, на котором импульсы сжатия и разрежения обозначены соответственно знаками (+) и (—). Итак, истекающая из щели струя рассматривается как клапан, управляемый со стороны трубы. Механизм обратной связи совершенно ясен. Специальные опыты, которые здесь по недостатку места не описываются, показывают, что струя действительно чрезвычайно чувствительна к боковому

напору и заметно изменяет свое направление при самых малых разностях давлений по обе стороны струи.

Заметим попутно, что описание явлений несколько не изменилось бы, если бы мы видоизменили устройство трубы, снабдив ее самым настоящим клапаном в виде подвижного язычка, как приблизительно показано на рис. 114.

Рассмотрим еще один механизм возбуждения автоколебаний в трубе. В 1939 г. Б. П. Константинов произвел в Научно-исследовательском институте музыкальной промышленности чрезвычайно эффектный опыт. Он создавал в открытой на свободном



Р и с. 113]

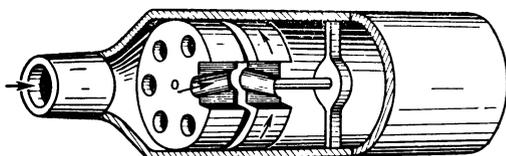
Р и с. 114

конце трубе вынужденные колебания при помощи обычной сирены, вставленной в другой конец трубы. По мере повышения давления высота звука, естественно, повышалась, пока не достигалась высота собственного тона трубы. При дальнейшем повышении давления высота звука, т. е. частота колебаний, оставалась неизменной; повышение давления вызывало лишь увеличение мощности колебаний. Это явление трактовалось тогда как стабилизация сирены вследствие захватывания. По существу же описанное устройство представляет собою типичную автоколебательную систему. Разберем ее действие.

Сирена состоит из двух дисков — статора и ротора, — снабженных одинаковым числом косо просверленных отверстий (рис. 115). Эти отверстия играют роль лопаток турбины, так что сирена представляет собою не что иное, как миниатюрную воздушную турбину. Но отверстия статора и ротора сирены играют еще и другую роль: они прерывают воздушную струю, проходящую через сирену. Воздух проходит через сирену, когда отверстия ротора и статора совпадают. Когда же при повороте ротора его отверстия окажутся против промежутков между отверстиями статора, прохождение воздуха прекратится. Таким образом, при вращении ротора сирены через нее проходит прерывистая (модулированная) струя воздуха. Частота модуляции пропорциональна числу отверстий и скорости вращения ротора. Последняя же тем больше, чем больше давление. Даже при неизменном давлении ротор сирены набирает скорость постепенно. Этим обусловлен всем известный воющий характер звука сирены. Итак, сирена есть клапан, управляемый вращательным движением ротора.

Рассмотрим теперь взаимоотношения сирены с примкнутой

к ней трубой. Начнем с момента, когда отверстия статора и ротора сирены совпадают. В этот момент сжатый воздух врывается в трубу и создает в ее начале импульс сжатия. Этот импульс, отразившись от открытого конца в виде импульса разрежения, возвращается назад к сирене. Здесь он может произвести различное действие, в зависимости от того, каковы скорость вращения сирены и время пробега импульса по трубе, так как от соотношения этих величин зависит относительное положение статора и ротора сирены к моменту возвращения импульса. Мы разберем два различных случая.



Р и с. 115

1. За время пробега импульса по трубе туда и обратно ротор поворачивается на величину шага отверстий. Другими словами, к моменту прибытия отраженного импульса отверстия ротора и статора совпадают. При этом для вращения ротора сирены создаются благоприятные условия, так как вследствие разрежения внутри трубы рабочая разность давлений, приводящая ротор во вращение, больше, чем если бы разрежения не было. Таким образом, при описанных условиях реакция со стороны трубы создает дополнительный момент, ускоряющий вращение ротора сирены.

2. За время пробега импульса в оба конца ротор сирены поворачивается на половину шага, так что отраженный импульс встречает закрытый конец. Не произведя никакого действия, импульс отражается, бежит к открытому концу, вновь отражается и возвращается вновь в виде импульса сжатия. Теперь он застает отверстия совпавшими и проход для воздуха открытым. Но так как давление внутри трубы повышено, то рабочая разность давлений меньше, чем при отсутствии колебаний. Следовательно, при описанных условиях колебательная реакция со стороны трубы создает дополнительный момент, тормозящий ротор сирены.

Два описанных случая относятся к скоростям ротора, различающимся вдвое. Совершенно ясно, что на этом интервале найдется такая скорость, при которой реакция трубы не будет оказывать ни тормозящего, ни ускоряющего действия. Это и будет установившаяся скорость, «стабилизованная» реакцией трубы. В самом деле, если бы ротор почему-либо пошел медленнее, реакция трубы подогнала бы его; если же он пошел бы быстрее, реакция трубы подтормозила бы его. Частота же звука, отвечающая этой стабильной скорости, определяется длиной трубы.

Таков механизм обратной связи в этой своеобразной автоколебательной системе. Что касается энергетических соотношений,

то в установившемся режиме трубы с сиреной к ней применимы те же соображения, что и к трубе, возбуждаемой при помощи язычка.

В заключение этого параграфа разберем еще один интересный тепловой механизм возбуждения автоколебаний. Речь идет о явлении, открытом Рийке в 1859 г. и описанном у Рэля [2, стр. 227] следующим образом:

«Когда кусок частой металлической сетки, натянутый на нижнюю часть трубы, открытой с обоих концов и расположенной вертикально, нагревается газовым пламенем, помещенном ниже ее, то почти непосредственно вслед за удалением пламени слышен звук значительной силы, длящийся несколько секунд. . . образование звука тесно связано с возникновением сквозной тяги, встречающей нагретую сетку.

Поддержание звука связано с переменной передачей тепла, являющейся результатом переменного движения воздуха через сетку, причем это движение складывается из равномерного движения вверх с движением попеременно вверх и вниз, обязанным колебанию».

Таково явление в общих чертах. Во всяком случае ясно, что источником энергии является нагретая сетка, отдающая свое тепло колеблющемуся воздуху.

К сожалению, мы не можем воспользоваться для исследования механизма данного явления тем же приемом, который применялся в предыдущих случаях, т. е. рассмотрением распространения и действия импульсов давления. Затруднение состоит вот в чем. Прямое действие источника энергии заключается в том, что сетка нагревает соприкасающийся с нею воздух и тем самым повышает его давление. Обратная же связь состоит в том, что тепло отдается воздуху в той или иной мере в зависимости от колебательной скорости воздуха в месте расположения сетки. Таким образом, действие каналов прямой и обратной связи обусловлено разными колебательными величинами, что сильно осложняет дело. Поэтому, отказываясь от детального разбора явления, мы ограничимся лишь общими соображениями.

Как сказано, переход тепла от сетки к воздуху вызывает повышение давления, имеющее колебательный характер. В трубе устанавливается половина волны давления с узлами по концам. Одновременно образуется стоячая полуволна скорости, имеющая узел посреди трубы и пучности по концам (см. рис. 108, а). Если мгновенная колебательная скорость направлена вниз, то условия нагрева будут неблагоприятны, так как на сетку будет попадать уже нагретый воздух, поднимающийся от сетки вверх. Если же колебательная скорость направлена вверх, то на сетку будет попадать свежий воздух снизу, и при этом получатся наиболее благоприятные условия для передачи тепла от сетки воздуху. Из этого следует, что при наличии колебаний тепло будет отбираться от сетки периодически, причем работой своеобразного теплового клапана управляет колебательная скорость. В этом и состоит механизм обратной связи в данной системе.

Очевидно, что для того чтобы обратная связь могла действовать, нельзя помещать сетку в узле скорости, т. е. посреди трубы. С другой стороны, для получения переменного давления нельзя ставить сетку и на конце трубы, так как там имеется узел давления. Значит нужно поместить горячую сетку где-то между узлами давления и скорости, скажем, на расстоянии четверти длины трубы от нижнего конца. Опыт подтверждает это заключение.

§ 21. Автоколебания, связанные с образованием вихрей

С незапамятных времен известно, что ветер порождает звуки. В частности, дуновение ветра может возбудить колебания струны. Это явление используется в так называемой эоловой арфе — рамке с настроенными в унисон струнами, подвешиваемой на открытом воздухе и издающей под влиянием ветра слабый мелодичный звук.

Первое количественное исследование эоловых тонов было сделано Струхалем (1878). Он изучал звукообразование при обдувании потоком воздуха неподвижных цилиндрических стержней и установил ряд простых закономерностей. Оказалось, что частота звука пропорциональна скорости потока и обратно пропорциональна диаметру стержня.

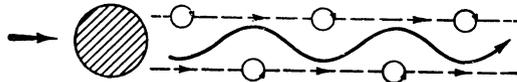
Кроме эоловых тонов давно уже наблюдались так называемые щелевые и клиновые тоны. При продувании воздуха через щель на фоне шума прослушивается музыкальный тон, впрочем довольно слабый и неустойчивый. Этот тон приобретает значительную силу и устойчивость, если перед щелью расположить клин, обращенный к щели острой гранью, так что выходящая из щели струя попадает на эту грань. Частота получаемого звука зависит от скорости потока, ширины щели и расстояния до клина.

Все эти явления получили объяснение после опубликования в 1912 г. исследований Кáрманна. Кáрман показал, что как при обтекании тела потоком, так и при истечении из щели образуются вихри, периодически срывающиеся с краев препятствия или щели и распространяющиеся затем в виде вихревых дорожек или цепочек. В каждой дорожке вихри располагаются на определенном расстоянии друг от друга и двигаются в направлении потока с определенной скоростью. Показано, что устойчивая форма вихревых дорожек отвечает поочередному отрыванию вихрей то с одной, то с другой стороны. Примерная картина вихревых дорожек, образующихся при обтекании потоком цилиндрического препятствия, показана на рис. 116.

Возникновение эоловых тонов объясняется с точки зрения представления о вихрях следующим образом: каждый раз при срыве вихря в области непосредственно за препятствием создается импульс давления. А так как срыв вихрей происходит периодически, то мы имеем дело с источником звука соответствующей частоты, расположенным около препятствия, позади него. Сама же вихревая дорожка источником звука не является [13].

Аналогичная картина возникает и при истечении воздуха из щели. Влияние клина может быть удовлетворительным образом объяснено путем рассмотрения взаимодействия клина с двумя вихревыми дорожками.

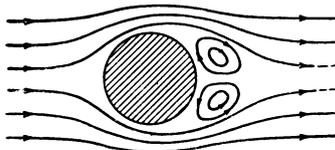
Механизм образования вихрей состоит в том, что позади препятствия образуется мертвое пространство, в котором среда находится в покое, т. е. не совершает поступательного движения в направлении потока. В пограничном слое между движущейся и покоящейся средами возникают касательные силы, обусловленные вязкостью среды. Эти силы приводят покоящуюся среду во



Р и с. 116

вращательное движение, как показано схематически на рис. 117, на котором изображены два симметрично расположенных зародыша вихрей. Благодаря небольшой начальной асимметрии один зародыш развивается за счет другого. Достигнув полного развития, вихрь срывается. Наступает очередь развития и срыва вихря с другой стороны, и так далее. Этот процесс в некоторых отношениях напоминает периодическое образование, созревание и срыв капель, о чем говорилось в § 17. Мы еще воспользуемся этой аналогией.

Но нас интересует главным образом не образование вихрей само по себе, а вихревое возбуждение колебаний различных механических и акустических колебательных систем. Наблюдалось и описано весьма большое число случаев такого возбуждения.



Р и с. 117

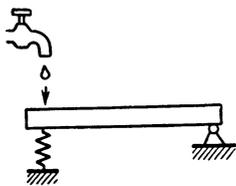
Об золотой арфе уже говорилось. Рэлей (1884) наблюдал колебания растопыренных пальцев при движении руки в воде, а также колебания движимого в воде гибкого стержня. В недавнее время Стрелков [14] производил опыты с маятником в форме шара, подвешенного на проволоке в аэродинамической трубе. Блохинцев [12] исследовал возбуждение резонаторов, помещая их в воздушный поток в аэродинамической трубе.

Во всех этих и многих других опытах отмечалось два обстоятельства. Во-первых, при колебаниях твердых тел (струны, стержня, шара) эти колебания совершаются в плоскости, перпендикулярной к направлению потока.

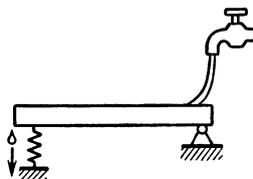
Во-вторых, во всех без исключения случаях наблюдается возбуждение механических и акустических колебательных систем в собственных частотах.

Первое обстоятельство легко понять, если связать колебания с возникновением вихрей: вихри срываются поочередно то с одной, то с другой стороны, вследствие чего создаются импульсы давления, действующие в поперечном направлении попеременно то в одну, то в другую сторону.

Первоначальное объяснение второго факта, естественно вытекающее из представления о периодическом срыве вихрей, состояло в том, что наблюдаемое явление есть не что иное, как обычный резонанс. Процесс вихреобразования доставляет периодическую вынуждающую силу, способную заметным образом раскачать данную колебательную систему только на собственной (резонансной) частоте.



Р и с. 118



Р и с. 119

В последнее время, в связи с развитием общего понимания автоколебаний, высказывается иная гипотеза, представляющаяся гораздо более правдоподобной. Эта гипотеза не отрицает роли вихреобразования в возбуждении колебаний, но предполагает, что существует и обратная связь, состоящая во влиянии колебательного движения системы на процесс образования и срыва вихрей. Таким образом, новое представление отличается от старого тем, что явление рассматривается не как вынужденное колебание, а как автоколебание. С этой точки зрения возбуждение в собственной частоте совершенно понятно.

К сожалению, об этих явлениях известно поразительно мало, принимая во внимание многолетнюю научную историю вопроса. Образование вихревых дорожек изучено только для двумерного случая, т. е. для случая обтекания тел в форме цилиндра или длинной пластинки. Картина вихреобразования при обтекании шара совершенно не исследована; высказано лишь предположение, что за шаром образуется винтовой вихрь. Что же касается деталей зарождения, развития и срыва вихрей и, в частности, связи этих явлений с движением обтекаемого тела, то об этом не известно решительно ничего.

Вопрос о связи вихреобразования с движением колебательной системы — вопрос об обратной связи — представляет исключительный интерес для исследования механизма автоколебаний и заслуживает подробного изучения.

А пока что для уяснения соотношения между старой и новой точками зрения на процесс возбуждения колебаний рассмотрим некоторую модель, в которой используется аналогия между образованием вихрей и образованием капель. Сама по себе эта

аналогия весьма поверхностна, но в нижеследующем ее возможности едва ли переоценены.

Пусть под кран, из которого капает вода, подставлен массивный желоб, закрепленный шарнирно на одном конце и снабженный пружиной, так что возможно колебательное движение желоба (рис. 118). Каждая падающая в желоб капля сообщает ему импульс силы. Капли падают периодически; если частота капель совпадает с собственной частотой желоба, то возможны вынужденные колебания его со значительной амплитудой. Совершенно ясно, что в этой модели движение желоба не оказывает никакого влияния на образование и срыв капель.

Видоизменим теперь модель, как показано на рис. 119. Теперь вода подается в желоб по вставленному в кран фитилю, а капли срываются с края желоба. При таком устройстве движение желоба уже непосредственно влияет на образование и срыв капель. Во-первых, вода будет стекать к концу желоба с ускорением, зависящим от наклона желоба и его угловой скорости. Приток воды к концу желоба, а следовательно, и рост капель зависят, таким образом, от движения желоба. Во-вторых, — и это самое существенное — срыв капель происходит под действием силы тяжести и добавляющейся к ней при движении желоба силы инерции, зависящей от линейного ускорения свободного конца. Нетрудно сообразить, что наиболее благоприятные условия для срыва капли создаются, когда желоб находится в крайнем нижнем положении. Действительно, в этом положении ускорение максимально и направлено вверх; сила, действующая на каплю и отрывающая ее, равна произведению массы капли на ускорение силы тяжести плюс максимальное значение колебательного ускорения. Говоря проще, желоб стряхивает с себя капли в нижнем своем положении. Итак, образование и срыв капель в определенной мере зависят от движения желоба ¹.

Возвращаясь к вихрям, рассмотрим снова рис. 117 и представим себе, что обтекаемое тело движется поперек потока. При этом в пограничном слое, прилегающем к его поверхности, возникнут обусловленные движением дополнительные вязкие силы. Направление этих сил таково, что они благоприятствуют развитию (а следовательно, и ускорению срыва) одного зародыша и, наоборот, препятствуют развитию другого зародыша. Так, например, при движении тела вверх (см. рис. 117) благоприятные условия развития создаются для верхнего зародыша. А так как вязкие силы пропорциональны скорости, то в системе могут создаться требуемые фазные соотношения, при которых возможны автоколебания. Эти простые общие соображения требуют, конечно, экспериментальной и теоретической проверки.

¹ В модели рис. 119 автоколебания невозможны, так как импульс силы, возникающий при срыве капли, действует в ненадлежащей фазе — в момент максимума ускорения, а не скорости. Для того чтобы автоколебания

По поводу вихревого механизма автоколебаний следует еще заметить, что наблюдается некоторое увлечение этими представлениями со времени первых работ о вихреобразовании. Это увлечение выражается в том, что вихреобразованием объясняют самовозбуждение и в тех случаях, когда роль вихрей вовсе не бесспорна. Так, например, автоколебания безъязычковой органичной трубы могут быть объяснены при помощи не вихревого, а струйного механизма, как это и сделано в § 20. Общеизвестное возбуждение низкого тона при вдувании воздуха под углом в горлышко пустой бутылки, вероятно, также должно быть отнесено за счет струйного механизма, по крайней мере при весьма малых скоростях потока, при которых этот вид колебаний возможен.

Практическое значение вихревого возбуждения колебаний, помимо звукообразования, состоит в том, что этот механизм может порождать мешающие нормальной работе или даже опасные колебания элементов конструкций. Речь может идти при этом о возбуждении колебаний как под действием естественного ветра, так и искусственного дутья (например, в мощных воздуходувных установках, в аэродинамических трубах и т. п.), а также о колебаниях элементов конструкции самолета.

§ 22. Флаттер [15]

В этом параграфе мы рассмотрим механизм автоколебательного явления, имеющего очень большое значение для самолетостроения. Речь пойдет об опасных колебаниях крыльев и хвостового оперения самолетов, возникающих при определенных скоростях полета и приводящих чрезвычайно быстро — за несколько секунд после возникновения — к полному разрушению и гибели самолета. Это явление носит профессиональное название **ф л а т т е р** (англ. flutter — трепетание).

Само собою разумеется, что изучению этого страшного явления было уделено самое серьезное внимание. В настоящее время его механизм достаточно выяснен, и всякая новая конструкция самолета рассчитывается так, чтобы флаттер не мог возникнуть.

Теория флаттера и связанные с ней расчеты сравнительно сложны; мы рассмотрим вопрос в самых общих чертах.

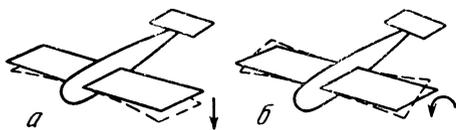
Механизм флаттера имеет чисто аэродинамическую природу. Энергия, необходимая для возбуждения автоколебаний, доставляется встречным потоком воздуха. Этот поток создает, как известно, подъемную силу, направленную вверх и поддерживающую самолет в воздухе. Другая сила, создаваемая потоком, — это направленная назад (т. е. против движения самолета) сила сопротивления.

Нас интересуют, однако, не эти постоянные силы, а те переменные силы, которые возникают при колебаниях крыла. Если

стали возможны, необходимо при помощи какого-либо приспособления задержать отделение капил от желоба.

эти силы направлены против скорости колебательного движения крыла, то они вызовут затухание колебаний (положительное сопротивление). Если же возникающие при колебаниях крыла переменные силы направлены в каждый момент по скорости колебательного движения, то такие силы будут раскачивать крыло и амплитуда колебаний будет нарастать вплоть до разрушения крыла (отрицательное сопротивление).

Таким образом, для того чтобы разобраться в механизме флаттера, нужно прежде всего выяснить вопрос о том, как возникают силы, совпадающие со скоростью по фазе, т. е. как образуется отрицательное механическое сопротивление.



Р и с. 120



Р и с. 121

Крыло, рассматриваемое как балка, заделанная одним концом, может совершать колебания двух основных видов: изгибные и крутильные (рис. 120, а и б). Определим силы, действующие на крыло, совершающее колебания изгиба во встречном потоке воздуха. При колебаниях изгиба сечение крыла, изображенное на рис. 121, движется вверх и вниз, оставаясь параллельным самому себе. Крыло несколько наклонено по отношению к направлению полета; угол α (рис. 121) называется углом атаки. Но если крыло, совершая колебания, движется в некоторый момент вверх, то это равносильно уменьшению угла атаки. Чтобы убедиться в этом, нужно представить себе траекторию частицы воздуха, скользящей по нижней поверхности крыла с учетом движения всего крыла вверх. Если же крыло движется вниз, то это равносильно увеличению угла атаки.

С другой стороны, подъемная сила зависит от угла атаки; в первом приближении приращение подъемной силы ΔF пропорционально приращению угла атаки.

Таким образом, при движении крыла в в е р х, когда угол атаки получает отрицательное приращение, возникает добавочная сила ΔF , направленная в н и з, т. е. против скорости колебательного движения. Когда же крыло движется в н и з, то угол атаки возрастает и возникает добавочная сила ΔF , направленная в в е р х.

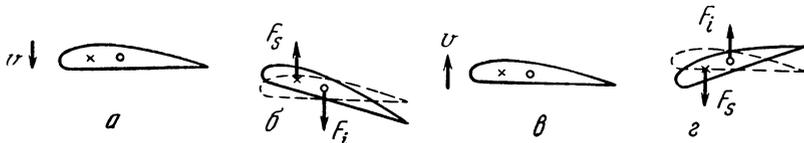
Итак, приращение подъемной силы направлено всегда против скорости. Следовательно, мы имеем положительное сопротивление колебательному движению, и чисто изгибные автоколебания крыла в полете возникнуть не могут.

Аналогичным образом можно убедиться в том, что невозможны и чисто крутильные автоколебания крыла. Это следует из того, что развивающиеся при крутильных колебаниях допол-

нительные аэродинамические моменты направлены против угловых скоростей сечения крыла.

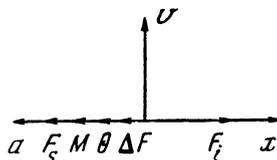
Для объяснения явления флаттера необходимо рассмотреть сложные изгибно-крутильные колебания крыла.

Оказывается, что чисто изгибные или чисто крутильные колебания крыла невозможны; один вид колебаний порождает другой, так что в действительности они существуют совместно. Иначе говоря, задача о колебаниях крыла есть задача о колебаниях связанных систем, или о колебании системы с двумя колебательными степенями свободы. Этим определяется относительная сложность проблемы флаттера крыла.



Р и с. 122

Р и с. 123



Разберем механизм связи обоих видов колебаний. Возьмем сперва колебания крыла в отсутствии потока. Если рассматривать данное сечение крыла, то можно отметить две точки, определяющие механические свойства конструкции: центр тяжести сечения и центр жесткости сечения. Понятие центра тяжести общеизвестно; центром жесткости называется точка приложения равнодействующей упругих сил, возникающих при деформации крыла. Оба названных центра не совпадают. На рис. 121 показано их примерное расположение.

Представим себе, что крыло изгибается и начинает с ускорением движение вниз. В таком случае возникнут силы инерции, направленные вверх; равнодействующая сил инерции будет приложена в центре тяжести. Вследствие этого появится крутящий момент, равный произведению равнодействующей сил инерции на расстояние между центром тяжести и центром жесткости. Этот момент вызовет закручивание крыла. Фазовые соотношения всех входящих в игру величин можно пояснить при помощи рис. 122 и векторной диаграммы рис. 123. На рис. 122 показаны четыре последовательных положения колеблющегося крыла: *a* — крыло движется вниз. Это — среднее положение, в котором смещение и ускорение равны нулю, а скорость имеет наибольшее значение; *б* — крайнее нижнее положение. В этом положении скорость равна нулю, а смещение и ускорение максимальны. На сечение действует сила упругости F_s , направленная вверх и приложенная в центре жесткости, и сила инерции F_i , направлен-

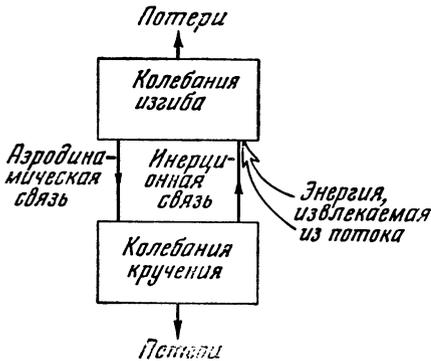
ная вниз и приложенная в центре тяжести. Силы F_s и F_i образуют пару, закручивающую крыло по часовой стрелке; v — среднее положение, крыло движется вверх; z — крайнее верхнее положение, направления сил и моментов изменились на обратные. На рис. 123 изображена векторная диаграмма для всех переменных величин. Здесь обозначено: x — смещение крыла (данного сечения), v — скорость, a — ускорение, F_s и F_i — соответственно силы упругости и инерции, M — крутящий момент, происхождение которого объяснено выше. Момент M вызывает поворот сечения на некоторый угол θ . Так как частота крутильных колебаний в действительных конструкциях выше частоты изгибных колебаний, то сопротивление крыла кручению имеет упругий (а не инерционный) характер. Поэтому (если пренебречь активным сопротивлением, обусловленным потерями при деформации) изменения угла θ совпадают по фазе с изменениями момента M , как и показано на диаграмме.

Посмотрим теперь, что получится, если колеблющееся описанным образом крыло поместить во встречный воздушный поток. При выбранных направлениях отсчета угол θ представляет собою положительное приращение угла атаки. Но, как говорилось выше, подъемная сила получает приращение ΔF , пропорциональное приращению угла атаки, т. е. находящееся с ним в фазе. Следовательно, переменная сила ΔF совпадает по фазе с ускорением; она не может поэтому совершать работу; она лишь влияет на период свободных колебаний крыла, уменьшая этот период. Итак, рассмотренная связь изгибных и крутильных колебаний крыла не содержит еще механизма автоколебаний.

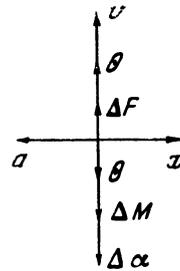
Мы установили, что изгибные колебания порождают крутильные. Аналогичным образом можно показать, что крутильные колебания вызывают изгибные. Таким образом, крыло представляет собою систему с двумя колебательными степенями свободы с инерционной связью между ними. Как всегда, в такого рода системе наблюдаются две собственные частоты, несколько отличающиеся от собственных частот для чисто изгибных и чисто крутильных колебаний. Хотя наличие связи между изгибными и крутильными колебаниями само по себе еще недостаточно для объяснения механизма флаттера, эта связь играет существенную роль в возникновении автоколебаний, как мы сейчас увидим.

Представим себе, что тем или иным способом мы возбудили изгибные колебания крыла, находящегося во встречном потоке. Результатом этих колебаний является периодическое изменение угла атаки, как было выяснено раньше. Но изменения угла атаки вызывают пропорциональные изменения аэродинамического момента. Переменный момент вызывает крутильные колебания. Последние же благодаря наличию инерционной связи усиливают колебания изгиба. Так образуется замкнутая автоколебательная система с отчетливой обратной связью и с преобразованием энергии потока в колебательную энергию.

Заметим, что здесь в отличие от предыдущего появляется новый вид связи между колебаниями изгиба и кручения, а именно, аэродинамическая связь. Ее отличие от ранее рассмотренной инерционной связи — отличие, чрезвычайно существенное с точки зрения автоколебаний, — состоит в том, что при инерционной связи закручивающий момент пропорционален у с к о р е н и ю изгибного движения крыла, тогда как при аэродинамической связи момент пропорционален с к о р о с т и этого движения.



◀ Р и с. 124



Р и с. 125

Энергетический обмен в описанном механизме поясняется схемой рис. 124.

Аналогичный механизм образуется, если исходить вначале из колебаний кручения. Колебания кручения вызывают периодическое изменение подъемной силы; изменения подъемной силы порождают колебания изгиба, а последние через инерционную связь могут усилить колебания кручения, и так далее.

Для того чтобы оба описанных механизма могли действовать, необходимо выполнение фазовых условий: необходимо, чтобы дополнительная сила, получаемая вследствие обратной связи, совпадала по фазе со скоростью данного типа колебательного движения.

Для первого механизма, т. е. для колебаний изгиба с обратной связью через крутильные колебания, фазовые соотношения иллюстрируются векторной диаграммой рис. 125. Основные векторы — смещение x , скорость v и ускорение a — те же, что и на рис. 123. Далее, $\Delta\alpha$ — приращение угла атаки, сдвинутое относительно скорости на 180° (так как при положительной скорости — вверх — приращение угла отрицательно, и наоборот). Добавочный аэродинамический момент ΔM совпадает по фазе с $\Delta\alpha$. Переменный угол закручивания θ совпадает по фазе с моментом ΔM — здесь сохранены те же предположения, что и при построении диаграммы рис. 123. Угловое ускорение $\dot{\theta}$ противоположно по фазе угловому смещению θ и совпадает по фазе со скоростью v . Наконец, переменная сила ΔF , обусловленная инерционной связью, совпадает по фазе с угловым ускорением, а следовательно, и со скоростью v колебаний изгиба. Мы убеждаемся, таким образом, что автоколебания возможны.

Энергия, извлекаемая колеблющимся крылом из встречного потока воздуха, быстро растет с увеличением скорости потока, т. е. скорости полета. Поэтому существует вполне определенная скорость полета, при которой поступление энергии сравнивается с потерями и возникают автоколебания. Эта скорость называется критической. У правильно рассчитанного самолета критическая скорость должна быть значительно выше максимальной скорости полета, включая и такие особые режимы полета, как пикирование, штопор и т. п. Выполнение этого условия гарантирует самолет от возникновения флаттера.

Критическая скорость в настоящее время надежно и сравнительно просто подсчитывается по конструктивным параметрам крыла и по данным продувок модели крыла в аэродинамической трубе. Однако неизменная тенденция к повышению максимальных скоростей самолетов заставляет рассмотреть общие пути к устранению флаттера.

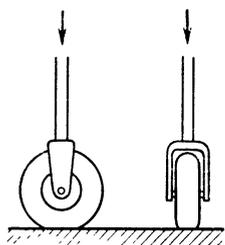
Если бы возможные механизмы автоколебаний исчерпывались сказанным выше, то радикальное устранение флаттера мы получили бы, если бы центры тяжести и жесткости совпадали в любом сечении крыла. Тогда не существовали бы инерционные связи между колебаниями кручения и изгиба, а, как было выяснено выше, чисто изгибные и чисто крутильные колебания могут быть только затухающими. Однако, во-первых, имеются еще аэродинамические подробности, в рассмотрение которых мы не входим, а во-вторых, указанное требование очень трудно осуществить конструктивно. Вместе с тем если не совмещение, то сближение обоих центров возможно, а тем самым уменьшается поступающая в крыло колебательная энергия и, следовательно, повышается критическая скорость.

Наряду с флаттером крыльев наблюдается еще флаттер, усложненный наличием элеронов (закрылков), флаттер хвостового оперения, флаттер винтов. Все эти явления сходны по физической природе с рассмотренными выше, все они вредны и опасны и подлежат безусловному устранению.

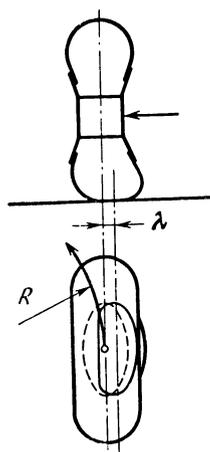
В заключение следует еще отметить, что явление флаттера встречается и в природе. Например, трепетание листьев на ветру имеет по существу ту же природу, что и флаттер. В особенности характерной моделью является осиновый лист (известно, что листья осины трепещут при самом слабом ветре) с его косым подвесом на длинном черенке, обладающем очень малой боковой жесткостью.

§ 23. Шимми

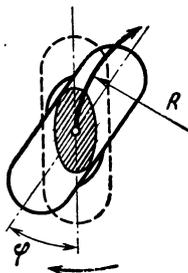
Этим профессиональным термином обозначается виляющее движение катящегося колеса. Шимми наблюдалось в передних колесах автомобилей ранних конструкций. В новейшее время проблема шимми возникла в связи с распространением самолетных шасси с носовым колесом. Этот вид шасси имеет серьезные преимущества с точки зрения условий посадки самолета на аэро-



Р и с. 126



Р и с. 127



Р и с. 128

дром [17]. Однако на первых порах внедрению этой рациональной конструкции препятствовало явление шимми, приводившее к ряду тяжелых аварий, пока не была разъяснена его природа и не были изысканы средства к его устранению. Нужно заранее признать, что явление слишком сложно, чтобы его можно было до конца объяснить качественно. Поэтому в данном параграфе будет дано только общее описание явления и характеристика факторов, от которых оно зависит. Взаимные же связи между этими факторами в том виде, как они установлены М. В. Келдышем [16], вынесены в добавление (см. стр. 369).

Переднее колесо самолета установлено на поворотной стойке (ноге) и принимает на себя определенную вертикальную нагрузку (рис. 126). В спокойном состоянии эта нагрузка сжимает пневматик (пневматическую шину), в результате чего он соприкасается с опорной плоскостью по некоторой площади, имеющей форму эллипса.

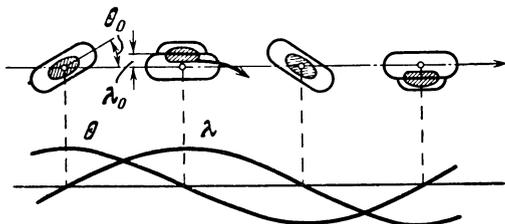
Явление шимми существенным образом зависит от упругих деформаций пневматика. Если отвлечься от влияния наклона колеса, то таких деформаций две: боковое смещение колеса λ и угол закручивания колеса относительно площади соприкосновения φ . Первый вид деформации показан на рис. 127, на котором штрихом отмечена эллиптическая площадь соприкосновения до деформации, сплошной линией — площадь соприкосновения, деформированная в результате бокового смещения.

Второй вид деформации поясняется рис. 128. На этом рисунке показана неизменная площадь соприкосновения (заштрихована) и измененное положение колеса при деформации вследствие закручивающего момента, действующего по часовой стрелке.

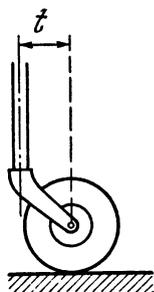
Основное кинематическое понятие, связанное с объяснением явления шимми, состоит в том, что обе вышеописанные деформации пневматика сообщают колесу

тенденцию двигаться по кривой, кривизна которой пропорциональна этим деформациям. Направления зависящих от деформаций траекторий колеса отмечены на рис. 127 и 128 кривыми стрелками; через R обозначен радиус кривизны траектории.

Этих представлений уже достаточно для описания одной из простейших разновидностей явления — так называемого кинематического шимми. Обозначим через θ угол между плоскостью обода колеса и направлением движения самолета, или,

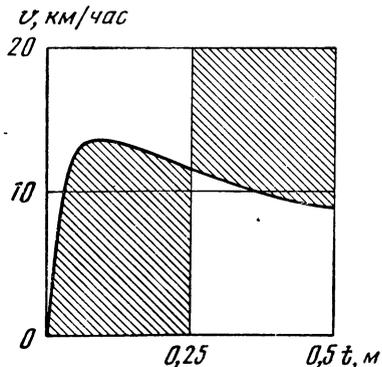


Р и с. 129

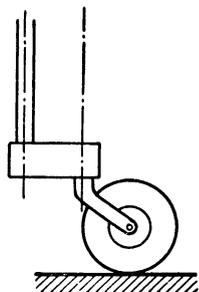


Р и с. 130

проще говоря, угол поворота стойки. Предположим, что в начальный момент колесо повернуто на угол θ_0 (рис. 129). При дальнейшем движении колесо сойдет с прямой, по которой движется ось стойки; начнет накапливаться деформация λ . Колесо станет поворачивать обратно, катясь по траектории переменной кривизны (эта траектория оказывается синусоидой). Наибольшая кривизна траектории получается в положении, когда боковое смещение λ наибольшее. При этом колесо располагается параллельно направлению движения. Затем под влиянием деформации λ оно возвращается на основную линию. Деформация при этом исчезает, но зато увеличивается угол θ , достигающий к концу полуцикла значения θ_0 . Таким образом, при описанных условиях колесо будет катиться по синусоиде, размахи которой не убывают. Любое случайное начальное возмущение вызывает непрекращающееся



Р и с. 131



Р и с. 132

в дальнейшем шимми. На рис. 129 кроме четырех последовательных положений колеса изображены также графики изменения величин λ и θ . Боковое смещение колеса изображено в сильно преувеличенном виде. При рассмотрении явления не принималась во внимание угловая деформация, так как на свободно вращающейся стойке (рис. 126) не могут развиваться какие-либо моменты, а динамику колеса мы пока не учитываем.

Радикальной мерой к устранению кинематического шимми является такое видоизменение конструкции стойки, при котором колесо вынесено относительно оси стойки, как показано на рис. 130. Вынос t считается положительным, если колесо сдвинуто по отношению к стойке назад. Наличие выноса сильно усложняет дело. Появляется момент относительно оси стойки, и, как следствие этого, возможны деформации закручивания. Они играют положительную роль: благодаря им колебания возможны только при самых малых выносах, да и то эти колебания оказываются затухающими. При больших же выносах движение вообще теряет колебательный характер; случайно отклонившееся колесо аperiodически возвращается на основную линию.

Основой интерес представляет, однако, не кинематическое, а динамическое шимми. При рассмотрении этого явления приходится учитывать кроме описанных факторов еще и момент инерции колеса относительно стойки. Колесо само по себе представляет колебательную систему, способную совершать крутильные колебания относительно стойки, причем параметрами, определяющими частоту этих колебаний, являются момент инерции колеса и момент упругости, обусловленный деформацией пневматика. С другой же стороны, механизм кинематического шимми остается в силе. Таким образом, явление состоит в автоколебаниях связанных систем, или системы с двумя степенями свободы: одна степень свободы динамическая, другая — кинематическая. Аналогичный случай нам встретился при рассмотрении флаттера; мы имели там связанные колебания двух видов: изгибные и крутильные. Но связи между обоими видами колебаний в случае флаттера проще и поддаются расчленению. Задача же о шимми сложнее; не пытаясь дать качественной картины самовозбуждения, отметим лишь, что применение стойки с выносом оказывается весьма радикальным средством борьбы с шимми и в общей постановке задачи, т. е. применительно к динамическому шимми.

На рис. 131 изображена диаграмма, заимствованная из работы Келдыша. На этой диаграмме по осям отложены скорость самолета v и вынос t . Площадь диаграммы разбита на области, показывающие, при каких условиях возможны автоколебания. Области устойчивости (т. е. области, в которых автоколебания невозможны) на диаграмме заштрихованы. Критическое значение t (на рисунке $t=0,25$), при котором смыкаются области устойчивости и неустойчивости, зависит от кинематических параметров деформации пневматика.

Первоначально основной мерой борьбы с шимми являлось изменение демпферов, гасящих вращательные колебания стойки. Конечно, демпферы — вполне универсальное средство предотвращения колебаний. Но желательно обойтись без них, так как они сильно усложняют и утяжеляют конструкцию. Рациональная конструкция стойки позволяет значительно облегчить демпферы или даже вовсе устранить их. В частности, очень эффективным видоизменением конструкции, данной на рис. 130, является стойка с двойным шарниром, устройство которой показано схематически на рис. 132. Колесо на такой стойке может совершать свободные колебания двоякого рода: колебания кручения, частота которых зависит от момента инерции колеса и момента упругости пневматика, и колебания бокового смещения, частота которых зависит от массы колеса и упругости пневматика. Анализ показывает, что если частота крутильных колебаний меньше, чем частота боковых колебаний, то шимми не возникает ни при каких скоростях самолета (без демпферов). Таким образом, эта конструкция при выполнении указанного условия представляет собою радикальное решение проблемы шимми.

§ 24. Параметрическое возбуждение

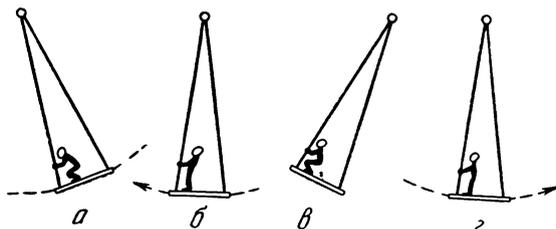
Энергия может быть вложена в систему не только прямым путем, который мы до сих пор рассматривали, т. е. не только путем периодического воздействия силы на колебательную систему.

Энергия может быть вложена в систему также путем периодического изменения какого-либо параметра системы при условии, что для этого изменения нужно совершать работу. А это условие выполняется только, если изменение параметра имеет надлежащую частоту и должным образом фазировано относительно движения системы.

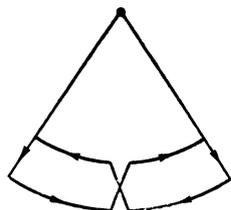
Общее представление о параметрическом возбуждении можно составить, рассматривая такой простейший пример, как раскачивание качелей. Для того чтобы раскачаться, человек, стоящий на доске качелей, приседает в крайних (верхних) положениях качелей (*a* и *e* на рис. 133) и резко распрямляется при проходе качелей через среднее (нижнее) положение (*b* и *г* на рис. 133). Центр тяжести системы при этом периодически поднимается и опускается. С учетом качания центр тяжести описывает траекторию, показанную на рис. 134.

Если бы человек приседал и распрямлялся на неподвижных качелях, то никакой работы он совершить не мог бы: работа, затраченная при поднятии центра тяжести, возвращалась бы при его опускании. Действительно, раскачать описанным способом неподвижные качели нельзя. Но если качели уже качаются, то при каждом цикле описанного выше сложного движения совершается работа, идущая на увеличение размаха качелей.

Дело в том, что, распрямляясь на д в и ж у щ и х с я качелях, человек развивает усилие, пропорциональное радиальному



Р и с. 133



Р и с. 134

(центростремительному) ускорению, и совершает соответствующую работу. Радиальное ускорение зависит от скорости. Оно имеет наибольшее значение при проходе качелей через среднее (нижнее) положение. Именно в этом положении человек распрямляется. В крайних же (верхних) положениях скорость, а с нею и радиальное ускорение равны нулю, и, приседая в этих положениях, человек никакой работы не производит¹. Работа же, зависящая от силы тяжести, за цикл равна нулю, как пояснено выше.

Рассматривая качели как математический маятник, мы можем трактовать происходящие при раскачивании качелей явления как результат периодического изменения д л и н ы маятника. Длина маятника — его основной п а р а м е т р. Таким образом, колебания возникают вследствие периодического изменения параметра колебательной системы, и это мы называем параметрическим возбуждением.

Другим примером параметрического возбуждения является так называемый опыт Мельде (1859). Опыт заключается в том, что натянутая струна прикрепляется одним концом к ножке колеблющегося камертона (рис. 135). В струне возбуждаются параметрические колебания. Изменяемым параметром является в данном случае н а т я ж е н и е с т р у н ы.

По поводу обоих приведенных примеров можно заметить, что параметрический способ возбуждения очень легко отличить от

¹ То же самое можно сказать и по-другому, а именно: человек, распрямляясь, уменьшает момент инерции качелей как системы, совершающей вращательное движение. Момент количества движения при этом сохраняется; значит возрастает угловая скорость. Энергия системы при этом также возрастает; ее прирост равен совершенной человеком при распрямлении работе. В крайних положениях скорость, а следовательно, и момент количества движения равны нулю; перемещение массы и связанное с этим изменение момента инерции не оказывают влияния на запас энергии в системе,

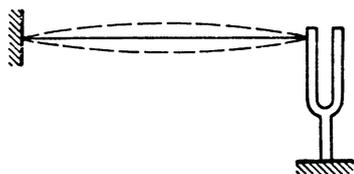
обычного. Оба примера являются механическими; в обоих наблюдаются механические колебания и действуют периодические механические силы. Но эти силы действуют не в направлении совершающихся колебаний, а в перпендикулярном направлении. Таким образом, сразу видно, что эти силы не могут непосредственно совершить работу над колебательной системой. Работа совершается и вкладывается в систему через посредство изменения параметра.

Всматриваясь внимательно в действие обеих описанных систем, можно подметить еще одно обстоятельство, очень характерное для параметрического возбуждения, а именно: частота изменения параметра вдвое больше частоты возбуждаемых колебаний. Действительно, человек на качелях совершает два полных цикла движения (два подъема и два приседания) за одно качание. В опыте с натянутой струной струна возбуждается, если ее собственная частота вдвое ниже частоты камертона.

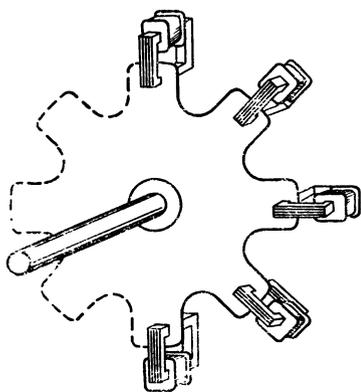
Аналогично обстоит дело и в других параметрических системах. Нужно лишь заметить, что отношение частоты колебаний к частоте изменения параметра может равняться не только $1/2$, но и 1 , $3/2$, 2 , $5/2$ и т. д. Однако легче всего параметрические колебания возбуждаются именно при отношении названных частот, равном $1/2$.

Можно возбудить параметрические колебания в электрическом контуре, если периодически изменять один из его параметров, т. е. емкость или индуктивность. Представим себе, что мы изменяем емкость колебательного контура при наличии колебаний следующим образом: раздвигаем обкладки конденсатора в моменты, когда напряжение на конденсаторе имеет наибольшее значение (что происходит дважды за период), и снова сдвигаем обкладки, когда напряжение на конденсаторе равно нулю. Очевидно, что, раздвигая обкладки, мы совершаем работу, которая идет на повышение напряжения (так как емкость уменьшается, а заряд остается неизменным). При сближении же обкладок работа не затрачивается и не возвращается, так как напряжение равно нулю. Аналогично обстоит дело и при периодическом изменении индуктивности, которую нужно уменьшать в моменты, когда ток имеет наибольшее значение, и увеличивать, когда ток переходит через нуль.

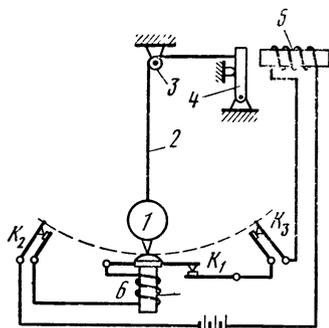
Академики Мандельштам и Папалекси, разработавшие теорию параметрического возбуждения [18], построили электрический параметрический генератор следующего устройства [19]. Статор генератора представляет собою группу катушек с разомкнутыми стальными сердечниками. Ротор представляет собою зубчатый металлический диск из немагнитного и хорошо проводящего металла (алюминий, медь). Число зубцов ротора равно числу катушек статора. Таким образом, при вращении ротора его зубцы то входят в зазоры сердечников, то оказываются в промежутках между ними (рис. 136). При введении проводящего металла в магнитное поле катушек индуктивность их уменьшается. Следова-



Р и с. 135



Р и с. 136



Р и с. 137

тельно, при вращении ротора индуктивность катушек будет периодически изменяться. Частота изменения индуктивности (частота модуляции) пропорциональна числу зубцов и числу оборотов ротора. Ротор приводится во вращение мотором с повышающей зубчатой передачей (в построенной модели число оборотов ротора доходило до 15 000 в минуту). Если теперь подключить к катушкам конденсатор и образовать, таким образом, колебательный контур, то, когда частота модуляции будет вдвое больше собственной частоты контура, в нем возникнут быстро нарастающие по амплитуде колебания. Если рост амплитуды ничем не ограничен, то переменное напряжение достигает таких больших значений, что пробивается изоляция конденсаторов. Во избежание этого при опытах генератор нагружался лампами накаливания. Лабораторная модель развивала колебательную мощность в несколько киловатт при частоте около 1000 гц. Для того чтобы колебания возбудились, нужно соблюдать определенное соотношение между относительным изменением параметра (глубиной модуляции) и затуханием контура. Чем затухание больше, тем больше и необходимая глубина модуляции. В этом требовании заключено по существу условие энергетического баланса.

Теперь заметим, что генератор Мандельштама и Папалекси является генератором с независимым параметрическим возбуждением. Это значит, что изменения параметра задаются извне, и частота этих изменений может быть какой угодно.

В опыте Мельде мы также имеем дело с независимым параметрическим возбуждением. В обоих случаях пужно подогнать частоты модуляции и колебательной системы к наивыгоднейшему соотношению, равному половине. В генераторе Мандельштама и Папалекси это достигается изменением числа оборотов ротора (при постоянной настройке контура), в опыте Мельде — изменением натяжения струны (при неизменной частоте камертона).

Можно, однако, представить себе и такую систему, в которой колебательная система сама управляет изменениями параметра через посредство соответствующим образом организованной обратной связи. Такая система будет уже обладать всеми чертами автоколебательной системы. Такого рода систему мы назовем автоколебательной системой с параметрическим возбуждением.

Качели являются именно такой системой. Однако канал обратной связи пролегает в этом случае через реакции человека. Чтобы не запутывать вопрос, мы попытаемся построить для примера схему по возможности простой автоколебательной системы с параметрическим возбуждением на основе математического маятника переменной длины.

Возможное устройство такого рода системы показано на рис. 137. Маятник 1 качается на нити 2 , перекинутой через блок 3 . Конец нити прикреплен к стальному рычагу 4 , могущему занимать два положения; в левом положении он удерживается ограничителем, в правое же положение переходит под действием электромагнита 5 . Имеется система контактов, управляемая движением маятника: контакт K_1 , блокируемый электромагнитом 6 , и два контакта K_2 и K_3 , находящихся в состоянии покоя в замкнутом положении. Все три контакта и оба электромагнита включены последовательно в цепь, питаемую батареей.

Действие системы таково. Маятник, проходя через среднее положение, прижимает контакт K_1 и замыкает цепь (этот момент показан на рис. 137). Контакт остается замкнутым и при дальнейшем движении маятника влево, будучи заблокирован электромагнитом 6 . При замыкании тока электромагнит 5 притягивает рычаг 4 , который, протягивая нить через блок, укорачивает маятник. Когда маятник доходит до контакта K_2 , он, задевая, размыкает его на короткое время. Блокирующий электромагнит отпускает контакт K_1 , а рабочий электромагнит отпускает рычаг, в результате чего маятник удлиняется. Цепь остается разомкнутой, пока маятник не вернется в среднее положение, где он снова замкнет контакт K_1 . На протяжении второго полупериода движения маятника все явления происходят в том же порядке, но с участием контакта K_3 вместо K_2 .

Итак, в описанной системе происходят изменения параметра — длины маятника — в точности таким же образом, как в качелях, так что траектория центра тяжести маятника, представленная на рис. 134, полностью относится и к данной системе.

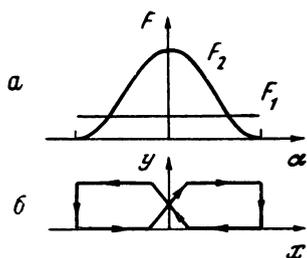


Рис. 138

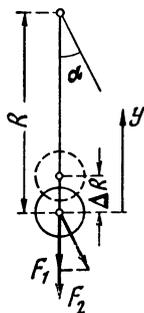


Рис. 139

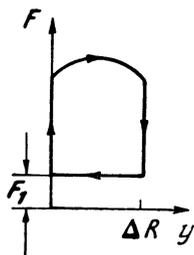


Рис. 140

Как видим, изменения параметра управляются через посредство обратной связи со стороны колебательной системы. Итак, описанная система может служить примером автоколебательной системы с параметрическим возбуждением.

Займемся теперь вопросом об энергетическом балансе параметрических колебаний. Рассмотрим этот вопрос на примере качелей или, что то же, параметрического маятника рис. 137.

Как уже отмечалось, источник энергии в системах с параметрическим возбуждением совершает работу при изменении параметра. Эту работу мы и должны определить, сравнив ее с энергией, теряемой при колебаниях.

Построим прежде всего график внешних сил, действующих вдоль маятника (других сил нет!), в зависимости от положения маятника. Угловое смещение маятника обозначим через α и будем отсчитывать его от среднего положения. На рис. 138, *a* изображены графики сил; F_1 — сила тяжести, точнее ее проекция на направление маятника (как показано на рис. 139); при малых амплитудах можно считать эту силу постоянной, т. е. не зависящей от положения маятника. F_2 — центробежная сила, зависящая от угловой скорости, а следовательно, и от положения маятника, как показано. На рис. 138, *b* дана зависимость смещения y центра тяжести вдоль маятника (считая от нижнего положения) от положения маятника x . Рис. 138, *b* представляет траекторию центра тяжести, т. е. повторяет в прямоугольных координатах рис. 134.

Внешняя сила F , действующая на подвес маятника, уравнивает сумму сил F_1 и F_2 , т. е. равна этой сумме по величине и противоположна по направлению. Сила F может совершить работу лишь на перемещениях по своему направлению, т. е. вдоль маят-

ника. Поэтому мы должны строить диаграмму работы в виде зависимости F от y . Беря соответствующие друг другу значения обеих этих величин с графиков рис. 138, получим диаграмму, изображенную на рис. 140. Для оценки вкладываемой в систему энергии нужно учитывать, что контур диаграммы рис. 140 обходится за один период колебания маятника дважды.

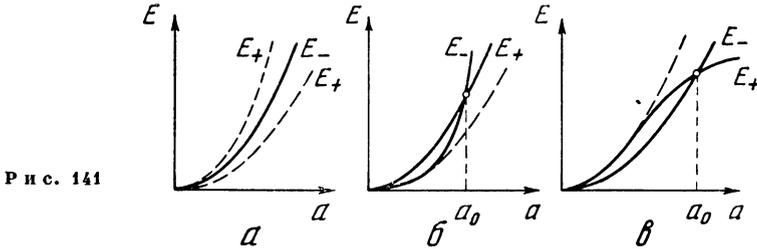


Рис. 141

Заметим, что площадь диаграммы зависит от величины ΔR , т. е. от смещения центра тяжести. Отношение ΔR к длине маятника R есть не что иное, как глубина модуляции параметра. Таким образом, вкладываемая в систему работа пропорциональна примерно первой степени глубины модуляции. Площадь диаграммы зависит также и от величины центробежной силы. Но центробежная сила пропорциональна квадрату амплитуды колебания маятника. Значит при постоянной глубине модуляции вкладываемая в систему энергия E_+ возрастает с амплитудой по параболическому закону.

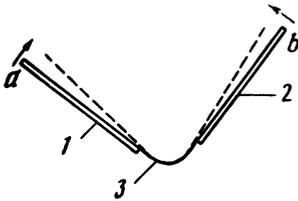


Рис. 142

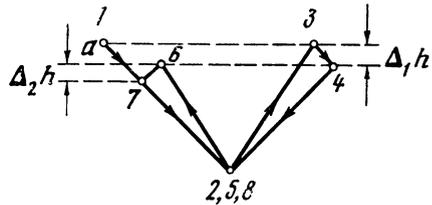


Рис. 143

С другой стороны, если нагрузка системы линейна, т. е. если, например, затухание маятника обусловлено вязким трением, то теряемая энергия E_- возрастает также пропорционально квадрату амплитуды. Стало быть, E_+ и E_- представляются двумя параболлами, которые нигде не пересекаются, кроме как в начале координат (рис. 141, а). Если парабола E_- лежит выше параболлы E_+ , то параметрические колебания не могут возникнуть. Если же парабола E_+ расположится выше E_- — этого можно достигнуть увеличением глубины модуляции — то колебания возбуждятся, но будут неограниченно нарастать. Для того чтобы получить пара-

метрические колебания с устойчивой амплитудой, необходимо что-нибудь из двух: либо чтобы нагрузка была нелинейна, т. е. чтобы E_- возрастала быстрее, чем по параболе (рис. 141, б), либо чтобы глубина модуляции не оставалась постоянной, а убывала бы с амплитудой; тогда E_+ будет расти медленнее, чем по параболе (рис. 141, в). Обе эти возможности могут быть использованы в реальном устройстве.

Так, например, человек, раскачивающийся на качелях, делает вначале очень энергичные движения, низко приседая. Когда же качели достигнут желаемого размаха, движения человека делаются еле заметными. Таким образом, в данном случае ограничение амплитуды достигается путем уменьшения глубины модуляции. Другой возможностью человек на качелях и не располагает.

§ 25. Одна механическая модель [20]

Рассмотрим одну механическую модель, относящуюся к параметрическим системам, но обладающую рядом своеобразных черт.

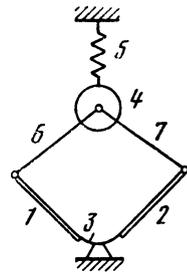
В этой модели энергия периодическими порциями отбирается от падающего с некоторой высоты тяжелого шарика. Модель представляет собою желоб, состоящий из трех звеньев (рис. 142). Прямые звенья 1 и 2 жестки, а звено 3 гибко. За счет изгиба звена 3 наклон звеньев 1 и 2 может изменяться, как показано на рис. 142 штрихом.

Предположим, что желоб неподвижен и что мы пускаем по нему в точке a тяжелый шарик. Скатившись по звену 1, шарик приобретает к концу спуска некоторую скорость; его потенциальная энергия перейдет в кинетическую. Поднимаясь по звену 2, он будет замедлять свое движение, и к концу подъема кинетическая энергия снова перейдет в потенциальную. Если бы потери на трение отсутствовали, то шарик поднялся бы на такую же высоту, с которой он первоначально был спущен. Таким образом, шарик совершал бы в желобе периодические колебания с неубывающей амплитудой.

Пусть теперь наклон звеньев 1 и 2 периодически изменяется следующим порядком: когда шарик находится в наинизшем положении, т. е. проходит звено 3, наклон звеньев 1 и 2 увеличивается, когда же шарик находится в наивысших точках, наклон звеньев 1 и 2 уменьшается. Если эти изменения наклона происходят скачкообразно, то легко сообразить, что шарик будет двигаться по траектории, изображенной на рис. 143. Путь шарика начинается в точке 1. Он сбегает в точку 2 по малому наклону. В этот момент наклон увеличивается и шарик избегает по более крутому склону до точки 3, находящейся на той же высоте, что и точка 1. В этот момент желоб опускается и шарик вместе с желобом переходит в точку 4, отдавая при этом часть своей потенциальной энергии, равную произведению веса шарика на отрезок $\Delta_1 h$ (рис. 143). После этого шарик пускается в обратный путь и,

пройдя наинизшее положение *5*, достигает точки *6* на той же высоте, что и *4*. Здесь желоб снова опускается и шарик снова отдает часть своей энергии, определяемую отрезком $\Delta_2 h$. Так происходит и дальнейшее движение шарика. Израсходовав весь свой первоначальный запас энергии, шарик вываливается в отверстие, сделанное в середине (в наинизшей точке) звена *3*.

Как видим, все это напоминает рассуждения о качелях. Переменным параметром является здесь наклон желоба. Период изменения этого параметра вдвое меньше времени, затрачиваемого шариком на полный цикл своего движения. Разница состоит в том, что мы не раскачиваем шарик (что также вполне возможно), а, наоборот, отнимаем от него его энергию периодическими порциями. Этими периодическими порциями можно раскачать некоторую колебательную систему. Такую систему можно выполнить, например, в виде, показанном на рис. 144. Здесь масса *4* подвешена на пружине *5* и соединена шарнирными тягами *6* и *7* со звеньями *1* и *2* желоба. Средняя часть гибкого звена *3* закреплена неподвижно. Очевидно, что собственная частота системы должна быть подобрана так, чтобы она была вдвое больше частоты колебания шарика.



Р и с. 144

Здесь, однако, нужно заметить, что частота колебания шарика не есть постоянная величина. По мере того как высота подъема шарика убывает, продолжительность цикла его движения сокращается. Поэтому оптимальные соотношения не могут быть выдержаны в точности, и это обстоятельство несколько усложняет дело, хотя и не исключает возможности возбуждения колебаний.

Поддерживать незатухающие колебания можно, вводя периодически в систему новые шарики. Потерявшие энергию шарики будут, как выше говорилось, вываливаться через нижнее отверстие желоба. Взамен их в игру будут вступать все новые и новые шарики со свежими запасами энергии. При таких условиях мы получим незатухающие колебания.

В предыдущем предполагалось, что шарик пускается на желоб во вполне определенной — «благоприятной» — фазе, а именно в тот момент, когда желоб опускается. Когда же шарик проходит наинизшее положение, желоб должен подниматься. При этих условиях шарик отдает энергию, как мы видели раньше. Но представим себе, что шарик попадает на желоб в другой фазе, а именно когда желоб поднимается. Тогда при проходе шарика через наинизшее положение желоб опустится, а в наивысшем положении шарика снова поднимется и так далее. При таких фазных соотношениях мы не только не получим от шарика энергию, но, наоборот, должны будем совершать работу подъема шарика вместе с желобом.

Следовательно, если бы шарики сыпались на желоб беспорядочно и попадали бы на него в произвольной фазе, то они с одина-

ковой вероятностью могли бы оказаться в благоприятной и в неблагоприятной фазах. В среднем энергия от них не отбиралась бы и колебания не могли бы поддерживаться.

Но тут выясняется одно замечательное свойство рассматриваемой модели: она самостоятельно производит рассортировку шариков по фазе. В самом деле, шарик неблагоприятной фазы, скатывающийся по большому наклону, т. е. при поднятом желобе, и вкатывающийся вверх по меньшему наклону, т. е. при опущенном желобе, сразу вылетает с желоба и выходит из игры. На желобе остаются только шарики благоприятной фазы. Система обладает, таким образом, фазовой избирательностью.

Из сказанного следует, что поддерживать незатухающие колебания можно посредством беспорядочно сыплющихся шариков; система сама рассортирует их по фазам, выбросит шарики неблагоприятной фазы и использует энергию шариков благоприятной фазы.

Устройство, представленное на рис. 144, является параметрической, но не автоколебательной системой. Соотношение частот должно быть заранее подобрано. Система была бы автоколебательной, если бы посредством какой-либо обратной связи движение шарика само управляло наклоном или, наоборот, колебательное движение желоба управляло движением шарика.

§ 26. Электронные генераторы сверхвысоких частот [21]

Описанная в предыдущем параграфе модель часто приводится в качестве механической аналогии так называемого генератора с тормозящим полем, или генератора с положительной сеткой, при помощи которого можно получить электрические колебания весьма высокой частоты (до диапазона дециметровых волн).

Аналогия действительно существует, как мы сейчас убедимся. Польза подобных аналогий часто оспаривается, но мы оставим эту методическую дискуссию в стороне, так как и электронное устройство и механическая модель представляют для нас самостоятельный интерес в качестве примеров параметрических систем.

Было обнаружено, что в обычном триоде могут возбуждаться при определенном режиме электрические колебания весьма высокой частоты. Период этих колебаний (так называемые колебания Баркгаузена-Курца) связан с временем пролета электронов в пространстве между электродами лампы.

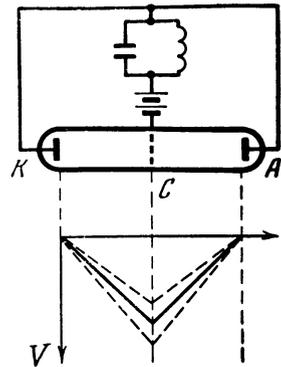
Схема включения триода показана на рис. 145. Анод и катод соединены вместе. Сетка имеет относительно катода высокий положительный потенциал. Кроме того, имеется еще и переменная разность потенциалов, развиваемая на контуре в цепи сетки. Диаграмма распределения потенциала в пространстве между электродами показана в нижней части рис. 145, где сплошной линией обозначено распределение потенциала, создаваемое батареей,

штриховыми линиями — распределение потенциала при наложении переменной составляющей напряжения. Следует заметить, что как в рассматриваемом случае, так и ниже наименования электродов имеют условный смысл: например, на рис. 145 анодом, т. е. электродом, имеющим положительный потенциал относительно катода, является в сущности сетка.

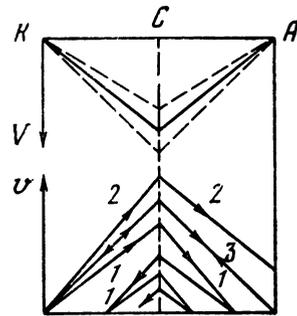
Предположим для начала, что переменная разность потенциалов отсутствует. Тогда вышедший из катода электрон попадает в ускоряющее поле в пространстве катод — сетка (K и C на рис. 145). Пройдя сетку, электрон попадает в поле обратного направления в пространстве сетка — анод (C и A на рис. 145). Это поле оказывает на электрон тормозящее действие; электрон теряет скорость и возвращается с ускорением к сетке. Не долетев до катода, он снова поворачивает назад и так далее. Таким образом, электрон может совершать колебательное движение между катодом и анодом. Если электрон при этом отдает свою энергию путем наведения тока во внешней цепи триода, то размах его колебаний уменьшается, пока он не попадет на сетку. На этом его индивидуальная история заканчивается; он выходит из игры.

Посмотрим теперь, что будет происходить, если на сетке имеется переменное напряжение, период которого равен времени пролета электрона от катода до анода. Рассмотрим при этом два различных случая.

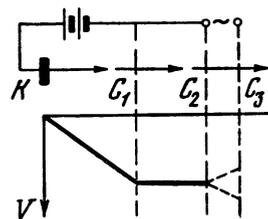
С л у ч а й 1. Электрон выходит из катода в такой момент, что на протяжении пути электрона к сетке поле, создаваемое переменной разностью потенциалов, тормозит электрон или, иначе говоря, на время пробега электроном пути катод — сетка приходится отрицательный полупериод напряжения на сетке (минус на сетке относи-



Р и с. 145



Р и с. 146



Р и с. 147

тельно катода). После же пролета электроном сетки наступает положительный полупериод переменного напряжения на сетке, в результате чего возникает дополнительное поле, опять-таки тормозящее электрон. Двигаясь в тормозящем поле, электрон совершает работу за счет собственного запаса энергии. Эта работа идет на создание наведенного тока во внешней цепи. Размах колебаний электрона с каждым пролетом убывает, и в конце концов он попадает на сетку. Описанные явления представляют собой историю электрона, вышедшего из катода в благоприятной фазе.

Случай 2. Электрон выходит из катода в такой момент, когда на протяжении его пути между катодом и сеткой действует положительный полупериод переменного напряжения на сетке. Когда же электрон попадает в пространство сетка—анод, наступает отрицательный полупериод переменного напряжения на сетке, действие которого уменьшает тормозящий эффект основного поля, т. е. опять-таки ускоряет электрон. В результате электрон, сохраняя некоторую скорость, попадает на анод и, таким образом, сразу выходит из игры. Его кинетическая энергия выделяется в виде тепла. Описанные явления относятся к электрону, вышедшему из катода в неблагоприятной фазе. Фаза эта неблагоприятна в том смысле, что энергия не только не отбирается от электрона, но, напротив, затрачивается на его ускорение.

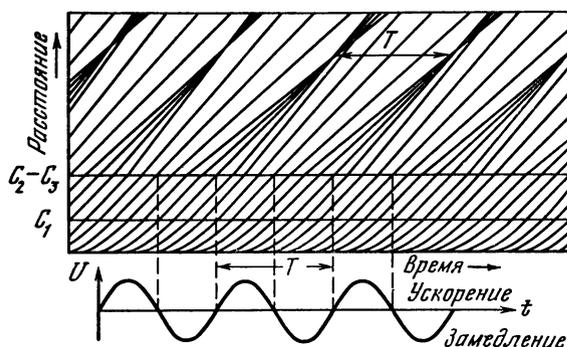
Как видим, в схеме происходит автоматический отбор электронов благоприятной фазы и возможно возбуждение колебаний за счет энергии электронов, выходящих из катода непрерывно в совершенно произвольные моменты времени.

Для наглядности различие в поведении электронов представлено на рис. 146 графиками зависимости скорости от положения электрона. В верхней части рисунка представлено распределение потенциала (то же, что и на рис. 145). В нижней части линия 1 изображает движение электрона благоприятной фазы, линия 2 — неблагоприятной, линия 3 относится к случаю, когда на сетке нет переменного напряжения.

Все описанные явления аналогичны механическим явлениям в модели, рассмотренной в предыдущем параграфе. Аналогия основана на том, что движение тяжелого шарика по наклонной плоскости совершается по тому же закону, что и движение электрона в электрическом поле. Наклон соответствует напряженности электрического поля. Но наклон есть градиент высоты, а напряженность поля — градиент потенциала. Отсюда сразу следует простой рецепт построения механических моделей, изображающих движение электронов: нужно, чтобы профиль лотка, по которому скатываются шарики, воспроизводил диаграмму распределения потенциала в пространстве, где движутся электроны. Сравнивая рис. 145 и рис. 142, мы видим, что это правило в данном случае соблюдено. Мы еще воспользуемся этими соображениями.

