

А.М.Кугушев, Н.С.Голубева • ОСНОВЫ РАДИОЭЛЕКТРОНИКИ

А.М.Кугушев, Н.С.Голубева

ОСНОВЫ РАДИО- ЭЛЕКТРОНИКИ



А. М. КУГУШЕВ, Н. С. ГОЛУБЕВА

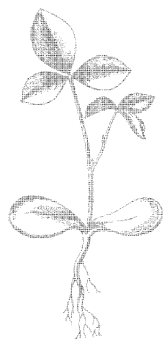
ОСНОВЫ РАДИОЭЛЕКТРОНИКИ

НЕЛИНЕЙНЫЕ ЭЛЕКТРОМАГНИТНЫЕ ПРОЦЕССЫ

*Допущено Министерством высшего
и среднего специального образования СССР
в качестве учебного пособия для студентов
радиотехнических специальностей вузов*



МОСКВА «ЭНЕРГИЯ» 1977



6Ф2
К88
УДК 621.37/.39

Рецензенты:

Ленинградский политехнический институт
им. М. И. Калинина (кафедра радиофизики),
Московский инженерно-физический институт (кафедра физики твердого тела).

Кугушев А. М. и Голубева Н. С.

К88 Основы радиоэлектроники (Нелинейные электромагнитные процессы). Учеб. пособие для студентов радиотехнических специальностей вузов. М., «Энергия», 1977.

400 с. с ил.

В книге изложены основы теории нелинейных электромагнитных процессов в средах (нелинейная электродинамика). Процессы описываются на основе уравнений электромагнитного поля и квантовой механики. Даются принципы работы квантовых генераторов и усилителей оптического диапазона и диапазона СВЧ (парамагнитных и молекулярных), а также электронных генераторов и усилителей СВЧ диапазона. В дополнительных главах приводится краткое изложение основ квантовой механики и методов решения нелинейных уравнений. Учебное пособие также может быть полезно аспирантам и инженерам.

К $\frac{30401-331}{051(01)-77}$ 178-77

6Ф2

© Издательство «Энергия», 1977.

ПРЕДИСЛОВИЕ

Современная радиоэлектроника — область всестороннего практического использования учения об электрических явлениях.

Технические средства радиоэлектроники являются основой не только всех видов электро- и радиосвязи, всего многообразия кибернетических устройств, автоматических систем управления и вычислительной техники, но и электроэнергетики.

Такое широкое использование электромагнитной энергии объясняется простотой преобразования ее из одного вида в другой, простотой передачи практически на любые расстояния и возможностью концентрирования с огромной плотностью.

В связи с этим дисциплина «Основы радиоэлектроники» должна быть основным теоретическим курсом в учебных планах подготовки инженеров по радиотехнике и электронной технике, а также инженеров смежных областей. В методическом отношении целесообразно разделить этот курс на отдельные части: линейные и нелинейные электромагнитные процессы.

Объем знаний, необходимых будущему специалисту, непрерывно возрастает. Поэтому необходимо избавляться от повторения одного и того же материала в различных курсах. Для этого изложение должно вестись главным образом дедуктивно, что позволяет глубже понимать процессы и значительно сократить объем изложения и отводимое время на изучение всей дисциплины.

Изучению линейных электромагнитных процессов посвящена ранее изданная книга авторов «Основы радиоэлектроники» (линейные электромагнитные процессы). М., «Энергия», 1969 г.

Настоящая книга содержит изложение нелинейных электромагнитных процессов в средах (нелинейная электродинамика). На основе этого изучаются преоб-

разование частоты, генерирование и усиление электромагнитных колебаний оптического и сверхвысокочастотного (СВЧ) диапазонов. Большое развитие в последние годы получила квантовая радиоэлектроника этих диапазонов. Однако учебной литературы для технических вузов, посвященной этим разделам, сравнительно мало. Это вызывает потребность написать книгу, дающую углубленное изложение фундаментальных основ нелинейных электромагнитных процессов в средах.

В книге дается анализ физических явлений в нелинейных средах и описание квантовых и СВЧ приборов с единой общей точки зрения.

Изложение нелинейных электромагнитных процессов основывается на нелинейных уравнениях электромагнитного поля. Нелинейность в эти уравнения вводится через параметры среды. Преобразование частоты, в частности возбуждение гармоник и комбинационных частот, самофокусировка и другие явления, при которых полная энергия электромагнитного поля не возрастает, рассматриваются как результат взаимодействия электромагнитного поля с пассивной средой.

Генерирование и усиление электромагнитных колебаний рассматриваются как взаимодействие электромагнитного поля с активной средой, в результате которого энергия электромагнитного поля увеличивается. В квантовых генераторах и усилителях оптического диапазона и СВЧ диапазона (в парамагнитных усилителях и молекулярных генераторах) это усиление происходит за счет внутренней энергии атомов и молекул. В вакуумных электронных генераторах и усилителях это усиление происходит за счет кинетической или потенциальной энергии электронов. При этом в обоих случаях первичным источником энергии, как правило, является источник постоянного тока. Только в последнее время начинает применяться получение электромагнитной энергии в виде почти когерентных колебаний оптического и коротковолнового участка СВЧ диапазонов путем непосредственного преобразования химической и тепловой энергии.

В книге не рассматриваются нелинейные электромагнитные процессы в цепях, к которым относятся микроэлектроника, включающая полупроводниковые дискретные и интегральные схемы, процессы в параметрических усилителях, а также процессы в обычных

нелинейных цепях, характеристиками которых являются интегральные величины — ток и напряжение. В настоящее время эти разделы радиоэлектроники настолько разрослись, что объединение их в одной книге с ограниченным объемом неосуществимо и в методическом отношении нецелесообразно.

Применяемые методы решения нелинейных уравнений и основные понятия квантовой механики изложены соответственно в Математическом и Физическом дополнениях, что дает возможность не нарушать логику и последовательность изложения основного материала. Эти дополнения приведены в конце книги. Ссылки на эти дополнения в тексте идут с буквами П1 (физическое) и П2 (математическое).

Книга соответствует учебному плану радиоэлектронных специальностей Московского высшего технического училища им. Н. Э. Баумана (МВТУ).

Авторы выражают свою благодарность профессорам Ю. А. Быковскому, М. И. Конторовичу и В. Ю. Петрунькину, доцентам Л. Н. Пахомову, В. А. Каратыгину и В. А. Розову, которые сделали много полезных замечаний при рецензировании рукописи книги.

Особенно благодарны авторы редактору книги доценту Московского энергетического института В. Г. Карташеву, который внимательно прочел и отредактировал рукопись, дал ряд ценных советов и указаний, что помогло улучшить содержание и стиль изложения книги.

Авторы будут признательны всем читателям за критические замечания по улучшению содержания книги, которые следует направлять по адресу: 113114. Москва, Шлюзовая наб., 10, издательство «Энергия».

Авторы

ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

- A_{nt} — коэффициент Эйнштейна для спонтанного излучения;
 a — обобщенная расстройка;
 \mathbf{B} — вектор магнитной индукции;
 $B_{вп}$ — вероятность перехода электрона из валентной зоны в зону проводимости;
 $B_{пв}$ — вероятность перехода электрона из зоны проводимости в валентную зону;
 B_{mn} — коэффициент Эйнштейна для индуцированного поглощения;
 B_n — нормальная составляющая вектора \mathbf{B} ;
 B_{nt} — коэффициент Эйнштейна для индуцированного излучения;
 \mathbf{B}_0 — вектор постоянной магнитной индукции;
 c — скорость света;
 \mathbf{D} — вектор электрической индукции;
 D_n — нормальная составляющая вектора \mathbf{D} ;
 e — заряд электрона;
 \mathbf{E} — вектор напряженности электрического поля;
 E_m — амплитуда напряженности электрического поля;
 \dot{E}_m — комплексная амплитуда напряженности электрического поля;
 $\dot{E}_m(n\omega)$ — комплексная амплитуда напряженности n -й гармоники электрического поля;
 \mathbf{E}_0 — вектор напряженности электрического поля отраженной волны;
 \mathbf{E}_0 — вектор напряженности постоянного электрического поля;
 \mathbf{F} — вектор силы;
 f — частота;
 $f_{дв}$ — вероятность наличия дырки в валентной зоне;

- $f_{\text{дп}}$ — вероятность наличия дырки в зоне проводимости;
 $f_{\text{эв}}$ — вероятность нахождения электрона в валентной зоне;
 $f_{\text{эп}}$ — вероятность нахождения электрона в зоне проводимости;
 G — коэффициент усиления;
 g — коэффициент вырождения;
 $g(\omega)$ — форм-фактор спектральной линии;
 \mathbf{H} — вектор напряженности магнитного поля;
 H_m — амплитуда напряженности магнитного поля;
 \dot{H}_m — комплексная амплитуда напряженности магнитного поля;
 $\dot{H}_m(n\omega)$ — комплексная амплитуда напряженности n -й гармоники магнитного поля;
 \mathbf{H}_0 — вектор напряженности постоянного магнитного поля;
 h, \hbar — постоянная Планка ($\hbar = h/2\pi$);
 $I(\mathbf{r}, t)$ — мгновенная интенсивность;
 $\langle I(\mathbf{r}, t) \rangle$ — средняя интенсивность;
 J — интенсивность, определяемая плотностью потока мощности излучения;
 \tilde{J} — нормированная интенсивность, определяемая плотностью потока мощности излучения;
 \mathbf{J} — вектор плотности тока;
 J_m — амплитуда плотности тока;
 \dot{J}_m — комплексная амплитуда плотности тока;
 $\dot{J}_m(n\omega)$ — комплексная амплитуда n -й гармоники плотности тока;
 $\mathbf{J}^{\text{ст}}$ — вектор плотности стороннего тока;
 $J_m^{\text{ст}}$ — амплитуда плотности стороннего тока;
 $\dot{J}_m^{\text{ст}}$ — комплексная амплитуда плотности стороннего тока;
 $\mathbf{J}_{\text{пол}}$ — вектор плотности тока поляризации;
 $\mathbf{J}_{\text{нам}}$ — вектор плотности тока намагниченности;
 k — постоянная Больцмана;
 \mathbf{k} — комплексный вектор постоянной распространения;
 $\dot{k}(n\omega)$ — комплексная постоянная распространения n -й гармоники;

- $\dot{k}_J(n\omega)$ — комплексная постоянная распространения n -й гармоники волны тока;
 $\dot{k}_M(n\omega)$ — комплексная постоянная распространения n -й гармоники волны намагниченности;
 $K_{ш}$ — коэффициент шума;
 $k_э$ — электронное волновое число;
 \dot{M} — вектор намагниченности;
 \dot{M}_m — комплексная амплитуда намагниченности;
 $\dot{M}_m(n\omega)$ — комплексная амплитуда n -й гармоники намагниченности;
 n — коэффициент преломления;
 N_n — населенность n -го уровня;
 N_ϕ — число фотонов;
 \mathbf{P} — вектор поляризации;
 P_m — амплитуда поляризации;
 \dot{P}_m — комплексная амплитуда поляризации;
 P — комплексная мощность;
 P_0 — действительная мощность;
 $\dot{P}_m(n\omega)$ — комплексная амплитуда n -й гармоники поляризации;
 $p_{0 \text{ нам}}$ — средняя плотность мощности, связанная с намагниченностью;
 $p_{0 \text{ пол}}$ — средняя плотность мощности, связанная с поляризацией;
 $p_{0 \text{ пров}}$ — средняя плотность мощности, связанная с проводимостью;
 $p_{0 \text{ погл}}^{\text{инд}}$ — средняя плотность мощности индуцированного поглощения;
 $p_{0 \text{ изл}}^{\text{инд}}$ — средняя плотность мощности индуцированного излучения;
 P_{nm} — вероятность индуцированного излучения;
 p_e — электрический дипольный момент частицы;
 Q — добротность;
 R — радиус;
 $R_{св}$ — сопротивление связи;
 T — абсолютная температура;
 t — время;
 $T_{ус}$ — эквивалентная шумовая температура усилителя;

T_1 — время «продольной» релаксации;
 T_2 — время «поперечной» релаксации;
 U_0 — постоянное напряжение;
 u — переменное напряжение;
 U_m — амплитуда напряжения;
 \mathbf{v} — вектор скорости;
 v_0 — постоянная составляющая скорости;
 v_{\sim} — переменная составляющая скорости;
 v_{ϕ} — фазовая скорость;
 v_m — комплексная амплитуда скорости;
 $v(n\omega)$ — скорость распространения n -й гармоники;
 W — энергия;
 w — плотность энергии;
 w_0 — средняя плотность энергии;
 $w_0(\omega_{nm})$ — средняя плотность энергии на частоте перехода ω_{nm} ;
 W_d — энергия дырки;
 W_e — энергия электрона;
 W_F — уровень Ферми;
 $W_{F\epsilon}$ — квазиуровень Ферми для электронов;
 W_{Fd} — квазиуровень Ферми для дырок;
 $W_{кин}$ — кинетическая энергия;
 $W_{пот}$ — потенциальная энергия;
 Z_0 — волновое сопротивление;
 $\alpha_{ус}$ — коэффициент усиления;
 α_n — постоянная затухания;
 β — фазовая постоянная;
 γ — гиромагнитное отношение;
 Γ — коэффициент отражения;
 $\gamma_{12}(\tau)$ — коэффициент когерентности;
 $\Gamma_{12}(\tau)$ — функция взаимной когерентности;
 ϵ — относительная диэлектрическая проницаемость;
 ϵ_0 — электрическая постоянная;
 ϵ_a — абсолютная диэлектрическая проницаемость;
 $\tilde{\epsilon}_a = \epsilon'_a - j\epsilon''_a$ — комплексная абсолютная диэлектрическая проницаемость;
 $\epsilon_a^n(n\omega)$ — линейная абсолютная диэлектрическая проницаемость на частоте $n\omega$;
 η — коэффициент полезного действия;

- θ — угол падения;
- θ_0 — угол отражения;
- ϑ — угол преломления;
- λ — длина волны;
- λ_p — резонансная длина волны;
- μ — относительная магнитная проницаемость;
- μ_0 — магнитная постоянная;
- μ_a — абсолютная магнитная проницаемость;
- $\tilde{\mu}_a = \mu'_a - j\mu''_a$ — комплексная абсолютная магнитная проницаемость;
- $\mu_a^n(n\omega)$ — линейная абсолютная магнитная проницаемость на частоте $n\omega$;
- ρ — объемная плотность заряда;
- $\bar{\rho}$ — то же в комплексной форме;
- $\rho_{св}$ — объемная плотность связанных зарядов;
- σ — удельная проводимость;
- $\sigma(\omega)$ — сечение радиационного перехода;
- $\sigma^L(n\omega)$ — линейная удельная проводимость на частоте $n\omega$;
- τ — время запаздывания;
- $\tau_{кор}$ — время когерентности;
- φ — фаза;
- $\chi_э$ — электрическая восприимчивость;
- χ_m — магнитная восприимчивость;
- ω — круговая частота;
- $\omega_{n,m}$ — частота квантового перехода;
- $\omega_{л}$ — частота линии рабочего перехода;
- ω_p — резонансная частота;
- ΔN — разность населенностей уровней;
- ΔN_0 — начальная разность населенностей уровней;
- $\Delta N_{0\text{ пор}}$ — пороговая начальная разность населенностей уровней;
- $\Delta N_{\text{стац}}$ — стационарная разность населенностей уровней;
- $\tilde{\Delta N}$ — нормированная разность населенностей уровней;
- ΔW — ширина запрещенной зоны;
- $\Delta\omega_p$ — ширина резонансной кривой резонатора;
- $\Delta\omega_{л}$ — ширина линии рабочего перехода;
- $\Delta\omega_{yc}$ — ширина полосы пропускания усилителя;

Общие математические обозначения

- $\mathbf{A}, \bar{\mathbf{A}}$ — вектор;
 A_i ($i=1, 2, 3$) — проекции вектора \mathbf{A} на оси декартовой системы координат;
 A_n — нормальная составляющая вектора \mathbf{A} ;
 A_τ — тангенциальная составляющая вектора \mathbf{A} ;
 $\dot{A} = A_m e^{j\varphi}$ — комплексная величина;
 $\dot{A}^* = A_m e^{-j\varphi}$ — сопряженное значение комплексной величины;
 \mathbf{e}_i ($i=1, 2, 3$) — орты (единичные векторы) декартовой системы координат;
 x_i ($i=1, 2, 3$) — оси декартовой системы координат;
 $J_n(x)$ — функция Бесселя n -го порядка;

Символы математических операций

- $[\mathbf{AB}]$ — векторное произведение;
 \mathbf{AB} — скалярное произведение;
 grad — градиент;
 div — дивергенция;
 rot — ротор;
 Δ — оператор Лапласа;
 \oint_S — интеграл по замкнутой поверхности S ;
 \int_V — интеграл по объему V .

Глава первая

ОСНОВНЫЕ УРАВНЕНИЯ И ХАРАКТЕРИСТИКИ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ И СРЕДЫ

1-1. НЕЛИНЕЙНЫЕ И ПАРАМЕТРИЧЕСКИЕ ЭЛЕКТРОМАГНИТНЫЕ ПРОЦЕССЫ В СРЕДАХ

Электромагнитные процессы в средах описываются уравнениями Максвелла *

$$\left. \begin{aligned} \operatorname{rot} \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}; \\ \operatorname{rot} \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}; \\ \operatorname{div} \mathbf{D} &= \rho; \\ \operatorname{div} \mathbf{B} &= 0 \end{aligned} \right\} \quad (1-1)$$

и уравнениями состояния для изотропной среды

$$\left. \begin{aligned} \mathbf{D} &= \epsilon_a \mathbf{E}; \\ \mathbf{B} &= \mu_a \mathbf{H}; \\ \mathbf{J} &= \sigma (\mathbf{E} + \mathbf{E}^{\text{ст}}). \end{aligned} \right\} \quad (1-2)$$

Здесь \mathbf{E} — напряженность электрического поля, В/м; $\mathbf{E}^{\text{ст}}$ — напряженность стороннего электрического поля; \mathbf{H} — напряженность магнитного поля, А/м; \mathbf{D} — электрическая индукция, Кл/м²; \mathbf{B} — магнитная индукция, Т; \mathbf{J} — плотность тока, А/м²; ρ — объемная плотность заряда, Кл/м³; ϵ_a и μ_a — абсолютные диэлектрическая и магнитная проницаемости, соответственно Ф/м и Г/м; σ — удельная проводимость среды, См/м.

В общем случае анизотропной среды параметры среды являются тензорами второго ранга и уравнения состояния среды имеют вид:

$$D_i = \epsilon_0 \epsilon_{ik} E_k;$$

* Теория электромагнитного поля подробно изложена в [1].

$$B_i = \mu_0 \mu_{ik} H_k;$$

$$J_i = \sigma_{ik} (E_k + E_k^{\text{ст}}).$$

В данной книге мы ограничимся рассмотрением изотропных сред.

Параметры линейной среды не изменяются под действием поля и уравнения состояния среды (1-2) оказываются линейными. Электромагнитный процесс описывается системой (1-1), которая в данном случае с учетом (1-2) представляет собой систему линейных дифференциальных уравнений с постоянными коэффициентами. Для таких уравнений выполняется принцип суперпозиции. Различные частотные составляющие поля распространяются независимо друг от друга. Амплитуды частотных составляющих распространяющегося поля пропорциональны соответствующим составляющим источника. Спектр распространяющегося поля неизменен, в нем нет составляющих, не содержащихся в спектре источника.

Электромагнитные процессы, протекающие в средах, свойства которых зависят от интенсивности электромагнитного поля, называются нелинейными. Нелинейные процессы связаны с нелинейными свойствами сред, которые проявляются в нелинейном взаимодействии вещества с распространяющимся электромагнитным полем. При этом в общем случае диэлектрическая проницаемость среды ϵ_a , магнитная проницаемость μ_a и проводимость σ зависят от величины электромагнитного поля и уравнения состояния среды (1-2) оказываются нелинейными (см. П 2, п. 1).

Нелинейность среды может определяться и инерционными свойствами вещества, проявляющимися в конечной, а не мгновенной скорости изменения параметров среды под действием распространяющегося электромагнитного поля и в зависимости этой скорости от напряженности поля. Реакция среды на воздействующее поле запаздывает на некоторое конечное время τ (время запаздывания), которое зависит от величины поля.

Среда называется нелинейной, если хотя бы один из ее параметров (диэлектрическая проницаемость, магнитная проницаемость, проводимость или скорость реакции среды) зависит от интенсивности действующего электромагнитного поля.

Электромагнитные процессы в нелинейных средах с учетом (1-2) описываются системой нелинейных дифференциальных уравнений (1-1). Принцип суперпозиции для таких уравнений не выполним. Электромагнитные поля, возбужденные различными источниками и распространяющиеся в нелинейной среде, взаимодействуют друг с другом. Изменение параметров среды под влиянием одной из частотных составляющих электромагнитного поля оказывает влияние на распространение других составляющих. Взаимодействие распространяющегося поля со средой приводит к существенному изменению поля. Характер этого изменения зависит от природы и свойств нелинейной среды, характера и интенсивности распространяющегося поля. При распространении в нелинейной среде в спектре электромагнитного поля появляются новые частоты, не содержащиеся в спектре источника. Таким образом, в среде происходит преобразование частоты. Этим нелинейные электромагнитные процессы принципиально отличаются от линейных, при которых возникновение новых частотных составляющих невозможно. В спектре распространяющегося поля появляются не только гармоники составляющих спектра источника, но и комбинационные частоты, соответствующие суммарным и разностным частотам. При этом амплитуды частотных составляющих распространяющегося поля оказываются непропорциональными амплитудам соответствующих частотных составляющих источника.

В оптическом диапазоне волн наиболее выраженными нелинейными свойствами обладают кристаллы дигидрофосфата калия (КДР), дигидрофосфата аммония (АДР), кальцита, кварца, а также жидкости — сероуглерод, нитробензол и т. п. В ферритах, полупроводниках и плазме в диапазоне СВЧ нелинейные эффекты наблюдаются при сравнительно слабых полях. Ферромагнетики и сегнетоэлектрики обладают большой нелинейностью на средних и низких частотах.

Наряду с нелинейными процессами в радиоэлектронике широкое применение нашли параметрические. Если параметры среды не зависят от величины распространяющегося поля, но изменяются во времени по определенному закону с помощью внешних сил (электрических, механических и других), то такая среда называется *параметрической* и явления, происходящие в ней — *параметрическими*. Параметрические процессы могут быть

обусловлены и периодическим изменением запаздывания.

Электромагнитный процесс в параметрической среде описывается системой линейных дифференциальных уравнений (1-1) с коэффициентами, зависящими от времени. Для таких уравнений выполняется принцип суперпозиции и составляющие распространяющегося поля не взаимодействуют друг с другом. При этом также наблюдается преобразование частоты. Спектр распространяющегося поля не зависит от его интенсивности, а определяется лишь спектром источника и законом изменения во времени параметров среды.

Нелинейные и параметрические процессы проявляются как обратное воздействие среды на распространяющееся поле. При распространении электромагнитного поля и в нелинейной, и в параметрической среде меняется спектр частот. Основное различие этих процессов состоит в том, что нелинейные процессы зависят от интенсивности распространяющегося поля, а параметрические не зависят.

Примерами нелинейных и параметрических процессов являются генерирование и усиление электрических колебаний, детектирование, умножение, деление и смешение частот.

В основе генерирования и усиления лежит взаимодействие электромагнитного поля с активной средой. В электронных приборах (триодах, клистронах, магнетронах, лампах бегущей волны и т. п.) энергия постоянного тока преобразуется в энергию высокой частоты в результате взаимодействия движущихся электронов с электромагнитным полем. Усиление или генерирование здесь происходит за счет кинетической энергии электронов, которая получается от источника постоянного тока. В квантовых генераторах и усилителях внутренняя энергия возбужденных атомов, молекул или ионов преобразуется в энергию электромагнитного излучения, а возбуждение частиц осуществляется за счет внешних источников энергии (электрических, тепловых и других видов).

1-2. ОСНОВНЫЕ УРАВНЕНИЯ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ

Электромагнитные процессы в нелинейных средах описываются системой уравнений Максвелла (1-1), которые следует решать совместно с нелинейными уравнениями состояния среды (1-2).

Для нелинейных изотропных сред уравнения состояния среды имеют следующий вид:

$$\left. \begin{aligned} \mathbf{D} &= \epsilon_0 \mathbf{E} + \mathbf{P} = \epsilon_0 \mathbf{E} + \mathbf{P}^{\text{л}} + \mathbf{P}^{\text{нл}} = \epsilon_a^{\text{л}} \mathbf{E} + \mathbf{P}^{\text{нл}}; \\ \mathbf{B} &= \mu_0 (\mathbf{H} + \mathbf{M}) = \mu_0 (\mathbf{H} + \mathbf{M}^{\text{л}} + \mathbf{M}^{\text{нл}}) = \\ &= \mu_a^{\text{л}} \mathbf{H} + \mu_0 \mathbf{M}^{\text{нл}}; \\ \mathbf{J} &= \mathbf{J}^{\text{л}} + \mathbf{J}^{\text{нл}} + \mathbf{J}^{\text{ст}} = \sigma^{\text{л}} \mathbf{E} + \mathbf{J}^{\text{нл}} + \mathbf{J}^{\text{ст}}. \end{aligned} \right\} \quad (1-3)$$

Здесь \mathbf{P} — вектор поляризации, Кл/м²; $\mathbf{P}^{\text{л}}$ и $\mathbf{P}^{\text{нл}}$ — соответственно части вектора \mathbf{P} , линейно и нелинейно зависящие от поля \mathbf{E} ; \mathbf{M} — вектор намагниченности, А/м; $\mathbf{M}^{\text{л}}$ и $\mathbf{M}^{\text{нл}}$ — соответственно линейная и нелинейная части вектора \mathbf{M} ; $\mathbf{J}^{\text{л}}$ и $\mathbf{J}^{\text{нл}}$ — линейная и нелинейная части плотности тока.

Уравнения Максвелла. Подставляя уравнения состояния среды (1-3) в (1-1), получаем систему уравнений Максвелла для нелинейных сред:

$$\left. \begin{aligned} \text{rot } \mathbf{H} &= \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}}{\partial t} + \mathbf{J}; \\ \text{rot } \mathbf{E} &= -\mu_0 \frac{\partial \mathbf{H}}{\partial t} - \mu_0 \frac{\partial \mathbf{M}}{\partial t}, \end{aligned} \right\} \quad (1-4)$$

или

$$\left. \begin{aligned} \text{rot } \mathbf{H} &= \epsilon_a^{\text{л}} \frac{\partial \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}^{\text{нл}}}{\partial t} + \sigma^{\text{л}} \mathbf{E} + \mathbf{J}^{\text{нл}} + \mathbf{J}^{\text{ст}}; \\ \text{rot } \mathbf{E} &= -\mu_a^{\text{л}} \frac{\partial \mathbf{H}}{\partial t} - \mu_0 \frac{\partial \mathbf{M}^{\text{нл}}}{\partial t}. \end{aligned} \right\} \quad (1-5)$$

Волновое уравнение. Для нахождения волнового уравнения напряженности электрического поля возьмем операцию rot от обеих частей второго уравнения системы (1-4):

$$\text{rot rot } \mathbf{E} = -\mu_0 \frac{\partial}{\partial t} \text{rot } \mathbf{H} - \mu_0 \frac{\partial}{\partial t} \text{rot } \mathbf{M}.$$

Подставляя сюда $\text{rot } \mathbf{H}$ из первого уравнения системы (1-4), получаем волновое уравнение

$$\text{rot rot } \mathbf{E} + \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = -\mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2} - \mu_0 \frac{\partial \mathbf{J}}{\partial t} - \mu_0 \frac{\partial}{\partial t} \text{rot } \mathbf{M}. \quad (1-6)$$

Первый и третий член правой части уравнения (1-6) характеризуют дополнительные источники поля в виде

токов поляризации и намагниченности, так как плотности этих токов определяются выражениями

$$\mathbf{J}_{\text{пол}} = \frac{\partial \mathbf{P}}{\partial t};$$

$$\mathbf{J}_{\text{нам}} = \text{rot } \mathbf{M},$$

где $\mathbf{J}_{\text{пол}}$ — плотность тока поляризации; $\mathbf{J}_{\text{нам}}$ — плотность тока намагниченности.

Действительно, ток поляризации определяется смещением связанных зарядов под действием переменного электрического поля.

Электрический момент элементарного диполя

$$\mathbf{p} = q\mathbf{l};$$

электрический момент
единицы объема

$$\mathbf{P} = \sum_{i=1}^n q_i \mathbf{l}_i,$$

где n — число диполей в единице объема.

Очевидно,

$$\frac{\partial \mathbf{P}}{\partial t} = \mathbf{v}_0 \sum_{i=1}^n q_i = \rho_{\text{св}} \mathbf{v}_0 = \mathbf{J}_{\text{пол}},$$

где \mathbf{v}_0 — средняя скорость смещения разноименных зарядов; $\rho_{\text{св}}$ — объемная плотность одноименных зарядов.

Элементарный магнитный диполь можно представить как ток i , протекающий по контуру, ограничивающему элементарную площадку ds . При этом магнитный момент элементарного диполя определится выражением

$$\mathbf{m} = i \, ds$$

или

$$\mathbf{m} = \mathbf{M} \, dV,$$

где dV — элементарный объем.

Если намагниченность среды однородна, то токи на общих границах соседних контуров, текущие в противо-

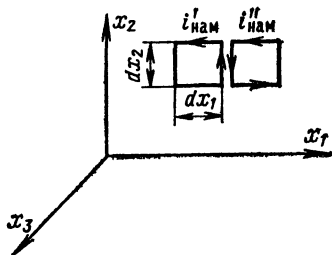


Рис. 1-1. К определению тока намагниченности.

положные стороны, взаимно компенсируются и суммарный магнитный ток равен нулю. Если намагниченность неоднородна, то токи, текущие в соседних контурах, неодинаковы, и компенсации не происходит. При этом суммарный магнитный ток не равен нулю.

Рассмотрим для простоты случай, когда намагниченность направлена по оси x_3 :

$$\mathbf{M} = (0, 0, M_3),$$

что соответствует расположению контуров элементарных магнитных диполей в плоскости $0x_1x_2$. Рассмотрим два контура с токами $i'_{\text{нам}}$ и $i''_{\text{нам}}$ (рис. 1-1). Если намагниченность среды неоднородна, то токи $i'_{\text{нам}}$ и $i''_{\text{нам}}$ неодинаковы:

$$i'_{\text{нам}} = \frac{m'}{ds} = \frac{M_3 dV}{ds} = \frac{M_3 dx_1 dx_2 dx_3}{dx_1 dx_2};$$

$$i''_{\text{нам}} = \frac{m''}{ds} = \frac{\left(M_3 + \frac{\partial M_3}{\partial x_1} dx_1 \right) dx_1 dx_2 dx_3}{dx_1 dx_2}.$$

Ток на общей границе контуров

$$i_{\text{нам}} = i''_{\text{нам}} - i'_{\text{нам}} = \frac{\partial M_3}{\partial x_1} dx_1 dx_3,$$

соответствующая плотность тока

$$J_{\text{нам}} = \frac{i_{\text{нам}}}{dx_1 dx_3} = \frac{dM_3}{dx_1}.$$

В общем случае

$$\mathbf{M} = (M_1, M_2, M_3),$$

$$\mathbf{J}_{\text{нам}} = \text{rot } \mathbf{M}.$$

Аналогичным путем можно получить волновое уравнение для вектора \mathbf{E} из системы (1-5):

$$\begin{aligned} & \text{rot rot } \mathbf{E} + \mu_a^\pi \sigma^\pi \frac{\partial \mathbf{E}}{\partial t} + \mu_a^\pi \epsilon_a^\pi \frac{\partial^2 \mathbf{E}}{\partial t^2} = \\ & = -\mu_a^\pi \frac{\partial^2 \mathbf{P}^{\pi\text{л}}}{\partial t^2} - \mu_a^\pi \frac{\partial \mathbf{J}^{\pi\text{л}}}{\partial t} - \mu_a^\pi \frac{\partial \mathbf{J}^{\text{ст}}}{\partial t} - \mu_0 \frac{\partial}{\partial t} \text{rot } \mathbf{M}^{\pi\text{л}}. \end{aligned} \quad (1-7)$$

Здесь члены $\frac{\partial^2 \mathbf{P}^{\pi\text{л}}}{\partial t^2}$, $\frac{\partial \mathbf{J}^{\pi\text{л}}}{\partial t}$, $\frac{\partial}{\partial t} \text{rot } \mathbf{M}^{\pi\text{л}}$ можно рассматривать как дополнительные источники поля, порождаю-

щие различные нелинейные эффекты (появление гармоник, смещение частот, выпрямление и т. д.).

Если свободные заряды отсутствуют, то

$$\operatorname{div} \mathbf{D} = 0$$

и

$$\operatorname{rot} \operatorname{rot} \mathbf{E} = \operatorname{grad} \operatorname{div} \mathbf{E} - \Delta \mathbf{E} = -\Delta \mathbf{E};$$

при этом волновое уравнение (1-6) будет иметь вид:

$$\Delta \mathbf{E} - \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2} + \mu_0 \frac{\partial \mathbf{J}}{\partial t} + \mu_0 \frac{\partial}{\partial t} \operatorname{rot} \mathbf{M}, \quad (1-8)$$

а уравнение (1-7)

$$\begin{aligned} \Delta \mathbf{E} - \mu_a^\perp \sigma^\perp \frac{\partial \mathbf{E}}{\partial t} - \mu_a^\perp \varepsilon_a^\perp \frac{\partial^2 \mathbf{E}}{\partial t^2} = \\ = \mu_a^\perp \frac{\partial^2 \mathbf{P}^{\text{нл}}}{\partial t^2} + \mu_a^\perp \frac{\partial \mathbf{J}^{\text{нл}}}{\partial t} + \mu_a^\perp \frac{\partial \mathbf{J}^{\text{сг}}}{\partial t} + \mu_0 \frac{\partial}{\partial t} \operatorname{rot} \mathbf{M}^{\text{нл}}. \end{aligned} \quad (1-9)$$

Уравнения Максвелла в символической форме. Если электромагнитное поле возбуждается монохроматическим током

$$\mathbf{J}^{\text{сг}} = \mathbf{J}_m^{\text{сг}} \cos(\omega t + \varphi),$$

то поле у источника также имеет монохроматический характер:

$$\mathbf{E} = \mathbf{E}_m \cos(\omega t + \varphi_E).$$

Поляризация нелинейного диэлектрика зависит от напряженности электрического поля

$$\mathbf{P} = \mathbf{P}(\mathbf{E})$$

и для изотропной среды может быть представлена в виде

$$\mathbf{P} = \mathbf{P}^\perp + \mathbf{P}^{\text{нл}}, \quad (1-10)$$

где

$$\begin{aligned} \mathbf{P}^\perp &= \varepsilon_0 \chi_1^\perp \mathbf{E}, \\ \mathbf{P}^{\text{нл}} &= \varepsilon_0 (\chi_{31} E^2 + \chi_{32} E^3 + \chi_{33} E^4 + \dots). \end{aligned}$$

Мы рассматриваем изотропный диэлектрик как упрощенную модель реальной среды, позволяющую сравнительно просто описать основные нелинейные эффекты в средах.

Для большинства диэлектриков (не обладающих сегнетоэлектрическими свойствами)

$$\frac{\chi_{32}}{\chi_{31}} \leq 10^{-11} \text{ м/В}$$

и

$$\chi_{31} E \gg \chi_{32} E^2 \gg \chi_{33} E^3 \gg \dots$$

Хотя в данном случае символический метод непосредственно не применим, в случае гармонического поля можно воспользоваться представлением

$$E = \frac{\dot{E} + \dot{E}^*}{2}, \quad (1-11)$$

где

$$\begin{aligned} \dot{E} &= \dot{E}_m e^{j\omega t} = E_m e^{j(\omega t + \varphi_E)}; \\ \dot{E}^* &= \dot{E}_m^* e^{-j\omega t} = E_m e^{-j(\omega t + \varphi_E)}. \end{aligned}$$

Подставляя (1-11) в выражение (1-10) для поляризации, получаем:

$$\begin{aligned} P = \epsilon_0 \left[\chi_3^l \frac{\dot{E} + \dot{E}^*}{2} + \frac{\chi_{31}}{4} (\dot{E}^2 + 2\dot{E}\dot{E}^* + \dot{E}^{*2}) + \right. \\ \left. + \frac{\chi_{32}}{8} (\dot{E}^3 + 3\dot{E}^2\dot{E}^* + 3\dot{E}\dot{E}^{*2} + \dot{E}^{*3}) + \right. \\ \left. + \frac{\chi_{33}}{16} (\dot{E}^4 + 4\dot{E}^3\dot{E}^* + 6\dot{E}^2\dot{E}^{*2} + 4\dot{E}\dot{E}^{*3} + \dot{E}^{*4}) + \dots \right]. \quad (1-12) \end{aligned}$$

Анализируя структуру выражения (1-12), можно заметить, что вследствие перемножения различных степеней \dot{E} и \dot{E}^* кроме составляющей с частотой ω возникают составляющие с удвоенной, утроенной и т. д. частотой, а также постоянная составляющая.

Рассмотрим, например, выражение $\dot{E}^2 + 2\dot{E}\dot{E}^* + \dot{E}^{*2}$, входящее в состав формулы (1-12). Поскольку

$$\dot{E}^2 = E_m^2 e^{j2(\omega t + \varphi_E)},$$

а

$$\dot{E}^{*2} = E_m^2 e^{-j2(\omega t + \varphi_E)},$$

то их сумма будет равна:

$$\dot{E}^2 + \dot{E}^{*2} = 2E_m^2 \cos(2\omega t + 2\varphi_E).$$

Найдем произведение $2\dot{E}\dot{E}^*$, входящее в состав этого же выражения:

$$2\dot{E}\dot{E}^* = 2E_m^2.$$

Таким образом, выражение $(\dot{E}^2 + 2\dot{E}\dot{E}^* + \dot{E}^{*2})$ содержит постоянную составляющую и составляющую с удвоенной частотой (вторую гармонику входного воздействия).

Аналогично этому, анализируя выражение $(\dot{E}^3 + 3\dot{E}^2\dot{E}^* + 3\dot{E}\dot{E}^{*2} + \dot{E}^{*3})$, нетрудно убедиться в том, что оно содержит составляющие с частотой ω и 3ω .

Группируя члены, соответствующие постоянной составляющей, первой гармонике, второй гармонике и т. д., преобразуем выражение (1-12) к следующему виду:

$$P = P^{\pi}(\omega) + \sum_{n=0}^{\infty} P^{n\pi}(n\omega),$$

где

$$P^{\pi}(\omega) = \epsilon_0 \frac{\chi_{\pi}^{\pi}}{2} (\dot{E} + \dot{E}^*)$$

— составляющая поляризации, линейно зависящая от амплитуды действующего поля и поэтому характеризующая лишь линейные эффекты;

$$P^{n\pi}(n\omega) = P_m^{n\pi}(n\omega) \cos n(\omega t + \varphi E)$$

— составляющие поляризации, нелинейно зависящие от действующего поля и поэтому определяющие нелинейные эффекты. В числе этих составляющих

$$P^{n\pi}(0) = \epsilon_0 \left(\frac{\chi_{\pi 1}^{\pi}}{2} E_m^2 + \frac{3}{8} \chi_{\pi 3}^{\pi} E_m^4 + \dots \right)$$

— постоянная составляющая поляризации. Комплексные амплитуды нелинейных составляющих поляризации с частотами ω , 2ω , 3ω ... определяются выражениями:

$$\left. \begin{aligned} P_m^{n\pi}(\omega) &= \epsilon_0 (\chi_{\pi 1}^{\omega} \dot{E}_m^2 \dot{E}_m^* + \chi_{\pi 2}^{\omega} \dot{E}_m^3 \dot{E}_m^{*2} + \dots); \\ P_m^{n\pi}(2\omega) &= \epsilon_0 (\chi_{\pi 1}^{2\omega} \dot{E}_m^2 + \chi_{\pi 2}^{2\omega} \dot{E}_m^3 \dot{E}_m^* + \dots); \\ P_m^{n\pi}(3\omega) &= \epsilon_0 (\chi_{\pi 1}^{3\omega} \dot{E}_m^3 + \chi_{\pi 2}^{3\omega} \dot{E}_m^4 \dot{E}_m^* + \dots), \\ &\dots \end{aligned} \right\} \quad (1-13)$$

где $\chi_{\pi k}^{n\omega}$ — коэффициенты, определяемые через коэффициенты $\chi_{\pi 1}$, $\chi_{\pi 2}$... разложения (1-10).

Пусть теперь возбуждающее поле содержит две гармонические составляющие:

$$E = E_{m1} \cos(\omega_1 t + \varphi_1) + E_{m2}(\cos(\omega_2 t + \varphi_2)). \quad (1-14)$$

Подставляя (1-14) в выражение для поляризации (1-10) и выполняя преобразования, аналогичные предельным выше, получаем выражение для поляризации в виде суммы гармонических составляющих с частотой ω_1 , ω_2 , $2\omega_1$, $2\omega_2$, ..., $\omega_1 + \omega_2$, $\omega_1 - \omega_2$, ..., $p\omega_1 \pm q\omega_2$, где p , q — целые числа. Составляющие с частотой $p\omega_1 \pm q\omega_2$ называются комбинационными составляющими, а частоты $\omega = p\omega_1 \pm q\omega_2$ — комбинационными частотами.

Амплитуды различных составляющих поляризации определяются выражениями

$$\left. \begin{aligned} \dot{P}_m^{\text{нл}}(\omega_1) &= \epsilon_0 (\chi_{\varepsilon 1}^{\omega_1} \dot{E}_{m1}^2 \dot{E}_{m1}^* + \chi_{\varepsilon 2}^{\omega_1} \dot{E}_{m1} \dot{E}_{m2} \dot{E}_{m2}^* + \\ &\quad + \chi_{\varepsilon 3}^{\omega_1} \dot{E}_{m1}^3 \dot{E}_{m1}^{*2} + \dots); \\ \dot{P}_m^{\text{нл}}(\omega_2) &= \epsilon_0 (\chi_{\varepsilon 1}^{\omega_2} \dot{E}_{m2}^2 \dot{E}_{m2}^* + \chi_{\varepsilon 2}^{\omega_2} \dot{E}_{m1} \dot{E}_{m1}^* \dot{E}_{m2} + \\ &\quad + \chi_{\varepsilon 3}^{\omega_2} \dot{E}_{m2}^3 \dot{E}_{m2}^{*2} + \dots); \\ \dot{P}_m^{\text{нл}}(\omega_1 + \omega_2) &= \epsilon_0 (\chi_{\varepsilon 1}^{\omega_1 + \omega_2} \dot{E}_{m1} \dot{E}_{m2} + \\ &\quad + \chi_{\varepsilon 2}^{\omega_1 + \omega_2} \dot{E}_{m1}^2 \dot{E}_{m1}^* \dot{E}_{m2} + \dots); \\ \dot{P}_m^{\text{нл}}(\omega_1 - \omega_2) &= \epsilon_0 (\chi_{\varepsilon 1}^{\omega_1 - \omega_2} \dot{E}_{m1} \dot{E}_{m2}^* + \\ &\quad + \chi_{\varepsilon 2}^{\omega_1 - \omega_2} \dot{E}_{m1}^2 \dot{E}_{m1}^* \dot{E}_{m2} + \dots); \\ \dot{P}_m^{\text{нл}}(2\omega_1) &= \epsilon_0 (\chi_{\varepsilon 1}^{2\omega_1} \dot{E}_{m1}^2 + \chi_{\varepsilon 2}^{2\omega_1} \dot{E}_{m1}^2 \dot{E}_{m2} \dot{E}_{m2}^* + \\ &\quad + \chi_{\varepsilon 3}^{2\omega_1} \dot{E}_{m1}^3 \dot{E}_{m1}^{*2} + \dots); \\ \dot{P}_m^{\text{нл}}(2\omega_2) &= \epsilon_0 (\chi_{\varepsilon 1}^{2\omega_2} \dot{E}_{m2}^2 + \chi_{\varepsilon 2}^{2\omega_2} \dot{E}_{m2}^2 \dot{E}_{m1} \dot{E}_{m1}^* + \\ &\quad + \chi_{\varepsilon 3}^{2\omega_2} \dot{E}_{m2}^3 \dot{E}_{m2}^{*2} + \dots) \end{aligned} \right\} \quad (1-15)$$

и т. д., где коэффициенты $\chi_{\varepsilon i}^{p\omega_1 \pm q\omega_2}$ также определяются через коэффициенты разложения (1-10) $\chi_{\varepsilon 1}$, $\chi_{\varepsilon 2}$...

При наличии в спектре возбуждающего поля трех и более составляющих с частотой ω_1 , ω_2 , ω_3 ... в спектре поляризации появятся более сложные комбинационные составляющие с частотой $p\omega_1 \pm q\omega_2 \pm r\omega_3 \pm \dots$, где p , q , r — целые числа.

Аналогичным путем можно проводить анализ процессов в нелинейных магнетиках. Представим намагниченность магнетика в виде

$$M = M^{\text{л}} + M^{\text{нл}},$$

где

$$M^{\text{л}} = \chi_m^{\text{л}} H, \\ M^{\text{нл}} = \chi_{m1} H^2 + \chi_{m2} H^3 + \chi_{m3} H^4 + \dots$$

В случае гармонического поля

$$H = H_m \cos(\omega t + \varphi_H).$$

Применяя представление

$$H = \frac{\dot{H} + \dot{H}^*}{2} = \frac{\dot{H}_m e^{j\omega t} + \dot{H}_m^* e^{-j\omega t}}{2},$$

получаем

$$M = \chi_m^{\text{л}} \frac{\dot{H} + \dot{H}^*}{2} + \frac{\chi_{m1}}{4} (H^2 + 2 \dot{H} \dot{H}^* + \dot{H}^{*2}) + \\ + \frac{\chi_{m2}}{8} (\dot{H}^3 + 3 \dot{H}^2 \dot{H}^* + 3 \dot{H} \dot{H}^{*2} + \dot{H}^{*3}) + \dots$$

или

$$M = M^{\text{нл}}(0) + M^{\text{л}}(\omega) + M^{\text{нл}}(\omega) + M^{\text{нл}}(2\omega) + M^{\text{нл}}(3\omega) + \dots$$

Таким образом, намагниченность нелинейного магнетика также содержит удвоенные, утроенные и т. д. частоты и постоянную составляющую. При действии сложного поля возникают комбинационные составляющие.

Аналогичным путем можно получить выражение для нелинейной плотности тока:

$$J = J^{\text{нл}}(0) + J^{\text{л}}(\omega) + J^{\text{нл}}(\omega) + J^{\text{нл}}(2\omega) + \dots$$

Согласно уравнениям Максвелла (1-5), а также волновым уравнениям (1-7) и (1-9) нелинейные составляющие поляризации, тока и намагничивания будут в свою очередь возбуждать гармоники поля, а если источник имеет несколько гармонических составляющих, то и комбинационные составляющие.

Итак, если источник монохроматический

$$J^{\text{ст}} = J_m^{\text{ст}} \cos(\omega t + \varphi),$$

то в нелинейной среде векторы \mathbf{E} , \mathbf{H} , $\mathbf{P}^{\text{нл}}$, $\mathbf{M}^{\text{нл}}$ и $\mathbf{J}^{\text{нл}}$ можно представить в виде суммы гармонических составляющих

$$\left. \begin{aligned} \mathbf{E} &= \frac{1}{2} \sum_{n=-\infty}^{\infty} \dot{\mathbf{E}}_m(n\omega) e^{jn\omega t}; \\ \mathbf{H} &= \frac{1}{2} \sum_{n=-\infty}^{\infty} \dot{\mathbf{H}}_m(n\omega) e^{jn\omega t}; \\ \mathbf{P}^{\text{нл}} &= \frac{1}{2} \sum_{n=-\infty}^{\infty} \dot{\mathbf{P}}_m^{\text{нл}}(n\omega) e^{jn\omega t}; \\ \mathbf{M}^{\text{нл}} &= \frac{1}{2} \sum_{n=-\infty}^{\infty} \dot{\mathbf{M}}_m^{\text{нл}}(n\omega) e^{jn\omega t}; \\ \mathbf{J}^{\text{нл}} &= \frac{1}{2} \sum_{n=-\infty}^{\infty} \dot{\mathbf{J}}_m^{\text{нл}}(n\omega) e^{jn\omega t}. \end{aligned} \right\} \quad (1-16)$$

Подставляя эти выражения в уравнения Максвелла (1-5), заменяя дифференцирование по времени умножением на $jn\omega$ и приравнивая величины, содержащие одинаковые частоты $n\omega$, получим бесконечную систему уравнений

$$\left. \begin{aligned} \text{rot } \dot{\mathbf{H}}_m(n\omega) &= jn\omega \left[\tilde{\epsilon}_a^n(n\omega) \dot{\mathbf{E}}_m(n\omega) + \dot{\mathbf{P}}_m^{\text{нл}}(n\omega) \right] + \\ &\quad + \dot{\mathbf{J}}_m^{\text{нл}}(n\omega) + \dot{\mathbf{J}}_m^{\text{ст}}(n\omega); \\ \text{rot } \dot{\mathbf{E}}_m(n\omega) &= -jn\omega \tilde{\mu}_a^n(n\omega) \dot{\mathbf{H}}_m(n\omega) - \\ &\quad - jn\omega \mu_0 \dot{\mathbf{M}}_m^{\text{нл}}(n\omega), \end{aligned} \right\} \quad (1-17)$$

где $n=0, \pm 1, \pm 2, \pm 3 \dots$

$$\tilde{\epsilon}_a^n(n\omega) = \epsilon_a^n(n\omega) - j \frac{\sigma}{n\omega}, \quad \tilde{\mu}_a^n(n\omega) = \mu_a^{n'}(n\omega) - j \mu_a^{n''}(n\omega),$$

$$\dot{\mathbf{J}}_m^{\text{ст}}(n\omega) = 0 \text{ при } n \neq 1.$$

Поле в нелинейной среде содержит большое число гармонических составляющих с частотами $n\omega$, которые взаимодействуют друг с другом. При взаимодействии, например, составляющих с частотами ω и 3ω возникают

комбинационные составляющие с частотами $3\omega + \omega = 4\omega$, $3\omega - \omega = 2\omega$, $3\omega - 2\omega = \omega$, $3\omega + 2\omega = 5\omega$ и т. д. Таким образом, все гармонические составляющие поля взаимосвязаны. Математически это описывается уравнениями (1-17), в правые части которых входят амплитуды $\mathbf{P}_m^{nl}(\omega)$, $\mathbf{M}_m^{nl}(\omega)$ и $\mathbf{J}_m^{nl}(\omega)$, причем каждая из этих амплитуд зависит не только от «своей» составляющей поля с частотой $n\omega$, но и от всех остальных гармонических составляющих. Таким образом, все уравнения (1-17) оказываются взаимосвязанными.

Если спектр источника поля содержит несколько гармонических составляющих, то в системе уравнений (1-17) кроме гармоник этих составляющих будут фигурировать комбинационные частоты. При этом все уравнения системы (1-17) также будут взаимосвязанными.

Волновое уравнение в символической форме. Подставляя (1-16) в волновое уравнение (1-7) и приравнявая члены, содержащие одинаковые частоты $n\omega$, получаем:

$$\begin{aligned} \text{rot rot } \dot{\mathbf{E}}_m(n\omega) - (n\omega)^2 \tilde{\epsilon}_a^n(n\omega) \tilde{\mu}_a^n(n\omega) \dot{\mathbf{E}}_m(n\omega) = \\ = (n\omega)^2 \tilde{\mu}_a^n(n\omega) \dot{\mathbf{P}}_m^{nl}(n\omega) - jn\omega \left[\tilde{\mu}_a^n(n\omega) \mathbf{J}_m^{nl}(n\omega) + \right. \\ \left. + \tilde{\mu}_a^n(n\omega) \mathbf{J}_m^{ct}(n\omega) + \mu_0 \text{rot } \dot{\mathbf{M}}_m^{nl}(n\omega) \right], \end{aligned} \quad (1-18)$$

где $n=0, \pm 1, \pm 2, \pm 3 \dots$;

$$\mathbf{J}_m^{ct}(n\omega) = 0 \text{ при } n \neq 1;$$

уравнение (1-9) переписывается в виде

$$\begin{aligned} \Delta \dot{\mathbf{E}}_m(n\omega) + (n\omega)^2 \tilde{\epsilon}_a^n(n\omega) \tilde{\mu}_a^n(n\omega) \dot{\mathbf{E}}_m(n\omega) = \\ = - (n\omega)^2 \tilde{\mu}_a^n(n\omega) \dot{\mathbf{P}}_m^{nl}(n\omega) + jn\omega \left[\tilde{\mu}_a^n(n\omega) \mathbf{J}_m^{nl}(n\omega) + \right. \\ \left. + \tilde{\mu}_a^n(n\omega) \mathbf{J}_m^{ct}(n\omega) + \mu_0 \text{rot } \dot{\mathbf{M}}_m^{nl}(n\omega) \right], \end{aligned} \quad (1-19)$$

где $n=0, \pm 1, \pm 2, \pm 3 \dots$;

$$\mathbf{J}_m^{ct}(n\omega) = 0 \text{ при } n \neq 1.$$

Если в спектре источника содержится несколько частот, то системы (1-18) и (1-19) будут содержать уравнения, написанные для гармоник этих частот и комбинационных частот. Системы (1-18) и (1-19) представ-

ляют собой системы бесконечного числа связанных уравнений.

При решении конкретных задач нелинейной электродинамики ограничиваются в зависимости от постановки и желаемой точности решения определенным номером гармоник, так как они быстро убывают с возрастанием номера. То же относится и к комбинационным частотам. При этом число уравнений сокращается, а сами нелинейные уравнения становятся приближенными.

Граничные условия. Полученные уравнения должны решаться с учетом граничных условий. В общем случае граничные условия для тангенциальных и нормальных составляющих поля [1] имеют вид:

$$\begin{aligned} E_{\tau(2)} &= E_{\tau(1)}, \quad D_{n(2)} - D_{n(1)} = \kappa; \\ H_{\tau(2)} - H_{\tau(1)} &= J_{\text{пов}}, \quad B_{n(2)} = B_{n(1)}, \end{aligned}$$

где $J_{\text{пов}}$ — плотность поверхностного тока, А/м; κ — плотность поверхностного заряда, Кл/м². Эти условия выполняются в любой момент времени, поэтому они должны выполняться для любой из гармоник и составляющих комбинационных частот. Для гармоник эти условия имеют вид:

$$\left. \begin{aligned} E_{\tau(2)}(n\omega) &= E_{\tau(1)}(n\omega); \\ H_{\tau(2)}(n\omega) - H_{\tau(1)}(n\omega) &= J_{\text{пов}}(n\omega); \\ D_{n(2)}(n\omega) - D_{n(1)}(n\omega) &= \kappa(n\omega); \\ B_{n(2)}(n\omega) &= B_{n(1)}(n\omega). \end{aligned} \right\} \quad (1-20)$$

Аналогичный вид имеют граничные условия и для комбинационных частот.

1-3. ЭНЕРГИЯ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ

Теорема Умова — Пойнтинга, выражающая закон сохранения энергии поля в интегральной форме, определяется выражением

$$\oint_s [\mathbf{E}\mathbf{H}] ds + \int_V \left(\mathbf{E} \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \frac{\partial \mathbf{B}}{\partial t} \right) dV + \int_V \mathbf{J}\mathbf{E} dV = 0, \quad (1-21)$$

где первый член определяет полную мощность, проходящую через поверхность s , ограничивающую исследуемый объем V ; второй член определяет изменение энер-

гии за единицу времени в объеме V , третий член — мощность, определяемую взаимодействием поля с проводящей средой.

Учитывая уравнения состояния среды

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P};$$

$$\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}),$$

формулу (1-21) можно переписать в виде

$$\oint_s [\mathbf{E}\mathbf{H}] d\mathbf{S} + \frac{\partial}{\partial t} \int_V \left(\frac{\epsilon_0 E^2 + \mu_0 H^2}{2} \right) dV + \int_V \mathbf{E} \frac{\partial \mathbf{P}}{\partial t} dV + \\ + \int_V \mu_0 \mathbf{H} \frac{\partial \mathbf{M}}{\partial t} dV + \int_V \mathbf{J} \mathbf{E} dV = 0.$$

Здесь

$$\frac{\partial}{\partial t} \int_V \left(\frac{\epsilon_0 E^2 + \mu_0 H^2}{2} \right) dV$$

— изменение энергии в единицу времени в вакууме;

$$\int_V \mathbf{E} \frac{\partial \mathbf{P}}{\partial t} dV$$

— изменение энергии в единицу времени в среде из-за поляризации;

$$\int_V \mu_0 \mathbf{H} \frac{\partial \mathbf{M}}{\partial t} dV$$

— изменение энергии в единицу времени в среде из-за намагничивания.

Плотность мощности, связанная с взаимодействием поля с проводящей средой, равна:

$$p_{\text{пров}} = \mathbf{J} \mathbf{E}. \quad (1-22)$$

Плотности мощности, связанные с процессами поляризации и намагничивания среды, определяются соответственно выражениями

$$\left. \begin{aligned} p_{\text{пол}} &= \mathbf{E} \frac{\partial \mathbf{P}}{\partial t}; \\ p_{\text{нам}} &= \mu_0 \mathbf{H} \frac{\partial \mathbf{M}}{\partial t}. \end{aligned} \right\} \quad (1-23)$$

В нелинейной среде в случае монохроматического источника

$$\mathbf{E} = \sum_{n=0}^{\infty} \mathbf{E}_m(n\omega) \cos [n\omega t - \varphi_E(n\omega)];$$

$$\mathbf{H} = \sum_{n=0}^{\infty} \mathbf{H}_m(n\omega) \cos [n\omega t - \varphi_H(n\omega)];$$

$$\mathbf{P} = \sum_{n=0}^{\infty} \mathbf{P}_m(n\omega) \cos [n\omega t - \varphi_P(n\omega)];$$

$$\mathbf{M} = \sum_{n=0}^{\infty} \mathbf{M}_m(n\omega) \cos [n\omega t - \varphi_M(n\omega)];$$

$$\mathbf{J} = \sum_{n=0}^{\infty} \mathbf{J}_m(n\omega) \cos [n\omega t - \varphi_J(n\omega)];$$

согласно (1-22) и (1-23)

$$p_{\text{пров}} = \sum_{q,r=0}^{\infty} \mathbf{J}_m(q\omega) \mathbf{E}_m(r\omega) \cos [q\omega t - \varphi_J(q\omega)] \times \\ \times \cos [r\omega t - \varphi_E(r\omega)], \quad (1-24)$$

$$p_{\text{пол}} = - \sum_{q,r=0}^{\infty} r\omega \mathbf{E}_m(q\omega) \mathbf{P}_m(r\omega) \times \\ \times \cos [q\omega t - \varphi_E(q\omega)] \sin [r\omega t - \varphi_P(r\omega)], \quad (1-25)$$

$$p_{\text{нам}} = - \mu_0 \sum_{q,r=0}^{\infty} r\omega \mathbf{H}_m(q\omega) \mathbf{M}_m(r\omega) \times \\ \times \cos [q\omega t - \varphi_H(q\omega)] \sin [r\omega t - \varphi_M(r\omega)]. \quad (1-26)$$

В приведенных выражениях с учетом волнового характера процессов

$$\varphi_E(n\omega) = \mathbf{k}(n\omega) \mathbf{r} + \varphi_E;$$

$$\varphi_H(n\omega) = \mathbf{k}(n\omega) \mathbf{r} + \varphi_H;$$

$$\varphi_P(n\omega) = \mathbf{k}_P(n\omega) \mathbf{r} + \varphi_P;$$

$$\varphi_M(n\omega) = \mathbf{k}_M(n\omega) \mathbf{r} + \varphi_M;$$

$$\varphi_J(n\omega) = \mathbf{k}_J(n\omega) \mathbf{r} + \varphi_J;$$

где $k(n\omega)$, $k_P(n\omega)$, $k_M(n\omega)$ и $k_J(n\omega)$ — постоянные распространения гармоник поля, поляризации, намагничённости и тока.

Слагаемые, стоящие под знаком суммы в выражениях (1-24) — (1-26) для мгновенных плотностей мощности, определяют обмен энергией между любыми гармониками поля и тока, поля и поляризации, поля и намагничённости. Если эти слагаемые положительны, то энергия от поля переходит в среду, если отрицательны — то от среды к полю. В каждой точке пространства этот процесс в зависимости от времени направлен в ту или другую сторону. В фиксированный момент времени в одних точках пространства процесс может быть направлен в одну сторону, в других — в другую.

Выражения для плотности мощности можно представить и в виде

$$p_{\text{пров}} = \frac{1}{2} \sum_{q,r=0}^{\infty} J_m(q\omega) E_m(r\omega) \{ \cos[(q+r)\omega t - \varphi_J(q\omega) - \varphi_E(r\omega)] + \cos[(q-r)\omega t - \varphi_J(q\omega) + \varphi_E(r\omega)] \}; \quad (1-27)$$

$$p_{\text{пол}} = -\frac{1}{2} \sum_{q,r=0}^{\infty} r\omega E_m(q\omega) P_m(r\omega) \{ \sin[(q+r)\omega t - \varphi_E(q\omega) - \varphi_P(r\omega)] - \sin[(q-r)\omega t - \varphi_E(q\omega) + \varphi_P(r\omega)] \}; \quad (1-28)$$

$$p_{\text{нам}} = -\frac{1}{2} \mu_0 \sum_{q,r=0}^{\infty} r\omega H_m(q\omega) M_m(r\omega) \times \\ \times \{ \sin[(q+r)\omega t - \varphi_H(q\omega) - \varphi_M(r\omega)] - \sin[(q-r)\omega t - \varphi_H(q\omega) + \varphi_M(r\omega)] \}. \quad (1-29)$$

Средние плотности мощности, связанные с проводимостью, поляризацией и намагничиванием среды, определяются соответственно выражениями

$$p_{\text{ср пров}} = \frac{1}{T} \int_0^T J E dt;$$

$$p_{0 \text{ пол}} = \frac{1}{T} \int_0^T \mathbf{E} \frac{\partial \mathbf{P}}{\partial t} dt;$$

$$p_{0 \text{ нам}} = \frac{1}{T} \int_0^T \mu_0 \mathbf{H} \frac{\partial \mathbf{M}}{\partial t} dt,$$

где время усреднения T много больше периода колебаний ($T \gg 2\pi/\omega$).

При усреднении по времени в выражениях (1-27), (1-28) и (1-29) отличны от нуля только слагаемые, не зависящие от времени, т. е. соответствующие $q-r=0$ и

$$\begin{aligned} p_{0 \text{ пров}} &= \frac{1}{2} \sum_{n=0}^{\infty} \mathbf{J}_m(n\omega) \mathbf{E}_m(n\omega) \cos \varphi_{JE}(n\omega) = \\ &= \sum_{n=0}^{\infty} \frac{1}{2} \operatorname{Re} \mathbf{J}(n\omega) \dot{\mathbf{E}}^*(n\omega) = \sum_{n=0}^{\infty} \frac{1}{2} \operatorname{Re} \mathbf{J}^*(n\omega) \dot{\mathbf{E}}(n\omega); \quad (1-30) \end{aligned}$$

$$\begin{aligned} p_{0 \text{ пол}} &= -\frac{1}{2} \sum_{n=0}^{\infty} n\omega \mathbf{E}_m(n\omega) \mathbf{P}_m(n\omega) \sin \varphi_{EP}(n\omega) = \\ &= \sum_{n=0}^{\infty} \frac{n\omega}{2} \operatorname{Im} \dot{\mathbf{E}}(n\omega) \dot{\mathbf{P}}^*(n\omega) = \\ &= -\sum_{n=0}^{\infty} \frac{n\omega}{2} \operatorname{Im} \dot{\mathbf{E}}^*(n\omega) \dot{\mathbf{P}}(n\omega); \quad (1-31) \end{aligned}$$

$$\begin{aligned} p_{0 \text{ нам}} &= -\frac{1}{2} \mu_0 \sum_{n=0}^{\infty} n\omega \mathbf{H}_m(n\omega) \mathbf{M}_m(n\omega) \sin \varphi_{HM}(n\omega) = \\ &= \mu_0 \sum_{n=0}^{\infty} \frac{n\omega}{2} \operatorname{Im} \dot{\mathbf{H}}(n\omega) \dot{\mathbf{M}}^*(n\omega) = \\ &= -\mu_0 \sum_{n=0}^{\infty} \frac{n\omega}{2} \operatorname{Im} \dot{\mathbf{H}}^*(n\omega) \dot{\mathbf{M}}(n\omega), \quad (1-32) \end{aligned}$$

где $\varphi_{JE}(n\omega) = \varphi_J(n\omega) - \varphi_E(n\omega)$, $\varphi_{EP}(n\omega) = \varphi_E(n\omega) - \varphi_P(n\omega)$, $\varphi_{HM} = \varphi_H(n\omega) - \varphi_M(n\omega)$ — сдвиг по фазе во

времени между $\mathbf{J}(n\omega)$ и $\mathbf{E}(n\omega)$, $\mathbf{E}(n\omega)$ и $\mathbf{P}(n\omega)$, $\mathbf{H}(n\omega)$ и $\mathbf{M}(n\omega)$ соответственно.

При $-\pi/2 < \varphi_{JE}(n\omega) < \pi/2$ слагаемые выражения (1-30) больше нуля и определяют мощность, поглощаемую средой из-за потерь проводимости, при $\pi/2 < \varphi_{JE}(n\omega) < 3/2\pi$ среда отдает энергию полю.

При $\pi < \varphi_{EP, HM}(n\omega) < 2\pi$ слагаемые выражений (1-31) и (1-32) больше нуля и определяют мощность, отдаваемую полем среде и расходуемую на увеличение поляризации и намагниченности среды, при $0 < \varphi_{EP, HM}(n\omega) < \pi$ соответствующие слагаемые меньше нуля и определяют мощность, отдаваемую средой распространяющему полю. Эффективный обмен энергией возможен лишь между одинаковыми гармониками поля и поляризации или намагниченности. Максимум энергии передается при сдвиге по фазе $\pm\pi/2$.

1-4. ПАССИВНЫЕ И АКТИВНЫЕ СРЕДЫ

Рассмотрим распространение плоской электромагнитной волны в однородной изотропной среде с комплексными диэлектрической и магнитной проницаемостями:

$$\tilde{\epsilon}_a = \epsilon'_a - j\epsilon''_a;$$

$$\tilde{\mu}_a = \mu'_a - j\mu''_a.$$

Постоянная распространения плоской волны в такой среде определяется выражением

$$\dot{k} = \omega \sqrt{\tilde{\epsilon}_a \tilde{\mu}_a} = \beta - j\alpha,$$

где β — фазовая постоянная; α — постоянная затухания.

Среда называется *пассивной*, если $\alpha > 0$ и она поглощает энергию распространяющегося электромагнитного поля.

Среда называется *активной*, если $\alpha < 0$ и она отдает энергию электромагнитному полю.

Если $\epsilon''_a > 0$ и $\mu''_a > 0$, то $\alpha > 0$ и среда пассивна.

Если $\epsilon''_a < 0$ и $\mu''_a < 0$, то $\alpha < 0$ и среда активна.

Если $\epsilon''_a > 0$, а $\mu''_a < 0$ или $\epsilon''_a < 0$, а $\mu''_a > 0$, то для определения характера среды надо определить знак величины

$$\alpha = -\operatorname{Im} \left[\omega \sqrt{(\epsilon'_a - j\epsilon''_a)(\mu'_a - j\mu''_a)} \right].$$

Примеры активных сред: 1) область электронного пучка, в которой электроны отдают энергию полю; 2) среда внутри конденсатора, емкость которого периодически изменяется (параметрические колебания); 3) активные среды, используемые в квантовых усилителях и генераторах.

При распространении плоской волны в активной среде в направлении оси x_3 электромагнитное поле определяется выражениями

$$E = E_m e^{|\alpha|x_3} \cos(\omega t - \beta x_3);$$

$$H = H_m e^{|\alpha|x_3} \cos(\omega t - \beta x_3 - \varphi).$$

При распространении волны на расстояние l напряженность поля увеличивается в $e^{|\alpha|l}$ раз, мощность поля в $e^{2|\alpha|l}$ раз. Величину $|\alpha|$ можно назвать коэффициентом усиления активной среды. При этом величина $G \approx \approx 8,69|\alpha|l$ определяет усиление в децибелах.

Поток электронов в переменном поле. Рассмотрим свойства некоторых сред более подробно. Для простоты полагаем, что поле имеет продольную электрическую составляющую, совпадающую с осью электронного пучка, и однородно по сечению пучка. Если электрическая составляющая поля совпадает с направлением движения электронов, то они тормозятся и энергия поля увеличивается за счет уменьшения кинетической энергии электронов. Происходит излучение, $\alpha < 0$, среда ведет себя как активная. При противоположном направлении электрической составляющей поля электроны ускоряются, кинетическая энергия их увеличивается, а энергия поля уменьшается. Электроны поглощают энергию поля, $\alpha > 0$, среда ведет себя как пассивная.