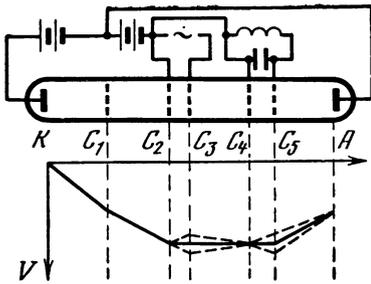
Обратимся теперь к сравнительно новому типу генератора сверхвысоких частот — к так называемому клистрону. Его действие основано на ином принципе, а именно на модуляции электронов по скорости, в результате чего получается электронный поток, модулированный по плотности. Рассмотрим подробно этот интересный механизм.

Пусть поток электронов, вышедших из катода K , ускоряется сеткой C_1 и проходит через две параллельные сетки C_2 и C_3 , между которыми приложена переменная разность потенциалов (рис. 147). Все электроны входят в пространство C_2-C_3 с одинаковой скоростью. Но в этом пространстве они ускоряются либо замедляются

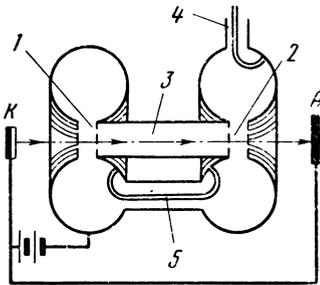


Р и с. 148

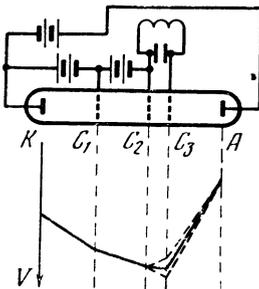
в зависимости от мгновенного значения поля между C_2 и C_3 . Таким образом, электроны выходят через C_3 уже с различными скоростями, причем чередование более быстрых и более медленных электронов происходит периодически в соответствии с изменениями напряжения между C_2 и C_3 . При дальнейшем движении электронов более быстрые догоняют более медленные. В результате этого в электронном потоке образуются густки, т. е. области с большей плотностью электронов. Процесс образования густков называют группированием, комбинацию сеток C_2 и C_3 — группирователем, а пространство, в котором сформировываются густки, называют пространством группирования или пространством дрейфа. Процесс группирования поясняется рис. 148. На этом графике представлены пути электронов, выходящих из катода через равные промежутки времени Δt . Ускоренные сеткой C_1 электроны подходят к группирователю C_2-C_3 также через равные интервалы. Здесь электроны получают различное ускорение в зависимости от поля в группирователе, изменения которого представлены синусоидой в нижней части рисунка. После пробега пространства группирования образуются густки электронов, ясно видимые в верхней части рисунка. Полученный таким образом электронный поток с периодически меняющейся плотностью можно использовать для наведения тока во внешней



Р и с. 149



Р и с. 150

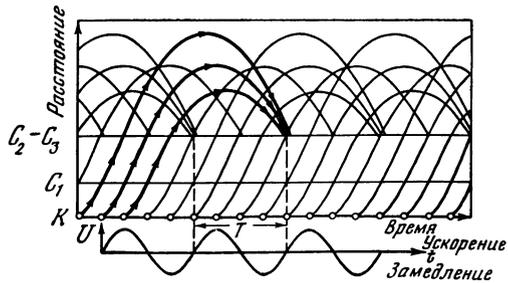


Р и с. 151

цепи, и мы приходим к схеме клистрона, показанной на рис. 149. К электродам, уже упомянутым выше, добавляются две сетки G_4 и G_5 , образующие так называемый улавливатель. Улавливатель располагается как раз там, где формируются электронные сгустки. Частота прохождения сгустков через улавливатель равна, как легко видеть из рис. 148, частоте напряжения на группирователе. К сеткам улавливателя присоединен контур, в котором наводится колебательный ток. Электронный поток, прошедший улавливатель, отводится «анодом» A . Потенциальная диаграмма изображена в нижней части рис. 149.

Конструктивное осуществление принципиальной схемы рис. 149 показано на рис. 150. Группирователь и улавливатель выполнены каждый в форме пары параллельных перфорированных дисков, входящих в виде сосредоточенной емкости в состав объемных резонаторов — так называемых румбаторнов. Эти резонаторы имеют тороидальную полость (на рис. 150 они показаны в разрезе).

Пространство дрейфа 3 находится в трубе, соединяющей оба



Р и с. 152

румбатрона. Коаксиальная линия 4 служит для вывода высокочастотных колебаний из клистрона. Оба румбатрона соединены петлей обратной связи 5. Действие этой обратной связи состоит в том, что колебания, наведенные в улавливателе модулированным по плотности электронным потоком, управляют самим процессом модуляции. Таким образом, изображенное на рис. 150 устройство представляет собою автоколебательную систему с параметрическим возбуждением. Параметром, как и в лампах с тормозящим полем, является напряженность электрического поля.

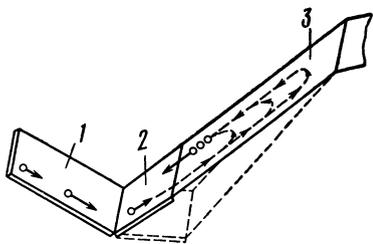


Рис. 153

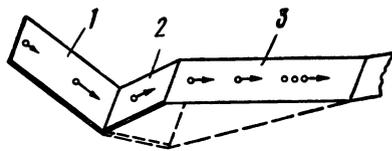


Рис. 154

Особое устройство имеет так называемый отражательный клистрон, предложенный в 1940 г. В. Ф. Коваленко [22]. В этом электронном приборе одна и та же пара сеток служит и в качестве группирователя, и в качестве улавливателя. Схема отражательного клистрона показана на рис. 151. Действие его состоит в том, что электрон, прошедший группирователь C_2-C_3 и получивший в нем то или иное дополнительное ускорение, попадает в пространство C_3-A , где действует поле обратного направления, так как анод A находится под отрицательным потенциалом по отношению к катоду и тем более по отношению к группирователю (см. потенциальную диаграмму на рис. 145). Соотношения подобраны так, что электроны, в том числе и наиболее быстрые, поворачивают обратно, не долетев до электрода A . Пространство C_3-A является пространством группирования. Самый процесс группирования происходит примерно так же, как и в прелетном клистроне рис. 149, несмотря на возвратное движение электронов. Процесс группирования в отражательном клистроне поясняется рис. 152. Электронные сгустки проходят через C_2-C_3 , которые на этот раз играют роль улавливателя. В системе могут возбуждаться колебания; эта система также есть автоколебательная система с параметрическим возбуждением. Однако в отличие от предыдущего случая обратная связь здесь не внешняя, а внутренняя: она осуществляется через электронный поток.

Для пояснения явлений в клистронах также часто прибегают к механической модели с тяжелыми шариками [23, стр. 140]. Процесс группирования можно иллюстрировать моделью, изобра-

женной на рис. 153. Шарики скатываются по лотку 1 и попадают на качающуюся площадку 2, периодически изменяющую свой наклон. По выходе с различной скоростью с площадки 2 шарики попадают на резиновую ленту 3, на которой происходит процесс группирования. Если теперь дать ленте 3 крутой наклон вверх, как показано на рис. 154, то мы получим модель отражательного клистрона. Шарики будут заворачивать обратно, и сгустки будут формироваться как раз перед попаданием шариков обратно на площадку 2. Можно подвесить площадку 2 на пружинах; образуется колебательная система, которая при соответствующей настройке будет раскачиваться периодически сбегаящими на нее шариками.

Модели, изображенные на рис. 153 и 154, построены в соответствии с потенциальными диаграммами клистронов (рис. 149 и 151).

§ 27. Заключение

В заключение повторим в виде кратких определений и тезисов основные установленные нами положения.

Автоколебаниями называют колебания, не вызываемые внешней периодической силой, — в этом их отличие от вынужденных колебаний.

Автоколебательная система — устройство, способное генерировать колебания на основе преобразования постоянного движения в колебательное и характеризующееся наличием источника энергии, колебательной системы, клапана, пропускающего энергию от источника к колебательной системе периодически порциями, и обратной связи со стороны колебательной системы, управляющей работой клапана.

Устойчивый режим автоколебаний определяется энергетическим балансом, т. е. равенством энергии, подводимой от источника к колебательной системе, и энергии, теряемой системой. При этом можно различать два случая: если автоколебания происходят с заданной частотой, то условие энергетического баланса определяет установившуюся амплитуду автоколебаний; если же автоколебания происходят с заданной амплитудой, то условие баланса определяет частоту автоколебаний.

В линейной системе устойчивые автоколебания вообще невозможны.

Различают мягкое и жесткое самовозбуждение. В первом случае система самостоятельно раскачивается из положения покоя. Во втором случае для возбуждения системы требуется конечной величины начальный толчок.

Мягкое и жесткое самовозбуждения связаны со свойством устойчивости; мягким самовозбуждением обладает система, динамически неустойчивая в положении покоя, жестким самовозбуждением — система, устойчивая в положении покоя.

Для того чтобы автоколебания были возможны, необходимо, чтобы в колебательную систему периодически вкладывалась энергия. Это означает, что переменная сила, возникающая в результате действия обратной связи и клапана, должна совершать за каждый цикл движения определенную работу. Но это невозможно, если сила есть однозначная функция положения системы. Для того чтобы работа могла быть совершена, необходима по меньшей мере двузначная зависимость силы от смещения.

Требуемая двузначность может быть двоякого рода — статическая и динамическая. Статической называется двузначность, которая обнаруживается в статических характеристиках зависимости силы от смещения, т. е. при сколь угодно медленных изменениях этих величин. Динамическая двузначность, наоборот, проявляется лишь в колебательном режиме и зависит от наличия в уравнениях системы членов с производными тех или иных порядков по времени.

Как статическая, так и динамическая двузначность могут трактоваться в колебательном режиме с точки зрения фазного сдвига между основными гармониками силы и смещения или силы и скорости. Работа совершается, если внешняя сила совпадает по фазе со скоростью или, что то же, если она опережает смещение по фазе на 90° .

При синусоидальном движении высшие гармоники силы работы не совершают и могут не приниматься во внимание.

Для исследования энергетических соотношений полезно построение диаграмм работы, представляющих собою графики в координатах сила — перемещение (или напряжение — заряд). При установившемся периодическом движении диаграмма работы представляет собою замкнутый контур. Площадь, охватываемая этим контуром, дает величину совершенной за цикл работы. Работа положительна или отрицательна (т. е. энергия вкладывается в систему или отбирается от нее) в зависимости от направления обхода диаграммы изображающей точкой.

В ряде случаев полезным оказывается понятие отрицательного сопротивления. Это понятие связано с фазовыми соотношениями. В обыкновенном — положительном — сопротивлении напряжение и ток или сила и скорость совпадают по фазе. Такое сопротивление является потребителем энергии. Если же на каком-либо элементе системы обнаруживается, что напряжение и ток противоположны по фазе, то такой элемент может являться источником колебательной энергии и обозначается как отрицательное сопротивление.

Если какой-либо элемент обладает падающей характеристикой зависимости напряжения от тока (или силы от скорости), то это значит, что его сопротивление, определяемое как отношение приращения напряжения к приращению тока, отрица-

тельно. Падающая характеристика и отрицательное сопротивление — понятия, выражающие одно и то же свойство системы.

Понятие отрицательного сопротивления имеет смысл только применительно к колебательным режимам. Нелинейный элемент может обладать отрицательным сопротивлением переменному току, но обязательно представляет положительное сопротивление для постоянного тока.

Падающая характеристика зависимости любых двух колебательных величин означает, что эти величины в противофазе (если их изменения лежат в пределах падающего участка).

При исследовании сложных автоколебательных систем и для выяснения роли тех или иных звеньев этих систем полезен анализ фаз. Под этим понимается рассмотрение фазовых соотношений в замкнутой цепи из колебательной системы и обратной связи, основанное на том, что результирующий фазный сдвиг после обхода любой замкнутой цепи равен нулю. Такого рода анализ удобно иллюстрировать построением векторных диаграмм, аналогичных диаграммам теории переменных токов.

Автоколебательные системы делятся на почти гармонические и релаксационные. Для первых характерна почти синусоидальная форма автоколебаний, для вторых — резко несинусоидальная и иногда разрывная. Имеется существенное различие в колебательных системах: в почти гармонических системах имеется колебательный комплекс из двух элементов (например, индуктивность и емкость), обменивающихся энергией; в релаксационной системе имеется один накопитель энергии (например, емкость), то запаасающий энергию, то отдающий ее. В соответствии с этим в релаксационной системе действие клапана характеризуется перепадом между двумя значениями энергии накопителя, при одном из которых клапан открыт, а при другом — закрыт.

Для релаксационных систем этого типа характерно также, что они работают в режиме заданной амплитуды, не зависящей от нагрузки, но определяемой величиной вышеупомянутого перепада.

При наличии на характеристике клапана кроме перепада еще и падающего участка возможен плавный переход релаксационных колебаний в почти гармонические при соответствующем постепенном изменении колебательной системы или режима.

Колебательные или задающие системы автоколебательных систем могут быть следующих родов: один накопитель (релаксационная система, вырожденный случай), колебательная система с одной степенью свободы, система с двумя или более колебательными степенями свободы (связанные системы), системы с распределенными постоянными.

Назовем прямым возбуждением такой механизм поддержания колебаний, когда периодическая сила действует непосредственно на колебательную систему и совершает над ней

работу. В отличие от этого возможно параметрическое возбуждение, состоящее в том, что сила периодически изменяет какой-либо параметр колебательной системы, совершая при этом работу. Характерным внешним признаком параметрического возбуждения является то, что частота модуляции (изменения параметра) может не совпадать с частотой возбуждаемых колебаний. В частности, легче всего параметрические колебания возбуждаются при отношении частот модуляции и колебаний, равном 2 : 1.

В предыдущем содержится уже целый ряд элементов класса и ф и к а ц и и автоколебательных систем, которая, однако, здесь не развивается.

При практической работе с автоколебаниями приходится решать две основные задачи: как получить колебания требуемой частоты, мощности и формы, если они нужны, или как устранить автоколебания, если они вредны. Основные рычаги управления — это фазовые соотношения и энергетический баланс.

ДОБАВЛЕНИЯ

1. Диаграммы работы и ортогональность функций

Мы пользовались диаграммами работы для наглядного графического пояснения энергетических соотношений. Аналитическое представление этих соотношений связано с одним общим свойством функций — с ортогональностью. Стало быть, диаграммы работы в принятой нами форме фигур Лиссажу могут служить геометрической интерпретацией свойства ортогональности. Взаимную связь всех этих понятий (ортогональность — энергетические соотношения — диаграммы работы — фигуры Лиссажу) мы постараемся здесь выяснить.

1. Энергия некоторого физического процесса может быть выражена интегралом

$$E = \int_{t_1}^{t_2} u(t)v(t) dt, \quad (1)$$

где u и v — две физические величины, выбранные так, что их произведение имеет смысл и размерность мощности (например, напряжение и ток; сила и скорость). Если ввести

$$w(t) = \int^t v(t) dt, \quad (2)$$

то формула (1) может быть переписана в виде

$$E = \int_{t_1}^{t_2} u(t) dw(t). \quad (3)$$

Здесь уже произведение uw имеет размерность работы (или энергии). Например, u и w могут означать силу и смещение (напряжение и заряд). Если процесс периодический, то интегрирование от t_1 до $t_1 + T$ дает нам энергию за один период.

2. Если теперь выбрать u и w за прямоугольные координаты и построить в них график процесса, исключив параметр t , то мы получим фигуру Лиссажу, которая является диаграммой работы, т. е. диаграммой, площадь которой позволяет непосредственно судить о величине энергии, а направление обхода — о направлении движения энергии, т. е. о том, вкладывается энергия в систему или отбирается от нее. Если площадь фигуры равна нулю, то это означает равенство нулю энергии. При этом нужно напомнить, что площади замкнутого контура присваивается тот или иной знак в зависимости от направления обхода, так что, например, фигура в форме симметричной восьмерки имеет нулевую площадь.

3. Определение ортогональности гласит: две функции $u(t)$ и $v(t)$ ортогональны друг другу на промежутке (t_1, t_2) , если

$$\int_{t_1}^{t_2} u(t)v(t) dt = 0. \quad (4)$$

Сопоставляя это с (1), мы видим, что равенство нулю энергии, ортогональность функций u и v и равенство нулю площади фигуры Лиссажу, представляющей нам диаграмму работы, — это лишь различные физические или математические аспекты одних и тех же соотношений.

Напомним теперь некоторые простейшие зависимости. Рассмотрим интеграл

$$I = \int_{-c}^c u(t)v(t) dt. \quad (5)$$

Если обозначить

$$f(t) = \int_0^t u(t)v(t) dt, \quad (6)$$

то значение интеграла (5) есть

$$I = f(c) - f(-c). \quad (7)$$

Важно отметить два частных случая: когда $f(t)$ четная функция, т. е.

$$f(t) = f(-t), \quad (8)$$

и когда $f(t)$ нечетная функция, т. е.

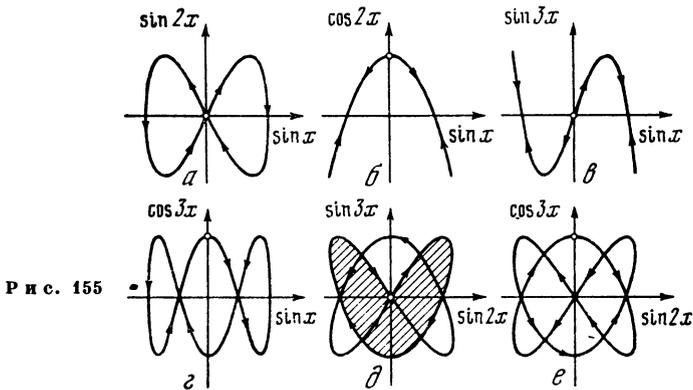
$$f(t) = -f(-t). \quad (9)$$

В первом случае мы получим равенство нулю интеграла (5) на любом промежутке $(-c, c)$ в силу равенства (8). Во втором же случае, т. е. когда функция нечетна, мы получим равенство нулю интеграла только при условии

$$f(\pm c) = 0, \quad (10)$$

откуда и можно найти значение c , при котором равенство удовлетворяется.

Эти соображения могут быть применены и к более общему виду функции на том основании, что всякая функция может быть раз-



ложена на четную и нечетную части при помощи элементарных соотношений

$$f(t) = f_1(t) + f_2(t);$$

$$f_1(t) = \frac{1}{2} [f(t) + f(-t)] \quad (\text{четная часть}),$$

$$f_2(t) = \frac{1}{2} [f(t) - f(-t)] \quad (\text{нечетная часть}).$$

4. Рассмотрим в качестве примера ортогональность тригонометрических функций. Она выражается, как известно, соотношениями

$$\int_{-\pi}^{\pi} \sin px \sin qxdx = 0 \quad (p \neq q), \quad (11)$$

$$\int_{-\pi}^{\pi} \sin px \cos qxdx = 0 \quad (\text{при любых } p \text{ и } q), \quad (12)$$

где p и q — неотрицательные целые числа. Соотношения (11) и (12) играют большую роль в теории рядов Фурье.

Воспользовавшись обозначениями (2) и (3), перепишем (11) и (12) в виде

$$\int_{-\pi}^{\pi} \sin px \, d(\cos qx) = 0 \quad (p \neq q), \quad (13)$$

$$\int_{-\pi}^{\pi} \sin px \, d(\sin qx) = 0 \quad (\text{при любых } p \text{ и } q) \quad (14)$$

или

$$\int_{-\pi}^{\pi} \cos px \, d(\cos qx) = 0.$$

Формулы (13) и (14) могут быть иллюстрированы при помощи фигур Лиссажу. Рассмотрим несколько примеров (рис. 155).

а) $u = \sin 2x$, $w = \sin x$. Фигура имеет форму восьмерки, площадь которой равна нулю, так как две равные петли, из которых состоит фигура, оббегаются в противоположных направлениях.

б) $u = \cos 2x$, $w = \sin x$.

в) $u = \sin 3x$, $w = \sin x$.

Оба эти случая отличаются тем, что контур вырождается в линию, пробегаемую за период в противоположных направлениях. Это означает ортогональность на любом промежутке. Выше указывалось, что такого рода ортогональность получается в том случае, когда

$$f(x) = \int_0^x u(x) v(x) \, dx = \int_0^x u(x) \, dw(x)$$

есть четная функция. Мы имеем, например, для рис. 155

$$f(x) = \int_0^x \cos 2x \, d(\sin x) = \cos x - \frac{2}{3} \cos^3 x,$$

т. е. вышеуказанное условие выполнено.

г) $u = \cos 3x$, $w = \sin x$. Фигура состоит из трех петель. Внутренняя петля обходится по часовой стрелке; площадь ее положительна. Две крайние петли обходятся против часовой стрелки. Результирующая площадь всей фигуры равна нулю.

д) $u = \sin 3x$, $w = \sin 2x$.

е) $u = \cos 3x$, $w = \sin 2x$.

Обе фигуры состоят из двух взаимно пересекающихся петель. Они совершенно сходны, однако разница заключается, во-первых, в положениях точек, отвечающих значению $x=0$ (на рис. 155 эта точка везде отмечена кружком), а во-вторых, в направлении обхода. Результирующие площади обеих фигур равны, конечно, нулю.

2. Уравнения шимми

Основное кинематическое соотношение, выражающее радиус кривизны траектории колеса через деформации пневматика, пишется в виде

$$d\theta/ds = 1/R = \alpha\lambda - \beta\varphi, \quad (1)$$

где s — пробегаемый колесом путь; θ — угол поворота стойки; R — радиус кривизны; λ — боковое смещение; φ — угол закручивания.

А. При отсутствии выноса моменты равны нулю и деформации закручивания нет. В этом случае имеем из (1)

$$d\theta/ds = \alpha\lambda. \quad (2)$$

Если добавить очевидное соотношение

$$d\lambda/ds = -\theta, \quad (3)$$

то получаем систему уравнений кинематического шимми. Исключая из системы (2) и (3) деформацию λ , получаем для угла θ уравнение второго порядка

$$d^2\theta/ds^2 + \alpha\theta = 0. \quad (4)$$

Величина α характеризует длину волны выписываемой колесом синусоиды.

Б. При наличии выноса положение усложняется. Мы получаем систему уравнений

$$d(\theta + \varphi)/ds = \alpha\lambda - \beta\varphi, \quad (5)$$

$$d\lambda/ds + t d\theta/ds = -(\theta + \varphi), \quad (6)$$

$$at\lambda + b\varphi = 0. \quad (7)$$

Уравнение (5) отличается от (2) тем, что, во-первых, кривизна выражается через производную по пути уже не угла поворота стойки, а суммарного угла, учитывающего закручивание колеса; во-вторых, в правой части имеются оба члена соотношения (1). Уравнение (6) отличается от (3) тем, что боковое смещение колеса уже не λ , а $\lambda + t\theta$; в правой части стоит также суммарный угол. Наконец, уравнение (7) есть уравнение моментов. При $t=0$ оно дает просто $\varphi=0$.

В. Переход к динамическому шимми состоит в том, что мы добавляем в уравнение моментов член, выражающий момент сил инерции, а также момент демпфера, если демпфер имеется. Мы получаем следующую систему

$$d(\theta + \varphi)/ds = \alpha\lambda - \beta\varphi. \quad (5)$$

$$d\lambda/ds = -(\theta + \varphi), \quad (8)$$

$$Jd^2\theta/d\tau^2 + h d\theta/d\tau = b\varphi. \quad (9)$$

Уравнение (8) получается из (6), если положить вынос $t=0$. В уравнении (9) производные по времени τ можно заменить производными по пути s на том основании, что

$$s = v\tau,$$

где v — скорость самолета. Сделав эту замену, получим вместо (9)

$$v^2 J d^2\theta/ds^2 + vhd\theta/ds = b\varphi. \quad (10)$$

Если исключить λ из системы (5), (8), (10), то получается

$$d^2\varphi/ds^2 + \beta d\varphi/ds + \alpha\varphi + d^2\theta/ds^2 + \alpha\theta = 0, \quad (11)$$

$$v^2 J d^2\theta/ds^2 + vhd\theta/ds - b\varphi = 0. \quad (12)$$

В таком виде уравнения ясно показывают наличие двух связанных систем (переменные φ и θ и члены связи).

Г. Уравнения динамического шимми при выносе, не равном нулю, пишутся в виде

$$d(\theta + \varphi)/ds = \alpha\lambda - \beta\varphi, \quad (5)$$

$$d\lambda/ds + t d\theta/ds = -(\theta + \varphi), \quad (6)$$

$$v^2 J d^2\theta/ds^2 + vhd\theta/ds = at\lambda + b\varphi. \quad (13)$$

В уравнении (13) по сравнению с (10) появился еще член, зависящий от t .

Д. Для колеса со свободным поперечным перемещением (как на рис. 132) в уравнениях появляется еще одна переменная — поперечное смещение x . Система уравнений принимает вид

$$d(\theta + \varphi)/ds = \alpha\lambda - \beta\varphi, \quad (5)$$

$$d\lambda/ds + dx/ds = -(\theta + \varphi), \quad (14)$$

$$v^2 J d^2\theta/ds^2 = b\varphi, \quad (15)$$

$$v^2 M d^2x/ds^2 = a\lambda. \quad (16)$$

Уравнение (14) отличается от (8) тем, что боковое смещение выражается суммой свободного перемещения вилки x и деформации λ . В уравнении (15) отброшено демпфирование. Уравнение (16) есть уравнение сил для поперечного движения; оно учитывает силу инерции колеса, масса которого M .

Исследование характера движения производится для всех приведенных случаев обычным способом — исследованием корней характеристического уравнения. Вещественные корни соответствуют аperiодическим режимам, комплексные — колебательным. Чисто мнимые корни отвечают незатухающим колебаниям (например, уравнение (4)). Комплексные корни с отрицательными вещественными частями отвечают затухающим колебаниям.

Самовозбуждающиеся автоколебания возможны только в тех случаях, в которых комплексные корни характеристического уравнения имеют положительные вещественные части. Эти со-

отношения позволяют выразить неравенствами условия невозможности автоколебаний, т. е. условия устойчивости. Так, например, условие устойчивости для случая Д — для колеса со свободным поперечным перемещением — состоит в том, чтобы

$$\omega_2 = \sqrt{b/J} < \omega_1 = \sqrt{a/M},$$

где ω_2 — частота крутильных колебаний колеса; ω_1 — частота колебаний бокового смещения. Если это условие выполнено, шимми не возникает ни при каких скоростях.

3. Автоколебания регуляторов

Регулятором называется устройство для автоматического поддержания постоянства режима той или иной машины (или аппарата) при изменении условий ее работы (например, при изменении нагрузки).

При известных условиях возможны автоколебания регулятора; это явление имеет, очевидно, для техники большое значение.

Современная теория регулирования представляет собою широко развитую отрасль науки; вопрос об автоколебаниях регуляторов является с точки зрения теории регулирования одним из многочисленных частных вопросов, разработанных до тонкостей даже для наиболее сложных случаев. Нет никакой возможности в рамках этой небольшой книги дать хотя бы общий обзор результатов, добытых теорией регулирования.

Но для нас регулятор является лишь одним из примеров автоколебательных систем. Поэтому уместно будет высказать некоторые общие соображения и разобрать в качестве примера простейшую систему регулятора.

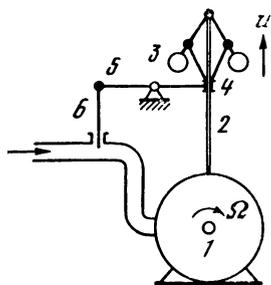
Прежде всего заметим, что автоколебания в регуляторах в принципе возможны потому, что в любом регуляторе имеется обратная связь. В самом деле, действие регулятора в том и состоит, что изменения регулируемой величины так или иначе воздействуют через посредство регулирующего органа на режим системы так, чтобы уменьшить изменения регулируемой величины. Таким образом, в системе машина—регулятор образуется замкнутая цепь прямой и обратной связи, так что общая схема машины с регулятором очень сходна с общей схемой автоколебательной системы.

Для некоторых систем регуляторов автоколебания составляют основу их действия. Это так называемые регуляторы прерывистого действия. Примером такого рода систем может служить терморегулятор, рассмотренный в § 19. Процесс прерывистого регулирования имеет колебательный характер, так что регулируемая величина (например, температура в случае терморегулятора) все время колеблется около заданного среднего значения.

Мы не будем больше заниматься этой разновидностью регуляторов и обратимся к таким регуляторам, которые отзываются на

изменение регулируемой величины соответствующим изменением режима. При этом процесс регулирования, отвечающий внезапному скачкообразному изменению условий работы, имеет аперiodический (что желательно во многих случаях) или затухающий колебательный характер. Возникновение автоколебаний в таких регуляторах недопустимо.

В качестве примера мы рассмотрим классическую схему центробежного регулятора, применяемого с давних пор для регулирования



Р и с. 156

скорости паровых машин. Схема центробежного регулятора дана на рис. 156. Машина 1 приводит во вращение вертикальный вал 2. Тяжелые шары 3, укрепленные на шарнирных рычагах, вращающихся вместе с валом 2, расходятся, стремясь двигаться прямолинейно, и поднимают муфту 4. Последняя через посредство рычага 5 передвигает заслонку 6, изменяющую количество пара, поступающего в машину, а следовательно, и вращающий момент. Известно, что при уменьшении нагрузки скорость машины склонна возрасти. Действие регулятора состоит в том,

что он препятствует этому, уменьшая подачу пара, так что благодаря наличию регулятора изменения скорости при изменении нагрузки гораздо меньше, чем если бы регулятора не было.

Описанный простейший регулятор относится к статическим регуляторам прямого действия. Исчерпывающая теория этого рода регуляторов дана еще в 1876 г. И. А. Вышнеградским [24]. Мы повторим здесь вкратце его теорию и выводы.

Рассматриваемая система есть связанная система, состоящая из машины и регулятора, причем машина влияет на регулятор, а регулятор на машину. Следовательно, система будет описываться двумя уравнениями с соответствующими членами связи.

У р а в н е н и е р е г у л я т о р а. В качестве переменной, описывающей движение регулятора, выберем смещение u муфты. При этом положительное направление u пусть отвечает положительному приращению скорости машины (т. е. $+u$ направлено вверх, см. рис. 156). В первом приближении регулятор представляет собою линейную систему с одной степенью свободы. Внешняя сила, действующая на эту систему, зависит от скорости машины; в том же приближении эта сила может считаться пропорциональной относительному изменению скорости машины. Таким образом, получаем уравнение

$$d^2u/dt^2 + 2\alpha du/dt + \omega_0^2 u = k_1 \Delta \Omega / \Omega_0. \quad (1)$$

Здесь α — показатель затухания, ω_0 — собственная частота регулятора, зависящая от приведенной к муфте массы поступа-

тельно движущихся частей регулятора, а также от возвращающей (квазиупругой) силы. Последняя зависит от веса шаров, а при наличии пружин — от их упругости. Через Ω_0 обозначена номинальная угловая скорость машины, а через $\Delta\Omega = \Omega - \Omega_0$ — ее приращение.

Уравнение машины. Машина также представляет собою систему с одной степенью свободы. Ее уравнение мы получим, приняв, что изменение внешнего момента вызывает ускорение машины. Уменьшение нагрузки эквивалентно увеличению вращающего момента. Но одновременно возникает противоположный дополнительный момент, обусловленный работой регулятора. Этот последний момент в первом приближении считается пропорциональным смещению муфты регулятора (а следовательно, и заслонки). Таким образом, уравнение машины запишется в виде

$$J d\Omega/dt = \Delta M - k_2 u. \quad (2)$$

Здесь J — момент инерции машины; ΔM — приращение момента, обусловленное изменением нагрузки; $-k_2 u$ — дополнительный момент, обусловленный работой регулятора.

Решая совместно (1) и (2), т. е. дифференцируя (1) по t и подставляя $d\Omega/dt$ из (2), получим для u уравнение третьего порядка

$$\frac{d^3 u}{dt^3} + 2\alpha \frac{d^2 u}{dt^2} + \omega_0^2 \frac{du}{dt} + \frac{k_1 k_2}{J \Omega_0} u = \frac{k_1}{\Omega_0 J} \Delta M. \quad (3)$$

Очевидно, что такое же уравнение можно получить и для Ω .

Свойства системы определяются, как обычно, исследованием корней характеристического уравнения. Автоколебания возможны, если имеются комплексные корни с положительными вещественными частями.

Для целей анализа Вышнеградский вводит вспомогательные величины $x = 2\alpha (k_1 k_2 / J \Omega_0)^{-1/2}$; $y = \omega_0^2 (k_1 k_2 / J \Omega_0)^{-1/2}$ и показывает, что условие устойчивости есть

$$xy > 1$$

или

$$2\alpha \omega_0^2 J \Omega_0 / k_1 k_2 > 1. \quad (4)$$

Отсюда сразу следует знаменитый тезис о том, что «без катаракта¹ нет регулятора». Действительно, при отсутствии катаракта α равно нулю (или во всяком случае очень мало), а тогда автоколебания неизбежны. Условие (4) показывает также, что

¹ Катарактом называется гаситель колебаний (демпфер), состоящий из цилиндра, заполненного маслом, в котором перемещается поршень с отверстием. Движению поршня противодействует сила вязкого трения, возникающего при продавливании масла через отверстие. Предполагается, что эта сила пропорциональна скорости поршня.

при недостаточном затухании, вносимом катарактом, полезно применять маховик (увеличение J), повышать коэффициент неравномерности (увеличение ω_0) и т. д. Дальнейший анализ корней характеристического уравнения позволяет установить условия, при которых процесс регулирования имеет аperiodический или колебательный характер. Результаты этого анализа представлены в графической форме на рис. 157. По осям отложены параметры x

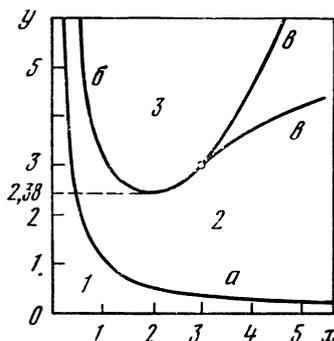


Рис. 157

и y . Площадь диаграммы делится на области. Область 1 — автоколебания, область 2 — колебательный затухающий процесс регулирования, область 3 — аperiodический процесс регулирования. Уравнения кривых, служащих границами между названными областями, таковы

а) $xy - 1 = 0$,

б) $2x^3 - 9xy + 27 = 0$,

в) $4(x^3 + y^3) - x^2y^2 - 18xy + 27 = 0$.

В заключение заметим, что общее качественное представление о возможности автоколебаний в описанной системе регулятора и о необходимости в связи с этим применения катаракта можно было бы сделать на основании элементарного анализа с применением приемов, описанных в § 14.

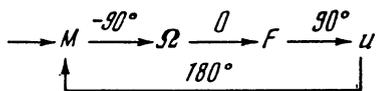
Составим функциональную схему машины с регулятором, введя кроме вышеперечисленных еще одну величину — силу F , действующую на регулятор и обусловленную изменением скорости машины, т. е.

$$F = k_1 \Delta \Omega / \Omega_0. \tag{5}$$

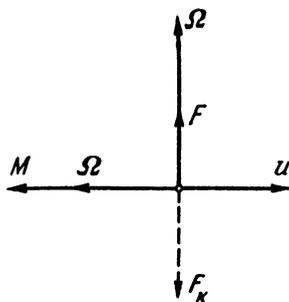
С введением этой величины схема может быть представлена рис. 158.

Предположим теперь, что происходят колебания. Тогда, разумея лишь первые гармоники переменных составляющих всех

обозначенных на схеме рис. 158 величин, можем построить для них векторную диаграмму рис. 159. Начинаем с вектора u . Вектор M в противофазе, так как увеличение u вызывает уменьшение M (падающая характеристика зависимости M от u). Момент совпадает по фазе с ускорением машины; значит, вектор скорости Ω отстает от M на 90° . F и Ω совпадают по фазе, что вытекает из (5). Мы видим на диаграмме, что F опережает u на 90° и что, следовательно,



Р и с. 158



Р и с. 159

автоколебания возможны. Совершенно очевидно, что сила F , возбуждающая автоколебания, может быть скомпенсирована большей и противоположно направленной силой F_k вязкого трения в катаракте.

4. Об RC генераторах синусоидальных колебаний

Вясним условия, при которых схема, составленная только из конденсаторов и сопротивлений (не считая, конечно, источника энергии и усилителя) оказывается способной генерировать синусоидальные колебания, т. е. приобретает свойства почти гармонической системы.

С математической точки зрения почти гармоническая система описывается уравнением второго порядка

$$\ddot{x} + 2\alpha\dot{x} + \omega_0^2x = 0 \quad (1)$$

и условие самовозбуждения есть (2)

$$\alpha \leq 0.$$

Порядок уравнения зависит от числа степеней свободы и от структуры отдельных ветвей схемы. Если ветвь содержит только R и C , то дифференциальное уравнение этой ветви имеет порядок не выше первого. Если же составить схему из нескольких ветвей, т. е. увеличить число степеней свободы, то порядок уравнения соответственно повысится. В частности, для схемы с двумя степенями свободы можно получить уравнение вида (1).

Рассмотрим в качестве примера схему рис. 160. Для этой схемы

$$\dot{U}_2 + 2\alpha U_2 + \omega_0^2 \int U_2 dt = \frac{1}{R_1 C_2} U_1, \quad (3)$$

где

$$\omega_0^2 = 1/R_1 R_2 C_1 C_2, \quad (4)$$

$$2\alpha = 1/R_1 C_1 + 1/R_2 C_2 + 1/R_1 C_2.$$

Уравнение (3) вполне аналогично уравнению для цепи и последовательно соединенных R , L и C .

Однако схема рис. 160 всегда аperiodична. Действительно, вычисляя относительное затухание $\kappa = \alpha/\omega_0$, найдем для данного случая

$$\kappa = \frac{1}{2} \left(\sqrt{\frac{R_2 C_2}{R_1 C_1}} + \sqrt{\frac{R_1 C_1}{R_2 C_2}} + \sqrt{\frac{R_2 C_1}{R_1 C_2}} \right). \quad (5)$$

Эта величина больше единицы для любой комбинации параметров, и, следовательно, свободные процессы в схеме не имеют колебательного характера.

Но если теперь мы введем в схему усилитель и образуем цепь обратной связи так, чтобы она влияла на затухание, то положение

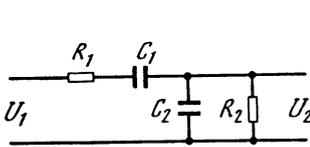


Рис. 160

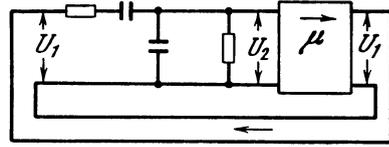


Рис. 161

радикально изменится. Схема с добавлением усилителя и обратной связи показана на рис. 161. Если положить просто

$$U_1 = \mu U_2, \quad (6)$$

где μ — коэффициент усиления по напряжению, то получится следующее уравнение:

$$U_1 + \left(\frac{1}{R_1 C_1} + \frac{1}{R_2 C_2} + \frac{1 - \mu}{R_1 C_2} \right) \dot{U}_1 + \frac{1}{R_1 R_2 C_1 C_2} U_1 = 0, \quad (7)$$

и теперь ясно видно, что при надлежащей величине коэффициента усиления возможно самовозбуждение, условие которого есть

$$\frac{1}{R_1 C_1} + \frac{1}{R_2 C_2} + \frac{1 - \mu}{R_1 C_2} \leq 0, \quad (8)$$

и что при этом условии будут генерироваться синусоидальные колебания с частотой ω_0 . Схема рис. 161 обладает всеми свойствами

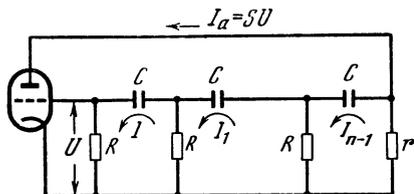
почти гармонической автоколебательной системы. Эта схема есть схема двухлампового генератора RC (так как условие (6) выполняется для четного числа каскадов).

Теперь рассмотрим одноламповую схему с цепочкой RC (рис. 162). Поведение схемы зависит от числа n звеньев цепочки. Положим $I_a = SU$.

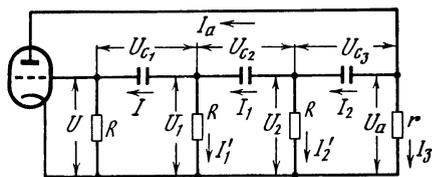
При $n=1$ получим уравнение для напряжения на сетке

$$\left(1 + \frac{r}{R} + rS\right) \dot{U} + \frac{1}{RC} U = 0, \quad (9)$$

т. е. уравнение первого порядка. Периодические синусоидальные колебания в такой схеме, очевидно, невозможны.



Р и с. 162



Р и с. 163

При $n=2$ уравнение принимает вид

$$\left(1 + 2\frac{r}{R} + rS\right) \dot{U} + \left(3 + \frac{r}{R}\right) \frac{1}{RC} \dot{U} + \frac{1}{R^2C^2} U = 0. \quad (10)$$

Это уравнение имеет хотя и второй порядок, но обратная связь не может обратить в нуль коэффициент при \dot{U} , и синусоидальные автоколебания в этой схеме также невозможны. Однако при достаточном значении S свободные процессы имеют затухающий колебательный характер.

Как известно, схема рис. 162 может генерировать синусоидальные колебания лишь при $n \geq 3$. При $n=3$, в частности, получается уравнение третьего порядка, из исследования коэффициентов которого следует, что схема может генерировать колебания с частотой

$$\omega_0 = \frac{1}{RC} \sqrt{\frac{R}{6R + 4r}} \quad (11)$$

и что условие самовозбуждения есть

$$rS \geq 29 + 23r/R + 4r^2/R^2. \quad (12)$$

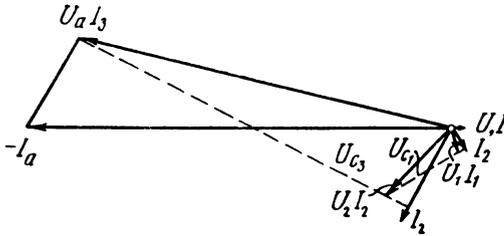
Эти соотношения приведены у К. Ф. Теодорчика [4, стр. 84—85]. Мы поясним их при помощи векторных диаграмм.

На рис. 163 изображена схема генератора с трехзвенной цепочкой. На схеме обозначены токи и напряжения. Выберем $R=r=1$;

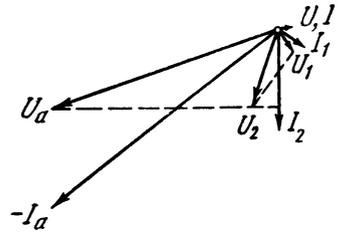
масштаб для токов и напряжений будет одинаковый. Из формулы (11) находим

$$1/\omega_0 C = \sqrt{10} = 3,16.$$

Начинаем построение (рис. 164) с вектора U ; этот же вектор изображает I . Из конца вектора U перпендикулярно к вектору I откладываем вектор $U_{c_1} = I/\omega_0 C$; сумма U и U_{c_1} дает U_1 . Этот же вектор изображает ток I'_1 . Складывая I и I'_1 , находим I_1 . Из конца вектора U_1 перпендикулярно к вектору I_1 откладываем $U_{c_2} = I_1/\omega_0 C$.



Р и с. 164



Р и с. 165

Складывая U_1 и U_{c_2} , получаем U_2 . Этот же вектор изображает I'_2 . Складывая I_1 и I'_2 , находим I_2 . Из конца вектора U_2 перпендикулярно к I_2 откладываем $U_{c_3} = I_2/\omega_0 C$ и находим U_a и I_3 . Сумма I_2 и I_3 дает вектор тока I_a с обратным знаком. Этот вектор в противофазе с U ; следовательно, \bar{U} и I_a совпадают по фазе, в чем и состоит условие баланса фаз, так как при этом может быть выполнено равенство $I_a = SU$. При другом значении частоты фазовые соотношения будут нарушены. Это иллюстрируется диаграммой рис. 165, построенной для $1/\omega C = 2$ (остальные параметры те же). Что касается баланса амплитуд, то диаграмма показывает необходимое значение крутизны, как отношения I_a к U . В нашем примере по формуле (12) имеем

$$S = I_a/U = 29 + 23 + 4 = 56.$$

Векторные диаграммы поясняют, почему не может генерировать схема с двухзвенной цепочкой. Дело в том, что ни при каких соотношениях между R и $1/\omega C$ вектор U_a не может при наличии всего двух звеньев развернуться относительно вектора \bar{U} на угол больше 180° (считая по часовой стрелке). А это необходимо для того, чтобы вектор I_a мог совпасть по направлению с U . При $n=3$ такой фазный разворот уже оказывается возможным; он тем более возможен при числе звеньев, большем трех.

Л и т е р а т у р а

Общие сочинения

1. Л. И. Мандельштам. Собрание трудов. Изд-во АН СССР, 1947.
2. Рэлея. Теория звука, т. II. Гостехиздат, 1944.
3. А. А. Андронов и С. Э. Хайкин. Теория колебаний, ч. 1. ОНТИ, 1937.

4. *К. Ф. Теодорчик*. Автоколебательные системы. Гостехиздат, 1952.
5. *А. А. Фельдбаум*. Введение в теорию нелинейных цепей. Энергоиздат, 1948.

К § 7

6. *Л. В. Арнольд*. Строительные пневматические инструменты. ОНТИ, 1936.
7. *А. Я. Шескин и С. Е. Гольдич*. Памятка по уходу за пневматическими молотками. Машиз, 1946.

К § 8

8. *С. П. Стрелков*. Маятник Фроуда. Журн. техн. физ. 3, 563 (1933).
9. *А. А. Харкевич*. О применении диаграмм сила--перемещение к исследованию колебаний. Журн. техн. физ., 12, 607, 1942.

К § 9

10. *А. И. Каширин*. Исследование вибраций при резании металла. Изд-во АН СССР, 1944.

К § 10

11. *Б. П. Константинов*. Об автоколебаниях и звукообразовании язычка гармонии. Журн. техн. физ., 9, 1820, 1939.

К § 21

12. *Д. И. Блохинцев*. Возбуждение резонаторов потоком воздуха. Журн. техн. физ., 1945, № 15.
13. *Д. И. Блохинцев*. Вихревой звук. Журн. техн. физ., 1945, № 15.
14. *С. П. Стрелков*. Опыт с колебанием шарового маятника в потоке воздуха. Журн. техн. физ., 9, 1763, 1939.

К § 22

15. *Е. П. Гроссман*. Флаттер. Труды ЦАГИ, № 284, 1937.

К § 23

16. *М. В. Келдыш*. Шимми переднего колеса трехколесного шасси. Труды ЦАГИ, № 564, 1945.
17. *И. И. Эскин*. Самолетные шасси с носовым колесом. Оборонгиз, 1947.

К § 24

18. *Л. И. Мандельштам и Н. Д. Папалекси*. О параметрическом возбуждении электрических колебаний. Журн. техн. физ., 4, 5, 1934.
19. *В. А. Лазарев*. О гетеропараметрическом возбуждении. Журн. техн. физ., 4, 30, 1934.

К §§ 25, 26

20. *Р. Сарбахер и В. Эдсон*. Техника сверхвысоких частот. Связьиздат, 1947.
21. *В. И. Калинин*. Генерирование дециметровых и сантиметровых волн. Связьиздат, 1948.
22. *В. Ф. Коваленко*. Введение в электронику сверхвысоких частот. «Советское радио», 1950.
23. *Е. М. Студенков*. Электронно-лучевые генераторы ультравысокочастотных колебаний. УФН, 23, 417, 1940.

К добавлению

24. *Д. К. Максвелл, И. А. Вышнеградский, А. Стодола*. Теория автоматического регулирования. Изд-во АН СССР, 1949.

УМНОЖАЮЩИЕ СХЕМЫ

В большом числе случаев необходимо располагать устройством, производящим перемножение двух функций времени. Говоря более определенно, требуется шестиполюсник с двумя входами, на которые подаются два независимых напряжения (или тока), и выходом, с которого снимается напряжение (или ток), пропорциональное произведению обоих входных напряжений. Умножение напряжений требуется, например, при преобразовании частоты и при модуляции¹. Оно оказывается необходимым в гетеродинных анализаторах спектра.

В новейшее время надобность его возникла в связи с построением корреляторов — приборов для получения функций корреляции различных случайных процессов.

Оказывается, что получить переменожение двух напряжений непосредственно не представляется возможным, так как неизвестны физические явления, в которых какая-либо переменная электрическая величина выражалась бы непосредственно произведением двух других взаимно независимых переменных величин.

Даже если привлечь неэлектрические величины, то число явлений, могущих быть использованными для указанных целей, не так уже велико. Среди них следует назвать прежде всего явления электростатического и электродинамического взаимодействия. Так, например, по закону Кулона сила притяжения двух зарядов пропорциональна их произведению. Аналогично сила взаимодействия между двумя проводниками с токами пропорциональна произведению токов.

В обоих этих примерах характерно, что произведению двух электрических величин пропорциональна механическая величина — сила взаимодействия. Поэтому, если желательно построить на основе этих явлений электрический шестиполюсник, то следует ввести в его состав электромеханический преобразователь.

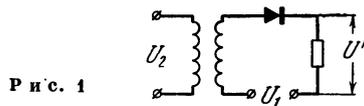
Принципиальную схему умножителя можно представить, например, в следующем виде: на противоположащие грани пьезоэлектрической пластинки наложены две системы параллельных

¹ Следует заметить, что умножение есть линейная операция, выполняемая в линейной системе с переменными параметрами. Широко распространено мнение, что модуляция — процесс принципиально нелинейный; это ошибочное мнение основано на том, что практические схемы модуляции в большинстве нелинейны.

проводов, обтекаемых независимыми токами. Сила взаимодействия между этими системами, пропорциональная произведению токов, сжимает пластинку и вызывает соответствующую разность потенциалов, которая и является выходной электрической величиной для данного умножающего шестиполюсника.

Умножающие схемы. В настоящее время умножение предпочитают осуществлять в различных смесительных и модуляторных схемах. Хотя все эти схемы общеизвестны, их генезис и возможности не всегда излагаются достаточно ясно; уделим внимание этой стороне вопроса.

Сущность умножения во всех смесителях и модуляторах сводится к тому, что два независимых напряжения действуют в одной и той же нелинейной цепи. При этих условиях можно получить на выходе схемы составляющую напряжения, пропорциональную



произведению входных напряжений. Но, кроме того, получаются и нежелательные составляющие в ассортименте, зависящем как от характеристики нелинейного элемента, так и от свойств схемы. Последовательным усложнением схемы можно добиться очищения выходного напряжения от лишних составляющих, но лишь до вполне определенного предела.

Начнем с простейшей схемы (рис. 1), в которой имеется нелинейный элемент с характеристикой

$$i = f(U). \quad (1)$$

В схеме действуют одновременно два внешних напряжения U_1 и U_2 ; с сопротивления снимается напряжение:

$$U' = R \left\{ f(U_0) + f'(U_0)(U_1 + U_2) + \frac{1}{2} f''(U_0)(U_1 + U_2)^2 + \frac{1}{6} f'''(U_0)(U_1 + U_2)^3 + \dots \right\}, \quad (2)$$

где U_0 — начальное напряжение, которое может, в частности, равняться нулю.

Перепишав (2) в виде

$$U' = R \{ a_0 + a_1(U_1 + U_2) + a_2(U_1^2 + \underline{2U_1U_2} + U_2^2) + a_3(U_1^3 + 3U_1^2U_2 + 3U_1U_2^2 + U_2^3 + \dots) \}, \quad (3)$$

мы видим, что член, содержащий произведение U_1U_2 (подчеркнут) получается только за счет наличия квадратичного члена в разложении (2). Кроме произведения, получается ряд нежелательных членов, некоторые из которых можно устранить следующим

приемом. Составив разность $U_3 = U' - U''$, где $U' = Rf(U_1 + U_2)$, $U'' = Rf(U_1 - U_2)$, получим

$$U_3 = 2R \{ a_1 U_2 + \underline{4a_2 U_1 U_2} + a_3 (3U_1^2 U_2 + U_2^3) + \dots \} \quad (4)$$

Эта операция осуществляется в балансной схеме рис. 2.

Можно сделать еще один шаг по пути усложнения схемы: возьмем вторую точно такую же балансную схему, но изменим в ней знаки обоих напряжений U_1 и U_2 на обратные. Тогда

$$U'_3 = 2R \{ -a_1 U_2 + \underline{4a_2 U_1 U_2} - a_3 (3U_1^2 U_2 + U_2^3) + \dots \}. \quad (4')$$

Сложив (4) и (4'), получим

$$U = U_3 + U'_3 = 8Ra_2 U_1 U_2 + \text{члены высших порядков}. \quad (5)$$

Описанные действия выполняются в кольцевой схеме рис. 3, представляющей собой совмещение двух схем рис. 2 с требуемым обращением знаков внешних напряжений.

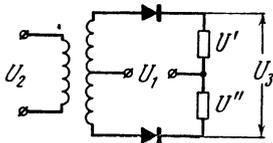


Рис. 2

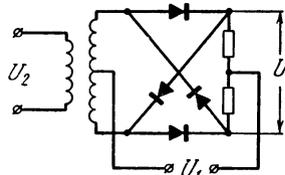


Рис. 3

Прием, примененный при синтезе схем рис. 2 и 3, состоит в том, что, изменяя знаки напряжений, мы изменяем соответственно знаки у тех или иных членов общего выражения (3). Составляя затем суммы или разности, мы устраняем соответствующие члены из состава результирующего напряжения. Посмотрим, нельзя ли идти по этому пути и далее.

Действие балансной схемы состоит в устранении всех членов, содержащих четные степени U_2 (см. формулу (4)). Что же касается кольцевой схемы, то она оставляет в результирующем напряжении только составляющие, содержащие произведение нечетных степеней U_1 и U_2 , так что (5) можно записать в виде

$$U = b_1 U_1 U_2 + b_2 U_1^3 U_2 + b_3 U_1 U_2^3 + b_4 U_1^3 U_2^2 + b_5 U_1^5 U_2 + \dots$$

или, вынося U_1, U_2 за скобку,

$$U = U_1 U_2 [b_1 + b_2 U_1^2 + b_3 U_2^2 + b_4 U_1^2 U_2^2 + b_5 U_1^4 + \dots]. \quad (6)$$

Последнее выражение показывает, что дальнейшее устранение нежелательных членов вышеописанным путем уже невозможно, так как в квадратных скобках стоят члены, зависящие только от четных степеней U_1 и U_2 , на знак которых изменение знака U_1 или U_2 не влияет. Таким образом, на кольцевой схеме чисто схемные возможности исчерпываются.

В конце концов дело обстоит так: при применении кольцевой схемы и при условии полной ее симметрии мы получаем чистое умножение лишь в том случае, когда $f(U)$ можно выразить степенным многочленом не выше третьей степени и обязательно содержащим член второй степени. Этот вывод обычно затемняется из-за плохой традиции делать все выкладки с тригонометрическими функциями.

Характеристика умножения. При испытании и налаживании всякой конкретной умножающей схемы необходимо описать ее свойства. В частности, желательно установить, насколько чисто схема производит требуемое умножение.

Общеизвестный способ испытания умножающей схемы состоит в том, что на нее подают два синусоидальных напряжения с частотами ω_1 и ω_2 ; схема производит чистое умножение, если на выходе не обнаруживается никаких частот, кроме $\omega_1 \pm \omega_2$, что следует из известных тригонометрических формул. Ниже излагается иной способ описания свойств умножающей схемы, дающий более полное и более наглядное представление о ней.

Мы привыкли оценивать свойства исследуемых нами систем по геометрическим образам, которые доставляют нам графические изображения тех или иных экспериментальных характеристик системы. Так, например, о линейности системы мы судим по форме характеристики, представляющей график зависимости выходной величины от входной. Отклонение от линейности легко обнаруживается и оценивается на глаз по отклонению графика от прямолинейности; прямая линия является в данном случае тем идеальным геометрическим образом, с которым мы сопоставляем реальную характеристику. Это сопоставление осуществляется тем легче и тем точнее, чем проще идеальный образ. Прямая несомненно относится к числу таких простейших образов на плоскости.

В качестве другого примера возьмем отклонение формы колебания от синусоидальной. Известно, что оценить на глаз небольшие отклонения графика от синусоиды очень трудно. Но можно выбрать более простой идеальный образ, например, окружность. Перестроив график колебания в полярных координатах, мы должны будем уже оценивать на глаз отклонения полурасположенной фигуры от правильной окружности, что является гораздо более чувствительным «реактивом».

Еще пример. Допустим, что мы желаем выбрать из семейства характеристик наиболее приближающуюся к параболе. Это трудно сделать на глаз. Но если изменить построение семейства, выбрав квадратичный масштаб по оси абсцисс, то искомая характеристика будет наиболее близкой к прямой, и можно без труда отличить ее от других.

Теперь перейдем к умножающей схеме. И здесь желательно найти такое графическое представление характеристик схемы, которое сразу позволило бы на глаз оценить ее достоинства. Речь идет опять о построении простейшего геометрического образа

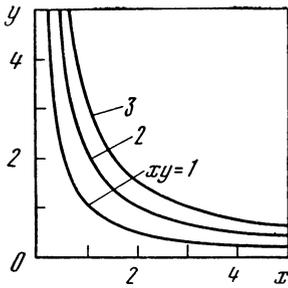


Рис. 4

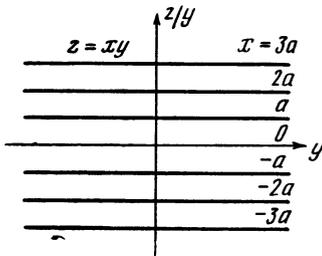


Рис. 5

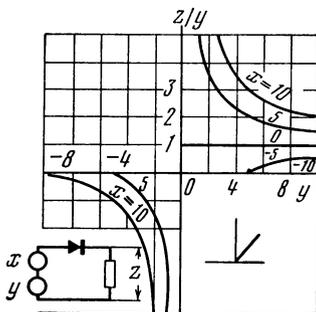


Рис. 6

для характеристик идеальной умножающей схемы. Однако в данном случае мы имеем дело с взаимозависимостью трех величин: двух входных и одной выходной. Поэтому следует выбрать образ в трехмерном пространстве. Если, например, отложить по трем взаимно перпендикулярным осям два сомножителя x и y и их произведение xy , то характеристическая поверхность будет поверхностью гиперboloида, сечения которого плоскостями, параллельными xoy , представляются семейством гипербол (рис. 4). Ясно, что такой гиперboloид в качестве идеального образа совершенно не годится. Простейшей поверхностью является плоскость; желательно найти также такие координаты или масштабы, чтобы характеристическая поверхность идеального умножителя представлялась плоскостью.

Естественно было бы применить в обычных декартовых координатах логарифмические или гиперболические масштабы. Но при этом начало координат уходит в бесконечность, так что эту возможность следует отбросить.

Очень простое решение задачи состоит в следующем. Пусть x и y — два сомножителя, z — их произведение. Если отложить x и y по двум декартовым осям, а по третьей отложить величину z/y , то получаемая поверхность есть плоскость. Плоскость эта может быть представлена своими сечениями на координатной плоскости $\frac{z}{y}oy$. Очевидно, что $z/y = x$, т. е. семейство характеристических линий представится параллельными оси y прямыми, как показано на рис. 5. Всякое отклонение реальных характеристик от такого рода картины, легко оце-

но, что $z/y = x$, т. е. семейство характеристических линий представится параллельными оси y прямыми, как показано на рис. 5. Всякое отклонение реальных характеристик от такого рода картины, легко оце-

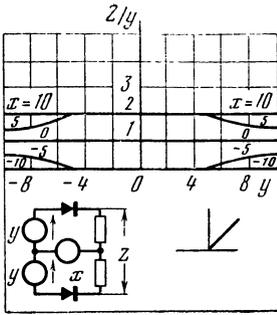


Рис. 7

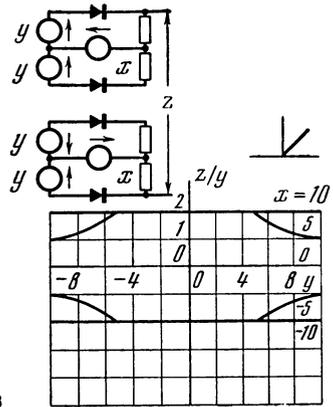


Рис. 8

ниваемое на глаз, указывает на несовершенство умножения.

Рассмотрим в качестве примера три схемы, приведенные на рис. 1—3, т. е. простую, балансную и кольцевую схему. Возьмем сначала ломаную характеристику вида

$$i = k(|U| + U).$$

Для простой схемы получим график рис. 6, показывающий, что ничего похожего на правильное умножение схема не дает. Характеристики балансной схемы изображены на рис. 7. Здесь видно, что семейство стало симметричным по отношению к y , но не x . Это является еще очень грубым нарушением. Наконец, на рис. 8, относящемся к кольцевой схеме, видна уже двусторонняя симметрия, однако правильного умножения нет вследствие влияния излома характеристики.

Возьмем теперь характеристику, выражаемую формулой $i = i_0 + a_1 U + a_2 U^2$. На рис. 9 представлены характеристики простой

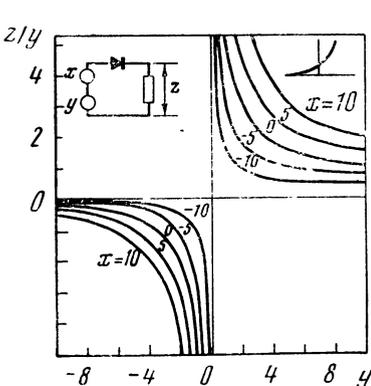


Рис. 9

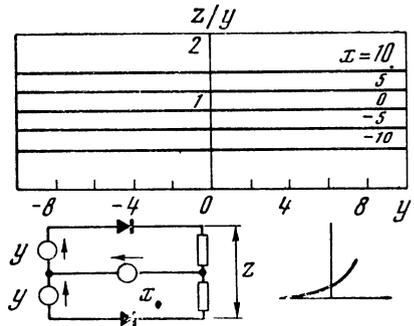


Рис. 10

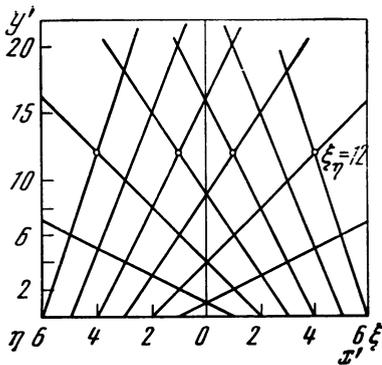


Рис. 11

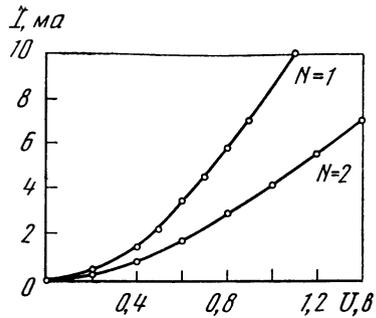


Рис. 12

схемы, на рис. 10 — балансной. Последний график показывает уже приближение к желаемым зависимостям; в смещении всего семейства вверх, нарушающем симметрию, проявляется наличие линейного члена в выходном напряжении. Переход к кольцевой схеме устраняет этот дефект; в кольцевой схеме мы получаем при квадратичной характеристике идеальное умножение, изображаемое графиком рис. 5.

Существует еще одна интересная возможность графического представления произведения, при которой характеристическая поверхность является плоскостью. Речь идет о построении такой

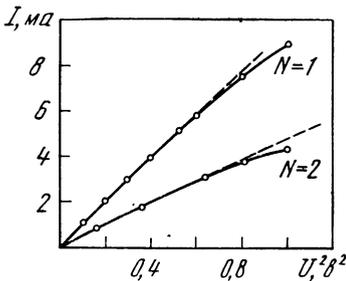


Рис. 13

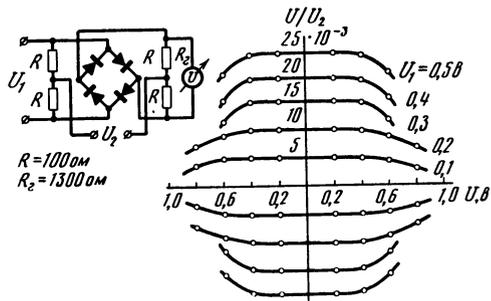


Рис. 14

системы координат, в которой по третьей оси можно было бы откладывать непосредственно произведение. Рассмотрим рис. 4. Если бы характеристическая поверхность была плоской, то ее сечения представлялись бы параллельными прямыми. Попробуем так деформировать координатную сетку рис. 4, чтобы гиперболы распрямлялись. Для этого нужно, очевидно, развернуть оси x и y на большой угол. Основой для построения новой системы координат служит то обстоятельство, что предельная гипербола се-

мейства, отвечающая уравнению $xy=0$, совпадает с осями. Для ее спрямления нужно повернуть оси y на 90° влево. Таким образом, в новой координатной системе $\xi\eta$ оси представляют собой две полу-прямые, исходящие из начала координат в противоположные стороны.

В результате такой деформации искажается соответственно и вся координатная сетка. Получается косоугольная система, представленная на рис. 11 и описываемая уравнениями

$$x' = \xi - \eta, \quad y' = \xi\eta,$$

где $x'y'$ — прямоугольные координаты. Линии равных произведений представляются в этой координатной системе параллельными прямыми. На рис. 11 показана для примера прямая $\xi\eta=12$, проходящая через точки с координатами (2.6), (3.4), (4.3) и (6.2). Следует отметить, что координатная поверхность $\xi\eta$ — двулистная.

Практический пример. При практическом применении умножающих схем в тех случаях, когда к качеству умножения предъявляются сколько-нибудь серьезные требования, рекомендуется прежде всего наладить и испытать схему на постоянных напряжениях с построением характеристического семейства. Основываясь на этом семействе, можно подобрать режимы схемы и определить допустимые пределы изменения перемножаемых величин.

Не проделав рекомендуемого испытания умножителя, например кольцевого модулятора, трудно даже представить себе, насколько характеристики устройства могут отличаться от желаемых.

Вместе с тем тщательная проверка и подбор элементов позволяют получить удовлетворительные результаты. В качестве примера приводятся данные, полученные Польшковским для германиевых детекторов ДГЦ-1. На рис. 12 представлены характеристики двух экземпляров в линейном масштабе, на рис. 13 — те же характеристики в квадратичном масштабе, а на рис. 14 — характеристики умножения для кольцевой схемы ($R=100$; $R_2=1300$ см; $U_1=0,5$ в). Рис. 13 показывает, что характеристики детекторов квадратичны с достаточным приближением до напряжения около 0,6 в. На рис. 14 ясно видна область удовлетворительного умножения.

НЕЛИНЕЙНЫЕ И ПАРАМЕТРИЧЕСКИЕ ЯВЛЕНИЯ В РАДИОТЕХНИКЕ

ВВЕДЕНИЕ

§ 1. Характеристика предмета

Эта книга посвящена нелинейным и параметрическим явлениям, используемым в радиотехнике.

Такая терминология относится к физическим особенностям изучаемых явлений, особенностям, находящим отражение в математических уравнениях, которыми мы описываем поведение тех или иных радиотехнических устройств.

Явления, с которыми нам приходится иметь дело в радиотехнике, описываются математическими уравнениями следующих трех видов:

1. Линейные уравнения с постоянными коэффициентами.
2. Линейные уравнения с переменными коэффициентами.
3. Нелинейные уравнения.

Свойства и способы решения этих трех видов уравнений существенно различны; этим и обусловлено традиционное деление курса радиотехники на «линейную» и «нелинейную» части.

Линейные уравнения с постоянными коэффициентами — самый простой тип уравнений, теория которых детально разработана и для решения которых существуют мощные общие методы (например, методы интегралов Фурье и Дюамеля). Эти общие методы основаны на применении принципа наложения (суперпозиции). Нужно заметить, что применимость принципа наложения служит *определением* линейной системы, т. е. системы, описываемой линейным уравнением как с постоянными, так и с переменными коэффициентами.

Общий вид однородного линейного уравнения с постоянными коэффициентами таков:

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = 0,$$

где $a_n, a_{n-1}, \dots, a_1, a_0$ — постоянные величины, откуда и название этого типа уравнений.

Линейные уравнения с переменными коэффициентами характеризуются тем, что коэффициенты уравнения *зависят от аргумента* t , т. е. являются заданными функциями времени. Теория этих уравнений более сложна и менее разработана. Этим главным образом оправдывается отнесение явлений, описываемых уравнениями с переменными коэффициентами, к «нелинейной» радиотехнике. С физической точки зрения уравнения с переменными коэффициентами описывают поведение систем, *параметры* которых так или иначе изменяются во времени; такие системы и происходящие в них явления будем кратко называть *параметрическими*. Примеры уравнений с переменными коэффициентами.

$$\frac{d^2y}{dt^2} + ty = 0,$$

$$\frac{d^2y}{dt^2} + (a + 2q \cos 2t) y = 0.$$

Последнее уравнение носит название уравнения Матё; оно играет большую роль в теоретической радиотехнике.

Наконец, нелинейные уравнения отличаются тем, что их коэффициенты *зависят от функции* или ее производных. Это наиболее трудный для решения тип уравнений; некоторые общие приемы будут рассмотрены ниже применительно к тем или иным радиотехническим проблемам.

Примеры нелинейных уравнений.

$$\frac{d^2y}{dt^2} + ay^2 = 0,$$

$$\frac{d^2y}{dt^2} + \frac{dy}{dt} y = 0,$$

$$\frac{d^2y}{dt^2} + f\left(\frac{dy}{dt}\right) + y = 0.$$

Последнее уравнение имеет довольно общий характер; оно встретится нам в теории лампового генератора.