Рассмотрим случай торможения электронов. Определим напряженность электрического поля выражением

$$E = e_3 E_m e^{i(\omega t - \vec{k}x_3)},$$

где $\vec{k} = \beta - j\alpha$ — постоянная распространения. При этом скорость электронов можно представить в виде суммы двух слагаемых:

$$v = e_3 v_0 + e_3 v_{\sim},$$

где v_0 — постоянная составляющая, определяемая постоянным электрическим полем, создаваемым внешним

источником; v_{\sim} — переменная составляющая, определяемая переменным полем:

$$\dot{v} = -v_m e^{j(\omega t - kx_3)}, \quad v_0 \gg v_m;$$

знак минус определяет торможение электрона.

В линейном приближении уравнение движения электрона имеет следующий вид:

$$m \frac{d^2 x_3}{dt^2} = eE$$

или

$$\frac{dv}{dt} = \frac{e}{m} E, \quad (1-33)$$

где e и m — соответственно заряд и масса электрона.

Запишем полную производную скорости $v = v(x_3, t)$

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x_3} \frac{dx_3}{dt} = \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x_3}$$

и найдем уравнение движения электрона в символической форме:

$$\frac{e}{m} \dot{E}_m \approx j\omega \dot{v}_m - jkv_0 v_m. \quad (1-34)$$

Из-за модуляции скорости электронов происходит модуляция плотности заряда (образование сгустков)

$$\dot{\rho} = \rho_0 + \dot{\rho}_{\sim},$$

где $\dot{\rho}_{\sim} = \dot{\rho}_m e^{j(\omega t - kx_3)}$ — переменная составляющая плотности электронного потока; $\rho_m \ll \rho_0$, а также модуляция плотности тока

$$\dot{J} = \rho_0 v_0 + \dot{J}_{\sim},$$

где

$$\dot{J}_{\sim} = (\rho_0 \dot{v}_m + \dot{\rho}_m v_0) e^{j(\omega t - kx_3)} = -\dot{J}_m e^{j(\omega t - kx_3)} \quad (1-35)$$

— переменная составляющая плотности тока.

Так как плотность тока J является функцией координаты x_3 , то согласно уравнению непрерывности ($\text{div } \mathbf{J} = -\partial \rho / \partial t$)

$$\frac{\partial \rho}{\partial t} = -\frac{\partial J}{\partial x_3}$$

или

$$\dot{\rho}_m = \frac{k}{\omega} \dot{J}_m. \quad (1-36)$$

Решая совместно уравнения (1-34), (1-35) и (1-36), получаем

$$\dot{J}_m \approx \frac{jk_3 \frac{\rho_0}{v_0} \frac{e}{m}}{(k_3 - k)^2} \dot{E}_m,$$

где $k_3 = \omega/v_0$ — величина, которую можно назвать волновым электронным числом.

Плотность комплексной мощности переменной составляющей тока определяется выражением

$$\dot{p} = \frac{1}{2} (\dot{J}_m^* \dot{E}_m) \approx - \frac{jk_3 \frac{\rho_0}{v_0} \frac{e}{m}}{2[(k_3 - \beta)^2 + 2j\alpha(k_3 - \beta) - \alpha^2]} E_m^2,$$

а плотность действительной мощности

$$\rho_0 = \text{Re } \dot{p} = - \frac{k_3 \frac{\rho_0}{v_0} \frac{e}{m} (k_3 - \beta) \alpha}{[(k_3 - \beta)^2 - \alpha^2]^2 + 4\alpha^2 (k_3 - \beta)^2} E_m^2. \quad (1-37)$$

При торможении электронов $\rho_0 < 0$, и электроны отдают энергию полю.

Согласно выражению (1-37) усиление электромагнитного поля ($\alpha < 0$) происходит при $k_3 < \beta$, или $v_0 > v_\phi$, т. е. скорость движения электронов должна превышать фазовую скорость распространения электромагнитной волны, чтобы сгустки электронов находились в тормозящем поле.

При $v_0 = v_\phi$ плотность действительной мощности $\rho_0 = 0$ и взаимодействие поля с потоком электронов отсутствует, что соответствует локализации электронов в области нулевого поля.

С другой стороны, плотность действительной мощности определяется выражением

$$\rho_0 = \frac{\omega \varepsilon_a''}{2} E_m^2. \quad (1-38)$$

Сравнивая (1-37) и (1-38), найдем величину мнимой части диэлектрической проницаемости ε_a'' :

$$\varepsilon_a'' = - \frac{2 \frac{\rho_0}{v_0^2} \frac{e}{m} (k_3 - \beta) \alpha}{[(k_3 - \beta)^2 - \alpha^2]^2 + 4\alpha^2 (k_3 - \beta)^2}. \quad (1-39)$$

Выражение (1-39) справедливо и для ускоряющего, и для тормозящего поля. Различие лишь в том, что в случае тормозящего поля $k_a < \beta$, $\epsilon''_a < 0$, и среда активная, а в ускоряющем поле $k_a > \beta$, $\epsilon''_a > 0$, и среда пассивная.

Квантовые двухуровневые среды. Кинетическая и потенциальная энергия свободных электронов в электромагнитном поле может принимать любые некантованные значения. Внутренняя энергия атомов и молекул твердого тела или газа кантована (см. П1, пункты 1, 2 и 3). При взаимодействии электромагнитного поля с этими частицами происходят их переходы с одного энергетического уровня на другой (П1, пункты 15 и 19). При этом частицы могут поглотить квант энергии электромагнитного поля и перейти в более высокое энергетическое состояние или отдать квант энергии и перейти в состояние с меньшей энергией. Электромагнитное поле также кантовано и состоит из дискретных квантов, но количество этих квантов огромно даже при малых мощностях, поэтому дискретный характер поля практически не проявляется.

Вследствие взаимодействия с собственным излучением система переходит из одного энергетического состояния n в другое m , излучив квант энергии

$$\hbar\omega_{nm} = W_n - W_m.$$

Эти переходы совершаются при отсутствии внешнего электромагнитного излучения и поэтому называются спонтанными (самопроизвольными).

Спонтанные переходы случайны и различные частицы излучают одновременно и независимо, поэтому фазы электромагнитных волн, излучаемых отдельными частицами, не согласованы друг с другом. Случайный характер имеет не только время излучения кванта, но и его поляризация и направление распространения. Это некогерентное излучение.

Рассмотрим систему частиц, имеющую энергетические уровни W_m и W_n ($W_n > W_m$). Если плотность мала, то взаимодействием атомов друг с другом можно пренебречь и безызлучательные переходы (П1, п. 7) не учитывать. Пусть в единице объема вещества в момент времени t некоторое количество частиц N_n находится в состояниях с энергией W_n . В течение промежутка времени dt часть частиц перейдет на уровень W_m . Если вероятность спонтанного перехода в единицу времени для од-

ной частицы обозначить через A_{nm} , то число частиц в единице объема, покинувших уровень с энергией W_n за время dt , определится выражением

$$dN_n = -A_{nm} N_n dt;$$

отсюда

$$N_n(t) = N_n(0) e^{-A_{nm}t} = N_n(0) e^{-\frac{t}{\Delta t_{nm}}},$$

где $N_n(0)$ — число частиц в единице объема вещества в состояниях с энергией W_n в момент времени $t=0$.

Для одной частицы вероятность нахождения на возбужденном уровне определяется выражением

$$f(t) = e^{-t/\Delta t_{nm}}. \quad (1.4)$$

Величина

$$\Delta t_{nm} = \frac{1}{A_{nm}}$$

определяет время, в течение которого число частиц, находящихся в возбужденном состоянии, убывает в e раз. Она определяет среднее время жизни частиц в возбужденном состоянии. Действительно, число частиц, проживших на уровне W_n время t , равно $A_{nm}N_n(t)$ — числу частиц, покидающих в этот момент этот уровень, и среднее время жизни определится выражением

$$\Delta t_{nm} = \frac{1}{N_n(0)} \int_0^{\infty} t A_{nm} N_n dt = \frac{1}{A_{nm}},$$

т. е. среднее время жизни Δt_{nm} есть величина обратная вероятности перехода.

Спонтанные переходы обуславливают конечность времени жизни атома или молекулы в возбужденном состоянии. Основными характеристиками процесса являются время жизни частицы в возбужденном состоянии или обратная величина — вероятность спонтанного перехода.

Плотность мощности спонтанного перехода определяется выражением

$$p_{\text{спонт}}(t) = \hbar \omega_{nm} A_{nm} N_n(t)$$

или

$$p_{\text{спонт}}(t) = \hbar \omega_{nm} A_{nm} N_n(0) e^{-A_{nm}t} = p_{\text{спонт}}(0) e^{-\frac{t}{\Delta t_{nm}}}, \quad (1.41)$$

т. е. мощность спонтанного излучения убывает во времени по экспоненциальному закону.

Величина A_{nm} называется *коэффициентом Эйнштейна для спонтанного излучения*. Она определяется средним числом фотонов спонтанно излучаемых одной частицей в 1 с и имеет размерность 1/с.

Излучение, вызываемое взаимодействием с веществом внешнего электромагнитного поля, частота которого близка к частоте квантового перехода (П1, п. 7), называется *индуцированным или вынужденным*. При этом частота, поляризация и направление распространения индуцированного излучения одинаково с полем, вызывающим это излучение, и внешнее поле усиливается без изменения характеристик колебания. Это когерентное излучение.

Вероятность индуцированного перехода в единицу времени (П1, п. 19) пропорциональна спектральной плотности энергии $\omega_0(\omega_{nm})$ вынуждающего излучения на частоте перехода и для одной частицы равна при индуцированном излучении

$$P_{nm} = B_{nm} \omega_0(\omega_{nm}). \quad (1-42)$$

Величина B_{nm} , $\text{м}^3 \cdot \text{Дж}^{-1} \cdot \text{с}^{-2}$, называется *коэффициентом Эйнштейна для индуцированного излучения*. Она определяется числом фотонов, испускаемых в среднем одной частицей в 1 с при единичной спектральной плотности вынуждающего излучения.

Вероятность излучательного перехода равна сумме вероятностей спонтанного и индуцированного переходов и для одной частицы определяется выражением

$$A_{nm} + B_{nm} \omega_0(\omega_{nm}),$$

при этом число переходов с уровня W_n на уровень W_m в момент времени t за промежуток dt в единице объема составит:

$$dN_n = -[A_{nm} + B_{nm} \omega_0(\omega_{nm})] N_n dt.$$

Внешнее поле может быть причиной не только индуцированных переходов с верхних уровней на нижние, сопровождаемые индуцированным излучением, но и причиной переходов с нижних уровней на верхние, приводящих к *резонансному поглощению* электромагнитной энергии полем веществом.

Вероятность перехода с уровня W_m на уровень W_n в единицу времени для одной частицы также пропорциональна спектральной плотности энергии на частоте перехода:

$$P_{mn} = B_{mn} \omega_0(\omega_{mn}).$$

Величина B_{mn} называется коэффициентом Эйнштейна для индуцированного поглощения и определяется числом фотонов в среднем поглощаемых одной частицей за 1 с при единичной спектральной плотности вынуждающего излучения.

Число переходов с уровня W_m на уровень W_n в единице объема вещества в момент t за промежуток времени dt определится выражением

$$dN_m = -B_{mn} \omega_0(\omega_{mn}) N_m dt,$$

где N_m — число частиц на уровне W_m в единице объема в момент времени t .

При тепловом равновесии, когда излучаемая и поглощаемая энергии одинаковы, одинаково и число частиц, переходящих с верхнего уровня на нижний и обратно:

$$[A_{nm} + B_{nm} \omega_0(\omega_{nm})] N_n = B_{mn} \omega(\omega_{mn}) N_m.$$

Из последнего соотношения и термодинамических соображений можно показать, что

$$g_n B_{nm} = g_m B_{mn},$$

где g_n и g_m — кратность вырождения уровней с энергией W_n и W_m (П1, п. 1);

$$A_{nm} = \frac{2\hbar\omega^3}{\pi c^3} B_{nm} = \frac{2\hbar\omega^3}{\pi c^3} \frac{g_m}{g_n} B_{mn}. \quad (1.43)$$

Если коэффициент преломления среды n отличен от единицы, то c заменяется на $v = c/n$.

В приведенных выше расчетах использовались коэффициенты Эйнштейна A_{nm} и B_{nm} , определяющие полные (интегральные по частоте) вероятности спонтанного и индуцированного излучения в единицу времени. В действительности вероятности перехода при спонтанном и индуцированном излучении зависят от частоты, на которой излучается фотон. Эта зависимость от частоты определяется формой спектральной линии (см. П1, п. 20). Спектральную плотность вероятности спонтанного перехода $A_{nm}(\omega)$ можно представить в виде

$$A_{nm}(\omega) = A_{nm} g(\omega),$$

причем

$$A_{nm} = \int_{-\infty}^{+\infty} A_{nm}(\omega) d\omega,$$

где $g(\omega)$ — форм-фактор спектральной линии. Аналогичное выражение можно записать для спектральной плотности вероятности индуцированного перехода:

$$B_{nm}(\omega) = B_{nm} g(\omega).$$

Очевидно, что соотношения между коэффициентами $A_{nm}(\omega)$ и $B_{nm}(\omega)$ те же самые, что и между A_{nm} и B_{nm} .

Плотность мощности индуцированного излучения при переходе с уровня W_n на уровень W_m определяется выражением

$$p_{0 \text{ изл}}^{\text{инд}} = B_{nm} g(\omega) \omega_0(\omega_{nm}) \hbar \omega_{nm} N_n. \quad (1-44)$$

Плотность мощности поглощения при переходе частиц с уровня W_m на уровень W_n :

$$p_{0 \text{ погл}}^{\text{инд}} = \frac{g_n}{g_m} B_{nm} g(\omega) \omega_0(\omega_{nm}) \hbar \omega_{nm} N_m. \quad (1-45)$$

Спонтанные переходы уменьшают населенность (см. П1, п. 4) только верхних уровней, индуцированные переходы происходят в обе стороны. При малой интенсивности вынуждающего поля интенсивность индуцированных переходов меньше интенсивности спонтанных и практически частицы переходят только на нижние уровни. При большой интенсивности вынуждающего поля интенсивность индуцированных переходов значительно превышает интенсивность спонтанных переходов и наблюдается тенденция к выравниванию населенности рассматриваемых верхнего и нижнего уровней.

В естественных условиях при равновесии между веществом и окружающей средой (в условиях теплового равновесия) распределение частиц по энергетическим уровням подчиняется *распределению Больцмана* (П1, п. 4)

$$\frac{N_n}{N_m} = \frac{g_n}{g_m} e^{-\frac{W_n - W_m}{kT}}.$$

Из этого распределения видно, что населенности уменьшаются с увеличением энергии уровня; с повыше-

нием температуры населенность верхних уровней растет, но все же остается меньше, чем населенность нижних уровней. Поэтому при нормальных условиях

$$\frac{N_n}{g_n} < \frac{N_m}{g_m}$$

и согласно (1-44) и (1-45)

$$p_{0 \text{ погл}}^{\text{инд}} > p_{0 \text{ изл}}^{\text{инд}},$$

т. е. вещество поглощает энергию электромагнитной волны. Среда, удовлетворяющая распределению Больцмана, является *пассивной*.

Чтобы при распространении волны в среде она не затухала, а усиливалась, необходимо, чтобы

$$p_{0 \text{ изл}}^{\text{инд}} > p_{0 \text{ погл}}^{\text{инд}}$$

и, следовательно,

$$\frac{N_n}{g_n} > \frac{N_m}{g_m}.$$

Неравновесные состояния, удовлетворяющие этому условию, называются состояниями с *инверсной населенностью*. Волна, частота которой соответствует частоте индуцированного перехода, при распространении в среде с инверсной населенностью усиливается, и поэтому такая среда является *активной*.

Спонтанные переходы не влияют на усиливающееся поле, но создают фон (шумы). Поскольку мощность этих шумов пропорциональна коэффициенту $A_{nm} = A_{nm}(\omega^3)$ [см. (1-43)], то она быстро растет с увеличением частоты и затрудняет прием полезного сигнала.

Состояния с инверсной населенностью удовлетворяют распределению Больцмана при «отрицательной температуре»

$$T = - \frac{W_n - W_m}{k \ln \frac{N_n}{g_n} \frac{g_m}{N_m}} < 0;$$

это понятие условно, оно эквивалентно понятию «инверсии населенностей».

Величина

$$\Delta N = N_n - \frac{g_n}{g_m} N_m > 0$$

определяет плотность инверсной населенности. Если вы-
рождение отсутствует,

$$\Delta N = N_n - N_m > 0.$$

Так как спонтанное излучение не влияет на усиление электромагнитного поля, распространяющегося в среде с инверсной населенностью, а вносит лишь шумы, то плотность мощности излучения согласно (1-44) и (1-45) определится выражением

$$p_{0 \text{ изл}} = p_{0 \text{ изл}}^{\text{инд}} - p_{0 \text{ погл}}^{\text{инд}} = B_{nm} g(\omega) \omega_0(\omega_{nm}) \hbar \omega_{nm} \Delta N. \quad (1-46)$$

С другой стороны, плотность действительной мощности в диэлектрике определяется выражением

$$p_{0\varepsilon} = \frac{\omega}{2} \varepsilon_a'' E_m^2.$$

В зависимости от знака ε_a'' это или поглощаемая, или излучаемая мощность. Отсюда согласно (1-46) для активного диэлектрика

$$\varepsilon'' = -B_{nm} g(\omega) \hbar \Delta N < 0, \quad (1-47)$$

так как $\omega_{nm} = \omega$ — частота распространяющегося поля,

$$\omega_0(\omega_{nm}) = \omega_0(\omega) = \frac{\varepsilon_0 E_m^2}{2}$$

— средняя плотность энергии. Для пассивного диэлектрика ε'' тоже определяется выражением (1-47), но $\varepsilon'' > 0$.

В случае магнитной среды

$$p_{0\mu} = \frac{\omega}{2} \mu_a'' H_m^2$$

и магнитная проницаемость активного магнетика определяется выражением

$$\mu'' = -B_{nm} g(\omega) \hbar \Delta N < 0; \quad (1-48)$$

для пассивного магнетика справедливо это же выражение, но $\mu'' > 0$.

Любое неравновесное распределение, если оно не подвергается возмущению, самопроизвольно изменяется до тех пор, пока не станет равновесным. Процессы, приводящие систему к равновесному состоянию, называются *релаксационными*. Если система приходит к равновесному состоянию за счет обмена энергией с окружаю-

щей средой, то релаксационные процессы называются *«продольными»*. Отклонение внутренней энергии от равновесного значения уменьшается во времени по экспоненциальному закону и определяется уравнением

$$\frac{d(W - W_{\text{равн}})}{dt} = -\frac{1}{T_1}(W - W_{\text{равн}}), \quad (1-49)$$

где T_1 — величина, обратная затуханию и называемая *«продольным» временем релаксации*.

Возвращение системы к равновесному состоянию за счет обмена энергией с окружающей средой определяется функцией

$$f_1(t) = \frac{1}{T_1} e^{-t/T_1}. \quad (1-50)$$

Система может находиться в неравновесном состоянии, не отличающемся от равновесного состояния энергией ($W = W_{\text{равн}}$), а отличающемся некоторой упорядоченностью, например между частицами может быть фазовая взаимосвязь. Релаксационные процессы, стремящиеся нарушить упорядоченность системы, называются *«поперечными»*. Они характеризуются экспоненциальным убыванием упорядоченности во времени и характеризуются *«поперечным» временем релаксации T_2* .

Возвращение системы к равновесному состоянию за счет разупорядоченности определяется функцией

$$f_2(t) = \frac{1}{T_2} e^{-t/T_2}. \quad (1-51)$$

Обмен энергией между системой и окружающей средой всегда приводит к разупорядоченности. Поэтому T_2 не может быть больше T_1 . В твердых средах $T_2 \ll T_1$, в жидкостях и газах T_2 и T_1 близки.

При больших напряженностях поля, поскольку вероятность индуцированного излучения и поглощения растет с ростом поля и значительно превышает вероятность релаксационного процесса (спонтанного излучения), населенности верхнего и нижнего уровней выравниваются. При этом доля мощности поглощаемого излучения по сравнению с мощностью всего электромагнитного поля падает, стремясь к некоторому пределу, определяемому скоростью, с которой частицы могут отдавать энергию окружающей среде, или скоростью ре-

лаксации. Этот эффект называется *насыщением*. При этом

$$\frac{N_n}{g_n} \approx \frac{N_m}{g_m}$$

и согласно (1-44) и (1-45)

$$p_{0 \text{ изл}}^{\text{инд}} \approx p_{0 \text{ погл}}^{\text{инд}};$$

доля поглощения вынуждающего электромагнитного поля мала, среда остается пассивной, но практически прозрачной для вынуждающего поля.

Полупроводниковые кристаллические среды. Энергетические уровни в полупроводнике образуют зоны. При температуре абсолютного нуля электроны последовательно заполняют все уровни. Верхняя, целиком заполненная зона называется *валентной*, так как ее заполняют валентные электроны. Следующая за валентной, более высокая зона разрешенных состояний при температуре абсолютного нуля пуста. Эта зона *проводимости* (см. П1, п. 5). Валентная зона и зона проводимости разделены зоной запрещенных значений энергий — *запрещенной зоной*. Ширина запрещенной зоны ΔW для полупроводников составляет 10^{-3} —3 эВ.

Если температура полупроводника отлична от нуля и достаточно высока, то электроны могут получить от кристаллической решетки энергию, достаточную для перехода из валентной зоны через запрещенную зону в зону проводимости. При этом в валентной зоне образуется *дырка*, которая ведет себя как положительно заряженная частица.

Одновременно с возникновением электронно-дырочных пар в полупроводнике может происходить и обратный процесс — спонтанный переход электронов из зоны проводимости на свободные уровни валентной зоны. При этом дырка и электрон проводимости исчезают, рекомбинируют. При перескоке энергия электрона уменьшается. Разность энергий

$$\hbar\omega = W_{\text{э}} - W_{\text{д}},$$

где $W_{\text{э}}$ и $W_{\text{д}}$ — энергии исчезнувшего электрона и дырки излучаются. Так как одновременно происходит множество перескоков электронов, не связанных друг с другом, то спонтанное излучение не поляризовано, не направлено и имеет широкий спектр.

В нормальных условиях, при отсутствии каких-либо внешних воздействий, электронно-дырочные пары возникают и исчезают в результате теплового движения и спонтанного излучения. Количество электронов и дырок при тепловом равновесии зависит от температуры и ширины запрещенной зоны и определяется *распределениями Ферми* [см. (П1-4) и (П1-5)].

Возникновение электронно-дырочных пар может происходить не только под действием тепловых колебаний решетки, но и под действием других причин, в частности, в результате облучения полупроводника электромагнитным полем. Если частота поля ω удовлетворяет условию

$$\hbar\omega > \Delta W,$$

где ΔW — ширина запрещенной зоны, то при поглощении энергии поля в полупроводнике возникают электронно-дырочные пары.

Одновременно под действием вынуждающего поля происходит переход электронов проводимости в валентную зону (вынужденная рекомбинация). Разность энергий $W_{\text{э}} - W_{\text{д}}$ определяет вынужденное излучение, частота, фаза, направление распространения и поляризация которого совпадают с вынуждающим полем.

Поглощение кванта энергии электромагнитного поля может произойти лишь в том случае, если в валентной зоне имеется электрон, а в зоне проводимости дырка, разность энергий которых $W_{\text{д}} - W_{\text{э}}$ соответствует кванту электромагнитного поля. Вероятность наличия такого электрона в валентной зоне $f_{\text{эв}}$ определяется распределением Ферми (П1-4). Вероятность наличия дырки в зоне проводимости $f_{\text{дп}}$ определяется (П1-5). Вероятность одновременного наличия электрона в валентной зоне и дырки в зоне проводимости определяется произведением $f_{\text{эв}}f_{\text{дп}}$. Плотность мощности, поглощаемой при переходах под действием поля, определяется выражением

$$\rho_{\text{полгл}}^{\text{нд}} = \hbar\omega N_{\text{ф}},$$

где $N_{\text{ф}}$ — число поглощаемых фотонов.

Очевидно,

$$N_{\text{ф}} = B_{\text{вп}} \rho_{\text{э}}(\omega_{\text{э}}) \rho_{\text{д}}(\omega_{\text{д}}) f_{\text{эв}} f_{\text{дп}},$$

где $B_{\text{вп}}$ — вероятность перехода электрона из валентной зоны в зону проводимости в единицу времени; $\rho_{\text{э}}(\omega_{\text{э}})$

и $\rho_d(\omega_d)$ — плотности состояний, соответствующих энергиям W_ε и W_d .

Окончательно,

$$\rho_{0 \text{ погл}}^{\text{инд}} = \hbar \omega B_{\text{вп}} \rho_\varepsilon(\omega_\varepsilon) \rho_d(\omega_d) f_{\text{эв}} f_{\text{дп}}.$$

Аналогично плотность мощности индуцированного излучения определяется выражением

$$\rho_{0 \text{ изл}}^{\text{инд}} = \hbar \omega B_{\text{пв}} \rho_\varepsilon(\omega_\varepsilon) \rho_d(\omega_d) f_{\text{эп}} f_{\text{дв}},$$

где $B_{\text{пв}}$ — вероятность перехода электрона проводимости в валентную зону в единицу времени; $f_{\text{эп}}$ — вероятность нахождения электрона в зоне проводимости; $f_{\text{дв}}$ — вероятность наличия дырки в валентной зоне.

Так как вероятности перехода сверху вниз и снизу вверх одинаковы [$B_{\text{вп}} = B_{\text{пв}}$, см. (П1-62)], то разность плотностей мощности индуцированного поглощения и излучения равна:

$$\rho_0 = \rho_{0 \text{ погл}}^{\text{инд}} - \rho_{0 \text{ изл}}^{\text{инд}} = \hbar \omega B_{\text{вп}} \rho_\varepsilon(\omega_\varepsilon) \rho_d(\omega_d) (f_{\text{эв}} f_{\text{дп}} - f_{\text{эп}} f_{\text{дв}}).$$

Учитывая, что

$$f_{\text{дп}} = 1 - f_{\text{эп}}, \quad f_{\text{дв}} = 1 - f_{\text{эв}},$$

получаем:

$$\rho_0 = \hbar \omega B_{\text{вп}} \rho_\varepsilon(\omega_\varepsilon) \rho_d(\omega_d) (f_{\text{эв}} - f_{\text{эп}}).$$

Согласно распределению Ферми (П1-4)

$$f_{\text{эв}} > f_{\text{эп}}$$

и $\rho_0 > 0$, т. е. среда, подчиняющаяся распределению Ферми, является пассивной и поглощает распространяющееся электромагнитное поле.

Чтобы среда излучала электромагнитное поле, необходимо, чтобы выполнялось условие

$$f_{\text{эв}} < f_{\text{эп}}. \quad (1-52)$$

Это условие соответствует инверсной населенности, так как в этом случае большинство нижних уровней зоны проводимости будет занято электронами, а большинство верхних уровней валентной зоны — дырками. Условие (1-52) согласно распределению (П1-4) можно переписать в виде

$$\frac{1}{e^{\frac{W_{\text{в}} - W_{F\text{в}}}{kT}} + 1} < \frac{1}{e^{\frac{W_{\text{п}} - W_{F\text{п}}}{kT}} + 1};$$

отсюда

$$W_{F_{\Pi}} - W_{F_{\text{В}}} > W_{\Pi} - W_{\text{В}}.$$

Так как $W_{\Pi} - W_{\text{В}} > \Delta W$, где ΔW — ширина запрещенной зоны, то

$$W_{F_{\text{Э}}} - W_{F_{\text{Д}}} > \Delta W. \quad (1-53)$$

Здесь $W_{F_{\text{Э}}} = W_{F_{\Pi}}$ и $W_{F_{\text{Д}}} = W_{F_{\text{В}}}$ — квазиуровни Ферми соответственно для электронов и дырок (см. П1, п. 6). Расстояния между этими квазиуровнями Ферми больше ширины запрещенной зоны и $W_{F_{\text{Э}}}$ лежит в зоне проводимости, а $W_{F_{\text{Д}}}$ — в валентной зоне. Полупроводник, удовлетворяющий условию (1-53), называется *вырожденным* (см. П1, п. 6).

Глава вторая

НЕЛИНЕЙНЫЕ ПРОЦЕССЫ В ПАССИВНЫХ СРЕДАХ

2-1. РАСПРОСТРАНЕНИЕ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ В БЕЗГРАНИЧНЫХ ДИЭЛЕКТРИЧЕСКИХ СРЕДАХ

Как известно, диэлектрик можно представить в виде системы связанных зарядов. Диэлектрик электрически нейтрален, поэтому

$$\text{div } \mathbf{E} = 0,$$

и распространение электромагнитного поля в диэлектрической среде с потерями без сторонних источников согласно (1-9) описывается уравнением

$$\Delta \mathbf{E} - \mu_0 \sigma^{\pi} \frac{\partial \mathbf{E}}{\partial t} - \mu_0 \varepsilon_a^{\pi} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \mathbf{p}^{\pi}}{\partial t^2}. \quad (2-1)$$

Механизмы, определяющие правую часть уравнения, связаны со смещением электронов, ионов и ориентацией молекул под действием распространяющегося поля.

Электронная поляризация определяется смещением электронов под действием поля относительно ядра, которое считаем неподвижным. Предполагается, что постоянный электрический момент отсутствует; магнитная составляющая поля, распространяющегося в диэлектрике, на смещение зарядов не влияет, заряды смещаются

под действием напряженности электрического поля E . Если диэлектрик изотропный и вектор E направлен по оси x_1 , то и смещение электронов будет происходить по этой оси. Пренебрегая взаимодействием индуцированных дипольных моментов, запишем уравнение движения электрона в скалярной форме

$$m \frac{d^2 x_1}{dt^2} = F \quad (2-2)$$

или

$$m \frac{d^2 x_1}{dt^2} + mg \frac{dx_1}{dt} + m\omega_0^2 x_1 + mbx_1^2 = eE, \quad (2-3)$$

где $mg \frac{dx_1}{dt}$ — сила, аналогичная силе трения, определяемая потерями в диэлектрике; $m\omega_0^2 x_1$ — сила, аналогичная силе упругости, определяемая действием ядра на электрон. Эта сила линейна при малых смещениях электрона, т. е. при малом значении напряженности поля. При больших напряженностях поля, при больших смещениях эта сила нелинейна и определяется выражением

$$m\omega_0^2 x_1 + mbx_1^2,$$

где e и m — заряд и масса электрона; ω_0 — резонансная частота колебания электронов, соответствующая оптическому диапазону волн, $\omega_0 = \sqrt{\frac{|k|}{m}}$; k — коэффициент упругости.

Умножив обе части уравнения (2-3) на e , разделив на m , а затем просуммировав по всем N электронам, смещающимся под действием поля в единице объема, получим:

$$\frac{d^2 P}{dt^2} + 2\gamma \frac{dP}{dt} + \omega_0^2 P + \beta P^2 = \frac{e^2 N}{m} E, \quad (2-4)$$

где $P = eNx_1$ — поляризация; $2\gamma = g$ — коэффициент, аналогичный коэффициенту трения, $\beta = b/eN$.

Рассмотрим распространение электромагнитного поля, возбуждаемого током

$$J^{\text{сг}} = J_m^{\text{сг}} \cos(\omega t + \varphi),$$

в нелинейной среде с потерями, не содержащей сторон-

них источников. Для этого случая система (1-19) имеет вид:

$$\begin{aligned} \Delta \dot{E}_m(n\omega) + (n\omega)^2 \tilde{\epsilon}_a^n(n\omega) \mu_0 \dot{E}_m(n\omega) = \\ = - (n\omega)^2 \mu_0 P_m^{nl}(n\omega), \end{aligned} \quad (2-5)$$

где $n=0, \pm 1, \pm 2, \pm 3 \dots$

Если поле распространяется в виде линейно-поляризованной плоской однородной волны в направлении оси x_3 , то уравнения (2-5) можно записать в скалярной форме и они будут иметь вид:

$$\frac{\partial^2 \dot{E}_m(n\omega)}{\partial x_3^2} + (n\omega)^2 \tilde{\epsilon}_a^n(n\omega) \mu_0 \dot{E}_m(n\omega) = - (n\omega)^2 \mu_0 P_m^{nl}(n\omega), \quad (2-6)$$

где $n=0, \pm 1, \pm 2, \pm \dots$

Для решения этих нелинейных уравнений воспользуемся методом последовательных приближений (см. П2 п. 3). В линейном (нулевом) приближении вместо системы (2-6) можно записать уравнение

$$\frac{\partial^2 \dot{E}_m(\omega)}{\partial x_3^2} + \omega^2 \tilde{\epsilon}_a^n(\omega) \mu_0 \dot{E}_m(\omega) = 0,$$

решение которого описывает бегущую волну:

$$E(\omega) = E_m e^{j[\omega t - \dot{k}(\omega)x_3]} = E_m e^{-\alpha x_3} e^{j[\omega t - \beta(\omega)x_3]},$$

где $\dot{k}(\omega)$ — постоянная распространения; $\dot{k}(\omega) = \beta(\omega) - j\alpha(\omega)$.

Уравнение (2-4) в линейном приближении имеет вид:

$$\frac{d^2 P}{dt^2} + 2\gamma \frac{dP}{dt} + \omega_0^2 P = \frac{e^2 N}{m} E,$$

или в символической записи

$$-\omega^2 \dot{P} + j\omega 2\gamma \dot{P} + \omega_0^2 \dot{P} = \frac{e^2 N}{m} \dot{E};$$

отсюда

$$\dot{P}^n(\omega) = \frac{e^2 N}{m(\omega_0^2 - \omega^2 + j2\gamma\omega)} \dot{E}.$$

В обычной форме

$$P^{\pi}(\omega) = \frac{\dot{P}^{\pi}(\omega) + \dot{P}^{\pi*}(\omega)}{2} = \frac{e^2 N}{m} \frac{E_m e^{-\alpha x_3}}{(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2} \times \\ \times [(\omega_0^2 - \omega^2) \cos(\omega t - \beta x_3) + 2\gamma\omega \sin(\omega t - \beta x_3)].$$

Учитывая, что

$$a \sin \omega t + b \cos \omega t = A \sin(\omega t + \varphi),$$

где

$$A = \sqrt{a^2 + b^2}, \quad \varphi = \arctg \frac{b}{a},$$

получаем:

$$P^{\pi}(\omega) = \frac{e^2 N}{m} \frac{E_m e^{-\alpha x_3}}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2}} \sin(\omega t + \varphi), \quad (2-7)$$

где

$$\varphi = \arctg \frac{\omega_0^2 - \omega^2}{2\gamma\omega}.$$

Электрическая восприимчивость среды в линейном приближении определяется выражением

$$\dot{\chi}_3^{\pi}(\omega) = \frac{\dot{P}^{\pi}(\omega)}{e_0 E} = \frac{e^2 N [(\omega_0^2 - \omega^2) - j2\gamma\omega]}{m\varepsilon_0 [(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2]},$$

а ее диэлектрическая проницаемость

$$\tilde{\varepsilon}^{\pi} = 1 + \frac{e^2 N [(\omega_0^2 - \omega^2) - j2\gamma\omega]}{m\varepsilon_0 [(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2]}$$

или

$$\tilde{\varepsilon}^{\pi} = 1 + \dot{\chi}_3^{\pi} = \varepsilon'^{\pi} - j\varepsilon''^{\pi},$$

где

$$\varepsilon'^{\pi} = 1 + \frac{e^2 N (\omega_0^2 - \omega^2)}{m\varepsilon_0 [(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2]}; \\ \varepsilon''^{\pi} = \frac{2e^2 N \gamma \omega}{m\varepsilon_0 [(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2]} > 0;$$

отсюда найдем проводимость среды

$$\sigma^{\text{л}}(\omega) = \omega \varepsilon_a^{\text{л}}(\omega) = \frac{2e^2 N \omega^2 \gamma}{m [(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2]} > 0$$

и плотность тока в ней:

$$J^{\text{л}}(\omega) = \frac{2e^2 N \gamma \omega^2 E_m e^{-\alpha x_3}}{m [(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2]} \cos[\omega t - \beta(\omega) x_3].$$

Зависимости $\varepsilon''^{\text{л}}$ и $\varepsilon'^{\text{л}}$ от частоты поля при условии $\gamma \ll \omega$ (малые потери) приведены на рис. 2-1 и 2-2. Кри-

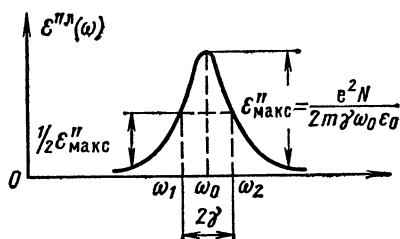


Рис. 2-1. Зависимость $\varepsilon''^{\text{л}}$ от частоты.

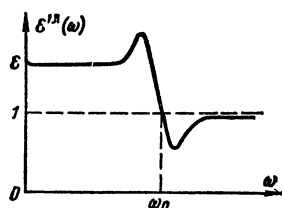


Рис. 2-2. Зависимость $\varepsilon'^{\text{л}}$ от частоты.

вая $\varepsilon''^{\text{л}}(\omega)$ определяет поглощение. Она имеет максимум при резонансе между полем и диполем ($\omega = \omega_0$). В этой точке

$$\varepsilon_{\text{макс}}^{\text{л}} = \frac{e^2 N}{2m\gamma\omega_0 \varepsilon_0}.$$

Кривая симметрична относительно ω_0 и имеет лоренцеву форму (см. П1, п. 20). Ширина кривой поглощения определяется из условия

$$\varepsilon^{\text{л}}(\omega_{1,2}) = \frac{1}{2} \varepsilon^{\text{л}}(\omega_0),$$

откуда

$$\Delta\omega = \omega_2 - \omega_1 = 2\gamma.$$

Согласно (1-47) для пассивной двухуровневой квантовой среды

$$\varepsilon_{\text{макс}}^{\text{л}} = B_{nm} \hbar (N_m - N_n) > 0, \quad N_n < N_m.$$

При малых напряженностях распространяющегося поля $\epsilon''_{\text{макс}}^{\text{л}}$ определяется разностью населенностей уровней n и m в условиях термодинамического равновесия (см. П1, п. 4). Частицы поглощают энергию распространяющегося поля и переходят с нижнего энергетического уровня на верхний. В результате населенность верхнего уровня должна увеличиваться, а нижнего уменьшаться и поглощение должно упасть. Однако релаксационные процессы восстанавливают термодинамическое равновесие, переводя избыточное (по сравнению с термодинамическим равновесием) количество частиц с верхнего уровня на нижний. При малых амплитудах распространяющегося поля релаксационные процессы идут быстрее и практически всегда существует термодинамическое равновесие. При этом поглощение, определяемое разностью $N_m - N_n$ населенностей нижнего и верхнего уровней, постоянно и не зависит от напряженности распространяющегося поля.

Для диэлектрика без потерь

$$\gamma = 0, \quad \varphi = \frac{\pi}{2}, \quad \alpha = 0$$

и поляризация согласно (2-7) определяется выражениями

$$P^{\text{л}}(\omega) = \frac{e^2 N}{m(\omega_0^2 - \omega^2)} E_m \cos[\omega t - k(\omega) x_3];$$

$$k(\omega) = \omega \sqrt{\epsilon_a^{\text{л}}(\omega) \mu_0},$$

откуда находим электрическую восприимчивость диэлектрика

$$\chi_3^{\text{л}}(\omega) = \frac{P^{\text{л}}}{\epsilon_0 E} = \frac{e^2 N}{\epsilon_0 m(\omega_0^2 - \omega^2)}$$

и его диэлектрическую проницаемость

$$\epsilon_a^{\text{л}}(\omega) = \epsilon_0 [1 + \chi_3^{\text{л}}(\omega)] = \epsilon_0 \left[1 + \frac{e^2 N}{\epsilon_0 m(\omega_0^2 - \omega^2)} \right].$$

Теперь рассмотрим процессы в нелинейной среде в первом приближении. Подставляя выражение для поля-

ризации в нулевом приближении (2-7) в уравнение (2-4), получаем:

$$\frac{d^2 P}{dt^2} + 2\gamma \frac{dP}{dt} + \omega_0^2 P = \frac{e^2 N}{m} E_m e^{-\alpha x_s} \cos [\omega t - \beta(\omega) x_s] - \\ - \frac{\beta e^4 N^2 E_m^2 e^{-2\alpha x_s}}{m^2 [(\omega_0^2 - \omega^2) + 4\gamma^2 \omega^2]} \sin^2 [\omega t - \beta(\omega) x_s + \varphi]. \quad (2-8)$$

Решение этого уравнения можно представить в виде

$$P = P^{\text{л}} + P^{\text{нл}}, \quad (2-9)$$

где $P^{\text{л}}$ определяется выражением (2-7), а $P^{\text{нл}}$ представляет собой нелинейную добавку. Учитывая

$$\sin^2 \vartheta = \frac{1 - \cos 2\vartheta}{2},$$

можно легко убедиться, что нелинейная добавка $P^{\text{нл}}$ содержит вторую гармонику и постоянную составляющую, и ее можно представить в виде

$$P^{\text{нл}} = \frac{\dot{P}(2\omega) + \dot{P}^*(2\omega)}{2} + P(0).$$

Подставляя (2-9) в (2-8), получаем уравнение для нелинейной добавки поляризации

$$-4\omega^2 [\dot{P}(2\omega) + \dot{P}^*(2\omega)] + j4\gamma\omega [\dot{P}(2\omega) + \dot{P}^*(2\omega)] + \\ + \omega_0^2 [\dot{P}(2\omega) + \dot{P}^*(2\omega)] + 2\omega_0^2 P(0) = \\ = - \frac{\beta e^4 N^2 (\dot{E} + \dot{E}^*)^2}{2m^2 [(\omega_0^2 - \omega^2) + 4\gamma^2 \omega^2]}; \quad (2-10)$$

отсюда найдем вторую гармонику поляризации

$$\dot{P}(2\omega) = - \frac{\beta e^4 N^2 \dot{E}^2 [(\omega_0^2 - 4\omega^2) - 4j\gamma\omega]}{2m^2 [(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2] [(\omega_0^2 - 4\omega^2)^2 + 16\gamma^2 \omega^2]},$$

или

$$P(2\omega) = \frac{\dot{P}(2\omega) + \dot{P}^*(2\omega)}{2} = \\ = - \frac{\beta e^4 N^2 E_m^2 e^{-2\alpha x_s}}{2m^2 [(\omega_0^2 - \omega^2) + 4\gamma^2 \omega^2] \sqrt{(\omega_0^2 - 4\omega^2)^2 + 16\gamma^2 \omega^2}} \times \\ \times \sin [2\omega t - \beta_P(2\omega) x_s + 2\varphi + \varphi'],$$

где

$$\varphi' = \operatorname{arctg} \frac{\omega_0^2 - 4\omega^2}{4\gamma\omega}, \beta_p(2\omega) = 2\beta(\omega)$$

— фазовая постоянная волны второй гармоники поляризации. Из последнего соотношения следует, что постоянная распространения второй гармоники волны поляризации равна:

$$\dot{k}_p(2\omega) = 2\dot{k}(\omega). \quad (2-11)$$

Найдем теперь электрическую восприимчивость диэлектрика:

$$\chi_3(2\omega) = - \frac{\beta e^4 N^2 [(\omega_0^2 - 4\omega^2) - 4j\gamma\omega]}{2m^2 \varepsilon_0 [(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2][(\omega_0^2 - 4\omega^2)^2 + 16\gamma^2 \omega^2]};$$

и, наконец, из уравнения (2-10) определим постоянную составляющую поляризации:

$$P(0) = - \frac{\beta e^4 N^2 E_m^2}{2m^2 \omega_0^2 [(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2]}.$$

В диэлектрике без потерь

$$\gamma = 0, \varphi = \varphi' = \pi/2, \alpha = 0,$$

поэтому

$$P(2\omega) = - \frac{\beta e^4 N^2 E_m^2}{2m^2 (\omega_0^2 - \omega^2)(\omega_0^2 - 4\omega^2)} \cos[2\omega t - \beta_p(2\omega) x_3],$$

$$\chi_3(2\omega) = - \frac{\beta e^4 N^2}{2\varepsilon_0 m^2 (\omega_0^2 - \omega^2)^2 (\omega_0^2 - 4\omega^2)};$$

$$P(0) = - \frac{\beta e^4 N^2 E_m^2}{2m^2 \omega_0^2 (\omega_0^2 - \omega^2)^2}.$$

Если возбуждающий ток имеет две составляющие с частотами ω_1 и ω_2

$$J^{\text{cr}} = J_{m1}^{\text{cr}} \cos \omega_1 t + J_{m2}^{\text{cr}} \cos \omega_2 t,$$

то поле и поляризация в линейном приближении также будут содержать две составляющие с частотами ω_1 и ω_2 :

$$E = E(\omega_1) + E(\omega_2);$$

$$P = P(\omega_1) + P(\omega_2).$$

В низшем нелинейном приближении кроме составляющих поляризации с частотами ω_1 и ω_2 появятся постоянная составляющая и составляющие с частотами $2\omega_1$, $2\omega_2$, $\omega_1 + \omega_2$ и $\omega_1 - \omega_2$.

Поляризация зависит от частоты и имеет резонансный характер. Резонанс возникает при $\omega_0 = \omega_1$, ω_2 , $2\omega_1$, $2\omega_2$, $\omega_1 \pm \omega_2$. Зависимость от частоты усиливается при приближении к одному из возможных резонансов. Если $\omega_1 - \omega_2 \approx \omega_0$ и $\omega_1 \gg \omega_0$, $\omega_2 \gg \omega_0$, удерживается лишь составляющая разностной частоты. При учете потерь поляризация при резонансе имеет конечное значение.

Поскольку электронная структура атома практически не зависит от температуры, то и электронная поляризация не зависит от температуры.

Согласно уравнению (2-6) волны нелинейной поляризации являются источниками полей: $P(0)$ — источником постоянного поля $E(0)$; $P(2\omega)$ — источником поля удвоенной частоты $E(2\omega)$.

Если поле возбуждается током

$$J^{\text{сг}} = J_{m1}^{\text{сг}} \cos \omega_1 t + J_{m2}^{\text{сг}} \cos \omega_2 t,$$

то уже в низшем нелинейном приближении получим составляющие поля

$$E(\omega_1), E(\omega_2), E(2\omega_1), E(2\omega_2), E(\omega_1 + \omega_2), E(\omega_1 - \omega_2).$$

Таким образом, при распространении поля в нелинейной среде происходит качественное его изменение в результате взаимодействия со средой. Процесс этот двусторонний; поле влияет на среду, среда в свою очередь вызывает качественное изменение поля. При этом наблюдается взаимодействие полей разных частот. Нелинейные эффекты, вызванные электронной поляризацией, проявляются вблизи резонансной частоты ω_0 , т. е. в оптическом диапазоне волн.

Нелинейные эффекты в оптическом диапазоне для большинства диэлектриков наблюдаются при напряженности поля, сравнимой с внутренними полями, действующими между заряженными частицами и имеющими величину порядка 10^8 В/м. Внешнее поле такой напряженности получено при фокусировке излучения лазера.

Поляризация $P(2\omega)$ возбуждает вторую гармонику поля $E(2\omega)$, которая удовлетворяет одному из уравнений системы (2-6)

$$\frac{\partial^2 \dot{E}(2\omega)}{\partial x_3^2} + 4\omega^2 \tilde{\epsilon}_a''(2\omega) \mu_0 \dot{E}(2\omega) = -4\omega^2 \mu_0 \dot{P}(2\omega). \quad (2-12)$$

Решение уравнения (2-12) равно сумме общего решения соответствующего однородного уравнения и частного решения неоднородного уравнения:

$$\dot{E}(2\omega) = \dot{E}_m'(2\omega) e^{i[2\omega t - \dot{k}(2\omega)x_3]} + \dot{E}_m''(2\omega) e^{i[2\omega t - \dot{k}_P(2\omega)x_3]}. \quad (2-13)$$

Первое слагаемое выражения (2-13) описывает поле бегущей волны с частотой 2ω и постоянной распространения

$$\dot{k}(2\omega) = 2\omega \sqrt{\tilde{\epsilon}_a''(2\omega) \mu_0}.$$

Второе слагаемое в выражении (2-13) представляет собой частное решение неоднородного уравнения (2-12). Оно описывает волну, постоянная распространения которой совпадает с постоянной распространения второй гармоники поляризации (2-11):

$$\dot{k}_P(2\omega) = 2\dot{k}(\omega) = 2\omega \sqrt{\tilde{\epsilon}_a''(\omega) \mu_0}.$$

Для определения амплитуды этой волны \dot{E}_m'' подставим второе слагаемое решения (2-13) в уравнение (2-12). После несложных преобразований окончательно получим:

$$\begin{aligned} \dot{E}(2\omega) = & \dot{E}_m'(2\omega) e^{i[2\omega t - \dot{k}(2\omega)x_3]} + \\ & + \frac{4\omega^2 \mu_0 \dot{P}_m(2\omega)}{\dot{k}_P^2(2\omega) - \dot{k}^2(2\omega)} e^{i[2\omega t - \dot{k}_P(2\omega)x_3]}. \end{aligned}$$

Из полученного выражения видно, что поле второй гармоники состоит из двух волн, первая из которых распространяется с постоянной распространения

$$\dot{k}(2\omega) = \beta(2\omega) - j\alpha(2\omega) = 2\omega \sqrt{\tilde{\epsilon}_a''(2\omega) \mu_0},$$

а вторая с постоянной распространения

$$\dot{k}_P(2\omega) = \beta_P(2\omega) - j\alpha_P(2\omega) = 2\omega \sqrt{\tilde{\epsilon}_a''(\omega) \mu_0} = 2\dot{k}(\omega).$$

Амплитуда второй волны определяется амплитудой волны поляризации $P(2\omega)$ и существенным образом зависит от соотношения величин $\dot{k}(2\omega)$ и $\dot{k}_P(2\omega)$.

Нелинейные эффекты, связанные с электронной поляризацией, малы. Однако по мере распространения волны и взаимодействия ее со средой эти эффекты могут накапливаться и их можно наблюдать в средах с малыми потерями. В средах с большими потерями эти эффекты могут затухнуть раньше, чем успеют развиться.

Пренебрегая потерями, получаем:

$$E(2\omega) = E_m'(2\omega) \cos[2\omega t - k(2\omega)x_3] + \\ + \frac{4\omega^2 \mu_0 P_m(2\omega)}{k_p^2(2\omega) - k^2(2\omega)} \cos[2\omega t - k_p(2\omega)x_3],$$

где

$$k(2\omega) = 2\omega \sqrt{\epsilon_a''(2\omega) \mu_0} = \beta(2\omega), \quad k_p(2\omega) = \\ = 2\omega \sqrt{\epsilon_a''(\omega) \mu_0} = \beta_p(2\omega).$$

При $k_p(2\omega) = k(2\omega)$ — амплитуда второй волны обращается в бесконечность. При учете потерь амплитуда имеет конечное значение.

Возрастание энергии поля $E(2\omega)$ определяется передачей энергии от поля основной частоты $E(\omega)$ через переизлучающую поляризацию $P(2\omega)$. Распространяющееся поле $E(\omega)$ возбуждает диполи вещества, которые создают поле, содержащее кроме составляющей основной частоты ее гармоники и постоянную составляющую. Энергообмен между полем и средой определяется выражением

$$p_{0 \text{ пол}} = \frac{1}{T} \int_0^T E \frac{\partial P}{\partial t} dt.$$

Обмен энергией между полем основной частоты $E(\omega)$ и волной поляризации $P(\omega)$ (линейный случай) согласно (1-31) определяется выражением

$$p_{0 \text{ пол}}(\omega) = \frac{\omega}{2} \operatorname{Im} \dot{E}(\omega) \dot{P}^*(\omega).$$

В случае среды с потерями

$$\dot{\chi}_s(\omega) = \chi_s'(\omega) - j\chi_s''(\omega)$$

и

$$\rho_0(\omega) = \frac{\omega \varepsilon_0 \chi_3''(\omega)}{2} E_m^2$$

определяет плотность действительной мощности, поглощаемой в диэлектрике.

При возбуждении волны поляризации удвоенной частоты $P(2\omega)$ полем основной частоты $E(\omega)$ энергообмен в среднем равен нулю (см. § 1-3). В этом случае мнимая часть нелинейной восприимчивости

$$\chi_3'(2\omega) = \chi_3'(2\omega) - j\chi_3''(2\omega)$$

не связана с поглощением энергии средой, а определяет лишь сдвиг фазы волны поляризации $P(2\omega)$ относительно поля основной частоты $E(\omega)$.

Энергообмен между полем $E(2\omega)$ и поляризацией $P(2\omega)$ согласно (1-31) определяется выражением

$$\begin{aligned} \rho_{0\text{ пол}}(2\omega) &= \omega \operatorname{Im} \dot{E}(2\omega) \dot{P}^*(2\omega) = \\ &= -\omega E_m(2\omega) P_m(2\omega) \sin \varphi_{EP}(2\omega), \end{aligned} \quad (2-14)$$

где

$$\varphi_{EP}(2\omega) = \beta(2\omega)x_3 - \beta_P(2\omega)x_3 + \varphi = \varphi_{EP}(x_3).$$

Знак выражения (2-14) периодически меняется с расстоянием. На участках, где $\rho_{0\text{ пол}}(2\omega) < 0$, энергия от волны поляризации $P(2\omega)$ (среды) передается волне поля $E(2\omega)$, там где $\rho_{0\text{ пол}}(2\omega) > 0$, энергия поля $E(2\omega)$ отдается волне поляризации $P(2\omega)$ (среде), т. е. энергия поля волны $E(2\omega)$ осциллирует в пространстве (рис. 2-3). Расстояние, на котором движение энергии происходит в одном направлении (от поля к среде или обратно), называется «длиной когерентности» и определяется изменением фазы $\varphi_{EP}(2\omega)$ на π , т. е.

$$[\beta(2\omega) - \beta_P(2\omega)] l_{\text{кор}} = \pi.$$

Поскольку $\beta_P(2\omega) = 2\beta(\omega)$, то

$$l_{\text{кор}} = \frac{\pi}{\beta(2\omega) - 2\beta(\omega)}. \quad (2-15)$$

Если в выражении (2-14)

$$\beta(2\omega) = \beta_P(2\omega), \quad (2-16)$$

$$\varphi_{EP} \neq \varphi(x_3)$$

и знак выражения (2-14) не изменяется, т. е. в среднем переход энергии осуществляется в одном направлении и при $0 < \varphi_{EP} < \pi$ энергия от волны поляризации $P(2\omega)$ переходит к волне поля $E(2\omega)$. По мере распространения энергия поля $E(2\omega)$ возрастает. Волна поляризации $P(2\omega)$ создается полем основной волны $E(\omega)$ и, таким образом, происходит перекачка энергии от волны поля

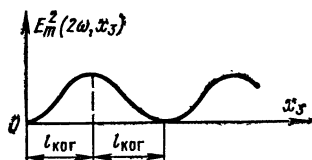


Рис. 2-3. Зависимость мощности второй гармоники от пути, пройденного волной в диэлектрике без потерь.

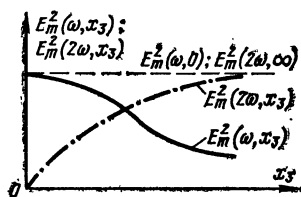


Рис. 2-4. Зависимость мощностей основной волны и второй гармоники от пути, пройденного волнами в диэлектрике без потерь, при выполнении условия волнового синхронизма.

$E(\omega)$ к волне поля $E(2\omega)$ через переизлучающую поляризацию $P(2\omega)$ (рис. 2-4). Физически это объясняется тем, что по мере распространения в создании поля $E(2\omega)$ участвуют все новые диполи, которые добавляют свое излучение к энергии распространяющейся волны $E(2\omega)$.

Учитывая, что

$$\beta_p(2\omega) = 2\beta(\omega),$$

перепишем выражение (2-16) в виде равенства

$$v(\omega) = v(2\omega), \quad (2-17)$$

называемого *условием волнового синхронизма* для второй гармоники, где $v(\omega)$ и $v(2\omega)$ — соответственно фазовые скорости волн $E(\omega)$ и $E(2\omega)$. При этом условии энергия поля основной волны $E(\omega)$ благодаря переизлучению, осуществляемому средой, по мере распространения переходит в энергию волны $E(2\omega)$.

При этом согласно выражению (2-15)

$$l_{\text{ког}} = \frac{\pi}{\beta(2\omega) - 2\beta(\omega)} = \infty.$$

Однако в изотропных средах из-за дисперсии условие волнового синхронизма (2-17) не выполняется. Это условие может быть выполнено в анизотропных средах с двойным лучепреломлением. В таких средах фазовая скорость волны зависит не только от частоты, но и от поляризации поля. При распространении электромагнитного поля в такой среде возникает обыкновенная волна, поляризация которой перпендикулярна приложенному постоянному полю в средах с управляемой анизотропией или оптической оси для одноосных кристаллов. Электрическая составляющая необыкновенной волны совпадает по направлению с постоянным полем или оптической осью кристалла. Если поле основной частоты $E(\omega)$ возбуждается таким образом, что поляризация его совпадает с поляризацией обыкновенной волны, то происходит передача энергии от основной волны к волне второй гармоники, которая является необыкновенной. Изменяя приложенное постоянное поле или направление распространения волны $E(\omega)$ в одноосном кристалле, добиваются выполнения условия волнового синхронизма, при котором передача энергии от волны $E(\omega)$ к $E(2\omega)$ происходит наиболее интенсивно. При этом длина когерентности равна бесконечности и мощность второй гармоники будет возрастать по мере распространения.

В обычных диэлектриках нелинейность мала и даже в сильных полях, создаваемых с помощью фокусировки луча в оптическом диапазоне электромагнитных волн,

$$\left| \frac{\rho_{\text{нл}}}{\rho_{\text{л}}} \right| \approx 10^{-3},$$

поэтому нелинейные эффекты будут проявляться при достаточно большой длине участка взаимодействия.

Кристаллы, используемые для получения гармоник, должны быть прозрачны для основной частоты и выделяемой гармоники, обладать высокой степенью однородности, достаточно большой нелинейностью и способностью к двойному лучепреломлению. Этим требованиям удовлетворяют кристаллы дигидрофосфата калия (KDP), дигидрофосфата аммония (ADP) и метаниобата

лития (LiNbO_2). При прохождении излучения лазера через эти кристаллы практически возможно получить вторую гармонику с эффективностью преобразования до 80%.

Аналогичным образом можно получить поле других гармоник. При наличии в возбуждающем поле двух составляющих с частотами ω_1 и ω_2 появляются комбинационные составляющие с частотами $n\omega_1 \pm m\omega_2$.

При достаточно больших напряженностях распространяющегося поля наблюдается зависимость поглощения от напряженности поля и уменьшение поглощения с увеличением напряженности. Поглощение определяется мнимой частью диэлектрической проницаемости. Согласно (1-47) в двухуровневых квантовых средах

$$\epsilon''_{a\text{макс}} = -\epsilon_0 B_{nm} \hbar (N_n - N_m).$$

При достаточно больших напряженностях поля населенность верхнего уровня возрастает, термодинамическое равновесие нарушается, а релаксационные процессы не успевают его восстановить. Разность населенностей $|N_n - N_m|$ падает, что приводит к уменьшению доли поглощения. Это типично нелинейный эффект насыщения.

Ионная поляризация связана со смещением противоположно заряженных ионов под действием электрического поля. Механизм ионной поляризации аналогичен электронной поляризации и ионная поляризация определяется аналогичными выражениями, в которых вместо массы электрона фигурирует приведенная масса молекулы

$$M = \frac{M_1 M_2}{M_1 + M_2},$$

где M_1 и M_2 — массы ионов. Вследствие большой массы ионов резонансная частота ионной поляризации значительно меньше резонансной частоты электронной поляризации и приходится на инфракрасный диапазон. В этом диапазоне наблюдается резко выраженная зависимость ионной электрической восприимчивости и поляризации от частоты (дисперсия). От температуры эти величины не зависят.

В диапазоне световых волн механизм ионной поляризации перестает действовать, так как относительно тяжелые ионы кристаллической решетки не «успевают» колебаться с частотой действующего поля.

Ориентационная поляризация связана с поворотом молекул, обладающих постоянным дипольным моментом, под действием электрического поля.

Молекулу, обладающую дипольным моментом, можно представить в виде диполя, образованного зарядами $+q$ и $-q$, находящимися на расстоянии l . В однородном электрическом поле на диполь действует пара сил (рис. 2-5), стремящаяся установить диполь по полю.

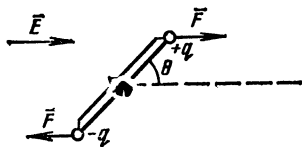


Рис. 2-5. Диполь в однородном электрическом поле.

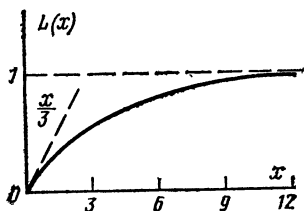


Рис. 2-6. Функция Ланжевена.

Потенциальная энергия диполя в электрическом поле равна:

$$U = -p_0 E \cos \theta,$$

где $p_0 = ql$ — постоянный дипольный момент молекулы; θ — угол между направлением дипольного момента (электрический момент диполя направлен от $-q$ к $+q$) и направлением вектора электрической напряженности.

Система всегда стремится к положению, соответствующему минимуму потенциальной энергии, т. е. стремится установиться по полю ($\theta = 0$). Однако из-за теплового движения не все диполи точно ориентируются по полю. Поляризацию вещества, связанную с ориентационным механизмом, можно представить в виде

$$P_{0 \text{ ориент}} = N p_0 \overline{\cos \theta},$$

где N — число диполей в единице объема; $\overline{\cos \theta}$ — среднее значение косинуса угла между полем и диполем, определяемое функцией Ланжевена:

$$\overline{\cos \theta} = L(x) = \text{cth } x - \frac{1}{x},$$

где $x = p_0 E / kT$; k — постоянная Больцмана. Из приведен-

ного (рис. 2-6) графика видно, что нелинейность проявляется при значениях x , соответствующих напряженности поля порядка 10^7 В/м и выше.

Ориентационная поляризация сильно зависит от температуры и возрастает с ее понижением.

Сегнетоэлектрики характеризуются наличием самопроизвольной поляризации: соседние диполи стремятся ориентироваться параллельно друг другу из-за взаимодействия между ними. Таким образом возникают целые

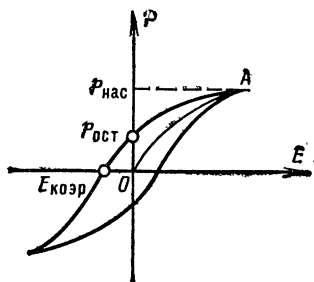


Рис. 2-7. Зависимость поляризации сегнетоэлектрика от напряженности поля.

макроскопические области (домены) с самопроизвольной поляризацией в определенном направлении. Поляризация сегнетоэлектриков зависит от приложенного поля E нелинейно. Под действием приложенного поля в сегнетоэлектрике могут наблюдаться следующие явления: изменение поляризации каждого домена по величине, изменение направления поляризации доменов, изменение размера доменов (если поляризация

доменов ориентирована параллельно или почти параллельно полю E , то они могут расти за счет смещения границ между доменами). Нелинейность сегнетоэлектриков проявляется в слабых полях.

Зависимость P от E имеет гистерезисный характер, обусловленный отставанием изменения поляризации от изменения электрического поля. Типичный вид петли гистерезиса показан на рис. 2-7. При увеличении напряженности поля поляризация вначале растет по кривой OA и стремится к насыщению $P_{н}$. При уменьшении напряженности поля до нуля диэлектрик остается в поляризованном состоянии, характеризуемом остаточной поляризацией $P_{ост}$. Чтобы свести поляризацию к нулю, необходимо приложить поле обратного направления $E_{к}$, называемое коэрцитивной силой.

Форма петли гистерезиса зависит от частоты поля и свойств сегнетоэлектрика.

Поляризация сегнетоэлектриков уменьшается с повышением температуры и при температурах, превышающих

так называемую сегнетоэлектрическую точку Кюри, са-
мопроизвольная поляризация исчезает.

Сегнетоэлектрики используются в конденсаторах, ку-
лон-вольтная характеристика которых нелинейна. Такие
конденсаторы называются варикондами.

2-2. РАСПРОСТРАНЕНИЕ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ В НЕОГРАНИЧЕННОЙ ПЛАЗМЕ

Плазма представляет собой ионизированный газ, со-
держащий заряженные частицы, нейтральные атомы и
молекулы. Обычно плазма электрически нейтральна, т. е.
на единицу ее объема приходится одинаковое число по-
ложительных и отрицательных зарядов, но в объемах,
линейные размеры которых сравнимы с величиной r_D ,
называемой радиусом Дебая, возможны флуктуации за-
ряда. Радиус Дебая определяется расстоянием, на кото-
ром происходит экранирование любого заряда плазмы
из-за группировки вокруг этого заряда противоположно
заряженных частиц. Нелинейность проявляется в плазме
при сравнительно небольших полях.

Так как плазма нейтральна, то

$$\operatorname{div} D = 0$$

и волновое уравнение (1-9) можно записать в виде

$$\Delta E - \mu_0 \epsilon_a^n \frac{\partial^2 E}{\partial t^2} = \mu_0 \frac{\partial J_{нл}}{\partial t} + \mu_0 \frac{\partial J^{ст}}{\partial t}. \quad (2-18)$$

Здесь $\epsilon_a^n = \epsilon_0 (1 - \omega_p^2 / \omega^2)$ — диэлектрическая проница-
емость, учитывающая индуцированный ток в линейном
приближении, $\omega_p = e \sqrt{\frac{n}{m \epsilon_0}}$ — собственная частота плаз-
мы, т. е. частота, с которой колеблются электроны около
своего положения равновесия после прекращения дей-
ствия поля; e — заряд электрона; m — масса электрона;
 n — концентрация плазмы; $J_{нл}$ — нелинейная плотность
индуцированного тока, определяющая нелинейные эф-
фекты.

При достаточно высокой частоте действующего поля,
когда выполняется условие

$$\omega_p \ll \omega,$$

значение диэлектрической проницаемости плазмы ϵ_a^n

почти совпадает со значением ϵ_0 . При этом волновое уравнение (2-18) будет иметь вид:

$$\Delta E - \mu_0 \epsilon_0 \frac{\partial^2 E}{\partial t^2} = \mu_0 \frac{\partial J^{nl}}{\partial t} + \mu_0 \frac{\partial J^{ct}}{\partial t}.$$

Пренебрегая движением тяжелых ионов, считаем, что ток в плазме определяется только движением электронов.

Уравнение движения электронов в плазме под действием распространяющегося поля будет иметь вид:

$$m \frac{dv}{dt} + mvv = eE + e[vB],$$

где v — скорость движения; ν — частота эффективных соударений электронов; член mvv определяет потери, поскольку при соударении с ионами или молекулой электрон передает импульс mv .

Так как на достаточно высоких частотах $\omega \gg \nu$, то членом mvv можно пренебречь. При этом получим

$$m \frac{dv}{dt} = eE + e[vB]. \quad (2-19)$$

Если сторонний источник имеет монохроматический характер

$$J = J_m^{ct} \cos \omega t,$$

то записанные для Фурье-компонент волновые уравнения согласно (1-19) будут иметь вид:

$$\Delta \dot{E}(n\omega) + (n\omega)^2 \mu_0 \epsilon_0 \dot{E}(n\omega) = jn\omega \mu_0 [J^{nl}(n\omega) + J^{ct}(n\omega)], \quad (2-20)$$

где $n=0, \pm 1, \pm 2, \pm 3, \dots$;

$$J^{ct}(n\omega) = 0 \text{ при } n \neq 1.$$

Пусть монохроматический сторонний источник возбуждает плоскую однородную волну, распространяющуюся в направлении оси x_3 , в плазме без потерь и сторонних источников. Волновое уравнение, определяющее распространяющееся в плазме поле, согласно (2-20) имеет вид:

$$\frac{\partial^2 \dot{E}(n\omega)}{\partial x_3^2} + (n\omega)^2 \mu_0 \epsilon_0 \dot{E}(n\omega) = jn\omega \mu_0 J^{nl}(n\omega). \quad (2-21)$$

В линейном приближении распространение волны характеризуется уравнением

$$\frac{\partial^2 \dot{\mathbf{E}}(\omega)}{\partial x_3^2} + \omega^2 \epsilon_0 \mu_0 \dot{\mathbf{E}}(\omega) = 0,$$

решение которого имеет вид:

$$\dot{\mathbf{E}}(\omega) = \mathbf{E}_m(\omega) e^{i[\omega t - k(\omega)x_3]},$$

где $k(\omega) = \omega \sqrt{\epsilon_0 \mu_0}$ — постоянная распространения в плазме, равная постоянной распространения в вакууме.