Математическая сторона дела интересует нас потому, что при помощи математического анализа мы можем провести количественное исследование различных радиотехнических устройств и, в конечном счете, построить систему *расчета* этих устройств. Что же касается физической и технической стороны дела, то нужно сказать, что к нелинейным и параметрическим явлениям относятся важнейшие для радиотехники явления и процессы. Вот краткий перечень этих процессов: генерирование электрических колебаний, умножение и деление частоты, выпрямление и детектирование, модуляция и преобразование частоты.

Итак, многие важные радиотехнические процессы относятся к нелинейным и параметрическим явлениям. Они описываются нелинейными уравнениями или уравнениями с переменными коэффициентами. Эти уравнения и составляют математический аппарат «нелинейной» радиотехники.

§ 2. Обзор методов нелинейной теории

Как указывалось в предыдущем параграфе, принцип наложения, составляющий основу линейной теории, неприменим для исследования явлений в нелинейных системах. Поясним коротко это положение.

Начнем со случаев линейной системы, описываемой линейным алгебраическим уравнением

$$y = ax.$$

Здесь x и y — функции времени; x означает воздействие на систему, а y — ее отклик. Принцип наложения состоит в том, что если на систему оказывается воздействие

$$x = x_1 + x_2,$$

то отклик системы будет

$$y = ax = a(x_1 + x_2) = y_1 + y_2,$$

т. е. действия обоих слагаемых x_1 и x_2 независимы. Это можно выразить краткой формулой: действие суммы причин равно сумме действий каждой причины, отдельно взятой. Так, например, ток в линейной цепи при действии нескольких э. д. с. определяется суммой токов, порождаемых каждой из э. д. с. в отсутствие других.

Принцип наложения лежит в основе общих методов решения многочисленных задач линейной теории. Мы настолько привыкаем к постоянному применению этого принципа, что, переходя к исследованию нелинейных проблем, нужно специально напомнить о неприменимости принципа наложения к нелинейным явлениям. Упустив это обстоятельство из вида, можно наделать грубых ошибок.

Поясним это на простом примере. Пусть дана нелинейная, а именно квадратичная зависимость

$$y = ax^2.$$

Пусть, далее,

$$x = x_1 + x_2.$$

Тогда

$$y_1 = ax_1^2, \quad y_2 = ax_2^2.$$

Но

$$y = a(x_1 + x_2)^2 = ax_1^2 + ax_2^2 + 2ax_1x_2 \neq y_1 + y_2,$$

т. е. квадрат суммы не равен сумме квадратов. Поэтому, например, при действии в нелинейной цепи суммы нескольких э. д. с. мы уже не можем свести решение задачи к суммированию токов, порождаемых каждой э. д. с. в отдельности; приходится учитывать сразу все составляющие э. д. с.

Понятно, что положение еще более усложнится, когда мы перейдем от линейных алгебраических соотношений к нелинейным дифференциальным уравнениям.

Вынужденный отказ от принципа наложения наносит тяжелый урон возможностям исследования. К сожалению, таких мощных и универсальных методов исследования, какими располагает линейная теория, нелинейная теория пока не знает. Арсенал орудий исследования нелинейных проблем составляет из ряда специализированных приемов и методов, имеющих различную силу и область применения. Ниже дана краткая предварительная характеристика этих методов; в дальнейшем нам предстоит познакомиться с ними в их действии применительно к различным радиотехническим проблемам.

1. Алгебраическая задача нахождения $y(t)$ по заданному $x(t)$ и по заданной нелинейной характеристике

$$y = f(x)$$

может быть всегда решена с любой степенью точности путем графического построения.

2. Та же задача может быть решена и аналитически, т. е. может быть найдена функция

$$y(t) = f[x(t)],$$

если дано аналитическое выражение характеристики системы

$$y = f(x).$$

При этих условиях возможно (но далеко не всегда) решение и дифференциального уравнения.

3. Нужно заметить, что очень часто характеристика системы известна нам из опыта, т. е. задана нам графиком или табличными данными. Нахождение аналитического выражения, достаточно точно представляющего действительную зависимость, является, таким образом, очень важной операцией в нелинейной теории. Эту операцию называют *аппроксимацией* нелинейной характеристики.

Аппроксимация должна по возможности удовлетворять следующим требованиям: а) аналитическое выражение должно представлять аппроксимируемую зависимость с достаточной точностью, б) оно должно допускать выполнение над ним требуемых математических операций и в) оно должно быть как можно более кратким. Существующие методы аппроксимации всем этим пожеланиям одинаково хорошо, к сожалению, не отвечают. Применяются следующие главнейшие методы: а) аппроксимация степенным многочленом, б) аппроксимация ломаной линией, в) аппроксимация трансцендентными функциями.

Первый метод состоит в том, что функция

$$y = f(x)$$

представляется в виде

$$y = a_0 + a_1x + a_2x^2 + \dots + a_nx^n = \sum_{k=0}^n a_kx^k.$$

Число членов степенного многочлена определяется требуемой точностью. Во всяком случае приближение должно быть достаточным, чтобы правильно отобразить хотя бы основные качественные черты явления. Это замечание относится, конечно, ко всякой аппроксимации.

Аппроксимация ломаной линией, называемая также кусочно-линейной аппроксимацией, состоит в том, что график действительной зависимости заменяется ломаной, тем лучше представляющей характеристику, чем больше берется точек излома. Аппроксимация ломаной особенно удобна для представления характеристик выпрямителей, работа лампы «с отсечкой» и т. п. Этот вид аппроксимации интересен тем, что позволяет воспользоваться специальным методом решения нелинейного дифференциального уравнения — так называемым «методом шивания». Сущность этого метода состоит в том, что поведение нелинейной системы представляется различными *линейными* уравнениями для различных интервалов значений переменной величины, в пределах которых нелинейная характеристика представлена отрезком прямой. Решения линейных уравнений нужно затем «сшить» на границах интервалов, т. е. приравнять конечные значения для одного интервала начальным значениям для другого.

Что касается аппроксимации трансцендентными функциями, например, экспоненциальными или гиперболическими, то этот метод позволяет получить очень компактные и в то же время дающие очень хорошее приближение аппроксимирующие выражения, но, к сожалению, дальнейшая аналитическая обработка этих выражений сильно затруднена.

4. Большую роль в нелинейной радиотехнике играют так называемые «квазилинейные» методы. Сущность этих методов сводится вкратце к тому, что уравнение нелинейной системы заменяется линейным по форме уравнением; однако в качестве коэф-

фициентов этого уравнения вводятся некоторые *средние* параметры, определенные с учетом имеющейся в действительности нелинейности. Поэтому средние значения параметров оказываются зависящими от величин, характеризующих режим системы. Квазилинейные методы особенно удобны для исследования *установившихся режимов* нелинейных систем, когда мы имеем дело с колебаниями, близкими к синусоидальным.

5. При исследовании *режимов установления* хорошую службу служит также относящийся к квазилинейным *метод медленно меняющихся амплитуд*. Общая идея этого метода состоит в том, что, исследуя вопросы установления колебаний в том или ином радиотехническом устройстве, мы интересуемся законом изменения амплитуды и частоты, предполагая, что эти величины изменяются медленно. Это обстоятельство позволяет нам заметно упростить задачу. При известных предположениях (подробнее об этом говорится ниже) мы получаем обычно уравнение порядка на единицу ниже исходного. Метод медленно меняющихся амплитуд применяется главным образом в случаях, когда наперед известно, что форма колебаний близка к синусоидальной.

6. Наиболее общим методом является, пожалуй, графический метод построения интегральных кривых данного нелинейного дифференциального уравнения. Это метод известен под названием *метода построения поля направлений*. Сущность и техника применения этого метода будут изложены ниже; пока заметим лишь, что метод позволяет решать нелинейные задачи наиболее общего характера; он годится как для установившихся, так и для неустановившихся режимов, как для синусоидальных, так и для резко несинусоидальных — так называемых релаксационных — колебаний. Удобный вариант метода, сводящий нахождение наклона интегральной кривой к чисто геометрическому построению, известен под названием *построения Льенара*.

7. Наконец, общим методом, позволяющим (иногда с исключительной простотой) сделать важные заключения общезначимого характера, является *энергетический метод*. Сущность этого метода состоит в рассмотрении энергетического баланса нелинейной системы, т. е. в прямом применении закона сохранения энергии. Этот метод не всегда позволяет разъяснить детали, но дает нам надежные общие соотношения, относящиеся к существу явления.

Таков вкратце перечень орудий исследования, которыми нам предстоит воспользоваться в дальнейшем.

НЕЛИНЕЙНЫЕ ЦЕПИ И ОСНОВНЫЕ НЕЛИНЕЙНЫЕ ПРОЦЕССЫ

В этой главе рассматриваются те вопросы нелинейной теории, математическое описание которых может быть (хотя бы в первом приближении) сведено к не дифференциальным (т. е. алгебраическим и трансцендентным) нелинейным уравнениям. Вопросы же, для рассмотрения которых требуется решение нелинейных *дифференциальных* уравнений и являющиеся, очевидно, более сложными, разбираются в последующих главах.

§ 3. Нелинейная цепь

Простейшая нелинейная цепь содержит, кроме источника э. д. с., нелинейное сопротивление, т. е. элемент, сопротивление которого не является постоянной величиной, а зависит от тока. Предполагается, что реактивных элементов, т. е. емкостей и индуктивностей, в цепи нет. Поэтому уравнение такой нелинейной цепи будет уравнением алгебраическим (а не дифференциальным).

Свойства нелинейного сопротивления выражаются непосредственно зависимостью тока I , протекающего через сопротивление, от напряжения U на его концах. Эта зависимость носит название вольтамперной характеристики и может быть записана как

$$I = f(U).$$

Задавшись напряжением в функции времени, можем найти ток также как функцию времени. Это делается или графически, или аналитически. Графическое построение показано на рис. 1, где в левой части изображены: сверху — характеристика нелинейного сопротивления $I=f(U)$, снизу — $U=U(t)$ (выбрано синусоидальное напряжение). График $I=I(t)$ получается построением, аналогичным построению третьей проекции по двум данным (на рис. 1 — справа). Построение ясно из чертежа. Заметим сразу, что *форма* кривой тока отличается от формы кривой напряжения. Это типично для нелинейной цепи и имеет очень большое значение. Формы кривых напряжения и тока были бы подобны только в случае линейной цепи; в этом случае вольтамперная характеристика представлялась бы прямой линией.

На рис. 1 на синусоидальное напряжение наложено еще постоянное напряжение U_0 . Оно определяет постоянную составляющую тока I_0 в отсутствие колебаний. Точка на характеристике с координатами (U_0, I_0) называется *рабочей точкой*. Изменяя U_0 , можно перемещать рабочую точку по характеристике. Выбор рабочей точки определяется желаемым результатом, как будет видно из дальнейшего. Отметим, что постоянная составляющая тока при

колебаниях (I' на рис. 1) отличается от тока покоя I_0 . Это также типично для нелинейной цепи.

Легко решается графически вопрос о распределении напряжений в цепи, в которой последовательно включены обычное (линейное, т. е. постоянное) сопротивление R и нелинейное сопротивление, определенное своей вольтамперной характеристикой. Соответствующее построение показано на рис. 2.

Вольтамперная характеристика нелинейного сопротивления построена, как обычно, в координатах (I, U). Из точки на оси абсцисс, соответствующей приложенному к цепи напряжению, проведена прямая под углом α к вертикальной оси. Эта прямая изображает вольтамперную характеристику линейного сопротивления R ; ее наклон ($\operatorname{tg} \alpha$) непосредственно выражает величину сопротивления и численно равен R .

Координаты точек пересечения прямой с вольтамперной характеристикой определяют, во-первых, ток в цепи, а во-вторых, распределение напряжений на элементах цепи, как показано на рис. 2 (U_R означает падение напряжения на линейном сопротивлении, а U' — на нелинейном).

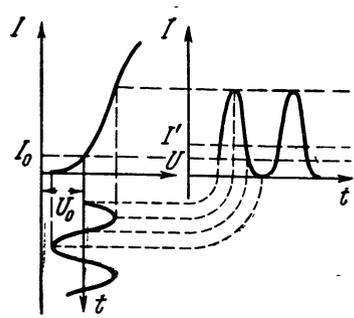


Рис. 1

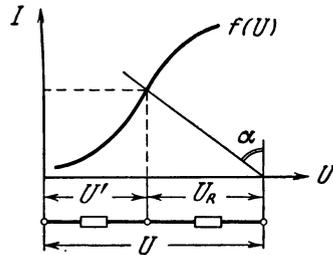


Рис. 2

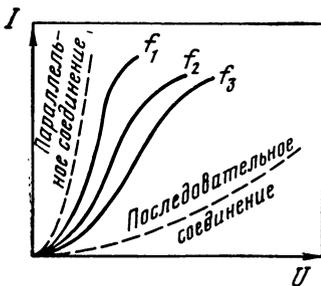


Рис. 3

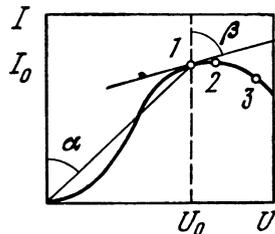


Рис. 4

Аналогично решается задача о распределении напряжений в случае соединения двух нелинейных сопротивлений.

Если число нелинейных сопротивлений, включенных в цепь, больше двух, то построение целесообразно изменить. На рис. 3 показаны три вольтамперные характеристики. Если нелинейные сопротивления включены параллельно, то вольтамперная характеристика результирующего сопротивления получается поординатным сложением характеристик. Если же сопротивления включены последовательно, то результирующая характеристика получается поабсциссным сложением всех характеристик.

Теперь нам нужно заняться вопросом об *определении* нелинейного сопротивления. Простое линейное сопротивление определяется как отношение падения напряжения на нем к протекающему через него току, т. е. $R=U/I$.

Это — постоянная величина, и определение не требует никаких дополнений. Но нелинейное сопротивление — величина переменная, и ее можно определить по-разному. На рис. 4 изображена вольтамперная характеристика нелинейного сопротивления. Если подать на сопротивление постоянное напряжение U_0 , то получим постоянный ток I_0 ; отношение этих величин можно назвать *сопротивлением постоянному току*

$$R_+ = \frac{U_0}{I_0} (= \operatorname{tg} \alpha).$$

Графически R_+ определяется наклоном секущей; ясно, что эта величина зависит от приложенного напряжения U_0 .

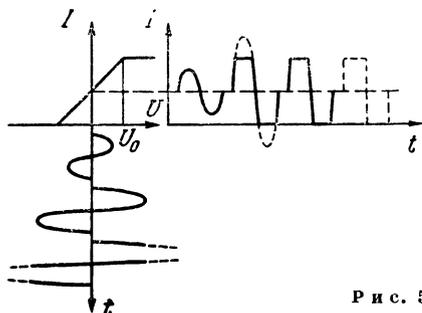
Положим теперь, что на постоянное напряжение наложено небольшое переменное напряжение ΔU . Приращение напряжения ΔU вызывает приращение тока ΔI . Вводят определение *сопротивления переменному току* как предела

$$R_{\sim} = \lim_{\Delta U \rightarrow 0} \frac{\Delta U}{\Delta I} = \frac{dU}{dI} = \frac{1}{f'(U)} (= \operatorname{tg} \beta).$$

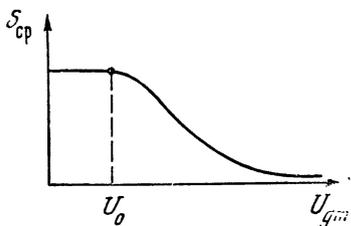
Определенную таким образом величину называют также дифференциальным сопротивлением. Графически сопротивление переменному току выражается наклоном касательной к вольтамперной характеристике в рабочей точке, как показано на рис. 4. Ясно, что сопротивления постоянному и переменному току не равны (они равны только для линейного сопротивления). Более того, сопротивление постоянному току всегда конечно и положительно, тогда как сопротивление переменному току может быть равно нулю, бесконечности (точка 2 на рис. 4, в которой касательная к кривой $f(U)$ горизонтальна, т. е. конечному приращению напряжения соответствует нулевое приращение тока) и даже становится отрицательным (точка 3 на рис. 4, в которой положительному приращению напряжения соответствует отрицательное приращение тока). Понятие отрицательного сопротивления играет в радиотехнике очень большую роль. Как мы увидим дальше,

отрицательное сопротивление может рассматриваться как источник энергии, тогда как положительное сопротивление является ее потребителем.

Кроме названных сопротивлений постоянному и переменному току, вводится еще одно важное определение, а именно определение *среднего сопротивления*. Это определение учитывает нелинейность особым образом. Положим, что через нелинейное сопротивление протекает синусоидальный ток. Вследствие нелинейности напряжение на сопротивление будет несинусоидально. Однако напряжение будет периодическим, и основная



Р и с. 5



Р и с. 6

частота его будет равна частоте тока. Если взять теперь амплитуду первой гармоники напряжения и отнести ее к амплитуде тока, то это отношение и определит то, что мы называем средним сопротивлением. Таким образом, определение среднего сопротивления связано с нахождением первой гармоники периодического отклика нелинейной системы на синусоидальное воздействие. К сожалению, среднее сопротивление не может быть определено графически по вольтамперной характеристике. Его находят чаще всего аналитически, располагая аппроксимирующим характеристикой выражением.

Все три приведенные выше определения могут применяться не только к сопротивлению, но и к другим параметрам, постоянным в линейной системе и переменным в нелинейной. Так, например, можно говорить о проводимости для постоянного тока, проводимости для переменного тока и средней проводимости. В радиотехнике широко применяется понятие *средней крутизны* электронной лампы. Эта величина определяется как отношение амплитуды первой гармоники анодного тока к амплитуде синусоидального напряжения на управляющей сетке. Чтобы составить себе представление о зависимости средней крутизны от амплитуды управляющего напряжения, рассмотрим рис. 5, где представлена характеристика триода, аппроксимированная ломаной линией. Рабочая точка выбрана посередине наклонного участка. При увеличении амплитуды управляющего напряжения крутизна остается постоянной, пока сеточное напряжение не выходит за пре-

делу прямолинейного участка характеристики. При этом анодный ток (переменная составляющая) синусоидален. При дальнейшем увеличении амплитуды напряжения на сетке синусоида анодного тока усекается сверху и снизу. В пределе мы получим в анодной цепи импульсы прямоугольной формы: амплитуда первой гармоники анодного тока остается неизменной при неограниченном росте амплитуды напряжения на сетке U_{gm} . Значит средняя крутизна S_{cp} стремится к нулю, как показано на рис. 6. В действительности характеристика лампы представляется кривой линией, а потому средняя крутизна изменяется плавно. В дальнейшем нам предстоит вычислить среднюю крутизну при степенной аппроксимации характеристики триода.

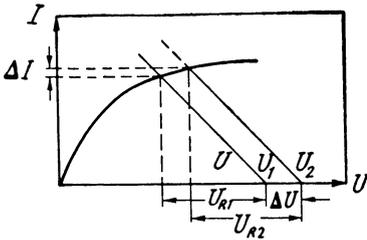
§ 4. Стабилизация тока и напряжения

Простейшим применением нелинейных сопротивлений является стабилизация тока и напряжения, широко применяемая в источниках питания радиотехнических устройств.

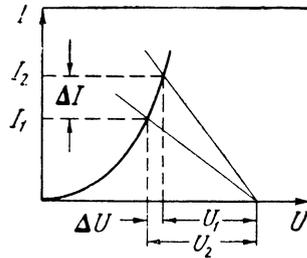
Задача стабилизации состоит в уменьшении колебаний тока и напряжения в условиях изменяющейся нагрузки или непостоянства напряжения сети. Если бы все сопротивления в цепи были линейны, то относительные изменения тока и напряжения были бы пропорциональны друг другу. Нарушить эту пропорциональность — в чем и состоит задача — можно только путем применения нелинейных элементов.

Положим, что мы желаем стабилизировать ток накала лампы, т. е. поддержать по возможности его постоянство при изменениях напряжения источника. Для этого нужно последовательно с нитью лампы включить нелинейное сопротивление, величина которого резко возрастает с увеличением тока и напряжения. При таких условиях избыток напряжения будет почти полностью падать на нелинейном сопротивлении, а напряжение на нити, т. е. напряжение накала, а, значит, и ток накала, будут изменяться незначительно. Вольтамперная характеристика стабилизатора должна, очевидно, иметь вид, показанный на рис. 7. Применяя построение рис. 2, можно проследить, как изменяется напряжение накала U_R при изменении питающего напряжения. Как видим, для стабилизации требуется, чтобы проводимость для переменного тока (т. е. отношение $\Delta I/\Delta U$) была по возможности малой, т. е. сопротивление переменному току должно быть по возможности велико. Сопротивление переменному току должно быть во всяком случае велико по сравнению с последовательно включенным линейным сопротивлением; в противном случае стабилизатор не окажет действия. Примером стабилизатора может служить *барретор*, представляющий собою железную нить, помещенную в наполненный водородом баллон. Требуемый вид вольтамперной характеристики получается в барреторе за счет того, что сопротивление металлов растет с повышением температуры, которая в свою очередь за-

висит от тока. Железо выбрано потому, что оно обладает относительно большим температурным коэффициентом сопротивления. В качестве стабилизатора небольших токов можно с успехом применять диод в режиме, близком к насыщению.



Р и с. 7



Р и с. 8

Стабилизатор напряжения должен, очевидно, обладать обратными свойствами. Его назначение состоит в поддержании постоянства напряжения при изменениях тока. Значит, отношение $\Delta U/\Delta I$, т. е. сопротивление переменному току, должно быть по возможности малым. Желательная характеристика стабилизатора показана на рис. 8, где дано построение, показывающее, что большим изменениям тока соответствуют малые изменения напряжения на нагрузке, включенной последовательно со стабилизатором.

Для стабилизации напряжения применяются главным образом газоразрядные лампы. Применяются также электронные схемы стабилизации.

§ 5. Преобразование спектра в нелинейной цепи

В § 3 уже отмечалось, что синусоидальное напряжение вызывает в нелинейной цепи несинусоидальный ток. Вообще форма тока отличается от формы напряжения. Обратно, если задан синусоидальный ток, то получается несинусоидальное напряжение.

Если рассмотреть вопрос со спектральной точки зрения, то оказывается, что спектр отклика нелинейной системы всегда отличается по составу от спектра воздействия; обычно спектр отклика богаче. Это общее положение, широко используемое в радиотехнике, мы выясним на нескольких простых примерах.

Пусть на нелинейное сопротивление с характеристикой $I=f(U)$ действует синусоидальное напряжение $U=U_m \sin \omega t$.

Найдем ток, для чего развернем характеристику в степенной ряд

$$I = \sum_{k=0}^{\infty} a_k U^k = \sum_{k=0}^{\infty} a_k U_m^k \sin^k \omega t.$$

При возведении синуса в k -ю степень мы получаем ряд колебаний кратных частот, причем наивысшая кратность равна k . Так, квадрат синуса дает постоянную составляющую (нулевую частоту) и вторую гармонику; куб дает первую и третью гармоники, и так далее. Приводим для справок табличку степеней синуса и косинуса, которая нам пригодится и в дальнейшем (табл. 1):

Таблица 1

$\sin^2 x = \frac{1}{2} (1 - \cos 2x)$	$\cos^2 x = \frac{1}{2} (1 + \cos 2x)$
$\sin^3 x = \frac{1}{4} (3\sin x - \sin 3x)$	$\cos^3 x = \frac{1}{4} (3\cos x + \cos 3x)$
$\sin^4 x = \frac{1}{8} (3 - 4\cos 2x + \cos 4x)$	$\cos^4 x = \frac{1}{8} (3 + 4\cos 2x + \cos 4x)$
$\sin^5 x = \frac{1}{16} (10\sin x - 5\sin 3x + \sin 5x)$	$\cos^5 x = \frac{1}{16} (10\cos x + 5\cos 3x + \cos 5x)$

Из сказанного следует, что при воздействии синусоидального напряжения на нелинейное сопротивление мы получаем в *общем случае* ток с бесконечным спектром, содержащим все гармоники. В частном случае характеристика может представляться не бесконечным рядом, а конечным степенным многочленом. В этом случае и ток будет иметь ограниченный спектр, причем номер наивысшей гармоники будет определяться степенью многочлена. Так, например, при квадратичной характеристике вида

$$I = a_0 + a_1 U + a_2 U^2$$

в составе тока не будет гармоник выше второй.

Рассмотрим теперь случай, когда на нелинейное сопротивление действуют два напряжения с разными частотами, т. е. пусть

$$U = U_1 + U_2 = U_{m1} \sin \omega_1 t + U_{m2} \sin \omega_2 t.$$

Подставляя это выражение в формулу характеристики, получим для тока

$$I = f(U_1 + U_2) = \sum_{k=0}^{\infty} a_k (U_{m1} \sin \omega_1 t + U_{m2} \sin \omega_2 t)^k.$$

Возводя двучлен в скобках в k -ю степень и группируя затем члены суммы, можно убедиться, что в составе тока образуются частоты, выражаемые общим соотношением

$$n\omega_1 \pm m\omega_2,$$

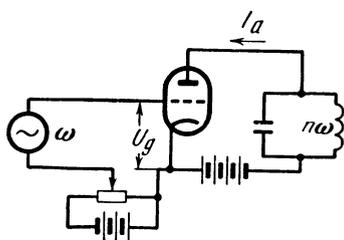
где m и n — любые целые числа, не исключая нуля. При $m=0$ мы получаем последовательность гармоник $n \omega_1$; при $n=0$ — гармоник $m \omega_2$. Частоты, отвечающие неравным нулю значениям m и n , называются *комбинационными частотами*. Они могут, очевидно, быть как выше, так и ниже исходных частот ω_1 и ω_2 .

§ 6. Умножение частоты

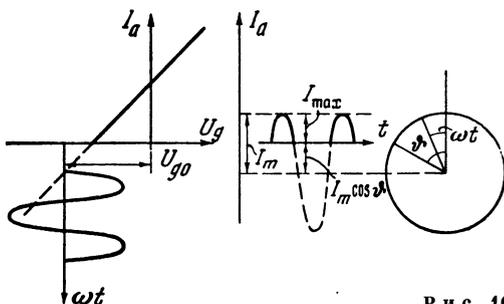
Примером использования обогащения спектра нелинейной системой может служить умножение частоты. Задача умножения частоты состоит в получении частоты, кратной исходной. Умножение частоты применяется в тех случаях, когда непосредственное генерирование требуемой частоты почему-либо неудобно.

Идея умножения крайне проста; она заключается в том, что напряжение исходной частоты действует в нелинейной цепи, в которой получается богатый гармониками ток. Гармоника желаемой кратности выделяется фильтром.

Возможная схема умножения показана на рис. 9. Синусоидальное напряжение с частотой ω приложено к сетке триода; контур в анодной цепи настроен на n -ю гармонику, которая и выделяется на контуре; остальные гармоники отфильтровываются, так как контур представляет для них малое сопротивление. На сетку



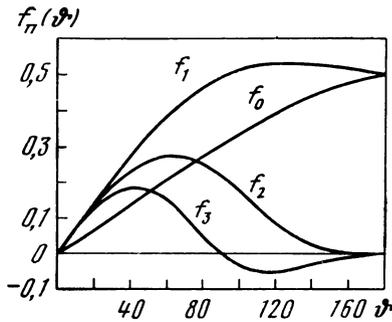
Р и с. 9



Р и с. 10

триода подается регулируемое отрицательное смещение, «достаточное, чтобы запереть триод. В результате этого происходит так называемая «работа с отсечкой».

Аппроксимируем характеристику триода ломаной линией, как показано на рис. 10; для наших целей это достаточно приближение. Выберем рабочую точку левее излома характеристики. При таком положении рабочей точки и синусоидальном напряжении на сетке в анодной цепи будут проходить лишь короткие импульсы тока. Остальная часть синусоиды (намеченная штрихом) будет отсечена — отсюда и название этого режима. Режим вполне характеризуется *углом отсечки*, т. е. тем значением аргумента ωt (отсчи-



Р и с. 11

таным от момента, когда анодный ток имеет максимальное значение), при котором ток обращается в нуль. Зная угол отсечки, можно разложить выражение для анодного тока в ряд Фурье. Анодный ток как функцию времени можно представить в виде

$$I_a = I_m (\cos \omega t - \cos \vartheta) \quad [I_a > 0]$$

(обозначения указаны на рис. 10). Для амплитуды n -й гармоники имеем по известной формуле

$$\begin{aligned} I_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} I_a(\omega t) \cos n\omega t d\omega t = \frac{2}{\pi} I_m \int_0^{\vartheta} \cos nx (\cos x - \cos \vartheta) dx = \\ &= \frac{1}{\pi} I_m \left[\frac{\sin(n+1)\vartheta}{n+1} + \frac{\sin(n-1)\vartheta}{n-1} - \frac{2 \cos \vartheta \sin n\vartheta}{n} \right] = I_m F_n(\vartheta). \end{aligned}$$

Практически может быть задано не I_m (определяемое амплитудой напряжения на сетке), а I_{\max} — пиковое значение анодного тока. Между этими величинами существует очевидная связь (см. рис. 10)

$$I_{\max} = I_m (1 - \cos \vartheta).$$

Выражая теперь амплитуду n -й гармоники анодного тока через пиковое значение анодного тока, имеем

$$I_n = I_{\max} \frac{F_n(\vartheta)}{1 - \cos \vartheta} = I_{\max} f_n(\vartheta).$$

Функции угла отсечки для различных n представлены графиками рис. 11. Эти графики показывают, как следует выбирать угол отсечки для получения наибольшей амплитуды желаемой гармоники. Так, например, при утроении частоты следует брать $\vartheta = 40^\circ$, при учетверении 30° и так далее.

§ 7. Линеаризация и нелинейные искажения

Нелинейные зависимости играют в радиотехнике двоякую роль. В целом ряде случаев нелинейность определяет самое существо процессов и используется для получения желаемого результата. Так обстоит дело, например, при умножении частоты, при детектировании и во многих других важных для радиотехники случаях. Но иногда нелинейность не только не требуется для осуществления того или иного процесса, но является вредной и нежелательной. Так обстоит, например, дело при усилении.

Типичное для нелинейной системы искажение формы и, следовательно, обогащение спектра колебаний необходимо для умножения частоты. Но при усилении задача состоит как раз в обратном — в сохранении по возможности исходной формы колебания. Все изменения формы колебания, происходящие вследствие нелинейности тех или иных элементов схемы, называются *нелинейными искажениями*. Необходимо стремиться к их уменьшению, т. е. к устранению влияния нелинейности.

Строго линейных систем не существует; всякая реальная система нелинейна. Таким образом, линейными мы называем системы, нелинейностью которых можно при данных обстоятельствах пренебречь. Возможность пренебречь нелинейностью зависит не только от характеристики системы, но и от режима работы. Так, например, при усилении стремятся расположить рабочую точку посередине прямолинейного участка характеристики лампы, а амплитуду управляющего напряжения ограничивают так, чтобы не выходить за пределы прямолинейного участка. В действительности характеристика всегда кривая; говоря о прямолинейном участке, мы подразумеваем участок с достаточно малой кривизной.

Замена нелинейной зависимости приближенной линейной, а также и меры, направленные к улучшению приближения, — называют вообще *линеаризацией*.

Понятие о линеаризации мы поясним сперва на частном примере. Пусть дана квадратичная характеристика $I = aU^2$. Эта нелинейная зависимость никак не может быть представлена приближенной линейной. Но положим, что напряжение U состоит из постоянной слагающей U_0 и переменной слагающей ΔU , которая нас и будет интересовать. Тогда

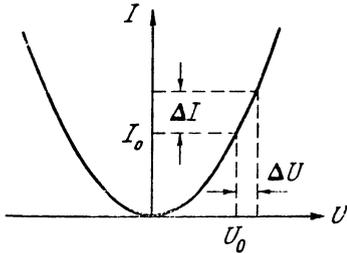
$$I = aU^2 = a(U_0 + \Delta U)^2 = a(U_0^2 + 2U_0\Delta U + \Delta U^2).$$

Как видим, ток имеет три слагающие: постоянную составляющую $I_0 = aU_0^2$; линейную составляющую $I_1 = 2aU_0\Delta U$ и квадратичную составляющую $I_2 = a\Delta U^2$.

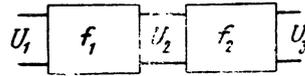
Постоянная составляющая нас обычно не интересует. Приращение тока запишем в виде

$$\Delta I = I_1 + I_2 = 2aU_0\Delta U + a\Delta U^2 = 2aU_0\Delta U \left(1 + \frac{1}{2} \cdot \frac{\Delta U}{U_0}\right).$$

Мы вправе отбросить второе слагаемое при условии $\Delta U/U_0 \ll 1$, т. е. линеаризация возможна с тем лучшим приближением, чем меньше ΔU и чем больше U_0 . Этот результат поясняется графиком рис. 12, из которого видно, что увеличение U_0 переводит рабочую точку на участок параболы с меньшей кривизной. Одновременно возрастает крутизна.



Р и с. 12



Р и с. 13

Уменьшение ΔU позволяет линеаризовать любую нелинейную зависимость¹. Чтобы показать это, положим, что задана нелинейная характеристика общего вида

$$I = f(U).$$

Разложим выражение для тока в ряд Тейлора около точки U_0

$$I = f(U_0) + f'(U_0)\Delta U + f''(U_0)\frac{\Delta U^2}{2} + f'''(U_0)\frac{\Delta U^3}{6} + \dots$$

Если ΔU — достаточно малая величина, то членами высших порядков можно пренебречь и записать линеаризованное выражение для тока в виде

$$I \cong f(U_0) + f'(U_0)\Delta U.$$

Постоянная составляющая определяется ординатой характеристики в рабочей точке, а крутизна — производной, т. е. наклоном касательной к характеристике в той же точке.

Может оказаться, что для данной характеристики $f(U)$ касательная к характеристике в рабочей точке горизонтальна, т. е. $f'(U_0) = 0$. В таком случае нужно сместить рабочую точку, изменив U_0 . Так, например, в только что рассмотренном случае квадратичной зависимости мы имели $f(U) = aU^2$, $f'(U_0) = 2aU_0$ и если бы мы взяли $U_0 = 0$, то получили бы $f'(0) = 0$.

Нелинейные искажения в почти линейной системе нужно каким-то образом оценивать и измерять. Для этого существует несколько методов; все они основаны на том, что нелинейность вызы-

¹ Однако существуют системы, линеаризованное описание явлений в которых лишено смысла. Эти исключительные случаи нам не встретятся.

вает обогащение спектра, которое не происходит в строго линейной системе.

Первый метод состоит в том, что на вход системы задается синусоидальное воздействие, а на выходе системы определяется действующее значение всех гармоник, кроме первой. Отношение этой величины к действующему значению первой гармоники (основной частоты) называется коэффициентом гармоник

$$k = \frac{\sqrt{I_2^2 + I_3^2 + \dots}}{I_1},$$

где через I_k обозначены амплитуды гармоник. Считают, что нелинейные искажения практически не ощущаются, если коэффициент гармоник не превосходит 1%. Техническими условиями на аппаратуру задается обычно коэффициент гармоник не свыше 5%.

Другой метод состоит в подаче на вход системы одновременно двух синусоидальных напряжений с различными частотами. При этом, как было показано в § 5, образуется богатый спектр комбинационных частот. Коэффициент, характеризующий нелинейные искажения, определяется так же, как и коэффициент гармоник.

В недавнее время предложен новый способ определения нелинейных искажений. Он основан на измерении суммарной мощности в некоторой полосе частот при подаче на вход воздействия со сплошным спектром, из которого предварительно вырезана фильтром та же самая полоса частот. Если бы система была линейна, то частоты, вырезанные на входе, отсутствовали бы и на выходе. В результате же нелинейности в пределах вырезанной полосы появляются комбинационные частоты и гармоники — так называемый продукт нелинейности. Этот продукт и измеряется. Новый метод ближе соответствует действительным условиям работы аппаратуры и дает более правильную оценку нелинейных искажений.

Следует еще указать, что нелинейные искажения можно *корректировать* аналогично тому, как корректируются частотные искажения. Положим, что имеется нелинейный четырехполюсник с характеристикой $U_2 = f_1(U_1)$. Добавим к этому четырехполюснику второй, служащий в качестве корректора (рис. 13). Его характеристика $U_3 = f_2(U_2)$ должна быть подобрана так, чтобы нелинейные искажения были устранены, т. е. чтобы было $U_3 = AU_1$. Это условие дает

$$\frac{U_3}{U_1} = \frac{U_3}{U_2} \cdot \frac{U_2}{U_1} = \frac{f_2[f_1(U_1)]}{U_1} = A$$

или

$$f_2[f_1(U_1)] = AU_1,$$

т. е. функция f_2 должна быть обратной¹ функции f_1 . Например, если $f_1(U_1) = aU_1^2$, то должно быть $f_2(U_2) = b\sqrt{U_2}$. При таких соотношениях мы получим $U_3 = b\sqrt{U_2} = b\sqrt{aU_1} = AU_1$.

§ 8. Выпрямление

Объектом выпрямления является переменный ток; цель выпрямления состоит в получении постоянного тока (например, для питания радиотехнических устройств).

Простейшая схема выпрямления показана на рис. 14. В схему включен нелинейный элемент, относительно которого предпола-

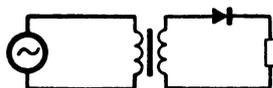


Рис. 14

гается, что он обладает конечной проводимостью в одном направлении и нулевой (или очень малой) проводимостью в другом. Направление, в котором выпрямитель пропускает ток, называется прямым, противоположное — обратным. Соответственно обозначаются и проводимости или сопротивления. Отношение прямой проводимости к обратной называется коэффициентом выпрямления.

Реальная и идеализированная характеристики выпрямителя показаны на рис. 15, а и б. Если в схеме рис. 14 при характеристике рис. 15, б действует синусоидальное напряжение, то в цепи протекают импульсы тока, отвечающие положительным полупериодам напряжения (рис. 16). Это есть так называемое однополупериодное

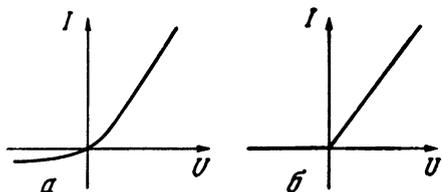


Рис. 15

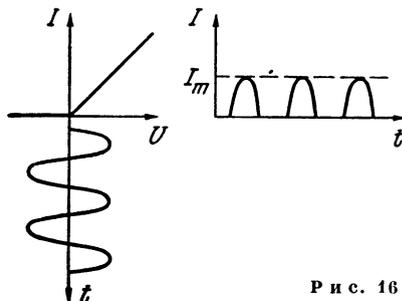


Рис. 16

или однократное выпрямление. В цепи протекает пульсирующий, т. е. не меняющий направления, но изменяющийся по величине ток. Он содержит требуемую постоянную составляющую, но, кроме того, также и переменную составляющую со сложным спектром.

¹ Напомним определение обратной функции. Если $y=f(x)$, то, решая это уравнение относительно x , получим $x=\varphi(y)$. Функция φ обратна функции f . При этом $\varphi[f(x)]/x=f[\varphi(y)]/y=1$.

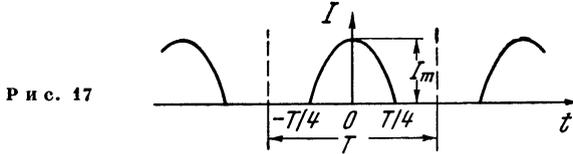
Постоянная составляющая равна

$$I_0 = \frac{1}{\pi} I_m.$$

Спектр переменной составляющей определим, разложив в ряд Фурье выражение для тока. Выражение для тока имеет вид

$$I = I_m \cos \omega t \left[-\frac{T}{4} < t < \frac{T}{4}, \text{ или } -\frac{\pi}{2} < \omega t < \frac{\pi}{2} \right].$$

На протяжении остальной доли периода ток равен нулю (рис. 17).



Таким образом,

$$I_n = \frac{1}{\pi} I_m \int_{-\pi/2}^{\pi/2} \cos x \cos nx dx = \frac{1}{\pi} I_m \left[\frac{\sin(n+1)x}{n+1} + \frac{\sin(n-1)x}{n-1} \right] \Big|_0^{\pi/2} =$$

$$\begin{aligned} & \frac{1}{2} I_m && \text{при } n = 1, \\ & = \frac{2}{\pi} I_m \frac{1}{n^2 - 1} && \text{при } n \text{ четном,} \\ & 0 && \text{при } n \text{ нечетном.} \end{aligned}$$

Итак, в составе выпрямленного тока имеется первая гармоника с амплитудой

$$I_1 = \frac{1}{2} I_m,$$

а кроме того, все четные гармоники, амплитуды которых равны

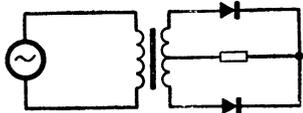
$$I_2 = \frac{2}{3\pi} I_m, \quad I_4 = \frac{2}{15\pi} I_m, \dots$$

Гораздо чаще применяется двухполупериодное (двухтактное) выпрямление. Оно производится схемой рис. 18. Действие выпрямителей в обеих ветвях схемы таково, что результирующую характеристику можно представить, как показано на рис. 19. Постоянная составляющая при двухполупериодном выпрямлении, очевидно, вдвое больше, чем при однополупериодном; в этом состоит одно из преимуществ двухполупериодного выпрямления. Что касается спектра тока, то первая гармоника в нем отсутствует. Это и понятно, так как основная частота выпрямленного тока (при условии симметрии обеих ветвей схемы) вдвое больше частоты питающего напряжения, и, следовательно, наименьшая частота в составе тока отвечает второй гармонике напряжения. Спектраль-

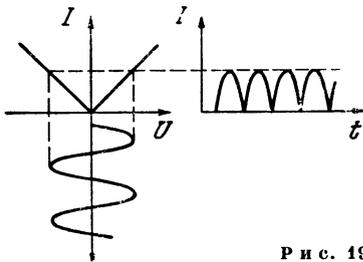
ный состав выпрямленного тока таков:

$$I_0 = \frac{2}{\pi} I_m, \quad I_2 = \frac{4}{3\pi} I_m, \quad I_4 = \frac{4}{15\pi} I_m.$$

Для очистки выпрямленного тока от переменной составляющей применяются обычные фильтры нижних частот. Типичная схема выпрямителя с двойным диодом изображена на рис. 20. Здесь выясняется еще одно преимущество двухполупериодного выпрямления. Дело в том, что основная частота переменной составляющей



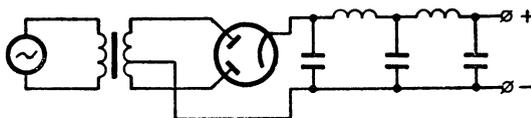
Р и с. 18



Р и с. 19

при двухполупериодном выпрямлении вдвое выше, чем при однополупериодном. Это позволяет соответственно повысить граничную частоту фильтра, а следовательно, облегчить его конструкцию, вдвое уменьшив как индуктивности, так и емкости. Фильтрация выпрямленного тока еще более облегчается при выпрямлении многофазного тока (трех- или шестифазного), так как при этом уменьшается величина пульсаций и возрастает их частота. Пульсации тем меньше, чем больше число фаз. Нужно пояснить, что мы не рассматриваем в этом параграфе работу выпрямителя в целом, включая фильтр. Дело в том, что схема выпрямителя с фильтром описывается нелинейным *дифференциальным* уравнением, решение которого представляет более сложную задачу. Мы вернемся к этому вопросу в § 31.

Следует еще добавить, что выпрямление часто применяется для измерения переменных напряжений и токов посредством электроизмерительных приборов постоянного тока (магнитоэлектрической



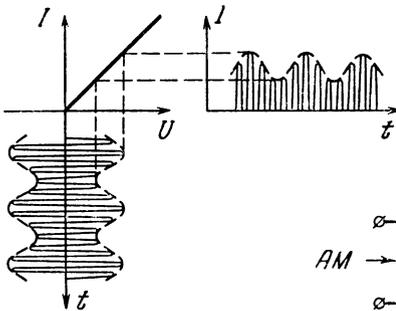
Р и с. 20

системы). Все современные многошкальные универсальные приборы постоянного и переменного тока являются детекторными приборами, т. е. снабжены выпрямительными элементами, что и позволяет пользоваться одним и тем же прибором для измерений как на постоянном, так и на переменном токе. Ламповый вольтметр — основной радиоизмерительный прибор — также представляет собою магнитоэлектрический прибор в комбинации с ламповой вы-

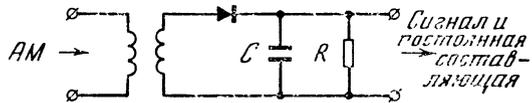
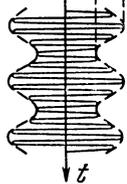
прямительной схемой. Однако требования к выпрямителю для измерительных целей несколько иные, чем требования к выпрямителю для целей питания. Дело в том, что измерительные приборы переменного тока должны измерять действующие значения измеряемых напряжений или токов, которые могут быть и несинусоидальными. А это условие выполняется лишь в том случае, если характеристика выпрямителя квадратична. За дальнейшими подробностями следует обратиться к курсу радиоизмерений.

§ 9. Детектирование

Объектом детектирования является модулированное колебание; цель детектирования состоит в извлечении из модулированного колебания сигнала, т. е. колебания низкой модулирующей частоты. В этом параграфе мы рассмотрим только детектирование колебаний, модулированных по амплитуде.



Р и с. 21



Р и с. 22

Процесс детектирования амплитудно-модулированных (АМ) колебаний имеет много общего с процессом выпрямления; различие состоит главным образом в объектах и желаемых результатах. Простейшая схема детектора содержит нелинейный элемент с такой же характеристикой, как и характеристика выпрямителя на рис. 15. Разберем сначала качественную картину явлений при детектировании. Пусть амплитудно-модулированное напряжение действует в цепи с нелинейным элементом — детектором, обладающим ломаной характеристикой (рис. 21). Ток в цепи детектора представляет собою импульсы высокой частоты с огибающей, воспроизводящей форму огибающей модулированного колебания. В составе тока получается несущая частота и ее гармоники. Кроме того, имеется постоянная составляющая, которая нас обычно не интересует. И, наконец, в составе тока появляется низкочастотная составляющая, которую мы стремимся выделить.

Нужно всячески подчеркнуть, что в составе модулированного колебания низкочастотной составляющей нет; как известно, например, при синусоидальной модуляции модулированное колеба-

ние состоит из трех синусоидальных колебаний: колебания несущей частоты и двух спутников, имеющих также высокие частоты.

Для выделения низкочастотной составляющей, появляющейся в результате детектирования, нужно отфильтровать высокие частоты и постоянную составляющую. Последняя устраняется обычно сама собою, так как не проходит через усилительный тракт, содержащий трансформаторы и переходные емкости. Что касается высоких частот, то они могут быть отфильтрованы простейшим звеном, состоящим из емкости и сопротивления. Схема детектора принимает вид, показанный на рис. 22. Она мало чем отличается — по виду — от схемы выпрямителя на рис. 20. Но схема рис. 22 выполняет другую задачу: она должна выделить низкую частоту сигнала, тогда как схема выпрямителя предназначена для устранения переменной составляющей и получения постоянного тока. Это различие определяет и различие в требованиях к фильтрам. Фильтрующее звено в схеме детектора на рис. 22 должно устранить высокую частоту, но сохранить низкую частоту. Для этого необходимо, чтобы сопротивление R было много больше емкостного сопротивления при высокой (несущей) частоте ω_0 и много меньше емкостного сопротивления при низкой частоте сигнала Ω . Это условие можно записать в виде неравенств

$$1/\omega_0 C \ll R \ll 1/\Omega C.$$

К тому же заключению можно прийти и иным путем. Можно рассматривать звено RC как интегрирующую схему, назначение которой состоит в *усреднении* тока. Однако схема должна усреднить ток по высокой частоте, но не по низкой, так как полное усреднение дало бы на выходе лишь постоянный ток. Эффект интегрирования или усреднения определяется постоянной времени звена, точнее соотношением между постоянной времени $\tau = RC$ и периодом тока. Усреднение происходит тогда, когда постоянная времени значительно превосходит период. Если же, напротив, постоянная времени значительно меньше периода, то усреднения не происходит, и напряжение на выходе звена следует за напряжением на входе. Исходя из этих соображений, можно записать условие правильной работы звена RC в следующей форме:

$$T_0 \ll \tau \ll T_1,$$

где T_0 и T_1 — соответственно периоды несущей частоты и частоты сигнала. Но это неравенство совпадает с предыдущим, если принять во внимание, что $T_0 = 2\pi/\omega_0$, $T_1 = 2\pi/\Omega$.

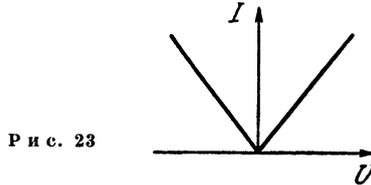
Часто применяется двухтактное детектирование. При этом несущая частота подвергается двухполупериодному выпрямлению, частота удваивается, и ее легче отделить.

Рассмотрим теперь процесс детектирования аналитически. Положим, что детектирование двухтактное и что характеристика детектора может быть записана в виде $I = k|U|$ (рис. 23). Возьмем случай синусоидальной модуляции частотой Ω несущей частоты ω_0 .

Мы имеем следующее выражение для модулированного напряжения:

$$U = U_m (1 + m \sin \Omega t) \sin \omega_0 t.$$

Для нахождения тока в цепи детектора мы должны взять модуль (абсолютное значение) этого выражения. Для этого нужно найти модуль каждого из сомножителей, так как модуль произведения равен произведению модулей сомножителей. Заметим, что коэффициент модуляции m есть положительная величина, не превосхо-



дящая единицы. Поэтому множитель $(1 + m \sin \Omega t)$ не может быть меньше нуля, а стало быть, остается найти только модуль последнего множителя. Таким образом,

$$I = kU_m (1 + m \sin \Omega t) |\sin \omega_0 t|.$$

Но $|\sin \omega_0 t|$ — это не что иное, как выпрямленная синусоида при двухполупериодном выпрямлении. Разложение этой функции в ряд Фурье известно

$$|\sin \omega_0 t| = \frac{2}{\pi} \left(1 - 2 \sum_{k=1}^{\infty} \frac{\cos 2k\omega_0 t}{4k^2 - 1} \right).$$

Итак,

$$\begin{aligned} I &= \frac{2}{\pi} kU_m (1 + m \sin \Omega t) \left(1 - 2 \sum_{k=1}^{\infty} \frac{\cos 2k\omega_0 t}{4k^2 - 1} \right) = \\ &= \frac{2}{\pi} kU_m \left[(1 + m \sin \Omega t) - 2 \sum_{k=1}^{\infty} \frac{\cos 2k\omega_0 t}{4k^2 - 1} - 2m \sum_{k=1}^{\infty} \frac{\sin \Omega t \cos 2k\omega_0 t}{4k^2 - 1} \right]. \end{aligned}$$

Как видим, цель детектирования достигнута: мы получили в составе тока слагаемое $(1 + m \sin \Omega t)$, т. е. модулирующий сигнал. Кроме того, в составе тока имеется ряд высоких частот: четные гармоники несущей частоты с частотами $2k\omega_0$ (первая сумма) и спутники с частотами $2k\omega_0 \pm \Omega$ (вторая сумма). Спектры модулированного колебания до и после детектирования изображены на рис. 24.

Мы рассматривали до сих пор детектирование при характеристике детектора в виде ломаной линии (рис. 21 или 23). Детектор с такого рода характеристикой называется «линейным». Само собой разумеется, что термин «линейный» имеет в виду не характеристику детектора в целом, а лишь форму ветвей характеристики по одну сторону (рис. 21) или по обе стороны (рис. 23) точки излома,

Рассматривая рис. 21, легко понять, что только «линейный» детектор не искажает формы огибающей. Для того чтобы показать, как влияет форма характеристики детектора на результат детектирования, рассмотрим в качестве примера детектирование амплитудно-модулированного колебания квадратичным детектором.

Пусть характеристика детектора дана соотношением

$$I = kU^2.$$

Подставляя в эту формулу выражение синусоидально модулированного по амплитуде колебания, находим

$$I = kU_m^2 (1 + m \sin \Omega t)^2 \sin^2 \omega_0 t = \frac{k}{2} U_m^2 \left[1 + \frac{m^2}{2} + \right. \\ \left. + 2m \sin \Omega t - \frac{m^2}{2} \cos 2\Omega t - \left(1 + \frac{m^2}{2} \right) \cos 2\omega_0 t - \right. \\ \left. - m \sin (2\omega_0 - \Omega) t + m \sin (2\omega_0 + \Omega) t + m^2 \cos 2(\omega_0 - \Omega) t + \right. \\ \left. + m^2 \cos 2(\omega_0 + \Omega) t \right].$$

Как видим, спектр получается ограниченным; он содержит только пять высоких частот, а именно $2\omega_0$, $2\omega_0 \pm \Omega$, $2(\omega_0 \pm \Omega)$.

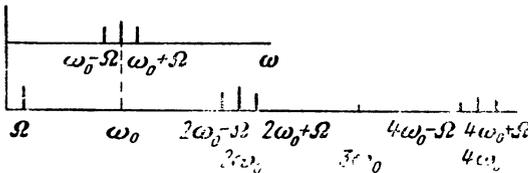


Рис. 24

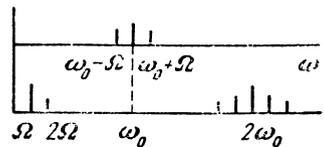
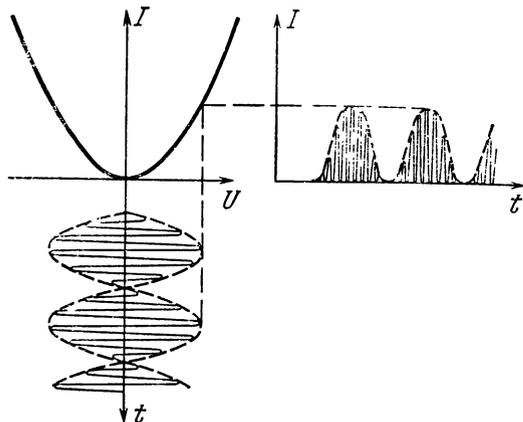


Рис. 25

Но сигнал оказывается искаженным: кроме частоты Ω в спектре детектированного колебания появляется также частота 2Ω , т. е. вторая гармоника сигнала, что означает искажение формы сигнала. Спектры модулированного колебания до и после квадратичного детектирования показаны на рис. 25.

Таким образом оказывается, как и следовало ожидать, что детектор с квадратичной характеристикой не пригоден для детектирования амплитудно-модулированных колебаний. Тем не менее квадратичное детектирование играет в радиотехнике большую роль. Оно применяется, например, для детектирования биений, когда желательно получить в результате детектирования чисто синусоидальное колебание разностной частоты. Именно так обстоит дело в гетеродинных генераторах звуковой частоты. Принцип действия этих генераторов состоит в том, что колебания двух генераторов высокой частоты складываются, образуя биения. Биения детектируются и получается колебание низкой частоты, равной разности двух высоких частот. Если сделать один из генераторов высокой частоты перестраивающимся, то можно получить плавно изменяющуюся в широких пределах низкую частоту. Так, напри-



Р и с. 26

мер, чтобы получить низкую частоту, изменяющуюся в интервале 0—10 кГц, можно взять два генератора высокой частоты: один, скажем, с частотой 100 кГц, а другой — перестраивающийся от 100 до 110 кГц, т. е. всего на 10%.

Очевидным преимуществом гетеродинных генераторов звуковой частоты является простота управления: весь диапазон перекрывается поворотом одной ручки, управляющей небольшим конденсатором переменной емкости в контуре одного из генераторов высокой частоты.

Разберем вопрос о детектировании биений. Для начала положим, что амплитуды обоих колебаний высокой частоты равны, т. е.

$$U = U_m (\sin \omega_1 t + \sin \omega_2 t).$$

Преобразуя, получим

$$U = 2U_m \sin \frac{\omega_1 + \omega_2}{2} t \cos \frac{\omega_1 - \omega_2}{2} t = 2U_m \sin \omega_0 t \sqrt{\frac{1}{2}(1 + \cos \Omega t)},$$

где $\omega_0 = (\omega_1 + \omega_2)/2$, $\Omega = \omega_1 - \omega_2$, т. е. через ω_0 обозначена средняя высокая частота, а через Ω — низкая частота биений. Форма подводимого к детектору напряжения показана на рис. 26.

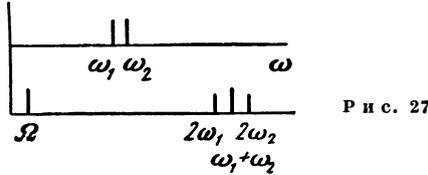
Следует заметить, что биения можно рассматривать как колебания высокой частоты ω_0 , модулированные частотой биений. Но, что существенно, модуляция несинусоидальна, модулирующий множитель имеет вид

$$\sqrt{\frac{1}{2}(1 + \cos \Omega t)}.$$

Если теперь применить квадратичное детектирование, то получится

$$\begin{aligned} I &= kU^2 = 2kU_m^2 (1 + \cos \Omega t) \sin^2 \omega_0 t = \\ &= kU_m^2 (1 + \cos \Omega t) (1 - \cos 2\omega_0 t) = kU_m^2 \left[1 + \cos \Omega t - \right. \\ &\quad \left. - \frac{1}{2} \cos 2\omega_1 t - \frac{1}{2} \cos 2\omega_2 t - \cos (\omega_1 + \omega_2) t \right], \end{aligned}$$

т. е. чисто синусоидальное колебание разностной частоты Ω , удвоенные исходные высокие частоты и их сумма. Спектр биений до и после квадратичного детектирования показан на рис. 27. Наше заключение о целесообразности квадратичного детектирования биений остается в силе и в том более общем случае, когда амплитуды



обоих колебаний высокой частоты оказываются не равными. Пусть будет

$$U = U_1 \sin \omega_1 t + U_2 \sin \omega_2 t.$$

На рис. 28 представлена векторная диаграмма, поясняющая сложение двух напряжений с неравными частотами. Мы имеем для амплитуды результирующего напряжения

$$U_m^2 = U_1^2 + U_2^2 - 2U_1U_2 \cos \alpha.$$

Но, как видно из чертежа, $\alpha = \pi - \Omega t$ и, следовательно, $-\cos \alpha = \cos \Omega t$. Таким образом,

$$U_m = U_1 \sqrt{1 + \frac{U_2^2}{U_1^2} + 2 \frac{U_2}{U_1} \cos \Omega t}.$$

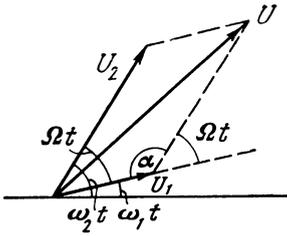
Эта формула представляет собою выражение для изменяющейся амплитуды, т. е. для огибающей биений при неравных амплитудах составляющих колебаний. При квадратичном детектировании, в результате возведения в квадрат, радикал пропадет, и мы получим и в этом случае чисто синусоидальное колебание разностной частоты Ω .

Следует заметить, что если амплитуды обоих напряжений сильно различаются, т. е. $U_2 \ll U_1$ или $U_2 \gg U_1$, то синусоидальное напряжение низкой частоты может быть получено и при линейном детектировании. Это следует из того, что, например, при $U_2 \ll U_1$

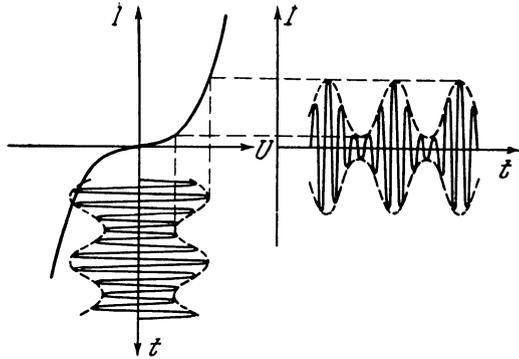
$$\begin{aligned} U_m &= U_1 \sqrt{1 + \frac{U_2^2}{U_1^2} + 2 \frac{U_2}{U_1} \cos \Omega t} \cong \\ &\cong U_1 \left(1 + 2 \frac{U_2}{U_1} \cos \Omega t\right)^{1/2} \cong U_1 \left(1 + \frac{U_2}{U_1} \cos \Omega t\right). \end{aligned}$$

К вопросу о квадратичном детектировании нужно еще добавить, что этот вид детектирования играет большую роль в схемах модуляции и преобразования частоты, о чем речь будет ниже.

Теперь мы должны обратиться к общим соображениям о свойствах характеристик детекторов. Характеристика детектора



Р и с. 28



Р и с. 29

должна быть нелинейной — это очевидно. Но этого мало. Если мы выбрали бы нелинейную характеристику вида

$$I = kU^3,$$

то никакого детектирования не получили бы, так как колебание, как показано на рис. 29, изменив свою форму, осталось симметричным относительно оси времени. Для детектирования годятся только такие нелинейные зависимости, которые несимметричны относительно оси U и которые, вследствие этого, дают постоянную составляющую тока при приложении переменного напряжения. Зависимости такого рода должны содержать четную часть.

Напомним определения.

Четной функцией называется функция, удовлетворяющая условию $f(x) = f(-x)$. Нечетная функция удовлетворяет условию $f(x) = -f(-x)$. Примеры четных функций: x^2 , x^4 , $\cos x$ и т. п. Примеры нечетных функций: x , x^3 , $\sin x$ и т. п.

Степенная функция x^n четна или нечетна в зависимости от того, четен или нечетен показатель степени n .

Детектор детектирует, а выпрямитель выпрямляет, лишь поскольку его характеристика содержит четную часть. Любая функция может быть разложена на четную и нечетную части. Пусть

$$f(x) = f_{\text{ч}}(x) + f_{\text{н}}(x),$$

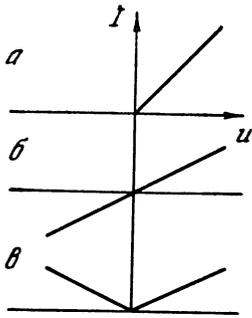
где $f_{\text{ч}}$ — четная часть; $f_{\text{н}}$ — нечетная. Изменим знак аргумента. Тогда по определению получим

$$f(-x) = f_{\text{ч}}x - f_{\text{н}}(x).$$

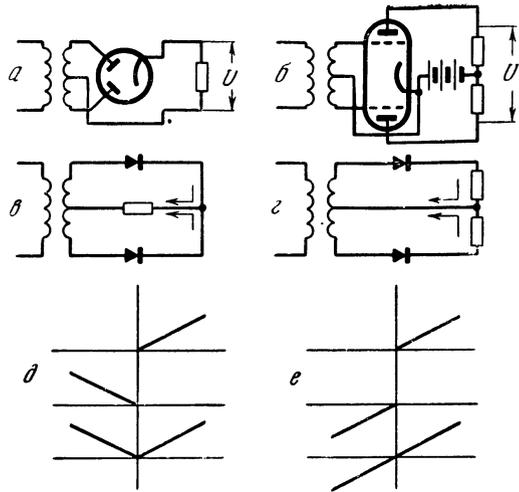
Складывая и вычитая, найдем

$$f_{\text{ч}}(x) = \frac{1}{2} [f(x) + f(-x)],$$

$$f_{\text{н}}(x) = \frac{1}{2} [f(x) - f(-x)].$$



Р и с. 30



Р и с. 31

В качестве примера приведем разложение на четную и нечетную части ломаной характеристики рис. 30, а. Аналитически эту характеристику можно записать в виде

$$I = \begin{cases} kU & U > 0, \\ 0 & U < 0. \end{cases}$$

Четная часть характеристики будет

$$I_{\text{ч}} = \frac{k}{2} |U|,$$

а нечетная

$$I_{\text{н}} = \frac{k}{2} U$$

(рис. 30, б и в). Нетрудно видеть, что, сложив (б) и (в), мы получим (а). Таким образом, аналитически ломаную характеристику на рис. 30, а можно представить формулой

$$I = \frac{1}{2} k (|U| + U),$$

которая годится для всех значений U как положительных, так и отрицательных.

Нечетная часть в-характеристике детектора или выпрямителя может содержаться — она не производит действия. Естественно, однако, пользоваться выпрямителями и детекторами с чисто четными характеристиками. Такими характеристиками обладают все симметричные двухтактные схемы — нечетные члены характеристики устраняются в этих схемах в результате того, что на обе ветви схемы выпрямляемое или детектируемое напряжение подается с противоположными знаками (в противофазе), а токи обеих

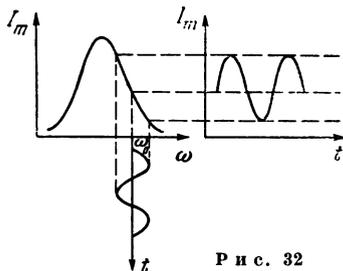
ветвей складываются. В этом смысле схема выпрямителя или детектора обладает свойствами, противоположными свойствам схемы двухтактного усилителя. В последнем напряжение также подается в обе ветви в противофазе, но снимаемое напряжение пропорционально не сумме, а разности токов. Таким образом, двухтактная схема выпрямителя использует только четную часть характеристики нелинейного элемента, а двухтактная схема усилителя использует только нечетную часть характеристики нелинейного элемента. Вот почему в двухтактной схеме усилителя выпадают все четные гармоники. Само собою разумеется, что из нечетных гармоник в случае усиления стремятся сохранить только первую — это соответствует неискаженному усилению, когда результирующая характеристика — прямая. В пояснение этих соображений на рис. 31, *a* и *b* показаны схемы двухтактного выпрямителя на двойном диоде и двухтактного усилителя на двойном триоде. На рис. 31, *в* и *г* даны те же схемы в обобщенном виде. Наконец, рис. 31, *д* и *е* показывают построение результирующих характеристик в предположении, что характеристика отдельного элемента — ломаная.

§ 10. Детектирование ЧМ и ИМ

В этом параграфе кратко излагаются общие принципы детектирования, применяемые при других (кроме АМ) видах модуляции, в частности, при частотной модуляции (ЧМ) и некоторых видах импульсной модуляции (ИМ).

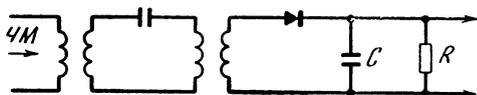
Обычный прием детектирования частотно-модулированных колебаний состоит в том, что ЧМ колебание сперва преобразовывается в АМ колебание, которое затем уже детектируется, как описано в § 9. Операцию преобразования ЧМ в АМ выполняет так называемый частотный детектор. В качестве простейшего частотного детектора может служить расстроенный относительно несущей частоты контур. При отсутствии модуляции, т. е. при $\omega = \omega_0$, рабочая точка располагается на скате резонансной кривой контура, как показано на рис. 32. В нижней части этого рисунка изображен закон изменения частоты во времени. При изменении частоты изменяется амплитуда тока I_m высокой частоты в контуре, как показано в правой части рисунка. Но изменение амплитуды — это и есть амплитудная модуляция. Таким образом, сняв с контура напряжение, остается продетектировать и отфильтровать его обычным образом. Принципиальная схема детектирования ЧМ принимает вид, показанный на рис. 33.

В современной технике применяются более совершенные схемы частотных детекторов. Дело в том, что во избежание нелинейных



Р и с. 32

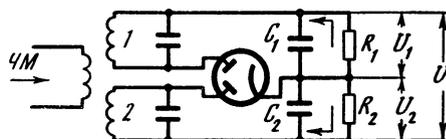
искажений характеристика частотного детектора должна быть прямолинейной. Другими словами, крутизна частотного детектора, выражаемая производной $dI_m/d\omega$, должна быть постоянной в пределах рабочего участка. При широкополосной ЧМ удовлетворить этому требованию при помощи простого контура затруднительно. Поэтому и применяются симметричные схемы, обладающие более благоприятными характеристиками.



Р и с. 33

Пример такой схемы приведен на рис. 34. Здесь имеются два контура 1 и 2, настраиваемых обычно на крайние значения частоты, т. е. на $\omega_0 + \Delta\omega$ и $\omega_0 - \Delta\omega$. Каждый из контуров преобразует ЧМ в АМ, как описано выше. АМ колебания детектируются соответствующими половинами двойного диода. Комбинации RC служат фильтрами. Низкочастотные напряжения U_1 и U_2 возникают на сопротивлениях R_1 и R_2 . Эти напряжения пропорциональны ординатам резонансных кривых соответственно контуров 1 и 2. Но так как эти напряжения на R_1 и R_2 противоположны по знаку (см. стрелки на схеме, указывающие направления токов), то с выхода схемы снимается *разность* напряжений U_1 и U_2 . Характеристика частотного детектора, т. е. зависимость выходного напряжения U от частоты ЧМ колебания на входе, получается путем вычитания двух резонансных кривых, как показано на рис. 35. Частотные детекторы описанного типа относятся к числу так называемых дискриминаторов (различителей).

Перейдем к импульсной модуляции. Основные виды импульсной модуляции, следующие: 1) модуляция импульсов по высоте (по «амплитуде») — АИМ, 2) модуляция импульсов по частоте следования — ЧИМ, 3) модуляция импульсов по длительности — ДИМ и 4) модуляция импульсов по фазе — ФИМ.



Р и с. 34

Детектирование АИМ сигнала ничем по существу не отличается от детектирования обычного амплитудно-модулированного колебания. Различие состоит лишь в составе высокочастотной части спектра, которая все равно должна быть отфильтрована.