В обычной форме

$$\mathbf{E}(\omega) = \mathbf{E}_m(\omega) \cos [\omega t - k(\omega) x_3].$$

Пусть вектор \mathbf{E} направлен вдоль оси x_1 :

$$\mathbf{E}(\omega) = \mathbf{e}_1 E_m(\omega) \cos [\omega t - k(\omega) x_3]; \quad (2-22)$$

тогда вектор \mathbf{H} направлен вдоль оси x_2 :

$$\mathbf{H}(\omega) = \mathbf{e}_2 H_m(\omega) \cos [\omega t - k(\omega) x_3]. \quad (2-23)$$

Электрическая и магнитная составляющие поля в плазме связаны соотношением

$$\sqrt{\mu_0} H = \sqrt{\epsilon_0} E. \quad (2-24)$$

Нелинейный эффект связан с лоренцевой силой $e[\mathbf{v}\mathbf{B}]$, так как скорость электронов \mathbf{v} зависит от напряженности поля. Отбрасывая ее, получаем линейное приближение уравнения (2-19):

$$m \frac{d\mathbf{v}}{dt} = e\mathbf{E}.$$

В символической форме, учитывая, что скорость электронов изменяется по тому же закону, что и поле, получаем:

$$j\omega m \dot{\mathbf{v}} = e \dot{\mathbf{E}};$$

отсюда

$$\dot{\mathbf{v}} = -j \frac{e}{\omega m} \dot{\mathbf{E}}$$

или с учетом (2-22)

$$\mathbf{v}(\omega) = \mathbf{e}_1 \frac{e}{\omega m} E_m \sin [\omega t - k(\omega) x_3]. \quad (2-25)$$

Перепишем уравнение (2-19) с учетом (2-22) — (2-24) в проекциях по осям координат:

$$\left. \begin{aligned} m \frac{dv_1}{dt} &= eE - \frac{e}{c} v_3 E; \\ m \frac{dv_2}{dt} &= 0; \\ m \frac{dv_3}{dt} &= \frac{e}{c} v_1 E. \end{aligned} \right\} \quad (2-26)$$

Подставляя (2-25) и (2-22) в третье уравнение системы (2-26), получаем нелинейное уравнение

$$m \frac{dv_3}{dt} = \frac{e^2}{cm\omega} E_m^2(\omega) \sin[\omega t - k(\omega) x_3] \cos[\omega t - k(\omega) x_3]$$

или

$$m \frac{dv_3}{dt} = \frac{e^2}{2cm\omega} E_m^2(\omega) \sin[2\omega t - 2k(\omega) x_3],$$

интегрируя которое, получаем в низшем нелинейном приближении

$$v_3(2\omega) = - \frac{e^2}{4cm^2\omega^2} E_m^2(\omega) \cos[2\omega t - 2k(\omega) x_3]. \quad (2-27)$$

Таким образом, вследствие нелинейности, обусловленной лоренцевой силой, возникает продольное колебание электронов с удвоенной частотой и продольный ток

$$\mathbf{J}^{\text{нл}}(2\omega) = - e_3 \frac{e^3 n}{4cm^2\omega^2} E_m^2(\omega) \cos[2\omega t - 2k(\omega) x_3]. \quad (2-28)$$

Появление тока вдоль оси x_3 объясняется тем, что в поле плоской монохроматической волны под действием силы Лоренца электроны совершают сложное движение типа восьмерки. Вначале электрон начинает двигаться по направлению электрической составляющей поля, появление скорости в направлении электрического поля приводит в действие силу Лоренца, вследствие чего появляется смещение, перпендикулярное \mathbf{E} и \mathbf{H} основной волны, т. е. ток в этом направлении.

Так как скорость модулирована в направлении x_3 , то концентрация электронов n также будет модулирована в этом направлении.

Таким образом, в нелинейном приближении поперечные волны поля создают продольные волны концентрации (плотности) электронов и тока плазмы. В продольном направлении из-за модуляции плотности происходит разделение зарядов. При этом возникает поле \mathbf{E} , стре-

мещаеся восстановить нейтральность плазмы. Под действием этого поля электроны приходят в движение, равенство положительных и отрицательных зарядов в какой-то момент времени восстанавливается, но по инерции электроны продолжают движение, появляется поле противоположного направления, и движение электронов повторяется в противоположном направлении. Эти колебания электронов в плазме называются плазменными или электростатическими.

Возникающее продольное поле имеет потенциальный характер, т. е.

$$\operatorname{rot} \mathbf{E} = 0,$$

и волновое уравнение (2-18) для плазмы, лишенной сторонних источников, будет иметь вид:

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = -\frac{1}{\varepsilon_0} \frac{\partial \mathbf{J}^{\text{пл}}}{\partial t}. \quad (2-29)$$

Подставляя $\mathbf{J}^{\text{пл}}$, определяемое выражением (2-28) в уравнение (2-29) и дважды интегрируя по времени, получаем выражение для напряженности продольного поля

$$\mathbf{E}(2\omega) = e_3 \frac{e^3 n}{8cm^2 \omega^3 \varepsilon_0} E_m^2(\omega) \sin[2\omega t - 2k(\omega)x_3], \quad (2-30)$$

которое соответствует излучению диполей, ориентированных по оси x_3 .

В результате действия поля $\mathbf{E}(2\omega)$ и силы Лоренца, определяемой основной поперечной волной, электроны будут двигаться в поперечном направлении. Действительно, подставляя выражения (2-27) и (2-22) в первое уравнение системы (2-26), получаем:

$$m \frac{dv_1}{dt} = eE(\omega) + \frac{e^3}{4c^2 m^2 \omega^2} E_m^3(\omega) \cos[2\omega t - 2k(\omega)x_3] \cos[\omega t - k(\omega)x_3].$$

Учитывая, что

$$\cos 2\theta \cos \theta = \frac{1}{2} [\cos 3\theta + \cos \theta],$$

получаем:

$$\frac{dv_1(3\omega)}{dt} = \frac{e^3}{8c^2 m^3 \omega^2} E_m^3(\omega) \cos[3\omega t - 3k(\omega)x_3],$$

откуда

$$v_1(3\omega) = \frac{e^3}{24c^2 m^3 \omega^3} E_m^3(\omega) \sin[3\omega t - 3k(\omega)x_3]. \quad (2-31)$$

Составляющая скорости электронов (2-31) образует поперечный ток

$$\mathbf{J}^{\text{пл}}(3\omega) = e_1 \frac{e^4 n}{24c^2 m^3 \omega^3} E_m^3(\omega) \sin[3\omega t - 3k(\omega)x_3],$$

который согласно (2-18) является источником излучения поперечного поля $\mathbf{E}(3\omega)$. Это поле при выполнении условия волнового синхронизма может достигнуть значительной величины за счет перекачки энергии поля $\mathbf{E}(\omega)$.

В результате действия поля $\mathbf{E}(3\omega)$ и силы Лоренца, определяемой основной поперечной волной, электроны будут двигаться в продольном направлении. Подставляя (2-22) и (2-31) в третье уравнение системы (2-26), получаем:

$$\frac{dv_3}{dt} = \frac{e^4}{24c^3 m^4 \omega^4} E_m^4(\omega) \sin[3\omega t - 3k(\omega)x_3] \cos[\omega t - k(\omega)x_3].$$

Учитывая, что

$$\sin 3\theta \cos \theta = \frac{1}{2} [\sin 4\theta - \sin 2\theta],$$

получаем $v_3 = v_3(4\omega)$ и соответственно $\mathbf{J}^{\text{пл}} = e_3 \mathbf{J}^{\text{пл}}(4\omega)$, $\mathbf{E} = e_3 \mathbf{E}(4\omega)$.

Таким образом, поперечные поля $\mathbf{E}(\omega)$, $\mathbf{E}(3\omega)$, $\mathbf{E}(5\omega)$ и т. д. и продольные поля $\mathbf{E}(2\omega)$, $\mathbf{E}(4\omega)$, $\mathbf{E}(6\omega)$ и т. д. не являются независимыми. Они взаимодействуют друг с другом через среду.

2-3. РАСПРОСТРАНЕНИЕ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ В НЕОГРАНИЧЕННОЙ МАГНИТНОЙ СРЕДЕ

В магнитной среде при отсутствии сторонних источников распространение электромагнитного поля определяется согласно (1-19) уравнением

$$\Delta \dot{\mathbf{E}}_m(n\omega) + (n\omega)^2 \epsilon_0 \tilde{\mu}_a^{\text{пл}}(n\omega) \dot{\mathbf{E}}_m(n\omega) = j n \omega \mu_0 \text{rot } \dot{\mathbf{M}}^{\text{пл}}(n\omega). \quad (2-32)$$

Свойства магнитных сред во многих отношениях формально сходны со свойствами диэлектрических (см. § 2-1). Наведенная намагниченность (диамагнетизм) аналогична наведенной поляризации. Существуют ато-

мы и молекулы с постоянными магнитными диполями, как атомы и молекулы с постоянными электрическими диполями. Имеются магнитные материалы, обладающие аналогично сегнетоэлектрикам спонтанной намагниченностью.

Особый интерес представляет распространение электромагнитного поля при наличии внешнего постоянного магнитного поля.

Парамагнитная среда характеризуется наличием частиц (атомы, молекулы, ионы), обладающих постоянным магнитным моментом. При отсутствии внешнего магнитного поля ориентация этих моментов хаотична и результирующая намагниченность равна нулю. При наличии внешнего магнитного поля элементарные магнитные моменты ориентируются по полю и намагниченность не равна нулю. При этом магнитные моменты частиц ориентируются под действием поля независимо друг от друга.

Намагниченность зависит от температуры и при постоянном магнитном поле уменьшается с увеличением температуры, так как тепловое движение частиц разрушает упорядоченность ориентации моментов. При фиксированной температуре намагниченность растет с увеличением напряженности внешнего магнитного поля, стремясь к пределу (насыщению), когда все элементарные моменты выстраиваются в одном направлении.

Магнитные свойства парамагнетиков связаны с наличием нескомпенсированных магнитных моментов электронов незаполненных оболочек (П1, п. 2). Примером парамагнетиков являются ионы группы железа Fe^{3+} , Cr^{3+} , Mn^{2+} , Co^{2+} ... с незаполненной внешней электронной оболочкой и ионы редкоземельных элементов с незаполненной внутренней оболочкой.

Магнитный момент частицы \mathbf{p}_m связан с механическим моментом $\mathbf{K}_{\text{мех}}$ соотношением

$$\mathbf{p}_m = -\gamma \mathbf{K}_{\text{мех}}, \quad (2-33)$$

где γ — гиромагнитное отношение.

В магнитном поле \mathbf{H} на частицу с магнитным моментом действует момент сил $\mu_0 [\mathbf{p}_m \mathbf{H}]$, стремящийся установить \mathbf{p}_m по полю. Изменение момента количества движения определяется выражением

$$\frac{d\mathbf{K}_{\text{мех}}}{dt} = \mu_0 [\mathbf{p}_m \mathbf{H}].$$

С учетом (2-33) получим

$$\frac{d\mathbf{p}_M}{dt} = -\mu_0\gamma [\mathbf{p}_M\mathbf{H}].$$

Усредняя по единице объема без учета взаимодействия частиц друг с другом и с окружающей средой, получаем уравнение движения вектора намагниченности \mathbf{M} без учета потерь

$$\frac{d\mathbf{M}}{dt} = -\mu_0\gamma [\mathbf{M}\mathbf{H}]. \quad (2-34)$$

При наличии лишь внешнего постоянного магнитного поля уравнение имеет вид:

$$\frac{d\mathbf{M}}{dt} = -\mu_0\gamma [\mathbf{M}\mathbf{H}_0].$$

Согласно этому уравнению вектор намагниченности \mathbf{M} прецессирует вокруг направления постоянного маг-

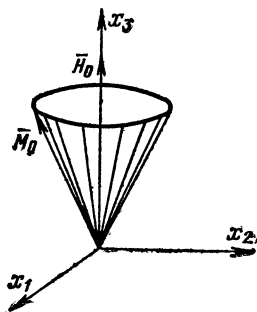


Рис. 2-8. Незатухающая прецессия вектора намагниченности \mathbf{M} в постоянном магнитном поле \mathbf{H}_0 .

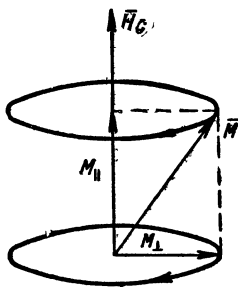


Рис. 2-9. Продольная M_{\parallel} и поперечная M_{\perp} составляющие магнитного момента \mathbf{M} .

нитного поля \mathbf{H}_0 с угловой скоростью $\omega_0 = \mu_0\gamma H_0$ (ларморова частота прецессии) и неизменным углом прецессии (рис. 2-8).

Намагниченность \mathbf{M} представляет собой сумму магнитных моментов всех частиц в единице объема. Эти моменты взаимодействуют друг с другом и с окружающей средой.

При отсутствии магнитного поля парамагнетик находится в состоянии теплового равновесия, при котором суммарный магнитный момент равен нулю. В магнитном поле происходит ориентация магнитных моментов частиц: часть из них устанавливается по полю, часть против поля. Результирующий момент \mathbf{M} в магнитном поле не равен нулю и соответствует новому равновесному состоянию. Процесс перехода от одного равновесного состояния к другому называется релаксацией. Установление нового равновесного состояния происходит не сразу, а по истечении некоторого времени — времени релаксации. При изменении магнитного поля \mathbf{H}_0 магнитный момент \mathbf{M} изменяется и парамагнетик переходит в новое равновесное состояние. Во время релаксации вектор \mathbf{M} прецессирует вокруг постоянного магнитного поля \mathbf{H}_0 (рис. 2-9) и его можно представить как сумму

$$\mathbf{M} = \mathbf{M}_{\parallel} + \mathbf{M}_{\perp},$$

где \mathbf{M}_{\parallel} и \mathbf{M}_{\perp} — составляющие магнитного момента \mathbf{M} , параллельная и перпендикулярная постоянному магнитному полю \mathbf{H}_0 соответственно.

В процессе релаксации происходит изменение составляющих \mathbf{M}_{\parallel} и \mathbf{M}_{\perp} . Вектор \mathbf{M} в равновесном состоянии направлен по полю, т. е. $\mathbf{M}_{\parallel} \approx \mathbf{M}$, а $\mathbf{M}_{\perp} = 0$. Изменение этих составляющих происходит с различной скоростью, поэтому различают продольную и поперечную релаксации.

Продольная релаксация (изменение \mathbf{M}_{\parallel}) связана с выстраиванием микроскопических магнитных моментов частиц по направлению поля. Продольная релаксация приводит к установлению равновесного для данной температуры и значения магнитного поля состояния. При этом большая часть магнитных моментов частиц направлена по полю, меньшая — против поля, что и определяет значение магнитного момента \mathbf{M} в равновесном состоянии. Равновесное распределение частиц по энергетическим уровням подчиняется распределению Больцмана (П1, п. 4), согласно которому наибольшее число частиц находится на нижнем уровне. Энергия частицы в магнитном поле связана с ориентацией магнитного момента. Минимуму энергии соответствует максимальная проекция магнитного момента на направление поля. Релаксационный процесс, приводящий систему в равновесное состояние, способствует увеличению частиц на нижнем

уровне и «выстраиванию» микроскопических магнитных моментов по полю.

Если равновесное распределение частиц нарушается, например, при поглощении энергии электромагнитного поля частоты $\omega = \omega_0$, частицы переходят на верхние уровни, ориентация их магнитных моментов изменяется. Релаксационный процесс возвращает частицы вниз, восстанавливая равновесие. При этом частицы отдают часть своей энергии в виде тепла окружающей среде. Так как магнитный момент парамагнетика связан со спинами частиц, а окружающая среда представляет собой атомы кристаллической решетки, то релаксация называется *спин-решеточной*. Пусть постоянное магнитное поле \mathbf{H}_0 направлено по оси x_3 , тогда изменение продольной составляющей вектора намагниченности $M_{\parallel} = M_z$ определяется уравнением, аналогичным (1-49):

$$\frac{d(M_z - M_0)}{dt} = -\frac{1}{T_1}(M_z - M_0), \quad (2-35)$$

где M_0 — равновесное значение, которому соответствует M_z после переориентации в магнитном поле; T_1 — время спин-решеточной или продольной релаксации.

Поперечная релаксация связана с прецессией магнитных моментов частиц вокруг направления постоянного магнитного поля \mathbf{H}_0 . Синхронная прецессия микроскопических моментов приводит к прецессии суммарного момента \mathbf{M} . Прецессию каждой частицы можно характеризовать вращением поперечной составляющей микроскопического момента в плоскости, перпендикулярной магнитному полю, прецессию вектора \mathbf{M} — вращением составляющей \mathbf{M}_{\perp} . Однако синхронность вращения, а следовательно, и фазировка моментов через некоторое время после начала прецессии нарушаются из-за разброса частот прецессии различных частиц. Дело в том, что частота прецессии частиц определяется не только внешним полем H_0 , но и полем, создаваемым соседними частицами,

$$\omega_0 = \mu_0 \gamma (H_0 + H_{\text{внутр}});$$

это поле $H_{\text{внутр}}$ достаточно велико (около $8 \cdot 10^4$ А/м) и неоднородно. Из-за неоднородности поля наблюдается значительный разброс в частотах прецессии различных частиц $\Delta\omega_0$.

Разброс частот определяет расфазировку магнитных моментов частиц. Время этой расфазировки (примерно $1/\Delta\omega_0$) и есть время поперечной релаксации. Так как при поперечной релаксации взаимодействуют спины соседних частиц, то релаксация называется *спин-спиновой*. При этом общая энергия всех парамагнитных частиц сохраняется, происходит лишь обмен энергией друг с другом.

В равновесном состоянии $M_{\perp} = 0$ и изменение этой составляющей определяется уравнениями

$$\left. \begin{aligned} \frac{dM_1}{dt} &= -\frac{M_1}{T_2}; \\ \frac{dM_2}{dt} &= -\frac{M_2}{T_2}, \end{aligned} \right\} \quad (2-36)$$

где T_2 — время поперечной или спин-спиновой релаксации.

Для твердых парамагнетиков

$$T_1 \ll T_2,$$

для жидкостей и газов

$$T_1 \approx T_2.$$

Движение вектора намагниченности \mathbf{M} в магнитном поле как под действием момента (2-34), так и под влиянием релаксационных процессов (2-35) и (2-36) определяется уравнением

$$\frac{d\mathbf{M}}{dt} = -\mu_0\gamma(\mathbf{M}\mathbf{H}) - \frac{M_1}{T_2}\mathbf{e}_1 - \frac{M_2}{T_2}\mathbf{e}_2 - \frac{M_3 - M_0}{T_1}\mathbf{e}_3. \quad (2-37)$$

Пусть поле \mathbf{H} представляет собой сумму

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{H}(\omega), \quad (2-38)$$

где \mathbf{H}_0 — постоянное магнитное поле, направленное вдоль оси x_3 ; $\mathbf{H}(\omega)$ — высокочастотное монохроматическое поле, перпендикулярное \mathbf{H}_0 , амплитуда этого поля $H_m(\omega) \ll H_0$.

Решение уравнения (2-37) будем искать методом последовательных приближений (П2, п. 3) в виде

$$\mathbf{M} = \mathbf{M}_0 + \mathbf{M}^{\perp}(\omega) + \mathbf{M}_1^{\text{пл}} + \mathbf{M}_2^{\text{пл}} + \dots, \quad (2-39)$$

где $\mathbf{M}_0 = \chi_{\text{м0}} \mathbf{H}_0$ — намагниченность в постоянном магнитном поле; $\chi_{\text{м0}}$ — магнитная восприимчивость в постоянном магнитном поле, т.е. статическая восприимчивость; $\mathbf{M}^{\text{л}}(\omega)$ — линейная часть намагниченности, определяемая переменным полем; $\mathbf{M}_1^{\text{нл}}, \mathbf{M}_2^{\text{нл}} \dots$ — нелинейные добавки соответствующего приближения.

Обычно

$$|\mathbf{M}_0| \gg |\mathbf{M}^{\text{л}}| \gg |\mathbf{M}_1^{\text{нл}}| \gg |\mathbf{M}_2^{\text{нл}}|.$$

Учитывая, что из-за параллельности \mathbf{H}_0 и \mathbf{M}_0

$$[\mathbf{M}_0 \mathbf{H}_0] = 0,$$

а при насыщении, когда высокочастотное поле мало по сравнению с постоянным,

$$M_3 - \dot{M}_0 \approx 0,$$

уравнение (2-37) с учетом (2-38) и (2-39) в линейном приближении будет иметь вид

$$\begin{aligned} \frac{d\mathbf{M}^{\text{л}}(\omega)}{dt} = & -\mu_0 \gamma [\mathbf{M}^{\text{л}}(\omega) \mathbf{H}_0] - \mu_0 \gamma [\mathbf{M}_0 \mathbf{H}(\omega)] - \\ & - \frac{M_1^{\text{л}}(\omega)}{T_2} \mathbf{e}_1 - \frac{M_2^{\text{л}}(\omega)}{T_2} \mathbf{e}_2. \end{aligned}$$

Перепишем это уравнение в символической форме в проекциях на оси декартовой системы координат:

$$\begin{aligned} j\omega \dot{M}_1^{\text{л}}(\omega) = & -\mu_0 \gamma [\dot{M}_2^{\text{л}}(\omega) H_0 - M_0 \dot{H}_2(\omega)] - \frac{\dot{M}_1^{\text{л}}(\omega)}{T_2}; \\ j\omega \dot{M}_2^{\text{л}}(\omega) = & -\mu_0 \gamma [M_0 \dot{H}_1(\omega) - \dot{M}_1^{\text{л}}(\omega) H_0] - \frac{M_2^{\text{л}}(\omega)}{T_2}, \\ j\omega \dot{M}_3^{\text{л}}(\omega) = & 0. \end{aligned}$$

Разрешая относительно составляющих намагниченности, получаем:

$$\left. \begin{aligned} \dot{M}_1^{\text{л}}(\omega) &= \dot{\chi}^{\text{м}} \dot{H}_1(\omega) - j\dot{\chi}_{\alpha}^{\text{м}} \dot{H}_2(\omega); \\ M_2^{\text{л}}(\omega) &= j\dot{\chi}_{\alpha}^{\text{м}} \dot{H}_1(\omega) + \dot{\chi}^{\text{м}} \dot{H}_2(\omega); \\ M_3^{\text{л}}(\omega) &= 0, \end{aligned} \right\} \quad (2-40)$$

где

$$\left. \begin{aligned} \dot{\chi}^{\text{м}} &= \chi_{\text{м}0} \frac{\omega_0^2}{\left(\omega_0^2 - \omega^2 + \frac{1}{T_2^2}\right) + 2j \frac{\omega}{T_2}}; \\ \dot{\chi}_{\alpha}^{\text{м}} &= -\chi_{\text{м}0} \frac{\omega_0 \left(\omega - j \frac{1}{T_2}\right)}{\left(\omega_0^2 - \omega^2 + \frac{1}{T_2^2}\right) + 2j \frac{\omega}{T_2}}; \\ \chi_{\text{м}0} &= \frac{M_0}{H_0}; \quad \omega_0 = \mu_0 \gamma H_0. \end{aligned} \right\} \quad (2-41)$$

Уравнения (2-40) могут быть представлены в тензорной форме:

$$\dot{M}_i^{\text{н}}(\omega) = \dot{\chi}_{ik}^{\text{м}}(\omega) \dot{H}_k(\omega),$$

где $\dot{\chi}_{ik}^{\text{м}}$ — тензор динамической магнитной восприимчивости.

Переменная составляющая вектора магнитной индукции определяется выражением

$$B_i(\omega) = \mu_0 \mu_{ik}(\omega) H_k(\omega),$$

где μ_{ik} — тензор динамической магнитной проницаемости

$$\mu_{ik} = \delta_{ik} + \chi_{ik}(\omega) \quad \delta_{ik} = \begin{cases} 1, & i = k, \\ 0, & i \neq k. \end{cases} \quad (2-42)$$

Дисперсия среды определяется действительной частью динамической магнитной проницаемости $\tilde{\mu}_a$, поглощение — мнимой. Согласно (2-42) и (2-41), разделяя действительную и мнимую части, получаем:

$$\tilde{\mu}_a = \mu'_a - j\mu''_a,$$

где

$$\mu'_a = \mu_0 \left[1 + \chi_{\text{м}0} \frac{\omega_0^2 \left(\omega_0^2 - \omega^2 + \frac{1}{T_2^2}\right)}{\left(\omega_0^2 - \omega^2 + \frac{1}{T_2^2}\right)^2 + 4 \frac{\omega^2}{T_2^2}} \right]; \quad (2-43)$$

$$\mu_a'' = \mu_0 \chi_{m0} \frac{\omega_0^2 2 \frac{\omega}{T_2}}{\left(\omega_0^2 - \omega^2 + \frac{1}{T_2^2}\right)^2 + 4 \frac{\omega^2}{T_2^2}}. \quad (2-44)$$

Максимальное значение мнимой части динамической проницаемости (2-44) соответствует резонансной частоте:

$$\omega_p = \omega_0 \sqrt{1 + \frac{1}{\omega_0^2 T_2^2}}.$$

При малых потерях

$$\frac{1}{\omega_0^2 T_2^2} \ll 1 \text{ и } \omega_p \approx \omega_0;$$

при этом

$$\left. \begin{aligned} \mu_a' &= \mu_0 \left[1 + \chi_{m0} \frac{\omega_0^2 (\omega_p^2 - \omega^2)}{(\omega_p^2 - \omega^2)^2 + 4 \frac{\omega^2}{T_2^2}} \right]; \\ \mu_a'' &= \mu_0 \chi_{m0} \frac{\omega_0^2 2 \frac{\omega}{T_2}}{(\omega_p^2 - \omega^2)^2 + 4 \frac{\omega^2}{T_2^2}}. \end{aligned} \right\} \quad (2-45)$$

Кривые зависимости μ' и μ'' от частоты при условии $1/T_2 \ll \omega$ (малые потери) приведены на рис. 2-10 и рис. 2-11. Кривая $\mu''(\omega)$ достигает максимума при $\omega = \omega_p$ и симметрична относительно $\omega = \omega_p$, т. е. имеет лоренцеву форму (П1, п. 20). Ширина резонансной кривой поглощения определяется по половинному уровню, т. е. из условия

$$\mu''(\omega_{1,2}) = \frac{1}{2} \mu''(\omega_p),$$

откуда

$$\Delta\omega = \omega_2 - \omega_1 = 2 \frac{1}{T_2}, \quad (2-46)$$

т. е. при малой амплитуде высокочастотного поля ширина линии поглощения определяется временем поперечной релаксации.

Резонансное поглощение высокочастотной энергии при наложении постоянного магнитного поля в средах, содержащих парамагнитные частицы, постоянный магнитный момент которых связан с нескомпенсированными магнитными моментами электронов, называется *электронным парамагнитным резонансом*.

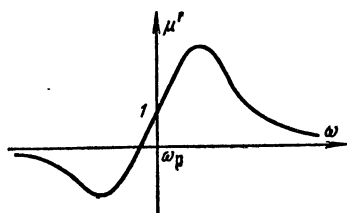


Рис. 2-10. Зависимость μ' от частоты для парамагнетика с потерями.

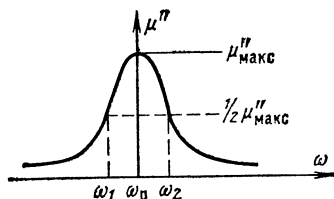


Рис. 2-11. Зависимость μ'' от частоты для парамагнетика с потерями.

В постоянном магнитном поле H_0 (П1, п. 2) происходит пространственное квантование полного момента атома J , и энергетический уровень, соответствующий определенному значению J , расщепляется на $2J+1$ подуровень с разными значениями проекции J на направление поля H_0 , определяемыми значениями M_J (эффект Зеемана).

Переходы, удовлетворяющие правилу отбора (см. П1, п. 7),

$$\Delta J = 0, \Delta M_J = \pm 1,$$

соответствуют СВЧ диапазону.

Расстояние по энергии между уровнями, удовлетворяющими правилу отбора, определяется выражением

$$\Delta W = g\mu_B H_0,$$

где g — фактор Ланде; μ_B — магнетон Бора: $\mu_B = \frac{e\hbar}{2m}$; e — заряд электрона; m — его масса.

Согласно (1-48) для двухуровневой квантовой среды

$$\mu''_{\text{макс}} = B_{nm} \hbar \Delta N = B_{nm} \hbar (N_m - N_n) > 0.$$

При малой амплитуде высокочастотного поля $\mu''_{\text{макс}}$ и, следовательно, поглощаемая мощность зависят от разности населенности уровней в состоянии термодинами-

ческого равновесия (П1, п. 4). Поглощение тем больше, чем больше частиц находится на нижнем уровне и чем меньше на верхнем.

При поглощении энергии частицы с нижнего уровня переходят на верхний и разность $N_m - N_n$, а следовательно, и поглощение должны уменьшаться. Однако при малой амплитуде высокочастотного поля релаксационные процессы успевают восстанавливать термодинамическое равновесие и поглощение, определяемое разностью населенностей нижнего и верхнего уровней, постоянно.

Если амплитуда высокочастотного поля достаточно велика, то релаксационные процессы не успевают восстанавливать термодинамическое равновесие между уровнями. Населенность верхнего уровня растет, а нижнего падает, и при очень большой амплитуде высокочастотного поля происходит выравнивание населенностей (насыщение), что приводит к уменьшению поглощения. Это нелинейный эффект.

Другие нелинейные эффекты (выпрямление, умножение частоты) можно исследовать, определяя следующие члены выражения (2-39): M_1^{nl} , M_2^{nl} и т. д.

Ферромагнитная среда характеризуется наличием микроскопических областей (доменов), обладающих самопроизвольной намагниченностью в определенном направлении. Магнитные свойства ферромагнетиков связаны с наличием нескомпенсированных магнитных моментов электронов внешних оболочек.

Особенности явлений в ферромагнетиках связаны с наличием сложной системы сильновзаимодействующих электронов. При этом в ферромагнетиках создается большая результирующая намагниченность и сильное внутреннее магнитное поле.

Ферромагнитные полупроводниковые среды — ферриты обладают малыми потерями и в них возможно распространение электромагнитного поля. Это распространение описывается уравнением (2-32).

Движение вектора намагниченности ферромагнетика может быть определено уравнением Ландау — Лифшица

$$\frac{d\mathbf{M}}{dt} = -\mu_0\gamma[\mathbf{M}\mathbf{H}] - \alpha \frac{\mu_0\gamma}{M^2}[\mathbf{M}[\mathbf{M}\mathbf{H}]], \quad (2-47)$$

где α — безразмерный параметр затухания.

Второй (релаксационный) член в правой части уравнения (2-47) характеризует вращательный момент, стремящийся возратить намагниченность в положение равновесия.

В постоянном магнитном поле \mathbf{H}_0 вектор намагниченности \mathbf{M} прецессирует вокруг направления \mathbf{H}_0 , оставаясь постоянным по величине. Действительно, умножая обе части уравнения (2-47) скалярно на \mathbf{M} , получаем:

$$\frac{dM^2}{dt^2} = 0,$$

т. е. $M = \text{const.}$

Это возможно, если конец вектора \mathbf{M} описывает спираль на поверхности сферы (рис. 2-12).

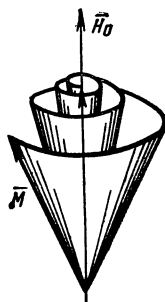


Рис. 2-12. Затухающая прецессия вектора намагниченности \mathbf{M} в постоянном магнитном поле \mathbf{H}_0 .

Пусть кроме постоянного магнитного поля $\mathbf{H}_0 = \mathbf{e}_3 H_0$ действует монохроматическое поле $\mathbf{H}(\omega)$ в плоскости, перпендикулярной \mathbf{H}_0 , т. е. результирующее поле имеет вид:

$$\mathbf{H} = \mathbf{H}_0 + \mathbf{H}(\omega), \quad (2-48)$$

где

$$\mathbf{H}(\omega) = \mathbf{e}_1 H_1(\omega) + \mathbf{e}_2 H_2(\omega), \\ H_m(\omega) \ll H_0.$$

Если $H_0 > H_s$, где H_s — напряженность поля насыщения, то намагниченность может быть описана выражением

$$\mathbf{M} = \mathbf{M}_0 + \mathbf{M}^l(\omega) + \mathbf{M}_1^{\text{нл}} + \mathbf{M}_2^{\text{нл}} + \dots, \quad (2-49)$$

где $\mathbf{M}_0 = \mathbf{e}_3 M_s$, M_s — намагниченность насыщения,

$$|\mathbf{M}_0| \gg |\mathbf{M}^l(\omega)| \gg |\mathbf{M}_1^{\text{нл}}| \gg |\mathbf{M}_2^{\text{нл}}| \gg \dots \quad (2-50)$$

Решая уравнение (2-47) с учетом (2-48) и (2-49) с помощью метода последовательных приближений, в линейном приближении получаем:

$$M_i^n(\omega) = \dot{\chi}_{ik}^m(\omega) \dot{H}_k(\omega), \quad (2-51)$$

где $\chi_{ik}^m(\omega)$ — тензор динамической магнитной восприимчивости

$$\left. \begin{aligned} \chi_{ik}^m(\omega) &= \left\| \begin{array}{ccc} \dot{\chi}_m - j\dot{\chi}_\alpha^m & 0 \\ j\dot{\chi}_\alpha^m & \chi_m & 0 \\ 0 & 0 & 0 \end{array} \right\|; \\ \dot{\chi}_m(\omega) &= \chi_{0m} \frac{\omega_p^2 + j\alpha\omega_0\omega}{\omega_p^2 - \omega^2 + 2j\alpha\omega_0\omega}; \\ \dot{\chi}_\alpha^m(\omega) &= \chi_{0m} \frac{\omega_0\omega}{\omega_p^2 - \omega^2 + 2j\alpha\omega_0\omega}, \end{aligned} \right\} \quad (2-52)$$

где $\omega_0 = \mu_0 \gamma H_0$, $\chi_{0m} = M_s/H_0$, $\omega_p = (1 + \alpha^2)^{1/2} \omega_0$.