То же относится и к ДИМ, хотя это на первый взгляд неочевидно. Но нужно учесть, что составляющая низкой частоты после обычного детектирования ИМ сигнала пропорциональна *площади*

импульса. Следовательно, безразлично, что именно будет изменяться в огибающей импульса при модуляции — высота или длительность. Таким образом, для детектирования как АИМ, так и ДИМ пригодна обычная схема детектора, как, например, на рис. 22.

Иначе обстоит дело с фазово-импульсной модуляцией. При ФИМ, как известно, площадь импульса остается неизменной. Модуляция состоит лишь в смещении импульса относительно тактовой точки. Поэтому для детектирования ФИМ нужно прибегнуть к предварительному преобразованию ФИМ в АИМ или ДИМ, или вообще в такую форму, в которой изменяется площадь импульса. Для преобразования ФИМ в ДИМ, может быть, например, применена схема, содержащая электронное реле, включаемое тактовыми

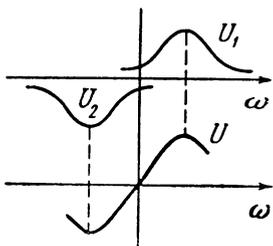


Рис. 35

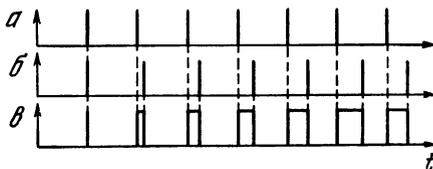


Рис. 36

импульсами и выключаемое фазово-модулированными импульсами (или наоборот), как показано на рис. 36, на котором означают: *a* — включающие тактовые импульсы, *b* — выключающие импульсы ФИМ, *v* — импульсы на выходе электронного реле.

Вопросы, связанные с детектированием при импульсно-кодовой модуляции, как более специальные, здесь не рассматриваются.

§ 11. Модуляция

Модуляция есть процесс, имеющий целью перенос спектра передаваемого (низкочастотного) сигнала в область высоких частот для того, чтобы сигнал мог быть передан радиотехническими средствами, т. е. посредством излучения электромагнитных волн. Такое преобразование необходимо потому, что эффективное излучение может быть получено только при высоких радиочастотах.

Модуляция состоит в том, что в результате определенного воздействия сигнала на колебание высокой частоты сигнал оказывается запечатленным в изменениях тех или иных параметров колебания высокой частоты. Высокая частота как бы *несет* в себе сигнал, сохраняя все его свойства, так что сигнал может быть снова выделен на приемном конце. Отсюда и выразительное название *несущая частота*. Процесс выделения сигнала называется детектированием; он рассмотрен в предыдущих параграфах.

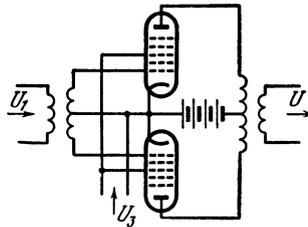
Модуляция в своей основе есть *линейный* процесс, осуществляемый в линейной системе с переменными параметрами.

Для начала мы поясним механизм амплитудной модуляции на примере крайне простой схемы, применявшейся в свое время в малоомощных телефонных передатчиках. Схема представлена на рис. 37. Генератор высокой частоты связан индуктивно с антенной. В антенну последовательно включен обычный угольный микрофон. Его действие состоит, как известно, в том, что под воздействием звуковых колебаний мембрана микрофона колеблется, и, надавливая на заключенный в микрофоне угольный порошок, изменяет его проводимость.

Предположим, что микрофон строго линеен, т. е. что его проводимость есть линейная функция звукового давления. Тогда для



Р и с. 37



Р и с. 38

проводимости цепи антенны имеем $Y = Y_0 + kP$. Пусть звуковое давление синусоидально $P = P_m \sin \Omega t$. Тогда для проводимости можем записать

$$Y = Y_0 + kP_m \sin \Omega t = Y_0 + Y_1 \sin \Omega t = Y_0 (1 + m \sin \Omega t),$$

где $m = Y_1/Y_0$ — коэффициент модуляции проводимости. Пусть, далее, электродвижущая сила, наводимая генератором в антенне, есть

$$E = E_m \sin \omega_0 t,$$

где ω_0 — высокая (несущая) частота. Ток в антенне равен произведению из э. д. с. на проводимость

$$I = EY = E_m Y_0 (1 + m \sin \Omega t) \sin \omega_0 t.$$

Мы получили обычное выражение амплитудно-модулированного колебания для случая синусоидальной модуляции частотой Ω .

Рассмотренная схема есть *линейная схема с переменным параметром*. Переменным параметром является проводимость.

Возвращаясь к выражению для амплитудно-модулированного колебания, перепишем его в более общем виде

$$I = I_m [1 + mf(t)] \sin \omega_0 t, \quad (1)$$

где $f(t)$ — произвольный передаваемый сигнал (предполагается лишь, что $|f(t)| < 1$). Как видим, процесс модуляции состоит в *перемножении* двух функций времени: $1 + mf(t)$ и $I_m \sin \omega_0 t$. Следовательно, модулятор в принципе должен представлять собою перемножающее устройство, т. е. устройство с двумя входами и од-

ним выходом, действие которого состоит в том, что при подаче на два входа двух функций $x(t)$ и $X_0(t)$ на выходе получается

$$y(t) = x(t) X_0(t).$$

Операция умножения на заданную функцию $X_0(t)$ линейна, как это видно из следующего соотношения:

$$\begin{aligned} y(t) &= [x_1(t) + x_2(t)] X_0(t) = \\ &= x_1(t) X_0(t) + x_2(t) X_0(t) = y_1(t) + y_2(t). \end{aligned}$$

Устройства, назначение которых состоит в перемножении двух функций времени, применяемые для модуляции и преобразования частоты, называются обычно смесителями. Термин этот нехорош, так как не выражает существа дела. Мы будем кратко называть устройство для перемножения двух функций *перемножителем*.

Итак, для получения обычной АМ нужно взять перемножитель и подать на его входы $x(t) = 1 + mf(t)$ и $X_0(t) = I_m \sin \omega_0 t$. В результате перемножения получится модулированное колебание (1).

Заметим, что выражение (1) можно переписать в виде

$$I = I_m \sin \omega_0 t + m I_m f(t) \sin \omega_0 t. \quad (2)$$

Первый член есть колебание несущей частоты, разложение второго члена дает боковые полосы модуляционного спектра. Наряду с обычной АМ применяется также передача без несущей частоты, получаемая при помощи так называемой *балансной* модуляции. Для получения АМ колебания без несущей частоты при помощи перемножителя достаточно устранить постоянную составляющую модулирующей функции, т. е. взять просто $x(t) = f(t)$. При этом на выходе перемножителя получим

$$y(t) = x(t) X_0(t) = I_m f(t) \sin \omega_0 t. \quad (3)$$

В спектре этого колебания несущая частота отсутствует; спектр состоит только из двух боковых полос.

Линейный перемножитель может быть построен на основе любого четырехполюсника с управляемым коэффициентом передачи. Так, если имеем вообще $y(t) = kx(t)$ и если мы можем изменять коэффициент передачи k по произвольно заданному закону $k = k(t)$, то и получаем перемножитель, действие которого описывается соотношением $y(t) = k(t)x(t)$.

Для примера рассмотрим работу в качестве перемножителя многосеточной лампы (в частности, геттода). В такой лампе, как известно, крутизна по третьей (так называемой «сигнальной») сетке зависит от напряжения на первой («гетеродинной») сетке. В идеальном случае зависимость крутизны от напряжения на первой сетке была бы линейной, т. е.

$$S_3 = b_0 + b_1 U_1. \quad (4)$$

Подавая на первую и третью сетки два перемножаемых напряжения $U_1(t)$ и $U_3(t)$, мы получили бы в анодной цепи ток

$$I_a = U_3 S_3 = U_3 (b_0 + b_1 U_1) = b_0 U_3 + b_1 U_1 U_3.$$

Как видим, чистого перемножения у нас не получилось бы из-за влияния постоянной составляющей крутизны в выражении (4). Но влияние этой составляющей можно устранить, собрав симметричную (балансную) схему с двумя лампами, вроде показанной на рис. 38 (включение вспомогательных сеток не показано, чтобы не затемнять схемы). Напряжение U_3 подается на третьи сетки обеих ламп с одинаковыми знаками (т. е. в фазе), а напряжение U_1 подается на первые сетки с противоположными знаками (т. е. в противофазе). Выходное напряжение U в силу встречного включения катушек в анодных цепях пропорционально разности анодных токов, т. е. пропорционально величине

$$(b_0 U_3 + b_1 U_1 U_3) - (b_0 U_3 - b_1 U_1 U_3) = 2b_1 U_1 U_3.$$

Таким образом, при условии линейной зависимости крутизны от напряжения схема рис. 38 работает как идеальный линейный перемножитель. В действительности (4) выполняется лишь приближенно, да и то в небольшом интервале напряжений. Тем не менее и при этих условиях балансная схема рис. 38 дает лучший результат, нежели простая (одноламповая) схема.

§ 12. Нелинейный способ получения АМ

Как показано в предыдущем параграфе, модуляция есть по существу линейная операция, выполняемая в линейной системе с переменным параметром. Операция модуляции есть не что иное, как операция *перемножения*.

Но большинство практических схем АМ работает по другому принципу. Дело в том, что построение правильно действующего линейного перемножителя затруднительно. Поэтому прибегают к косвенному способу перемножения, сущность которого состоит в том, что *сумму* двух перемножаемых напряжений заставляют действовать в нелинейной цепи. При этом в составе тока получается составляющая, пропорциональная произведению обоих напряжений. Прежде всего покажем качественно, что получение таким путем АМ колебания возможно. Возьмем два напряжения: напряжение несущей частоты $U_1 = U_m \sin \omega_0 t$ и модулирующее напряжение $U_2 = (1 + m \sin \Omega t)$. Включим источники обоих напряжений последовательно в цепь с нелинейным элементом. Пусть характеристика этого элемента представлена ломаной линией. Форма тока в цепи может быть получена обычным построением, показанным на рис. 39. Нетрудно видеть, что в составе тока получается АМ колебание. Кроме него получаются еще постоянная составляющая, составляющая низкой частоты и ряд высокочастотных составляющих. Все эти лишние составляющие должны быть отфильтрованы.

Разберем теперь подробнее механизм нелинейного перемножения. Пусть характеристика нелинейного элемента представлена степенным рядом

$$I = a_0 + a_1 U + a_2 U^2 + a_3 U^3 + a_4 U^4 + \dots$$

В нашем случае

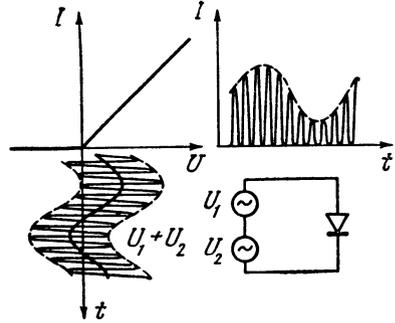
$$U = U_1 + U_2,$$

так что

$$\begin{aligned} I = a_0 + a_1(U_1 + U_2) + a_2(U_1 + U_2)^2 + a_3(U_1 + U_2)^3 + \\ + a_4(U_1 + U_2)^4 + \dots = a_0 + a_1U_1 + a_1U_2 + a_2U_1^2 + \\ + a_2U_2^2 + \overline{2a_2U_1U_2} + a_3U_1^3 + a_3U_2^3 + 3a_3U_1^2U_2 + \\ + 3a_3U_1U_2^2 + a_4U_1^4 + \dots \end{aligned} \quad (1)$$

Мы видим, что нужный нам член, содержащий произведение U_1U_2 , получается в составе тока *только* за счет квадратичного члена вольтамперной характеристики нелинейного элемента. Все остальные члены дают лишние составляющие, так что мы имеем нечистое перемножение, и получаемый ток требует последующей очистки, хотя бы путем фильтрации.

Однако возможно путем последовательного усложнения схемы добиться исключения ряда лишних членов и получить, таким образом, более совершенное перемножение.



Р и с. 39

Обратимся к рис. 40. Исходная схема с двумя напряжениями U_1 и U_2 , действующими в нелинейной цепи, показана на рис. 40, а. Совместим две такие схемы, как показано на рис. 40, б, так, чтобы они имели общую ветвь, в которую включено напряжение U_1 . Напряжение, снимаемое со схемы, равно $U = R(I_1 - I_2)$. Заметим, что напряжение U_2 подано в схему дважды и притом так, что в верхней половине действует суммарное напряжение $U_1 + U_2$ (мгновенные направления напряжений обозначены стрелками), а в нижней половине — разностное $U_1 - U_2$. Для токов получаем

$$\begin{aligned} I_1 = a_0 + a_1(U_1 + U_2) + a_2(U_1 + U_2)^2 + \\ + a_3(U_1 + U_2)^3 + a_4(U_1 + U_2)^4 + \dots, \\ I_2 = a_0 + a_1(U_1 - U_2) + a_2(U_1 - U_2)^2 + \\ + a_3(U_1 - U_2)^3 + a_4(U_1 - U_2)^4 + \dots \end{aligned}$$

Составляя разность токов, получим

$$\begin{aligned} I_1 - I_2 = 2a_1U_2 + 4a_2U_1U_2 + 6a_3U_1^2U_2 + 2a_3U_2^3 + \\ + 8a_4U_1^3U_2 + 8a_4U_1U_2^3 + \dots \end{aligned} \quad (2)$$

(Предполагается, что схема симметрична, т. е. что коэффициенты a_k в разложениях токов I_1 и I_2 одинаковы.) Как видим, в результате вычитания целый ряд членов выпал. Сохранились только члены,

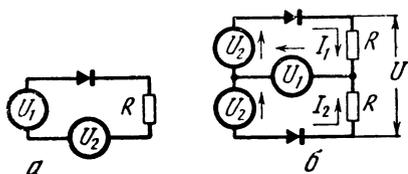
содержащие нечетные степени U_2 . Если бы характеристика нелинейного элемента была строго квадратичной, т. е.

$$I = kU^2,$$

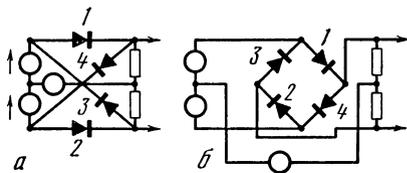
то мы получили бы чистое перемножение по формуле

$$k(U_1 + U_2)^2 - k(U_1 - U_2)^2 = 4kU_1U_2.$$

Схемы типа рис. 40, б называются *балансными*. Они применяются, в частности, для получения модулированного колебания, не содержащего несущей частоты. Действительно, если U_1 есть напряжение несущей частоты, то из (2), как видим, оно выпадает. Такого рода модуляция называется балансной модуляцией.



Р и с. 40



Р и с. 41

Идея усовершенствования схемы рис. 40, а, приводящая к балансной схеме рис. 40, б, сводится к тому, что, изменяя знак одного из напряжений и производя вычитание токов, мы избавляемся от целого ряда излишних членов. По этому пути можно идти и далее. Возьмем теперь две одинаковые балансные схемы, изменим в одной из них знаки *обоих* напряжений U_1 и U_2 на обратные и образуем напряжение, равное сумме напряжений, получаемых на каждой из схем в отдельности. Такой результат достигается в схеме рис. 41, а, представляющей собою совмещение двух балансных схем. Снимаемое с этой схемы напряжение можно представить в виде

$$U = U' + U'',$$

где U' и U'' — напряжения, развиваемые каждой из совмещенных балансных схем в отдельности. Для этих двух напряжений мы можем записать (см. (2))

$$U' = R(I'_1 - I'_2) = 2R(a_1U_2 + 2a_2U_1U_2 + 3a_3U_1^2U_2 + a_3U_2^3 + 4a_4U_1^3U_2 + 4a_4U_1U_2^3 \dots),$$

$$U'' = R(I''_1 - I''_2) = 2R(-a_1U_2 + 2a_2U_1U_2 - 3a_3U_1^2U_2 - a_3U_2^3 + 4a_4U_1^3U_2 + 4a_4U_1U_2^3 + \dots).$$

Выражение для U'' отличается от выражения для U' тем, что знаки обоих напряжений U_1 и U_2 заменены на обратные. В результате сложения получаем

$$U = U' + U'' = 4R(a_2U_1U_2 + 2a_4U_1^3U_2 + 2a_4U_1U_2^3 + \dots). \quad (3)$$

Это выражение показывает, что достигнуто дальнейшее устранение лишних членов. В частности, если характеристика нелинейного элемента выражается многочленом степени не выше третьей, т. е.

$$I = a_0 + a_1 U + a_2 U^2 + a_3 U^3,$$

то схема рис. 41, а (при условии полной ее симметрии) дает идеальное перемножение. Можно перерисовать эту схему в виде рис. 41, б, как ее чаще всего изображают. Получается как бы замкнутое кольцо из четырех нелинейных элементов. Поэтому схему рис. 41 называют *кольцевым модулятором*. Эта схема имеет широкое применение в качестве модулятора и преобразователя частоты.

Следует заметить, что дальнейшее улучшение схемы тем же способом, последовательное применение которого привело нас сперва к балансной, а затем к кольцевой схеме, невозможно. В самом деле, рассмотрение формулы (3) показывает, что в составе напряжения, снимаемого с кольцевой схемы, сохраняются только члены, содержащие произведения нечетных степеней U_1 и U_2 . Следовательно, при изменении знака одного или обоих напряжений *все* члены выражения (3) сохраняют одинаковый знак и ни один из этих членов не может быть устранен в отдельности сложением или вычитанием. Поэтому построение более сложной схемы путем совмещения двух кольцевых схем ничего не дает.

§ 13. Практические схемы модуляции

Рассмотрим простейшие и наиболее употребительные принципиальные схемы модуляторов.

А. Амплитудная модуляция

Для обычной амплитудной модуляции применяются главным образом две схемы: схема сеточной модуляции и схема анодной модуляции. Схема сеточной модуляции показана на рис. 42. Напряжение несущей частоты U_1 и модулирующее напряжение U_2 подаются на сетку триода. Блокировочная емкость C_1 обеспечивает прохождение высокой частоты в обход обмотки низкочастотного трансформатора. Модуляция получается за счет нелинейности характеристики триода $I_a = f(U_g)$. В составе анодного тока получают лишние составляющие. Их отфильтровывает контур LC , настроенный на несущую частоту. Ширина полосы пропускания контура берется немного больше ширины модуляционного спектра, т. е. двойной ширины спектра сигнала. При таком условии контур отфильтровывает как постоянную составляющую и низкую частоту, так и высокочастотные составляющие, засоряющие модуляционный спектр.

Схема анодной модуляции показана на рис. 43. Ее отличие от схемы сеточной модуляции состоит в том, что напряжение несу-

щей частоты U_1 и модулирующее напряжение U_2 действуют в разных цепях лампы L_1 ; напряжение U_1 приложено непосредственно к сетке, а напряжение U_2 подается в анодную цепь. Так как действие напряжения в анодной цепи, грубо говоря, в μ раз слабее действия такого же напряжения в цепи сетки, то модулирующее напряжение предварительно усиливается вспомогательной лампой L_2 . С принципиальной точки зрения действие схем рис. 42 и 43 одинаково, но схема анодной модуляции рис. 43 имеет чисто технические преимущества, обсуждаемые в курсе передатчиков.

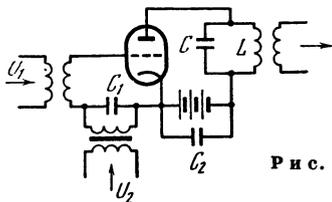


Рис. 42

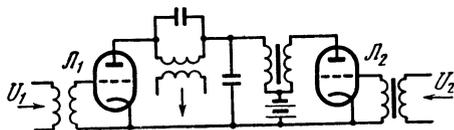


Рис. 43

Б. Частотная модуляция

Для получения частотной модуляции необходимо изменять частоту задающего генератора, что проще всего сделать, изменяя настройку контура задающего генератора. В простейшей схеме радиотелефонной модуляции это достигается включением конденсаторного микрофона непосредственно в контур задающего генератора. В современной технике используется главным образом схема, действие которой основано на применении так называемой *реактивной лампы*. Реактивная лампа представляет собою устройство, входное сопротивление которого практически чисто реактивно, и изменяет свою величину в соответствии с изменениями подводимого модулирующего напряжения. Теория реактивной лампы в основе своей крайне проста и сводится к следующему. Составим схему рис. 44 и запишем выражение для тока I

$$I = \frac{U}{R + j\omega L}.$$

Если выбрать параметры R и L так, чтобы было $R \gg \omega L$, то

$$I \cong U/R.$$

Напряжение на сетке

$$U_g = j\omega LI \cong j\omega \frac{L}{R} U.$$

Анодный ток (если пренебречь реакцией анода) равен

$$I_a = S U_g = j\omega \frac{LS}{R} U.$$

Положим теперь, что параметры подобраны так, что

$$I \ll I_a,$$

т. е. что входной ток полностью определяется анодным током I_a (а составляющей I пренебрегается). Тогда мы можем найти проводимость схемы между точками 1 и 2, разделив входной ток на приложенное напряжение

$$Y = I_a/U \cong j\omega LS/R = j\omega C_s,$$

где эквивалентная емкость $C_s = LS/R$. Таким образом, оказывается, что при сделанных допущениях входная проводимость схемы имеет чисто емкостный характер, или, короче, что схема ведет себя, как емкость¹.

Теперь остается сделать емкость управляемой. Это легко сделать, так как емкость зависит от крутизны. Выбрав рабочую точку на квадратичном участке характеристики триода, мы получим изменение крутизны, а следовательно, и эквивалентной емкости, пропорциональное изменению напряжения на сетке. В самом деле, если

$$I_a = a_0 + a_1U + a_2U^2,$$

то

$$S = dI_a/dU = a_1 + 2a_2U.$$

Принципиальная схема модулятора ЧМ с реактивной лампой показана на рис. 45. На этой схеме L_1 — генераторная лампа; LC — задающий контур; L_2 — подключенная параллельно кон-

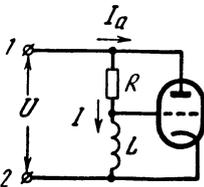


Рис. 44

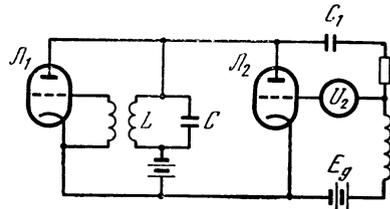


Рис. 45

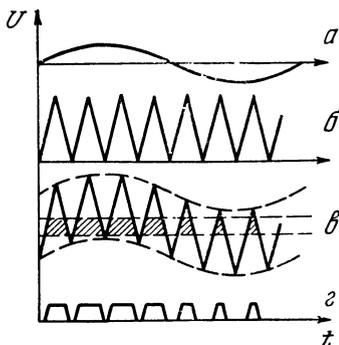
туру реактивная лампа; U_2 — модулирующее напряжение, подаваемое на сетку реактивной лампы; C_1 — емкость, защищающая сетку реактивной лампы от анодного напряжения; E_g — напряжение смещения для установки рабочей точки.

В. Импульсная модуляция

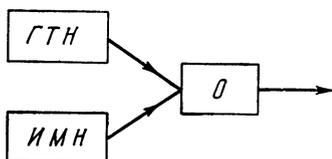
Мы не будем останавливаться на АИМ², так как, во-первых, для этого вида модуляции пригодны те же способы, что и для обычной АМ, а во-вторых, АИМ малоупотребительна в настоящее время вследствие своей малой помехоустойчивости. Наибольшее применение имеют ДИМ и ФИМ, механизм которых мы и рас-

¹ Изменив схему фазовращателя RL , можно получить индуктивную, а также
² любую комплексную входную проводимость.

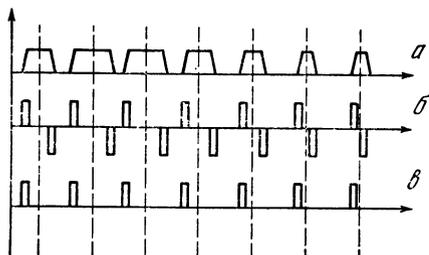
См. стр. 418.



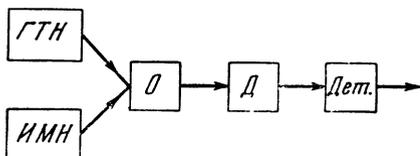
Р и с. 46



Р и с. 47



Р и с. 48



Р и с. 49

смотрим. Простейший и наиболее распространенный способ получения последовательности импульсов, модулированных по фазе или по длительности, основан на наложении модулирующего напряжения на периодическое напряжение треугольной формы с основной частотой, равной частоте следования немодулированной последовательности импульсов. Полученная сумма напряжений подвергается дальнейшей обработке: ограничению, дифференцированию и т. п. Начнем с модуляции по длительности. На рис. 46, а изображено модулирующее напряжение, на рис. 46, б — треугольное напряжение, на рис. 46, в — результат сложения обоих напряжений. Если теперь ограничить результирующее напряжение сверху и снизу, как показано штриховыми линиями на рис. 46, в, то получатся трапецидальные импульсы, модулированные по длительности. Кругизной скатов легко управлять, изменяя соотношение между размахом треугольного напряжения и полосой ограничения. Таким образом, скелетная схема ДИМ состоит из генератора треугольного напряжения (ГТН на рис. 47), источника модулирующего напряжения (ИМН) и ограничителя (О). Схемы названных элементов никаких принципиальных особенностей не содержат, и мы ограничиваемся скелетной схемой. Для получения ФИМ можно использовать ДИМ как промежуточный продукт. Возьмем последовательность импульсов, модулированных по длительности, в том виде, в каком эта по-

следовательность нами получена вышеописанным способом, и продифференцируем ее (рис. 48, б). Получим прямоугольные импульсы противоположной полярности; положительные импульсы соответствуют переднему скату дифференцируемых трапецеидальных импульсов, отрицательные — заднему скату. Если теперь пропустить полученную последовательность через детектор, который отсекает отрицательные импульсы, то получится (рис. 48, в) последовательность прямоугольных¹ импульсов, модулированных по фазе, т. е. смещенных относительно тактовых точек, отмеченных на рис. 48 вертикальными линиями. Таким образом, скелетная схема ФИМ отличается от схемы ДИМ тем, что к ней добавляется дифференцирующее звено (Д) и детектор (Дет. на рис. 49). Варианты и подробности схем импульсной модуляции рассматриваются в соответствующих специальных курсах.

§ 14. Преобразование частоты

Рассмотренное выше преобразование спектра при АМ состоит в том, что низкочастотный спектр сигнала смещается по шкале частот в область радиочастот. Такое смещение может рассматриваться как частный случай более общей линейной операции, называемой в радиотехнике преобразованием частоты.

Под преобразованием частоты подразумевается вообще смещение спектра по шкале частот в ту или другую сторону, т. е. в область как более высоких, так и более низких частот. Предполагается, что при таком смещении или переносе спектра он не подвергается искажению.

Для преобразования частоты применяется общий прием, состоящий в том, что данное колебание (подлежащее преобразованию) умножается на синусоидальное колебание, доставляемое отдельным генератором — гетеродином. Этот прием называют часто гетеродинированием.

Сущность преобразования состоит в том, что при перемножении тригонометрических функций разных аргументов получаются два члена, один из которых зависит от суммы, другой — от разности аргументов, например

$$\sin x \sin y = \frac{1}{2} [\cos(x - y) - \cos(x + y)].$$

Рассмотрим для примера преобразование, состоящее в смещении модуляционного спектра из области радиочастот в область так называемых промежуточных частот (порядка сотен килогерц). На рис. 50, а изображен исходный спектр, имеющий несущую частоту ω_0 и две сплошные боковые полосы. Рис. 50, б изображает

¹ Практически импульсы не могут быть прямоугольными. Их форма также приближается к трапецеидальной, т. е. скаты импульсов имеют конечную крутизну, определяемую шириной полосы пропускания всей схемы модулятора.

спектр колебания гетеродина — одну спектральную линию с частотой ω_1 . На рис. 50, *в* представлен спектр на выходе перемножителя; он состоит из двух модуляционных спектров, имеющих несущие частоты $\omega_0 - \omega_1$ и $\omega_0 + \omega_1$. На рис. 50, *г* изображена частотная характеристика полосового фильтра промежуточной частоты (ФПЧ) и, наконец, на рис. 50, *д* — спектр на выходе фильтра; на рис. 51 изображена скелетная схема устройства, выполняющего

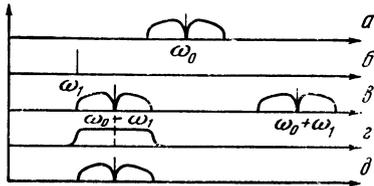


Рис. 50

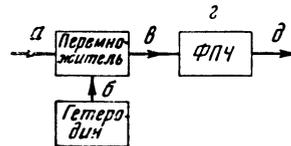


Рис. 51

все описанные действия, в результате которых спектр исходного сигнала оказывается перемещенным с частоты ω_0 на частоту $\omega_0 - \omega_1$. Буквы на рис. 51 соответствуют рис. 50.

Заметим теперь, что если мы будем плавно изменять частоту гетеродина ω_1 , то преобразованный спектр будет плавно же перемещаться по шкале частот, так как средняя частота спектра $\omega_{\text{пр}} = \omega_0 - \omega_1$. На возможности такого перемещения основан принцип действия супергетеродинного приемника. Принцип этот состоит в том, что на входе стоит перемножитель с гетеродином по схеме рис. 51; настройка осуществляется изменением частоты гетеродина. Усиление производится на промежуточной частоте, причем усилитель промежуточной частоты (УПЧ) имеет большое число ступеней, снабженных фильтрами промежуточной частоты.

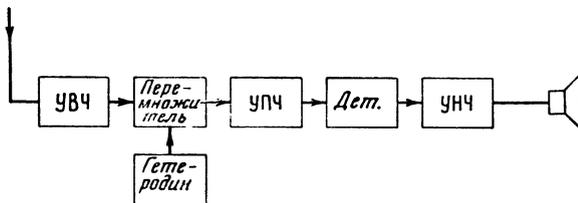


Рис. 52

Самое главное в этой схеме это то, что фильтры настроены раз навсегда на одну определенную полосу. В супергетеродинном приемнике не фильтры (или контуры) настраиваются на частоту принимаемого сигнала, а, наоборот, принимаемый сигнал настраивается на частоту УПЧ путем вышеописанного перемещения спектра сигнала по шкале частот. Это дает неоценимые преимущества. Во-первых, настройка приемника производится одной ручкой (управляющей частотой гетеродина) независимо от числа

ступеней усиления. Во-вторых, фильтры ступеней УПЧ, не подвергающиеся перестройке, могут быть сделаны по сложной схеме, обеспечивающей минимальные искажения при максимальной избирательности. Нужно, впрочем, заметить, что схема супергетеродинного приемника обычно усложняется введением УВЧ — усиления высокой частоты (до перемножителя), однако и в таком виде эта схема является наилучшей схемой приемника. Она показана в скелетной форме на рис. 52. Кроме уже упомянутых элементов, на ней указаны детектор (Дет.) и усилитель низкой частоты (УНЧ).

Во многих радиотехнических устройствах как на передающем, так и на приемном конце системы радиосвязи прибегают к многократному преобразованию частоты.

В заключение заметим еще, что возможен специальный случай преобразования частоты, когда гетеродин настраивается в точности на частоту ω_0 . При этом модуляционный спектр смещается в область низких частот ($\omega_{np}=0$). Это есть случай так называемого синхронного детектирования, о котором речь будет позднее (§ 38).

Г л а в а II

ГЕНЕРИРОВАНИЕ КОЛЕБАНИЙ

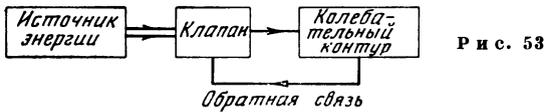
§ 15. Автоколебания

Всякое радиотехническое устройство представляет собою длинную цепь, составленную из линейных и нелинейных звеньев. Вдоль этой цепи проходят электрические колебания, передаваемые от звена к звену в усиленном или так или иначе преобразованном виде в зависимости от назначения данного звена. Существенно, что в большинстве звеньев колебания возникают лишь под воздействием колебаний в предшествующем звене. Так, например, на выходе данной ступени усилителя колебания возникают лишь при подаче на вход данной ступени изменяющегося напряжения с выхода предшествующей ступени. При таких обстоятельствах мы называем колебания *вынужденными*.

Но если следовать от конца к началу сложной радиотехнической цепи, мы всегда дойдем до звена, в котором колебания не являются вынужденными, в котором они впервые зарождаются. Такого рода самостоятельно возникающие колебания носят название *автоколебаний*. Звено, в котором возникают автоколебания, называется по физической терминологии *автоколебательной системой*. С точки зрения техники это есть *генератор* колебаний. Автоколебательная система подчинена, разумеется, закону сохранения энергии; всякий электрический генератор можно рас-

смагивать как преобразователь энергии источника постоянного тока в энергию электрических колебаний.

Рассмотрим типичный механизм автоколебаний; основные черты этого механизма мы найдем во многих конкретных схемах генераторов. Начнем со свободных колебаний в простейшем контуре. Если бы контур был без потерь, то в нем происходили бы незатухающие колебания. Но во всяком реальном контуре имеются потери; кроме того, мы рассматриваем генератор как источник колебательной энергии, и, следовательно, некоторое количество энергии отбирается во внешние цепи, в нагрузку генератора. Поэтому простой контур не может служить генератором; запас энергии, который был ему первоначально сообщен, быстро исчерпается, и колебания затухнут. Следовательно, для получения незатухающих колебаний необходимо все время пополнять запас энергии в контуре, по мере ее расходования.



Итак, в состав генератора, кроме колебательного контура, должен входить *источник энергии*. Этим источником является батарея или выпрямитель — вообще источник постоянного напряжения. Для того чтобы поддержать колебания в контуре, необходимо, чтобы энергия подавалась в контур периодически, отдельными порциями, в такт с уже происходящими колебаниями; поток энергии должен быть пульсирующим. Орган, управляющий поступлением энергии в контур, мы назовем *клапаном*. Для того чтобы работа клапана была согласована с колебаниями в контуре, нужно чтобы работа клапана в свою очередь управлялась самим контуром. Механизм управления клапаном со стороны контура называется *обратной связью*. Таким образом, получается замкнутая цепь взаимодействия, которая и необходима для получения автоколебаний. Скелетная схема автоколебательной системы показана на рис. 53.

Покажем практическое осуществление этой схемы на примере простейшего лампового генератора, схема которого представлена на рис. 54. Потери энергии в контуре LC , обусловленные наличием сопротивления R , восполняются за счет импульсов анодного тока. Источником энергии служит анодная батарея. Роль клапана играет сетка триода, напряжение на сетке управляет анодным током. В свою очередь напряжение на сетке определяется колебаниями в контуре. Цепь взаимодействия замыкается через катушку L_1 в анодной цепи триода, индуктивно связанную с катушкой контура L ; коэффициент взаимной индукции обозначен через M . Таким образом, все перечисленные выше элементы автоколебатель-

ной системы в схеме рис. 54 налицо. Однако нужно заметить, что хотя эта схема *может* генерировать колебания, но не обязательно *будет* их генерировать, если не выполнить некоторые обязательные условия. В дальнейшем нам предстоит сделать подробный анализ работы генератора. Нам нужно будет получить ответ на ряд основных вопросов, как, например: 1) при каких условиях возникнут автоколебания? 2) какова будет амплитуда установившихся автоколебаний?

На второй вопрос нельзя ответить, не разобравшись в той роли, которую играет нелинейность автоколебательной системы.

Предварительные соображения по этому поводу даны в следующем параграфе.

§ 16. Энергетика автоколебаний; динамическая устойчивость

Мощность, *теряемая* в сопротивлении R (рис. 54), равна

$$P_- = \frac{d^2}{2} \cdot \frac{U_m^2}{R},$$

т. е. пропорциональна квадрату амплитуды напряжения на контуре. Обозначим мощность, *поступающую* в контур за счет источника энергии, через P_+ . Если бы вся схема генератора, включая работу клапана и действие обратной связи, была линейна, то и мощность P_+ росла бы с квадратом амплитуды. Изобразим графически зависимости теряемой и получаемой контуром мощностей от амплитуды, выбрав для амплитуд квадратичный масштаб. В таком масштабе оба графика будут представлены для линейной системы прямыми линиями, проходящими через начало координат (рис. 55). Если прямая P_+ лежит выше прямой P_- , то это означает, что контур получает больше энергии, чем теряет, т. е. что запас энергии в контуре возрастает, а это в свою очередь означает рост амплитуды колебания. Если же, наоборот, прямая P_+ лежала бы ниже P_- , то это означало бы, что расход энергии превосходит приход, и при таком условии в контуре были бы возможны только затухающие колебания.

В условиях рис. 55 амплитуда росла бы неограниченно. Практически этого не может быть: всякая реальная схема нелинейна. В рассматриваемом случае нелинейность может быть обусловлена, например, насыщением триода или влиянием сеточного тока. Благодаря влиянию нелинейности получаемая мощность P_+ растет медленнее, чем с квадратом амплитуды, график искривляется и принимает вид рис. 56. При некотором значении амплитуды графики P_+ и P_- пересекаются. Точка пересечения отвечает *балансу энергии*, т. е. точному равенству теряемой и получаемой энергии. Очевидно, что только при условии баланса энергии запас энергии в контуре остается неизменным, а следовательно, амплитуда колебаний — постоянной. Таким образом, абсцисса

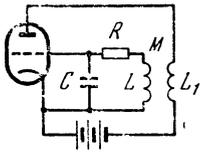


Рис. 54

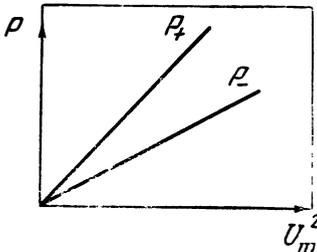


Рис. 55

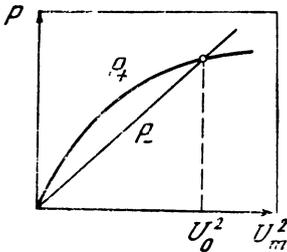


Рис. 56

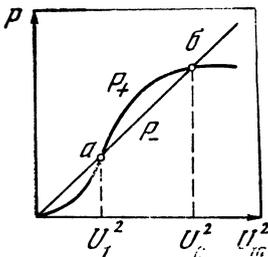


Рис. 57

точки пересечения на рис. 56 непосредственно определяет установившуюся амплитуду колебания U_0 .

Итак, наличие нелинейности есть обязательное свойство автоколебательной системы, способной генерировать колебания с неизменной (установившейся) амплитудой.

Баланс энергии может достигаться и не при одном значении амплитуды. Рассмотрим рис. 57. Здесь имеются две точки пересечения, — спрашивается, которая из них определяет установившуюся амплитуду колебания? Нам придется ввести понятие об *устойчивости* колебательного движения, которую будем называть *динамической устойчивостью* (в отличие от статической устойчивости, которая характеризует не движение, а состояние равновесия). Мы будем называть колебательное движение *устойчивым*, если имеется тенденция к сохранению установившейся амплитуды колебания. Пусть некоторое внешнее воздействие изменило амплитуду колебания, а затем это воздействие было прекращено, и система была предоставлена самой себе. Если теперь в системе восстановится первоначальное значение амплитуды, то мы скажем, что система *динамически устойчива*. Если же, напротив, амплитуда будет и далее изменяться в том же направлении уже без внешнего воздействия, то система *динамически неустойчива*. Легко видеть, что понятия динамической и статической устойчивости во многом сходны. Из приведенного определения сразу вытекает и способ испытания на устойчивость: достаточно вывести систему из данного состояния и посмотреть, как она будет себя вести в дальнейшем.

Применим этот способ к исследованию системы, соотношения в которой характеризуются графиками рис. 57. Вопрос сводится теперь к тому, устойчивы или неустойчивы режимы, определяемые

точками пересечения a и b . Начнем с точки a и предположим, что по какой-либо причине амплитуда стала меньше U_1 . В этой области прямая P_- лежит выше кривой P_+ , что означает, что система теряет энергии больше, чем получает. Поэтому амплитуда будет убывать и далее, пока колебания совсем не затухнут. Если же взять амплитуду больше чем U_1 , то будем иметь $P_+ > P_-$, и амплитуда будет возрастать. Отсюда следует, что точка a хотя и отвечает условию баланса энергии, но не является устойчивой: при малейшем изменении амплитуды появится тенденция к дальнейшему ее изменению в ту же сторону. Если амплитуда станет больше U_1 , то она будет затем нарастать и далее, пока не достигнет значения U_0 . Дальнейший рост амплитуды на этом прекратится, так как при $U_m > U_0$ мы имеем снова $P_- > P_+$. Таким образом, точка b является устойчивой точкой, и ее абсцисса определяет установившуюся амплитуду U_0 .

Итак, различие между системами, свойства которых отображены графиками рис. 56 и 57, состоит в том, что первая самостоятельно раскачивается до установившейся амплитуды U_1 , вторая же нуждается в первоначальном толчке, который создал бы амплитуду, превосходящую U_1 , после чего дальнейшее нарастание амплитуды до U_0 произойдет уже автоматически. Первая система называется системой с *мягким самовозбуждением*, или просто мягкой системой; вторая — системой с *жестким самовозбуждением*, или жесткой системой. Следует еще заметить, что различие между мягкой и жесткой системой сводится к тому, устойчива или неустойчива система в состоянии покоя (т. е. когда амплитуда равна нулю). Рассуждая, как выше, относительно начала координат (которое ведь также является точкой баланса), легко установить, что жесткая система имеет устойчивое состояние покоя, а мягкая — неустойчивое.

§ 17. Условие самовозбуждения генератора

Если не задаваться вопросом о величине установившейся амплитуды, а интересоваться лишь тем, возбуждятся в данной системе автоколебания или нет, то такая постановка вопроса приведет нас к линейной задаче. Действительно, мы интересуемся при этом лишь начальной стадией явления, когда колебания в генераторе только что возникли, т. е. когда амплитуды еще очень малы, тогда как нелинейность проявляется лишь при значительных амплитудах. Поэтому математическая формулировка задачи о самовозбуждении приведет нас к *линейному* дифференциальному уравнению. Составим такое уравнение для генератора по схеме рис. 58. Если обозначить: I — ток в контуре, I_a — переменная составляющая анодного тока, $U = U_g$ — напряжение на контуре, равное напряжению на сетке, то будем иметь уравнение напряжений в контуре

$$L \frac{dI}{dt} + RI + \frac{1}{C} \int Idt = M \frac{dI_a}{dt}. \quad (1)$$

Здесь в правой части стоит э. д. с., наводимая в контуре катушкой связи. С другой стороны,

$$I_a = S_0 U, \quad U = \frac{1}{C} \int Idt,$$

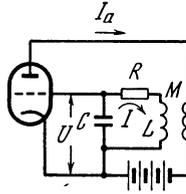
где S_0 — крутизна триода. Учитывая эти соотношения, получаем из (1)

$$LC \frac{d^2U}{dt^2} + RC \frac{dU}{dt} + U = MS_0 \frac{dU}{dt}$$

или

$$\frac{d^2U}{dt^2} + (2\alpha - \omega_0^2 MS_0) \frac{dU}{dt} + \omega_0^2 U = 0, \quad (2)$$

где, как обычно, $\alpha = R/2L$, $\omega_0 = 1/\sqrt{LC}$ соответственно означают показатель затухания и резонансную частоту. Мы видим, что при



Р и с. 58

наличии обратной связи показатель затухания изменяется: он убывает, так как величина $\omega_0^2 MS_0$ положительна. Показатель α характеризует скорость затухания колебаний, т. е. скорость расходования запаса энергии в контуре. Увеличивая обратную связь, можно достичь равенства

$$2\alpha - \omega_0^2 MS_0 = 0.$$

При этом результирующий (т. е. с учетом действия обратной связи) показатель затухания обратится в нуль. Это означает, что колебания уже не будут затухать. Отрицательный же показатель затухания означает, что колебания будут нарастать по экспоненциальному закону. Итак, условие

$$2\alpha - \omega_0^2 MS_0 \leq 0 \quad (3)$$

есть условие возникновения колебаний, т. е. *условие самовозбуждения* (система предполагается мягкой). Вводя затухание d контура, можно переписать это условие в виде

$$M \geq d/\omega_0 S_0. \quad (4)$$

Таким образом, чтобы генератор возбудился при данных затухании и настройке контура и при данном триоде, нужно увеличивать обратную связь (т. е. коэффициент взаимоиндукции M между контурной катушкой и катушкой обратной связи).

Условие самовозбуждения можно истолковать еще несколько иначе. Перепишем уравнение (2) в виде

$$L \frac{d^2 U}{dt^2} + \left(R - \frac{MS_0}{C} \right) \frac{dU}{dt} + \frac{1}{C} U = 0.$$

Теперь можно представить дело так, что обратная связь вносит в контур *отрицательное сопротивление*

$$R' = -MS_0/C \quad (5)$$

и условие самовозбуждения с этой точки зрения состоит в том, что результирующее сопротивление становится отрицательным, т. е.

$$R + R' = R - MS_0/C \leq 0. \quad (6)$$

Такая трактовка вопроса очень распространена в радиотехнике.

§ 18. Комплексная обратная связь

Обратная связь в генераторе вносит в контур отрицательное активное сопротивление. Такого рода обратную связь называют *положительной* в отличие от *отрицательной обратной связи*, широко применяемой в усилителях и других радиотехнических устройствах для исправления характеристик и уменьшения искажений. Но обратная связь может рассматриваться и в более общем виде. Возьмем к примеру схему рис. 59. Здесь имеется контур *LCR*. Усиленное напряжение снимается с сопротивления в анодной цепи триода и через *фазовращатель* Φ вводится обратно в контур; обусловленное обратной связью дополнительное напряжение U' может иметь произвольную фазу и произвольную амплитуду. Это и есть наиболее общий случай обратной связи. На рис. 60, *а* показана векторная диаграмма напряжений в контуре. Если разделить все напряжения на общий ток, то можно перейти от диаграммы напряжений к совершенно подобной ей диаграмме сопротивлений, как это делается в теории переменных токов. Диаграмма сопротивлений изображена на рис. 60, *б*. Она соответствует эквивалентной схеме рис. 61, на которой обозначены также все напряжения. Теперь мы видим, что действие обратной связи в общем случае может быть истолковано как внесение в схему *произвольного* комплексного сопротивления. Характер этого сопротивления зависит от фазы обратной связи. При $\varphi = 0$ мы имеем $Z' = R'$, результирующее активное сопротивление возрастает; это есть отрицательная обратная связь. При $\varphi = \pi$ получаем $Z' = -R'$, т. е. обратная связь вносит отрицательное активное сопротивление; это есть положительная обратная связь. Если при этом R' по абсолютному значению больше R , то результирующее сопротивление отрицательно, и в системе возникают автоколебания. Таким образом, при положительной обратной связи условие $|R'| > R$ есть условие самовозбуждения.

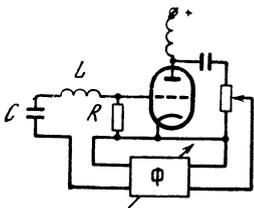


Рис. 59

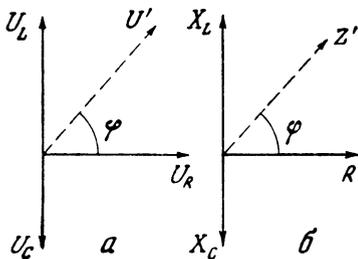


Рис. 60

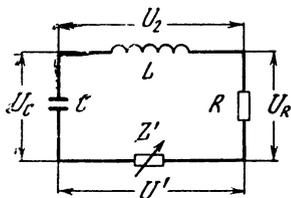


Рис. 61

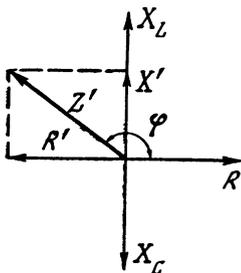


Рис. 62

Рассмотрим теперь случай произвольного угла φ . Пусть значение φ заключено между $\pi/2$ и π (рис. 62). Тогда вносимое обратной связью сопротивление Z' комплексно и может быть выражено как $Z' = R' + jX'$, что с геометрической точки зрения означает построение двух проекций вектора Z' , как показано на рисунке. Если $|R'| > R$, то система самовозбудится. Что же касается X' , то, как видно из диаграммы, наличие X' равносильно увеличению индуктивного сопротивления X_L (или, что то же самое, уменьшению емкостного сопротивления X_C). Иначе говоря, при описываемых условиях обратная связь как бы увеличивает индуктивность или емкость. Следствием этого является изменение настройки контура и при возбуждении в системе колебаний они будут иметь меньшую частоту, чем резонансная частота $\omega_0 = 1/\sqrt{LC}$ самого контура. Если $\pi < \varphi < 3\pi/2$, то обратная связь будет повышать частоту колебаний. Отсюда следует, что во избежание влияния обратной связи на частоту нужно позаботиться, чтобы в генераторе фаза обратной связи была точно равна π ($\varphi = 180^\circ$; «чистая» положительная обратная связь)¹. С другой же стороны, все описанные соотношения можно использовать для построения схемы частотной модуляции. Если фаза лежит в пределах

¹ В генераторе по схеме рис. 54 это условие выполнено. Если схема не генерирует, то первое, что нужно сделать, — это поменять концы у одной из катушек (не меняя ее положения), так как возможно, что катушка включена неправильно, и вместо положительной обратной связи получается отрицательная.

— $\pi/2 < \varphi < \pi/2$, то влияние обратной связи на настройку сохраняется, но автоколебания уже невозможны, так как составляющая R' дает отрицательную обратную связь, т. е. увеличивает результирующее сопротивление контура.

§ 19. Инженерные критерии устойчивости

Вопрос о возможности возникновения автоколебаний в некоторой схеме интересует нас в радиотехнике не только с точки зрения построения генераторов. Генератор есть источник колебаний; автоколебания являются основной его функцией. Другие же звенья радиотехнических устройств, как, например, усилители, преобразователи частоты, детекторы и т. п. работают в режиме вынужденных колебаний, и возникновение в этих звеньях автоколебаний в большинстве случаев не только нежелательно, но совершенно недопустимо. Поэтому вопрос о возможности возникновения автоколебаний в данной схеме имеет в радиотехнике довольно широкое значение. Вопрос этот ставится как вопрос об устойчивости в том смысле, что схема устойчива, если в нормальном для нее режиме автоколебания в ней не возникают, и наоборот. Таким образом, всякий генератор¹ принципиально должен быть неустойчивым, а все остальные схемы, назначением которых не является генерирование колебаний, должны быть устойчивыми. Следует пояснить, что речь идет здесь об устойчивости начального состояния системы, а не об устойчивости уже существующих автоколебаний — см. по этому поводу § 16.

Общий вопрос об устойчивости, т. е. о том, могут или нет в данной системе возбудиться автоколебания, решается на основании некоторых общих признаков, выражаемых с помощью так называемых *критериев устойчивости*. Существует несколько критериев устойчивости, по-разному отображающих единые свойства системы. В основе всех известных критериев лежит исследование корней характеристического уравнения системы.

Дело в том, что решение обыкновенного дифференциального уравнения с постоянными коэффициентами

$$\sum_{k=0}^n a_k \frac{d^k y}{dt^k} = 0$$

имеет вид

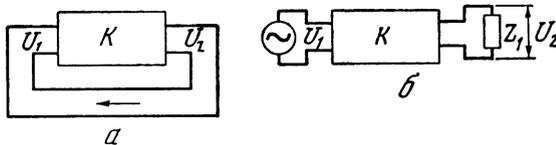
$$y(t) = \sum_{k=1}^n A_k e^{\gamma_k t},$$

где A_k — постоянные интегрирования, определяемые из начальных условий, а γ_k — корни характеристического уравнения.

¹ В этом параграфе здесь и дальше имеется в виду генератор с мягким возбуждением.

Система будет устойчивой, если вещественные части всех корней отрицательны. Физически это означает, что все колебания, возникшие в системе по тем или иным причинам, будут затухающими. Возможно, не решая характеристического уравнения, т. е. не находя корней, установить, выполняется ли вышеуказанное требование. Это можно сделать при помощи критерия Рауса-Гурвица, основанного на свойствах определителей, составленных из коэффициентов a_k .

Критерий Рауса-Гурвица имеет чисто аналитический характер. Существуют и другие критерии, которым придана форма



Р и с. 63

графических построений, что делает их более пригодными для инженерной практики. Мы приведем здесь без доказательства описание двух наиболее известных и распространенных критериев устойчивости, а именно критерия Найквиста и критерия Михайлова.

Начнем с некоторых предварительных соображений. Положим, что имеется некоторая схема, например усилитель, изображенный на рис. 63, *а* в виде четырехполюсника, и пусть вход и выход усилителя соединены между собою накоротко, как показано на рисунке, так что имеется обратная связь с выхода на вход. При таких условиях вполне возможно, что усилитель загенерирует, т. е. что возбудятся автоколебания. Заметим, что если в *замкнутой* схеме рис. 63, *а* действительно существуют установившиеся автоколебания, то переменное напряжение на выходе в точности равно переменному напряжению на входе по той простой причине, что вход и выход непосредственно соединены между собою. Если схема генерирует синусоидальные колебания, то равенство напряжений означает равенство комплексных амплитуд. Иначе говоря, амплитуды напряжений равны, а фазовый сдвиг равен нулю. Это условие известно под названием условия *баланса фаз и амплитуд*.

Разорвем теперь цепь обратной связи и попытаемся воспроизвести в разомкнутой схеме такой же режим, который имелся в замкнутой схеме. Для этого нужно подать на вход переменное напряжение от независимого генератора. На выходном же конце нужно нагрузить четырехполюсник сопротивлением, равным входному сопротивлению четырехполюсника. Если теперь подать от генератора напряжение той же частоты, которую схема генерировала при наличии обратной связи, то прежние условия работы будут полностью сохранены и окажется, что $\dot{U}_1 = \dot{U}_2$ (\dot{U} означает ком-

плексную амплитуду напряжения). Теперь предположим, что частота генератора в схеме рис. 63, б плавно изменяется от 0 до ∞ . Для каждого значения частоты можно определить комплексное отношение

$$K(\omega) = \dot{U}_2 / \dot{U}_1.$$

Это отношение в случае усилителя выражает не что иное, как коэффициент усиления (по напряжению); в случае же про-

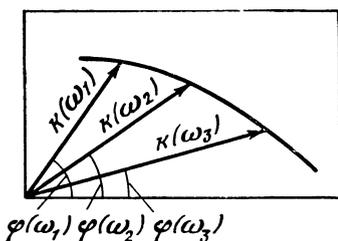


Рис. 64

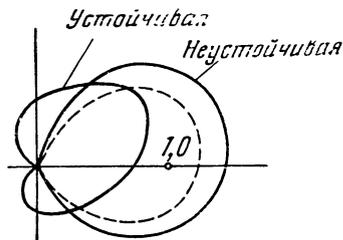


Рис. 65

извольного четырехполюсника величину K называют коэффициентом передачи. Коэффициент передачи, представляющий собою отношение комплексных амплитуд, есть комплексная величина, которую можно записать в виде

$$K(\omega) = ke^{j\varphi},$$

где k — вещественное отношение амплитуд, а φ — фазовый сдвиг между напряжениями на входе и на выходе. Условие баланса фаз и амплитуд выражается через эти две величины двумя равенствами: $k=1$, $\varphi=0$. Важно заметить, что, во-первых, эти равенства являются *точными*, а во-вторых, в установившемся режиме *оба* равенства обязательно выполняются.

Будем теперь строить K как вектор на комплексной плоскости и следить за траекторией конца этого вектора при изменении частоты. Эту траекторию называют амплитудно-фазовой характеристикой. Построение показано на рис. 64. Во многих случаях коэффициент передачи при $\omega=0$ и при $\omega=\infty$ равен нулю. В этих случаях амплитудно-фазовая характеристика образует замкнутый контур; точки начала и конца обхода контура совпадают с началом координат. Каждая точка амплитудно-фазовой (а-ф.) характеристики отвечает определенному значению частоты.

Если *установившиеся* автоколебания возможны, то а.-ф. характеристика проходит через точку с полярными координатами $k=1$, $\varphi=0$ — так называемая «точка (1, 0)», отвечающая условию баланса фаз и амплитуд. При этом будут генерироваться колебания с частотой, отвечающей положению вектора k вдоль вещественной оси ($\varphi=0$).

Если при $\varphi=0$ $k < 1$, то это означает, что колебания в системе могут быть только затухающими. Если же $k > 1$, то это значит, что колебания не только возникнут, но будут нарастать. Отсюда и следует критерий Найквиста: замкнутая система будет неустойчивой, т. е. в ней возникнут автоколебания, если а.-ф. характеристика разомкнутой системы охватывает точку (1, 0).

На рис. 65 показаны примеры а.-ф. характеристик для устойчивой и неустойчивой систем. Амплитудно-фазовая характеристика может быть снята экспериментально по схеме рис. 63, а либо вычислена, если схема и параметры четырехполюсника известны.

Амплитудно-фазовые характеристики не только могут дать ответ на вопрос об устойчивости системы, но позволяют также определить частоту и амплитуду установившихся колебаний. Мы не затрагивали до сих пор вопроса о том, при каких амплитудах входного напряжения снимается а.-ф. характеристика. Если бы исследуемый четырехполюсник был линейным, то коэффициент передачи K не зависел бы от амплитуды. При таких условиях $k > 1$ при $\varphi=0$ означало бы неограниченное нарастание амплитуды автоколебаний, так как *установившийся* режим с неизменной амплитудой может быть только при $k=1$. В действительности всякая система нелинейна, а потому по мере увеличения амплитуды коэффициент передачи убывает. Контур а.-ф. характеристики при этом *стягивается* (штрих на рис. 65) и если первоначально точка (1, 0) была охвачена этим контуром ($k > 1$), то рано или поздно контур пройдет через точку (1, 0). Амплитуда, при которой это случится, и есть установившееся значение амплитуды. Одновременно определяется и значение частоты, генерируемой системой в установившемся режиме. Теперь мы можем уточнить вышеприведенную формулировку критерия Найквиста в том смысле, что если речь идет об устойчивости системы в состоянии покоя (т. е. когда колебаний еще нет), то а.-ф. характеристика должна сниматься при самых малых амплитудах входного напряжения.

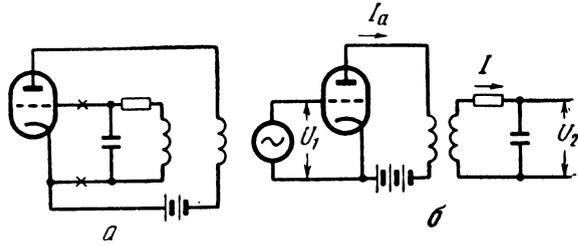
В качестве примера выведем с помощью критерия Найквиста условие самовозбуждения для генератора с контуром в цепи сетки, уже рассмотренного выше (см. § 15). На рис. 66, а дана замкнутая схема генератора, на рис. 66, б — та же схема в разомкнутом виде (цепь обратной связи разорвана в точках, отмеченных на рис. 66, а крестиками). Найдем коэффициент передачи для схемы рис. 66, б. Мы имеем $I_a = S U_1$, $e = j \omega M I_a$,

$$I = e / [R + j(\omega L - 1/\omega C)], \quad U_2 = I / j\omega C,$$

откуда

$$K = \frac{U_2}{U_1} = \frac{MS}{C} \cdot \frac{1}{R + j(\omega L - 1/\omega C)}. \quad (1)$$

Р и с. 66



Найдем модуль и аргумент этого выражения

$$k = |K| = \frac{MS}{C} \frac{1}{\sqrt{R^2 + (\omega L - 1/\omega C)^2}}, \quad (2)$$

$$\varphi = \arctg \frac{1/\omega C - \omega L}{R}. \quad (3)$$

Уравнение а.-ф. характеристики в полярных координатах имеет вид

$$k = \frac{MS}{RC} \cdot \frac{1}{\sqrt{1 + \operatorname{tg}^2 \varphi}} = \frac{MS}{RC} \cos \varphi. \quad (4)$$

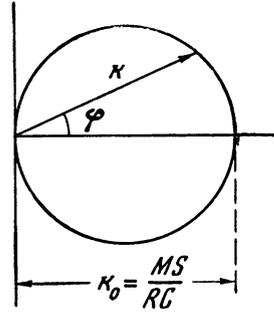
Таким образом, график а.-ф. характеристики в рассматриваемом случае имеет вид окружности, проходящей через начало координат и расположенной симметрично относительно вещественной оси (рис. 67). Диаметр окружности равен $k_0 = MS/RC$. Легко видеть, что условие охвата точки $(1, 0)$ выполняется при

$$k_0 = \frac{MS}{RC} \geq 1, \quad (5)$$

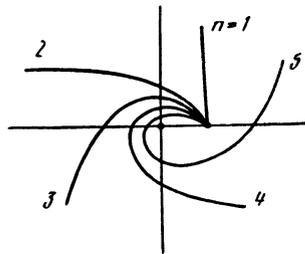
а это и есть условие самовозбуждения, полученное другим путем в § 17. Что касается генерируемой частоты, то для ее определения нужно приравнять φ нулю. Это дает (см. (3))

$$\omega = \omega_0 = 1/\sqrt{LC}.$$

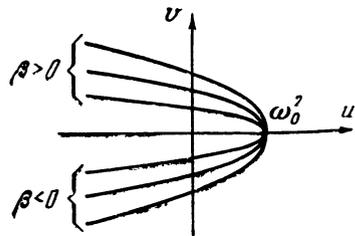
В более общем случае мы будем иметь нелинейное выражение,



Р и с. 67



Р и с. 68



Р и с. 69

в котором k будет функцией как частоты, так и амплитуды. В этом случае, составляя два уравнения, выражающие условие баланса фаз и амплитуд, т. е. $k=1$, $\varphi=0$, мы найдем из них две искомые величины: амплитуду и частоту установившихся колебаний.

Перейдем к критерию Михайлова. Его отличие от критерия Найквиста состоит прежде всего в том, что он приспособлен для исследования устойчивости замкнутых систем. Составляется уравнение данной системы с учетом обратной связи в операторной форме и приводится к виду

$$D(p)x = 0,$$

где x — колебательная величина (ток, напряжение), относительно которой составлено уравнение. Затем подстановкой $p=j\omega$ образуется комплексная величина $D(j\omega)$. Эта величина представляется вектором на комплексной плоскости. Частоте придаются все значения от 0 до ∞ . При этом конец вектора D описывает траекторию, называемую годографом Михайлова. Вид этого годографа и определяет устойчивость системы на основании следующего правила: для того чтобы система была устойчива, необходимо и достаточно, чтобы при изменении частоты от 0 до ∞ вектор D , беря начало на положительной вещественной полуоси, вращался только против часовой стрелки, нигде не обращаясь в нуль; при этом для системы, описываемой уравнением n -го порядка, вектор поворачивается при изменении частоты в указанном интервале на угол $n\pi/2$.

Это правило и есть формулировка критерия Михайлова.

На рис. 68 изображены годографы для устойчивых систем различных порядков.

Покажем применение критерия Михайлова на примере, в качестве которого возьмем все тот же генератор по схеме рис. 66, а. Для этого генератора в § 17 было выведено следующее уравнение:

$$\frac{d^2U}{dt^2} + \frac{1}{L} \left(R - \frac{MS_0}{C} \right) \frac{dU}{dt} + \omega_0^2 U = 0,$$

или в операторной форме

$$(p^2 + 2\beta p + \omega_0^2) U = D(p) U = 0,$$

где для сокращения обозначено

$$2\beta = \frac{1}{L} \left(R - \frac{MS_0}{C} \right).$$

По физическому смыслу β есть результирующий (т. е. с учетом действия обратной связи) показатель затухания. Заменяя p на $j\omega$, получим

$$D(j\omega) = \omega_0^2 - \omega^2 + j2\omega\beta = u + jv.$$

Уравнение годографа удобнее в данном случае выразить в прямоугольных координатах u и v . Мы имеем $u = \omega_0^2 - \omega^2$, $v = 2\beta\omega$. Исключая ω , получим $u = \omega_0^2 - v^2/4\beta^2$. Это есть уравнение пара-

болы. Мы рассматриваем только одну ветвь параболы, отвечающую положительным значениям ω . Учитывая это, получим для различных значений семейство кривых, представленное на рис. 69. Кривые, лежащие выше вещественной оси, удовлетворяют условию устойчивости; они соответствуют положительным значениям β . Система теряет устойчивость при $\beta=0$, т. е. при $R=MS_0/C$.

Кривые, лежащие ниже оси, соответствуют уже отрицательным значениям β ; применяя к ним вышеуказанное правило, видим, что эти кривые выражают неустойчивые состояния системы, так как вектор D поворачивается (при следовании его конца по кривой) по часовой стрелке.

§ 20. Установившийся режим генератора

В этом параграфе мы будем исследовать установившийся режим лампового генератора, пользуясь элементарной квазилинейной теорией.

Сущность квазилинейной теории генератора состоит в том, что нелинейность, ограничивающая амплитуду автоколебаний, учитывается в уравнении генератора особым образом, а именно: в качестве одного из параметров в уравнение вводится *средняя крутизна* триода, которая есть не постоянная величина, а *функция амплитуды* колебаний. Однако для неизменной амплитуды, т. е. для незатухающих и не нарастающих колебаний средняя крутизна также остается неизменной. Отсюда следует, что в *установившемся* режиме со средней крутизной можно обращаться как с постоянной величиной, а стало быть, уравнение генератора принимает форму линейного уравнения с постоянными коэффициентами. Отсюда и названия «квазилинейная (как бы линейная) теория», «квазилинейное уравнение». Из сказанного следует, что квазилинейная теория приспособлена в первую очередь для исследования установившихся режимов. Впрочем с ее помощью можно разрешать и более сложные вопросы.

Мы будем вычислять методами квазилинейной теории установившуюся амплитуду колебаний в ламповом генераторе. Объектом исследования послужит опять генератор с контуром в цепи сетки. Составим уравнение напряжений для контура

$$L \frac{dI}{dt} + RI + \frac{1}{C} \int Idt = M \frac{dI_a}{dt}. \quad (1)$$

Предположим, что напряжение и ток в контуре синусоидальны. Это — обычное для квазилинейной теории предположение. Оно означает, что мы пренебрегаем влиянием гармоник наводимой в контуре э. д. с., т. е. влиянием гармоник анодного тока. Практически это предположение оправдывается тем, что контур генератора, обладая большой добротностью, а следовательно, высокой избирательностью, отфильтровывает гармоники. Таким образом, мы можем в правой части заменить анодный ток его первой гар-

моникой. После этого возможен переход к комплексной форме уравнения (1) путем замены дифференцирования по времени умножением на $j\omega$

$$(j\omega L + R + 1/j\omega C) I = j\omega M I_{a_1},$$

или, вводя

$$\dot{U} = \frac{1}{j\omega C} I,$$

$$\left(\omega_0^2 - \omega^2 + j\omega \frac{R}{L}\right) \dot{U} = j\omega M I_{a_1}. \quad (2)$$

Здесь I_{a_1} — амплитуда первой гармоники анодного тока. Воспользуемся определением средней крутизны

$$S_0 = |I_{a_1} / \dot{U}| = I_{a_1} / \dot{U}$$

и получим

$$\omega_0^2 - \omega^2 + j\omega \frac{1}{L} \left(R - \frac{MS_0}{C}\right) = 0.$$

Это равенство может рассматриваться как уравнение частоты. Разрешая его относительно частоты, найдем, что частота комплексна. Известно, что это значит: комплексное значение частоты выражает нарастающие или затухающие по экспоненциальному закону колебания. Но мы рассматриваем установившийся режим, при котором амплитуда колебаний неизменна. В таком режиме частота может иметь только вещественное значение. Значит мнимый член в выражении (2) должен равняться нулю. Отсюда сразу находим частоту установившихся колебаний

$$\omega = \omega_0$$

при условии

$$R - MS_0/C = 0. \quad (3)$$

Это последнее соотношение очень похоже по внешнему виду на формулу (6) § 17, но имеет совершенно другой смысл. *Неравенство*

$$R - MS_0/C \leq 0,$$

выведенное в § 17, выражает условие самовозбуждения. Входящая в него величина S_0 есть *начальная крутизна*, или крутизна в рабочей точке, определяемая как

$$S_0 = (dI_a/dU)_{U=E_0},$$

где E_0 — абсцисса рабочей точки на характеристике лампы. Выражение же (3) представляет собою *равенство*, обязательно выполняемое при установившихся колебаниях с конечной амплитудой. Входящая в (3) величина S_0 есть *средняя крутизна*, зависящая не только от выбора рабочей точки, но и от амплитуды. Но если S_0 есть функция амплитуды, то (3) может рассматриваться как урав-

нение, решая которое мы можем найти установившуюся амплитуду. Именно это мы сейчас и сделаем.

Первый шаг состоит в том, что мы выбираем подходящую аппроксимацию характеристики лампы. Для наших целей достаточно представить характеристику многочленом пятой степени

$$I_a = a_0 + a_1 U + a_3 U^3 + a_5 U^5. \quad (4)$$

Для получения дифференциальной крутизны

$$S = dI_a/dU$$

достаточно было бы продифференцировать (4) по U . Но нам нужно найти среднюю крутизну. Положим для этого $U = U_m \cos \omega t$. Тогда

$$I_a = a_0 + a_1 U_m \cos \omega t + a_3 U_m^3 \cos^3 \omega t + a_5 U_m^5 \cos^5 \omega t.$$

Пользуясь табл. 1 степеней тригонометрических функций (стр. 400), получим

$$I_a = a_0 + a_1 U_m \cos \omega t + \frac{1}{4} a_3 U_m^3 (3 \cos \omega t + \cos 3\omega t) + \\ + \frac{1}{16} a_5 U_m^5 (10 \cos \omega t + 5 \cos 3\omega t + \cos 5\omega t).$$

Отсюда амплитуда первой гармоники (т. е. коэффициент при $\cos \omega t$)

$$I_{a_1} = a_1 U_m + \frac{3}{4} a_3 U_m^3 + \frac{5}{8} a_5 U_m^5.$$

Средняя крутизна по определению

$$S_e = \frac{I_{a_1}}{U_m} = a_1 + \frac{3}{4} a_3 U_m^2 + \frac{5}{8} a_5 U_m^4. \quad (5)$$

Итак, мы получили выражение для средней крутизны как функции амплитуды. Из (3) получаем

$$S_e - \frac{RC}{M} = 0, \quad (6)$$

или

$$\frac{5}{8} a_5 U_m^4 + \frac{3}{4} a_3 U_m^2 + a_1 - \frac{RC}{M} = 0. \quad (7)$$

Это есть биквадратное уравнение, легко разрешаемое относительно интересующей нас амплитуды U_m .

Не останавливаясь на этом, представим уравнение (6) в графической форме. Будем откладывать оба члена левой части (6), т. е. величины S_e и RC/M по оси ординат, а амплитуду напряжения U_m — по оси абсцисс. RC/M не зависит от U_m , это — постоянная величина, график которой есть горизонтальная прямая. Прямую эту мы будем называть прямой обратной связи, так как ее расположение зависит от глубины обратной связи, выражаемой через коэффициент взаимоиндукции M . Средняя крутизна зависит от

амплитуды, и график этой зависимости будет выражаться некоторой кривой, форма которой зависит от знаков и величин коэффициентов a_k . Положим для начала, что оба коэффициента a_3 и a_5 отрицательны. Тогда средняя крутизна, имеющая при $U_m=0$ значение $S_c=a_1=S_0$, с увеличением U_m монотонно убывает. График для этого случая показан на рис. 70. Если прямая обратной связи лежит выше кривой S_c (штрих на рис. 70), то равенство (6) не выполняется ни при каких значениях амплитуды, и автоколебания вообще невозможны. Если же увеличить обратную связь, то прямая обратной связи опустится и будет пересекаться с кри-

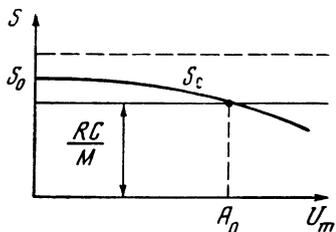


Рис. 70

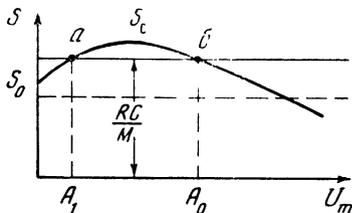


Рис. 71

вой S_c в некоторой точке, как показано на рис. 70. Точка пересечения отвечает равенству (6); абсцисса этой точки непосредственно дает установившуюся амплитуду A_0 . Если обратная связь будет увеличена, то прямая обратной связи опустится еще ниже, точка пересечения сдвинется вправо. Это означает, что амплитуда установившихся колебаний увеличится. Если же мы будем уменьшать обратную связь, то точка пересечения будет двигаться влево, пока не дойдет до оси ординат. При этом $S_c=S_0$, и это положение соответствует границе области самовозбуждения; амплитуда при этом обращается в нуль; при дальнейшем уменьшении обратной связи условие самовозбуждения не будет уже выполнено и автоколебания не возникнут.

Вся описанная картина относится к *мягкой* системе. В самом деле, так как $S_c < S_0$, то, если выполнено условие самовозбуждения $RC/M - S_0 \leq 0$, обязательно возникают автоколебания с вполне определенной установившейся амплитудой.

Рассмотрим теперь другой случай: пусть $a_3 > 0$, $a_5 < 0$. При малых амплитудах квадратичный член (см. формулу (5)) преобладает, а потому вначале S_c *растет* с увеличением амплитуды. В дальнейшем же член, содержащий четвертую степень амплитуды, а потому растущий быстрее квадратичного, оказывает решающее влияние, и средняя крутизна, достигнув максимума, начинает убывать. Кривая средней крутизны для рассматриваемого случая показана на рис. 71. Пусть прямая обратной связи занимает положение, показанное на рисунке сплошной прямой. Условие самовозбуждения $S_0 > RC/M$ при этом не выполнено, и автоколебания не возникнут. Однако имеются две точки пересе-

чения кривой S_c с прямой обратной связи. Для обсуждения состояния системы в этих точках введем термины *избыток* и *недостаток крутизны*, относящиеся к неравенствам $S_c \geq RC/M$. Эти термины имеют прямую связь с энергетическими соотношениями: избыток крутизны ($S_c > RC/M$) означает, что система получает энергии больше, чем теряет, а недостаток крутизны ($S_c < RC/M$) — что, наоборот, расход энергии больше прихода.

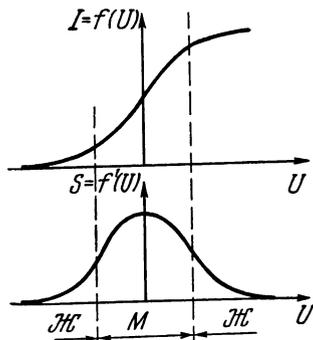
Рассмотрим состояние системы в точке a . Если амплитуда по какой-либо причине больше U_1 , то будем иметь избыток крутизны, и амплитуда будет возрастать и далее. Если же амплитуда станет меньше U_1 , то вследствие недостатка крутизны она будет убывать и далее, пока колебания не затухнут. Таким образом, точка a есть неустойчивая точка. В точке же b увеличение амплитуды вызывает недостаток крутизны, а уменьшение амплитуды — избыток крутизны. Таким образом, точка b есть устойчивая точка; абсцисса этой точки определяет установившуюся амплитуду. Графики рис. 70 изображают свойства жесткой системы: (для того чтобы получить установившиеся колебания, нужно дать системе такой начальный толчок, чтобы возникли колебания с амплитудой, большей U_1 . Дальнейшее увеличение амплитуды вплоть до установившегося значения U_0 произойдет автоматически. Графики рис. 70 и 71 изображают по существу те же энергетические соотношения, что и графики рис. 56 и 57. Если увеличить обратную связь настолько, что будет выполнено условие самовозбуждения (штриховая прямая на рис. 71), то система делается мягкой независимо от формы кривой средней крутизны.

Можно получить как мягкий, так и жесткий режим при одном и том же триоде путем соответствующего выбора рабочей точки. Если рабочая точка выбрана в точке перегиба характеристики

$$I_a = f(U_g),$$

то средняя крутизна будет монотонно убывать с увеличением амплитуды (рис. 70). Если же взять рабочую точку на загибе характеристики, то средняя крутизна будет сначала возрастать, а затем убывать. Аналитически условие получения того или иного режима выражается знаком коэффициента при члене третьей степени в разложении характеристики в степенной ряд. Если обозначить напряжение смещения через E_0 (абсцисса рабочей точки), то имеем по формуле Тейлора

$$I = f(E_0 + U) = f(E_0) + f'(E_0) \frac{U}{1!} + f''(E_0) \frac{U^2}{2!} + f'''(E_0) \frac{U^3}{3!} + \dots$$



Р и с. 72

Таким образом,

$$a_3 = \frac{1}{3!} f'''(E_0),$$

т. е. коэффициент a_3 (см. (4)), знак которого определяет мягкость или жесткость системы, пропорционален третьей производной функции $f(U)$ в рабочей точке $U=E_0$.

Равенство $f'''(E_0)=0$ определяет границу между мягким и жестким режимами. Но при $f'''=0$ имеем максимум или минимум f'' , что соответствует точке перегиба f' . С другой же стороны, $f'(U_g)$ есть не что иное, как дифференциальная крутизна. Итак, границы между мягким и жестким режимом проходят через точки перегиба характеристики крутизны

$$S = f'(U_g) = dI_a/dU_g,$$

как показано на рис. 72 вертикальными прямыми.

§ 21. Установление колебаний в генераторе

После рассмотрения установившегося режима генератора переходим к процессу установления: мы должны проследить развитие колебательного явления от момента включения схемы и до того времени, когда амплитуда и частота практически уже не будут меняться, т. е. когда можно считать колебание уже установившимся.

Задача сводится к тому, чтобы проследить изменение во времени величин, характеризующих колебание, т. е. в первую очередь амплитуды, а затем и частоты. Форму колебаний мы будем полагать близкой к синусоидальной.

Задача будет решаться методом, известным под названием *метода медленно меняющихся амплитуд*. По существу своему этот метод относится к квазилинейным; возможность его применения к данной задаче основана на том, что амплитуда, хотя и не постоянная, как при установившемся режиме, предполагается *медленно* изменяющейся. Критерий «медленности» будет дан ниже.

Запишем уравнение генератора в виде

$$L \frac{dI}{dt} + RI + \frac{1}{C} \int I dt = M \frac{dI_a}{dt}. \quad (1)$$

Этим уравнением мы уже пользовались в §§ 17 и 20. Но в § 17 мы полагали $I_a = S_0 U$ и свели, таким образом, уравнение (1) к линейному уравнению, имея в виду найти условие самовозбуждения. В § 20 мы составили уравнение для первой гармоники в комплексной форме, введя в уравнение среднюю крутизну; наша задача состояла в исследовании установившегося режима. Теперь же мы придадим уравнению (1) общую нелинейную форму.

Зависимость анодного тока от напряжения на сетке представляется нелинейной функцией

$$I_a = f(U).$$

Запишем правую часть (1) в виде

$$M \frac{dI_a}{dt} = M \frac{dI_a}{dU} \cdot \frac{dU}{dt}.$$

Но

$$dI_a/dU = df(U)/dU = S(U)$$

есть дифференциальная крутизна, выражаемая графически наклоном касательной к характеристике триода в данной точке (т. е. в точке с абсциссой U). Вводя, как и прежде,

$$U = \frac{1}{C} \int Idt,$$

получим вместо (1) исходное уравнение генератора в виде

$$\frac{d^2U}{dt^2} + \frac{1}{L} \left[R - \frac{M}{C} S(U) \right] \frac{dU}{dt} + \omega_0^2 U = 0. \quad (2)$$

Положим, что рассматривается мягкая система. В этом случае нелинейная характеристика может аппроксимироваться степенным многочленом третьей степени

$$I_a = a_0 + a_1 U + a_3 U^3.$$

Дифференцируя, находим выражение для крутизны

$$S = dI_a/dU = a_1 + 3a_3 U^2.$$

Подставляя это в (1), получаем нелинейное уравнение мягкой системы в виде

$$d^2U/dt^2 + (2\beta_0 + \gamma U^2) dU/dt + \omega_0^2 U = 0. \quad (3)$$

Здесь $\beta_0 = a - \omega_0^2 M S_0 / 2$ — начальный показатель затухания (с учетом действия обратной связи); γ — просто сокращенное обозначение $\gamma = 3 \omega_0^2 M a_3$.

Способ решения уравнения (3) состоит в том, что мы предполагаем решение в виде

$$U = A(t) \sin \omega t. \quad (4)$$

Величина A есть переменная «амплитуда» колебания; функция $A(t)$ выражает то, что в радиотехнике называется *огibaющей*. Что касается неизвестной пока частоты ω , то мы будем считать ее постоянной¹.

Продифференцируем предполагаемое решение (3) дважды

$$U' = A' \sin \omega t + A \omega \cos \omega t,$$

$$U'' = A'' \sin \omega t + 2A' \omega \cos \omega t - A \omega^2 \sin \omega t$$

¹ Это лишь первое приближение. В действительности в процессе установления меняется как амплитуда, так и частота колебаний. Но изменения частоты обычно невелики; учтя эти изменения, мы получим во втором приближении лишь незначительную поправку на закон изменения амплитуды.

и подставим значения U и его производных в (2)

$$A' \sin \omega t + 2A' \omega \cos \omega t - A \omega^2 \sin \omega t + \\ + (2\beta_0 + \gamma A^2 \sin^2 \omega t)(A' \sin \omega t + A \omega \cos \omega t) + \omega_0^2 A \sin \omega t = 0.$$

Производя тригонометрические преобразования и приравнивая порознь нулю коэффициенты при $\sin \omega t$ и $\cos \omega t$, получим два уравнения

$$A'' + \left(2\beta_0 + \frac{3}{4} \gamma A^2\right) A' + (\omega_0^2 - \omega^2) A = 0, \\ 2A' + \left(2\beta_0 + \frac{1}{4} \gamma A^2\right) A = 0. \quad (5)$$

При этом члены, содержащие тройной аргумент ($3\omega t$), отброшены в силу исходного предположения о синусоидальности колебания (см. формулу (4)).

Уравнения (5) могут быть еще упрощены. Дело в том, что члены этих уравнений имеют различный порядок малости. Мы предполагаем, что функция $A(t)$, выражающая изменение амплитуды, есть *медленная* функция. Это значит, что изменения амплитуды совершаются значительно медленнее, чем само колебательное изменение напряжения. Функция $f_1(t)$ считается медленной по сравнению с какой-либо другой функцией $f_2(t)$, если ее производная $f_1'(t)$ много меньше производной $f_2'(t)$ (для сравнения нужно предварительно нормировать производные, например, разделив их на среднеквадратичные значения самих функций).

Производная функции, выражающей колебание, т. е. $U = A \sin \omega t$, не превосходит по абсолютной величине ωA . Производная же огибающей $A(t)$ обозначена A' , и условие медленности $A(t)$ может быть записано в виде $|A'| \ll \omega A$. Другими словами, A' есть величина первого порядка малости по сравнению с ωA . Но если так, то вторая производная функции $A(t)$ будет уже величиной второго порядка малости по сравнению со второй производной U , имеющей порядок $\omega^2 A$.

На основании этих соображений можно вычеркнуть из первого уравнения (5) член A'' , как имеющий второй порядок малости по сравнению с последним членом.

В результате получим вместо (5)

$$\left(2\beta_0 + \frac{3}{4} \gamma A^2\right) A' + (\omega_0^2 - \omega^2) A = 0, \\ 2A' + \left(2\beta_0 + \frac{1}{4} \gamma A^2\right) A = 0. \quad (6)$$

Уравнения, получаемые путем вычеркивания членов высших порядков малости, называются *укороченными* уравнениями. Укорачивание уравнений является общим приемом метода медленно меняющихся амплитуд. Он позволяет понизить порядок уравнений и упростить их строение. Впрочем в рассматриваемом случае эти

преимущества останутся неиспользованными, так как мы воспользуемся ниже только вторым из двух уравнений (6), не подвергшимся укорочению.

Уравнения (6) являются нелинейными, так как коэффициенты зависят от амплитуды. Найдем из этих уравнений прежде всего установившиеся амплитуду и частоту.

В установившемся режиме $A' = 0$, $A = A_0$. Первое уравнение дает сразу

$$\omega = \omega_0,$$

а из второго находим

$$A_0 = \sqrt{-8\beta_0/\gamma} \quad (7)$$

или, выражая β_0 и γ через исходные параметры,

$$A_0 = \frac{2}{\sqrt{3}} \sqrt{\frac{a_1}{a_3} - \frac{2\alpha}{\omega_0^2 M a_3}} = \frac{2}{\sqrt{3}a_3} \sqrt{S_0 - \frac{RC}{M}}. \quad (8)$$

Это то же значение, которое мы нашли бы из уравнения (7) § 20, положив $a_5 = 0$. Как видим, амплитуда колебаний определяется нелинейностью, выражаемой в нашем случае коэффициентом a_3 ; амплитуда зависит также от собственного затухания контура, начальной крутизны и глубины обратной связи.

Если присмотреться к формуле (8), то видно, что амплитуда тем меньше, чем больше нелинейность (т. е. чем больше коэффициент a_3); амплитуда тем больше, чем с большим запасом выполнено условие самовозбуждения, т. е. условие $S_0 > RC/M$.

Теперь найдем закон установления амплитуды, т. е. функцию $A(t)$, воспользовавшись вторым уравнением (6). Для решения этого уравнения умножим его почленно на A , после чего можно записать

$$\frac{dA^2}{dt} + \left(2\beta_0 + \frac{1}{4} \gamma A^2\right) A^2 = 0.$$

Разделим переменные

$$\frac{dA^2}{\left(2\beta_0 + \frac{1}{4} \gamma A^2\right) A^2} = -dt.$$

Разложим дробь в левой части на простейшие

$$\frac{dA^2}{A^2} - \frac{d\left(2\beta_0 + \frac{1}{4} \gamma A^2\right)}{2\beta_0 + \frac{1}{4} \gamma A^2} = -2\beta_0 dt.$$

Интегрируя, получим

$$\ln A^2 - \ln\left(2\beta_0 + \frac{1}{4} \gamma A^2\right) = -2\beta_0 t + C$$

или

$$\ln\left(\frac{1}{4}\gamma + 2\beta_0\frac{1}{A^2}\right) = 2\beta_0 t - C.$$

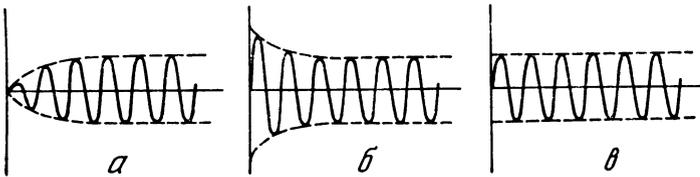
Отсюда

$$\frac{\gamma}{8\beta_0} + \frac{1}{A^2} = \frac{1}{2\beta_0} e^{-C} e^{2\beta_0 t} = B e^{2\beta_0 t}$$

и, наконец (см. (7)),

$$A(t) = A_0 \sqrt{1 + D e^{2\beta_0 t}}.$$

В этом выражении $D = BA_0^2$ есть постоянная интегрирования, определяемая начальными условиями. В зависимости от значе-



Р и с. 73

ния D амплитуда в первый момент может быть меньше или больше установившегося значения A_0 или даже в точности равняться ему (при $D=0$). Величина β_0 есть начальный результирующий показатель затухания. Величина эта во всяком случае отрицательна — в этом состоит условие самовозбуждения. Учитывая это, мы видим, что с течением времени экспоненциальный член под радикалом убывает, и мы получаем в пределе (при $t \rightarrow \infty$) установившееся значение амплитуды, выражаемое формулой (7).

На рис. 73 изображен процесс установления колебаний для различных начальных условий. Физический смысл тех или иных начальных условий легче раскрыть, опираясь на графическое представление процесса установления, как это будет сделано в § 23.

§ 22. Инерционная нелинейность

В генераторе, установившийся режим которого рассмотрен в § 20, рост амплитуды автоколебаний ограничивается нелинейностью характеристики триода $I_a = f(U_g)$. Эта характеристика выражает зависимость между мгновенными значениями анодного тока и напряжения на сетке. Поэтому даже при чисто синусоидальном напряжении на сетке переменная составляющая анодного тока принципиально несинусоидальна. Напряжение на контуре (а следовательно, и на сетке) практически синусоидально только за счет фильтрующего действия самого контура. Переход к квазилинейной

форме уравнения состоит в том, что мы составляем уравнение для первых гармоник тока и напряжения и вводим среднюю крутизну, определяемую как отношение амплитуд первых гармоник тока и напряжения.

Нелинейность в автоколебательной системе обязательна. Возможно, однако, построить генератор на другой основе. Возможность эта состоит во введении в состав автоколебательной системы нелинейного параметра, значение которого зависит не от мгновенных, а от действующих (или амплитудных) значений тока или напряжения. Такого рода нелинейность называется *инерционной*.

Примером устройства, обладающего инерционной нелинейностью, может служить *термистор* (термосопротивление). Сопротивление термистора зависит от его температуры, а температура определяется балансом мощностей — мощности, выделяемой в термисторе в форме джоулева тепла, и мощности, отдаваемой термистором в окружающую среду через посредство конвекции и излучения. Скорость процесса теплообмена между термистором и окружающей средой характеризуется постоянной времени, которая тем больше, чем больше теплоемкость термистора и чем меньше теплопроводность в окружающую среду. Постоянная времени имеет порядок единиц и даже десятков секунд, т. е. на много порядков превосходит период высокочастотного колебания. Поэтому если нагрев термистора производится током высокой частоты, то температура термистора на протяжении периода не может следовать за изменениями мгновенной мощности и остается постоянной с высокой степенью точности. Таким образом, можно принять, что сопротивление термистора есть функция *амплитуды* (а не мгновенного значения) тока или напряжения, причем важно, что это предположение может быть выполнено с любой степенью точности (так как технически легко увеличить постоянную времени до весьма значительных величин). А это в свою очередь означает, что получаемые автоколебания синусоидальны со столь же высокой степенью точности. Это является существенной особенностью систем с инерционной нелинейностью.

Рассмотрим теперь генератор, отличающийся от ранее рассмотренного тем, что триод предполагается идеально линейным, т. е. характеристика его предполагается линейной с постоянной крутизной S_0 . Необходимая для ограничения амплитуды нелинейность вводится в систему путем включения в контур в качестве активного сопротивления термистора с зависящим от температуры (а следовательно, от амплитуды колебания) сопротивлением R_t . Термисторы бывают металлические и полупроводниковые; у первых сопротивление с температурой растет, у вторых — падает; первые нужно включить в контур последовательно, вторые — параллельно. Будем в обоих случаях считать, что сопротивление R_t выражает эквивалентное последовательное сопротивление контура. В таком случае мы можем воспользоваться комплексной

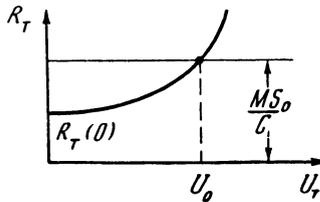
формой квазилинейного уравнения из § 20, переписав его в виде

$$\omega_0^2 - \omega^2 + j\omega \frac{1}{L} \left(R_T - \frac{MS_0}{C} \right) = 0.$$

Для генератора с термистором, описываемого этим соотношением, частота колебаний ω равна ω_0 при любой амплитуде, а следовательно, не меняется в процессе установления колебаний. Условие самовозбуждения запишется в виде

$$R_T(0) < MS_0/C,$$

где $R_T(0)$ — сопротивление термистора при отсутствии колебаний, определяемое только температурой окружающей среды.



Р и с. 74

Амплитуда в установившемся режиме находится из уравнения

$$R_T(U_m) = MS_0/C.$$

Характеристика термистора $R_T(U_m)$ должна, разумеется, быть известна. Уравнение можно решить графически, как показано на рис. 74. Возможность получения с помощью инерционной нелинейности синусоидальных установившихся автоколебаний имеет большое значение. Во-первых, во многих случаях важно иметь хорошую форму колебаний, свободную от гармоник. Во-вторых, наличие гармоник вызывает смещение частоты. Этот эффект тем заметнее, чем меньше добротность контура: применение инерционной нелинейности позволяет получить синусоидальные колебания и в случае контура с невысокой добротностью. Последнее особенно важно для так называемых RC -генераторов (см. § 27), цепи которых, составленные только из емкостей и сопротивлений, эквивалентны аperiодическим контурам.

§ 23. Фазовая плоскость

Обычное графическое представление колебательных явлений в какой-либо системе состоит в том, что все изменяющиеся во времени величины (например, токи, напряжения) выражаются соответствующими функциями времени. Графики этих функций — это осциллограммы.

Но можно поступить и иначе, а именно: взять в качестве прямоугольных координат мгновенные значения переменных величин, совокупность которых полностью определяет состояние системы.

Мы получим, таким образом, некоторое пространство, каждая точка которого представляет состояние системы в данный момент времени. Такое пространство называется *фазовым*.

Число измерений фазового пространства определяется порядком дифференциального уравнения системы. Известно, что для решения уравнения n -го порядка нужно задать в качестве начальных условий значения для начального момента функции и ее производных до $(n-1)$ -го порядка, т. е. всего n независимых величин. Нам придется иметь дело преимущественно с уравнениями второго порядка. В этом случае число измерений фазового пространства равно двум, т. е. фазовое пространство в этом случае есть *фазовая плоскость*. Колебательные явления в системе, описываемой уравнением второго порядка, могут быть представлены диаграммами на фазовой плоскости.

Рассмотрим, к примеру, свободные колебания в линейном контуре без потерь. Дифференциальное уравнение для этого случая имеет вид

$$L \frac{dI}{dt} + \frac{1}{C} \int I dt = 0 \quad (1)$$

или

$$\frac{d^2Q}{dt^2} + \omega_0^2 Q = 0. \quad (2)$$

Решение этого уравнения таково

$$Q = M \sin \omega_0 t + N \cos \omega_0 t, \quad (3)$$

где M и N — произвольные постоянные, определяемые начальными условиями. Для тока имеем

$$I = \frac{dQ}{dt} = M \omega_0 \cos \omega_0 t - N \omega_0 \sin \omega_0 t. \quad (4)$$

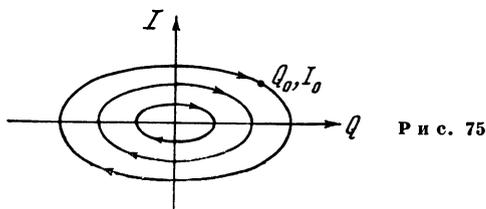
Возьмем ток и заряд в качестве координат фазовой плоскости и выразим одну величину непосредственно через другую, т. е. исключим параметр t . Для этого разделим (4) на ω_0 , возведем в квадрат и сложим с квадратом (3). Это даст уравнение

$$\frac{1}{\omega_0^2} I^2 + Q^2 = M^2 + N^2,$$

которое есть уравнение семейства эллипсов (рис. 75). Заметим, что, изменив масштаб, например, отложив по оси абсцисс $\omega_0 Q$ вместо Q , мы получим семейство окружностей. Каждая точка на одном из эллипсов изображает состояние системы в данный момент (так как координаты точки суть мгновенные значения тока и напряжения). Поэтому точка на диаграмме называется изображающей точкой. С течением времени точка перемещается, описывая линию, называемую фазовой траекторией. Если на фазовой плоскости изображен периодический процесс, то фазовая траектория будет

представлять собою замкнутый контур, обегаемый изображающей точкой за время, равное одному периоду. Семейство фазовых траекторий, отображающее возможные в данной системе колебательные явления, называют «фазовым портретом» данной системы. Так, рис. 75 является фазовым портретом линейного контура без потерь.

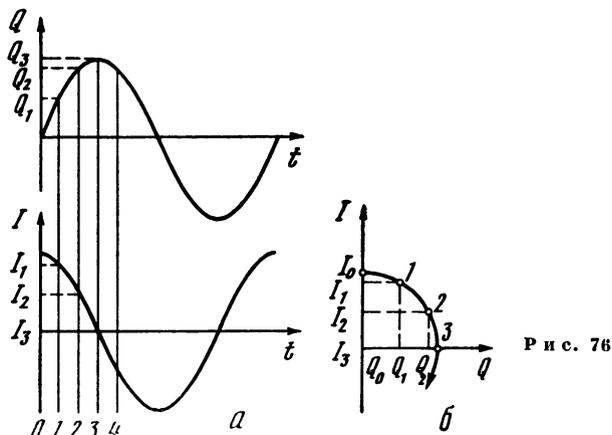
На рис. 75 выбраны координаты Q и $I = dQ/dt$. При таком выборе координат изображающая точка движется по часовой стрелке, как отмечено на рис. 75.



Р и с. 75

Мы получили фазовый портрет на рис. 75 аналитическим путем. Покажем, как производится построение фазового портрета по осциллограммам. На рис. 76, а изображены осциллограммы заряда и тока. Если теперь в моменты, отмеченные и занумерованные на рисунке, взять отсчеты мгновенных значений Q и I (т. е. ординаты соответствующих кривых) и перенести их в качестве координат изображающей точки на рис. 76, б, то можно построить фазовую траекторию. Построение подтверждает, что изображающая точка движется по часовой стрелке.

Поясним теперь роль начальных условий. Начальными условиями для уравнения второго порядка вида (2) являются, как известно, значения функции (Q) и ее первой производной ($I = dQ/dt$) в момент, принимаемый за начало отсчета времен (в частности,



Р и с. 76

в момент $t = 0$). Следовательно, задание начальных условий определяет исходные координаты изображающей точки. На рис. 75 свободное колебание представляется эллипсом, проходящим через точку, отображающую своими координатами Q_0 и I_0 начальные условия.

Обратимся к контуру с потерями. Уравнение такого контура имеет вид

$$\frac{d^2Q}{dt^2} + 2\alpha \frac{dQ}{dt} + \omega_0^2 Q = 0. \quad (5)$$

Решение этого уравнения при начальных условиях $Q(0) = 0$, $I(0) = I_0$ таково

$$Q = \frac{I_0}{\omega_1} e^{-\alpha t} \sin \omega_1 t, \quad (6)$$

где $\omega_1 = \sqrt{\omega_0^2 - \alpha^2}$ — собственная частота. Найдем ток

$$I = \frac{I_0}{\omega_1} e^{-\alpha t} (-\alpha \sin \omega_1 t + \omega_1 \cos \omega_1 t).$$

При малом показателе затухания α можно пренебречь первым членом в скобках и принять

$$I \cong I_0 e^{-\alpha t} \cos \omega_1 t. \quad (7)$$

Для исключения времени t из (6) и (7) представим зависимость между током и зарядом в полярных координатах. Запишем

$$\begin{aligned} I &= I_0 e^{-\alpha t} \cos \omega_1 t = \rho \cos \vartheta, \\ \omega_1 Q &= I_0 e^{-\alpha t} \sin \omega_1 t = \rho \sin \vartheta, \end{aligned}$$

т. е. положим

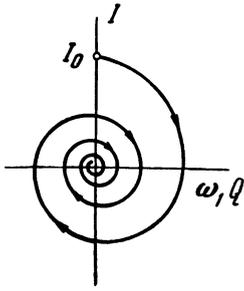
$$\rho = I_0 e^{-\alpha t}, \quad \vartheta = \omega_1 t,$$

откуда

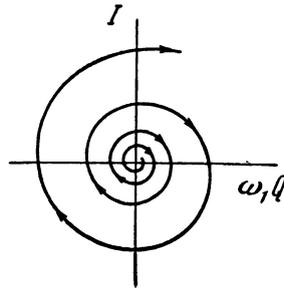
$$\rho = I_0 e^{-\alpha \vartheta / \omega_1}.$$

Это и есть уравнение фазовой траектории в полярных координатах, представляющее свертывающуюся логарифмическую спираль (рис. 77).

При других начальных условиях получим другие спирали. Все они будут вложены друг в друга и все будут сходиться в точке $I = Q = 0$. Каждая спираль изображает затухающее колебание. Семейство же спиралей с различными начальными условиями образует фазовый портрет контура с потерями. В заключение данного параграфа рассмотрим еще случай контура, в котором действием обратной связи затухание сделано отрицательным. Подобная система описывается уравнением, отличающимся от (5) только знаком при члене с первой производной заряда по времени. Решение уравнения представится формулами (6) и (7), а уравнение фазовой траектории — формулой (8), причем во всех трех формулах надо лишь заменить знак $(-)$ на $(+)$. Это означает, что мы



Р и с. 77



Р и с. 78

имеем колебания с экспоненциально нарастающей амплитудой, а фазовая траектория представляет собой раскручивающуюся спираль (рис. 78). Если бы система действительно была линейной, то амплитуда нарастала бы неограниченно; этому соответствует неограниченный рост радиуса спирали.

§ 24. Фазовые портреты генераторов

В предыдущем параграфе рассмотрены в качестве примеров линейные системы. Между тем нас интересуют явления в генераторах, являющихся заведомо нелинейными системами, так как в линейной системе автоколебания с установившейся амплитудой, как нам уже известно, невозможны.

Очередная задача состоит в построении фазовых портретов генераторов. Начнем со случая мягко возбуждающегося генератора, как более простого. Фазовый портрет должен дать полную картину возможных режимов генератора: установившегося режима и процессов установления колебаний при любых начальных условиях.

Мы не будем выводить уравнений фазовых траекторий; ограничимся качественным рассуждением. Для этого нужно прежде всего пояснить, как отображается на фазовых диаграммах то или иное значение затухания.

Мы имеем для простого контура уравнение спирали

$$\rho = I_0 e^{-\frac{\alpha}{\omega_1} \vartheta}.$$

Возьмем два значения радиуса ρ , соответствующих изменению ϑ на 2π , т. е. радиусы двух соседних витков спирали (рис. 79)

$$\rho_1 = I_0 e^{-\frac{\alpha}{\omega_1} \vartheta}, \quad \rho_2 = I_0 e^{-\frac{\alpha}{\omega_1} (\vartheta + 2\pi)},$$

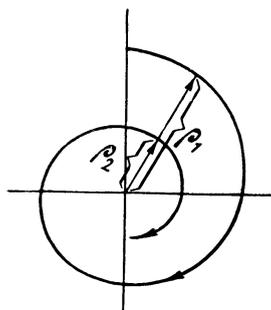
и составим отношение этих радиусов

$$\frac{\rho_1}{\rho_2} = e^{\frac{\alpha}{\omega_1} 2\pi} = e^{\alpha T} \cong e^{\pi d},$$

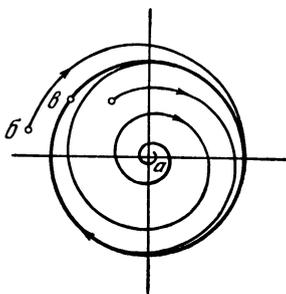
где d — затухание контура. Таким образом, радиус спирали изменяется от витка к витку в геометрической прогрессии¹, знаменатель которой зависит от затухания. То же относится и к раскручивающейся спирали.

Рассмотрим теперь явления в мягком генераторе. В начальной стадии самовозбуждения, когда колебания еще малы, явления могут с удовлетворительной точностью описываться линейным уравнением (2) § 17. Если условие самовозбуждения выполнено, то показатель затухания отрицателен, и процесс самовозбуждения представляется раскручивающейся спиралью, как на рис. 78. Однако по мере роста амплитуды колебаний начинает сказываться нелинейность. Ее влияние проявляется в том, что затухание становится меньше по абсолютной величине; радиус спирали замедляет свой рост. Наконец, достигается энергетический баланс. При этом затухание становится равным нулю, дальнейший рост амплитуды прекращается. Изображающая точка описывает замкнутый контур в виде эллипса (или окружности при соответствующем выборе масштабов). Это и есть установившийся режим генератора.

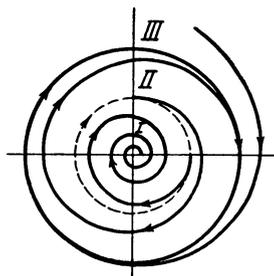
Если задать генератору начальные условия, представляемые некоторой точкой *вне* контура, соответствующего установившемуся режиму, то затухание будет положительным, и изображающая точка будет перемещаться по свертывающейся спирали, радиус которой будет убывать все медленнее, пока траектория не вернется на тот же контур установившегося режима. Все это изображено на рис. 80, представляющем собою фазовый портрет генератора с мягким возбуждением. Замкнутый контур, отображающий ус-



Р и с. 79



Р и с. 80



Р и с. 81

¹ Свойство логарифмической спирали. Для так называемой архимедовой спирали радиус изменяется по арифметической прогрессии.

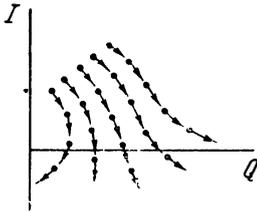
тановившееся периодическое колебание (жирная линия на рис. 80), называется в теории колебаний предельным циклом. Для фазового портрета на рис. 80 характерно, что любая фазовая траектория, где бы она ни начиналась, навертывается в конце концов на предельный цикл. Это означает, что всякое нарушение установившегося режима выравнивается автоматически после устранения вызвавшей это нарушение причины. При таких условиях предельный цикл называется устойчивым. Легко видеть теперь, что три случая установления колебаний, изображенные на рис. 73, соответствуют на рис. 80 положениям исходной точки фазовой траектории: a — в начале координат, c — на предельном цикле и b — вне его.

Обратимся к жесткому генератору. Этот случай отличается от только что рассмотренного тем, что при малых амплитудах затухание положительно, и фазовые траектории сворачиваются к началу координат. Существует значение амплитуды, при котором затухание первый раз обращается в нуль. Этому значению амплитуды соответствует свой предельный цикл. Однако цикл этот неустойчив. Фазовые траектории не навертываются на него, а, наоборот, сходя с предельного цикла, либо свертываются во внутрь к началу координат, либо развертываются в области вне неустойчивого предельного цикла. В этой области затухание отрицательно, а амплитуда растет, пока затухание снова не обратится в нуль. Здесь образуется второй предельный цикл, на этот раз уже устойчивый. Отличие устойчивого предельного цикла от неустойчивого с точки зрения фазовых диаграмм состоит в том, что фазовые траектории стремятся к устойчивому циклу и сбегают с неустойчивого. Все сказанное поясняется фазовым портретом жесткого генератора, изображенного на рис. 81. Фазовая плоскость может быть разбита на три области: I — внутри неустойчивого цикла, II — между устойчивым и неустойчивым циклами и III — вне устойчивого цикла. Автоколебания не возбуждятся, если начальные условия изображаются точкой в области I , и возбуждятся, если исходная точка лежит в области II или III . Разница между двумя последними случаями будет состоять в том, что в первом случае амплитуда будет устанавливаться, возрастая до установившегося значения, а во втором, — убывая до того же значения. Различие между мягким и жестким генераторами, представляемое рис. 80 и 81, ранее было пояснено рис. 70 и 71 и рис. 56 и 57. Полезно сравнить между собою все три пары рисунков, так как они по-разному отображают одни и те же соотношения.

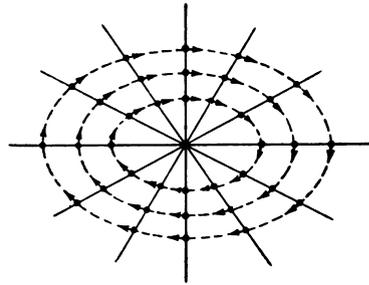
§ 25. Графические методы

Фазовые диаграммы позволяют просто и наглядно представить в графической форме различные колебательные явления, а также свойства тех или иных систем (пассивных цепей, автоколебательных систем).

Но роль фазовых диаграмм этим не ограничивается. Построения на фазовой плоскости служат не только в качестве иллюстраций — они составляют также основу графических методов решения нелинейных уравнений. При этом существенно, что графическое решение не требует упрощающих предположений, например предположения малой нелинейности, синусоидального решения и т. п.

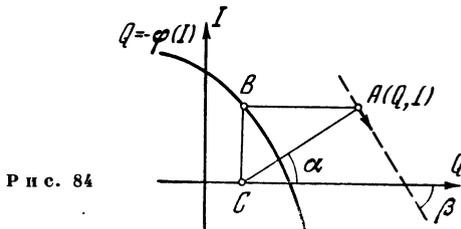


Р и с. 82



Р и с. 83

Когда говорится о решении нелинейной задачи путем графического построения на фазовой плоскости, то имеется в виду, что решение получается в виде фазового портрета. Фазовый портрет дает нам все необходимые сведения о поведении системы, описываемой данным нелинейным уравнением: на фазовом портрете видны все возможные режимы при любых начальных условиях; в частности, видны предельные циклы, если они существуют; по фазовому портрету можно судить об устойчивости предельных циклов: форма предельных циклов выражает форму колебаний и т. д.



Р и с. 84

Одним из наиболее универсальных графических методов является метод построения поля направлений. Он позволяет построить фазовый портрет системы, описываемой нелинейным дифференциальным уравнением второго порядка следующего вида:

$$Q'' + \varphi(Q') + \psi(Q) = 0. \quad (1)$$

Введем

$$I = dQ/dt.$$

Уравнение (1) переписется в виде

$$\frac{dI}{dt} + \varphi(I) + \psi(Q) = 0. \quad (2)$$

Разделим (2) на I

$$\frac{dI}{dQ} = -\frac{\varphi(I) + \psi(Q)}{I}. \quad (3)$$

В этом соотношении dI/dQ означает наклон касательной к фазовой траектории в данной точке, т. е. в точке с координатами Q и I . Беря любую точку, мы можем, зная ее координаты, вычислить правую часть (3) и определить таким образом направление фазовой траектории в данной точке. Вычислив значения dI/dQ для многих точек, мы можем построить на фазовой плоскости так называемое поле направлений, как показано на рис. 82. Ориентируясь на стрелочки, показывающие направление фазовой траектории в данной точке, можно построить семейство траекторий, т. е. фазовый портрет, с любой желаемой степенью точности, которая, очевидно, повышается с увеличением числа точек. Удобно вести построение, опираясь на линии, вдоль которых направление траекторий (т. е. наклон касательной к траектории) остается неизменным. Эти линии называются изоклинами (линиями равных наклонов). Уравнение изоклин имеет вид

$$\frac{\varphi(I) + \psi(Q)}{I} = \text{const.}$$

Поясним это на простейшем примере линейной системы, описываемой уравнением

$$dI/dt + \omega_0^2 Q = 0.$$

В этом случае

$$\varphi(I) = 0, \quad \psi(Q) = \omega_0^2 Q$$

и уравнение изоклин

$$Q/I = \text{const},$$

что представляет собою уравнение пучка прямых, проходящих через начало координат. Построение соответствующего фазового портрета показано на рис. 83 — это портрет контура без потерь.

В ряде случаев мы можем ограничиться уравнением менее общего вида, чем (1), а именно уравнением

$$Q'' + \varphi(Q') + Q = 0. \quad (4)$$

Именно к такой форме можно привести уравнение большинства генераторов, с которыми мы имеем дело в радиотехнике. Графическое решение уравнения (4) может быть выполнено без вспомогательных вычислений путем чисто графического построения,

известного под названием *построения Льенара*. Обоснование этого построения состоит в следующем. Мы имеем вместо (3)

$$\frac{dI}{dQ} = -\frac{\varphi(I) + Q}{I}. \quad (5)$$

Построим на фазовой плоскости кривую

$$Q = -\varphi(I) \quad (6)$$

(рис. 84). Из произвольной точки A с координатами (Q, I) проведем параллельно оси Q прямую до пересечения с кривой (6). Из точки пересечения B опускаем перпендикуляр на ось Q . Основание перпендикуляра обозначено через C . Мы имеем

$$\overline{BA} = Q + \varphi(I), \quad \overline{BC} = I, \quad \frac{\overline{BC}}{\overline{BA}} = \frac{I}{Q + \varphi(I)}.$$

Но последнее отношение есть не что иное, как $\operatorname{tg} \alpha$ — тангенс угла, образуемого прямой CA с осью Q . Для построения же прямой, тангенс угла наклона которой к оси Q равен

$$\operatorname{tg} \beta = \frac{dI}{dQ} = -\frac{\varphi(I) + Q}{I} = -\operatorname{ctg} \alpha,$$

нужно построить перпендикуляр к CA в точке A . Этот перпендикуляр и дает направление фазовой траектории в точке A . Итак, построение Льенара сводится к следующему: 1) после приведения уравнения к виду (4) строится характеристическая кривая (6); 2) строится прямоугольный треугольник ABC ; 3) строится перпендикуляр к гипотенузе AC , который и дает искомое направление траектории в точке A .

Покажем построение Льенара на примере лампового генератора, нелинейное уравнение которого имеет вид (см. § 21)

$$U'' + \frac{1}{L} \left(R - \frac{MS}{C} \right) U' + \omega_0^2 U = 0 \quad (7)$$

или

$$U'' + f(U) U' + \omega_0^2 U = 0 \quad (8)$$

(так как крутизна S зависит от напряжения U). Для приведения (8) к виду (4) введем прежде всего безразмерное («собственное») время

$$\tau = \omega_0 t.$$

Тогда

$$\frac{dU}{dt} = \frac{dU}{d\tau} \cdot \frac{d\tau}{dt} = \omega_0 \frac{dU}{d\tau}, \quad \frac{d^2U}{dt^2} = \omega_0^2 \frac{d^2U}{d\tau^2},$$

мы получим

$$U'' + \frac{1}{\omega_0} f(U) U' + U = 0. \quad (9)$$

Здесь штрихи означают уже дифференцирование по τ . Теперь введем новую переменную

$$x = \int U d\tau \quad (10)$$

и определим некоторую функцию $F(x')=F(U)$ соотношением

$$F(x') = \frac{1}{\omega_0} \int f(U) dU. \quad (11)$$

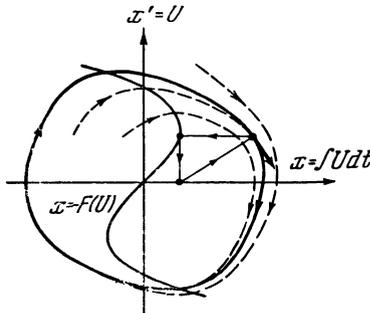
Проинтегрируем (9) по τ

$$\int U'' d\tau + \frac{1}{\omega_0} \int f(U) U' d\tau + \int U d\tau = U' + \frac{1}{\omega_0} \int f(U) dU + \int U d\tau.$$

Заменив теперь U на x на основании (10), получим

$$x'' + F(x') + x = 0, \quad (12)$$

т. е. уравнение, в точности совпадающее с (4). Координатами фазовой плоскости у нас будут теперь служить величины x и $x'=U$,



Р и с. 85

т. е. напряжение и интеграл от напряжения. Нам нужно прежде всего построить характеристическую кривую

$$x = -F(U).$$

Мы имеем

$$\begin{aligned} F(U) &= \frac{1}{\omega_0} \int f(U) dU = \frac{1}{\omega_0 L} \int \left(R - \frac{MS}{C} \right) dU = \\ &= \frac{1}{\omega_0 L} \left[RU - \frac{M}{C} \int S(U) dU \right]. \end{aligned} \quad (13)$$

Рассмотрим случай мягкого возбуждения, для которого

$$S(U) = a_1 + 3a_3 U^2 = S_0 + 3a_3 U^2 \quad (14)$$

(см. § 21). Подставляя (14) в (13), находим

$$F(U) = \frac{1}{\omega_0 L} \left[\left(R - \frac{MS_0}{C} \right) U - \frac{M}{C} a_3 U^3 \right]. \quad (15)$$

Кривая

$$x = -F(U)$$

есть, таким образом, кубическая парабола, график которой нанесен на рис. 85. Дальнейшее построение производится графическим способом, как показано на том же рисунке. Мы получаем уже известный нам фазовый портрет мягкого генератора. Однако следует заметить, что мы узнали нечто большее, чем то, что дала нам квазилинейная теория: мы получили *форму* колебания, которая в нелинейной системе должна, конечно, отличаться от синусоидальной. На портрете рис. 85 это находит свое отражение в том, что предельный цикл отличается от эллипса.

§ 26. Некоторые схемы генераторов

Говоря о ламповом генераторе, мы до сих пор пользовались только одной классической схемой с контуром в цепи сетки и с индуктивной обратной связью. Это делалось для того, чтобы продемонстрировать различные подходы и методы на одном и том же объекте исследования.

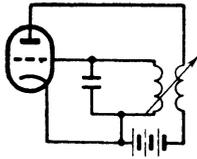
Рассмотрим теперь некоторые другие схемы генераторов и воспользуемся случаем для разъяснения одного принципиального вопроса, относящегося к обратной связи. Схема генератора с контуром в цепи сетки повторена на рис. 86; на рис. 87 изображена схема генератора с контуром в цепи анода. Не отличается принципиально от обеих этих схем схема с автотрансформаторной связью, показанная на рис. 88; эта схема часто называется трехточечной.

Обратная связь может быть не только индуктивной, но и емкостной. Для подачи обратной связи на сетку может использоваться емкость сетка—анод. Если она недостаточна, то между сеткой и анодом подключается дополнительный конденсатор обратной связи, как показано на рис. 89.

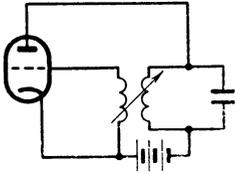
Схемы рис. 86—89 обладают тем общим признаком, что в них имеется явная цепь обратной связи, причем глубина обратной связи может регулироваться путем изменения того или иного параметра схемы — коэффициента связи в схемах рис. 86 и 87, коэффициента трансформации в схеме рис. 88, емкости сетка—анод в схеме рис. 89. Но существует ряд генераторов, в которых обратная связь действует по скрытым, неявным каналам, не могущим быть изображенными на схеме в виде отдельной цепи обратной связи. Обратная связь существует и в таких генераторах, потому что, как было показано в § 15, обратная связь является необходимым звеном всякой автоколебательной системы. Для полного понимания действия генераторов с неявной обратной связью ее механизм должен быть раскрыт.

В качестве примеров генераторов с неявной обратной связью рассмотрим вкратце динатронный и транзитронный генераторы.

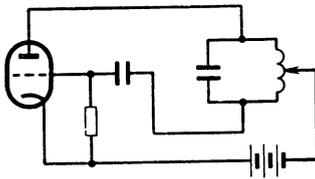
Действие динатронного генератора основано на так называемом динатронном эффекте, состоящем в том, что при попадании



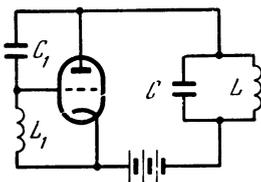
Р и с. 86



Р и с. 87



Р и с. 88



Р и с. 89

электронов на анод возникает вторичная эмиссия; электрон, прилетающий на анод из катода, выбивает из анода новые электроны. Если на один первичный электрон приходится несколько вторичных, то составляющая тока, обусловленная вторичными электронами, преобладает над током первичной эмиссии. В результате этого анодный ток не только убывает, но может даже изменить знак. Очевидно, что каждый первичный электрон может выбить тем больше вторичных, чем больше скорость первичного электрона. Скорость же электрона зависит от ускоряющего напряжения. В качестве диатрона может быть использован любой подходящий триод.

Следует заметить, что электрод, называемый в триоде анодом, при использовании триода в качестве диатрона не является уже анодом; назначение этого электрода состоит в испускании вторичных электронов; чтобы подчеркнуть это обстоятельство, анод триода, работающего в диатронном режиме, называют иногда динодом. Что же касается сетки триода, то в диатронном режиме ей задается высокий положительный потенциал, так что в диатроне она является анодом.

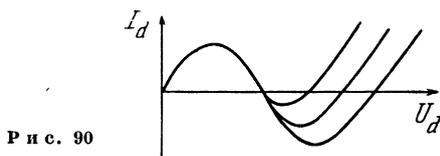
Характеристики диатрона, т. е. зависимость тока I_a в цепи динода от напряжения U_a на нем, показаны на рис. 90. Три кривые рис. 90 относятся к разным напряжениям на аноде; чем выше анодное напряжение, тем правее располагается характеристика.

Как видим, вольтамперная характеристика диатрона имеет падающий участок, так что диатрон является отрицательным сопротивлением, чем и обусловлена возможность его использования в качестве генератора. Схема диатронного генератора на триоде изображена на рис. 91.

Крутизна падающего участка вольтамперной характеристики, а следо-

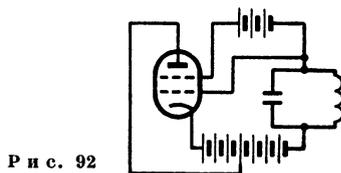
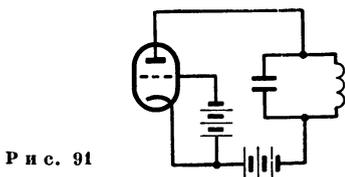
вательно, и величина отрицательного сопротивления зависит от тока эмиссии. В триоде ток эмиссии можно менять только путем изменения накала. Поэтому в качестве динатрона предпочитают применять тетрод, в котором током эмиссии можно управлять посредством потенциала на первой сетке.

Механизм обратной связи в динатронном генераторе крайне прост: он состоит в том, что напряжение на контуре, будучи подано



непосредственно на динод, управляет током динода. Таким образом, обратная связь действует по той же цепи, по которой осуществляется питание контура. Надлежащие фазовые соотношения обеспечиваются работой на падающем участке характеристики.

Обратимся к транзитронному генератору. Возьмем тетрод с ограниченной полной эмиссией (т. е., например, с вольфрамовым катодом) и включим его по схеме рис. 92. На вторую сетку подано отрицательное напряжение, и ток в ней отсутствует. Что же касается тока первой сетки I_1 , то в сумме с анодным током I_a он должен равняться току эмиссии. Таким образом, любое изменение тока I_1 может происходить только за счет перераспределения тока эмиссии между первой сеткой и анодом. При определенных соотношениях увеличение положительного напряжения на первой



и, следовательно, на второй сетках (они соединены между собою через батарею) может вызвать увеличение анодного тока, а следовательно, уменьшение тока I_1 . Таким образом, характеристика зависимости тока первой сетки от напряжения на ней может иметь падающий участок, т. е. между первой сеткой и катодом может получиться отрицательное сопротивление. Обратная связь в транзитронном генераторе состоит в том, что напряжение на контуре управляет перераспределением токов. Обратная связь действует одновременно по двум цепям, а именно по цепям обеих сеток.

Практические схемы транзитронных генераторов собираются на пентодах или гексодах, в которых первая сетка служит для ограничения полного тока эмиссии. Изменяя напряжение на первой сетке, можно управлять величиной отрицательного сопротивления.

§ 27. *RC*-генераторы синусоидальных колебаний

В рассмотренных до сих пор генераторах генерируемая частота определяется контуром из *LC*, представляющим собою колебательную систему в составе генератора. Возникает вопрос о том, при каких условиях возможно получить тот же результат, применяя схемы, составленные только из сопротивлений и емкостей.

Как нам уже известно, уравнение генератора имеет вид уравнения второго порядка, в котором действием обратной связи коэффициент при первой производной обращается в нуль. Следовательно, если мы составим схему из *R* и *C*, описываемую таким же уравнением, то это будет означать, что построенная нами схема обладает такими же свойствами, как и схема обычного генератора. (Такое заключение основывается на предположении, что составляемые нами дифференциальные уравнения правильно отображают реальные свойства схемы.)

Следовательно, для построения генератора *RC* мы должны прежде всего составить такую схему из сопротивлений и емкостей, которая описывалась бы дифференциальным уравнением второго порядка, аналогичным уравнению обычного контура. Но, как известно, простая цепь, содержащая только *R* и *C*, описывается уравнением первого порядка. Стало быть, мы должны перейти от простой цепи к сложной, обладающей по меньшей мере двумя степенями свободы¹.

Рассмотрим схему рис. 93. Уравнение, связывающее напряжение на входе и на выходе, имеет вид

$$U_2' + 2\alpha U_2 + \omega_0^2 \int U_2 dt = \frac{1}{R_1 C_2} U_1, \quad (1)$$

где

$$\omega_0^2 = 1/R_1 R_2 C_1 C_2, \quad (2)$$

$$2\alpha = 1/R_1 C_1 + 1/R_2 C_2 + 1/R_1 C_2. \quad (3)$$

Уравнение (1) совершенно аналогично уравнению цепи из последовательно включенных *R*, *C* и *L*. Если теперь ввести в схему усилитель и дать обратную связь с выхода на вход, то получится схема, способная генерировать синусоидальные (точнее, близкие к синусоидальным) колебания (рис. 94).

¹ Число степеней свободы в электрической цепи определяется числом независимых токов, которые могут протекать по цепи. Число степеней свободы может также определяться минимальным числом разрывов цепи, прекращающих всякое токопрохождение.

Составим уравнение схемы рис. 94. Пусть коэффициент усиления усилителя есть

$$K = U_1/U_2 \quad (4)$$

(заметим, что мы сохранили обозначения рис. 93, так что для усилителя U_1 является выходным напряжением, U_2 — входным). Выражая на основании (4) U_1 через U_2 в правой части (1) и перенося все в левую часть, получим

$$U_1'' + \left(\frac{1}{R_1 C_1} + \frac{1}{R_2 C_2} + \frac{1-K}{R_1 C_2} \right) U_1' + \frac{1}{R_1 R_2 C_1 C_2} U_1 = 0 \quad (5)$$

— обычное уравнение генератора. Условие самовозбуждения получим, подчинив коэффициент при первой производной условию

$$1/R_1 C_1 + 1/R_2 C_2 + (1-K)/R_1 C_2 \leq 0. \quad (6)$$

Очевидно, что для выполнения условия самовозбуждения нужно иметь достаточно большой коэффициент усиления K .

По поводу схемы рис. 94 следует еще заметить, что соотношение (4) выполняется только при четном числе ступеней усиления, так как каждая ступень производит обращение фазы на 180° . Поэтому при нечетном числе ступеней мы получим отрицательную обратную связь вместо требуемой положительной. Стало быть, необходимое число ламп в усилителе схемы рис. 94 не менее двух, и эту схему часто называют схемой двухлампового RC -генератора.

К построению схемы RC -генератора можно подойти и иначе. Всякая схема, приводимая к скелетной схеме автоколебательной

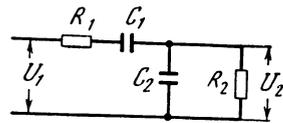


Рис. 93

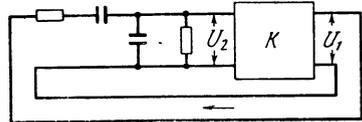


Рис. 94

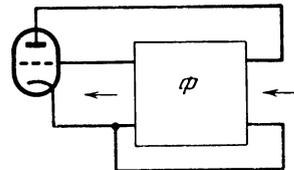


Рис. 95

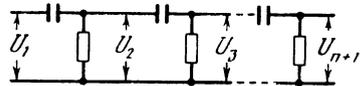


Рис. 96

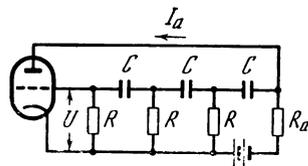


Рис. 97

системы рис. 53, будет генерировать установившиеся по частоте и амплитуде колебания, если устойчиво выполняется условие баланса амплитуд и фаз. Первое условие определяет амплитуду, второе — частоту установившихся колебаний. Условие баланса амплитуд обеспечивается подбором надлежащего усиления. Для выполнения же условия баланса фаз необходимо, чтобы при *данной* частоте (т. е. при той же частоте, которую генерирует или должен генерировать генератор) получалась положительная обратная связь. Так, схема рис. 95, в которой обратная связь подается на сетку триода через фазовращающий четырехполюсник Φ , может генерировать частоту, для которой фаза напряжения при прохождении через четырехполюсник сдвигается *ровно* на 180° (так как сама лампа осуществляет поворот фазы еще на 180°). Эти соображения указывают путь построения однолампового RC -генератора. В качестве фазовращателя может быть использована лестничного типа схема, состоящая из нескольких звеньев RC , как показано на рис. 96. Фаза поворачивается на некоторый угол при переходе через каждое звено. Однако в каждом звене совершается поворот фазы на угол *меньше* 90° . Таким образом, для получения сдвига *ровно* на 180° необходима цепочка не менее чем из трех звеньев. В результате схема однолампового RC -генератора принимает вид рис. 97. Полагая

$$I_a = SU$$

и составляя уравнение схемы, можно вывести условие самовозбуждения

$$R_a S \geq 29 + 23R_a/R + 4R_a^2/R^2$$

и выражение для частоты

$$\omega_0 = \frac{1}{RC} \sqrt{\frac{R}{6R + 4R_a}}$$

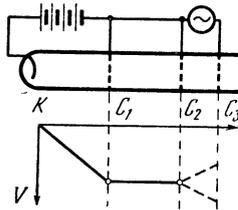
Следует пояснить, что для схемы с тремя звеньями получается уравнение третьего порядка. Из трех корней характеристического уравнения один отрицателен и действителен, два же остальных комплексны; они и определяют частоту и затухание.

§ 28. Генераторы сверхвысоких частот

Описанные выше схемы генераторов непригодны для генерирования колебаний сверхвысокой частоты. Это обусловлено, во-первых, тем, что на сверхвысокой частоте существенное влияние начинают оказывать паразитные параметры схемы и лампы¹. Изменяя способ монтажа и конструкцию деталей, в частности, применяя специальные конструкции ламп, сильно отличающиеся

¹ Паразитными называются параметры, не учитываемые принципиальной схемой первого приближения, как то: емкости проводов и элементов схемы между собою и на землю, индуктивности проводов и т. п.

от обычных, можно получить при помощи обычных принципиальных схем частоты до порядка $5 \cdot 10^8$ гц. Во-вторых, при сверхвысоких частотах приходится учитывать время пролета электрона в пространстве между электродами лампы, так как это время становится сравнимым с периодом колебания. Это последнее обстоятельство создает для обычного генератора неустранимые затруднения принципиального характера. Но, с другой стороны, это же

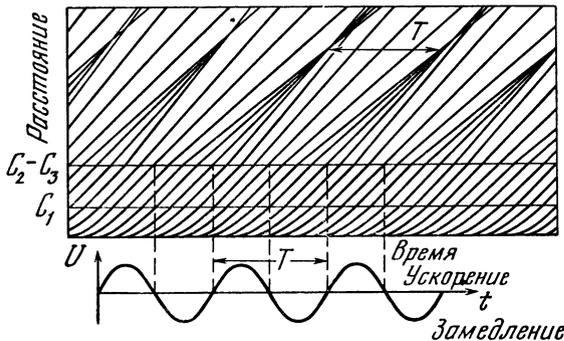


Р и с. 98

обстоятельство может быть использовано для построения ряда генераторов, в которых конечность времени пролета электронов лежит в самой основе механизма генерации.

В современной технике с. в. ч. применяются главным образом три вида генераторов: 1) клистронный генератор, 2) магнетронный генератор и 3) генератор на лампе обратной волны.

Мы не будем разбирать теорию этих генераторов. Но для того, чтобы дать понятие о некоторых принципах использования конечной скорости электронов для генерирования колебаний с. в. ч., приведем краткое качественное описание явлений в клистроне.



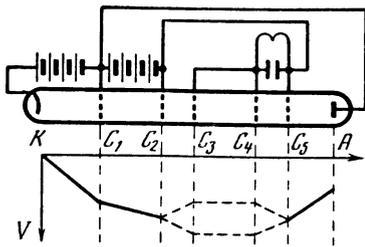
Р и с. 99

Основа действия клистрона состоит в *модуляции* электронного потока *по скорости* с последующим автоматическим преобразованием модуляции по скорости в *модуляцию по плотности* электронного потока. Термин «модуляция» употребляется здесь просто в том смысле, что некоторый параметр изменяется в соответствии с оказываемым на него воздействием.

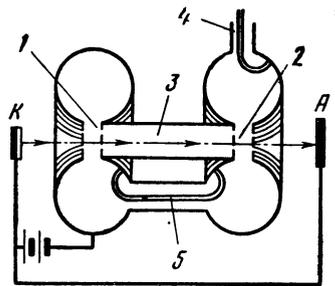
Рассмотрим рис. 98, на котором изображена пока только часть клистрона. Кроме катода K , имеются еще три электрода в форме сеток. Сетка C_1 играет роль ускоряющего электрода; ей сообщен положительный потенциал относительно катода. Между сетками C_2 и C_3 приложено переменное напряжение. Соответствующая потенциальная диаграмма изображена в нижней части рис. 98. Электроны, выйдя из катода, движутся на участке $K-C_1$ с ускорением и, накопив одинаковую скорость, проходят с этой скоростью участок C_1-C_2 , после чего попадают в промежуток C_2-C_3 . Здесь действует переменное поле, которое либо замедляет, либо ускоряет электроны. Таким образом, по выходе из промежутка C_2-C_3 электронный поток оказывается модулированным по скорости, т. е. электроны движутся с разными скоростями. В результате этого в пространстве правее C_3 более быстрые электроны догоняют более медленные и образуются так называемые электронные сгустки, т. е. области с большой плотностью электронов. Легко сообразить, что сгустки располагаются вдоль оси клистрона периодически, образуя бегущую волну плотности. Длина этой волны, т. е. расстояние между соседними сгустками, равна $\lambda = v/f$, где v — средняя скорость до модуляции; f — частота модуляции. Процесс образования сгустков называется группированием; комбинация сеток C_2-C_3 — группирователем; пространство, в котором формируются сгустки, — пространством группирования или пространством дрейфа. Процесс группирования поясняется рис. 99, в нижней части которого изображено изменение во времени переменного напряжения на группирователе. В верхней части рисунка графиком в координатах расстояние—время изображено движение электронов. Предполагается для простоты, что электроны выходят из катода через равные промежутки времени. Получив одинаковые скорости после прохождения ускоряющего поля, электроны поступают в группирователь C_2-C_3 , где скорости электронов увеличиваются или уменьшаются в зависимости от знака напряжения на группирователе (см. синусоиду в нижней части рисунка). График движения электронов, выходящих из группирователя, представляется прямыми с разным наклоном: более крутые линии соответствуют более быстрым электронам.

На рисунке видно, что благодаря различному наклону прямые, представляющие движение электронов, образуют сходящиеся пучки, что и является графическим пояснением формирования электронных сгустков.

Образующийся в пространстве группирования электронный поток переменной плотности может быть использован для наведения тока во внешней цепи. Для этого понадобятся еще две сетки C_4 и C_5 , к которым может быть подключен настроенный на соответствующую частоту контур. Если же теперь снять напряжение с контура и подать на группирователь, осуществив таким образом обратную связь, то мы получим схему клистронного генератора, изображенную на рис. 100. Отработанные электроны попа-



Р и с. 100

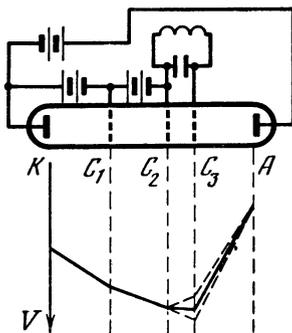


Р и с. 101

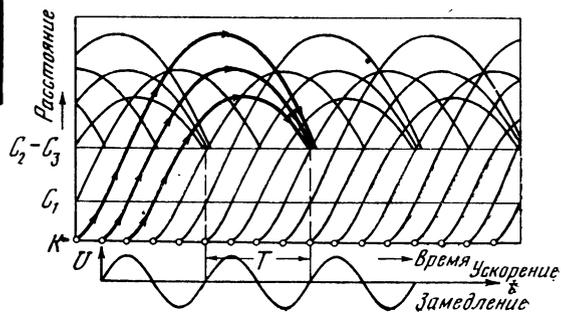
дают на коллектор A , а с него на катод. Комбинация сеток C_4 и C_5 носит название улавливателя. Для того чтобы система возбуждалась, необходима надлежащая фаза обратной связи. При заданном конструкцией расстоянии между группирователем и улавливателем фазу обратной связи можно регулировать изменением ускоряющих напряжений на C_1 и C_2 .

Конструктивное оформление клистрона с двумя резонаторами показано на рис. 101, на котором 1 и 2 означают объемные резонаторы тороидальной формы (на рисунке они показаны в разрезе). Группирователь и улавливатель образованы парами перфорированных дисков, входящими в качестве емкостей в объемные резонаторы. Пространство группирования выполнено в виде трубы 3. Еще одна труба содержит петлю обратной связи 5. Для вывода переменного напряжения служит петля 4.

Наиболее распространен так называемый *отражательный клистрон*. Его отличие от только что рассмотренной системы состоит в том, что группирователь и улавливатель совмещены. Электронный поток, пройдя через пару сеток, выполняющих роль группирователя, тормозится отрицательным полем и поворачивает обратно, проходя снова через ту же пару сеток, которая на этот раз играет роль улавливателя. Схема отражательного клистрона и соответствующая потенциальная диаграмма даны на рис. 102. Процесс группирования в отражательном клистроне поясняется



Р и с. 102



Р и с. 103

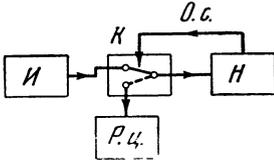
рис. 103. Этот рисунок отличается от рис. 99 тем, что линии, представляющие движение электронов по выходе из группирователя, не прямые, а параболы (что соответствует движению в ускоряющем поле обратного направления). Электроны возвращаются к C_2-C_3 ; вследствие различия начальных скоростей образуются сгустки. Необходимо, очевидно, настроить систему так, чтобы сгустки образовались в зоне C_2-C_3 . Следует заметить, что в отражательном клистроне цепь обратной связи в явном виде отсутствует; обратная связь осуществляется через электронный поток. Настройка отражательного клистрона может производиться не только изменением ускоряющих положительных потенциалов на C_1 и C_2 , но и отрицательным потенциалом на отражательном электроде A .

§ 29. Релаксационные автоколебания

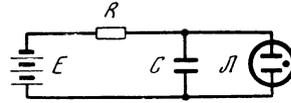
Мы рассматривали до сих пор генераторы колебаний, близких по форме к синусоидальным. Такие автоколебания носят название *почти-гармонических*. Но в радиотехнике применяются также автоколебания, резко отличные по форме от синусоидальных; такие автоколебания, отличающиеся иногда внезапными скачкообразными изменениями тока или напряжения, или скорости изменения этих величин — что изображается графически осциллограммой в виде ломаной линии, — называются *релаксационными* автоколебаниями. Существенное различие формы почти-гармонических и релаксационных автоколебаний обусловлено различием в устройстве и действии генераторов этих видов колебаний. В состав генератора почти-гармонических колебаний входит обычно контур, составленный из L и C (или эквивалентная контуру схема из R и C), т. е. колебательная система, содержащая два (по крайней мере) реактивных элемента, способных накапливать энергию. Колебательный процесс в контуре состоит с энергетической точки зрения в том, что запасенная в контуре энергия то сосредоточивается полностью в магнитном поле индуктивности, то переходит в электрическое поле емкости. Такая перекачка энергии происходит периодически; если потери отсутствуют, то колебание совершается по строго синусоидальному закону; при наличии же потерь амплитуда убывает со временем по экспоненциальному закону в связи с тем, что запас энергии все время убывает. Если же потери энергии восполняются, то колебания делаются незатухающими, а форма их лишь постольку отличается от синусоидальной, поскольку это зависит от ограничивающей амплитуду нелинейности.