Разделяя действительные и мнимые части второго выражения (2-52) для динамической восприимчивости, получаем:

$$\dot{\chi}_m = \chi'_m - j\chi''_m,$$

где

$$\chi'_m = \chi_{0m} \frac{\omega_p^2 (\omega_p^2 - \omega^2) + 2\alpha^2 \omega_0^2 \omega^2}{(\omega_p^2 - \omega^2)^2 + 4\alpha^2 \omega_0^2 \omega^2}; \quad (2-53)$$

$$\chi''_m = \chi_{0m} \frac{\alpha \omega_0 \omega (\omega_p^2 + \omega^2)}{(\omega_p^2 - \omega^2)^2 + 4\alpha^2 \omega_0^2 \omega^2}. \quad (2-54)$$

Тензор динамической проницаемости определяется выражением

$$\mu_{ik} = \delta_{ik} + \chi_{ik}^m. \quad (2-55)$$

Действительная часть динамической проницаемости согласно (2-55) и (2-53) равна:

$$\mu' = 1 + \chi_{0m} \frac{\omega_p^2 (\omega_p^2 - \omega^2) + 2\alpha^2 \omega_0^2 \omega^2}{(\omega_p^2 - \omega^2)^2 + 4\alpha^2 \omega_0^2 \omega^2}, \quad (2-56)$$

мнимая часть согласно (2-53)

$$\mu'' = \chi_{0m} \frac{\alpha \omega_0 \omega (\omega_p^2 + \omega^2)}{(\omega_p^2 - \omega^2)^2 + 4\alpha^2 \omega_0^2 \omega^2}. \quad (2-57)$$

Зависимость μ' от частоты характеризует дисперсионные свойства среды. Характерный вид кривой $\mu'(\omega)$ приведен на рис. 2-10. Величина μ'' определяет поглощение. Типичный вид зависимости μ'' от частоты показан на рис. 2-11. При резонансе ($\omega = \omega_p$) μ'' достигает максимального значения:

$$\mu''_{\text{макс}}(\omega_p) = \chi_{0m} \frac{\omega_p}{2\alpha\omega_0}.$$

Резонансная кривая симметрична относительно резонансной частоты ω_p , т. е. имеет лоренцеву форму (П1, п. 20). Ширина резонансной кривой определяется по половинному уровню.

Резонансное поглощение высокочастотной энергии в ферромагнитных средах при наложении внешнего постоянного магнитного поля называется *ферромагнитным резонансом*.

Поглощение определяется переходами между соседними уровнями, возникающими при расщеплении уровней атома в постоянном магнитном поле (эффект Зеемана).

Ферромагнетики ведут себя как линейные среды лишь при небольшой напряженности высокочастотного поля. При больших амплитудах поля возникает нелинейная зависимость между составляющими переменного поля и намагниченности.

Условие постоянства вектора намагниченности \mathbf{M} при $H_0 > H_s$ можно записать в виде

$$M_1^2 + M_2^2 + M_3^2 = M^2 = M_s^2,$$

где M_s — намагниченность насыщения, причем $\mathbf{M}_s = \mathbf{e}_3 M_s$ (рис. 2-13).

Так как обычно $M_1, M_2 \ll M_s$, то

$$M_3 \approx M_s - \frac{1}{2M_s} (M_1^2 + M_2^2). \quad (2-58)$$

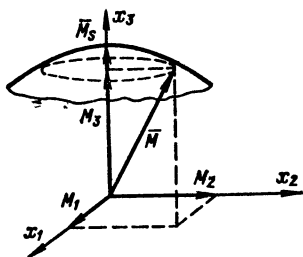


Рис. 2-13. К определению нелинейной намагниченности.

Пусть высокочастотное поле имеет составляющие

$$\left. \begin{aligned} H_1(\omega) &= H_{m1}(\omega) \cos \omega t; \\ H_2(\omega) &= H_{m2}(\omega) \cos(\omega t + \varphi). \end{aligned} \right\} \quad (2-59)$$

Для упрощения расчетов рассмотрим среду без потерь. При этом уравнение движения вектора намагниченности (2-47) будет иметь вид:

$$\frac{d\mathbf{M}}{dt} = -\mu_0 \gamma [\mathbf{M}\mathbf{H}], \quad (2-60)$$

а его решения в линейном приближении согласно (2-51) и с учетом (2-59)

$$\left. \begin{aligned} M_1(\omega) &= \chi_m H_{m1}(\omega) \cos \omega t + \chi_\alpha^m H_{m2}(\omega) \sin(\omega t + \varphi); \\ M_2(\omega) &= -\chi_\alpha^m H_{m1}(\omega) \sin \omega t + \chi_m H_{m2}(\omega) \cos(\omega t + \varphi); \\ M_3(\omega) &= 0. \end{aligned} \right\} \quad (2-61)$$

Первое нелинейное приближение получим, подставив (2-61) в (2-58). После несложных преобразований имеем:

$$M_3 = M_s - \Delta M_3 + M_{m3}(2\omega) \cos(2\omega t + \varphi),$$

где

$$\begin{aligned} \Delta M_3 &= \frac{1}{4M_s} \{ (\chi_m^2 + \chi_\alpha^{m2}) [H_{m1}^2(\omega) + H_{m2}^2(\omega)] + \\ &+ 4\chi_m \chi_\alpha^m \sin \varphi H_{m1}(\omega) H_{m2}(\omega) \} \end{aligned} \quad (2-62)$$

— постоянная составляющая;

$$\begin{aligned} M_{m3}(2\omega) \cos(2\omega t + \varphi) &= -\frac{1}{4M_s} \{ (\chi_m^2 - \chi_\alpha^{m2}) \times \\ &\times [H_{m1}^2(\omega) \cos 2\omega t + H_{m2}^2(\omega) \cos(2\omega t + \varphi)] \} \end{aligned} \quad (2-63)$$

— составляющая удвоенной частоты.

Таким образом, поперечное переменное магнитное поле приводит к уменьшению постоянной намагниченности и появлению переменной намагниченности, направленной по оси x_3 и изменяющейся с удвоенной частотой.

Уменьшение постоянной намагниченности, зависящее по квадратичному закону от амплитуд переменного поля, как видно из формулы (2-62), происходит при любой поляризации переменного поля.

В случае линейной поляризации, когда, например, $H_{m2}(\omega) = 0$, согласно (2-62)

$$\Delta M_3 = \frac{1}{4M_s} (\chi_m^2 + \chi_\alpha^2) H_{m1}^2(\omega).$$

В случае круговой поляризации $\varphi = +\pi/2$, $H_{m1}(\omega) = -H_{m2}(\omega) = H_m(\omega)$ и согласно (2-62)

$$\Delta M_3 = \frac{1}{2M_s} (\chi_m \mp \chi_\alpha)^2 H_m^2(\omega), \quad (2-64)$$

где знак «—» соответствует левому вращению (по часовой стрелке, если смотреть вдоль направления распространения), знак «+» — правому. Поэтому для правого вращения эффект уменьшения постоянной намагниченности сильнее, чем для левого. Наиболее интенсивно этот эффект происходит в области ферромагнитного резонанса и может быть использован для детектирования модулированного сигнала СВЧ.

Удвоение частоты определяется выражением (2-63). Косинусы и синусы удвоенной частоты входят в это выражение линейно. Применяя символический метод, перепишем (2-63) в виде

$$\dot{M}_3(2\omega) = -\frac{1}{4M_s} (\chi_m^2 - \chi_\alpha^2) [\dot{H}_{m1}^2(\omega) + \dot{H}_{m2}^2(\omega)]. \quad (2-65)$$

Таким образом, переменная намагниченность содержит гармонику удвоенной частоты, направленную вдоль постоянного магнитного поля. Согласно (2-32) эта гармоника возбудит электромагнитное поле частоты 2ω , магнитная составляющая которого направлена по оси x_3 . Наличие переменной составляющей $H_3(2\omega)$ изменит вид проекций уравнения (2-60) на оси x_1 и x_2 . Решая эти уравнения относительно компонент переменной намагниченности M_1 и M_2 , перпендикулярных направлению постоянного магнитного поля, получаем составляющие намагниченности частоты 3ω .

Однако если переменное магнитное поле имеет круговую поляризацию, т. е.

$$\dot{H}_{m1}(\omega) = \pm j\dot{H}_{m2}(\omega),$$

то согласно (2-65)

$$M_3(2\omega) = 0,$$

т. е. гармоника намагниченности удвоенной частоты отсутствует. При линейной поляризации поля составля-

ющая намагниченности удвоенной частоты $M_3(2\omega)$ максимальна.

Возникновение составляющей намагниченности частоты 2ω можно объяснить следующим образом. Вектор намагниченности прецессирует вокруг направления постоянного магнитного поля, оставаясь постоянным по величине. В общем случае составляющие переменной намагниченности в плоскости, перпендикулярной \vec{H}_0 , не равны друг другу и \vec{M} прецессирует по эллиптической траектории (рис. 2-14). Поскольку вектор \vec{M} не изменяет свою величину, то его проекция на направление

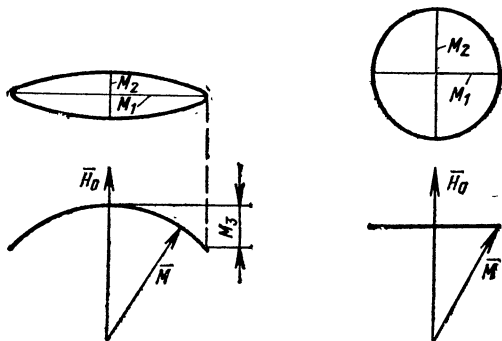


Рис. 2-14. Возникновение намагниченности $M_3(2\omega)$ в общем случае.

Рис. 2-15. Отсутствие намагниченности $M_3(2\omega)$ в случае круговой поляризации поля.

\vec{H}_0 изменяется с частотой, вдвое большей частоты переменного поля. Если $H_{m1} = \pm j H_{m2}$, т. е. в плоскости, перпендикулярной \vec{H}_0 , переменное магнитное поле имеет круговую поляризацию, то $M_{m1} = \pm j M_{m2}$ и, как видно из рис. 2-15, составляющая удвоенной частоты будет отсутствовать.

При воздействии на ферромагнитную среду составляющих поля с частотами ω_1 и ω_2 наряду с умножением частоты наблюдается появление комбинационных частот $n\omega_1 \pm m\omega_2$, и в первую очередь суммарных и разностных частот $\omega_1 \pm \omega_2$.

Рассмотрим поле, имеющее две частотные составляющие с круговой поляризацией $H(\omega_1)$ и $H(\omega_2)$:

$$\vec{H} = \vec{H}(\omega_1) + \vec{H}(\omega_2).$$

В проекциях на координатные оси это поле можно записать в виде

$$\left. \begin{aligned} H_1 &= H_m(\omega_1) \cos \omega_1 t + H_m(\omega_2) \cos \omega_2 t; \\ H_2 &= \pm H_m(\omega_1) \sin \omega_1 t \pm H_m(\omega_2) \sin \omega_2 t. \end{aligned} \right\} \quad (2-66)$$

Знаки « \pm » в последнем выражении указывают на направление вращения. Знак « $-$ » соответствует правому вращению, знак « $+$ » — левому. Формулы (2-66) определяют поля, имеющие одинаковое направление вращения. Очевидно, в случае полей, имеющих противоположное вращение, знаки « \pm » перед одним из слагаемых второго выражения (2-66) заменяется на « \mp ».

Подставляя (2-66) в (2-61), получаем:

$$\begin{aligned} M_1 &= [\chi_m(\omega_1) \mp \chi_\alpha^m(\omega_1)] H_m(\omega_1) \cos \omega_1 t + \\ &+ [\chi_m(\omega_2) \mp \chi_\alpha^m(\omega_2)] H_m(\omega_2) \cos \omega_2 t. \end{aligned} \quad (2-67)$$

$$\begin{aligned} M_2 &= \pm [\chi_m(\omega_1) \mp \chi_\alpha^m(\omega_1)] H_m(\omega_1) \sin \omega_1 t \pm \\ &\pm [\chi_m(\omega_2) \mp \chi_\alpha^m(\omega_2)] H_m(\omega_2) \sin \omega_2 t. \end{aligned} \quad (2-68)$$

Здесь знаки « \pm » перед квадратными скобками определяют направление вращения, верхний знак в квадратных скобках соответствует левому вращению, нижний — правому.

Подставляя (2-67) и (2-68) в (2-58), получаем:

$$M_3 = M_s - \Delta M_3 + M_{m3}(\omega_1 \mp \omega_2) \cos(\omega_1 \mp \omega_2) t,$$

где

$$\begin{aligned} \Delta M_3 &= \frac{1}{2M_s} \{ [\chi_m(\omega_1) \mp \chi_\alpha^m(\omega_1)]^2 H_m^2(\omega_1) + \\ &+ [\chi_m(\omega_2) \mp \chi_\alpha^m(\omega_2)]^2 H_m^2(\omega_2) \} \end{aligned} \quad (2-69)$$

— постоянная составляющая:

$$\begin{aligned} &M_{m3}(\omega_1 \mp \omega_2) \cos(\omega_1 \mp \omega_2) t = \\ &= \frac{H_m(\omega_1) H_m(\omega_2)}{M_s} [\chi_m(\omega_1) \mp \chi_\alpha^m(\omega_1)] [\chi_m(\omega_2) \mp \\ &\mp \chi_\alpha^m(\omega_2)] \cos(\omega_1 \mp \omega_2) t \end{aligned} \quad (2-70)$$

— составляющая разностной и суммарной частот.

В формулах (2-69) и (2-70), как и ранее, верхний знак в квадратных скобках соответствует левому вращению, нижний — правому, разностная частота

$\omega_1 - \omega_2$ — одинаковому направлению вращения полей, суммарная $\omega_1 + \omega_2$ — противоположному.

Согласно формуле (2-69) уменьшение постоянной намагниченности является аддитивным процессом [сравните с формулой (2-64)]. В случае действия двух переменных полей этот процесс определяется двумя составляющими, обусловленными каждым полем в отдельности. Эффект наиболее существен для правой круговой поляризации и в области ферромагнитного резонанса, где величина $\chi_m(\omega) + \chi_\alpha^m(\omega)$ значительна.

Намагниченность разностной частоты согласно (2-70) максимальна при правой круговой поляризации обоих полей в области ферромагнитного резонанса хотя бы одного из полей, когда величина $\chi_m(\omega) + \chi_\alpha^m(\omega)$ для этого поля максимальна. Намагниченность суммарной частоты согласно (2-70) максимальна в области ферромагнитного резонанса для поля правой круговой поляризации, когда $\chi_m(\omega) + \chi_\alpha^m(\omega)$ максимальна.

Согласно (2-32) составляющие намагниченности $M_3(\omega_1 + \omega_2)$ и $M_3(\omega_1 - \omega_2)$ возбуждают соответственно электромагнитные поля $E(\omega_1 + \omega_2)$ и $E(\omega_1 - \omega_2)$, магнитные составляющие которых направлены по оси x_3 .

Если действующие магнитные поля $H(\omega_1)$ и $H(\omega_2)$ поперечные, но имеют линейную или эллиптическую поляризацию, то кроме постоянной составляющей, составляющих частот $\omega_1 \pm \omega_2$ появятся также составляющие удвоенных частот $2\omega_1$ и $2\omega_2$. Составляющие частот $\omega_1 \pm \omega_2$ появятся вследствие того, что поля линейной и эллиптической поляризации можно представить как суперпозицию полей круговой поляризации, вращающихся в противоположные стороны.

2-4. САМОФОКУСИРОВКА И САМОКАНАЛИЗАЦИЯ ЭНЕРГИИ ЭЛЕКТРОМАГНИТНОГО ПОЛЯ

Под действием электромагнитного поля большой интенсивности диэлектрическая проницаемость среды изменяется. Для диэлектрика нелинейность диэлектрической проницаемости определяется не только электронной и ориентационной поляризуемостью, но и давлением электромагнитной волны, изменяющей плотность вещества (особенно в газах), а следовательно, и диэлектрическую проницаемость. Нагрев среды под действием

электромагнитного поля также приводит к изменению плотности вещества, а следовательно, и диэлектрической проницаемости.

Электромагнитную энергию можно передать на расстояние с помощью направленного излучения в виде потока энергии, ограниченного в поперечном направлении. В этом же направлении изменяется и напряженность поля в потоке.

Влияние нелинейности на распространение потока энергии можно исследовать с помощью нелинейных волновых уравнений (2-1). Однако из-за сложности решения этой задачи ограничимся качественным рассмотрением.

Под действием электромагнитного поля большой интенсивности диэлектрическая проницаемость среды изменяется приблизительно по закону

$$\epsilon_a = \epsilon_0 (\epsilon^n + \epsilon_2 E^2), \quad (2-71)$$

причем при любой реально достижимой напряженности поля выполняется условие

$$\epsilon_2 E^2 \ll 1;$$

тогда показатель преломления определяется выражением

$$n = \sqrt{\epsilon} = n^1 + n_2 E^2. \quad (2-72)$$

Так как напряженность поля в потоке убывает от оси к краям, то коэффициент преломления при $n_2 > 0$ также убывает в радиальном направлении. При этом фаза поля на периферии потока согласно выражению

$$e^{-j\omega \sqrt{\epsilon_a \mu_0} z} = e^{-j \frac{\omega n}{c} z}$$

будет опережать фазу поля на оси. Вследствие этого происходит искривление фазового фронта, что эквивалентно действию собирающей линзы (рис. 2-16). Энергия концентрируется на оси потока (фокусируется), что приводит к увеличению напряженности поля и коэффициента преломления на оси потока и к дальнейшему увеличению фокусировки (*самофокусировка*). В отличие от фокусирующей линзы, когда поток после фокуса снова расходится, в нелинейной среде происходит захват потока в канал (*самоканализация*). Дальнейшее фокусирование ограничивается другими нелинейными процессами, противодействующими фокусировке. Диа-

метр канала обычно составляет несколько длин волн. Напряженность поля внутри канала может оказаться настолько большой (примерно 10^8 В/см), что происходит разрушение материала.

В нелинейной среде при большой плотности мощности в потоке (порядка мегаватт на квадратный сантиметр в оптическом диапазоне волн) создаются условия, при которых поток не будет расходиться и приближенно для монохроматического поля его можно рассматри-

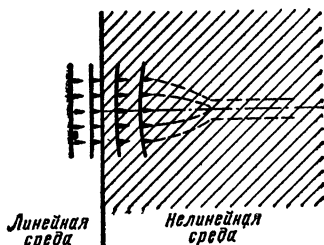


Рис. 2-16. Самофокусировка и самоканализация энергии.

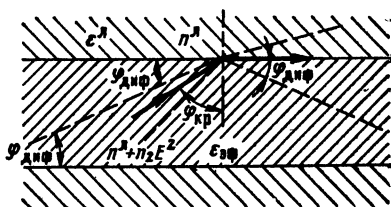


Рис. 2-17. К самоканализации энергии.

вать как распространяющийся внутри канала, заполненного диэлектрической средой с проницаемостью

$$\epsilon_{эф} = \epsilon^л + \frac{1}{2} \epsilon_2 E_m^2, \quad (2-73)$$

представляющей диэлектрическую проницаемость (2-71), усредненную во времени. Канал окружен линейной средой с диэлектрической проницаемостью $\epsilon^л$ (рис. 2-17). При этом любой луч, падающий на границу сред под углом, большим критического, определяемого выражением

$$\sin \varphi_{кр} = \frac{n^л}{n^л + \frac{1}{2} n_2 E_m^2} = 1 - \frac{\epsilon_2}{4\epsilon^л} E_m^2, \quad (2-74)$$

удовлетворяет условию полного внутреннего отражения и не выходит за пределы канала.

С другой стороны, при слабой нелинейности

$$\sin \varphi_{кр} = \cos \left(\frac{\pi}{2} - \varphi_{кр} \right) \approx 1 - \frac{(\pi/2 - \varphi_{кр})^2}{2};$$

отсюда согласно (2-74)

$$\frac{\pi}{2} - \varphi_{кр} \approx \sqrt{\frac{\varepsilon_2}{2\varepsilon_1}} E_m. \quad (2-75)$$

Если излучатель имеет апертуру в виде круга диаметром $2R$, то при распространении в линейной среде дифракционная расходимость определяется выражением

$$\varphi_{диф} = \frac{1,22\lambda}{2R}, \quad (2-76)$$

где λ — длина волны в данной среде.

Очевидно, в нелинейной среде при выполнении условия

$$\varphi_{диф} \leq \frac{\pi}{2} - \varphi_{кр} \quad (2-77)$$

поток будет иметь форму цилиндра и расходимости не будет. Условие (2-77) называется условием самоканализации. Это условие согласно (2-75) и (2-76) может быть представлено в виде

$$\frac{1,22\lambda}{2R} \leq \sqrt{\frac{\varepsilon_2}{2\varepsilon_1}} E_m. \quad (2-78)$$

Средняя мощность, передаваемая при канализации,

$$P_0 \approx \frac{\pi R^2 E_m^2}{2Z_0}, \quad (2-79)$$

где $Z_0 = \sqrt{\mu_0/\varepsilon_{эф}}$ — волновое сопротивление среды в канале.

Наименьшая или критическая мощность, при которой возникает явление самоканализации, определяется согласно (2-79) и (2-78) выражением

$$P_{0кр} \approx \frac{(1,22\lambda)^2 \pi \varepsilon_1}{4Z_0 \varepsilon_2}. \quad (2-80)$$

Если $P_0 < P_{0кр}$, то поток будет расходящимся, при $P_0 \approx P_{0кр}$ наблюдается самоканализация потока; при $P_0 > P_{0кр}$ плотность мощности и напряженность в канале сильно возрастают, что может привести к распаду канала на отдельные области (нити) самоканализации.

Если среда такова, что в выражении (2-72) $n_2 < 0$, то при распространении потока энергии цилиндрической формы, в котором напряженность падает от оси

к краям, коэффициент преломления будет возрастать в радиальном направлении. По мере распространения потока в такой нелинейной среде поверхности равных фаз будут все больше и больше выгибаться, что эквивалентно рассеивающей линзе (дефокусировка).

Аналогичные явления наблюдаются при распространении узких направленных пучков электромагнитных волн в плазме ионосферы. Пучок может фокусироваться или образовывать волноводный канал в верхних слоях ионосферы (свыше 250 км), где концентрация падает с ростом мощности волны. В нижних слоях (до 250 км) концентрация электронов возрастает под воздействием поля и пучки дефокусируются.

2.5. НАКЛОННОЕ ПАДЕНИЕ НА ГРАНИЦУ ИЗОТРОПНЫХ ЛИНЕЙНОЙ И НЕЛИНЕЙНОЙ ДИЭЛЕКТРИЧЕСКИХ СРЕД БЕЗ ПОТЕРЬ

Рассмотрим падение плоской однородной волны, распространяющейся в линейной среде и падающей

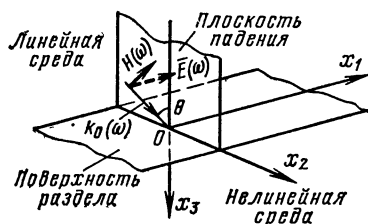


Рис. 2-18. Наклонное падение электромагнитной волны на границу линейной и нелинейной сред.

под углом θ на плоскую границу раздела с нелинейной средой. Считаем падающую волну монохроматической, пренебрегая влиянием гармоник из-за слабой нелинейности второй среды.

Поле в линейной среде $E_{(1)}$, $H_{(1)}$ определяется как сумма падающей E , H и отраженной E_0 , H_0 волн:

$$\begin{aligned} E_{(1)} &= E + E_0, \\ H_{(1)} &= H + H_0. \end{aligned}$$

Поле в нелинейной среде определяется полем преломленной волны

$$E_{(2)}, H_{(2)}.$$

Рассмотрим падение волны с линейной горизонтальной поляризацией. Вертикально поляризованная волна исследуется аналогичным образом. Все другие возможные виды поляризации можно рассмотреть как супер-

позицию волн горизонтальной и вертикальной поляризации.

В случае плоской границы раздела удобно совместить ее с одной из плоскостей координат, например, с плоскостью Ox_1x_2 . В случае линейной поляризации для упрощения решения удобно совместить другую координатную плоскость, например Ox_2x_3 (рис. 2-18) с плоскостью падения. Тогда для горизонтальной поляризации с осью x_1 совпадает направление вектора \mathbf{E} , для вертикальной — направление вектора \mathbf{H} и выражения, определяющие поле, упростятся.

Электромагнитное поле в диэлектрике без потерь и сторонних источников согласно (2-1) определяется волновым уравнением

$$\Delta \mathbf{E} - \mu_0 \epsilon_a^\pi \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \mathbf{P}_{\text{нл}}}{\partial t^2}, \quad (2-81)$$

где $\mathbf{P}_{\text{нл}}$ определяется согласно (2-4) уравнением

$$\frac{d^2 P}{dt^2} + \omega_0^2 P + \beta P^2 = \frac{e^2 N}{m} E. \quad (2-82)$$

Рассмотрим линейное приближение. Поле падающей волны определяется волновым уравнением

$$\Delta \mathbf{E} - \mu_0 \epsilon_{a1}^\pi \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0,$$

которое для монохроматического поля можно записать в виде

$$\Delta \dot{\mathbf{E}}(\omega) + k_{(1)}^2(\omega) \dot{\mathbf{E}}(\omega) = 0 \quad (2-83)$$

и его решение

$$\dot{\mathbf{E}}(\omega) = \mathbf{E}_m(\omega) e^{j[\omega t - \mathbf{K}_{(1)}(\omega) \mathbf{R}]},$$

где

$$k_{(1)}(\omega) = \omega \sqrt{\mu_0 \epsilon_{a1}(\omega)}.$$

Поле падающей волны горизонтальной поляризации в нашем случае будет иметь вид (см. [1]):

$$\dot{\mathbf{E}}(\omega) = \mathbf{e}_1 E_m(\omega) e^{j[\omega t - k_{(1)}(\omega)(x_2 \sin \theta + x_3 \cos \theta)]}, \quad (2-84)$$

$$\begin{aligned} \dot{\mathbf{H}}(\omega) = j \frac{\text{rot } \dot{\mathbf{E}}(\omega)}{\omega \mu_0} = H_m(\omega) (\mathbf{e}_2 \cos \theta - \mathbf{e}_3 \sin \theta) \times \\ \times e^{j[\omega t - k_{(1)}(\omega)(x_2 \sin \theta + x_3 \cos \theta)]}, \end{aligned}$$

где $k_{(1)} = \omega \sqrt{\epsilon_{a1}(\omega) \mu_0}$ — постоянная распространения в первой среде; $\epsilon_{a1}(\omega)$ — диэлектрическая проницаемость первой среды; μ_0 — ее магнитная проницаемость.

Так как среда без потерь, то электрическая и магнитная составляющие поля совпадают по фазе:

$$E_m(\omega) = Z_{01}(\omega) H_m(\omega),$$

где $Z_{01}(\omega) = \sqrt{\frac{\mu_0}{\epsilon_{a1}(\omega)}}$ — волновое сопротивление первой среды.

В линейном приближении поле отраженной и прошедшей волны будет также определяться уравнением (2-83), решения которого будут иметь вид, аналогичный выражениям (2-84).

Отражение и преломление плоских волн на границе раздела двух сред в линейном приближении подробно проанализировано в [1]. Результаты этого анализа сводятся к следующему:

Угол падения равен углу отражения:

$$\theta = \pi - \theta_0. \quad (2-85)$$

Угол падения θ и угол преломления ϑ связаны соотношением

$$\frac{\sin \vartheta}{\sin \theta} = \frac{k_{(1)}(\omega)}{k_{(2)}(\omega)} = \sqrt{\frac{\epsilon_1(\omega)}{\epsilon_2(\omega)}}. \quad (2-86)$$

Амплитуда отраженной волны $E_{m0}(\omega)$ и амплитуда преломленной волны $E_{m(2)}(\omega)$ определяются формулами Френеля, которые для горизонтальной поляризации имеют следующий вид:

$$\left. \begin{aligned} E_{m0}(\omega) &= E_m(\omega) \frac{\sqrt{\epsilon_1(\omega)} \cos \theta - \sqrt{\epsilon_2(\omega)} \cos \vartheta}{\sqrt{\epsilon_1(\omega)} \cos \theta + \sqrt{\epsilon_2(\omega)} \cos \vartheta}; \\ E_{m(2)}(\omega) &= E_m(\omega) \frac{2 \sqrt{\epsilon_1(\omega)} \cos \theta}{\sqrt{\epsilon_1(\omega)} \cos \theta + \sqrt{\epsilon_2(\omega)} \cos \vartheta}, \end{aligned} \right\} \quad (2-87)$$

откуда находим выражения для коэффициента отражения $\Gamma_E(\omega)$ и коэффициента преломления $P_E(\omega)$:

$$\left. \begin{aligned} \Gamma_E(\omega) &= \frac{E_{m0}(\omega)}{E_m(\omega)} = \frac{\sqrt{\epsilon_1(\omega)} \cos \theta - \sqrt{\epsilon_2(\omega)} \cos \vartheta}{\sqrt{\epsilon_1(\omega)} \cos \theta + \sqrt{\epsilon_2(\omega)} \cos \vartheta}; \\ P_E(\omega) &= \frac{E_{m(2)}(\omega)}{E_m(\omega)} = \frac{2 \sqrt{\epsilon_1(\omega)} \cos \theta}{\sqrt{\epsilon_1(\omega)} \cos \theta + \sqrt{\epsilon_2(\omega)} \cos \vartheta}. \end{aligned} \right\} \quad (2-88)$$

Перейдем к анализу нелинейных эффектов. Если поле возбуждается монохроматическим источником, то волновое уравнение (2-1) можно переписать для Фурье-компонент в виде

$$\Delta \dot{E}(n\omega) + (n\omega)^2 \epsilon_a^n(n\omega) \mu_0 \dot{E}(n\omega) = - (n\omega)^2 \mu_0 \dot{P}^{nl}(n\omega), \quad (2-89)$$

где $n=0, \pm 1, \pm 2, \dots$

В низшем нелинейном приближении согласно уравнению (2-82) поле $E_{(2)}(\omega)$ возбуждает в нелинейной среде волну нелинейной поляризации $P_{(2)}(2\omega)$, определяемую выражением

$$\dot{P}(2\omega) = P_{m(2)}(2\omega) e^{j[2\omega t - k_P(2\omega)r]},$$

где $k_P(2\omega)$ — постоянная распространения волны нелинейной поляризации, равная

$$k_P(2\omega) = 2k_{(2)}(\omega) = 2\omega \sqrt{\epsilon_{a2}^n(\omega) \mu_0}. \quad (2-90)$$

Амплитуда второй гармоники волны поляризации $P_{m(2)}(2\omega)$ согласно (1-10) с учетом (2-88) определяется выражением

$$P_{m(2)}(2\omega) = \epsilon_0 \chi_2(2\omega) E_{m(2)}^2(\omega) = \epsilon_0 \chi_2(2\omega) P_E^2(\omega) E_m^2(\omega). \quad (2-91)$$

Волна поляризации $P_{(2)}(2\omega)$ является источником поля $E_{(2)}(2\omega)$. Согласно волновому уравнению (2-89), ограничиваясь низшим приближением для второй гармоники, получаем:

$$\Delta E(2\omega) + 4\omega^2 \epsilon_a^n(2\omega) \mu_0 E(2\omega) = - 4\omega^2 \mu_0 \dot{P}(2\omega), \quad (2-92)$$

где

$$2\omega \sqrt{\epsilon_a^n(2\omega) \mu_0} = k_{(2)}(2\omega)$$

— постоянная распространения волны второй гармоники в нелинейной среде.

Решение неоднородного уравнения (2-92) равно сумме общего решения однородного и частного решения неоднородного уравнений. Очевидно, в изотропной среде направление вектора $E_{(2)}(\omega)$ совпадает с направлением второй гармоники вектора поляризации $P_{(2)}(2\omega)$ и направлением второй гармоники вектора напряжен-

ности поля $\mathbf{E}_{(2)}(2\omega)$. Таким образом, аналогично решению уравнения (2-12) получим:

$$\begin{aligned}\dot{\mathbf{E}}_{m(2)}(2\omega) &= \mathbf{e}_1 E_{m(2)}(2\omega) e^{-jk_{(2)}(2\omega)r} + \\ &+ \mathbf{e}_1 \frac{4\omega^2 \mu_0 P_{m(2)}(2\omega)}{k_P^2(2\omega) - k_{(2)}^2(2\omega)} e^{-jk_P(2\omega)r}\end{aligned}$$

или

$$\begin{aligned}\mathbf{E}_{m(2)}(2\omega) &= \mathbf{e}_1 E_{m(2)}(2\omega) e^{-jk_{(2)}(2\omega)[x_2 \sin \vartheta(2\omega) + x_3 \cos \vartheta(2\omega)]} + \\ &+ \mathbf{e}_1 \frac{P_{m(2)}(2\omega)}{\varepsilon_{a2}(\omega) - \varepsilon_{a2}(2\omega)} e^{-jk_P(2\omega)[x_2 \sin \vartheta_P(2\omega) + x_3 \cos \vartheta_P(2\omega)]} \quad (2-93)\end{aligned}$$

где $\vartheta(2\omega)$ — угол между положительным направлением оси x_3 и направлением распространения волны $\mathbf{E}_{(2)}(2\omega)$; $\vartheta_P(2\omega)$ — угол между положительным направлением оси x_3 и направлением распространения волны поляризации $\mathbf{P}_{(2)}(2\omega)$.

Магнитная составляющая поля удвоенной частоты во второй среде определяется выражением

$$\begin{aligned}\dot{\mathbf{H}}_{m(2)}(2\omega) &= j \frac{\text{rot } \dot{\mathbf{E}}_{m(2)}(2\omega)}{2\omega\mu_0} = \frac{E_{m(2)}(2\omega)}{Z_{02}(2\omega)} [\mathbf{e}_2 \cos \vartheta(2\omega) - \\ &- \mathbf{e}_3 \sin \vartheta(2\omega)] e^{-jk_{(2)}(2\omega)[x_2 \sin \vartheta(2\omega) + x_3 \cos \vartheta(2\omega)]} + \\ &+ \frac{P_{m(2)}(2\omega)}{Z_{02}[\varepsilon_{a2}(\omega) - \varepsilon_{a2}(2\omega)]} [\mathbf{e}_2 \cos \vartheta_P(2\omega) - \mathbf{e}_3 \sin \vartheta_P(2\omega)] \times \\ &\times e^{-jk_P(2\omega)[x_2 \sin \vartheta_P(2\omega) + x_3 \cos \vartheta_P(2\omega)]}. \quad (2-94)\end{aligned}$$

На границе раздела сред согласно (1-20) выполняются условия

$$E_{\tau(1)}(2\omega) = E_{\tau(2)}(2\omega), \quad H_{\tau(1)}(2\omega) = H_{\tau(2)}(2\omega).$$

Так как граничные условия удовлетворяются в любой момент времени, то отсюда непосредственно следует, что в первой среде существует «отраженная» волна второй гармоники. Поле «отраженной» волны второй гармоники описывается выражением

$$\dot{\mathbf{E}}_0(2\omega) = \mathbf{e}_1 E_{m0} e^{j\{2\omega t - k_0(2\omega)[x_2 \sin \theta_0(2\omega) + x_3 \cos \theta_0(2\omega)]\}}, \quad (2-95)$$

где $k_0(2\omega) = 2\omega \sqrt{\varepsilon_{a1}^n(2\omega) \mu_0}$ — постоянная распространения этой волны, $\theta_0(2\omega)$ — угол отражения.