Совершенно иной механизм генератора релаксационных колебаний. Он характеризуется тем, что имеется единственный (в простейшем случае) накопитель энергии. Происходящий в релаксационном генераторе периодический процесс состоит в том, что накопитель энергии набирает возрастающий запас энергии, пока

не будет достигнут некоторый определенный высший уровень. Затем некоторый клапан разряжает накопитель на цепь разряда, пока запас энергии не достигнет определенного низшего уровня. Здесь клапан снова срабатывает и включает накопитель на зарядку. Сказанное поясняется блок-схемой рис. 104, на которой означают: *И* — источник энергии, *Н* — накопитель, *К* — клапан,



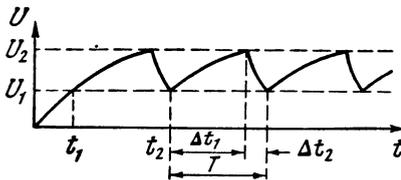
Р и с. 104



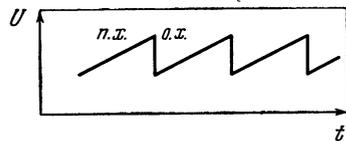
Р и с. 105

пан, *О. с.* — обратная связь, *Р. ц.* — разрядная цепь. Из приведенного общего описания уже следует, что для релаксационного генератора характерно наличие клапана, могущего занимать два фиксированных положения, одно из которых соответствует включению накопителя на зарядку, а второе — на разрядку. Клапан управляется обратной связью с накопителя; действие обратной связи состоит в том, что переключение клапана производится при достижении двух фиксированных уровней энергии — высшего и низшего. Отсюда сразу следует, что размах релаксационных колебаний не зависит от энергетического баланса, а определяется полностью перепадом между указанными двумя уровнями, при которых срабатывает клапан.

Все эти общие свойства релаксационного генератора легко уяснить себе, рассмотрев простейший пример — релаксационный генератор с газоразрядной лампой. Схема генератора изображена на рис. 105. Источник энергии — батарея *Е* — заряжает через



Р и с. 106



Р и с. 107

сопротивление *R* конденсатор *С*, играющий в этой схеме роль накопителя. В качестве клапана применена газоразрядная лампа *Л*; она же является цепью, на которую разряжается накопитель. Действие схемы таково: после включения батареи конденсатор начинает заряжаться. Когда напряжение на нем достигнет напряжения U_2 — «напряжения зажигания» лампы *Л*, — в последней произойдет ионизация заполняющего ее газа и сопротивление лампы, которое до этого было весьма большим, резко упадет до малой величины *r*. Конденсатор начнет быстро разряжаться на

это сопротивление. Когда же напряжение упадет до значения U_1 , при котором ионизация прекращается, сопротивление лампы снова резко возрастет, и возобновится процесс зарядки конденсатора. Ход явления представлен осциллограммой рис. 106, изображающей зависимость напряжения на конденсаторе от времени. Размах колебаний зависит только от перепада напряжения, т. е. от разности $U_2 - U_1$. Что же касается частоты, то она зависит от постоянных времени цепей зарядки и разрядки, а также от питающего напряжения. Определим период колебания. Он складывается из двух отрезков Δt_1 и Δt_2 (см. рис. 106), первый из которых определяет время зарядки, второй — разрядки. Мы имеем для процесса зарядки

$$U = E(1 - e^{-t/\tau_1}),$$

где $\tau_1 = RC$ — постоянная времени зарядной цепи. Можем составить два равенства

$$U_1 = E(1 - e^{-t_1/\tau_1}), \quad U_2 = E(1 - e^{-t_2/\tau_1})$$

(см. рис. 106), из которых находим

$$t_1 = \tau_1 \ln \frac{E}{E - U_1}, \quad t_2 = \tau_1 \ln \frac{E}{E - U_2},$$

а отсюда и время зарядки

$$\Delta t_1 = t_2 - t_1 = RC \ln \frac{E - U_1}{E - U_2}.$$

Это выражение показывает, что чем больше E при данных U_1 и U_2 , тем меньше отношение под знаком логарифма отличается от единицы и тем меньше время зарядки. Процесс разрядки выражается формулой

$$U = U_2 e^{-\frac{1}{\tau_2}(t-t_2)},$$

где $\tau_2 = rC$ — постоянная времени разрядной цепи. При $t = t_3$ достигается значение напряжения $U = U_1$. Мы получаем

$$U_1 = U_2 e^{-\frac{1}{\tau_2}(t_3-t_2)} = U_2 e^{-\frac{\Delta t_2}{\tau_2}},$$

откуда

$$\Delta t_2 = rC \ln \frac{U_2}{U_1}.$$

Теперь мы получим период колебания, складывая Δt_1 и Δt_2 ,

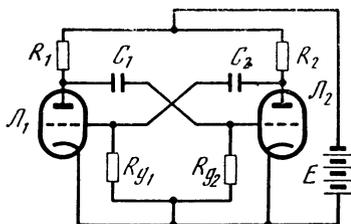
$$T = \Delta t_1 + \Delta t_2 = C \left(R \ln \frac{E - U_1}{E - U_2} + r \ln \frac{U_2}{U_1} \right).$$

Как видим, период колебания зависит от постоянных времени цепей зарядки и разрядки, т. е. от величин RC и rC , от соотношения напряжений зажигания и потухания U_2 и U_1 и от питающего

напряжения E . Если взять E и R достаточно большими, а r и C достаточно малыми, то форма колебания приближается к показанной на рис. 107. Напряжение такой формы называется «пилообразным» и имеет в радиотехнике широкое применение в качестве *развертывающего* напряжения в электронной осциллокопии, в телевидении, радиолокации и др. Развертывающему напряжению предъявляют обычно требования высокой степени линейности на участке прямого хода (п. х. на рис. 107) и минимальной длительности обратного хода (о. х.). Рассмотренная нами простейшая схема рис. 105 плохо удовлетворяет обим этим требованиям. Линейность прямого хода можно значительно улучшить, включив вместо сопротивления R ограничитель тока, например диод, работающий в режиме насыщения. Однако для генерирования пилообразного развертывающего напряжения предпочитают пользоваться специальными, более совершенными релаксационными генераторами.

§ 30. Релаксационные генераторы

Одной из наиболее ранних схем релаксационных генераторов является симметричный *мульти vibrator*¹, не утративший значения и до наших дней. Схема симметричного мульти vibratorа показана на рис. 108. Обе половины схемы можно рассматривать как два отдельных генератора, связанных между собою так, что один из генераторов выполняет функции клапана, управляющего работой другого.



Р и с. 108

Переходя к описанию работы схемы, начнем с такого ее состояния, когда токи в анодных цепях обеих ламп в точности равны. Это есть состояние равновесия, однако неустойчивого. В самом деле, пусть ток первой лампы немного возрастет. При этом напряжение на лампе станет меньше первоначального, и конденсатор C_1 начнет разряжаться. Разрядный ток, протекая по сопротивлению $R_{г2}$, создаст отрицательное напряжение на сетке второй лампы. Вследствие этого напряжение на аноде возрастет, а анодный ток уменьшится. Благодаря повышению напряжения на аноде нач-

¹ Т. е. генератор колебаний со многими частотами (имеется в виду богатство спектра). Мульти vibrator применялся в свое время для градуировки волномеров; опорными точками служили частоты многочисленных гармоник.

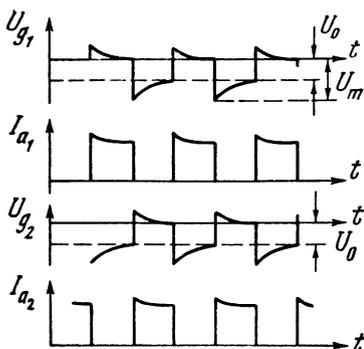


Рис. 109

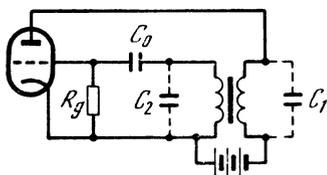


Рис. 110

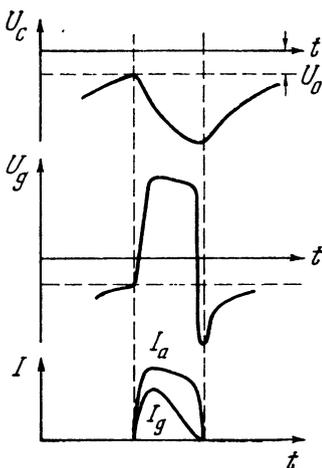


Рис. 111

нется подзарядка конденсатора C_2 . Зарядный ток, протекая по сопротивлению R_{g1} , вызовет положительное напряжение на сетке первой лампы, а следовательно, дальнейшее увеличение тока первой лампы.

Это рассуждение показывает, что рассмотренное нами исходное состояние, характеризующееся равенством токов и напряжений в обеих половинах схемы, действительной неустойчиво.

Рост тока первой лампы и нарастание отрицательного напряжения на сетке второй лампы происходит очень быстро — скачком. В результате первая лампа окажется насыщенной, а вторая запертой большим отрицательным напряжением. Теперь конденсатор C_1 продолжает разряжаться через L_1 и R_{g2} ; разрядный ток, а с ним и отрицательное напряжение на сетке L_2 убывают, пока лампа L_2 не отойдет. В этот момент напряжение на ней падает, конденсатор C_2 начинает разряжаться, и происходит скачкообразное изменение токов и напряжений, описанное выше, но с переменной ролей обеих половин схемы. Изменения во времени анодных токов и напряжений на сетках в обеих половинах схемы изображены на рис. 109. Вследствие симметрии схемы длительность импульсов анодного тока в обеих половинах одинакова. Таким образом, форма анодного тока в каждой лампе приближается к прямоугольной. Впрочем, изменив параметры, можно получить также форму, приближающуюся к треугольной. Форма колебания зависит от постоянных времени разрядки и зарядки. Из схемы рис. 108 видно, что постоянная времени разрядки равна

$$\tau_p = C \left(R_g + \frac{R_i R}{R_i + R} \right),$$

а постоянная времени зарядки

$$\tau_s = C \left(R + \frac{R_g r_g}{R_g + r_g} \right),$$

где C — емкость; R — сопротивление в цепи анода; R_i — сопротивление лампы (катод — анод); R_g — сопротивление в цепи сетки; r_g — сопротивление лампы (катод — сетка). Обычно $\tau_p \gg \tau_s$; при этом форма колебаний близка к прямоугольной. Зная постоянную времени, можно найти период колебания. За половину периода напряжение на сетке убывает от наибольшего значения U_m до напряжения запираания U_0 . Таким образом,

$$U_0 = U_m e^{-\frac{1}{2} \cdot \frac{T}{\tau_p}},$$

откуда

$$T = 2\tau_p \ln \frac{U_m}{U_0}.$$

Напряжение запираания U_0 — это параметр лампы. Что же касается U_m , то в начале разрядки, т. е. непосредственно после скачка тока, оно равно

$$U_m = E \frac{1}{1 + R_i/R + R_i/R_g}.$$

Итак, период зависит от параметров лампы, от параметров схемы и от питающего напряжения.

В качестве второго примера релаксационного генератора рассмотрим вкратце так называемый блокинг-генератор, применяемый главным образом для получения весьма коротких импульсов (длительностью до долей микросекунды). Схема блокинг-генератора изображена на рис. 110. Она состоит из триода и трансформатора, связывающего цепи анода и сетки. Емкости, образующие с индуктивностями обмоток колебательные контуры, в схеме не предусмотрены; однако необходимо учитывать распределенные емкости, эквиваленты которых намечены на схеме штрихом. Схему рис. 110 можно рассматривать как вырожденную схему обычного генератора с контурами в цепях анода и сетки; для схемы рис. 110 характерно, что процессы в цепях имеют аperiодический характер и перемежаются резкими скачками (эти скачки, называемые в иностранной литературе «блокинг», и обуславливают наименование рассматриваемого типа генератора).

Действие схемы рассмотрим с момента, когда конденсатор C_0 заряжен, на сетке — отрицательный потенциал относительно катода и лампа заперта. Конденсатор C_0 разряжается через сопротивление R_g ; постоянная времени процесса разрядки равна $R_g C_0$. Когда напряжение достигнет значения U_0 , лампа отперется; возникнет анодный ток, а благодаря сильной положительной об-

ратной связи на сетке появится положительное напряжение, обуславливающее дальнейший быстрый рост анодного тока. Этот процесс, протекающий с очень большой скоростью, приводит к тому, что напряжение на аноде падает почти до нуля (это и есть «блокинг»). Когда результирующее напряжение на сетке (равное сумме отрицательного напряжения на C_0 и положительного напряжения, наводимого обратной связью) упадет до нуля, появится сеточный ток, и конденсатор начнет быстро заряжаться через промежуток сетка — катод с постоянной времени $C_0 r_g$. В этой стадии процесса напряжение на сетке положительно и превосходит напряжение на аноде; крутизна лампы при этом мала, и изменения напряжения на сетке не оказывают большого влияния на анодный ток. Но в дальнейшем, после того, как напряжение на сетке, достигнув наибольшего значения, начинает убывать, крутизна увеличивается. Наконец, достигаются соотношения, благоприятные для очередного скачка: уменьшение напряжения на сетке вызывает уменьшение анодного тока, что вызывает дальнейшее уменьшение сеточного напряжения, и лампа с очень большой скоростью запирается. После этого весь описанный цикл повторяется. На рис. 111 показано изменение во времени напряжения на конденсаторе U_c , на сетке U_g и сеточного и анодного токов I_g и I_a .

Условие самовозбуждения блокинг-генератора может быть записано в виде (при условии, что C_1 и C_2 очень малы по сравнению с C_0)

$$kS_a - 1/nR_i - n/R_g - nS_g + 1/r_g > 0,$$

где S_a и S_g — соответственно крутизны по аноду и по сетке; R_i и r_g — внутренние сопротивления катод — анод и катод — сетка; n — коэффициент трансформации; k — коэффициент связи (близкий к единице в рассматриваемом случае). Минимальная длительность импульсов зависит от индуктивности рассеяния трансформатора и паразитной емкости C_1 . Период следования импульсов определяется в основном постоянной времени разрядки конденсатора C_0 ; приближенно $T \cong R_g C_0$.

Г л а в а III

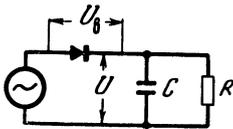
НЕЛИНЕЙНЫЕ СИСТЕМЫ ПОД ВНЕШНИМ ВОЗДЕЙСТВИЕМ

В главе I рассматривались цепи с нелинейными сопротивлениями и главнейшие радиотехнические процессы, основанные на явлениях в таких цепях. В настоящей главе рассматриваются более сложные явления, происходящие в нелинейных цепях с реактивными сопротивлениями, а также вынужденный режим некоторых автоколебательных и потенциально-автоколебательных

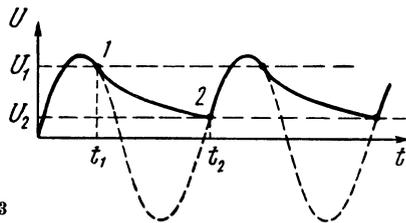
систем. Последний термин означает систему, могущую при известных обстоятельствах генерировать колебания, но находящуюся в исходном положении (при отсутствии внешнего воздействия) в устойчивом состоянии («недовозбужденная» система).

§ 31. Работа выпрямителя

Выпрямительная схема всегда содержит того или иного устройства фильтр, служащий для ослабления переменной составляющей выпрямленного напряжения. Таким образом, в реальной выпрямительной схеме должны содержаться реактивные сопротивления. Простейшая схема представлена на рис. 112. Найдем выходное напряжение U .



Р и с . 112



Р и с . 113

Положим для простоты, что характеристика выпрямителя задана в виде следующей зависимости проводимости от напряжения

$$Y = \begin{cases} 0 & \text{при } U_b < 0, \\ \infty & \text{при } U_b > 0. \end{cases}$$

(В действительности прямая и обратная проводимости конечны, но различаются в хорошем выпрямителе во много раз.)

Если сопротивление выпрямителя постоянно в некотором интервале напряжений, то это значит, что в пределах этого интервала явления описываются линейным уравнением. В соседнем же интервале может действовать тоже линейная, но уже другая зависимость. В нашем случае работа схемы описывается следующими линейными уравнениями:

$$U = E \quad (U_b > 0), \quad (1)$$

$$U + RI = 0 \quad (U_b < 0). \quad (2)$$

Напряжение на выпрямителе равно, очевидно,

$$U_b = E - U.$$

Беря

$$E = E_m \sin \omega t$$

и выражая в (2) ток через напряжение, получим

$$U = E_m \sin \omega t \quad (E > U),$$

$$\frac{dU}{dt} + \frac{1}{RC} U = 0 \quad (E < U).$$

Периодическое решение нашей нелинейной задачи мы получим, учитывая, что напряжение U не может измениться скачком; следовательно, мы должны значение напряжения внутри одного интервала (т. е. при $E > U$) приравнять соответствующему значению напряжения внутри другого интервала (т. е. при $E < U$) на общей их границе (т. е. при $E = U$). Это и есть метод сшивания.

Качественная картина явления, совершенно очевидная заранее, представлена на рис. 113. Напряжение изменяется по синусоидальному закону до $t=t_1$; в этот момент скорость убывания напряжения на конденсаторе становится меньше скорости убывания синусоидального напряжения E . Выпрямитель запирается, и конденсатор разряжается на сопротивление R . Процесс разряда продолжается, пока в момент $t=t_2$ U и E не сравняются. В этот момент выпрямитель отпирается, и дальнейшее изменение напряжения происходит по синусоидальному закону вплоть до $t=t_1+T$.

Процесс разряда происходит по экспоненциальному закону

$$U = U_1 e^{-\alpha(t-t_1)}, \quad (3)$$

где $\alpha=1/RC$. Неизвестными являются координаты точек 1 и 2, т. е. величины U_1 , U_2 , t_1 и t_2 . Для их определения мы должны располагать четырьмя уравнениями.

Первое уравнение связывает U_1 и t_1

$$U_1 = E_m \sin \omega t_1. \quad (4)$$

Второе уравнение получим, приравняв в этой же точке производные экспоненты (3) и синусоиды (2). Это дает

$$-\alpha U_1 = \omega E_m \cos \omega t_1. \quad (5)$$

Далее имеем, с одной стороны,

$$U_2 = E_m \sin \omega t_2, \quad (6)$$

с другой же стороны, должно быть

$$U_2 = U_1 e^{-\alpha(t_2-t_1)}. \quad (7)$$

Система уравнений (4)–(7) определяет искомые величины. Из (4) и (5) находим

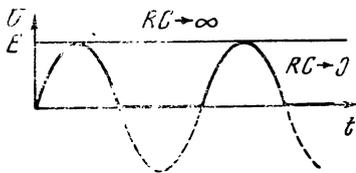
$$\operatorname{tg} \omega t_1 = -\omega/\alpha.$$

Выражая \sin через tg и подставляя в (4), получаем

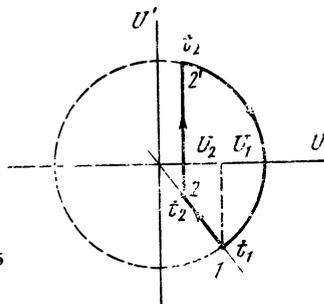
$$U_1 = \frac{E_m}{\sqrt{1 + \frac{\alpha^2}{\omega^2}}}.$$

Для t_2 , приравняв правые части (6) и (7), получим аналитически неразрешимое трансцендентное уравнение. Но графически это уравнение разрешается очень просто: из точки 1 с уже известными координатами U_1, t_1 нужно провести экспоненту (3) до пересечения с синусоидой в точке 2. Координаты этой точки и дают искомые значения U_2 и t_2 . В зависимости от значения постоянной времени $1/\alpha = RC$ будем иметь ход напряжения в пределах, указанных на рис. 114.

Построим диаграмму напряжения в координатах U, U' . Изображением синусоиды на плоскости U, U' является окружность



Р и с. 114



Р и с. 115

(при согласованных масштабах). Изображение экспоненты мы получим, продифференцировав (3)

$$U' = -\alpha U_1 e^{-\alpha(t-t_1)}. \quad (8)$$

Разделив (8) на (3), найдем

$$U'/U = -\alpha.$$

Таким образом, экспоненциальный закон изменения напряжения в координатах U, U' представляется прямой, проходящей через начало координат с наклоном $-\alpha$.

Диаграмма имеет вид, показанный на рис. 115. Отрезок 1—2 соответствует экспоненте. В точке 2 происходит скачок производной; изображающая точка мгновенно перебрасывается из положения 2 в положение 2'. Далее изображающая точка движется по дуге окружности 2'—1, соответствующей во времени синусоидальному закону изменения напряжения. Предельные случаи, представленные на рис. 114, изобразятся на диаграмме так: при $RC \rightarrow 0$ наклонная прямая займет вертикальное положение и диаграмма выродится в полуокружность; при $RC \rightarrow \infty$ прямая расположится горизонтально, и диаграмма стянется в одну точку на оси с координатой $U = E_m$.

Режим $RC \rightarrow \infty$ употребляется, в частности, при построении пик-вольтметров — приборов, измеряющих максимальное значение напряжения.

§ 32. Резонанс в нелинейной цепи

Явление резонанса в цепи, содержащей нелинейные элементы, имеет существенные особенности. Мы рассмотрим резонанс в последовательной цепи, содержащей нелинейную емкость.

Составим уравнение напряжений для последовательной цепи из R , L и C , считая, что R и L — постоянные величины, а C зависит от заряда

$$LQ'' + RQ' + \frac{1}{c(Q)} Q = E_m \sin \omega t$$

или

$$Q'' + 2\alpha Q' + F(Q) = B \sin \omega t. \quad (1)$$

Будем считать, что $F(Q)$ — нечетная функция, и положим

$$F(Q) = \omega_0^2 (Q + \lambda Q^3). \quad (2)$$

Таким образом, степень нелинейности во всех последующих соотношениях определяется параметром λ ; при $\lambda \rightarrow 0$ мы возвращаемся к обычным линейным зависимостям.

Будем искать установившееся решение (1) в виде

$$Q = a \sin \omega t + b \cos \omega t, \quad (3)$$

где a и b — постоянные величины. Амплитуда заряда

$$A = Q_m = \sqrt{a^2 + b^2}.$$

Подставляя решение (3) в (1) и приравнявая порознь коэффициенты при $\sin \omega t$ и $\cos \omega t$, получим два уравнения

$$(\omega_0^2 - \omega^2) a - 2\alpha \omega b + \frac{3}{4} \lambda \omega_0 a A^2 = B,$$

$$(\omega_0^2 - \omega^2) b + 2\alpha \omega a + \frac{3}{4} \lambda \omega_0 b A^2 = 0.$$

Умножая поочередно на a и b , складывая и вычитая, получим уравнение, содержащее только $A^2 = a^2 + b^2$,

$$\left(\frac{3}{4} \lambda\right)^2 \omega_0^4 A^6 + \frac{3}{2} \lambda \omega_0^2 (\omega_0^2 - \omega^2) A^4 + \\ + [(\omega_0^2 - \omega^2)^2 + \omega_0^2 d^2] A^2 - B^2 = \varphi(A^2, \omega^2) = 0. \quad (4)$$

Решив это уравнение относительно A , мы получили бы амплитуду заряда как функцию амплитуды внешней э. д. с. B и частоты ω , т. е. выражение для резонансной кривой.

Но здесь возникает затруднение, связанное с тем, что уравнение (4) — кубическое относительно A^2 . Поэтому его удобнее решать не относительно A^2 , а относительно ω^2 , так как для ω^2 уравнение (4) является квадратным (а для ω — биквадратным). Переписав (4) в виде

$$\omega^4 - 2\omega_0^2 \left(1 + \frac{3}{4} \lambda A^2 - \frac{1}{2} d^2\right) \omega^2 + \omega_0^4 \left(1 + \frac{3}{4} \lambda A^2\right)^2 - \frac{B^2}{A^2} = 0,$$

найдем решение этого уравнения:

$$\omega^2 = \omega_0^2 \left[\left(1 + \frac{3}{4} \lambda A^2 - \frac{1}{2} d^2 \pm \sqrt{\frac{B^2}{A^2} - d^2 \left(1 + \frac{3}{4} \lambda A^2 \right) + \frac{1}{4} d^4} \right) \right].$$

То обстоятельство, что мы нашли не амплитуду как функцию частоты, а, наоборот, частоту как функцию амплитуды, не меняет сути дела; такая замена с точки зрения графика сводится лишь к перемене местами осей абсцисс и ординат.

Высокая степень уравнения для A указывает на то, что решение может быть многозначным. Поэтому полезно исследовать зависимость между A^2 и ω^2 путем рассмотрения крайних положений касательных к кривой, выражающей эту зависимость. Производную можно выразить через частные производные неявной функции (4); именно ¹

$$\frac{dA^2}{d\omega^2} = - \frac{\partial \varphi / \partial \omega^2}{\partial \varphi / \partial A^2}. \quad (5)$$

Касательная горизонтальна, когда производная обращается в нуль, т. е. когда

$$\partial \varphi / \partial \omega^2 = 0. \quad (6)$$

Это равенство определяет положение максимумов резонансной кривой. Касательная принимает вертикальное положение, когда производная (5) обращается в бесконечность, т. е. при условии

$$\partial \varphi / \partial A^2 = 0. \quad (7)$$

Найдем обе частные производные; приравняв их нулю, получим уравнения

$$\omega^2 - \omega_0^2 \left(1 + \frac{3}{4} \lambda A^2 - \frac{1}{2} d^2 \right) = 0, \quad (8)$$

$$3 \left(\frac{3}{4} \lambda \right)^2 \omega_0^4 A^4 + 4 \left(\frac{3}{4} \lambda \right) (\omega_0^2 - \omega^2) A^2 + (\omega_0^2 - \omega^2)^2 + d^2 = 0. \quad (9)$$

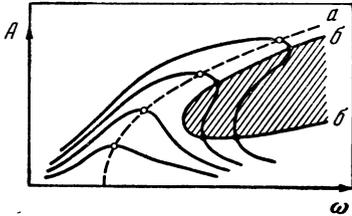
Уравнение (8) есть уравнение кривой, на которой располагаются максимумы семейства резонансных кривых. Уравнение (9) есть уравнение кривой, представляющей место точек, где резонансная кривая проходит вертикально.

Семейство резонансных кривых изображено на рис. 116. Как видим, кривые наклонены вправо, причем наклон тем больше, чем больше максимальная амплитуда, так как при этом сильнее проявляется нелинейность. (Заметим, что при $\lambda < 0$ получился бы наклон влево.) На графике нанесены кривые, уравнения которых выражены формулами (8) (кривая *a*) и (9) (кривая *b*). Максимумы резонансных кривых отмечены кружками; точки, в которых каса-

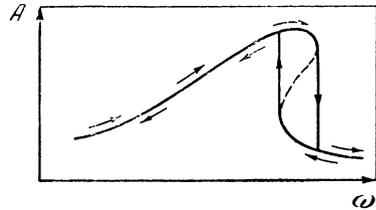
¹ Рассматривая зависимость между A^2 и ω^2 (вместо зависимости между A и ω), мы не меняем характера особых точек кривой; изменяется лишь масштаб графика, т. е. линейный масштаб заменяется квадратичным.

тельные к резонансным кривым вертикальны, отмечены крестиками.

При экспериментальном снятии нелинейных резонансных кривых типа рис. 116 происходят следующие явления. При повышении частоты амплитуда колебаний нарастает, достигает максимума, затем убывает, пока не будет достигнуто значение, соответствующее точке с вертикальной касательной. Здесь происходит срыв: амплитуда колебания скачком убывает до значения, соответствующего нижней ветви кривой (рис. 117), а в дальнейшем убы-



Р и с. 116



Р и с. 117

вает монотонно. Если же снимать кривую путем уменьшения частоты, то сперва амплитуда нарастает монотонно, а затем срывается на верхнюю ветвь кривой, как показано на том же рис. 117. Таким образом, скачок амплитуды происходит при разных значениях частоты в зависимости от того, повышается частота или понижается. Участок резонансной кривой между точками срыва (штриховая линия на рис. 117) соответствует неустойчивым режимам, не могущим наблюдаться при снятии резонансной кривой. Следовательно, область, ограничиваемая кривой b , есть область неустойчивости; на рис. 116 эта область заштрихована.

Аналогичные явления наблюдаются и при нелинейной индуктивности, например при наличии в цепи катушки индуктивности с ферромагнитным сердечником. Резонансные явления в такой цепи называют обычно феррорезонансом.

В заключение заметим, что резонанс в нелинейной цепи возможен и на дробной частоте, т. е. на частоте, равной некоторой доле частоты внешнего напряжения. Другими словами, резонанс в нелинейной цепи может применяться для деления частоты. Так, например, в технике имеют применение феррорезонансные делители частоты в отношении 1 : 3.

§ 33. Автоколебательная система под внешним воздействием

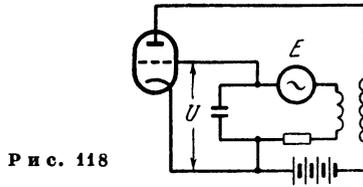
В этом параграфе мы выведем некоторые общие соотношения, которые в дальнейшем будут использованы при рассмотрении ряда технических вопросов. Речь пойдет об автоколебательной системе, на которую действует внешняя электродвижущая сила.

Всякая автоколебательная система нелинейна; поэтому при воздействии на нее внешней э. д. с. возникает ряд своеобразных нелинейных явлений, используемых в радиотехнике.

Мы составим уравнение для схемы рис. 118, представляющей собою генератор с контуром в цепи сетки. Отличие этой схемы от ранее рассматривавшейся состоит в том, что в контуре действует еще внешняя э. д. с. $E = E_m \cos \omega t$. Обозначив напряжение на сетке через U , получим:

$$U'' + 2\beta U' + \omega_0^2 U = \omega_0^2 E_m \cos \omega t, \quad (1)$$

где, как и раньше, обозначено $\omega_0^2 = 1/LC$, $2\beta = 2\alpha - \omega_0^2 MS$.



Р и с. 118

Пусть характеристика лампы аппроксимирована многочленом 3-й степени, тогда

$$S = a_1 + 3a_3 U^2, \quad 2\beta = 2\beta_0 + \gamma U^2,$$

где $2\beta_0 = 2\alpha - \omega_0^2 MS_0$, $\gamma = 3\omega_0^2 Ma_3$. Используя эти соотношения и обозначения, перепишем (1) в виде

$$U'' + (2\beta_0 + \gamma U^2) U' + \omega_0^2 U = \omega_0^2 E_m \cos \omega t. \quad (2)$$

Решение уравнения (2) будем искать в виде

$$U = A \sin(\omega t + \varphi), \quad (3)$$

где A и φ — неизвестные амплитуда и фаза, являющиеся в общем случае функциями времени, что следует иметь в виду при дифференцировании. Так, первая производная (3) будет

$$U' = A' \sin(\omega t + \varphi) + A(\omega + \varphi') \cos(\omega t + \varphi).$$

Подставим решение (3) в (2); при подстановке $U^2 U'$ отбросим члены, содержащие третью гармонику. Отбросим также член, содержащий A'' (так как A — медленная функция по сравнению с $\sin \omega t$). После этого разложим правую часть (2), по формуле

$$\cos \omega t \equiv \cos \varphi \cos(\omega t + \varphi) + \sin \varphi \sin(\omega t + \varphi)$$

и приравняем порознь коэффициенты при $\sin(\omega t + \varphi)$ и $\cos(\omega t + \varphi)$ в правой и левой частях. Это даст нам два уравнения

$$\left(2\beta_0 + \frac{3}{4} \gamma A^2\right) A' + (\omega_0^2 - \omega_1^2) A = \omega_0^2 E_m \sin \varphi,$$

$$2\omega_1 A' + \left(2\beta_0 + \frac{1}{4} \gamma A^2\right) \omega_1 A = -\omega_0^2 E_m \cos \varphi. \quad (4)$$

Здесь обозначено

$$\omega_1 = \frac{d}{dt} (\omega t + \varphi) = \omega + \varphi'.$$

Мы будем в дальнейшем заниматься главным образом установившимся режимом, когда

$$A' = \varphi' = 0.$$

При этом условии получаем из (4) уравнения для установившихся амплитуды и фазы

$$\begin{aligned} (\omega_0^2 - \omega^2) A &= \omega_0^2 E_m \sin \varphi, \\ \left(2\beta_0 + \frac{1}{4} \gamma A^2\right) \omega A &= -\omega_0^2 E_m \cos \varphi. \end{aligned} \quad (5)$$

При отсутствии внешней э. д. с., т. е. при $E_m = 0$, первое уравнение даст нам частоту автоколебаний

$$\omega = \omega_0,$$

а второе уравнение — уже известное нам (см. § 21) значение установившейся амплитуды автоколебаний мягкого генератора

$$A_0 = \sqrt{-\frac{8\beta_0}{\gamma}} = \sqrt{\frac{4}{3} \frac{a_1 - RC/M}{a_3}}. \quad (6)$$

Возводя уравнения (5) в квадрат и складывая, получим

$$\left[(\omega_0^2 - \omega^2)^2 + \omega^2 \left(2\beta_0 + \frac{1}{4} \gamma A^2\right)^2 \right] A^2 = \omega_0^4 E_m^2. \quad (7)$$

Это соотношение позволяет найти амплитуду установившихся колебаний, которая зависит: 1) от амплитуды внешней э. д. с. E_m , 2) от расстройки $\omega_0^2 - \omega^2$, 3) от начального затухания $\beta_0 = \alpha - \omega_0^2 MS_0/2$ и 4) от нелинейности характеристики, выражаемой коэффициентом $\gamma = 3\omega_0^2 Ma_3$. Если найти из (7) амплитуду A как функцию частоты, то получим выражение для своего рода резонансных кривых. Здесь, однако, мы столкнемся с тем же затруднением, что и в предыдущем параграфе при исследовании нелинейного резонанса; относительно A мы получаем уравнение 6-й степени. Это затруднение преодолевается тем же приемом — обращением интересующей нас зависимости.

При некоторых соотношениях решения уравнения (7) неустойчивы. Границу области неустойчивости можно определить как место точек, в которых касательные к резонансным кривым вертикальны, — аналогично тому, как это делалось в предыдущем параграфе.

Положим, что расстройка мала, т. е. $\omega \cong \omega_0$ и обозначим

$$\Delta\omega = \omega_0 - \omega.$$

Воспользовавшись, кроме того, соотношением (6), можно преобразовать (7) к виду

$$\left[\frac{\Delta\omega^2}{\beta_0} + \left(1 - \frac{A^2}{A_0^2} \right)^2 \right] \frac{A^2}{A_0^2} = \frac{\omega_0^2 E_m^2}{4\beta_0^2 A_0^2}, \quad (8)$$

или, обозначая для краткости $\Delta\omega/\beta_0 = x$, $A^2/A_0^2 = y$, $\omega_0^2 E_m^2 / 4\beta_0^2 A_0^2 = E$, к виду

$$y [x^2 + (1 - y)^2] - E = F(x, y) = 0. \quad (9)$$

Таким образом, уравнение резонансных кривых (7) приведено к виду (9); в этой форме уравнение связывает переменные x и y . Первая переменная пропорциональна расстройке, вторая есть квадрат относительной амплитуды (т. е. амплитуды A при наличии внешней э. д. с., отнесенной к амплитуде A_0 автоколебаний при отсутствии внешней э. д. с.). График y как функции x изображает в известном масштабе резонансные кривые. Уравнение (9) является кубическим относительно y и квадратным относительно x . Поэтому для вычислений удобнее разрешать его относительно x . Геометрический смысл подобного обращения был разъяснен выше.

Исследуем производную

$$\frac{dy}{dx} = - \frac{\partial F / \partial x}{\partial F / \partial y}.$$

Уравнение координат максимумов мы получим, приравняв производную нулю, что даст условие $\partial F / \partial x = 0$. Мы имеем

$$\partial F / \partial x = 2yx = 0. \quad (10)$$

Отсюда следует, что максимумы располагаются на вертикальной прямой $x=0$, что соответствует отсутствию расстройки, т. е. $\omega = \omega_0$; ось x , уравнение которой есть $y=0$, является асимптотой семейства резонансных кривых.

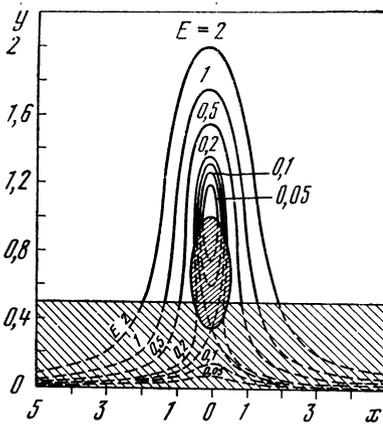
Производная dy/dx обращается в бесконечность (т. е. касательные к резонансным кривым вертикальны) в точках, для которых

$$\partial F / \partial y = 0.$$

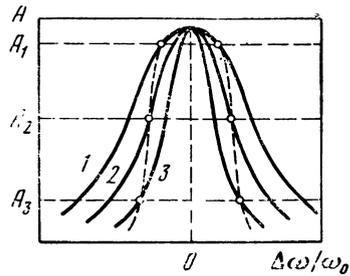
Это равенство является границей области устойчивых решений. Дифференцируя (9) по y , находим

$$\begin{aligned} \partial F / \partial y = x^2 + (1 - y)^2 - 2y(1 - y) = x^2 + \\ + (1 - y)(1 - 3y) = 0. \end{aligned} \quad (11)$$

Это уравнение эллипса с центром в точке $x=0$, $y=2/3$ и с вертикальной полуосью, равной $1/3$. Таким образом, верхняя горизонтальная касательная к эллипсу проходит на высоте $y=1$. Область внутри эллипса соответствует неустойчивым режимам. Подробное исследование устойчивости показывает, что кроме уже найденного нами существует еще одно условие устойчивости, а именно: $y > 1/2$, т. е. $A > A_0/\sqrt{2}$. Таким образом, при малых амплитудах «вы-



Р и с. 119



Р и с. 120

нужденных» колебаний режим оказывается неустойчивым. Горизонтальная прямая

$$y = 1/2 \quad (12)$$

является нижней границей области устойчивости.

Графики функции (9) совместно с граничными линиями (11) и (12) представлены на рис. 119. Области неустойчивости заштрихованы.

Нужно заметить, что для пояснения характера зависимости (7) прибегают часто к нестрогому, но зато очень наглядному приему. Прием этот состоит в том, что формула (7) переписывается в виде

$$A = \frac{E_m}{\sqrt{\left(\frac{\omega_0^2 - \omega^2}{\omega_0^2}\right)^2 + \frac{\omega^2}{\omega_0^2} \left(\frac{2\beta_0}{\omega_0} + \frac{1}{4} \frac{\gamma}{\omega_0} A^2\right)^2}},$$

или, полагая расстройку малой,

$$A \cong \frac{E_m}{\sqrt{\varepsilon^2 + \left(d_0 + \frac{1}{4} \frac{\gamma}{\omega_0} A^2\right)^2}}, \quad (13)$$

где $\varepsilon = 2\Delta\omega/\omega_0$ — обычное обозначение относительной расстройки, $d_0 = 2\beta_0/\omega_0$ — начальное затухание. Последняя формула рассматривается как выражение для резонансной кривой, отличающееся от обычного тем, что затухание зависит от амплитуды (второй член под корнем).

При такой трактовке можно подойти к построению зависимости (13), исходя из семейства обычных резонансных кривых при различном затухании. Такое семейство изображено на рис. 120. При больших амплитудах (значение A_1 на рис. 120) затухание велико; такому значению затухания соответствует резонансная кривая 1. При меньшей амплитуде A_2 затухание соответственно меньше, и мы переходим на более узкую резонансную кривую 2.

Еще меньшей амплитуде A_3 и еще меньшему затуханию соответствует резонансная кривая \mathcal{Z} . Соединяя отмеченные на рисунке точки штриховой линией, получаем очертание резонансной кривой.

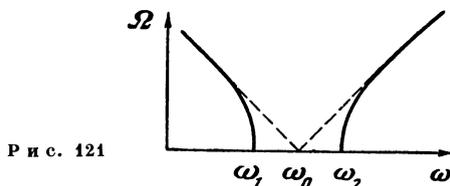
Заметим в заключение этого параграфа, что условие устойчивости на основании первого уравнения (5) может быть выражено через угол φ (определяющий сдвиг фазы между внешней э. д. с. и свободным автоколебанием) в виде

$$-\pi/2 < \varphi < \pi/2.$$

§ 34. Захватывание

Если в линейной системе существуют одновременно два колебания с частотами ω и ω_0 , то эти два колебания интерферируют и образуют биения. Частота биений равна абсолютной разности частот, т. е. $\Omega = |\omega - \omega_0|$. Если ω_0 — постоянная, то с изменением ω частота биений изменяется по закону, представленному двумя штриховыми прямыми на рис. 121.

В автоколебательной системе, находящейся под воздействием внешней э. д. с., также наблюдаются биения, но при изменении частоты явление протекает иначе. При возрастании ω частота биений сперва убывает (по мере приближения ω к ω_0) по линейному закону. Затем частота биений начинает убывать все быстрее (сплошная линия на рис. 121) и при некотором значении $\omega = \omega_1$



Р и с. 121

биения срываются; генератор при этом генерирует частоту ω_1 , несмотря на то, что настроен на частоту ω_0 . При дальнейшем повышении частоты до значения $\omega = \omega_2 > \omega_0$ это положение сохраняется: генерируется частота, равная частоте внешней э. д. с.; биения отсутствуют. Наконец, при $\omega > \omega_2$ снова возникают биения¹.

¹ По этому поводу следует заметить, что самый термин *биения* подразумевает наличие двух отдельных колебаний, результатом сложения которых являются биения. В случае нелинейной автоколебательной системы мы имеем дело с единым нелинейным процессом, характеризующимся вне полосы захватывания периодическим изменением амплитуды. Вдали от полосы захватывания картина явления в самом деле такова, как если бы имелись две составляющие процесса — «автоколебание» с частотой ω_0 и «вынужденное колебание» с частотой ω . Однако такое представление имеет формальный характер, так как эти составляющие, вообще говоря, физически неразделимы. Поэтому термины «биения», «режим биений», хотя и применяются к рассматриваемым явлениям, но имеют здесь условный смысл.

Описанное явление, состоящее в том, что внешняя э. д. с. навязывает генератору свою частоту, носит название *захватывания*. Частотный интервал $\omega_2 - \omega_1$, в пределах которого частота генератора следует за частотой внешней э. д. с., называется *полосой захватывания*. Наша ближайшая задача состоит в том, чтобы выяснить, какова ширина этой полосы и от чего она зависит.

Для решения этой задачи обратимся к картине явлений, представленной резонансными кривыми рис. 119. Устойчивый режим, представленный на рис. 119 сплошными линиями, — это и есть режим захватывания, когда происходят вынужденные колебания с частотой внешней э. д. с. Переход в неустойчивую область означает возникновение биений. Таким образом, ширина полосы захватывания (интервал $\omega_2 - \omega_1$ на рис. 121) есть не что иное, как ширина резонансной кривой на границе области неустойчивости. Исходя из этих соображений, мы теперь легко определим ширину полосы захватывания.

Уравнение резонансных кривых рис. 119 есть

$$y [x^2 + (1 - y)^2] - E = 0. \quad (1)$$

При малых значениях внешней э. д. с., т. е. при малых E , кривые опираются на вершину эллипса на высоте $y=1$. Подставляя это значение в (1), находим

$$x^2 = E$$

или, раскрывая сокращенные обозначения (см. § 33),

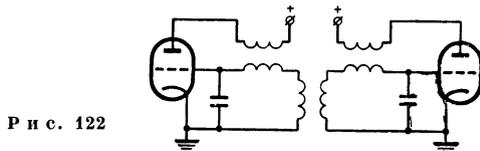
$$2\Delta\omega/\omega_0 = E_m/A_0.$$

Величина слева — это и есть относительная ширина полосы захватывания. Она равна, оказывается, отношению амплитуды E_m внешней э. д. с. к амплитуде A_0 напряжения автоколебания при отсутствии внешней э. д. с. Считая величину A_0 постоянным параметром рассматриваемой системы, заключаем, что ширина полосы захватывания тем больше, чем сильнее внешнее воздействие.

При больших E_m резонансные кривые опираются уже не на эллипс, а на горизонтальную прямую $y=1/2$. Подставляя это значение в (1), найдем $x^2 = 2E - 1/4$ или $2\Delta\omega/\omega_0 = \sqrt{2E_m^2/A_0^2 - \beta_0^2/\omega_0^2}$. Но второй член под корнем выражает квадрат начального (т. е. с учетом действия обратной связи) затухания системы. Это во всяком случае очень малая величина, так что можно положить в рассматриваемом случае $2\Delta\omega/\omega_0 \cong \sqrt{2}E_m/A_0$. Для промежуточных значений E_m , когда резонансные кривые опираются на скаты эллипса, можно положить $2\Delta\omega/\omega_0 = kE_m/A_0$, где, как мы установили, $1 < k < \sqrt{2}$. Точное значение k может быть найдено, в случае надобности, путем совместного решения уравнений (9) и (11) предыдущего параграфа.

Явление захватывания используется в радиотехнике для синхронизации генератора внешней э. д. с. или для взаимной син-

хронизации двух генераторов. Если два генератора связаны друг с другом, как показано на схеме рис. 122, то каждый из генераторов наводит в контуре другого э. д. с., являющуюся для этого другого генератора внешней. Работа схемы описывается системой из двух уравнений типа (1) (§ 33), причем правые части играют роль членов связи. Если оба генератора одинаковы, но настроены

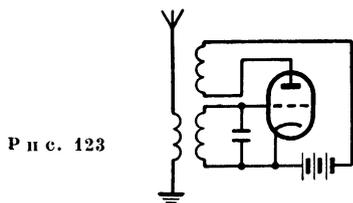


Р и с. 122

на разные частоты, то они могут взаимно синхронизироваться на некоторой промежуточной частоте, имеющей в общем случае два возможных значения. Если же один из генераторов является более мощным, то он может рассматриваться как источник независимой э. д. с. и, следовательно, захватывает менее мощный генератор, заставляя его работать на своей частоте.

§ 35. Регенеративный приемник

Схема простейшего регенеративного приемника, изображенная на рис. 123, сводится к схеме рис. 118. Однако в отличие от ранее рассмотренного случая регенеративный приемник характеризуется тем, что он работает в недовозбужденном режиме, хотя и вблизи порога возбуждения, так как при этом получается наибольшее усиление.



Р и с. 123

Уравнение регенеративного приемника совпадает с уравнением (7) § 33

$$\left[(\omega_0^2 - \omega^2)^2 + \omega^2 \left(2\beta_0 + \frac{1}{4} \gamma A^2 \right)^2 \right] A^2 = \omega_0^4 E_m^2. \quad (1)$$

Рассмотрим прежде всего вопрос о чувствительности приемника при точной настройке на принимаемую частоту, т. е. при $\omega_0 = \omega$. При этом условии из (1) получим

$$\left(2\beta_0 + \frac{1}{4} \gamma A^2 \right) A = \omega_0 E_m. \quad (2)$$

Увеличивая обратную связь, можно довести приемник почти до самовозбуждения. При этом начальный показатель затухания $\beta_0 = \alpha - \omega_0^2 MS/2$ близок к нулю, и мы имеем

$$\frac{1}{4} \gamma A^3 = \omega_0 E_m,$$

откуда

$$A = \left(4 \frac{\omega_0}{\gamma} E_m \right)^{1/3} = k E_m^{1/3}. \quad (3)$$

Найдем чувствительность (или, что то же, усиление) приемника, определив ее как $\nu = dA/dE_m$. Дифференцируя (3), получаем $\nu = k E_m^{-2/3}$. Как видим, при E_m , стремящемся к нулю, чувствительность стремится к бесконечности. Таким образом, регенеративный приемник обладает переменной чувствительностью тем большей, чем слабее принимаемый сигнал. Это является, конечно, выгодным свойством рассматриваемого приемника.

Практически предельно-достижимая чувствительность ограничивается собственным шумом приемника, который вблизи порога самовозбуждения становится очень сильным благодаря большому усилению.

Регенеративный приемник сыграл в свое время большую роль благодаря исключительной простоте и высокой чувствительности. Однако регенеративный приемник обладает и очень существенными недостатками: неустойчивостью, плохой избирательностью. К тому же при работе на самом пороге возбуждения часто возникает генерация, создающая помеху для близко расположенных приемников. Поэтому регенеративный приемник был вытеснен более совершенными приемниками, а регенеративная схема сохранила значение лишь в качестве так называемого регенеративного фильтра.

§ 36. Импульсная синхронизация

В § 33 рассматривалась синхронизация генератора синусоидальным внешним напряжением.

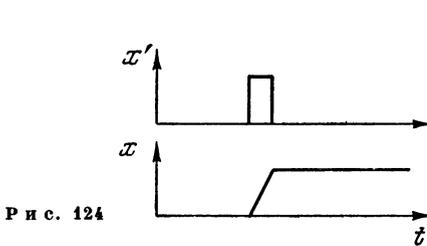
В этом параграфе будет рассмотрена синхронизация периодическими короткими импульсами. Этот случай допускает очень наглядное качественное объяснение процесса синхронизации, и мы воспользуемся этой возможностью; аналитического исследования импульсной синхронизации мы делать не будем.

Мы исходим по-прежнему из схемы рис. 118, но предполагаем теперь, что внешняя э. д. с. представляет собою периодическую последовательность коротких импульсов произвольной формы. Для объяснения механизма синхронизации мы представим явление в виде диаграмм на фазовой плоскости. Прежде всего нам нужно дать графическое представление действия импульса напряжения.

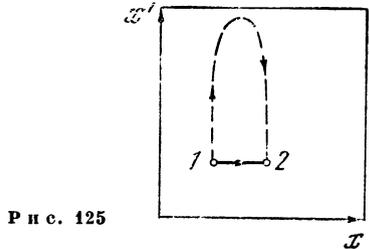
Возьмем в качестве переменных, откладываемых по координатным осям те же величины, что и в примере в конце § 25, а именно:

$$x = \int U dt, \quad x' = U.$$

Пусть импульс характеризуется тем, что напряжение начинает изменяться от некоторого начального значения (в частности, равного нулю), а затем по прошествии небольшого времени снова возвращается к исходному значению. Таким образом, по окончании импульса напряжение $U=x'$ остается неизменным, а интеграл



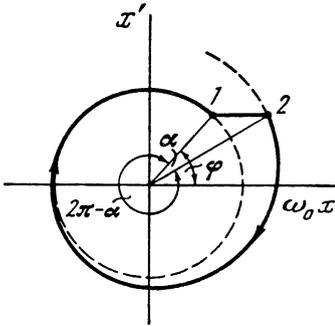
Р и с. 124



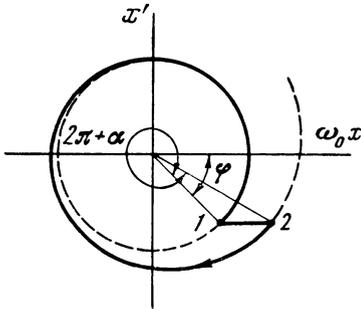
Р и с. 125

от напряжения получает приращение, определяемое площадью импульса. Эти соотношения поясняются на рис. 124 для случая прямоугольного импульса. Если теперь представить те же соотношения на фазовой плоскости в координатах x, x' , то изображающая точка, находившаяся первоначально в положении 1, перейдет под действием импульса в новое положение 2, описав при этом намеченную штрихом траекторию (рис. 125). Время перехода равно длительности импульса. Мы будем считать, что импульс настолько короток, что изображающая точка перебрасывается из 1 в 2 практически мгновенно. Что касается формы траектории, то она зависит, конечно, от формы импульса. Но нас, собственно, не интересует вид траектории; для дальнейшего существенно лишь то, что изображающая точка смещается в конечном счете параллельно оси x (так как напряжение до и после импульса одинаково) на отрезок, пропорциональный площади импульса. Это смещение изображающей точки и выражает действие импульса. Мы обозначим его через h .

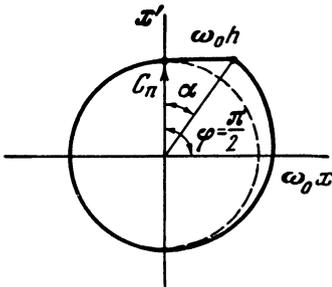
Теперь обратимся к действию синхронизирующих импульсов на работающий генератор. При отсутствии импульсов режим генератора выражается устойчивым предельным циклом. При наличии же импульсов должен образоваться замкнутый цикл, включающий в себя скачок, обусловленный действием импульса. Такой цикл может образоваться лишь из отрезка спирали, как показано на рис. 126. По оси абсцисс отложено $\omega_0 x$ вместо x для того, чтобы уравнять масштабы по обеим осям. При таком согласованном масштабе предельный цикл имеет форму окружности. Будем считать, что изображающая точка движется по спирали с постоянной угло-



Р и с. 126



Р и с. 127



Р и с. 128

вой скоростью и что поворот на 2π соответствует периоду автоколебаний $T_0 = 2\pi/\omega_0$. Тогда из рис. 126 сразу видно, что изображающая точка обходит замкнутый цикл, отмеченный на рис. 126 сплошной линией, за время $T < T_0$. Это следует из того, что угловое перемещение точки равно $2\pi - \alpha < 2\pi$, а следовательно, на такое перемещение затрачивается меньшее время. Участок скачка (1—2) проходит, по нашему предположению, мгновенно. Рис. 126 соответствует, стало быть, случаю, когда импульсы синхронизируют генератор на частоте $\omega > \omega_0$. Фазовая диаграмма для случая синхронизации на частоте $\omega < \omega_0$ изображена на рис. 127. Здесь изображающая точка проходит за цикл по спирали угловой путь $2\pi + \alpha > 2\pi$ и затрачивает на это время $T > T_0$. Подобного рода диаграммы позволяют ответить на вопрос о ширине полосы захватывания. Будем увеличивать угол φ , определяющий фазу синхронизирующего импульса относительно свободного колебания генератора. Наибольшее значение угла α , т. е. наибольшая расстройка $\Delta\omega = |\omega - \omega_0|$, получится, очевидно, при $|\varphi| = \pi/2$.

Соотношения, получаемые при этом условии, поясняются рис. 128. Мы имеем

$$\operatorname{tg} \alpha = \omega_0 n / C_0,$$

где C_0 — амплитуда автоколебаний, соответствующая предельному циклу. В действительности α малый угол (на наших рисунках быстрота свертывания спирали грубо преувеличена). Поэтому $\operatorname{tg} \alpha \cong \alpha$. Но α есть мера изменения периода (против периода авто-

колебаний, равного 2π в угловой мере). Таким образом, $\alpha/2\pi = \Delta\omega/\omega_0$ и, следовательно,

$$\frac{\Delta\omega}{\omega_0} \cong \frac{\omega_0}{2\pi} \cdot \frac{h}{C_0}.$$

Это есть наибольшая относительная расстройка в одну сторону. Полная же относительная ширина полосы захватывания равна

$$\frac{2\Delta\omega}{\omega_0} \cong \frac{\omega_0}{\pi} \cdot \frac{h}{C_0} \cong \frac{\omega h}{\pi C_0}.$$

Можно еще преобразовать этот результат, учитывая, что $\omega h/\pi$ есть не что иное, как амплитуда первой гармоники в разложении периодической последовательности коротких импульсов в ряд Фурье. Эта составляющая, имеющая частоту ω , и производит синхронизирующее действие. Обозначая

$$\frac{1}{\pi} \omega h = E_1,$$

получим для полосы захватывания

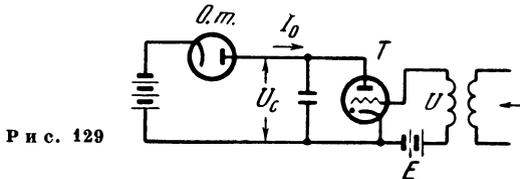
$$2\Delta\omega/\omega_0 \cong E_1/C_0,$$

т. е. тот же результат, что и в § 34.

Заметим в заключение, что связь между расстройкой и углом φ , очевидная из рис. 126 и 127, — это та же самая связь, которая указывается уравнениями (5) § 33.

§ 37. Синхронизация релаксационного генератора

Разберем очень простой механизм синхронизации релаксационного генератора внешним синусоидальным напряжением. В качестве генератора выберем схему с тиратроном (рис. 129). Эта схема отличается от схемы релаксационного генератора с двухэлектродной газоразрядной лампой (рис. 105) тем, что напряжение зажигания U_2 зависит от напряжения на сетке тиратрона T ,



Р и с. 129

а именно напряжение зажигания тем выше, чем больше отрицательное напряжение на сетке. Напряжение потухания U_1 практически неизменно. Заметим, что в схему введен ограничитель тока O . т., так что зарядный ток предполагается постоянным, а следовательно, напряжение на конденсаторе в процессе зарядки нарастает линейно.

При отсутствии внешнего напряжения процесс в генераторе протекает, как было описано в § 29: конденсатор заряжается до напряжения зажигания U_2 ; при этом значении напряжения возникает большая ионная проводимость, и конденсатор быстро (практически мгновенно) разряжается до напряжения потухания U_1 ; затем возобновляется зарядка конденсатора.

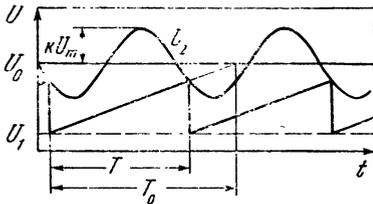


Рис. 130

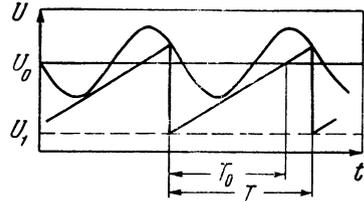


Рис. 131

Включим теперь на сетку внешнее синхронизирующее синусоидальное напряжение

$$U = U_m \sin \omega t.$$

Напряжение зажигания U_2 будет теперь переменной величиной. Положим, что напряжение зажигания изменяется пропорционально напряжению на сетке. Будем иметь

$$U_2(t) = k(E + U_m \sin \omega t) = U_0 + kU_m \sin \omega t,$$

где U_0 — напряжение зажигания при отсутствии переменного внешнего напряжения (т. е. при постоянном напряжении E). Зажигание происходит при

$$U_c = U_2,$$

т. е. когда напряжение на конденсаторе достигнет напряжения зажигания. В этот момент происходит скачок напряжения. Напряжение на конденсаторе растет во время зарядки по линейному закону

$$U_c = U_1 + \frac{I_0}{C} t.$$

На рис. 130 построен график колебания, который изображает периодический процесс, и мы видим, что период колебания получается равным периоду внешнего напряжения T и *неравным* периоду свободных колебаний генератора T_0 . Иначе говоря, внешнее напряжение захватывает частоту генератора; происходит синхронизация генератора внешним напряжением.

Выведем основные соотношения для установившегося синхронизированного колебания. Если вести отсчет времени от скачка напряжения, то напряжение зажигания нужно записать в виде

$$U_2 = U_0 - kU_m \sin(\omega t + \varphi).$$

Следующий скачок произойдет при

$$U_c = U_0 - kU_m \sin(\omega T + \varphi) = U_0 - kU_m \sin \varphi.$$

Но, с другой стороны, в момент $t = T$

$$U_c = U_1 + \frac{I_0}{C} T = U_1 + \alpha T.$$

Таким образом,

$$U_1 + \alpha T = U_0 - kU_m \sin \varphi.$$

Воспользовавшись соотношением

$$U_1 + \alpha T = U_0.$$

получим

$$\alpha(T_0 - T) = kU_m \sin \varphi$$

или

$$\frac{T_0 - T}{T_0} = \frac{kU_m}{U_0 - U_1} \sin \varphi, \quad (1)$$

или

$$\frac{T}{T_0} = 1 - \frac{kU_m}{U_0 - U_1} \sin \varphi. \quad (2)$$

При $\sin \varphi < 0$ период получает положительное приращение, т. е. увеличивается. Соответствующий график изображен на рис. 131. Если считать, что относительное изменение периода невелико, то его можно приравнять относительному изменению частоты, т. е. положить (см. (1))

$$\frac{\omega_0 - \omega}{\omega_0} \cong \frac{kU_m}{U_0 - U_1} \sin \varphi.$$

Крайние значения $\sin \varphi$ равны ± 1 ; отсюда находим полную ширину полосы захватывания

$$\frac{2\Delta\omega}{\omega_0} \cong \frac{2kU_m}{U_0 - U_1}.$$

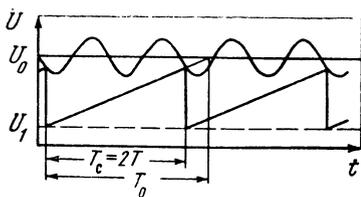
Исследование устойчивости синхронизированного режима показывает, что условие устойчивости выражается через фазу синхронизации точно так же, как в § 33, а именно:

$$-\pi/2 < \varphi < \pi/2.$$

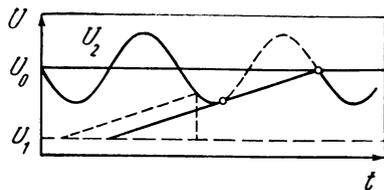
Выведенные соотношения требуют еще уточнения. Во-первых, нужно заметить, что равенство

$$U_c = U_0 - kU_m \sin \varphi,$$

из которого мы определяем момент скачка напряжения, может удовлетворяться не только при $t = T$, но и при $t = nT$, где n —



Р и с. 132

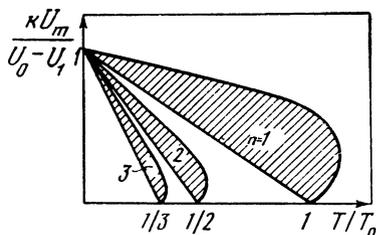


Р и с. 133

любое целое число. Таким образом, в формулах (1) и (2) следует писать в общем случае

$$T_c = nT,$$

где T_c — период синхронизированных колебаний. На рис. 132 в пояснение сказанного приведен график синхронизированного колебания при $k=2$. Это называется иногда синхронизацией на субгармонике. Во-вторых, при заданных U_0 , U_1 , α , kU_m и T некоторые режимы неосуществимы, так как прямая U_c касается



Р и с. 134

синусоиды U_2 и не может попасть на известный участок синусоиды. Это показано на рис. 133, на котором недостижимый участок синусоиды отмечен штрихом.

Принимая все это во внимание, можно построить график, представляющий области возможных режимов синхронизации при различных кратностях n (рис. 134); эти области на рисунке заштрихованы.

Глава IV

ПАРАМЕТРИЧЕСКИЕ ЯВЛЕНИЯ

В этой главе будут рассмотрены некоторые явления и процессы в системах с переменными (зависящими от времени) параметрами, математическое описание которых сводится к линейным уравнениям с переменными коэффициентами. Теория этих уравнений относительно сложна и мало разработана. Имеется лишь достаточно полная теория уравнения Хилла

$$x'' + [a + qf(t)]x = 0,$$

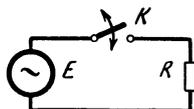
где $f(t)$ — периодическая функция. Частный вид уравнения Хилла

$$x'' + [a + 2q \cos 2t] x = 0$$

называется уравнением Матё. Краткие сведения из теории уравнения Матё будут даны в § 40. В предшествующих же параграфах рассматривались только параметрические явления, описываемые алгебраическими уравнениями.

§ 38. Синхронное детектирование

Исследование синхронного детектирования мы начнем с простейшего примера синхронного выпрямления. Пусть имеется цепь, состоящая из источника переменного напряжения E , сопротивления нагрузки R и синхронного ключа K (рис. 135).



Р и с. 135

Ключ имеет независимое управление, однако синхронное с частотой источника. Действие ключа состоит в том, что он попеременно то замыкает цепь на время, равное половине периода переменного напряжения, то размыкает цепь на такое же время. Таким образом, проводимость цепи равна либо нулю, либо $1/R$ в зависимости от того, разомкнут или замкнут ключ. При таких обстоятельствах можно получить в цепи выпрямленный ток. Эта возможность поясняется графиками рис. 136.

Необходимо сразу же подчеркнуть, что такое выпрямление принципиально отличается от обычного. При обычном выпрямлении проводимость цепи есть также переменная величина, но *проводимость зависит от напряжения* на выпрямителе. Таким образом, схема обычного выпрямителя есть *нелинейная* цепь. В рассмотренном же здесь случае синхронного выпрямителя переменная *проводимость зависит от времени* и, следовательно, мы имеем дело с цепью с переменными параметрами, которая может быть описана *линейным* уравнением с переменными коэффициентами.

Особенности синхронного выпрямления сразу обнаружатся, если мы заметим, что при независимой (хотя и синхронной) работе ключа фаза коммутации произвольна. Легко представить себе, что если фаза коммутации изменится на 180° , то изменится полярность выпрямленного напряжения на сопротивлении R . Если же сдвинуть фазу на 90° или 270° , то выпрямления вообще не получится. Это пояснено графиками рис. 137. Как видим, при сдвиге на 90° , т. е. во времени на $T/4$, ток в цепи не имеет

постоянной составляющей. Такого рода зависимость выпрямляющего действия от фазы характерна для синхронного выпрямления и для синхронного детектирования. Мы будем еще обсуждать вытекающие из этого свойства следствия.

Обратимся теперь к синхронному детектированию АМ колебания. Задача детектирования состоит в том, чтобы восстановить низкочастотный модулирующий сигнал. Пусть дано модулированное напряжение вида

$$E = E_0 [1 + mf(t)] \sin \psi_0 t. \quad (1)$$

Подадим это напряжение на цепь, в которую включено переменное сопротивление (рис. 138), изменяющееся периодически с той же

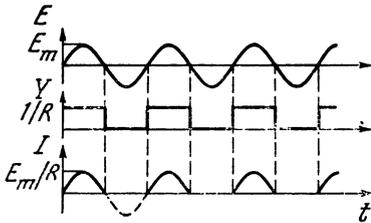


Рис. 136

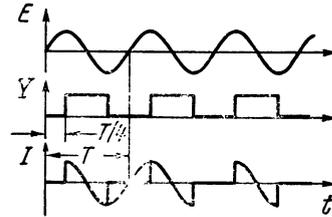


Рис. 137

частотой ω_0 , так что проводимость цепи может быть записана в виде

$$Y = Y_0 (1 + n \sin \omega_0 t), \quad (2)$$

т. е. предположим, что проводимость меняется синхронно и синфазно с напряжением несущей частоты. Для тока в цепи получим, перемножив (1) и (2),

$$I = EY = E_0 Y_0 \left\{ [1 + mf(t)] \sin \omega_0 t - \right. \\ \left. - \frac{n}{2} [1 + mf(t)] \cos 2\omega_0 t + \frac{n}{2} [1 + mf(t)] \right\}.$$

Первое слагаемое представляет собою исходное модулированное колебание с высокочастотным спектром, расположенным симметрично около частоты ω_0 . Второе слагаемое имеет аналогичный (но в $n/2$ раз меньший по плотности) спектр, расположенный около частоты $2\omega_0$. И, наконец, третье слагаемое является модулирующим низкочастотным сигналом. Высокочастотные слагаемые легко отделяются соответствующими фильтрами, которые должны предусматриваться в схеме любого детектора. Не менее легко освободиться и от постоянной составляющей. После такой очистки мы получим составляющую тока

$$I = \frac{1}{2} nm E_0 Y_0 f(t), \quad (3)$$

что и представляет собою желаемый результат детектирования.

Посмотрим теперь, как повлияет на результат нарушение синфазности. Пусть в отличие от предыдущего

$$Y = Y_0 [1 + n \sin(\omega_0 t + \varphi)]. \quad (4)$$

Перемножая (1) и (4), найдем

$$I = EY = E_0 Y_0 [1 + mf(t)] [1 + n \sin(\omega_0 t + \varphi)] \sin \omega_0 t.$$

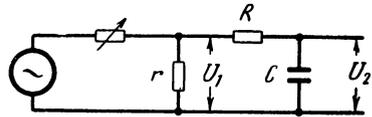
Выполнив действия и отбросив высокочастотные и постоянную составляющие, найдем, что интересующая нас составляющая тока, содержащая сигнал $f(t)$, равна

$$I = \frac{1}{2} nm E_0 Y_0 f(t) \cos \varphi. \quad (5)$$

Как видим, выгоднее всего вести детектирование при $\varphi=0$ (или π), при этом (5) переходит в (3). Если же $\varphi = \pm \pi/2$, то детектирования не получается.