Найдем магнитную составляющую поля «отраженной» волны второй гармоники

$$\begin{aligned} \dot{H}_0(2\omega) &= \frac{j \operatorname{rot} \dot{E}_0(2\omega)}{2\omega \mu_0} = \\ &= \frac{E_{m0}(2\omega)}{Z_{01}(2\omega)} [\mathbf{e}_2 \cos \theta_0(2\omega) - \mathbf{e}_3 \sin \theta_0(2\omega)] \times \\ &\times e^{j \{2\omega t - k_0(2\omega)[x_2 \sin \theta_0(2\omega) + x_3 \cos \theta_0(2\omega)]\}} \end{aligned} \quad (2-96)$$

На границе раздела ($x_3=0$) из условия непрерывности касательных составляющих электрического поля согласно выражениям (2-93) и (2-95) для второй гармоники получим

$$\begin{aligned} E_{m2}(2\omega) e^{-jk_{(2)}(2\omega) \sin \vartheta(2\omega) x_2} + \frac{P_{m(2)}(2\omega)}{\varepsilon_{a2}(\omega) - \varepsilon_{a2}(2\omega)} \times \\ \times e^{-jk_P(2\omega) \sin \vartheta_P(2\omega) x_2} = E_{m0}(2\omega) e^{-jk_0(2\omega) \sin \theta_0(2\omega) x_2}. \end{aligned} \quad (2-97)$$

Граничные условия выполняются в любой точке границы раздела. Равенство (2-97) выполняется при любом x_2 , если

$$k_{(2)}(2\omega) \sin \vartheta(2\omega) = k_P(2\omega) \sin \vartheta_P(2\omega) = k_0(2\omega) \sin \theta_0(2\omega),$$

а так как

$$k_P(2\omega) \sin \vartheta_P(2\omega) = 2k_{(2)}(\omega) \sin \vartheta = 2k_{(1)}(\omega) \sin \theta,$$

то

$$\begin{aligned} 2k_{(1)}(\omega) \sin \theta = k_P(2\omega) \sin \vartheta_P(2\omega) = k_{(2)}(2\omega) \sin \vartheta(2\omega) = \\ = k_0(2\omega) \sin \theta_0(2\omega), \end{aligned} \quad (2-98)$$

где

$$\begin{aligned} k_{(1)}(\omega) &= \omega \sqrt{\varepsilon_{a1}(\omega) \mu_0}, \quad k_{(2)}(\omega) = \omega \sqrt{\varepsilon_{a2}(\omega) \mu_0}, \\ k_P(2\omega) &= 2k_{(2)}(\omega), \end{aligned}$$

$$k_{(2)}(2\omega) = 2\omega \sqrt{\varepsilon_{a2}(2\omega) \mu_0}, \quad k_0(2\omega) = 2\omega \sqrt{\varepsilon_{a1}(2\omega) \mu_0}.$$

Исходя из (2-98), получим выражения, представляющие собой *модифицированные законы Снеллиуса*

$$\left. \begin{aligned} \sin \vartheta(2\omega) &= \frac{2k_{(1)}(\omega)}{k_{(2)}(2\omega)} \sin \theta = \sqrt{\frac{\varepsilon_1(\omega)}{\varepsilon_2(\omega)}} \sin \theta; \\ \sin \vartheta_P(2\omega) &= \frac{2k_{(1)}(\omega)}{k_P(2\omega)} \sin \theta = \sqrt{\frac{\varepsilon_1(\omega)}{\varepsilon_2(\omega)}} \sin \theta; \\ \sin \theta_0(2\omega) &= \sqrt{\frac{\varepsilon_1(\omega)}{\varepsilon_1(2\omega)}} \sin \theta. \end{aligned} \right\} \quad (2-99)$$

Таким образом, в общем случае направление распространения волны второй гармоники в нелинейной

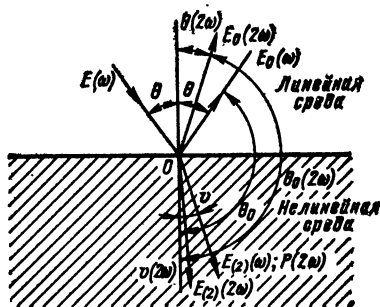


Рис. 2-19. Распространение волн основной частоты и второй гармоники при падении на границу линейной и нелинейной сред.

среде $E_{(2)}(2\omega)$ отличается от совпадающих направлений распространения преломленной волны основной частоты $E_{(2)}(\omega)$ и волны нелинейной поляризации $P_{(2)}(2\omega)$; направление распространения «отраженной» волны второй гармоники $E_0(2\omega)$, возникающей в первой (линейной) среде не совпадает с направлением распространения отраженной волны основной частоты $E_0(\omega)$ (рис. 2-19).

Направление распространения волн второй гармоники в линейной и нелинейной средах зависит от их дисперсионных свойств.

Если первая среда — вакуум (линейная недиспергирующая среда), то согласно (2-99) получим

$$\left. \begin{aligned} \sin \vartheta(2\omega) &= \frac{\sin \theta}{\sqrt{\varepsilon_2(2\omega)}}; \\ \sin \vartheta_P(2\omega) &= \frac{\sin \theta}{\sqrt{\varepsilon_2(\omega)}}; \\ \sin \theta_0(2\omega) &= \sin \theta, \end{aligned} \right\} \quad (2-100)$$

т. е. отраженная волна второй гармоники в вакууме распространяется в том же направлении, что и отраженная волна основной частоты. Волна нелинейной поляризации $P_{(2)}(2\omega)$ в нелинейной среде распространяется в том же направлении, что и преломленная волна основной частоты, но волна второй гармоники поля в общем случае, когда $\epsilon_2(2\omega) \neq \epsilon_2(\omega)$ распространяется в другом направлении и лишь в случае $\epsilon_2(2\omega) = \epsilon_2(\omega)$ или нормального падения распространяется в том же направлении.

Соотношения между амплитудами волн вторых гармоник также получаются из граничных условий (1-20)

$$E_{\tau 1}(2\omega) = E_{\tau(2)}(2\omega);$$

$$H_{\tau(1)}(2\omega) = H_{\tau(2)}(2\omega)$$

или согласно (2-97) и выражениям (2-93) — (2-96), (2-90) и (2-91)

$$E_{m(2)}(2\omega) - E_{m0}(2\omega) = - \frac{4\omega^2 \mu_0 \epsilon_0 \chi_2(2\omega) P_E^2(\omega)}{4k_{(2)}^2(\omega) - k_{(2)}^2(2\omega)} E_m^2(\omega), \quad (2-101)$$

$$\begin{aligned} \sqrt{\epsilon_2(2\omega)} E_{m(2)}(2\omega) \cos \vartheta(2\omega) + \sqrt{\epsilon_1(2\omega)} E_{m0}(2\omega) \cos \theta(2\omega) = \\ = - \frac{4\omega^2 \mu_0 \epsilon_0 \chi_2(2\omega) \sqrt{\epsilon_2(\omega)} P_E^2(\omega)}{4k_{(2)}^2(\omega) - k_{(2)}^2(2\omega)} E_m^2(\omega) \cos \vartheta, \quad (2-102) \end{aligned}$$

где $\theta(2\omega) = \pi - \theta_0(2\omega)$.

Решая эти уравнения совместно, получаем выражения для амплитуд «отраженной» и «преломленной» волн второй гармоники:

$$\begin{aligned} E_{m0}(2\omega) = E_m^2(\omega) \frac{\chi_2(2\omega) P_E^2(\omega)}{\epsilon_2(\omega) - \epsilon_2(2\omega)} \times \\ \times \frac{\sqrt{\epsilon_2(2\omega)} \cos \vartheta(2\omega) - \sqrt{\epsilon_2(\omega)} \cos \vartheta}{\sqrt{\epsilon_2(2\omega)} \cos \vartheta(2\omega) + \sqrt{\epsilon_1(2\omega)} \cos \theta(2\omega)}, \quad (2-103) \end{aligned}$$

$$\begin{aligned} E_{m(2)}(2\omega) = - E_m^2(\omega) \frac{\chi_2(2\omega) P_E^2(\omega)}{\epsilon_2(\omega) - \epsilon_2(2\omega)} \times \\ \times \frac{\sqrt{\epsilon_2(\omega)} \cos \vartheta + \sqrt{\epsilon_1(2\omega)} \cos \theta(2\omega)}{\sqrt{\epsilon_2(2\omega)} \cos \vartheta(2\omega) + \sqrt{\epsilon_1(2\omega)} \cos \theta(2\omega)}, \quad (2-104) \end{aligned}$$

которые представляют собой *модифицированные формулы Френеля* для поля горизонтальной поляризации.

Умножая числитель и знаменатель выражения (2-103) на $\sqrt{\varepsilon_2(2\omega) \cos \vartheta(2\omega) + \sqrt{\varepsilon_2(\omega) \cos \vartheta}}$ с учетом законов преломления (2-86) и (2-99), получаем:

$$E_{m0}(2\omega) = E_m^2(\omega) \frac{\chi_2(2\omega) P_E^2(\omega)}{[\sqrt{\varepsilon_2(2\omega) \cos \vartheta(2\omega) + \sqrt{\varepsilon_2(\omega) \cos \vartheta}}] \times} \\ \times \frac{\chi_2(2\omega) P_E^2(\omega)}{[\sqrt{\varepsilon_2(2\omega) \cos \vartheta(2\omega) + \sqrt{\varepsilon_1(2\omega) \cos \vartheta(2\omega)}}]}. \quad (2-104)$$

Таким образом, амплитуда «отраженной» волны практически не зависит от рассогласования фазовых скоростей волн, распространяющихся в нелинейном диэлектрике. «Отраженное» поле второй гармоники определяется поверхностным слоем нелинейной среды толщиной порядка длины волны. Так как преобразование энергии на границе мало, реакцией поверхностного слоя на падающую волну можно пренебречь и амплитуду падающей волны можно считать постоянной.

Полное поле второй гармоники в нелинейной среде определяется выражением (2-93). Согласно (2-99)

$$\mathbf{k}_p(2\omega) \mathbf{r} - \mathbf{k}_{(2)}(2\omega) \mathbf{r} = \\ = \frac{2\omega}{c} x_3 [\sqrt{\varepsilon_2(\omega) \cos \vartheta_p(2\omega)} - \sqrt{\varepsilon_2(2\omega) \cos \vartheta(2\omega)}],$$

следовательно, с учетом (2-101) выражение (2-93) можно представить в виде

$$\dot{\mathbf{E}}_{(2)}(2\omega) = \mathbf{e}_1 \left\{ E_{m0}(2\omega) + \frac{\mathbf{P}_{m(2)}(2\omega)}{\varepsilon_{a2}(\omega) - \varepsilon_{a2}(2\omega)} \times \right. \\ \times \left[e^{-j \frac{2\omega}{c} x_3 [\sqrt{\varepsilon_2(\omega) \cos \vartheta_p(2\omega)} - \sqrt{\varepsilon_2(2\omega) \cos \vartheta(2\omega)}]} - 1 \right] \Bigg\} \times \\ \times e^{j\{\omega t - k_{(2)}(2\omega) [x_2 \sin \vartheta(2\omega) + x_3 \cos \vartheta(2\omega)]\}}.$$

Для значений x_3 , удовлетворяющих условию

$$\frac{2\omega}{c} x_3 [\sqrt{\varepsilon_2(\omega) \cos \vartheta_p(2\omega)} - \sqrt{\varepsilon_2(2\omega) \cos \vartheta(2\omega)}] \ll 1,$$

согласно приближенной формуле

$$e^{-x} \approx 1 - x$$

получим

$$\dot{E}_{(2)}(2\omega) = e_1 \left\{ E_{m0}(2\omega) - j \frac{P_{m(2)}(2\omega)}{\varepsilon_{a2}(\omega) - \varepsilon_{a2}(2\omega)} \frac{2\omega}{c} x_3 \times \right. \\ \left. \times [V \varepsilon_2(\omega) \cos \vartheta_P(2\omega) - V \varepsilon_2(2\omega) \cos \vartheta(2\omega)] \right\} e^{j[\omega t - k_{(2)}(2\omega)r]}$$

Умножим числитель и знаменатель полученного выражения на $V \varepsilon_2(\omega) \cos \vartheta_P(2\omega) + V \varepsilon_2(2\omega) \cos \vartheta(2\omega)$ и с учетом выражения (2-99) получим:

$$E_{(2)}(2\omega) = e_1 \left\{ E_{m0}(2\omega) - \right. \\ \left. - j \frac{P_{m(2)}(2\omega) \frac{2\omega}{c} x_3}{V \varepsilon_2(\omega) \cos \vartheta_P(2\omega) + V \varepsilon_2(2\omega) \cos \vartheta(2\omega)} \right\} e^{j[\omega t - k_{(2)}(2\omega)r]}. \quad (2-105)$$

Таким образом, начальная амплитуда «преломленной» волны второй гармоники (при $x_3=0$) равна амплитуде «отраженной» волны, составляющая, сдвинутая относительно нелинейной поляризации, возрастает пропорционально расстоянию x_3 от границы при согласовании фазовых скоростей или при согласовании нормальных составляющих скоростей:

$$V \varepsilon_2(2\omega) \cos \vartheta(2\omega) = V \varepsilon_2(\omega) \cos \vartheta_P(2\omega), \quad (2-106)$$

так как касательные составляющие согласованы в силу граничных условий. Соотношение (2-106) выполняется при

$$V \varepsilon_2(2\omega) = V \varepsilon_2(\omega), \quad \cos \vartheta(2\omega) = \cos \vartheta_P(2\omega),$$

т. е. при условии волнового синхронизма

$$k_P(2\omega) = k_{(2)}(2\omega).$$

Так как практически все среды обладают дисперсией и

$$\varepsilon_2(\omega) \neq \varepsilon_2(2\omega),$$

то согласование возможно лишь в анизотропных средах (см. § 2-1) с помощью двойного лучепреломления.

Если условия волнового синхронизма не выполнены, то амплитуда напряженности поля будет осциллировать

с изменением x_3 . Расстояние l между соседними максимумами напряженности поля определяется из условия

$$l \frac{2\omega}{c} [\sqrt{\epsilon_2(\omega)} \cos \vartheta_P(2\omega) - \sqrt{\epsilon_2(2\omega)} \cos \vartheta(2\omega)] = 2\pi;$$

это расстояние соответствует удвоенной длине когерентности.

«Преломленная» волна второй гармоники представляет собой неоднородную плоскую волну. Плоскости равных амплитуд не совпадают с плоскостями равных фаз, параллельны поверхности раздела и определяются выражением

$$x_3 = \text{const.}$$

Амплитуда напряженности поля «преломленной» волны второй гармоники во много раз больше амплитуды напряженности поля «отраженной» волны второй гармоники, так как в создании поля «отраженной» волны участвуют лишь поверхностный слой нелинейной среды толщиной порядка длины волны, а в создании «преломленной» волны по мере ее распространения принимают участие все новые и новые диполи нелинейной среды.

Глава третья

УСИЛЕНИЕ И ГЕНЕРИРОВАНИЕ В АКТИВНЫХ КВАНТОВЫХ СРЕДАХ

3-1. КВАНТОВОЕ УСИЛЕНИЕ И ГЕНЕРИРОВАНИЕ

При распространении в пассивной среде с потерями электромагнитное поле поглощается и амплитуда поля с расстоянием убывает по экспоненциальному закону

$$E_m(x) = E_m e^{-\alpha_n x},$$

где α_n — постоянная затухания. Для диэлектрика с малыми потерями α_n определяется выражением

$$\alpha_n = \frac{\omega \epsilon''}{2cn}, \quad (3-1)$$

где n — коэффициент преломления.

При распространении в активной квантовой среде электромагнитное поле усиливается, амплитуда поля

возрастает с расстоянием по экспоненциальному закону

$$E_m(x) = E_m e^{\alpha_{yc} x}.$$

Здесь α_{yc} — коэффициент квантового усиления. Его можно определить по формуле (3-1), рассматривая активную среду как диэлектрик с малыми отрицательными потерями ($\epsilon'' < 0$). С учетом (1-47) для двухуровневой квантовой среды получим:

$$\alpha_{yc}(\omega) = \frac{\omega \hbar}{2cn} B_{nm} g(\omega) \left(N_n - \frac{g_n}{g_m} N_m \right). \quad (3-2)$$

Для усиления электромагнитного поля необходимо, чтобы $\alpha_{yc}(\omega) > 0$, т. е. создание инверсной населенности, когда выполняется условие

$$\left(N_n - \frac{g_n}{g_m} N_m \right) > 0.$$

Для создания инверсной населенности необходим источник энергии, который переводил бы частицы в возбужденное состояние (источник возбуждения), аналогичный источнику постоянного напряжения (источник питания), сообщаящему электронам кинетическую или потенциальную энергию и применяемого в обычных электронных лампах. Создание избытка возбужденных частиц на верхнем уровне по отношению к уровню, лежащему ниже, осуществляется различными методами, которые будут рассмотрены ниже.

В квантовой среде наряду с индуцированными происходят спонтанные переходы, фаза, поляризация и частота которых не связаны с распространяющимся полем, и поэтому на усиление электромагнитной волны не влияющие. Спонтанные переходы являются источником шумов. Уровень шумов, определяемый спонтанным излучением, согласно (1-43) в диапазоне сверхвысоких частот много меньше, чем в оптическом.

Коэффициент квантового усиления можно выразить через коэффициент Эйнштейна для спонтанного излучения A_{nm} (1-43):

$$\alpha_{yc}(\omega) = \frac{\pi c^2}{4\omega^2} A_{nm} g(\omega) \left(N_n - \frac{g_n}{g_m} N_m \right). \quad (3-3)$$

На частоте перехода $\omega = \omega_{nm}$ согласно (П1-68) коэффициент усиления максимален. Для газообразных сред с учетом (П1-71)

$$\alpha_{yc}(\omega_{nm}) = \frac{\sqrt{\pi} c^2}{4\omega^2 \Delta\omega_D} A_{nm} \left(N_n - \frac{g_n}{g_m} N_m \right),$$

где $\Delta\omega_D$ — доплеровский сдвиг частоты, пропорциональный ω_{nm} .

Полагая, что населенность верхнего уровня рабочего перехода много больше населенности нижнего уровня и при создании инверсной населенности произведение $N_n A_{nm}$ поддерживается постоянным независимо от частоты перехода, получим:

$$\alpha_{yc}(\omega_{nm}) \sim \frac{1}{\omega_{nm}^2 \Delta\omega_D} \sim \lambda_{nm}^3, \quad (3-4)$$

т.е. коэффициент квантового усиления быстро падает при уменьшении длины волны перехода.

В случае активного магнетика, определяя α_{yc} с помощью формулы

$$\alpha = \frac{\omega\mu''}{2cn}$$

и формулы (1-48), получаем то же выражение (3-2).

В реальной активной среде наряду с усилением наблюдаются потери за счет рассеяния на неоднородностях среды, резонансного поглощения из-за наличия других уровней, частота перехода между которыми равна частоте рабочего перехода ω_{nm} , и т.д. Характеризуя эти потери коэффициентом потерь α_n и учитывая их при распространении электромагнитной волны в активной среде, получаем:

$$E_m(x) = E_m e^{(\alpha_{yc} - \alpha_n)x}.$$

Для усиления электромагнитного поля в активной среде необходимо, чтобы коэффициент квантового усиления был не только больше нуля, но и больше коэффициента потерь активной среды, т.е.

$$\alpha_{yc} > \alpha_n; \quad (3-5)$$

иначе говоря, мощность излучения должна превышать мощность потерь:

$$P_{изл} > P_{пот} \quad (3-6)$$

(это выражение иногда называют *условием усиления*).

Коэффициент усиления α_{yc} характеризует усиление на единицу длины и остается постоянным при неболь-

ших напряженностях распространяющегося поля. При увеличении напряженности поля по мере распространения коэффициент усиления среды уменьшается, так как согласно (3-2) $\alpha_{ус}$ зависит от инверсной населенности. При увеличении напряженности распространяющегося поля увеличивается вероятность индуцированных переходов между уровнями рабочего перехода [см. формулу (1-42)] и населенность уровней выравнивается. При

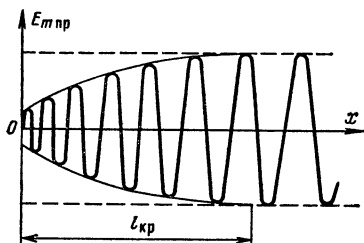


Рис. 3-1. Усиление в активной среде.

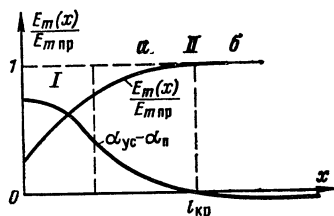


Рис. 3-2. Зависимость коэффициента усиления $\alpha_{ус}$ и амплитуды распространяющегося поля от расстояния, пройденного волной в активной среде.

некоторой предельной напряженности распространяющегося поля $E_{m пр}(x)$ наступает насыщение и усиление прекращается, амплитуда распространяющегося поля не изменяется (рис. 3-1). Уменьшение $\alpha_{ус}$ с ростом амплитуды поля E_m ограничивает максимальное значение напряженности поля $E_{m пр}$, которое может быть получено в процессе усиления. В той части активной среды, где напряженность поля равна предельному значению, усиления нет, так как режим среды близок к режиму полного насыщения. Напряженность поля достигает предельного значения при прохождении в среде некоторого критического пути $l_{кр}$. Использование для усиления активной среды с длиной, больше критической, нецелесообразно, так как энергия источника возбуждения расходуется для возбуждения области, соответствующей $l > l_{кр}$, а распространяющееся поле в этой области не возрастает. Зависимости $\alpha_{ус}$ и E_m от расстояния, пройденного волной в среде, приведены на рис. 3-2. Область I соответствует малым значениям напряженности поля, при которых инверсная населенность практически не изменя-

ется и коэффициент α_{yc} постоянен. Так как значение α_{yc} обычно мало, то при постоянном α_{yc} в этой области

$$E_m(x) = [1 + (\alpha_{yc} - \alpha_n)x] E_m(0),$$

т. е. область *I* является областью линейного режима усиления.

В области *II* можно выделить две подобласти. В области *II, а* равным приращениям пройденного расстояния соответствуют все меньшие и меньшие приращения напряженности поля. Это область нелинейного режима усиления. При $l = l_{кр}$ амплитуда поля достигает предельного значения и увеличение ее с расстоянием прекращается. Область *II, б* является областью полного насыщения.

Усиление в активной среде можно характеризовать коэффициентом усиления по мощности G . Если l — расстояние, которое проходит волна в активной среде, и излучение по всей длине происходит равномерно, то

$$G = e^{2(\alpha_{yc} - \alpha_n)l}.$$

В радиодиапазоне $\alpha_{yc} \approx 10^{-2} \text{ см}^{-1}$, а величиной α_n можно пренебречь. Для усиления волны в 100 раз необходимо, чтобы она прошла расстояние в 5 м. Практически усилитель такой длины не реален. Эффективную длину усилителя можно увеличить, возвращая усиленный сигнал обратно на вход усилителя и заставляя его многократно проходить один и тот же путь (*положительная обратная связь*). Обратная связь осуществляется с помощью резонатора. Резонатор накапливает энергию, излученную частицами, и с помощью этой энергии влияет на излучение последующих частиц. Усилитель, использующий положительную обратную связь, называется *регенеративным*.

При достаточно сильной обратной связи, когда почти весь усиленный сигнал возвращается обратно, усиление может достигнуть большого значения. В этом случае даже при отсутствии внешнего сигнала возможно усиление случайного спонтанного излучения, в результате чего усилитель превратится в генератор. При этом излучение активных частиц должно быть достаточным для компенсации потерь и отвода энергии из резонатора, т. е. мощность излучения активных частиц должна удовлетворять условию генерирования

$$P_{\text{изл}} > P_{\Sigma}, \quad (3-7)$$

где P_{Σ} — сумма мощности потерь и мощности излучаемой во внешнее пространство.

Если энергия излучения активных частиц превышает внутренние потери в резонаторе и среде, но меньше полных потерь, включающих потери на излучение во внешнее пространство, т. е.

$$P_{\text{пот}} < P_{\text{изл}} < P_{\Sigma}, \quad (3-8)$$

то генерирование не возникает, однако сигнал, взаимодействуя с активной средой, усилится и выйдет из резонатора с большей амплитудой. Происходит регенеративное усиление сигнала.

3-2. КВАНТОВЫЕ ГЕНЕРАТОРЫ И УСИЛИТЕЛИ

Развитие современной радиоэлектроники характеризуется:

1) освоением новых диапазонов электромагнитных волн: инфракрасного (0,4 мкм — 0,76 мкм), видимого (0,76—0,4 мкм), ультрафиолетового (0,4 мкм — 2 нм) и рентгеновского (2 нм — 6 пм);

2) отходом от методов классической радиоэлектроники, использовавшей потоки свободных электронов, и переходом к использованию свойств вещества и связанных (внутри вещества) электронов.

Квантовые генераторы и усилители основаны на использовании электромагнитного излучения молекул или атомов вещества, возникающего в результате взаимодействия с полем. Излучение в различных участках электромагнитного спектра связано с различным характером движения частиц и изменением их энергии. Излучение в диапазоне СВЧ связано с изменением вращательной энергии молекул и переходами между зеемановскими подуровнями, в инфракрасном диапазоне — с изменением энергии колебаний атомов в молекуле, в оптическом — с изменением энергии электронов атома при их переходах с одного энергетического уровня на другой, в рентгеновском — с переходом электронов из внешних оболочек в глубинные, в гамма-излучении — с ядерными переходами.

Квантовые усилители и генераторы нашли широкое применение в науке и технике. Квантовые генераторы

СВЧ диапазона характеризуются очень высокой стабильностью частоты (10^{-11} и выше). Используемые в качестве стандартов частоты эти генераторы позволили измерять время с недостижимой прежде точностью, что очень важно для астрономических и многих других научных и практических исследований. Квантовые усилители СВЧ диапазона имеют исключительно низкий уровень шумов и применение их в радиоприемных устройствах позволяет повысить чувствительность приемников более чем на два порядка. Такие усилители широко применяются в радиоастрономии, космической связи, радиолокации, радионавигации и т. д.

Квантовые приборы оптического диапазона появились позднее квантовых приборов СВЧ диапазона и были созданы в результате непосредственного использования принципов и техники квантовых усилителей и генераторов СВЧ диапазона.

Высокая степень когерентности излучения оптических квантовых генераторов и усилителей позволяет распространить радиотехнические методы управления, связи, локации и навигации на оптический диапазон.

Сдвиг в сторону коротких волн вызван требованиями повышения числа каналов и помехоустойчивости связи, создания локаторов с чрезвычайно высокой разрешающей способностью (принципиально не осуществимой на СВЧ), создания следящих и навигационных систем высокой угловой точности.

С уменьшением длины волны используемого квантового перехода коэффициент квантового усиления быстро падает [см. выражение (3-4)] и для рентгеновского диапазона и диапазона γ -излучений оказывается слишком малым, чтобы выполнялось условие усиления (3-6) или генерирования (3-7). Для усиления излучения этих диапазонов инверсная населенность должна быть намного выше, чем в оптическом диапазоне, что может быть достигнуто с помощью очень мощных источников возбуждения. В настоящее время проблема создания таких источников практически решена для рентгеновского диапазона. В диапазоне γ -излучений может быть использована ядерная реакция, но обеспечить необходимую мощность пока не представляется возможным.

Квантовый генератор — это источник когерентного излучения, основанный на использовании вынужденного излучения и обратной связи. Когерентное излучение

квантового генератора — результат создания инверсной населенности энергетических уровней под действием стороннего источника энергии возбуждения.

Квантовый усилитель — усилитель, основанный на использовании вынужденного излучения для усиления электромагнитных волн. Сторонний источник возбуждения используется для создания инверсной населенности с тем, чтобы при воздействии на активную среду слабого внешнего сигнала, обладающего определенной частотой, фазой, поляризацией и направленностью, возникло вынужденное излучение с теми же параметрами, но значительно большей мощностью.

Квантовый генератор состоит из трех элементов: активной среды, источника возбуждения и резонатора (рис. 3-3). Квантовое усиление возможно в двух режимах: режиме бегущей волны (без обратной связи) и регенеративном режиме с положительной обратной связью, для создания которой используются резонаторы.

Обычные электронные усилители и генераторы имеют аналогичные элементы: активная среда — поток электронов, кинетическая энергия которых определяется сторонним источником (источником питания), обратная связь, компенсирующая потери и поддерживающая колебания определенной частоты.

В оптических квантовых генераторах и усилителях обратная связь осуществляется с помощью открытых резонаторов, представляющих собой два зеркала с большими коэффициентами отражения, между которыми помещается активная среда. В квантовых приборах СВЧ диапазона обратная связь осуществляется объемными резонаторами.

В связи с тем что в квантовых генераторах и усилителях используется индуцированное излучение частиц вещества, квантовые приборы, работающие в диапазоне СВЧ, называют мазерами. Слово *maser* образовано из первых букв слов фразы *microwave amplification by stimulated emission of radiation*, что означает усиление на СВЧ посредством индуцированного излучения. Квантовые генераторы и усилители оптического диапазона на-

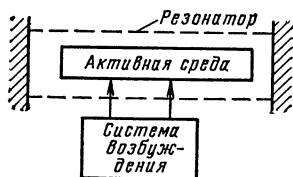


Рис. 3-3. Схема квантового генератора и регенеративного усилителя.

зываются лазерами. Слово *laser* образовано от фразы *light amplification by stimulated emission of radiation* — усиление света посредством индуцированного излучения.

3-3. СОЗДАНИЕ ИНВЕРСНОЙ НАСЕЛЕННОСТИ В СРЕДАХ

Возбудить среду можно не только с помощью стороннего электромагнитного поля, которое она поглощает, но и пропуская через нее потоки электронов или элект-

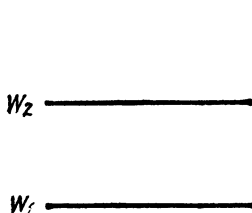


Рис. 3-4. Двухуровневая среда.

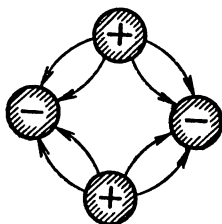


Рис. 3-5. Квадрупольный конденсатор.

рический ток, нагревая, с помощью химических реакций и т. д.

Рассмотрим двухуровневую систему (рис. 3-4). Нижний уровень W_1 соответствует стационарному невозбужденному состоянию, верхний W_2 — возбужденному. При тепловом равновесии (П1, п. 4) населенность уровней удовлетворяет соотношению

$$N_{2\text{равн}} < N_{1\text{равн}}.$$

При нагревании населенность верхнего уровня увеличивается, но остается меньшей нижнего. Инверсную населенность можно создать, удалив частицы, находящиеся на нижнем уровне. Это возможно, если частицы имеют электрические или магнитные дипольные моменты. Для сортировки частиц используются неоднородные электрические и магнитные поля. Например, в молекулярном генераторе на аммиаке инверсная населенность осуществляется за счет сортировки молекул в неоднородном электрическом поле. Сначала аммиак подогревается в специальной печи. После этого молекулы аммиака, имеющие различные энергетические уровни, про-

пускаются через неоднородное электрическое поле, создаваемое квадрупольными конденсаторами (рис. 3-5). В неоднородном электрическом поле на молекулу действует сила, которая для возбужденных молекул направлена в сторону слабого поля, для невозбужденных — в сторону сильного. При этом происходит сортировка молекул. Возбужденные молекулы концентрируются

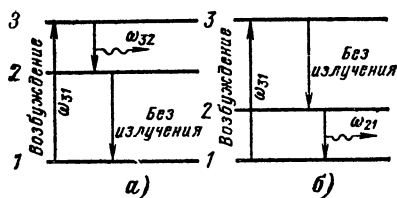


Рис. 3-6. Трехуровневые среды в диапазоне СВЧ.

вдоль оси конденсатора, невозбужденные отклоняются к его краям.

Пучок отсортированных возбужденных молекул обладает инверсной населенностью

$$N_2 > N_1.$$

Если вдоль такого пучка пропустить излучение соответствующей длины волны ($\lambda = 1,27$ см), то произойдет его усиление.

В квантовых генераторах и усилителях оптического и СВЧ диапазонов практическое применение нашли трехуровневые и четырехуровневые квантовые среды, возбуждение которых осуществляется сторонним полем излучения («накачкой»).

Рассмотрим трехуровневую среду (рис. 3-6) в диапазоне СВЧ. Считаем, что все переходы разрешены. При тепловом равновесии без облучения

$$N_{1\text{равн}} > N_{2\text{равн}} > N_{3\text{равн}}.$$

Распределение по уровням в случае отсутствия вырождения подчиняется уравнению (П1, п. 4)

$$\frac{N_{n\text{равн}}}{N_{m\text{равн}}} = e^{-\frac{W_n - W_m}{kT}}.$$

В диапазоне СВЧ при нормальных температурах

$$\frac{W_n - W_m}{kT} \ll 1$$

и

$$\frac{N_{зравн}}{N_{1равн}} \approx 1 - \frac{W_3 - W_1}{kT}; \quad (3-9)$$

$$\frac{N_{2равн}}{N_{1равн}} \approx 1 - \frac{W_2 - W_1}{kT}. \quad (3-10)$$

Если среду облучить сторонним полем частоты

$$\omega_{31} = \frac{W_3 - W_1}{\hbar},$$

соответствующей переходу между уровнями 1 и 3, то часть частиц перейдет на уровень 3 с уровня 1. Одновременно будут происходить индуцированные переходы с верхнего уровня 3 на нижний 1, сопровождающиеся излучением энергии. В диапазоне СВЧ согласно (1-43) спонтанное излучение мало. При достаточной интенсивности возбуждающего поля, так как вероятность индуцированного перехода пропорциональна плотности энергии возбуждения (1-42) и вероятности переходов с верхнего уровня на нижний и обратно одинаковы (П1-62), происходит выравнивание населенностей уровней 1 и 3. Будем считать, что населенность уровня 2 при этом не меняется. Тогда

$$N_1 = N_3 = \frac{N_{1равн} + N_{зравн}}{2} = \frac{N_{1равн}}{2} \left(1 + \frac{N_{зравн}}{N_{1равн}} \right)$$

и с учетом (3-9)

$$N_1 = N_3 = N_{1равн} \left(1 - \frac{W_3 - W_1}{2kT} \right); \quad (3-11)$$

при этом

$$N_2 = N_{1равн} \left(1 - \frac{W_2 - W_1}{kT} \right). \quad (3-12)$$

Инверсная населенность между уровнями 3 и 2 может быть получена при условии

$$N_3 > N_2,$$

или согласно (3-11) и (3-12)

$$W_2 - W_1 > \frac{W_3 - W_1}{2},$$

т. е. уровень 2 должен быть ближе к уровню 3, чем к уровню 1 (рис. 3-6, а).