Р и с. 138



Р и с. 139

Наконец, нам необходимо выяснить, что получится при нарушении условия синхронности. Здесь обнаружится, что синхронный детектор обладает *избирательностью*. Для исследования этого свойства рассмотрим конкретную схему с простейшим фильтром в виде RC — звена (рис. 139). Предполагается, что сопротивление r , с которого снимается напряжение U_1 , достаточно мало, так что присоединение параллельно ему звена RC не изменяет режима основной схемы синхронного выпрямителя. Мы имеем

$$U_1 = Ir = EYr. \quad (6)$$

Положим, что

$$E = E_0 \sin \omega_0 t, \quad (7)$$

$$Y = Y_0 (1 + n \sin \omega t), \quad (8)$$

где ω — произвольная частота. Для комплексной амплитуды напряжения на выходе при некоторой частоте Ω имеем

$$\dot{U}_2 = \dot{U}_1 \frac{1}{1 + j\Omega RC}. \quad (9)$$

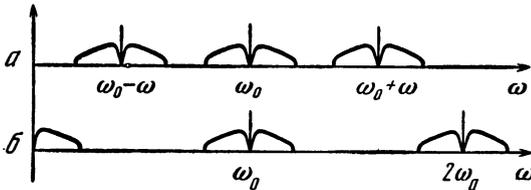
Найдем U_1

$$\begin{aligned} U_1 &= EYr = E_0 Y_0 r (1 + n \sin \omega t) \sin \omega_0 t = \\ &= E_0 Y_0 r \left[\sin \omega_0 t + \frac{n}{2} \cos(\omega_0 - \omega) t - \frac{n}{2} \cos(\omega_0 + \omega) t \right]. \end{aligned}$$

Таким образом, на входе звена RC мы имеем три напряжения с частотами ω_0 , $\omega_0 - \omega$ и $\omega_0 + \omega$ и комплексными амплитудами соответственно $E_0 Y_0 r$, $j \frac{n}{2} E_0 Y_0 r$ и $-j \frac{n}{2} E_0 Y_0 r$. Подставляя эти величины в формулу (9), получим следующие значения комплексных амплитуд выходного напряжения U_2 для трех перечисленных выше составляющих

$$\frac{E_0 Y_0 r}{1 + j \omega_0 RC}, \quad \frac{j \frac{n}{2} E_0 Y_0 r}{1 + j (\omega_0 - \omega) RC}, \quad \frac{-j \frac{n}{2} E_0 Y_0 r}{1 + j (\omega_0 + \omega) RC}.$$

При небольшой расстройке, т. е. при $\omega \cong \omega_0$ и при большой постоянной времени, т. е. при $\omega_0 RC \gg 1$, можно пренебречь первой и



третьей величинами по сравнению со второй, так что приблизительно

$$|\dot{U}_2| \cong \frac{E_0 Y_0 r n}{2RC} \cdot \frac{1}{\sqrt{(\omega_0 - \omega)^2 + 1/RC^2}}.$$

Но это выражение совпадает с выражением для обычной резонансной кривой. Особенность рассматриваемого случая состоит в том, что роль затухания играет величина $1/RC$. Следовательно, чем больше постоянная времени RC , тем острее «резонансная кривая» синхронного детектора. Однако максимум чувствительности получается, когда разностная частота $\omega_0 - \omega$ обращается в нуль, т. е. при синхронном режиме.

Использовать асинхронный режим для приема как угодно модулированных сигналов нельзя. Действительно, пусть вместо синусоидального напряжения (7) имеется амплитудно-модулированное напряжение

$$E = E_0 [1 + mf(t)] \sin \omega_0 t.$$

Составляя произведение

$$\begin{aligned} U_1 &= EYr = E_0 Y_0 r [1 + mf(t)] (1 + n \sin \omega t) \sin \omega_0 t = \\ &= E_0 Y_0 r [1 + mf(t)] \left[\sin \omega_0 t + \right. \\ &\quad \left. + \frac{n}{2} \cos (\omega_0 - \omega) t - \frac{n}{2} \cos (\omega_0 + \omega) t \right], \end{aligned}$$

мы видим, что спектр напряжения U_1 воспроизводит спектр E на частотах ω_0 , $\omega_0 - \omega$ и $\omega_0 + \omega$ в качестве несущих, как показано

на рис. 140, а. Как видим, при соотношениях рис. 140, а получается не детектирование, а преобразование частоты. При малых расстройках, когда $\omega_0 - \omega$ меньше ширины спектра модулирующей функции, возникают искажения. И лишь при $\omega = \omega_0$, т. е. при синхронном режиме, получается правильное детектирование (рис. 140, б).

В заключение заметим, что если входное напряжение E — периодическое, то, изменяя частоту ω , можно поочередно выделить и измерить гармонические составляющие E . При этом синхронный детектор работает в сущности как гармонический анализатор. Разрешающая способность такого анализатора, т. е. способность различить и измерить порознь близко расположенные на шкале частот гармонические составляющие, зависит от постоянной времени RC и может быть сделана очень большой.

§ 39. Фазовое детектирование и разделение

Отмечавшаяся в предыдущем параграфе чувствительность синхронного детектора к фазе детектируемого напряжения может иметь технические применения.

Для начала рассмотрим прием радиотелеграфных сигналов при фазовой модуляции. Речь идет о сигнале, в котором при нажатии и отжатии телеграфного ключа фаза несущей частоты меняется скачком, как показано на рис. 141. График первой строки показывает изменение фазы при передаче буквы «А» кодом

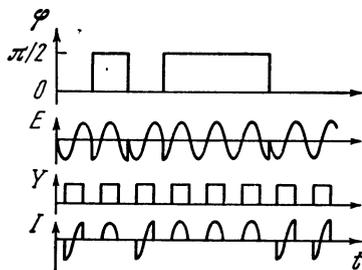


Рис. 141

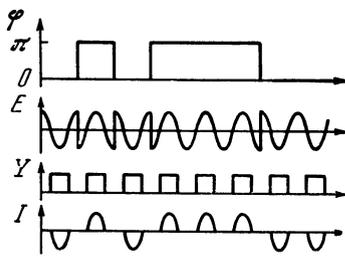


Рис. 142

Морзе. Во второй строке изображено модулированное напряжение. В третьей строке дан график проводимости — предполагается, что применено простейшее устройство с синхронным ключом рис. 135. В четвертой строке изображен получаемый ток. После усреднения, т. е. фильтрации, мы воспроизведем с тем или иным приближением переданный телеграфный сигнал. Нужно сразу пояснить, что на рисунке на каждую телеграфную точку приходится только один период несущей частоты, что не соответствует действительности (такое соотношение выбрано только для того, чтобы не потерять наглядности рисунка). В действительности

несущая частота может быть, скажем, в 10^4 раз больше частоты манипуляции, т. е. на одну точку приходится порядка 10^4 периодов несущей частоты. Если модулировать изменением фазы не на $\pi/2$, а на π , то получится картина, показанная на рис. 142 и не требующая пояснений. Итак, цепь с периодически изменяющейся проводимостью может применяться в качестве фазового детектора.

Обратимся теперь к фазовому разделению. Задача разделения сигналов возникает при многоканальной связи, когда по одной линии передаются одновременно несколько разных сигналов. На приемном конце многоканальной системы связи необходимо разделить поступающую смесь сигналов и выделить сигнал, принадлежащий каждому данному каналу. При частотной системе многоканальной связи для передачи сигналов, принадлежащих различным каналам, применяют отдельные несущие (или поднесущие) для каждого канала. При этом можно расположить спектры сигнала каждого канала на шкале частот так, что они не будут перекрываться. Разделение на приемном конце производится обычными полосовыми фильтрами (так называемая «расфильтровка»). При фазовом разделении в качестве несущих применяются колебания одинаковой частоты, но сдвинутые по фазе. Рассмотрим вопрос о возможности фазового разделения аналитически.

Пусть сигнал одного канала есть $f_1(t)$, а другого — $f_2(t)$. Промодулируем одним сигналом напряжение несущей частоты вида

$$E_1 = E_{01} \sin \omega_0 t,$$

а для другого сигнала возьмем несущее напряжение вида

$$E_2 = E_{02} \sin (\omega_0 t + \psi).$$

Тогда на приемном конце получим сложное модулированное напряжение

$$\begin{aligned} E &= E_{01} [1 + m_1 f_1(t)] \sin \omega_0 t + E_{02} [1 + m_2 f_2(t)] \sin (\omega_0 t + \psi) = \\ &= a(t) \sin \omega_0 t + b(t) \sin (\omega_0 t + \psi). \end{aligned} \quad (1)$$

Подадим это напряжение в цепь, проводимость которой меняется по закону

$$Y = Y_0 (1 + n \sin \omega_0 t). \quad (2)$$

Для нахождения тока перемножим (1) и (2)

$$\begin{aligned} I &= EY = Y_0 (1 + n \sin \omega_0 t) [a(t) \sin \omega_0 t + \\ &+ b(t) \sin (\omega_0 t + \psi)] = Y_0 \left[a(t) \sin \omega_0 t + b(t) \sin (\omega_0 t + \psi) + \frac{n}{2} a(t) - \right. \\ &- \frac{n}{2} a(t) \cos 2\omega_0 t + \frac{n}{2} b(t) \cos \psi - \frac{n}{2} b(t) \cos \psi \cos 2\omega_0 t + \\ &\left. + \frac{n}{2} b(t) \sin \psi \sin 2\omega_0 t \right]. \end{aligned}$$

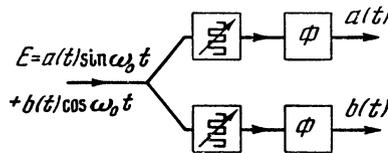
Как видим, кроме составляющих высоких частот, группирующихся около ω_0 и $2\omega_0$, в составе тока имеются составляющие, пропорциональные $a(t)$ и $b(t)$. Однако нам нужно, чтобы в результате обработки модулированного напряжения приемным устройством сигналы $a(t)$ и $b(t)$ были выделены порознь. Предыдущая формула показывает, что мы достигнем желаемого результата при $\psi = \pi/2$. Тогда после отфильтровки высокочастотных составляющих получим

$$I = \frac{1}{2} Y_0 a(t).$$

Итак, для того чтобы при помощи синхронного детектирования разделить два канала, необходимо взять несущие колебания в виде

$$E_1 = E_{10} \sin \omega_0 t, \quad E_2 = E_{02} \cos \omega_0 t.$$

На приемном конце нужно поставить два синхронных детектора, проводимости которых изменяются также по закону синуса и косинуса. При этих условиях мы получим на выходах фильтров разделенные сигналы $a(t)$ и $b(t)$, как показано на скелетной схеме рис. 143. Из всего сказанного следует, что описанным путем можно



Р и с. 143

осуществить всего лишь двухканальную систему передачи. Если же изменение фазы используется в дополнение к любому другому способу разделения (например, частотному или временному), то возможное число каналов может быть удвоено.

Следует заметить в заключение, что осуществление синхронного детектирования и разделения наталкивается на ряд специфических технических трудностей, которые мы здесь обсуждать не будем.

§ 40. Краткие сведения об уравнении Матъё¹

Общее решение уравнения Матъё

$$x'' + (a + 2q \cos 2\tau) x = 0 \quad (1)$$

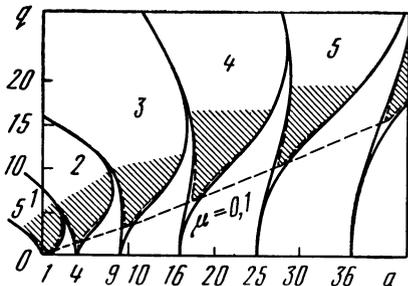
складывается из двух независимых решений и может быть записано в виде

$$x = x_1 + x_2 = A e^{\mu\tau} \Phi_1(\tau) + B e^{-\mu\tau} \Phi_2(\tau). \quad (2)$$

Здесь A и B — произвольные постоянные; Φ_1 и Φ_2 — периодические функции времени периода π или 2π . Эти периодические

¹ Этот параграф имеет чисто справочный характер. Здесь приведены в готовом виде (без выводов) только используемые в дальнейшем данные.

функции могут быть выражены соответствующими тригонометрическими рядами, но не эта сторона дела будет нас интересовать. С точки зрения наших задач наибольшее значение имеет вопрос об *устойчивости* решений уравнения Матъё. Легко видеть, что если μ — действительное число, не равное нулю, то одно из слагаемых решения (2) неограниченно растет с увеличением τ , а это и означает, что решение неустойчиво. Показатель μ зависит от коэффициентов a и q . Можно показать, что границы областей устойчивости определяются условием $\mu=0$. Графическое представление областей устойчивости на плоскости a, q дано на рис. 144.



Р и с. 144

Области неустойчивости (заштрихованные на рисунке) ограничены кривыми, сходящимися на оси a к точкам $a=1, 4, 9, 16$ и т. д. Картина симметрична относительно оси a .

К уравнению (1) сводится уравнение контура без потерь с периодически изменяемым параметром, например емкостью. Но уравнение реального контура имеет первоначально вид

$$y'' + 2xy' + (b + 2q \cos 2\tau) y = 0, \quad (3)$$

и задача состоит в том, чтобы свести и это уравнение к стандартной форме (1). Для этого возьмем

$$y = e^{-x\tau} x. \quad (4)$$

В результате такой подстановки мы и получим уравнение (1), в котором нужно положить

$$a = b - x^2. \quad (5)$$

Сопоставляя (2) и (4), мы видим, что решение уравнения (3) имеет вид

$$y = Ae^{(-x+\mu)\tau} \Phi_1(\tau) + Be^{(-x-\mu)\tau} \Phi_2(\tau). \quad (6)$$

Решение будет неустойчивым уже не при $|\mu| > 0$, а при

$$-x + |\mu| > 0, \quad (7)$$

т. е. при

$$|\mu| > x,$$

Следовательно, для уравнения (3) области неустойчивости будут ограничены уже не кривыми $\mu=0$, а кривыми

$$|\mu| = \kappa.$$

Эти кривые уже не доходят до оси a ; они спускаются к оси a языками, наинизшие точки которых отстоят от оси тем дальше, чем больше μ . На рис. 144 штриховыми линиями отмечены области неустойчивости для уравнения (3) при $\kappa=0, 1$, т. е. когда уравнение граничных линий имеет вид

$$\mu = \kappa = 0,1.$$

Система теряет устойчивость легче всего при наименьших значениях q , т. е. в условиях, соответствующих минимуму граничных кривых. Поэтому нам нужно найти координаты минимумов. Точное решение этой задачи очень громоздко. Но если полагать, что в уравнении (1) $a \gg q$ (это как раз тот случай, который нам встретится в дальнейшем), то можно воспользоваться приближенными соотношениями. Приближение состоит в том, что мы отбрасываем члены с высокими степенями q в степенных рядах, выражающих интересующие нас зависимости; мы полагаем, что точка минимума лежит посредине горизонтального отрезка, соединяющего две кривые $\mu=0$ в области неустойчивости. Тогда получаются следующие данные (m означает номер области неустойчивости, a_m и q_m — координаты минимума граничной кривой):

Таблица 2

m	1	2	3	4	5
a_m	$1 - \mu^2$	$4 + \frac{8}{3}\mu$	$9 + (6\mu)^{2/3}$	$16 + \frac{16}{5}(2\mu)^{1/2}$	$25 + \frac{4}{3}(45\mu)^{2/5}$
q_m	2μ	$4\mu^{1/2}$	$4(6\mu)^{1/3}$	$4(72\mu)^{1/4}$	$8(45\mu)^{1/5}$

Точки с указанными в табл. 2 координатами при $\mu=0,1$ отмечены на рис. 144. Они лежат на кривой, проходящей через начало и мало отличающейся от прямой.

Это значит, что при данном κ для нарушения устойчивости требуется тем большее изменение параметра, чем больше a . Приведенных сведений уже достаточно для рассмотрения параметрического возбуждения.

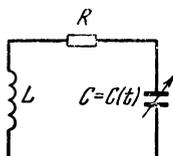
§ 41. Параметрическое возбуждение

В этом параграфе мы рассмотрим механизм колебаний, могущих возникнуть в контуре, если какой-либо его параметр (например, емкость или индуктивность) периодически изменяется.

Нужно сразу подчеркнуть, что мы имеем здесь дело с явлением совершенно особого рода. Обычно колебания в контуре возникают при действии той или иной электродвижущей силы. В рассматриваемом же случае никаких электродвижущих сил нет — схема опыта показана на рис. 145.

Поэтому сразу возникает вопрос о том, откуда берется энергия, необходимая для возбуждения и поддержания колебаний. Общие энергетические рассуждения и в данном случае позволяют раскрыть механизм явления.

Дело заключается в том, что для изменения емкости (мы будем в дальнейшем рассматривать именно этот частный случай)



Р и с. 145

необходимо, вообще говоря, совершить определенную работу против сил притяжения между обкладками. Представим себе, что колебания в контуре уже существуют и что на конденсаторе имеется переменное напряжение. Будем раздвигать обкладки в момент, когда напряжение максимально, и сближать их, когда напряжение равно нулю. Тогда за каждый цикл мы будем совершать определенную работу, т. е. вкладывать в систему некоторую энергию в механической форме. В системе происходит преобразование механической энергии в электрическую, и таким образом и осуществляется приток энергии в систему извне.

Для того чтобы пояснить механизм преобразования механической энергии в электрическую, напомним известные соотношения электростатики. Пусть изолированный конденсатор емкостью C заряжен до напряжения U между обкладками. Энергия, запасенная в электрическом поле конденсатора, равна $\mathcal{E} = U^2 C / 2$. При этом механическая сила, с которой обкладки притягиваются друг к другу, равна $F = \mathcal{E} / 2d$, где d — расстояние между обкладками. Если раздвинуть обкладки на Δd , то будет совершена механическая работа

$$\Delta A = F \Delta d = \frac{1}{2} \mathcal{E} \frac{\Delta d}{d}.$$

Соответствующее приращение должна получить и потенциальная энергия электрического поля. Перепишем выражение для энергии в виде

$$\mathcal{E} = Q_0^2 / 2C,$$

где $Q_0 = UC$ — начальный заряд, остающийся при описываемом опыте неизменным (так как конденсатор изолирован). Емкость

конденсатора равна $C = \epsilon S/d$, где ϵ — диэлектрическая проницаемость; S — площадь обкладок. Используя это соотношение, получим

$$\mathcal{E} = Q_0^2 d / 2\epsilon S.$$

Если изменится d , то энергия получит приращение

$$\Delta \mathcal{E} = \frac{Q_0^2}{2\epsilon S} \Delta d = \frac{Q_0^2}{2C} \cdot \frac{\Delta d}{d} = \frac{1}{2} \mathcal{E} \frac{\Delta d}{d},$$

равное, как видим, совершенной работе ΔA . Увеличение энергии конденсатора выразится в повышении напряжения на нем на величину

$$\Delta U = U \Delta d / d.$$

Таким образом, относительное изменение напряжения равно относительному изменению расстояния между обкладками¹.

После этих общефизических рассуждений вернемся к схеме рис. 145. Мы можем составить следующее уравнение напряжений:

$$LQ'' + RQ' + \frac{1}{C(t)}Q = 0. \quad (1)$$

Если емкость изменяется около среднего значения C_0 по закону

$$C(t) = C_0 + C_1 f(t) = C_0 \left[1 + \frac{C_1}{C_0} f(t) \right] = C_0 [1 + n f(t)],$$

то при $n = C_1/C_0 \ll 1$ можем переписать (1) в виде

$$Q'' + 2\alpha Q' + \omega_0^2 [1 - n f(t)] Q = 0, \quad (2)$$

где $\omega_0^2 = 1/LC_0$. Заменяем переменную по формуле

$$Q = y e^{-\alpha t}.$$

Это избавит нас от члена с первой производной

$$y'' + [\omega_1^2 - \omega_0^2 n f(t)] y = 0,$$

где $\omega_1^2 = \omega_0^2 - \alpha^2$ — квадрат частоты свободных колебаний контура с постоянными параметрами L , R и C_0 . Пусть теперь

$$f(t) = \cos \omega t.$$

Обозначая $\omega t = 2\tau$ и заменяя производную по t производной по τ , получим

$$y'' + 4 \left[\frac{\omega_1^2}{\omega^2} + \frac{\omega_0^2}{\omega^2} n \cos 2\tau \right] y = 0, \quad (3)$$

а это и есть уравнение Матрё. Теперь задача нахождения условий, при которых возможно возбуждение колебаний в контуре, сво-

¹ Не следует забывать, что все вышеприведенные соотношения получены из линеаризованных соотношений и представляют собой первое приближение, справедливое при $\Delta d/d \ll 1$. Такой же порядок малости имеют и $\Delta U/U$ и $\Delta \mathcal{E}/\mathcal{E}$.

дится к математической задаче определения соотношений, при которых уравнение Матё имеет неустойчивые решения. Решение этой задачи известно (см. § 40): неустойчивое решение получается при

$$x = 2\alpha/\omega < \mu. \quad (4)$$

Мы имеем следующие значения коэффициентов:

$$q = n2\omega_0^2/\omega^2, \quad a = 4\omega_1^2/\omega^2$$

или, если затухание невелико и можно пренебречь его влиянием,

$$a \cong 4\omega_0^2/\omega^2.$$

Таким образом, коэффициент q выражает относительное изменение параметра, а коэффициент a — отношение частот: частоты колебаний в контуре и частоты модуляции. Возбуждение легче всего возникает в точках минимумов граничных кривых. Для первой области неустойчивости мы имеем (см. § 40) $a=1-\mu^2$, $q=2\mu$. Так как $\mu \ll 1$, то можно положить $a=1$, откуда

$$\omega_0 = \omega/2,$$

т. е. колебания легче всего возбуждятся на частоте, равной половине частоты модуляции. Подставляя значение $\mu=q/2$ в условие возбуждения (4), получим

$$\frac{2\alpha}{\omega} < \frac{1}{2}q = n \frac{\omega_0^2}{\omega^2}.$$

Преобразовывая это неравенство, получим

$$\frac{2\alpha}{\omega} = d\omega_0/\omega < n\omega_0^2/\omega^2.$$

Но так как у нас $\omega_0/\omega=1/2$, то окончательно условие возбуждения запишется в виде

$$n > 2d. \quad (5)$$

Итак, при данном затухании контура параметрические колебания в нем возникнут при периодическом изменении емкости при условии достаточно большого относительного изменения емкости. Чем меньше затухание, тем меньше необходимое относительное изменение

$$n = C_1/C_0.$$

Мы вывели условие возбуждения для первой области неустойчивости при $\omega_0 \cong \omega/2$. Возбуждение возможно и для других соотношений частот, т. е. для других областей неустойчивости.

Однако при других отношениях частот труднее возбудить колебание, т. е. требуется большая глубина модуляции. Для вто-

рой области неустойчивости имеем координаты минимума граничной кривой

$$q = 4\mu^{1/2}, \quad a = 4 + \frac{8}{3}\mu.$$

Положим приближенно $a = 4$, т. е. $\omega_0 = \omega$. Подставляя в (4)

$$\mu = \frac{1}{16}q^2 = \frac{1}{4}n^2 \frac{\omega_0^4}{\omega^4} = \frac{1}{4}n^2,$$

получим

$$d < \frac{1}{4}n^2,$$

т. е.

$$n > \sqrt{2d}.$$

Для третьей области неустойчивости координаты минимума

$$q = 4(6\mu)^{1/2}, \quad a = 9 + (6\mu)^{2/3} \cong 9,$$

откуда $4\omega_0^2/\omega^2 = 9$, $\omega_0 = 3\omega/2$, $\mu = q^2/384$.

Подставляя в (4), получим

$$n = \frac{8}{3\sqrt[3]{3}}\sqrt[3]{d} \cong 1,84\sqrt[3]{d}$$

и так далее.

Простейший результат, выражаемый формулой (5), можно было бы получить, опираясь не на теорию уравнения Матъё, а на энергетические соображения, основа которых у нас уже подготовлена. Мы сделаем сейчас энергетический вывод условия возбуждения, так как этот вывод помогает уяснить физическую картину явлений при параметрическом возбуждении колебаний.

Полная энергия \mathcal{E} в колебательном контуре равна сумме энергий, запасенных в магнитном поле индуктивности, (\mathcal{E}_L) и в электрическом поле емкости (\mathcal{E}_C). При отсутствии потерь и при отсутствии внешнего воздействия запас энергии остается неизменным; энергия лишь периодически перекачивается из индуктивности в емкость и обратно, в чем, в сущности, и состоит процесс свободных колебаний. При наличии потерь, т. е. с учетом активного сопротивления R , запас энергии убывает. Потеря энергии \mathcal{E}_- , отнесенная ко времени и усредненная за период, дает теряемую мощность, равную

$$P_- = \left(\frac{d\mathcal{E}_-}{dt} \right) = \frac{1}{2} I_m^2 R,$$

где черта означает усреднение за период, а $I_m = \omega_0 Q_m = \omega_0 C U_m$ — амплитуда тока, которую можно также выразить через амплитуду заряда Q_m и через амплитуду напряжения на конденсаторе U_m .

Таким образом,

$$P_- = \frac{1}{2} \omega_0^2 C^2 R U_m^2 = \omega_0^2 R C \mathcal{E} = \omega_0 d \mathcal{E}, \quad (6)$$

где d — затухание контура. Колебания могут быть поддержаны, если теряемая мощность восполняется извне. Это и достигается периодическим изменением емкости. Мы имеем

$$\mathcal{E} = \mathcal{E}_L + \mathcal{E}_c.$$

Найдем приращение энергии, обусловленное только изменением емкости

$$d_c \mathcal{E} = \frac{\partial \mathcal{E}}{\partial C} dC,$$

так как \mathcal{E}_L не зависит от C ,

$$\mathcal{E}_c = U^2 C / 2,$$

так что

$$d_c \mathcal{E} = \frac{1}{2} U^2 dC.$$

Разделив обе части на dt , получим

$$\frac{d_c \mathcal{E}}{dt} = \frac{1}{2} U^2 \frac{dC}{dt}. \quad (7)$$

Мы предполагаем, что у нас имеется синусоидальное колебание, так что

$$U = U_m \cos \omega_0 t.$$

Подставляя это значение в (7), найдем

$$\frac{d_c \mathcal{E}}{dt} = \frac{1}{4} U_m^2 (1 + \cos 2\omega_0 t) \frac{dC}{dt}. \quad (8)$$

Теперь заметим, что среднее значение $d_c \mathcal{E} / dt$ есть не что иное, как мощность P_+ , поступающая в контур. Для того чтобы эта мощность не была равна нулю, необходимо в формуле

$$C(t) = C_0 + C_1 f(t)$$

положить

$$f(t) = \sin 2\omega_0 t,$$

тогда

$$dC/dt = 2\omega_0 C_1 \cos 2\omega_0 t. \quad (9)$$

Действительно, подставив (9) в (8), получим

$$\begin{aligned} \frac{d_c \mathcal{E}}{dt} &= \frac{1}{2} U_m^2 \omega_0 C_1 (\cos 2\omega_0 t + \cos^2 2\omega_0 t), \\ P_+ &= \left(\overline{\frac{d_c \mathcal{E}}{dt}} \right) = \frac{1}{4} U_m^2 \omega_0 C_1 = \frac{1}{4} U_m^2 \omega_0 C_0 \frac{C_1}{C_0} = \frac{1}{2} n \omega_0 \mathcal{E}. \end{aligned} \quad (10)$$

Для того чтобы колебания не затухали, необходимо, очевидно, выполнение неравенства

$$P_+ \geq P_-.$$

Отсюда с помощью формул (6) и (10) находим

$$n \geq 2d,$$

что совпадает с (5).

В заключение заметим, что мы рассматривали задачу о параметрическом возбуждении колебаний как линейную задачу; мы нашли приближенное условие самовозбуждения при $\omega_0/\omega = 1/2$, но ничего не можем пока сказать об установившейся амплитуде колебаний. Установившийся режим может быть определен только при учете нелинейности. Так как в схеме рис. 145 нелинейных элементов нет, то амплитуда колебаний теоретически может нарастать неограниченно. Практически в такого рода схеме напряжение быстро достигает таких больших значений, что пробиваются конденсаторы. Поэтому в схему приходится вводить нелинейные сопротивления специально для ограничения амплитуды колебаний.

§ 42. Деление частоты

В ряде случаев необходимо, располагая генератором известной частоты, получить колебание более низкой частоты, в частности, в n раз более низкой. Операцию получения частоты в n раз ниже исходной называют делением частоты. Чаще всего встречается деление на два, т. е. получение половинной частоты.

Деление частоты, т. е. возникновение дробных частот — так называемых субгармоник — возможно при нелинейном резонансе и в некоторых других нелинейных схемах. В этом параграфе мы рассмотрим устройство и действие параметрического делителя частоты.

Составим схему из источника синусоидального напряжения

$$E = E_m \cos 2\omega t$$

и переменной проводимости

$$Y = Y_0 (1 + m \sin \omega t).$$

Сила тока в такой цепи

$$I = EY = E_m Y_0 \left(\cos 2\omega t + \frac{m}{2} \sin 3\omega t - \frac{m}{2} \sin \omega t \right).$$

Как видно, в составе тока появилась компонента с частотой, равной половине частоты источника. Кроме того, имеются составляющие с частотой источника 2ω и с полуторной частотой 3ω . Если отфильтровать эти составляющие и использовать колебание с частотой ω для *управления* проводимостью, то получим принципиальную схему делителя, показанную на рис. 146. Таким

образом, в схеме делителя должна иметься обратная связь, при помощи которой производится изменение параметра (в частности, проводимости) с дробной частотой.

В качестве управляемой проводимости можно использовать крутизну триода, зависящую от напряжения на сетке. Если характеристика лампы задана в виде

$$I_a = f(U),$$

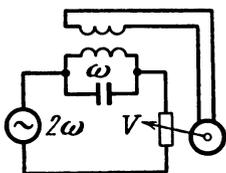
то при наличии квадратичного члена в выражении функции f мы получим линейную зависимость крутизны от напряжения на сетке. Пусть

$$f(U) = a_0 + a_1 U + a_2 U^2.$$

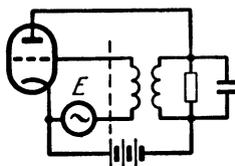
Тогда

$$S = \frac{dI_a}{dU} = a_1 + 2a_2 U.$$

Простейшая схема с использованием этой возможности изображена на рис. 147. Эта схема отличается от схемы обычного



Р и с. 146



Р и с. 147

генератора с контуром в цепи анода лишь наличием источника переменного напряжения E в цепи сетки. Схема рис. 147 в принципе не отличается от схемы регенеративного приемника и вообще от схемы потенциально-автоколебательной системы, находящейся под воздействием внешнего источника. Источник имеет частоту 2ω . Контур настроен на половинную частоту ω . Напряжение на сетке равно сумме напряжения E и напряжения, наводимого в катушке связи. Для того чтобы напряжение половинной частоты действительно могло вырабатываться, необходимо выполнение известных условий, которые мы сейчас в простейшей форме выведем.

Схема рис. 147 может рассматриваться как своего рода генератор половинной частоты. Для выяснения условий возбуждения этого генератора воспользуемся условием баланса фаз и амплитуд. Разорвем для этого схему рис. 147 по штриховой линии и представим ее, как показано на рис. 148. Положим, что на вход подано напряжение

$$U_1 = U_m \sin \omega t.$$

На сетке будем иметь

$$U_g = U_1 + E = U_m \sin \omega t + E_m \cos 2\omega t.$$

Ток в анодной цепи будет

$$I_a = f(U_g) = a_0 + a_1(U_1 + E) + a_2(U_1 + E)^2.$$

Если контур настроен на частоту ω , то составляющие напряжения с другими частотами будут практически отсутствовать, а для напряжения с частотой ω будем иметь

$$U = I_a R = R(a_1 + E_m a_2) U_1.$$

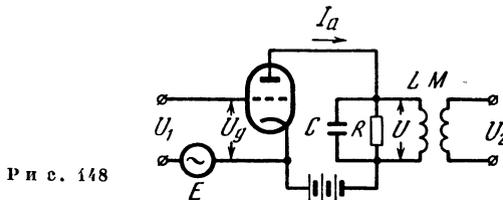
Выходное напряжение равно

$$U_2 = \frac{M}{L} U = kU,$$

где $k = M/L$ — коэффициент связи. Таким образом,

$$U_2 = kR(a_1 + E_m a_2) U_1.$$

Для того чтобы колебания частоты ω могли поддерживаться в *замкнутой* цепи (т. е. схеме рис. 147, к которой мы вернемся,



соединив накоротко входные и выходные зажимы схемы рис. 148), необходимо выполнение условия

$$U_2 \geq U_1,$$

откуда

$$kR(a_1 + E_m a_2) \geq 1. \quad (1)$$

Это и есть условие возбуждения схемы рис. 147, работающей в качестве делителя. Но по поводу этого условия нужно сразу сделать одно существенное замечание. Дело в том, что, как видно из формулы (1), схема может возбуждаться и при $E=0$, т. е. при условии

$$kR a_1 \geq 1. \quad (2)$$

Смысл этого соотношения станет совершенно ясным, если взглянуть на схему рис. 147: если устранить источник напряжения E , то мы будем иметь схему самого обыкновенного генератора, способного генерировать колебания с частотой ω (на которую настроен контур). Формула (2) не что иное, как условие самовозбуждения для такого генератора. Следовательно, для того чтобы схема рис. 147 возбуждалась в качестве делителя и не

возбуждалась в качестве генератора, необходимо одновременное соблюдение двух условий ¹

$$\begin{aligned} kR(a_1 + E_m a_2) &\geq 1, \\ kRa_1 &< 1. \end{aligned} \quad (3)$$

Нужно сделать еще несколько замечаний. Во-первых, условие баланса фаз в нашем анализе выполнено автоматически, так как, с одной стороны, мы выбрали $U_1 = U_m \sin \omega t$, $E = E_m \cos 2 \omega t$ (т. е. заранее взяли надлежащие начальные фазы для обоих напряжений), а с другой стороны, предположили, что контур настроен точно в резонанс. Более подробное исследование показывает, что деление возможно в пределах некоторой расстройки. Во-вторых, при нашей постановке задачи мы в состоянии только установить условие возбуждения, но не можем найти установившейся амплитуды. Дело в том, что, учитывая только квадратичный член характеристики (необходимый для получения деления), мы сводим параметрическую задачу к линейной. Установившийся режим возможен, как известно, только в нелинейной системе. Для исследования установившегося режима необходимо учесть по меньшей мере член 3-й степени в выражении для характеристики триода. Такого рода более детальное исследование составляет теорию так называемого «резонанса 2-го рода».

В заключение следует еще заметить, что одновременное выполнение обоих условий (3) практически стеснительно. Возможность самовозбуждения схемы рис. 147 как обычного генератора следует считать ее недостатком. От этого недостатка свободны так называемые регенеративные делители частоты, которым посвящен следующий параграф.

§ 43. Регенеративное деление частоты

Характерной чертой параметрического возбуждения является то, что могут возникать колебания с частотой, не равной частоте изменения параметра. В частности, как мы видели, легче всего возбуждаются колебания с частотой

$$\omega_0 = \omega/2,$$

где ω — частота изменения параметра. Это обстоятельство может быть использовано для деления частоты.

Рассмотрим скелетную схему рис. 149. На вход подается напряжение U_1 . Оно поступает на множитель Π , к которому подводится также напряжение U_4 . Действие множителя определяется тем, что выходное напряжение U_2 пропорционально произведению двух входных напряжений U_1 и U_4 . Напряжение U_2

¹ Возможно деление и при возбужденном генераторе. В этом случае механизм явления иной: речь идет уже о захватывании на субгармонике. Этот случай мы разбирать не будем.

фильтруется фильтром Φ , усиливается усилителем $У$ и поступает на выход. Выходное напряжение U_3 через фазовращатель Φ в. подается обратно на перемножитель.

Покажем, что в такой схеме возможно деление частоты в отношении 1 : 2. Пусть на вход подано напряжение

$$U_1 = \sin 2\omega t.$$

Предположим, что деление частоты происходит, так что

$$U_4 = \sin \omega t.$$

Тогда

$$U_2 = U_1 U_4 = \sin \omega t \sin 2\omega t = \frac{1}{2} \cos \omega t - \frac{1}{2} \cos 3\omega t.$$

Фильтр погасит составляющую тройной частоты, так что на выходе фильтра получим напряжение $\frac{1}{2} \cos \omega t$. Пусть усилитель дает усиление по напряжению вдвое. Тогда

$$U_3 = \cos \omega t,$$

т. е. на выходе имеем напряжение половинной частоты по отношению к частоте на входе. Если теперь фазовращатель поворачивает фазу напряжения на 90° , то получим

$$U_4 = \sin \omega t$$

в соответствии с нашим предположением. Схемы типа рис. 149 называются регенеративными делителями частоты. Этот термин указывает на обязательное наличие обратной связи с выхода схемы на ее вход.

Нужно заметить, что приведенное рассуждение показывает лишь, что деление частоты при помощи регенеративного делителя возможно. Но это рассуждение не позволяет найти условия, при которых деление действительно происходит. Оно не позволяет также обнаружить существенные черты явления, одна из которых

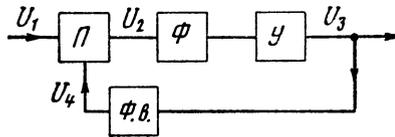


Рис. 149

состоит в том, что деление начинается только при вполне определенном значении входного напряжения U_1 . Для того чтобы разобраться в этих важных подробностях, рассмотрим какую-либо конкретную схему, например показанную на рис. 150. В этой схеме в качестве перемножителя применен кольцевой модулятор (см. § 12); мы будем полагать, что он производит чистое перемножение поданных на него напряжений. Составим уравнение токов для контура в анодной цепи триода

$$C_0 \frac{dU}{dt} + \frac{1}{R} U + \frac{1}{L} \int U dt = S U_s, \quad (1)$$

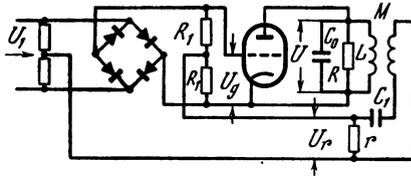
где S — крутизна. Э. д. с., наводимая в катушке обратной связи, равна

$$e = M \frac{dI_L}{dt} = \frac{M}{L} U. \quad (2)$$

Эта э. д. с. действует в цепи $C_1 r$; напряжение на r равно

$$U_r \cong r C_1 \frac{de}{dt} \quad (3)$$

(при условии, что напряжение на емкости много больше U_r). Комбинация $r C_1$ выполняет роль фазовращателя. Напряжение U ,



Р и с. 150

подается на перемножитель; это и есть то напряжение, которое в предыдущем было обозначено через U_4 . Далее, напряжение, снимаемое с кольцевого модулятора, равно (см. § 12)

$$U_2 = U_g = 4R_1 a_2 U_1 U_4. \quad (4)$$

Из (2), (3) и (4) получаем

$$U_g = 4a_2 r R_1 C_1 \frac{M}{L} U_1 \frac{dU}{dt} = k U_1 \frac{dU}{dt}. \quad (5)$$

Подставляя это значение в уравнение (1) и объединяя члены, содержащие dU/dt , получим уравнение с переменным коэффициентом

$$(C_0 - k S U_1) \frac{dU}{dt} + \frac{1}{R} U + \frac{1}{L} \int U dt = 0. \quad (6)$$

Это — уравнение токов в контуре, в котором емкость изменяется в соответствии с входным напряжением. Если входное напряжение изменяется по закону

$$U_1 = U_m \sin 2\omega t,$$

то для действующей емкости имеем

$$C(t) = C_0 - k S U_m \sin 2\omega t = C_0 \left(1 - \frac{k S}{C_0} U_m \sin 2\omega t \right).$$

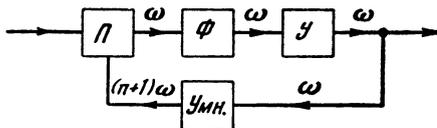
Теперь становится ясным, что механизм регенеративного деления частоты сводится к параметрическому возбуждению колебаний. Условие возбуждения $n > 2d$ в нашем случае выглядит так

$$\frac{k S}{C_0} U_m > 2d$$

или

$$U_m > \frac{1}{2} \cdot \frac{d}{a_2 \omega_0^2 r R_1 C_1 M S}. \quad (7)$$

Напомним смысл входящих в эту формулу величин: ω_0 — резонансная частота контура, равная $(LC_0)^{-1/2}$; d — затухание контура; a_2 — коэффициент при U^2 в выражении для характеристики нелинейного элемента кольцевого модулятора; R_1 — сопротивление нагрузки модулятора (следует напомнить, что при выводе формулы (4) это сопротивление предполагалось малым по сравнению с внутренним сопротивлением самого модулятора); r и C_1 — сопротивление и емкость в цепи обратной связи (напомним, что должно быть $rC_1 \ll 1/\omega_0$); M — коэффициент взаимной индукции.



Р и с. 151

Формула (7) позволяет найти минимальную амплитуду входного напряжения, при которой схема делителя приходит в действие.

В предыдущем анализе мы стремились подобрать наиболее благоприятную фазу обратной связи, для чего в цепь обратной связи был введен соответствующий фазовращатель. Это необязательно; делитель будет работать и при других фазовых соотношениях, хотя для возбуждения потребуются соответственно увеличенное входное напряжение. В практических схемах специальный фазовращатель обычно не ставится.

Возможны многочисленные варианты схем регенеративных делителей. Укажем здесь на существенное видоизменение схемы рис. 149, приводящее к делителю, способному выполнять деление частоты в любое число раз. Для этого в цепь обратной связи вводится умножитель (УМН) частоты в $n+1$ раз (рис. 151). При работе делителя на выходе получается частота ω , если на вход подается частота $n\omega$. На выходе умножителя в цепи обратной связи получается частота $(n+1)\omega$. Эта частота попадает на вход перемножителя. При перемножении колебаний с частотами $n\omega$ и $(n+1)\omega$ получаются: суммарная частота $(2n+1)\omega$ и нужная нам разностная частота $(n+1)\omega - n\omega = \omega$, которая выделяется фильтром и поступает на выход.

Возможно, конечно, получить деление высокой кратности и другим путем, а именно, включая делители друг за другом ступенями. При таком использовании делителей, делящих, например, частоту в отношении 1 : 2, можно получить деление в отношении 1 : 2^m , где m — число ступеней.

ПЕРЕДАЧА СИГНАЛОВ МОДУЛИРОВАННЫМ ШУМОМ

1. Общий способ образования сигнала связи состоит в том, что выбирается некоторый переносчик, т. е. процесс, характеризующийся некоторым числом постоянных параметров, а затем производится модуляция, т. е. изменение одного из параметров переносчика в соответствии с передаваемым сообщением.

В качестве переносчиков применяются: постоянный ток, синусоидальное колебание и периодическая последовательность импульсов. Синусоидальное колебание определяется тремя параметрами: амплитудой, частотой и фазой; при модуляции каждого из этих параметров получится три вида модуляции: АМ, ЧМ и ФМ. Периодическая последовательность прямоугольных импульсов определяется высотой импульсов, их длительностью, частотой следования и положением импульсов относительно таковых точек. Соответствующие виды модуляции известны как АИМ, ДИМ, ЧИМ и ФИМ. Для радиоимпульсов добавляются еще некоторые виды модуляции, связанные с изменением параметров высокочастотного заполнения.

2. Возможности построения сигнала связи отнюдь не исчерпываются названными. В качестве переносчика может применяться любой длительный процесс, определяемый теми или иными параметрами, могущими изменяться в результате модуляции. Мы рассмотрим в общих чертах одну из возможностей, а именно применение в качестве переносчика шума, представляющего собой стационарный случайный процесс.

В качестве постоянных параметров случайного процесса $\epsilon(t)$ могут применяться моменты его распределения. Первый момент есть просто среднее значение, которое мы будем считать равным нулю. Интересующие нас параметры мы получим из рассмотрения второго момента.

В общем виде второй момент — это функция автокорреляции $B(\tau)$ процесса $\epsilon(t)$. В частности, при $\tau=0$ $B(0)=P_0$. По физическому смыслу эта величина есть не что иное, как мощность процесса $\xi(t)$. Для стационарного процесса P_0 есть постоянная величина. Ясно, что этот параметр можно модулировать и притом теми же техническими средствами, которые применяются при обычной АМ или АИМ. Итак, модулируя шум по мощности, мы можем получить сигнал, способ образования которого по аналогии

с существующей терминологией назовем амплитудно-шумовой модуляцией (АШУМ).

Можно получить иные виды модуляции, модулируя те или иные параметры, определяющие вид функции корреляции $B(\tau)$. Но нам удобнее будет перейти к спектру мощности, связанному с функцией корреляции известным соотношением

$$G(\omega) = \frac{2}{\pi} \int_0^{\infty} B(\tau) \cos \omega \tau d\tau. \quad (1)$$

Пусть, в частности, переносчик выбран в виде белого шума, из которого выделена полоса от ω_1 до ω_2 , т. е. пусть

$$G(\omega) = \begin{cases} G_0 & \omega_1 < \omega < \omega_2 \\ 0 & \text{вне указанной полосы.} \end{cases} \quad (2)$$

В этом случае параметрами переносчика являются граничные частоты ω_1 и ω_2 , и модуляция может состоять в изменении каждого из этих параметров. Если же менять одновременно ω_1 и ω_2 , но так, чтобы разность $\omega_2 - \omega_1 = \Delta \omega$ оставалась постоянной, то модуляция будет характеризоваться тем, что спектр шума, не меняя своей формы, будет перемещаться по шкале частот, следуя за изменениями передаваемого сообщения. Такой вид модуляции можно назвать частотно-шумовой модуляцией (ЧаШУМ). Аналога фазовой модуляции для шума не существует, но можно построить сколько угодно видов шумовой модуляции, используя спектры более сложной формы или моменты высших порядков.

3. Ограничиваясь рассмотрением простейшей — амплитудно-шумовой — модуляции, отметим, что при ее применении получается сигнал с новыми средствами, а для его образования требуются необычные технические средства. В частности, существенно меняется способ генерирования переносчика. Генератор ограниченного по полосе белого шума представляется в виде шумового генератора (дающего в принципе шум с однородным, вплоть до частот порядка 10^{12} гц, спектром), соответствующего полосового фильтра и полосового усилителя. Такая схема (в особенности на свч) существенно отличается от обычной схемы для генерирования синусоидального колебания.

Один из первых вопросов, которые здесь возникают, это вопрос о спектре амплитудно-модулированного шума. Запишем напряжение для модулированного шума в виде

$$\zeta(t) = [1 + mx(t)] \xi(t), \quad (3)$$

где $\xi(t)$ — переносчик, $x(t)$ — модулирующая функция, также представляющая собой случайный процесс. Определим спектр

мощности процесса $\zeta(t)$. Для этого найдем сначала функцию корреляции

$$\begin{aligned} B_{\zeta}(\tau) &= M\zeta(t)\zeta(t+\tau) = \\ &= M[1+mx(t)]\xi(t)[1+mx(t+\tau)]\xi(t+\tau) = \\ &= M[1+mx(t)+mx(t+\tau)+m^2x(t)x(t+\tau)]\xi(t)\xi(t+\tau) = \\ &= [1+m^2B_x(\tau)]B_{\xi}(\tau) \end{aligned}$$

(на том основании, что $M_x=0$ и $x(t)$ и $\xi(t)$ независимы).

Теперь находим спектр

$$\begin{aligned} G_{\zeta}(\omega) &= \frac{1}{\pi} \int_0^{\infty} B_{\zeta}(\tau) \cos \omega\tau d\tau = \frac{2}{\pi} \int_0^{\infty} B_{\xi}(\tau) \cos \omega\tau d\tau + \\ &+ \frac{2}{\pi} m^2 \int_0^{\infty} B_x(\tau) B_{\xi}(\tau) \cos \omega\tau d\tau. \end{aligned}$$

Первый член — это спектр переносчика. Для интеграла во втором члене на основании теоремы о спектре произведения имеем

$$\int_0^{\infty} B_x(\tau) B_{\xi}(\tau) \cos \omega\tau d\tau = \frac{1}{2} \int_{-\infty}^{\infty} G_{\xi}(\nu) G_x(\omega - \nu) d\nu. \quad (4)$$

Итак, окончательно, спектр модулированного шума (3) равен

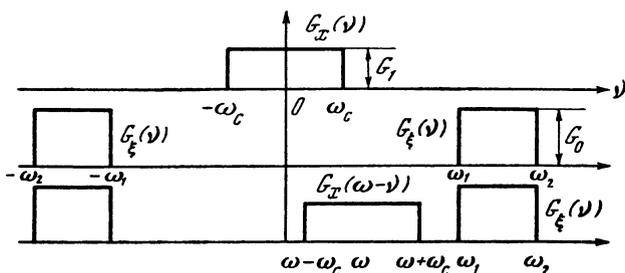
$$G_{\zeta}(\omega) = G_{\xi}(\omega) + \frac{m^2}{\pi} \int_{-\infty}^{\infty} G_{\xi}(\nu) G_x(\omega - \nu) d\nu. \quad (5)$$

Известно, что ширина спектра произведения функций равна сумме ширин спектров сомножителей. Рассмотрим в качестве примера случай, когда $G_{\xi}(\omega)$ определяется соотношениями (2), а спектр модулирующего процесса $x(t)$ задан в виде

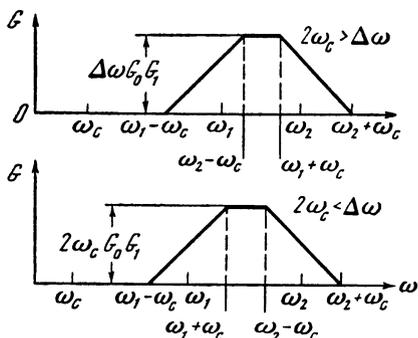
$$G_x(\omega) = \begin{cases} G_1 & 0 < \omega < \omega_c, \\ 0 & \omega_c < \omega < \infty. \end{cases} \quad (6)$$

Спектр произведения не равен нулю для тех значений ω , для которых спектры $G_{\xi}(\nu)$ и $G_x(\omega - \nu)$ перекрываются. При этом нужно помнить, что спектр мощности $G(\omega)$ — четная функция частоты. На рис. 1 изображены сомножители подынтегрального выражения в (5). Выполняя интегрирование, находим для спектра, выражаемого вторым членом (5), соотношения, представленные графически на рис. 2.

4. Среди особенностей сигнала АШУМ заслуживают внимания его свойства по отношению к замиранию. Имеется ввиду общее замирание интерференционного происхождения, обусловленное многолучевым распространением.



Р и с. 1



Р и с. 2

Будем рассматривать для простоты схематизированный случай двух лучей с временным сдвигом τ ; интенсивность обоих лучей будем считать одинаковой. При сложении двух синусоидальных колебаний мощностью P_0 каждое будем иметь для мощности результирующего колебания

$$P = 2P_0(1 + \cos \omega_0 \tau), \quad (7)$$

т. е. при $\omega_0 \tau = (2n+1)\pi$ результирующая мощность обращается в нуль. При сложении же двух шумовых переносчиков, сдвинутых во времени на τ , результирующая мощность будет

$$P = M[\xi(t) + \xi(t + \tau)]^2 = M\xi^2(t) + M\xi^2(t + \tau) + 2M\xi(t)\xi(t + \tau)$$

или

$$P = 2P_0[1 + B(\tau)/B(0)] = 2P_0[1 + b(\tau)], \quad (8)$$

где $b(\tau)$ — нормированная функция автокорреляции. Формула (8) показывает, что чем меньше корреляция, тем меньше шумовой сигнал подвержен замиранию. Корреляция же тем меньше, чем шире спектр.

Рассмотрим пример. Пусть переносчик задан спектром (2). Для функции корреляции имеем

$$B(\tau) = G_0 \int_{\omega_1}^{\omega_2} \cos \omega \tau d\omega = \frac{G_0}{\tau} (\sin \omega_2 \tau - \sin \omega_1 \tau) = \frac{2G_0}{\tau} \sin \frac{1}{2} \Delta\omega \tau \cos \omega_0 \tau,$$

где $\omega_0 = (\omega_1 + \omega_2)/2$ — центральная частота спектра. Используя (8), находим

$$\eta = \frac{P}{2P_0} = 1 + \frac{\sin \frac{1}{2} \Delta\omega\tau}{\frac{1}{2} \Delta\omega\tau} \cos \omega_0\tau.$$

Наименьшее значение (не равное, однако, нулю, как в случае синусоидального переносчика) будем иметь при $\cos \omega_0\tau = -1$, так что

$$\eta_{\min} = 1 - \frac{\sin \frac{1}{2} \Delta\omega\tau}{\frac{1}{2} \Delta\omega\tau}.$$

Это соотношение можно еще упростить на том основании, что синус не превосходит единицы по абсолютной величине. Поэтому

$$\eta_{\min} \leq 1 - 2/\Delta\omega\tau. \quad (9)$$

Эта формула показывает, что замирание может быть сделано неощутимым, если $\Delta\omega\tau$ достаточно велико. Так, например, допуская $\eta_{\min} \geq 0,8$, найдем $\Delta\omega\tau \geq 2/(1 - \eta_{\min}) = 10$, откуда

$$\Delta f = \frac{1}{2\pi} \Delta\omega = \frac{1}{\pi\tau(1 - \eta_{\min})} = \frac{1,6}{\tau}.$$

Статистика показывает [1], что значения $\tau < 0,2$ мсек (на линиях длиной от 1500 км и выше) практически не наблюдаются. Подставляя это наименьшее значение, находим

$$\Delta f \geq \frac{1,6 \cdot 10^3}{0,2} = 8 \cdot 10^3 \text{ гц.}$$

Итак, при переносчике в виде полосы белого шума шириной 8 кГц мощность суммарного сигнала при двухлучевом приеме будет убывать в результате замирания не более чем на 20% (в среднеквадратичное значение напряжения не больше чем на 10%).

Конечно, расширение полосы усиливает влияние избирательного замирания; все же описанный метод борьбы с общим замиранием может в известных условиях оказаться более выгодным, чем, например, прием на разнесенные антенны.

5. Следует пояснить принципиальное различие между обычными видами модуляции и модуляцией статистических параметров. Дело в том, что, например, в случае синусоидального переносчика амплитуда, частота и фаза — это действительно постоянные параметры, т. е. величины, мгновенные значения которых (в отсутствие модуляции) неизменны. Что же касается мощности случайного процесса, то это средняя величина, определяемая (путем усреднения по времени) как

$$P_0 = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi^2(t) dt.$$

Мгновенная мощность, т. е. $\xi(t)$ — случайная величина. Предположим для определенности, что речь идет о телеграфном приемнике и что прием осуществляется при помощи квадратичного детектора с последующим интегрированием, так что на выходе измеряется энергия

$$E_T = \int_0^T \xi^2(t) dt.$$

Это также случайная величина. Обозначим ее среднее значение через E_0 , а дисперсию через σ^2 . В таком случае относительное среднеквадратичное отклонение величины E_T от среднего значения E_0 может быть выражено отношением σ/E_0 . По существу дела, флуктуации E_T можно рассматривать как помеху, наложенную на «истинное» значение E_0 . Вопрос о величине флуктуаций энергии исследован Райсом [2]. Для рассмотренного нами случая спектра вида (2) относительная величина флуктуаций убывает с ростом произведения $\Delta f T$. При больших значениях этого произведения может применяться асимптотическое выражение

$$\sigma/E_0 \sim (\Delta f T)^{-1/2}.$$

Для телеграфного приема время интегрирования T — это, очевидно, длительность элементарной посылки. Длительность эта для различных телеграфных систем составляет примерно от 5 до 25 мсек. Принимая $T=10$ мсек и $\Delta f=10$ кГц, найдем

$$\sigma/E_0 \sim (10 \cdot 10^3 \cdot 10 \cdot 10^{-3})^{-1/2} = 0,1.$$

Мы не будем заниматься вычислением вероятности ошибки, так как для этого помимо распределения вероятностей величины E_T нужно еще знать характеристики конкретной аппаратуры, которая может содержать пороговые ограничители, регенераторы и т. п.

Л и т е р а т у р а

1. *Е. Л. Черенкова*. Искажения телеграфных сигналов при передаче на коротких волнах. Связьиздат, 1955.
2. *S. O. Rice*. Filtered thermal noise fluctuation of energy as a function of interval length. JASA, 1943, 14, 4.

О ВЫЧИСЛЕНИИ СПЕКТРОВ СЛУЧАЙНЫХ ПРОЦЕССОВ

В технической литературе последнего времени неоднократно излагались вопросы, относящиеся к спектральному представлению случайных процессов. Однако в этом изложении встречается ряд неясностей и неточностей, некоторые из которых стали традиционными. Между тем практическая важность понятия о спектре случайного процесса настолько велика, что приведение в некоторый порядок методических вопросов, связанных с этим понятием, представляется весьма желательным.

1. Основное определение спектра стационарного процесса вытекает из теоремы Хинчина и гласит

$$G(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} B(\tau) e^{-i\omega\tau} d\tau \quad (1)$$

(множитель $1/\pi$ необязателен; постоянные множители во всех последующих соотношениях произвольны, но подчинены выбранной системе обозначений). Функцию $G(\omega)$ называют в различных источниках статистическим спектром или энергетическим спектром, или спектром мощности случайного процесса.

2. Функция корреляции $B(\tau)$, входящая в определение (1), в свою очередь определяется соотношением

$$B(\tau) = M(\xi\xi_{\tau}), \quad (2)$$

где $\xi = \xi(t)$, $\xi_{\tau} = \xi(t-\tau)$ — значения процесса $\xi(t)$ для двух моментов, отстоящих друг от друга на τ . Для стационарного процесса по определению $B(\tau)$ не зависит от t . Поэтому не зависит от t и $G(\omega)$.

3. В силу эргодичности стационарного процесса среднее по множеству с вероятностью единица равно среднему по времени для некоторой реализации случайного процесса. Поэтому вместо (2) можно записать

$$B(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) \xi(t-\tau) dt, \quad (3)$$

а для спектра может быть выведено равноценное выражение

$$G(\omega) = \frac{1}{\pi} \lim_{T \rightarrow \infty} \frac{1}{T} |S_T(\omega)|^2, \quad (4)$$

где

$$S_T(\omega) = \int_0^T \xi(t) e^{-i\omega t} dt \quad (5)$$

означает текущий спектр случайного процесса $\xi(t)$.

Определение (3) приводится очень часто в качестве основного и единственного, что неправомерно, так как основным определением является (2). По поводу соотношений (3) и (4) следует сделать два существенных замечания, определяющих их роль и значение.

Первое. Соотношение (3) и (4) имеют силу только для стационарных (хотя бы в широком смысле) процессов. Между тем ряд авторов без всяких оговорок применяют их к процессам, заведомо нестационарным (примеры будут приведены ниже).

Второе. Соотношения (3) и (4) могут служить основой для техники (приближенного) экспериментального определения $G(\omega)$ и $B(\tau)$. Но эти соотношения не могут быть применены для вычисления $G(\omega)$ и $B(\tau)$ по той простой причине, что нам никогда не задается аналитически реализация случайного процесса $\xi(t)$. При аналитическом рассмотрении случайный процесс всегда задается теми или иными распределениями. Поэтому прямое вычисление на основе (3) и (4) невозможно, и если пересмотреть вычисления, приводимые в различных местах, то оказывается, что, исходя из определений (3) и (4), авторы фактически на том или ином этапе прибегают к неизбежному усреднению по множеству.

4. Для вычисления функции корреляции $B(\tau)$ на основе ее определения (2), которое может быть переписано в виде

$$B(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \varphi(x_1, x_2, \tau) dx_1 dx_2, \quad (6)$$

необходимо знать двумерную плотность вероятностей как функцию x_1, x_2 и параметра τ .

Приведем в качестве примера стационарный процесс, характеризующийся тем, что случайная величина принимает с равной вероятностью значения $\pm a$, причем перемена знака происходит в случайные моменты. Задается распределение нулей как вероятность $p(n, \tau)$ того, что на интервале τ окажется ровно n нулей.

Двумерная плотность вероятностей для такого процесса может быть выражена формулой

$$\begin{aligned} \varphi(x_1, x_2, \tau) = & \frac{1}{2} [\delta(x_1 - a, x_2 - a) + \\ & + \delta(x_1 + a, x_2 + a)] \sum_{k=0}^{\infty} p(2k, \tau) + \\ & + \frac{1}{2} [\delta(x_1 - a, x_2 + a) + \delta(x_1 + a, x_2 - a)] \sum_{k=0}^{\infty} p(2k + 1, \tau). \end{aligned}$$

При интегрировании этого выражения нужно учесть, что

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = f(x_0).$$

Здесь δ — символ дельта-функции. Вычисление для случая пуассонова распределения, т. е.

$$p(n, \tau) = \frac{(\mu\tau)^n}{n!} e^{-\mu\tau},$$

дает

$$B(\tau) = a^2 e^{-2\mu\tau},$$

где μ — средняя частота нулей. Этот известный результат получается обычно иным путем.

5. Спектр случайного процесса $G(\omega)$ и функция корреляции $B(\tau)$ связаны друг с другом парой преобразований Фурье. Эти преобразования симметричны, из чего можно заключить, что любая формула для вычисления $B(\tau)$ должна соответствовать симметричная формула для вычисления $G(\omega)$. Пересматривая ранее приведенные формулы, мы видим, что симметрия неполна: некоторые формулы не имеют парных. Выведем недостающие.

Определим случайную функцию

$$\xi_t(x) = \begin{cases} \xi(x) & x < t \\ 0 & x > t \end{cases}$$

и введем текущий спектр

$$S_t(\omega) = \int_{-\infty}^t \xi(x) e^{-i\omega x} dx = \int_{-\infty}^{\infty} \xi_t(x) e^{-i\omega x} dx.$$

Запишем обратное преобразование

$$\xi_t(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_t(\omega) e^{i\omega x} d\omega.$$

Умножим это выражение на $\xi_t(x - \tau) dx$ и проинтегрируем в бесконечных пределах

$$\int_{-\infty}^{\infty} \xi_t(x) \xi_t(x - \tau) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \xi_t(x - \tau) dx \int_{-\infty}^{\infty} S_t(\omega) e^{i\omega x} d\omega.$$

Изменив порядок интегрирования в правой части, получим

$$\int_{-\infty}^{\infty} \xi_t(x) \xi_t(x - \tau) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |S_t(\omega)|^2 e^{i\omega\tau} d\omega.$$

Изменим пределы интегрирования в левой части на основании определения функции $\xi_t(x)$: верхним пределом будет t . Дифференцирование по t дает

$$\xi(t)\xi(t-\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\partial}{\partial t} |S_t(\omega)|^2 e^{i\omega\tau} d\omega.$$

Беря среднее по множеству, получаем

$$M[\xi(t)\xi(t-\tau)] = B(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} M\left(\frac{\partial}{\partial t} |S_t(\omega)|^2\right) e^{i\omega\tau} d\omega.$$

Таким образом,

$$G(\omega) = \frac{1}{\pi} M\left(\frac{\partial}{\partial t} |S_t(\omega)|^2\right). \quad (7)$$

Заметим, что величина в скобках — та самая величина, которую Пейдж [1] предложил в качестве мгновенного спектра мощности. Аналогичным определением пользуется Ламперд [2].

Возьмем теперь формулу (1) и введем в нее значение $B(\tau)$ согласно (6). Мы получим

$$G(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \psi(x_1, x_2, \omega) dx_1 dx_2, \quad (8)$$

где

$$\psi(x_1, x_2, \omega) = \int_{-\infty}^{\infty} \varphi(x_1, x_2, \tau) e^{-i\omega\tau} d\tau.$$

Теперь мы можем собрать все формулы для $B(\tau)$ и $G(\omega)$:

$$\begin{array}{l|l} B(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} G(\omega) e^{i\omega\tau} d\omega = & G(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} B(\tau) e^{-i\omega\tau} d\tau = \\ = M[\xi(t)\xi(t-\tau)] = & = M\left(\frac{\partial}{\partial t} |S_t(\omega)|^2\right) = \\ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 \varphi(x_1, x_2, \tau) dx_1 dx_2 = & = \frac{1}{2\pi} \int_{-\infty}^{\infty} x_1 x_2 \psi(x_1, x_2, \omega) dx_1 dx_2 = \\ = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t)\xi(t-\tau) dt. & = \lim_{T \rightarrow \infty} \frac{1}{T} |S_T(\omega)|^2. \end{array}$$

6. В литературе рассмотрен ряд примеров вычисления спектров случайных процессов. Один из таких процессов упомянут в п. 4; назовем этот процесс «обобщенным телеграфным сигналом».

Другим примером может служить несколько более сложный процесс, у которого не только расположение нулей, но и значение функции в интервале между нулями случайно. Назовем такой процесс «сигналом следящей системы». Оба названных процесса являются стационарными, и спектры этих процессов могут вычисляться непосредственно по приведенным выше формулам. Но наряду с указанными процессами рассматривается и ряд других, некоторые из которых не являются стационарными не только в узком, но и в широком смысле. Вот перечень этих процессов:

«Телеграфный сигнал», принимающий с равной вероятностью значения $\pm a$ с переменной знака только в фиксированные равноотстоящие моменты.

«Сигнал АИМ» (амплитудно-импульсная модуляция), представляющий собой последовательность равноотстоящих во времени импульсов одинаковой формы, но случайной величины.

«Сигнал ФИМ» (фаза-импульсная модуляция), представляющий собой последовательность импульсов равной величины и одинаковой формы, но смещенных относительно равноотстоящих моментов на случайные интервалы.

«Сигнал АМ», представляющий собой синусоидальное колебание со случайной амплитудой, но постоянной фазой.

«Сигнал ФМ» (или ЧМ), представляющий собой синусоидальное колебание со случайной фазой, но постоянной амплитудой.

Нестационарные процессы из числа названных рассматриваются обычно без всяких оговорок (в этом был повинен и автор).

Для удобства читателя здесь приводятся ссылки на литературу (только книжную): обобщенный телеграфный сигнал [3, стр. 118—119], [4, стр. 49], [5, стр. 194—196], [6, стр. 150—151]; сигнал следящей системы [7, стр. 372—376], [6, стр. 151—154]; телеграфный сигнал [3, стр. 119—120], [7, стр. 349—350]; сигнал АИМ [4, стр. 46—47], [5, стр. 189—192], [6, стр. 154—157]; сигнал ФИМ [4, стр. 47—48], [5, стр. 192—194], [6, стр. 158]; сигнал АМ и ФМ [8, стр. 190—192, 197—214].

7. В конце концов нужно договориться, что мы будем понимать под спектром нестационарного случайного процесса. В таком определении ощущается самая настоятельная надобность, потому что, как видно из предыдущего, ряд практически важных сигналов связи относится к нестационарным процессам; с другой стороны, в существовании спектра, например АМ колебания, у инженеров никаких сомнений никогда не возникало.

Естественный подход к проблеме вытекает из представления о случайном процессе как функции двух переменных $\xi(\theta, t)$, где θ — элемент множества. При фиксированном t получаем случайную величину — значение случайного процесса в данный момент; при фиксированном θ получаем ту или иную реализацию случайного процесса. Среднее значение функции двух переменных получается в общем случае усреднением по обоим переменным.

Усреднение по одной переменной дает, вообще говоря, величину, зависящую от второй переменной. Таким образом, для нестационарного процесса среднее по множеству зависит от времени, а среднее по времени зависит от реализации. Для стационарного эргодического процесса среднее по времени совпадает со средним по множеству; это значит, что оба средних, будучи равны друг другу, не зависят ни от времени, ни от реализации.

В случае нестационарного процесса для получения спектров и функций корреляции, не зависящих ни от времени, ни от реализации, необходимо, очевидно, выполнить усреднение дважды (т. е. как по времени, так и по множеству); порядок усреднения безразличен¹. Полученные таким образом функции следовало бы в отличие от обычных называть средним спектром и средней функцией корреляции. Для этих средних функций сохраняется взаимная связь через пару трансформаций Фурье. В самом деле, для зависящих от времени средних по ансамблю это доказано Лампердом; остается выполнить усреднение по времени и мы получим

$$\overline{B(\tau, t)} = \frac{1}{2} \int_{-\infty}^{\infty} \overline{G(\omega, t)} e^{i\omega\tau} d\omega,$$

$$\overline{G(\omega, t)} = \frac{1}{\pi} \int_{-\infty}^{\infty} \overline{B(\tau, t)} e^{-i\omega\tau} d\tau,$$

где

$$B(t, \tau) = M [\xi(t) \xi(t - \tau)].$$

Итак, для вычисления средних спектров и функций корреляции нестационарных процессов можно воспользоваться приведенными в п. 5 формулами, добавив в них второе усреднение в произвольном порядке. Именно к этому сводится методика вычислений, фактически выполняемых в цитированных работах.

В качестве примера применения этих соображений вычислим среднюю функцию корреляции телеграфного сигнала $\xi(t) = \pm a$ с возможным переходом через нуль при $t = t_k = k\tau_0$, где k — любое целое число. Это — нестационарный процесс; его двумерная плотность вероятностей зависит от времени

$$\frac{1}{2} [\delta(x_1 - a, x_2 - a) + \delta(x_1 + a, x_2 + a)] \dots$$

$$\varphi(x_1, x_2, \tau, t) = \begin{matrix} 0 < t < \tau_0 - \tau \\ \frac{1}{4} [\delta(x_1 - a, x_2 - a) + \delta(x_1 + a, x_2 + a) + \\ + \delta(x_1 - a, x_2 + a) + \delta(x_1 + a, x_2 - a)] \\ \tau_0 - \tau < t < \infty \end{matrix}$$

¹) Двойное усреднение применено к рассматриваемому классу нестационарных процессов Бунимовичем [9].

Все моменты, конечно, также будут зависеть от времени. Но так как усреднение можно производить на любом этапе, то мы усредним по времени самое распределение

$$\varphi(x_1, x_2, \tau) = \overline{\varphi(x_1, x_2, \tau, t)} = \frac{1}{2} [\delta(x_1 - a, x_2 - a) + \delta(x_1 + a, x_2 + a)] \quad 0 < \tau < \tau_0,$$

$$\varphi(x_1, x_2, \tau) = \overline{\varphi(x_1, x_2, \tau, t)} = \frac{1}{4} [\delta(x_1 - a, x_2 - a) + \delta(x_1 + a, x_2 + a) + \delta(x_1 - a, x_2 + a) + \delta(x_1 + a, x_2 - a)] \quad \tau_0 < \tau < \infty.$$

Подставив в (6) и выполнив интегрирование, найдем

$$\overline{B(\tau)} = \begin{cases} a^2(\tau_0 - \tau)/\tau_0 & 0 < \tau < \tau_0, \\ 0 & \tau_0 < \tau < \infty. \end{cases}$$

8. Для некоторых нестационарных процессов достаточно одного усреднения. Рассмотрим к примеру АМ-колебание

$$\eta(t) = \xi(t) \cos \omega_0 t,$$

где $\xi(t)$ — стационарный процесс. Составим произведение

$$\eta(t) \eta(t + \tau) = \xi(t) \xi(t + \tau) \cos \omega_0 t \cos \omega_0 (t + \tau) \quad (9)$$

и усредним его по множеству

$$B_\eta(\tau, t) = \frac{1}{2} B_\xi(\tau) [\cos \omega_0 (2t + \tau) + \cos \omega_0 \tau].$$

Последующее усреднение по времени дает

$$\overline{B_\eta(\tau, t)} = \frac{1}{2} B_\xi(\tau) \cos \omega_0 \tau,$$

откуда находится и спектр

$$\overline{G_\eta(\omega, t)} = \frac{1}{4} [G_\xi(\omega_0 + \omega) + G_\xi(\omega_0 - \omega)].$$

Но можно поступить и иначе: усреднить (9) сразу по времени. Это даст

$$\begin{aligned} \overline{\eta(t) \eta(t + \tau)} &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) \xi(t + \tau) \cos \omega_0 t \cos \omega_0 (t + \tau) dt = \\ &= \frac{1}{2} \cos \omega_0 \tau \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) \xi(t + \tau) dt + \\ &+ \frac{1}{2} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \xi(t) \xi(t + \tau) \cos \omega_0 (2t + \tau) dt. \end{aligned}$$

Для стационарного $\xi(t)$ первый член дает $\frac{1}{2} B_{\xi}(\tau) \cos \omega_0 \tau$, а второй, как это можно показать, равен нулю.

Итак, в рассматриваемом случае усреднение по времени дает неслучайную величину, тогда как усреднение по множеству дает функцию времени. Определение и общие свойства нестационарных процессов, для которых получают такого рода соотношения, требуют уточнения.

Л и т е р а т у р а

1. *C. H. Page*. Instantaneous power spectra. *J. Appl. Phys.*, 23, 1, 1952, 103—106.
2. *D. G. Lampard*. Generalization of the Wiener-Khintchine theorem to non-stationary processes. *J. Appl. Phys.*, 25, 6, 1954, 862—803.
3. *С. О. Райс*. Теория флюктуационных шумов. Сб.: Теория передачи электрических сигналов при наличии помех. ИЛ, 1953.
4. *Дж. Лоусон, Дж. Уленбек*. Пороговые сигналы. «Советское радио», 1952.
5. *Цянь-Сюэ-Сэнь*. Техническая кибернетика. ИЛ, 1956.
6. *В. В. Солодовников*. Введение в статистическую динамику систем автоматического управления. Гостехиздат, 1952.
7. *Х. Джеймс, Н. Никольс, Р. Филлипс*. Теория следящих систем. ИЛ, 1953.
8. *И. С. Гоноровский*. Радиосигналы и переходные явления в радиопежах. Связьиздат, 1954.
9. *В. И. Бунимович*. Флюктуационные процессы в радиоприемных устройствах. «Советское радио», 1951.

О ВОЗМОЖНОСТЯХ СЖАТИЯ СПЕКТРА

Проблема сжатия спектра в настоящее время еще далека от полного решения; это относится в особенности к теоретической стороне дела. Совершенно неясен критерий «сжимаемости» процесса и характер преобразования, которому нужно подвергнуть процесс в целях сжатия его спектра.

Существует, однако, один класс процессов, относительно которого можно сказать кое-что определенное и, возможно, не без пользы для дальнейшего развития теории и техники. Речь идет о модулированных процессах.

Вкратце дело сводится к тому, что при любом виде модуляции спектр модулированного процесса шире, чем спектр модулирующего процесса. Поэтому в рассматриваемом случае сжатие спектра осуществляется путем выделения модулирующего процесса. Система передачи представляется в следующем виде: на передающей стороне из подлежащего передаче процесса выделяется модулирующий процесс; этот процесс (имеющий спектр более узкий, чем исходный) передается по каналу связи; на приемной стороне исходный процесс восстанавливается путем модуляции переносчика, порождаемого местным генератором. Заметим, что ситуация обратна той, которая имеется при обычной радиопередаче, когда на передающей стороне производится модуляция, а на приемной — детектирование, т. е. выделение низкочастотного модулирующего процесса.

Интерес к модулированным процессам определяется не только тем, что мы давно уже знаем довольно много о свойствах этих процессов и об операциях модуляции и детектирования; важно и то, что некоторые действительные процессы, передача которых представляет существенный интерес для связи, являются по своей природе именно модулированными процессами. Такова, в частности, речь.

Речь может быть описана с физической точки зрения как процесс, в образовании которого участвуют по меньшей мере два переносчика — звук голосовых связок, представляющий собою периодическую последовательность импульсов, и звук выдуваемого воздуха, представляющий собою шум с более или менее равномерным спектром. Названные переносчики подвергаются одновременно нескольким видам модуляции. Звук голосовых связок модулируется по амплитуде и по частоте (т. е. изменяется сила

и высота основного тона); оба переносчика подвергаются модуляции типа частотно-шумовой в ротовой полости (по поводу этого термина [1]), чем определяется механизм образования формант. Оба переносчика, далее, подвергаются амплитудной модуляции в процессе формирования согласных.

Не вдаваясь в дальнейшие подробности, можно считать, что речь представляет собою составной (т. е. состоящий из нескольких процессов) сложномодулированный (т. е. с применением нескольких видов модуляции) процесс.

Существенно, что все модулирующие процессы имеют сравнительно очень узкие спектры, ширины которых имеют порядок десятков герц; в совокупности они занимают полосу 100—200 *гц*. Передав модулирующие процессы по такому узкому каналу, можно восстановить (синтезировать) речь на приемной стороне путем модуляции соответствующих переносчиков. В этом и состоит идея «вокодера», высказанная и осуществленная Дадли почти 20 лет тому назад [2].

Передача посредством полосового анализа. Основная идея вокодера, состоящая в выделении и передаче модулирующих процессов, была в самом начале высказана, хотя и в несколько иных терминах, но достаточно отчетливо. С тех пор появилось несколько вариантов, известных под названием формантных, фонемных, полосовых вокодеров. Ниже рассматриваются возможности одной из таких систем, представленной в схематизированном виде.

Рассмотрим систему, на передающей стороне которой имеется набор полосовых фильтров, перекрывающий весь спектр подлежащего передаче процесса. На выходе каждого полосового фильтра будет наблюдаться мгновенный спектр, т. е. функция времени и номер фильтра. Мгновенный спектр — это преобразование Фурье с весом, причем весовая функция привязана к текущему времени: вид этой функции зависит от характеристики на выходе каждого данного фильтра [3, § 6]. Таким образом, напряжение на выходе каждого данного фильтра будет зависеть только от времени. Если это напряжение меняется медленно и если число фильтров невелико, то совокупность выходных напряжений займет узкую полосу. Для восстановления исходного процесса на приемной стороне имеются шумовой генератор, аналогичный набор фильтров и система амплитудных модуляторов, управляющих интенсивностью в каждой полосе. Опишем теперь действие системы аналитически.

Пусть исходный процесс обозначен $x(t)$; это есть напряжение, подаваемое одновременно на входы всех фильтров. Напряжение на выходе i -го фильтра равно

$$y_i(t) = \int_{-\infty}^t x(\tau) g_i(t - \tau) d\tau, \quad (1)$$

где $g(t)$ — импульсная реакция фильтра. Для узкополосного фильтра

$$g_i(t) = \frac{\Delta\omega}{\pi} S(t) \cos \omega_i(t). \quad (2)$$

Здесь $\Delta\omega$ — ширина полосы прозрачности, предполагаемая для всех фильтров одинаковой; ω_i — центральная частота полосы прозрачности i -го фильтра. Медленная (по сравнению с $\cos \omega_i t$) функция $S(t)$ зависит от характеристики выбранного фильтра. Так, например, для идеального полосового фильтра с частотной характеристикой

$$K_i(\omega) = \begin{cases} 1 & \omega_i - \frac{1}{2}\Delta\omega < \omega < \omega_i + \frac{1}{2}\Delta\omega, \\ 0 & \text{вне указанной полосы} \end{cases}$$

имеем

$$S(t) = \frac{\sin \frac{1}{2} \Delta\omega t}{\frac{1}{2} \Delta\omega t}.$$

Таким образом,

$$\begin{aligned} y_i(t) &= \frac{\Delta\omega}{\pi} \int_{-\infty}^t x(\tau) S(t-\tau) \cos \omega_i(t-\tau) d\tau = \\ &= \frac{\Delta\omega}{\pi} \operatorname{Re} e^{i\omega_i t} \int_{-\infty}^t x(\tau) S(t-\tau) e^{-i\omega_i \tau} d\tau \end{aligned}$$

или

$$y_i(t) = \frac{\Delta\omega}{\pi} \operatorname{Re} F(\omega_i, t) e^{i\omega_i t}. \quad (3)$$

Функция

$$F(\omega_i, t) = \int_{-\infty}^t x(\tau) S(t-\tau) e^{-i\omega_i \tau} d\tau. \quad (4)$$

является мгновенным спектром и выражает спектральную плотность в i -й полосе. С другой стороны, эта функция может рассматриваться как модулирующая функция, как это ясно видно из (3). Совокупность функций $F(\omega_i, t)$ и передается по каналу связи.

Обратимся к приемной стороне. Узкополосный шум на выходе каждого полосового фильтра может быть представлен в виде [4, гл. IV]

$$\xi_i(t) = a(t) \cos [\omega_i t + \varphi(t)],$$

где $a(t)$ и $\varphi(t)$ — медленные случайные процессы. Иначе говоря, процесс на выходе узкополосного фильтра может быть пред-

ставлен, как синусоидальное колебание с частотой ω_i , модулированное случайным образом как по амплитуде, так и по фазе. Но чем полоса уже, тем слабее модуляция; на существование рассуждения не повлияет, если мы будем считать каждое отдельное напряжение, подвергнутое модуляции на приемной стороне, чисто синусоидальным. Технически это означает применение для восстановления сигнала не шумового генератора с набором полосовых фильтров, а набора генераторов синусоидальных колебаний с частотами ω_i . Каждое такое колебание должно быть промодулировано переданной по каналу связи модулирующей функцией $F(\omega_i, t)$. Тогда на приемной стороне получаются в точности такие же модулированные колебания $y_i(t)$, которые имелись на передающей стороне (3). Для восстановления сигнала остается эти колебания сложить

$$y(t) = \sum_{i=1}^n y_i(t) = \frac{\Delta\omega}{n} \sum_i F(\omega_i, t) e^{i\omega_i t} = \\ = \frac{\Delta\omega}{n} \sum \operatorname{Re} e^{i\omega_i t} \int_{-\infty}^t x(\tau) S(t-\tau) e^{-i\omega_i \tau} d\tau$$

или

$$y(t) = \frac{\Delta\omega}{\pi} \int_{-\infty}^t x(\tau) S(t-\tau) \left(\operatorname{Re} \sum_i e^{i\omega(t-\tau)} \right) d\tau. \quad (5)$$

Очевидно, что $y(t)$ отличается от $x(t)$, т. е. что передаваемый сигнал воспроизводится с искажением. Конечно, можно вычислить отклонение y от x на основе того или иного критерия, например, критерия квадратичной ошибки. Однако если имеется в виду задача речи, то такие вычисления не имеют смысла, так как объективные критерии субъективного качества воспроизведения речи до сих пор неизвестны.

Важно другое, а именно то, что при неограниченном увеличении числа n частотных полос с одновременным уменьшением их ширины $\Delta\omega$ мы получаем стремление y к x в пределе. Действительно,

$$\lim_{\Delta\omega \rightarrow 0} S(t) = 1,$$

сумма заменяется интегралом, и мы имеем из (5)

$$y(t) \xrightarrow[\substack{\Delta\omega \rightarrow 0 \\ n \rightarrow \infty}]{} \frac{1}{\pi} \int_{-\infty}^t x(\tau) \left(\operatorname{Re} \int_0^{\infty} e^{i\omega(t-\tau)} d\omega \right) d\tau = \\ = \frac{1}{\pi} \int_{-\infty}^t x(\tau) d\tau \int_0^{\infty} \cos \omega(t-\tau) d\omega.$$

Но внутренний интеграл есть не что иное, как δ -функция. Таким образом,

$$y(t) \xrightarrow[\substack{\Delta\omega \rightarrow 0 \\ n \rightarrow \infty}]{\quad} \int_{-\infty}^t x(\tau) \delta(t - \tau) d\tau = x(t).$$

Теперь положение определилось: если мы и сокращаем полосу частот путем передачи не $x(t)$, а совокупности модулирующих функций $F(\omega_i, t)$, то это сокращение происходит за счет искажения сигнала (тем большего, чем больше сокращение, т. е. чем меньше n). Для точного воспроизведения сигнала этим методом нужно передать весь спектр полностью (к этому и сводится вышеописанный предельный переход); никакого сокращения при этом, понятно, не получится.

Чтобы сделать ситуацию еще более очевидной, представим себе, что передается фазово-модулированное колебание

$$x(t) = \sin[\omega_0 t + \Delta\varphi f(t)].$$

Такой сигнал вызовет изменяющиеся во времени отклики на выходе любого набора фильтров. Суммирование конечного числа таких откликов никогда не даст в точности сигнала $x(t)$. Между тем для совершенно точного воспроизведения сигнала $x(t)$ на приемной стороне достаточно передать по каналу связи только одну модулирующую функцию $f(t)$. Это и будет в данном случае наилучшим способом передачи, при котором требуется наименьшая полоса частот.

О представлении составных процессов. Всякий процесс представить можно в виде

$$x(t) = \tilde{a}(t) \cos \tilde{\varphi}(t). \quad (6)$$

Это представление становится однозначным, если воспользоваться понятием аналитического сигнала, т. е. если положить

$$x(t) = \operatorname{Re} \tilde{a}(t) e^{i\tilde{\varphi}(t)} = \operatorname{Re} X(t). \quad (7)$$

Можно рассматривать $\tilde{a}(t)$ и $\tilde{\varphi}(t)$ как модулирующие функции, и спрашивается: можно ли и при каких условиях сэкономить на полосе, передавая вместо $x(t)$ модулирующие функции $\tilde{a}(t)$ и $\tilde{\varphi}(t)$?

Рассмотрим сначала пример. Пусть дан процесс, представляющий собою сумму двух простых АМ-колебаний

$$x(t) = x_1(t) \cos \omega_1 t + x_2(t) \cos \omega_2 t. \quad (8)$$

Ширина спектра этого процесса не превосходит удвоенной суммы ширин спектров функций x_1 и x_2 . Если передавать только модулирующие функции x_1 и x_2 , то потребуется вдвое меньшая полоса.

Представим теперь (8) в форме (6). Для «амплитуды» (огibaющей) $a(t)$ имеем

$$a(t) = \sqrt{x_1^2 + x_2^2 + 2x_1x_2 \cos(\omega_1 - \omega_2)t}. \quad (9)$$

Вместо мгновенной фазы $\varphi(t)$ будем рассматривать мгновенную частоту $\omega(t)$, определяемую как

$$\omega(t) = \frac{d\varphi}{dt} = \text{Im} \frac{1}{X} \cdot \frac{dx}{dt}. \quad (10)$$

Нужно заметить, что применение понятия мгновенной частоты требует известной осмотрительности; серьезный и содержательный анализ связанных с этим понятием представлений можно найти в работе Дитля [5].

Для мгновенной частоты находим

$$\omega(t) = \frac{1}{a^2(t)} [\omega_1 x_1^2 + \omega_2 x_2^2 + (\omega_1 + \omega_2) x_1 x_2 \cos(\omega_1 - \omega_2)t]. \quad (11)$$

Задача состоит теперь в том, чтобы оценить ширины спектров функций $a(t)$ и $\varphi(t)$. Ширины спектров функций $a(t)$ и $x(t)$ имеют одинаковый порядок. Поэтому важно найти оценку ширины спектра мгновенной частоты.

Формула (11) состоит из двух сомножителей. Ширина спектра произведения не менее суммы ширин спектров сомножителей [3, добавление I]. Для выражения в квадратных скобках получается такая же оценка ширины спектра, как и для $a^2(t)$.

Займемся отдельно множителем

$$\frac{1}{a^2(t)} = \frac{1}{x_1^2 + x_2^2 + 2x_1x_2 \cos(\omega_1 - \omega_2)t} = f_1(t).$$

Перепишем это выражение в виде

$$f_1(t) = \frac{1}{x_1^2 + x_2^2} \cdot \frac{1}{1 + \frac{2x_1x_2}{x_1^2 + x_2^2} \cos \Omega t}.$$

Напомним, что x_1 и x_2 — заданные медленные функции времени. Будем, однако, считать их постоянными. Тогда дело сведется к нахождению спектра периодической функции

$$f_2(t) = \frac{1}{1 + a \cos \Omega t}.$$

Истинный спектр будет отличаться от линейчатого спектра функции f_2 тем, что дискретные линии будут размыты тем более, чем больше дисперсии x_1 и x_2 ; однако общая тенденция спектра (т. е. закон убывания огibaющей спектра) при сделанном упрощении не изменится. Так как $f_2(t)$ — четная функция, то

$$C_k = \frac{2}{T} \int_{-T/2}^{T/2} \frac{\cos k\Omega t dt}{1 + a \cos \Omega t} = \frac{2}{\pi} \int_0^\pi \frac{\cos kx dx}{1 + a \cos x} = \frac{2}{\sqrt{1 - a^2}} \left(\frac{\sqrt{1 - a^2} - 1}{a} \right)^k.$$

Если $\alpha \ll 1$, то спектр быстро убывает (по геометрической прогрессии со знаменателем α). Если же $\alpha \simeq 1$ (что соответствует случаю $x_1 \simeq x_2$), то величина в круглых скобках сколь угодно мало отличается от единицы, и спектр убывает сколь угодно медленно. Это означает, что ширина спектра функции f_2 , а следовательно, и ширина спектра мгновенной частоты ω может быть сколь угодно большой, как бы мы ни определили ширину спектра.

Рассмотренный пример позволяет сделать некоторые заключения общего характера. Мы видим, что представление $a(t) \cos \varphi(t)$ с последующим использованием для передачи функций $a(t)$ и $\varphi(t)$ (или $\omega = \varphi'$) — вовсе невыгодно применительно к составным процессам. Под составным понимается процесс, представляющий собою в простейшем случае сумму простых модулированных процессов (как в нашем примере), а в более общем понимании — некоторую функцию от нескольких простых процессов. Под простыми понимаются процессы вида $f(t) \cos [\omega t + \varphi(t)]$, где $f(t)$ и $\varphi(t)$ — медленные функции сами по себе «несжимаемые». Для АМ-процессов можно ввести несколько более общее определение вида $f(t) \eta(t)$, где $\eta(t)$ — переносчик, которым может служить не только синусоидальное колебание, но и другая периодическая функция (например, периодическая последовательность импульсов), или даже быстрый стационарный процесс.

Неуспех попыток осуществить систему, описанную Марку и Даге [6], объясняется, по-видимому, тем, что представление (6) применимо к речи, которая является заведомо составным процессом¹.

На основании сказанного путь к сжатию спектра сигнала (основанный на представлении о сигнале, как о модуляционном процессе) состоит в следующем: прежде всего нужно разложить данный процесс на простые модулированные процессы, затем выделить модулирующие функции и передать их по каналу. На приемной стороне сигнал синтезируется в обратном порядке.

Итак, нужно найти математический метод разложения процесса на простые, и разработать адекватную техническую процедуру. Совершенно неясно пока, как это сделать; нужно признать, что даже корректная постановка задачи невозможна, пока не будет дано четкое определение того, что выше было названо простым процессом.

Однако если иметь в виду, в частности, речь, то дело облегчается тем, что можно а priori перечислить образующие речь простые процессы, опираясь на знание устройства и действия речевого аппарата.

¹ Упомянутое в [6] сжатие спектра путем деления частоты, уменьшающего индекс, имеет место, пока индекс остается большим. Однако при сколь угодно малом индексе ширина спектра ЧМ-колебания не может быть, как известно, меньше удвоенной ширины спектра модулирующей функции.

Л и т е р а т у р а

1. А. А. Харкевич. Передача сигналов модулированным шумом. Электросвязь, 1957, № 11.
2. H. Dudley. Remaking speech. JASA, 11, 2, 1939, 169—177.
3. А. А. Харкевич. Спектры и анализ. Гостехиздат, 1957.
4. В. И. Бунимович. Флуктуационные процессы в радиоприемных устройствах. «Советское радио», 1951.
5. А. Dittl. Einschwingvorgänge bei Frequenzmodulation Hft. EA, 64, 6, 1954, 184—193.
6. П. Марку, Ж. Даге. Новые методы передачи речи. Теория передачи сообщений. ИЛ, 1957.

О ТЕОРЕМЕ КОТЕЛЬНИКОВА

Теорема Котельникова (в иностранной литературе *sampling theorem*) в формулировке автора [1] гласит (теорема II): «Любую функцию $F(t)$, состоящую из частот от 0 до f_1 , можно непрерывно передавать с любой точностью при помощи чисел, следующих друг за другом через $1/2f_1$ сек». Обобщением этого предложения является теорема V: «Любую функцию $F(t)$, состоящую из частот от f_1 до f_2 , можно передавать с любой точностью при помощи чисел, передаваемых друг за другом через $1/2(f_2 - f_1)$ сек». Таким образом, для определения функции на интервале T требуется $n = 2TW$ чисел, где $W = f_2 - f_1$.

Теорема II получается из теоремы V, если положить нижнюю граничную частоту спектра равной нулю. В дальнейшем для упрощения мы будем иметь в виду именно этот случай.

Утверждения вышеприведенных теорем легко получить в наглядной форме, рассматривая представление функции на любом интервале посредством тригонометрических многочленов, т. е. усеченных рядов Фурье. Однако интересно другое представление (составляющее содержание теоремы I), а именно:

«Любую функцию $F(t)$, состоящую из частот от 0 до f_1 периодов в секунду, можно представить рядом

$$F(t) = \sum_{-\infty}^{\infty} D_k \frac{\sin \omega_1 (t - k/2f_1)}{t - k/2f_1}, \quad (1)$$

где k — целое число; $\omega_1 = 2\pi f_1$; D_k — постоянные, зависящие от $F(t)$. И, наоборот, любая функция $F(t)$, представленная рядом (1), состоит лишь из частот от 0 до f_1 периодов в секунду».

Теорема II является прямым следствием теоремы I.

Переписав (1) в принятых в настоящее время обозначениях

$$f(t) = \sum_{k=-\infty}^{\infty} f(k\Delta t) \frac{\sin \omega_c (t - k\Delta t)}{\omega_c (t - k\Delta t)}, \quad (2)$$

заметим, что величины $f(k\Delta t)$ представляют собою не что иное, как мгновенные значения функции $f(t)$, отсчитанные в моменты $t = k\Delta t$. Интервал Δt связан с верхней граничной частотой $\omega_c = 2\pi f_c$ соотношением

$$\Delta t = \pi/\omega_c = 1/2f_c.$$

Легко понять значение теоремы Котельникова для техники связи, она позволяет заменить передачу непрерывной функции передачей дискретной последовательности. Теорема эта лежит, таким образом, в основе всех видов импульсной связи.

Однако за последнее время в связи с развитием теории случайных процессов и вероятностных аспектов теории связи возникла необходимость в пересмотре значения теоремы Котельникова.

Дело в том, что теорема Котельникова как точное утверждение справедлива только для функций, спектральная плотность которых равна нулю для всех частот вне некоторого конечного интервала, называемого обычно шириной спектра. Между тем, если рассматривать передаваемую функцию как случайный процесс, каковым она в действительности является, то предположение о строгой ограниченности спектра приводит к несколько неожиданному результату. Оказывается [2], что случайный процесс, спектральная плотность которого равна нулю на конечном интервале, является процессом сингулярным¹. Это значит, что значения процесса могут быть предсказаны с любой точностью на любое время вперед. А это, в свою очередь, означает, что процессы со строго ограниченным спектром не могут переносить информацию (в том смысле, что течение процесса не сопровождается приростом информации). Иначе говоря, никакой процесс, представляющий собой сигнал связи, принципиально не может иметь ограниченный спектр. А если так, то теорема Котельникова имеет в виду объекты, не представляющие интереса для техники связи, и роль и значение теоремы требуют критической переоценки. Этим и объясняется появление в 1956—1957 гг. нескольких работ, посвященных теореме Котельникова.

Самое естественное при описанных обстоятельствах — это рассматривать теорему Котельникова не как точное утверждение, относящееся к функциям с ограниченным спектром, а как приближенное утверждение, относящееся к функциям с неограниченным спектром. Можно, оставляя в стороне теоретико-информационные аспекты проблемы, рассматривать вопрос о приближенном представлении функции с неограниченным спектром при помощи ряда (2) и о погрешности такого приближения.

Такую постановку вопроса мы находим в работе Турбовича [3]. Пусть $f(t)$ — функция с неограниченным спектром, а ω_c — произвольно выбранное значение частоты, принимаемое в качестве граничного. Показано, что, сохраняя соотношение (3), функцию можно представить в виде суммы трех слагаемых

$$f(t) = f_1(t) + f_2(t) + f_3(t),$$

где $f_1(t)$ выражается рядом (2), т. е. не содержит составляющих с частотами выше ω_c , спектр функции $f_2(t)$ неограничен, а спектр функции $f_3(t)$ не содержит составляющих с частотами ниже ω_c .

¹ Это не относится к процессам, спектральная плотность которых обращается в нуль в отдельных точках.

Если значение ω_c выбрано так, что при $|\omega| > \omega_c$ спектральная плотность функции $f(t)$ мала, то слагаемые $f_2(t)$ и $f_3(t)$ малы по сравнению с $f_1(t)$ и можно записать

$$f(t) \approx f_1(t) = \sum_{k=-\infty}^{\infty} f(k\Delta t) \frac{\sin \omega_c(t - k\Delta t)}{\omega_c(t - k\Delta t)}. \quad (4)$$

Напомним, что $f(t)$ имеет неограниченный спектр, а $f_1(t)$ — спектр, ограниченный сверху частотой ω_c .

Для погрешности приближения (4) получаются оценки

$$\frac{E_c}{E} \leq \frac{[f(t) - f_1(t)]^2}{[f(t)]^2} \leq (3 + Q) E_c/E. \quad (5)$$

Оцениваемая величина — это нормированный интеграл квадрата ошибки, т. е. отклонения $f_1(t)$ от $f(t)$; E — полная энергия спектра; E_c — энергия, содержащаяся в высокочастотной части спектра, т. е. полосе частот выше ω_c . Выражение (5) можно переписать в виде

$$E_c \leq \Delta E \leq (3 + Q) E_c,$$

где $\Delta E = \overline{[f(t) - f_1(t)]^2} = \int_{-\infty}^{\infty} [f(t) - f_1(t)]^2 dt$ — величина, которую

можно назвать энергией ошибки (так как $E = \int_{-\infty}^{\infty} f^2(t) dt$ — полная энергия).

Величина Q зависит от закона убывания спектральной плотности для частот, выше граничной. В работе рассмотрены случаи убывания спектральной плотности по экспоненциальному и степенному законам. Если убывание происходит достаточно быстро, то Q — малая величина; при этом энергия ошибки заключена между E_c и $3E_c$.

Сходным образом ставится задача в работе Немировского. Здесь рассматривается последовательность модулированных по амплитуде отсчетами функции $f(t)$ [с неограниченным спектром $F(\omega)$] импульсов произвольной формы $g(t)$ [со спектром $G(\omega)$], т. е. последовательность вида

$$[f'(t)] = \sum_{k=-\infty}^{\infty} f(k\Delta t) g(t - k\Delta t).$$

Эта последовательность пропускается (на приемной стороне) через фильтр с коэффициентом передачи $K(\omega)$; на выходе фильтра получается функция

$$[f(t)] = \sum_{k=-\infty}^{\infty} f(k\Delta t) r(t - k\Delta t), \quad (6)$$

где $r(t)$ — отклик фильтра на импульс $g(t)$. Функция $[f(t)]$, очевидно, отличается от исходной функции $f(t)$; требуется оценить ошибку. Для этого составляется выражение спектра $[F(\omega)]$ функции $[f(t)]$ в виде

$$[F(\omega)] = \Delta t R(\omega) F(\omega) + \Delta t R(\omega) \sum_{m=-\infty}^{\infty} F(2m\omega_c + \omega). \quad (7)$$

В этом выражении

$$R(\omega) = G(\omega) K(\omega)$$

[т. е. спектр $r(t)$]; сумма берется по всем m , кроме $m=0$: член, соответствующий этому значению, выписан отдельно в качестве первого слагаемого. Сначала рассматривается случай

$$R(\omega) = \begin{cases} 1/\Delta t & |\omega| < \omega_c, \\ 0 & |\omega| > \omega_c, \end{cases} \quad (8)$$

т. е. идеализированный случай, когда $R(\omega)$ имеет вид характеристики идеального фильтра нижних частот.

В этом случае

$$[f(t)] = \sum_{k=-\infty}^{\infty} f(k\Delta t) \frac{\sin \omega_c(t - k\Delta t)}{\omega_c(t - k\Delta t)}, \quad (9)$$

но это, конечно, не означает, что $[f(t)] = f(t)$; дело в том, что $f(t)$ — функция с неограниченным спектром, а потому (9) есть приближенное представление функции $f(t)$; функция $[f(t)]$ есть не что иное, как функция $f_1(t)$ из работы Турбовича [см. формулу (4)].

Ошибку приближения Немировский оценивает при помощи спектров. Для рассматриваемого случая, отвечающего условию (8), из (7) сразу находится спектр ошибки

$$\Delta F(\omega) = F(\omega) - [F(\omega)] = \begin{cases} -\sum_m F(2m\omega_c + \omega) & |\omega| < \omega_c, \\ F(\omega) & |\omega| > \omega_c. \end{cases}$$

Ясно, что ошибка тем меньше, чем меньше спектральная плотность $F(\omega)$ при $|\omega| > \omega_c$, т. е. при частотах выше той, которая принята в качестве граничной. В работе даны оценки ошибки в общем виде, а также для случая, когда спектральная плотность при $|\omega| > \omega_c$ убывает по закону $|\omega|^{-n}$. В этом случае

$$\frac{E_c}{E} \leq \frac{\Delta E}{E} \leq (1+k) \frac{E_c}{E}$$

обозначения взяты те же, что в формуле (5). Коэффициент k зависит только от n . С возрастанием n $k \rightarrow 2$. Таким образом, оценки Немировского и Турбовича совпадают.

Что касается общего случая произвольного $R(\omega)$, то его предлагается рассматривать с принятием некоторых упрощений, а именно: полагая, что $R(\omega)$ мало отличается от (8), что поэтому первый член (7) дает хорошее приближение для $F(\omega)$ и что, следовательно, ошибка определяется вторым членом (7). В этих предположениях без труда получаются общие оценки, зависящие от вида функций $F(\omega)$ и $R(\omega)$.

В заключение ставится (но не исследуется) вопрос о построении разложений по функциям, отличным от $\sin x/x$.

В работе Железнова [4] ставится с самого начала более общая задача найти наилучшее (с точки зрения некоторого определенного критерия верности) представление случайной функции $u(t)$ разложением вида

$$v(t) = \sum_k u_k f(t - k\tau_1), \quad (10)$$

где u_k — случайные величины, связанные с $u(t)$ и могущие, в частности, представлять собой значения $u(t)$ в определенные моменты времени.

Физически подобное разложение осуществляется в схеме АИМ, в которой последовательность коротких импульсов (с периодом следования τ_1) модулируется по амплитуде напряжением $u(t)$, пропущенным через фильтр Φ_0 .

Модулированная последовательность пропускается через фильтр Φ_1 с импульсной реакцией $f(t)$, на выходе которого получаем функцию $v(t)$. Такую схему автор называет схемой класса А; в схеме класса Б отсутствует фильтр Φ_0 . Утверждается, что наилучшая верность в схеме А получается, когда фильтры Φ_0 и Φ_1 одинаковы, и что схема А дает наилучшую верность при любом заданном фильтре Φ_1 [т. е. при любой заданной функции $f(t)$, по которой, согласно (10), разлагается $v(t)$].

Для вышеописанных преобразований справедливы некоторые теоремы, первая из которых (теорема II оригинала) утверждает, что наилучшим для схемы А является разложение по функции $f(t)$, являющейся первой собственной функцией интегрального уравнения

$$\int_0^{\tau_1} R(t - t') f_n(t') dt' = \lambda_n^2 f_n(t), \quad (11)$$

где $R(\tau)$ — функция корреляции; λ_n — собственное значение уравнения. Имеется в виду обычный квадратичный критерий верности. При указанном наилучшем разложении среднеквадратичная ошибка равна

$$v_0 = 1 - \lambda_1^2 / \tau_1.$$

В случае неотрицательной корреляции для ошибки получаются следующие оценки

$$1 - \frac{2}{\tau_1} \int_0^{\tau_1} R(\tau) d\tau \leq \nu_0 \leq 1 - \frac{2}{\tau_1} \int_0^{\tau_1} \left(1 - \frac{\tau}{\tau_1}\right) R(\tau) d\tau. \quad (12)$$

Что касается схемы Б (теорема III), то для нее наилучшим является разложение по функции корреляции; точнее говоря, должно быть

$$f(t) = \begin{cases} R(t) & |t| < \tau_1/2, \\ 0 & |t| > \tau_1/2. \end{cases}$$

При таком разложении ошибка равна

$$\nu_0 = 1 - \frac{2}{\tau_1} \int_0^{\frac{\tau_1}{2}} R^2(\tau) d\tau.$$

По смыслу обеих вышеприведенных теорем функция $f(t)$ (а следовательно, фильтры) должна подбираться в зависимости от характера сигнала, так как функция $f(t)$ определяется функцией корреляции сигнала. Можно поставить вопрос о том, чтобы выбрать функцию $f(t)$ раз навсегда. Так, если выбрать

$$f(t) = \begin{cases} 1 & 0 \leq t \leq \tau_1, \\ 0 & \text{вне указанного интервала,} \end{cases}$$

то ошибка будет равна (теореме IV)

$$\nu_0 = 1 - \frac{2}{\tau_1} \int_0^{\tau_1} \left(1 - \frac{\tau}{\tau_1}\right) R(\tau) d\tau, \quad (13)$$

что как раз равно верхней грани ошибки, оцениваемой для наилучшего разложения [см. формулу (12)].

Что касается разложения Котельникова (2), то его можно трактовать, как результат преобразования процесса с неограниченным спектром, выполняемого схемой А с применением в качестве Φ_0 и Φ_1 идеальных фильтров нижних частот с граничной частотой ω_c .

Ошибка при этом равна (теорема V)

$$\nu_0 = 1 - \frac{2}{\tau_1} \int_0^{\tau_1} \left(1 - \frac{\tau}{\tau_1}\right) R(\tau) d\tau, \quad (14)$$

где τ_0 — интервал корреляции (предполагается, что функция корреляции равна нулю при $|t| > \tau_0$); $\tau_1 = \pi/\omega_c$.

И, наконец, рассматривается разложение (4) в качестве приближенного представления процесса с неограниченным спектром.

Ошибка приближения получается ровно вдвое больше значения, даваемого формулой (14) (теорема IV).

В заключение формулируется положение, выражающее обобщенное понимание теоремы Котельникова:

«Непрерывные квазистационарные сигналы $u(t)$ с неограниченным спектром можно передавать при помощи чисел, следующих друг за другом через промежуток τ_1 сек, с верностью, сколь угодно близкой в предельной верности ν_0 , если промежуток τ_1 не превосходит интервала корреляции τ_0 , а длительность сигналов T много больше интервала корреляции τ_0 » (сравни с формулировкой теоремы Котельникова в начале обзора).

Приведенная формулировка подводит методологический итог всем работам, содержание которых было предсказано выше. Существо дела сводится, как видим, к тому, что теорема Котельникова в применении к процессам с неограниченным спектром рассматривается как приближенное утверждение. До сих пор оставалось, однако, нерассмотренным теоретико-информационное содержание теоремы Котельникова. Этот пробел восполняется докладом Колмогорова [5].

Будем исходить из понятия ϵ -энтропии («эпсилон-энтропии»). ϵ -энтропия $H_\epsilon(\xi)$ — это количество информации в случайном объекте ξ , наблюдаемом с точностью ϵ ; это понятие совпадает с шенноновской «скоростью создания сообщений».

В простейшем случае

$$H_\epsilon(\xi) \approx n \log \frac{1}{\epsilon}, \quad (15)$$

где n — число измерений объекта ξ (т. е. ξ рассматривается как n -мерный случайный вектор). Соотношение (15) можно получить из элементарных соображений для квантованного процесса с шагом квантования ϵ).

Оказывается, что и в значительно более общем случае определяющим членом в выражении для ϵ -энтропии является выражение в правой части (15). Имея это в виду, определим теперь число измерений n . Для этого воспользуемся спектральным представлением ϵ -энтропии (для нормального процесса)

$$H_\epsilon(\xi) = \frac{T}{2} \int \lg \frac{f_{\xi\xi}(\lambda)}{\Theta^2} d\lambda. \quad (16)$$

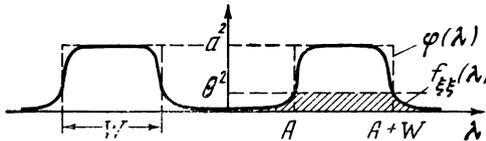
В этой формуле $f_{\xi\xi}$ означает спектральную плотность процесса ξ_1 , T — длительность рассматриваемого отрезка процесса, интеграл берется по интервалу частот, для которого $f_{\xi\xi}(\lambda) > \Theta^2$. Для установления связи между Θ и ϵ нужно учесть, что по физическому смыслу ϵ^2 есть мощность ошибки; ее можно, следовательно, определить по отбрасываемой части спектра, т. е.

$$\epsilon^2 = \int_{-\infty}^{\infty} \min(\Theta^2, f_{\xi\xi}(\lambda)) d\lambda. \quad (17)$$

Теперь обратимся к непосредственно интересующему нас случаю «практически-ограниченного» спектра, т. е. к случаю, когда спектральная плотность может быть приближенно представлена функцией

$$f_{\xi\xi}(\lambda) \approx \varphi(\lambda) = \begin{cases} a^2 & A \leq |\lambda| \leq A+W, \\ 0 & \text{вне указанных интервалов.} \end{cases}$$

График спектральной плотности дан на рисунке. Здесь отмечено



значение Θ^2 ; заштрихованная площадь представляет значение ϵ^2 согласно (17). Легко видеть, что (при не слишком малых ϵ)

$$\epsilon^2 \approx 2W\Theta^2$$

и

$$H_\epsilon(\xi) \approx TW \lg \frac{a^2}{\Theta^2} = TW \lg \frac{2Wa^2}{\epsilon^2},$$

или

$$H_\epsilon(\xi) \approx 2TW \left(\lg \frac{1}{\epsilon} + \lg \sqrt{2W} a \right). \quad (18)$$

Сравнивая первый член этого выражения с (15), мы видим, что величина $2TW$ определяет размерность (т. е. число координат) n .

Таким образом, с теоретико-информационной точки зрения теорема Котельникова определяет число измерений пространства, в котором задается n -мерный вектор, отображающий с требуемой точностью непрерывный процесс $\xi(t)$, причем точность выражается через ϵ -энтропию. Полученные соотношения непосредственно показывают, что с увеличением W (т. е. с увеличением числа измерений) $H_\epsilon(\xi)$, вообще говоря, возрастает (если ϵ неизменно).

Подводя итоги, можно сказать, что наше теперешнее понимание смысла теоремы Котельникова довольно существенно отличается от первоначального. Теорема Котельникова рассматривается в применении к случайным процессам с неограниченным спектром (тогда как первоначально она была сформулирована для детерминированных функций со строго ограниченным спектром): она рассматривается как приближенное (а не точное) утверждение, позволяющее: 1) дать приближенное выражение процесса в форме ряда по запаздывающим функциям или 2) дать приближенную оценку числа измерений случайного вектора, исходя из заданной точности в определении ϵ -энтропии.

Из всего этого отнюдь не следует, что теорема Котельникова утратила свое значение; напротив, она получила более глубокий теоретический смысл и более широкое практическое применение.

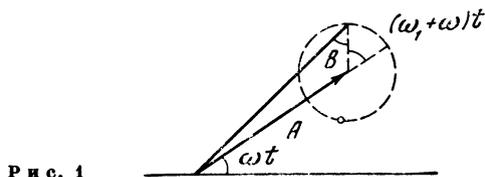
Л и т е р а т у р а

1. *В. А. Котельников*. О пропускной способности «эфира» и проволоки в электросвязи. Материалы к I Всесоюзному съезду по вопросам технической реконструкции дела связи и развития слаботочной промышленности. Всесоюзный энергетический комитет, 1933.
2. *А. Н. Колмогоров*. Стационарные последовательности в гильбертовском пространстве. Бюлл. МГУ, 2, 6, 1941. См. также *А. М. Яглом*. Введение в теорию стационарных случайных функций. УМН, 7, 5 (51), 1952.
3. *И. Т. Турбович*. Некоторое обобщение теоремы Котельникова. «Радиотехника», 11, № 4, 1956.
4. *Н. А. Железнов*. Исследование ортогональных разложений стохастических сигналов с неограниченным спектром. «Радиотехника», 12, 11, 1957.
5. *А. Н. Колмогоров*. Теория передачи информации. Изд-во АН СССР, 1965.

ПАРАЗИТНАЯ МОДУЛЯЦИЯ, ВЫЗЫВАЕМАЯ МАЛОЙ АДДИТИВНОЙ ПОМЕХОЙ ¹

Постановка задачи. Если на немодулированный переносчик налагается помеха, то возникает паразитная модуляция. Так, в случае синусоидального переносчика $f(t) = A \cos \omega_0 t$ наложение синусоидальной помехи $f_1(t) = B \cos \omega_1 t$ приводит к появлению паразитной модуляции как по амплитуде, так и по углу, что поясняется хорошо известной векторной диаграммой (рис. 1).

В случае малой помехи, т. е. при $B \ll A$, наибольшие паразитные приращения амплитуды, фазы и частоты составляют соответственно B , B/A и $(\omega_0 - \omega_1)B/A$. От этих соотношений можно перейти к более общему виду помехи путем соответствующего интегрирования по спектру.



Р и с. 1

Мы ставим себе более общую задачу: найти паразитные изменения параметров переносчика произвольного вида при малой аддитивной помехе, заданной своим спектром.

Пусть переносчик представляет собой функцию

$$f = f(a_1, a_2, a_3, \dots, t), \quad (1)$$

где a_k — параметры, которые могут модулироваться. При отсутствии модуляции параметры a_k постоянны.

Пусть на немодулированный переносчик накладывается малая помеха, представляющая собой стационарный случайный процесс $\xi(t)$ с нулевым средним и известным вторым моментом (т. е. спектром или функцией корреляции). В результате сложения получаем

$$F(t) = f(a_1, a_2, \dots, t) + \xi(t). \quad (2)$$

Подберем теперь функцию

$$F_1(t) = f(a_1 + \delta a_1, a_2 + \delta a_2, \dots, t), \quad (3)$$

¹ Совместно с Э. Л. Блохом.

которая в определенном смысле наименее уклоняется от F . Здесь δa_k означают паразитные приращения, которые получают параметры a_k в результате наложения помехи. Величины δa_k случайны. Как будет показано, их средние значения равны нулю, и задача сведется к нахождению средних квадратов или дисперсий.

В качестве меры меры уклонения F_1 от F выберем его средний квадрат. т. е. введем

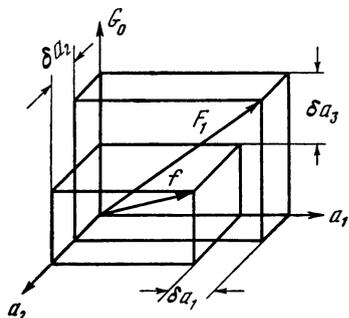
$$d = \overline{[F_1(t) - F(t)]^2}. \quad (4)$$

Для нахождения δa_k нужно минимизировать d .

После того как будут найдены δa_k , точнее говоря, их дисперсии $D \delta a_k$ нужно сравнить их с изменениями параметров, получаемыми в результате полезной модуляции. Обозначив наибольшие полезные изменения через Δa_k , можно записать для отношения сигнал/помеха на выходе приемника соответствующего вида модуляции

$$\rho_k = (\Delta a_k)^2 / D \delta a_k. \quad (5)$$

Геометрическая интерпретация. Представим переносчик (1) вектором в пространстве параметров P , т. е. в пространстве, в котором по координатным осям откладываются значения параметров. При отсутствии модуляции конец вектора f занимает в P вполне определенное положение. Модуляция по какому-либо параметру вызывает перемещение конца вектора f вдоль отрезка прямой, параллельной соответствующей оси. Длина этого отрезка



Р и с. 2

равна $2\Delta a_k$. При наложении помехи вектора f случайным образом смещается во всех направлениях.

Нахождение функции F_1 (3) есть не что иное, как отображение функции F (2) на пространство параметров P . Правило отображения состоит в минимизации квадратичного уклонения (4).

Существенно, что пространство P анизотропно. Это значит, что данная помеха вызывает различные приращения различных параметров. Поэтому можно поставить важный для техники вопрос о наиболее помехоустойчивом виде модуляции, т. е. о выборе из числа параметров переносчика такого, для которого ρ_k (5) имеет наибольшее значение. Трехмерная модель пространства P пред-

ставлена на рис. 2, где в направлении соответствующих осей отложены паразитные приращения параметров.

Вывод общих соотношений. Запишем (4) в виде

$$d = \frac{1}{T} \int_0^T [F_1(t) - F(t)]^2 dt = \frac{1}{T} \int_0^T [\delta f(t) - \xi(t)]^2 dt. \quad (6)$$

Здесь $\delta f(t) = F_1(t) - f(t)$.

Если $f(a_1, a_2, \dots, t)$ имеет непрерывные частные производные по параметрам a_1, a_2, \dots , то в случае малой помехи имеем

$$\delta f(t) = \sum_{k=1}^n \frac{\partial f}{\partial a_k} \delta a_k. \quad (7)$$

Интервал усреднения T берется равным интервалу корреляции полезного сигнала, т. е.

$$T = 1/2F,$$

где F — ширина спектра модулирующей функции. На протяжении этого интервала полезный сигнал, а следовательно, и пропорциональное ему полезное приращение параметра могут считаться постоянными, чем и определяется выбор полос пропускания и постоянных времени в цепях приемника.

Будем теперь минимизировать d , для чего составим частные производные по паразитным приращениям и приравняем их нулю

$$\begin{aligned} \frac{\partial d}{\partial a_k} &= \frac{1}{T} \int_0^T \frac{\partial}{\partial a_k} \left(\sum_i \frac{\partial f}{\partial a_i} \delta a_i - \xi \right)^2 dt = \\ &= \frac{2}{T} \int_0^T \left(\sum_i \frac{\partial f}{\partial a_i} \cdot \frac{\partial f}{\partial a_k} \delta a_i - \frac{\partial f}{\partial a_k} \xi \right) dt = 0. \end{aligned}$$

Отсюда

$$\frac{1}{T} \int_0^T \sum_i \frac{\partial f}{\partial a_i} \frac{\partial f}{\partial a_k} \delta a_i dt = \frac{1}{T} \int_0^T \frac{\partial f}{\partial a_k} \xi dt$$

или

$$\sum_{i=1}^n b_{ik} \delta a_i = \frac{1}{T} \int_0^T \frac{\partial f}{\partial a_k} \xi dt, \quad (8)$$

где

$$b_{ik} = \frac{1}{T} \int_0^T \frac{\partial f}{\partial a_i} \frac{\partial f}{\partial a_k} dt. \quad (9)$$

Разрешая систему (8) относительно искомых δa_i , находим

$$\delta a_i = \frac{1}{T} \int_0^T y_i(t) \xi(t) dt \quad (10)$$

и, следовательно, математическое ожидание δa_i равно нулю, если $M\xi(t) = 0$. Дисперсия этих величин равна

$$D\delta a_i = M(\delta a_i)^2 = M \frac{1}{T^2} \int_0^T \int_0^T y_i(t) y_i(t_1) \xi(t) \xi(t_1) dt dt_1$$

или

$$D\delta a_i = \frac{1}{T^2} \int_0^T \int_0^T y_i(t) y_i(t_1) B_\xi(t - t_1) dt dt_1, \quad (11)$$

где $B_\xi(\tau) = M[\xi(t)\xi(t + \tau)]$ — функция корреляции помехи,

$$y_i(t) = \frac{1}{D} \sum_{k=1}^n D_{i,k} \frac{\partial f}{\partial a_k}, \quad (12)$$

D — определитель матрицы $(b_{i,k})$; $D_{i,k}$ — соответствующие алгебраические дополнения.

Формула (11) неудобна для вычисления, так как обычно задается полоса пропускания, и интегрирование желательно вести в частотной области. Вводя

$$B(\tau) = \frac{1}{2} \int_{-\infty}^{\infty} G(\omega) e^{i\omega\tau} d\omega,$$

где $G(\omega)$ — спектр мощности, и, пользуясь определением спектра,

$$S_y(\omega) = \int_{-\infty}^{\infty} y(t) e^{-i\omega t} dt,$$

получим вместо (11) расчетную формулу

$$D\delta a_i = \frac{1}{2T^2} \int_{-\infty}^{\infty} G_\xi(\omega) \Phi_{y_i}^2(\omega) d\omega, \quad (13)$$

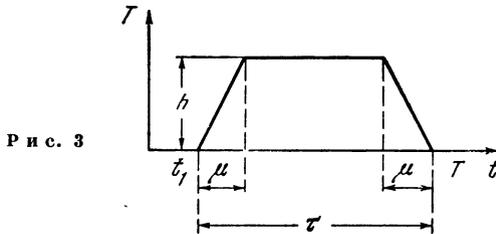
где $\Phi^2(\omega) = S(\omega)S^*(\omega)$ (звездочка означает комплексно-сопряженную величину). Функция $\Phi_{y_i}^2(\omega)$ выражает спектр энергии функции $y(t)$.

Пример. Не останавливаясь на синусоидальном переносчике, возьмем в качестве примера менее банальный случай, а именно периодическую последовательность трапецеидальных импульсов,

характеризующихся тремя параметрами; высотой h , моментом начала (фазой) t_1 и длительностью τ . График функции

$$f = f(h, t_1, \tau, t)$$

на протяжении одного периода T показан на рис. 3.



Р и с. 3

Длительность фронтов, обозначенная через μ , считается неизменным параметром. Аналитическую запись функции f можно представить в виде

$$f = \frac{h}{\mu} [(t - t_1) \sigma(t - t_1) - (t - t_1 - \mu) \sigma(t - t_1 - \mu) - (t - t_1 - \tau + \mu) \sigma(t - t_1 - \tau + \mu) + (t - t_1 - \tau) \sigma(t - t_1 - \tau)].$$

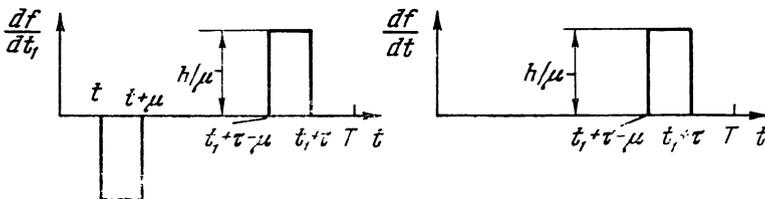
Дифференцируя это выражение, находим все три частных производные по параметрам

$$\partial f / \partial h = \frac{1}{h} f;$$

$$\partial f / \partial t_1 = \frac{h}{\mu} [-\sigma(t - t_1) + \sigma(t - t_1 - \mu) + \sigma(t - t_1 - \tau + \mu) - \sigma(t - t_1 - \tau)];$$

$$\partial f / \partial \tau = \frac{h}{\mu} [\sigma(t - t_1 - \tau + \mu) - \sigma(t - t_1 - \tau)].$$

Графики двух последних производных представлены на рис. 4.



Р и с. 4

Определяем b_{ik} по формуле (9). Матрица (b_{ik}) имеет вид

$$(b_{ik}) = \begin{pmatrix} \frac{\mu}{T} \left(\frac{\tau}{\mu} - \frac{4}{3} \right) & 0 & \frac{h}{2T}, \\ 0 & 2 \frac{h^2}{T\mu} & \frac{h^2}{T\mu}, \\ \frac{h}{2T} & \frac{h^2}{T\mu} & \frac{h^2}{T\mu}. \end{pmatrix}$$

Отсюда находим

$$D = \frac{h^4}{T^2\mu} \left(\frac{\tau}{\mu} - \frac{11}{6} \right)$$

и алгебраические дополнения D_{ik}

$$(D_{ik}) = \begin{array}{ccc} \frac{h^4}{T^2\mu^2} & \frac{h^3}{2T^2\mu} & \frac{-h^3}{T^2\mu} \\ \frac{h^3}{2T^2\mu} & \frac{h^2}{T^2} \left(\frac{\tau}{\mu} - \frac{19}{12} \right) & -\frac{h^2}{T^2} \left(\frac{\tau}{\mu} - \frac{4}{3} \right) \\ -\frac{h^3}{T^2\mu} & -\frac{h^2}{T^2} \left(\frac{\tau}{\mu} - \frac{4}{3} \right) & \frac{2h^2}{T^2} \left(\frac{\tau}{\mu} - \frac{4}{3} \right). \end{array}$$

Будем полагать, что $\alpha = \tau/\mu \gg 1$, т. е. длительность фронтов очень мала по сравнению с длительностью импульса, и в дальнейшем будем делать вытекающие из этого упрощения. Теперь можно записать для функций $y_i(t)$:

$$y_1(t) = \frac{T}{\tau} \left(f_I + \frac{1}{2} f_{II} - f_{III} \right),$$

$$y_2(t) = \frac{T}{h} \left(\frac{\mu}{2\tau} f_I + f_{II} - f_{III} \right),$$

$$y_3(t) = -\frac{T}{h} \left(\frac{\mu}{\tau} f_I + f_{II} - 2f_{III} \right),$$

где

$$f_I = \frac{1}{h} f; \quad f_{II} = \frac{\mu}{h} \cdot \frac{\partial f}{\partial t_1}, \quad f_{III} = \frac{\mu}{h} \cdot \frac{\partial f}{\partial \tau}.$$

Для спектров этих трех нормированных функций (с точностью до общего комплексного множителя, равного по модулю единице), имеем соответственно

$$S_I = \frac{2}{\omega^2\mu} \left[\cos \omega \left(\frac{\tau}{2} - \mu \right) - \cos \omega \frac{\tau}{2} \right] \approx \frac{2}{\omega} \sin \omega \frac{\tau}{2},$$

$$S_{II} = \frac{2}{i\omega} \left[\cos \omega \left(\frac{\tau}{2} - \mu \right) - \cos \omega \frac{\tau}{2} \right] \approx -i2\mu \sin \omega \frac{\tau}{2},$$

$$S_{III} = \frac{1}{i\omega} \left[\cos \omega \left(\frac{\tau}{2} - \mu \right) - \cos \omega \frac{\tau}{2} \right] - \frac{1}{\omega} \left[\sin \omega \left(\frac{\tau}{2} - \mu \right) - \sin \omega \frac{\tau}{2} \right] \approx \mu \left(\cos \omega \frac{\tau}{2} - i \sin \omega \frac{\tau}{2} \right).$$

Пусть помеха представляет собой белый шум со спектральной плотностью G_0 , постоянной в пределах полосы пропускания $0 < \omega < \Omega$. Тогда имеем вместо (13)

$$D\delta a_i = \frac{G_0}{T^2} \int_0^{\Omega} \Phi_y^2(\omega) d\omega.$$

Опустим подробности дальнейших вычислений, в которых для упрощения, кроме неравенства $\mu/\tau \ll 1$, использовано еще условие $\Omega\tau \gg 1$ (так как полоса пропускания Ω выбирается обычно так, что $\Omega\mu \approx 0,4$, откуда и следует $\Omega\tau = \Omega\mu\tau/\mu \approx 0,4\tau/\mu \gg 1$). В конце концов получаем

$$\begin{aligned} D\delta h &= N_0/2\tau_0, & D\delta t_1 &= 0,064N_0\mu/h_0^2, \\ D\delta\tau &= 0,128N_0\mu/h_0^2, \end{aligned}$$

где $N_0 = 2\pi G_0$ — спектральная плотность мощности, отнесенная к 1 гц, нуликами отмечены начальные (немодулированные) значения параметров. Остается ввести полезные приращения. Для АИМ имеем $\Delta h = h_0$ (т. е. амплитуда h может меняться от нуля до удвоенного значения). Для ФИМ t_1 может меняться от нуля до $T/2$ в обе стороны, если $t_{10} = T/2$ и если скважность

$$\beta = T/\tau \gg 1.$$

При этих условиях $\Delta t_1 = T/2$. Наконец, в случае ДИМ, выбирая $\tau_0 = T/2$ (т. е. $\beta_0 = 2$), имеем $\Delta\tau = T/2$. Пользуясь этими значениями, находим по формуле (5) отношение сигнал/помеха для амплитудно-импульсной модуляции

$$\rho_A = (\Delta h)^2/D\delta h = 2h_0^2\tau_0/N_0 = 2E/N_0,$$

где $E = h_0^2\tau$ — энергия импульса. Для ФИМ имеем

$$\rho_\Phi \approx 3,76T^2h_0^2/N_0\mu = 3,76\beta_0^2\alpha_0 E/N_0,$$

где β_0 — начальная скважность; $\alpha_0 = \tau_0/\mu$ — относительная крутизна фронтов. Наконец, для ДИМ

$$\rho_D \approx 1,88T^2h_0^2/N_0\mu = 1,88\beta_0^2\alpha_0 E/N_0 = 7,52\alpha_0 E/N_0.$$

Итак, для отношения сигнал/помеха для трех рассматриваемых видов модуляции можно записать

$$\rho_A : \rho_\Phi : \rho_D = 1 : 1,88\beta_0^2\alpha_0 : 3,76\alpha_0,$$

откуда легко сделать заключение о значительных преимуществах ФИМ.

Помехоустойчивость ФИМ и ДИМ возрастает с увеличением относительной крутизны $\alpha_0 = \tau_0/\mu$, т. е. с увеличением полосы пропускания. Для ФИМ, оказывается, большую роль играет и начальная скважность β_0 .

Возможность компенсации паразитной модуляции. Паразитные приращения δa_i параметров переносчика, будучи вызваны одной и той же причиной, оказываются коррелированными между собой. Это обстоятельство указывает на принципиальную возможность компенсации паразитной модуляции. Идея состоит в том, что паразитные приращения немодулированных (т. е. не используемых при данном виде модуляции) параметров применяются

для устранения (или уменьшения) паразитных приращений модулируемых параметров. В дальнейшем будем называть модулируемые параметры информационными, а немодулированные — контрольными. Пусть имеется m -информационных параметров a_1, a_2, \dots, a_m . Остальные $n-m$ параметров a_{m+1}, \dots, a_n немодулируются и могут быть использованы в качестве контрольных. Информационный параметр a_i получает вследствие наложения помехи паразитное приращение δa_i .

Образует линейную комбинацию паразитных приращений контрольных параметров

$$\widetilde{\delta a}_i = \sum_{s=1}^{n-m} K_{is} \delta a_{m+s}$$

и составим разность

$$\delta a_i - \widetilde{\delta a}_i = \varepsilon_i.$$

Если бы $\widetilde{\delta a}_i$ в точности равнялось δa_i , то возможна была бы полная компенсация паразитной модуляции. Но так как мы имеем дело со случайными величинами, то лучшее, что мы можем сделать, это минимизировать дисперсию разности ε_i . Мы имеем

$$D\varepsilon_i = M(\delta a_i - \widetilde{\delta a}_i)^2 = M\left(\delta a_i - \sum_{s=1}^{n-m} K_{is} \delta a_{m+s}\right)^2.$$

Для коэффициентов K_{is} , минимизирующих эту величину, получаем систему уравнений

$$\sum_{\nu=1}^{n-m} K_{i\nu} \sqrt{D\delta a_{m+s} D\delta a_{m+\nu}} r_{m+s, m+\nu} = \sqrt{D\delta a_i D\delta a_{m+s}} r_{i, m+s},$$

$$(s = 1, 2, \dots, n-m),$$

где $r_{ij} = M(\delta a_i \delta a_j) / \sqrt{D\delta a_i D\delta a_j}$ — нормированный коэффициент корреляции случайных величин δa_i и δa_j .

Эти соотношения сильно упрощаются, когда выбирается только один контрольный параметр a_n . В этом случае $\widetilde{\delta a}_i = K_i \delta a_n$ и $K_i = \sqrt{D\delta a_i / D\delta a_n} r_{in}$ ($i = 1, 2, \dots, n-1$).

Такой же результат получается и в том случае, когда двумерное распределение случайных величин δa_i и δa_n является нормальным и в качестве δa_i выбирается наивероятнейшее значение при условии, что δa_n известно. Компенсация тем лучше, чем больше по абсолютному значению коэффициенты корреляции r_{in} .

Для вычисления коэффициентов корреляции служит формула

$$M(\delta a_i \delta a_j) = \frac{1}{T^2} \int_0^T \int_0^T y_i(t) y_j(t_1) B_\xi(t - t_1) dt dt_1 \quad (14)$$

или, в спектральном представлении,

$$M(\delta a_i \delta a_j) = \frac{1}{2T^2} \int_{-\infty}^{\infty} G_{\xi}(\omega) S_{y_i}(\omega) S_{y_j}^*(\omega) d\omega \quad (15)$$

(эти две формулы являются обобщением формул (11) и (13)).

Вычислим коэффициенты корреляции для приращения параметров импульсной модуляции из примера предыдущего параграфа, т. е. для величин δh , δt_1 и $\delta \tau$. Так как в полосе Ω $G_{\xi}(\omega) = G_0 = \text{const}$, то вместо (15) имеем

$$M(\delta a_i \delta a_j) = \frac{G_0}{T^2} \int_0^{\Omega} \text{Re} \{ S_{y_i}(\omega) S_{y_j}^*(\omega) \} d\omega$$

и вычисление (использующее те же приближения, что и выше) дает

$$M(\delta h \delta t_1) = 0,531 \frac{\mu}{\tau} \cdot \frac{N}{h}, \quad M(\delta t_1 \delta \tau) = 0,064 \frac{\mu N}{h^2},$$

$$M(\delta \tau \delta h) = -\frac{\mu}{\tau} \cdot \frac{N}{h}.$$

Нормировка дает следующие значения коэффициентов корреляции:

$$r_{ht_1} = 2,96 \sqrt{\mu/\tau}, \quad r_{t_1} = -0,707, \quad r_{\tau h} = -3,96 \sqrt{\mu/\tau}.$$

Величины r_{ht_1} и $r_{\tau h}$ зависят от относительной крутизны фронтов импульса τ/μ . Это значит, что в пределе при переходе к прямоугольным импульсам (что соответствует неограниченному расширению полосы) корреляция между паразитной модуляцией по амплитуде и по остальным двум параметрам разрушается. Это объясняется тем, что при $\mu \rightarrow 0$ паразитные приращения δt_1 и $\delta \tau$ также стремятся к нулю, как видно из формул предыдущего параграфа.

Вместе с тем между δt_1 и $\delta \tau$ существует постоянная (не зависящая от μ) и довольно сильная корреляция. Это означает возможность эффективной компенсации паразитной модуляции параметров t_1 и τ при использовании одного из них в качестве информационного, а другого — в качестве контрольного.

Мы обсуждали возможность компенсации помехи при наличии значительной корреляции между паразитными приращениями параметров. Нужно еще заметить, что если такая корреляция отсутствует или очень мала, то и в этом случае можно уменьшить влияние помехи, но на основе совершенно иного принципа. Если модулировать одним и тем же сигналом несколько разных параметров при условии, что паразитные приращения этих параметров статистически независимы, то дело сводится к образованию нескольких независимых каналов. В этом случае можно прибегнуть к методу накопления, складывая выходные сигналы всех

каналов. Отношение сигнал/помеха при этом возрастает, как известно, в N раз, где N — число каналов (т. е. число модулируемых параметров). Наличие корреляции уменьшает выигрыш, получаемый за счет накопления, но зато открывает возможность вышеописанной компенсации.

Замечания по поводу случая большой помехи. Мы ограничились рассмотрением малой помехи, так как в этом случае возможна линеаризация задачи, выражаемая формулой (7). Отбрасывание условия малости помехи приводит в общем случае к чрезвычайному усложнению исследования, так как вместо (7) мы должны были бы исходить из точного выражения

$$\delta f(t) = f(a_1 + \delta a_1, a_2 + \delta a_2, \dots, a_n + \delta a_n, t) - \\ - f(a_1, a_2, \dots, a_n, t).$$

В этих условиях, вероятно, рентабельнее получить интересующие нас зависимости экспериментально или путем использования вычислительных машин.

Это замечание относится также и к случаю малой помехи, когда некоторые из $\partial f/\partial a_k$ имеют бесконечные разрывы, что имеет место, например, для сигналов в виде прямоугольных импульсов.

СОДЕРЖАНИЕ

О детектировании биений	5
О специальной частотной характеристике для оценки искаженной формы кривой	13
О применении критерия квадратичной погрешности к оценке линейных искажений	22
Об экспериментальном подборе параметров сложных систем	39
О коррекции искажений, вносимых оптической щелью	53
О диаграммах сила—перемещение в применении к исследованию колебаний	65
Вынужденные колебания системы, ударяющейся об ограничитель	
О расчете одного рода корректирующих схем	78
Спектры и анализ	87
Глава I. Спектры	87
Глава II. Анализ	149
Глава III. Спектры случайных процессов	198
Добавления	236
Литература	251
Автоколебания	253
Предисловие	253
Введение	254
§ 1. Автоколебательная система	254
§ 2. Энергетика автоколебаний	257
§ 3. Диаграммы работы	260
§ 4. Статическая двузначность	265
§ 5. Динамическая двузначность	267
§ 6. Поршневые двигатели	271
§ 7. Пневматические ударные инструменты	277
§ 8. Фрикционный маятник	280
	565

§ 9. Автоколебания резца	283
§ 10. Гармонный язычок,	287
§ 11. Часы	290
§ 12. Роль фазы обратной связи	295
§ 13. Падающая характеристика и отрицательное сопротивление	297
§ 14. Анализ фаз	301
§ 15. Релаксационные колебания	306
§ 16. Фрикционные автоколебания	311
§ 17. Примеры гидравлических систем	315
§ 18. Релаксационные и почти гармонические колебания	318
§ 19. Термопрерыватель	321
§ 20. Автоколебания в трубах	325
§ 21. Автоколебания, связанные с образованием вихрей	334
§ 22. Флаттер	338
§ 23. Шимми	343
§ 24. Параметрическое возбуждение	347
§ 25. Одна механическая модель	354
§ 26. Электронные генераторы сверхвысоких частот	356
§ 27. Заключение	362
Добавления	365
Литература	378
Умножающие схемы	380
Нелинейные и параметрические явления в радиотехнике	388
Введение	388
Глава I. Нелинейные цепи и основные нелинейные процессы	394
Глава II. Генерирование колебаний	431
Глава III. Нелинейные системы под внешним воздействием	482
Глава IV. Параметрические явления	502
Передача сигналов модулированным шумом	524
О вычислении спектров случайных процессов	530
О возможностях сжатия спектра	538
О теореме Котельникова	546
Паразитная модуляция, вызываемая малой аддитивной помехой	555

**Александр Александрович
Харкевич**

**ИЗБРАННЫЕ ТРУДЫ
В ТРЕХ ТОМАХ**

Том 2

**ЛИНЕЙНЫЕ
И НЕЛИНЕЙНЫЕ СИСТЕМЫ**

**Утверждено к печати
Институтом проблем
передачи информации**

**Редактор издательства
Н. Н. Соколова
Художник
В. Дмитрияди
Технический редактор
Р. М. Денисова**

Сдано в набор 4/Х 1972 г. Подписано к печати 7/II 1973 г. Формат
60×90¹/₁₆. Бумага № 2. Усл. печ. л. 35,5. Уч.-изд. л. 32,5. Т-02726.
Тираж 4100 экз. Тип. зак. 1460. Цена 2 р. 48 к.

Издательство «Наука», 103717 ГСП, Москва К-62, Подсосенский пер. 21
1-я типография издательства «Наука» 199034, Ленинград, 9-я линия, 12

1874



PUBLISHED BY