Инверсная населенность между уровнями 2 и 1 может быть получена при условии

$$N_2 > N_1$$

или согласно (3-11) и (3-12)

$$W_2 - W_1 < \frac{W_3 - W_1}{2},$$

т. е. уровень 2 должен быть ближе к уровню 1, чем к уровню 3 (рис. 3-6, б).

Создание инверсной населенности накладывает определенные требования на времена релаксации. Релаксационные процессы в диапазоне СВЧ в основном определяются безызлучательными переходами, спонтанное излучение согласно (1-43) мало. Для выравнивания населенностей уровней 1 и 3 время релаксации между этими уровнями $(T_1)_{13}$ должно быть достаточно велико, иначе для выравнивания населенностей потребуются слишком большие интенсивности возбуждающего поля, что приводит к нагреванию активной среды и, как следствие этого, к увеличению шумов.

При создании инверсной населенности между уровнями 3 и 2 желательно, чтобы

$$(T_1)_{32} > (T_1)_{21}.$$

Увеличение $(T_1)_{32}$ приводит к увеличению числа частиц на уровне 3 за счет уменьшения числа безызлучательных переходов с этого уровня. Уменьшение $(T_1)_{21}$ приводит к уменьшению населенности уровня 2, так как чем меньше это время, тем больше частиц покидает этот уровень за счет безызлучательных переходов. Все это приводит к увеличению разности $N_3 - N_2$.

Если инверсная населенность создается между уровнями 2 и 1, желательно, чтобы

$$(T_1)_{32} < (T_1)_{21}.$$

В обоих рассматриваемых случаях частота источника возбуждения

$$\omega_{31} = \frac{W_3 - W_1}{\hbar}$$

должна более чем вдвое превышать частоту усиливаемого излучения

$$\omega_{32} = \frac{W_3 - W_2}{\hbar} \text{ или } \omega_{21} = \frac{W_2 - W_1}{\hbar},$$

что накладывает известные ограничения на этот метод получения инверсной населенности. С повышением частоты трудности создания мощного источника возбуждения в диапазоне СВЧ возрастают.

Возбуждение четырехуровневых сред в диапазоне СВЧ целесообразно производить на двух частотах. При этом стороннее поле одной частоты ω_{31} обогащает верхний уровень рабочего перехода, поле другой частоты ω_{42}

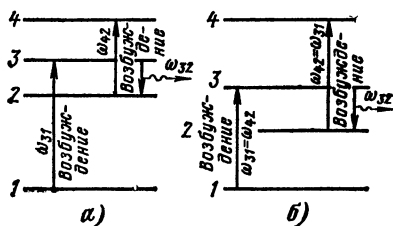


Рис. 3-7. Четырехуровневые среды в диапазоне СВЧ.

обедняет нижний уровень (рис. 3-7, а). Инверсная населенность создается между уровнями 3 и 2. Если $\omega_{31} = \omega_{42}$, то инверсная населенность в четырехуровневой среде создается полем возбуждения одной частоты (рис. 3-7, б).

В оптическом диапазоне $\frac{W_n - W_m}{kT} \gg 1$ и согласно распределению Больцмана в равновесном состоянии практически все частицы находятся на нижнем (основном) энергетическом уровне. Спонтанное излучение согласно (1-43) играет в оптическом диапазоне большую роль, чем в диапазоне СВЧ. Время жизни частиц в возбужденном состоянии порядка 10^{-6} — 10^{-7} с и уравнивать населенности уровней возможно лишь при очень больших мощностях источников поля излучения. В оптическом диапазоне энергетические уровни могут сливаться в широкие энергетические полосы, что позволяет использовать немонахроматические источники возбуждения.

Рассмотрим трехуровневую среду (рис. 3-8). Уровень 3 представляет собой полосу, что существенно повышает эффективность источника возбуждения, перебрасывающего частицы с уровня 1 на уровень 3. Уровень 2 является метастабильным. Время жизни частиц на этом уровне много больше, чем на уровне 3, и частицы с уровня 3 быстро без излучения переходят на уровень 2, накапливаясь на нем. Инверсная населенность создается

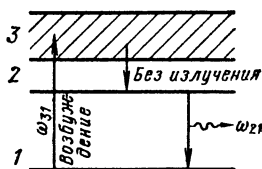


Рис. 3-8. Трехуровневая среда в оптическом диапазоне.

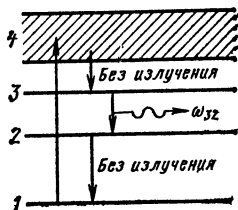


Рис. 3-9. Четырехуровневая среда в оптическом диапазоне.

между уровнями 2 и 1. Если через такую среду пропустить электромагнитное поле частоты

$$\omega_{21} = \frac{W_2 - W_1}{\hbar},$$

то произойдет усиление электромагнитного поля.

Желательно, чтобы уровень 2 находился на близком расстоянии от уровня 3, так как при этом энергия, передаваемая среде при безызлучательных переходах, мала и среда не разогревается.

Принципиальным недостатком этого метода создания инверсной населенности является необходимость использования источников возбуждения большой мощности, так как трудно создать инверсную населенность относительно уровня 1, на котором при отсутствии возбуждения находятся практически все частицы.

Четырехуровневые среды (рис. 3-9) в оптическом диапазоне требуют значительно меньшей мощности источника возбуждения. Уровень 4 представляет полосу. В результате возбуждения частицы переводятся с уровня 1 на уровень 4. Уровень 3 является метастабильным, время жизни на нем возбужденных частиц достаточно велико. Частицы с уровня 4 без излучения переходят на

уровень 3 и накапливаются на нем за счет большого времени жизни. Инверсная населенность создается между уровнем 3 и уровнем 2, который находится на значительном расстоянии от уровня 1 и поэтому при термодинамическом равновесии его заселенность близка к нулю. Создать инверсную населенность относительно этого пустого уровня гораздо легче, чем относительно сильно заселенного уровня 1.

3.4. УСИЛЕНИЕ БЕГУЩЕЙ ВОЛНЫ. ОБЩАЯ ТЕОРИЯ

Работу квантовых генераторов и усилителей можно рассматривать с позиций классической теории электромагнитного поля, так как, во-первых, мы имеем дело с взаимодействием поля с веществом и, во-вторых, излучение или поглощение определяется большим числом фотонов, когда классическое приближение достаточно точно. При этом активная среда рассматривается как квантовый ансамбль частиц (П1, п. 11) и в целом теория является полуклассической.

В квантовых усилителях бегущей волны потери в активной среде характеризуются распределенной проводимостью σ .

Волновое уравнение с учетом отсутствия свободных зарядов, согласно (1-8) будет иметь вид:

$$\Delta \mathbf{E} - \mu_0 \sigma \frac{\partial \mathbf{E}}{\partial t} - \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2}. \quad (3-13)$$

Обычно активная среда представляет собой совокупность двух компонент: основы, состоящей из частиц, не принимающих участие в излучательных квантовых переходах, и частиц активатора, определяющих излучение. При этом вектор поляризации удобно представить в виде

$$\mathbf{P} = \mathbf{P}' + \mathbf{P}^a,$$

где \mathbf{P}' — поляризация среды без учета частиц активатора; \mathbf{P}^a — поляризация, связанная с частицами активатора. Согласно (1-3)

$$\mathbf{P}' = \mathbf{D} - \epsilon_0 \mathbf{E} = \epsilon_a \mathbf{E} - \epsilon_0 \mathbf{E}, \quad (3-14)$$

где $\epsilon_a = \epsilon \epsilon_0$ — диэлектрическая проницаемость среды без учета частиц активатора.

Подставляя (3-14) в (3-13), получаем:

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} - v^2 \Delta \mathbf{E} + \frac{\sigma}{\epsilon_a} \frac{\partial \mathbf{E}}{\partial t} = - \frac{1}{\epsilon_a} \frac{\partial^2 \mathbf{P}^a}{\partial t^2}, \quad (3-15)$$

где $v = \frac{1}{\sqrt{\epsilon_a \mu_a}}$ — скорость распространения электромагнитной волны в среде без учета частиц активатора.

Зависимость поляризации P^a от поля определим квантовомеханическим путем. Активную компоненту будем рассматривать как совокупность квантовомеханических систем с двумя энергетическими уровнями W_1 и W_2 , считая, что другие уровни достаточно далеко удалены от этих двух и не взаимодействуют с ними ($W_1 < W_2$), (рис. 3-4).

Стационарные состояния 1 и 2 описываются волновыми функциями вида (П1-6)

$$\psi_{01} e^{i \frac{W_1}{\hbar} t} \text{ и } \psi_{02} e^{i \frac{W_2}{\hbar} t},$$

которые являются решениями уравнения (П1-11). Под действием электрического поля E системы переходят в нестационарные состояния, описываемые уравнением (П1-40):

$$-i\hbar \frac{\partial \psi}{\partial t} = [\hat{H}_0 + \hat{U}(t)] \psi, \quad (3-16)$$

где \hat{H}_0 — гамильтониан без учета взаимодействия, а

$$\hat{U}(t) = -\mathbf{p}_e E$$

при дипольном взаимодействии поля с частицей.

Решение (3-16) будем искать в виде волновой функции $\psi(t)$, зависящей от времени, которую, пренебрегая взаимодействием с другими уровнями, можно представить в виде

$$\psi(t) = a(t) \psi_{01} + b(t) \psi_{02}. \quad (3-17)$$

Подставляя (3-17) в (3-16) с учетом (П1-41), получаем:

$$\begin{aligned} -i\hbar \left(\frac{da}{dt} \psi_{01} + \frac{db}{dt} \psi_{02} \right) &= W_1 a \psi_{01} + W_2 b \psi_{02} - \\ &- \mathbf{p}_e E a \psi_{01} - \mathbf{p}_e E b \psi_{02}. \end{aligned} \quad (3-18)$$

Умножая обе части (3-18) на ψ_{01}^* и интегрируя по конфигурационному пространству частицы, т. е. совокупности координат, определяющих положение частицы

и ее элементарных частей в пространстве, с учетом ортогональности волновых функций (П1-14) получаем:

$$\frac{da}{dt} = \frac{j}{\hbar} (W_1 a - p_e E b), \quad (3-19)$$

где $p_e = p_{12} = \int_V \psi_{01}^* p_e \psi_{02} dV = p_{12}^*$ — матричный элемент эквивалентного дипольного момента частицы (см. П1, п. 16).

Аналогично, умножая (3-18) на ψ_{02}^* и интегрируя, получаем:

$$\frac{db}{dt} = \frac{j}{\hbar} (W_2 b - p_e E a). \quad (3-20)$$

Согласно (П1-46) матрица плотности имеет вид:

$$\|\rho_{ik}\| = \begin{vmatrix} aa^* & ab^* \\ a^*b & bb^* \end{vmatrix}. \quad (3-21)$$

Здесь недиагональные элементы определяют дипольный момент, диагональные — вероятность населенности уровней.

Среднее значение вектора поляризации P^a определяется выражением

$$P^a = \tilde{p}_e N,$$

где N — число частиц активатора в единице объема; \tilde{p}_e — средний дипольный момент частицы. Для простоты будем считать, что у матрицы дипольного момента отличны от нуля только недиагональные элементы [см. (П1-37)], т. е. $p_{12} = p_{21} = p_e$. Тогда согласно (П1-47), с учетом (3-21), получим:

$$P^a = (ab^* + a^*b) p_e N. \quad (3-22)$$

Продифференцируем выражение (3-22) по времени

$$\frac{dP^a}{dt} = \left(\frac{da}{dt} b^* + a \frac{db^*}{dt} + \frac{da^*}{dt} b + a^* \frac{db}{dt} \right) p_e N$$

и преобразуем этот результат с учетом (3-19) и (3-20). В результате получим:

$$\begin{aligned} \frac{dP^a}{dt} = & \frac{j}{\hbar} (W_1 ab^* - p_e E |b|^2 - W_2 ab^* + \\ & + p_e E |a|^2 - W_1 a^* b + p_e E |b|^2 + W_2 a^* b - p_e E |a|^2) p_e N = \end{aligned}$$

$$= j \left(\frac{W_1 - W_2}{\hbar} ab^* + \frac{|a|^2 - |b|^2}{\hbar} p_e E - \frac{W_1 - W_2}{\hbar} a^* b - \frac{|a|^2 - |b|^2}{\hbar} p_e E \right) p_e N = -j\omega_{21} (ab^* - a^* b) p_e N, \quad (3-23)$$

где $\omega_{21} = \frac{W_2 - W_1}{\hbar}$ — частота перехода.

Вторая производная поляризации определяется выражением

$$\frac{d^2 P^a}{dt^2} = -j\omega_{21} \left(\frac{da}{dt} b^* + a \frac{db^*}{dt} - \frac{da^*}{dt} b - a^* \frac{db}{dt} \right) p_e N$$

или с учетом (3-19) и (3-20)

$$\frac{d^2 P^a}{dt^2} = -\omega_{21}^2 (ab^* + a^* b) p_e N + \frac{2\omega_{21}}{\hbar} (|a|^2 - |b|^2) N E p_e^2. \quad (3-24)$$

Согласно физическому смыслу элементов матрицы плотности величина

$$(|b|^2 - |a|^2) N = \Delta N' \quad (3-25)$$

представляет собой разность населенностей верхнего и нижнего уровней при наличии электромагнитного поля.

С учетом (3-22) и (3-25) выражение (3-24) можно переписать в виде

$$\frac{d^2 P^a}{dt^2} = -\omega_{21}^2 P^a - \frac{2\omega_{21}}{\hbar} \Delta N' E p_e^2. \quad (3-26)$$

Производная по времени выражения (3-25) имеет вид:

$$\frac{d\Delta N'}{dt} = \left(\frac{db}{dt} b^* + b \frac{db^*}{dt} - \frac{da}{dt} a^* - a \frac{da^*}{dt} \right) N,$$

или с учетом (3-19) и (3-20)

$$\frac{d\Delta N'}{dt} = -\frac{2j}{\hbar} (ab^* - a^* b) p_e E N.$$

Сравнивая с (3-23), имеем

$$\frac{d\Delta N'}{dt} = \frac{2}{\hbar\omega_{21}} E \frac{dP^a}{dt}. \quad (3-27)$$

Уравнения (3-26) и (3-27) описывают изменение поляризации P^a и разности населенностей уровней $\Delta N'$ под действием электромагнитного поля. Однако эти изменения происходят не только из-за электромагнитного поля, но и за счет релаксационных процессов, действие кото-

рых определяется функциями распределения (1-50) и (1-51):

для $\Delta N'$

$$f_1(\tau) = \frac{1}{T_1} e^{-\tau/T_1};$$

для P^a

$$f_2(\tau) = \frac{1}{T_2} e^{-\tau/T_2}.$$

Обычно в твердых диэлектриках релаксация инверсной населенности, определяемая временем продольной релаксации T_1 , происходит значительно медленнее, чем релаксация поляризации, определяемая временем поперечной релаксации T_2 , т. е. $T_1 \gg T_2$. Это объясняется тем, что релаксация поляризации связана с изменением фазы волновой функции частицы, а релаксация населенности связана с существенным изменением энергии частицы. В газах при низком давлении $T_1 \approx T_2$.

Из-за процессов релаксации частицы, принимающие участие в процессе усиления в данный момент t , имеют различное время τ взаимодействия с полем. Выделяя из полной производной по времени производную по τ , получаем:

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{\partial}{\partial \tau} = \frac{\partial}{\partial t} + \frac{1}{T_{1,2}} \quad (3-28)$$

и

$$\frac{d^2}{dt^2} = \left(\frac{\partial}{\partial t} + \frac{1}{T_{1,2}} \right)^2. \quad (3-29)$$

Согласно (П1-26) среднее макроскопическое значение

$$\left. \begin{aligned} P^a &= \int_0^t P^a(\tau, t) f_2(\tau) d\tau - P_0^a; \\ \Delta N' &= \int_0^t \Delta N'(\tau, t) f_1(\tau) d\tau - \Delta N_0, \end{aligned} \right\} \quad (3-30)$$

где P_0^a и ΔN_0 — начальные значения, которые имеют величины P^a и ΔN до начала процесса усиления (при $\tau = 0$). Обычно $P_0^a = 0$.

Величина ΔN_0 определяет усиление среды. Среда усиливает, если она находится в состоянии инверсной

населенности, которая создается с помощью стороннего источника возбуждения. Эта населенность при отсутствии распространяющегося поля \mathbf{E} определяется выражением

$$\Delta N_0 = (|b^0|^2 - |a^0|^2) N,$$

где $|a^0|^2 = \rho_{11}^0$, $|b^0|^2 = \rho_{22}^0$ — диагональные элементы матрицы плотности при отсутствии поля \mathbf{E} .

Усредняя выражение (3-26) и (3-27), с учетом (3-28) и (3-29) согласно (3-30) получаем:

$$\begin{aligned} \frac{\partial^2 \mathbf{P}^a}{dt^2} + \frac{2}{T_2} \frac{\partial \mathbf{P}^a}{\partial t} + \omega_{21} \left(1 + \frac{1}{\omega_{21}^2 T_2^2} \right) \mathbf{P}^a = \\ = - \frac{2\omega_{21}}{\hbar} (\Delta N' + \Delta N_0) \mathbf{E} \rho_e^2, \\ \frac{\partial \Delta N'}{\partial t} + \frac{\Delta N'}{T_1} = \frac{2}{\hbar \omega_{21}} \mathbf{E} \left(\frac{\partial \mathbf{P}^a}{\partial t} + \frac{\mathbf{P}^a}{T_2} \right). \end{aligned}$$

Обозначая через

$$\Delta N = \Delta N' + \Delta N_0$$

разность населенностей уровней, определяемую системой возбуждения и полем \mathbf{E} , получаем:

$$\begin{aligned} \frac{\partial^2 \mathbf{P}^a}{\partial t^2} + \frac{2}{T_2} \frac{\partial \mathbf{P}^a}{\partial t} + \omega_{21} \left(1 + \frac{1}{\omega_{21}^2 T_2^2} \right) \mathbf{P}^a = - \frac{2\omega_{21}}{\hbar} \Delta N \mathbf{E} \rho_e^2, \\ \frac{\partial \Delta N}{\partial t} + \frac{\Delta N - \Delta N_0}{T_1} = \frac{2}{\hbar \omega_{21}} \mathbf{E} \left(\frac{\partial \mathbf{P}^a}{\partial t} + \frac{\mathbf{P}^a}{T_2} \right). \end{aligned}$$

Согласно последнему уравнению скорость изменения разности населенностей уровней зависит от напряженности распространяющегося поля \mathbf{E} , т. е. нелинейность квантовой среды имеет *инерционный характер*.

Вместе с уравнением (3-15) эти уравнения образуют замкнутую систему

$$\left. \begin{aligned} \frac{\partial^2 \mathbf{E}}{\partial t^2} - v^2 \Delta \mathbf{E} + \frac{\sigma}{\varepsilon_a} \frac{\partial \mathbf{E}}{\partial t} &= - \frac{1}{\varepsilon_a} \frac{\partial^2 \mathbf{P}^a}{\partial t^2}; \\ \frac{\partial^2 \mathbf{P}^a}{\partial t^2} + \frac{2}{T_2} \frac{\partial \mathbf{P}^a}{\partial t} + \omega_{21} \left(1 + \frac{1}{\omega_{21}^2 T_2^2} \right) \mathbf{P}^a &= \\ &= - \frac{2\omega_{21}}{\hbar} \Delta N \mathbf{E} \rho_e^2; \\ \frac{\partial \Delta N}{\partial t} + \frac{\Delta N - \Delta N_0}{T_1} &= \frac{2}{\hbar \omega_{21}} \mathbf{E} \left(\frac{\partial \mathbf{P}^a}{\partial t} + \frac{\mathbf{P}^a}{T_2} \right). \end{aligned} \right\} \quad (3-31)$$

Система уравнений (3-31) вместе с начальными условиями определяют распространение усиливаемого поля в активной среде.

Если поле представляет собой плоскую волну линейной поляризации, распространяющуюся в направлении оси x , то уравнения (3-31) можно представить в скалярном виде

$$\left. \begin{aligned} \frac{\partial^2 E}{\partial t^2} - v^2 \frac{\partial^2 E}{\partial x^2} + \frac{\sigma}{\epsilon_a} \frac{\partial E}{\partial t} &= - \frac{1}{\epsilon_a} \frac{\partial^2 P^a}{\partial t^2}; \\ \frac{\partial^2 P^a}{\partial t^2} + \frac{2}{T_2} \frac{\partial P^a}{\partial t} + \omega_{21}^2 \left(1 + \frac{1}{\omega_{21}^2 T_2^2} \right) P^a &= \\ &= - \frac{2\omega_{21}}{\hbar} \Delta N E p_e^2; \\ \frac{\partial \Delta N}{\partial t} + \frac{\Delta N - \Delta N_0}{T_1} &= \frac{2}{\hbar \omega_{21}} E \left(\frac{\partial P^a}{\partial t} + \frac{P^a}{T_2} \right). \end{aligned} \right\} \quad (3-32)$$

Практически любой электромагнитный процесс характеризуется спектром частот. Если напряженность поля E и поляризация P^a описываются абсолютно интегрируемыми функциями, то для них можно определить преобразование Фурье:

$$\begin{aligned} \dot{E}(x, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}(x, \omega) e^{j\omega t} d\omega, \\ \dot{S}(x, \omega) &= \int_0^{\infty} \dot{E}(x, t) e^{-j\omega t} dt; \end{aligned}$$

аналогичный вид имеют преобразования для поляризации P^a .

Введя среднюю частоту спектра ω_0 , получим:

$$\dot{E}(x, t) = \frac{1}{2\pi} e^{j\omega_0 t} \int_{-\infty}^{\infty} \dot{S}(x, \omega) e^{j(\omega - \omega_0)t} d\omega = \dot{E}_m(x, t) e^{j\omega_0 t},$$

где $\dot{E}_m(x, t)$ — комплексная амплитуда, изменяющаяся во времени и пространстве

$$\dot{E}_m(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \dot{S}(x, \omega) e^{j(\omega - \omega_0)t} d\omega.$$

В общем случае

$$\dot{E}_m(x, t) = E_m(x, t) e^{i\varphi(x, t)}.$$

Монохроматическое поле частоты ω_0 определяется выражением

$$\dot{E}(x, t) = \dot{E}_m(x) e^{i\omega_0 t},$$

где $\dot{E}_m(x)$ — комплексная амплитуда, не зависящая от времени.

Излучение квантовых приборов СВЧ характеризуется очень узким спектром частот $\Delta\omega$. Отношение $\Delta\omega/\omega \ll \ll 1$ и комплексная амплитуда $\dot{E}_m(x, t)$ медленно изменяются во времени по сравнению с $e^{i\omega_0 t}$. Такое поле называется квазимонохроматическим. Излучение лазеров характеризуется достаточно широким спектром частот $\Delta\omega$, однако отношение $\Delta\omega/\omega_0 \ll 1$ и это излучение можно также рассматривать как квазимонохроматическое.

В диапазоне СВЧ и оптическом диапазоне при распространении поля в усиливающей среде изменения амплитуды и фазы волны поля и волны поляризации малы на расстояниях порядка длины волны и за время порядка периода. Поэтому для решения системы (3-32) можно воспользоваться методом медленно меняющихся амплитуд и фаз (П2, п. 5):

$$E = E_m(t, x) \cos[\omega t - kx + \varphi(t, x)] = E_m \cos \Phi;$$

$$P^a = P_m^a(t, x) \cos[\omega t - kx + \psi(t, x)] = P_m^a \cos \Psi,$$

где $E_m(t, x)$, $P_m^a(t, x)$, $\varphi(t, x)$, $\psi(t, x)$ — медленно меняющиеся во времени и пространстве функции; $\Phi = \omega t - kx + \varphi$; $\Psi = \omega t - kx + \psi$; $k = \omega/v$.

Вычислим производные

$$\frac{\partial E}{\partial t}, \frac{\partial^2 E}{\partial t^2}, \frac{\partial^2 E}{\partial x^2}, \frac{\partial P^a}{\partial t}, \frac{\partial^2 P^a}{\partial t^2},$$

пренебрегая вследствие малости членами, содержащими вторые производные амплитуд и фаз и произведения их первых производных:

$$\begin{aligned} \frac{\partial E}{\partial t} &= \frac{\partial E_m}{\partial t} \cos \Phi - \left(\omega + \frac{\partial \varphi}{\partial t} \right) E_m \sin \Phi, \\ \frac{\partial^2 E}{\partial t^2} &= -2\omega \frac{\partial E_m}{\partial t} \sin \Phi - \omega \left(\omega + 2 \frac{\partial \varphi}{\partial t} \right) E_m \cos \Phi; \end{aligned}$$

$$\frac{\partial^2 E}{\partial x^2} = 2k \frac{\partial E_m}{\partial x} \sin \Phi - k \left(k - 2 \frac{\partial \varphi}{\partial x} \right) E_m \cos \Phi;$$

$$\frac{\partial P^a}{\partial t} = \frac{\partial P_m^a}{\partial t} \cos \Psi - \left(\omega + \frac{\partial \psi}{\partial t} \right) P_m \sin \Psi;$$

$$\frac{\partial^2 P^a}{\partial t^2} = -2\omega \frac{\partial P_m^a}{\partial t} \sin \Psi - \omega \left(\omega + 2 \frac{\partial \psi}{\partial t} \right) P_m^a \cos \Psi.$$

Подставим полученные значения в (3-32):

$$\begin{aligned} & \left[-2\omega \frac{\partial E_m}{\partial t} - 2\omega v \frac{\partial E_m}{\partial x} - \frac{\sigma}{\varepsilon_a} \left(\omega + \frac{\partial \varphi}{\partial t} \right) E_m \right] \sin \Phi - \\ & - \left[-2\omega \left(\frac{\partial \varphi}{\partial t} + v \frac{\partial \varphi}{\partial x} \right) E_m - \frac{\sigma}{\varepsilon_a} \frac{\partial E_m}{\partial t} \right] \cos \Phi = \\ & = 2 \frac{\omega}{\varepsilon_a} \frac{\partial P_m^a}{\partial t} \sin \Psi + \frac{\omega}{\varepsilon_a} \left(\omega + 2 \frac{\partial \psi}{\partial t} \right) P_m^a \cos \Psi, \\ & \left[-2\omega \frac{\partial P_m^a}{\partial t} - \frac{2}{T_2} \left(\omega + \frac{\partial \psi}{\partial t} \right) P_m^a \right] \sin \Psi + \left[-\omega \left(\omega + \right. \right. \\ & + 2 \frac{\partial \psi}{\partial t} \left. \right) P_m^a + \frac{2}{T_2} \frac{\partial P_m^a}{\partial t} + \omega_{21}^2 \left(1 + \frac{1}{\omega_{21}^2 T_2^2} \right) P_m^a \left. \right] \cos \Psi = \\ & = -\frac{2\omega_{21}}{\hbar} \Delta N p_e^2 E_m \cos \Phi, \end{aligned}$$

$$\begin{aligned} \frac{\partial \Delta N}{\partial t} + \frac{\Delta N - \Delta N_0}{T_1} &= \frac{2}{\hbar \omega_{21}} E_m \cos \Phi \left[\left(\frac{\partial P_m^a}{\partial t} + \right. \right. \\ & + \frac{P_m^a}{T_2} \left. \right) \cos \Psi - \left(\omega + \frac{\partial \psi}{\partial t} \right) P_m^a \sin \Psi \left. \right]. \end{aligned}$$

Учитывая тригонометрические соотношения

$$\Psi = \Phi + (\psi - \varphi);$$

$$\Phi = \Psi + (\varphi - \psi);$$

$$\left. \begin{aligned} \cos \Psi &= \cos (\psi - \varphi) \cos \Phi - \sin (\psi - \varphi) \sin \Phi; \\ \sin \Psi &= \cos (\psi - \varphi) \sin \Phi + \sin (\psi - \varphi) \cos \Phi; \\ \cos \Phi &= \cos (\psi - \varphi) \cos \Psi + \sin (\psi - \varphi) \sin \Psi; \\ \sin \Phi &= \cos (\psi - \varphi) \sin \Psi - \sin (\psi - \varphi) \cos \Psi; \\ 2 \sin \Phi \cos \Phi &= \sin 2\Phi; \\ \cos^2 \Phi &= \frac{1 + \cos 2\Phi}{2} \end{aligned} \right\} (3-33)$$

и приравнивая друг к другу коэффициенты при соответствующих тригонометрических функциях и свободные члены, получаем:

$$\begin{aligned}
 & \frac{\partial E_m}{\partial t} + v \frac{\partial E_m}{\partial x} + \frac{\sigma}{2\omega\epsilon_a} \left(\omega + \frac{\partial \varphi}{\partial t} \right) E_m = \\
 & = -\frac{1}{\epsilon_a} \left[\frac{\partial P_m^a}{\partial t} \cos(\psi - \varphi) - \frac{1}{2} \left(\omega + 2 \frac{\partial \psi}{\partial t} \right) P_m^a \sin(\psi - \varphi) \right], \\
 & \quad \left(\frac{\partial \varphi}{\partial t} + v \frac{\partial \varphi}{\partial x} \right) E_m - \frac{\sigma}{2\epsilon_a \omega} \frac{\partial E_m}{\partial t} = \\
 & = -\frac{1}{\epsilon_a} \left[\frac{\partial P_m^a}{\partial t} \sin(\psi - \varphi) + \frac{1}{2} \left(\omega + 2 \frac{\partial \psi}{\partial t} \right) P_m^a \cos(\psi - \varphi) \right]; \\
 & \quad \frac{\partial P_m^a}{\partial t} + \frac{1}{\omega T_2} \left(\omega + \frac{\partial \psi}{\partial t} \right) P_m^a = \frac{\omega_{21}}{\hbar \omega} \Delta N p_e^2 E_m \sin(\psi - \varphi); \\
 & \quad \frac{1}{2} \left(\omega + 2 \frac{\partial \psi}{\partial t} \right) P_m^a - \frac{1}{\omega T_2} \frac{\partial P_m^a}{\partial t} - \frac{\omega_{21}^2}{2\omega} \left(1 + \frac{1}{\omega_{21}^2 T_2^2} \right) P_m^a = \\
 & \quad = \frac{\omega_{21}}{\omega \hbar} \Delta N p_e^2 E_m \cos(\psi - \varphi); \\
 & \quad \frac{\partial \Delta N}{\partial t} + \frac{\Delta N - \Delta N_0}{T_1} = \frac{E_m}{\hbar \omega_{21}} \left[\left(\frac{\partial P_m^a}{\partial t} + \frac{P_m^a}{T_2} \right) \times \right. \\
 & \quad \times \cos(\psi - \varphi) - \left. \left(\omega + \frac{\partial \psi}{\partial t} \right) P_m^a \sin(\psi - \varphi) \right].
 \end{aligned}$$

Поскольку в оптическом диапазоне и диапазоне СВЧ выполняются условия

$$\begin{aligned}
 \omega & \gg \frac{\partial \varphi}{\partial t}, \quad \omega \gg \frac{\partial \psi}{\partial t}, \quad \omega P_m^a \gg \frac{\partial P_m^a}{\partial t}, \\
 \omega E_m & \gg \frac{\partial E_m}{\partial t}, \quad \frac{1}{T_2} \ll \omega,
 \end{aligned}$$

частота распространяющегося поля должна быть близка к частоте квантового перехода

$$\omega \approx \omega_{21};$$

полагая для диэлектрических сред с малыми потерями постоянную затухания равной

$$\alpha_n = \omega \sqrt{\mu_0 \epsilon_a} \frac{\operatorname{tg} \delta_a}{2} = \frac{\sigma}{2\epsilon_a v},$$

получаем:

$$\left. \begin{aligned} \frac{\partial E_m}{\partial t} + v \frac{\partial E_m}{\partial x} + v \alpha_n E_m &= \frac{\omega}{2\epsilon_a} P_m^a \sin(\psi - \varphi); \\ \left(\frac{\partial \varphi}{\partial t} + v \frac{\partial \varphi}{\partial x} \right) E_m &= - \frac{\omega}{2\epsilon_a} P_m^a \cos(\psi - \varphi); \\ \frac{\partial P_m^a}{\partial t} + \frac{P_m^a}{T_2} &= \frac{\rho_e^2}{\hbar} \Delta N E_m \sin(\psi - \varphi); \\ \left[\frac{\partial \psi}{\partial t} + (\omega - \omega_{21}) \right] P_m^a &= \frac{\rho_e^2}{\hbar} \Delta N E_m \cos(\psi - \varphi); \\ \frac{\partial \Delta N}{\partial t} + \frac{\Delta N - \Delta N_0}{T_1} &= - \frac{1}{\hbar} P_m^a E_m \sin(\psi - \varphi). \end{aligned} \right\} \quad (3-34)$$

Решение этой системы уравнений позволяет определить амплитуды и фазы поляризации и напряженности поля, а также величину инверсной населенности.

3-5. УСИЛЕНИЕ В РЕЗОНАТОРЕ. ОБЩАЯ ТЕОРИЯ

В квантовом регенеративном усилителе или генераторе потери определяем распределенной проводимостью σ . Эта величина характеризует потери в активной среде, в стенках или зеркалах резонатора и потери, связанные с излучением энергии. В этом случае электромагнитное поле, как и в случае усилителя бегущей волны, определяется первым уравнением (3-31)

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} - v^2 \Delta \mathbf{E} + \frac{\sigma}{\epsilon_a} \frac{\partial \mathbf{E}}{\partial t} = - \frac{1}{\epsilon_a} \frac{\partial^2 \mathbf{P}^a}{\partial t^2}, \quad (3-35)$$

где \mathbf{P}^a — вектор поляризации, связанный с частицами активатора; ϵ_a — диэлектрическая проницаемость среды без учета частиц активатора, $v = \frac{1}{\sqrt{\epsilon_a \mu_0}}$.

Однако решение этого уравнения при наличии резонатора следует искать в виде пространственных стоячих волн

$$\mathbf{E}(x_1, x_2, x_3, t) = \sum_n U_n(x_1, x_2, x_3) E_n(t),$$

где $U_n(x_1, x_2, x_3)$ — функция, определяющая пространственное распределение поля n -го типа в резонаторе.

Обычно в резонаторах лазеров и мазеров поле зависит лишь от продольной координаты и слабо зависит от поперечных координат, поэтому для любого типа колебаний можно записать

$$E(x_3, t) = \sum_n E_n(t) \sin k_n x_3, \quad (3-36)$$

где n — номер типа колебаний; $k_n = n\pi/L$ — продольное волновое число; L — длина резонатора.

Подставляя (3-36) в (3-35), получаем:

$$\sum_n \sin k_n x_3 \left[\frac{d^2 E_n(t)}{dt^2} + \frac{\sigma}{\epsilon_a} \frac{dE_n(t)}{dt} + v^2 k_n^2 E_n(t) \right] = - \frac{1}{\epsilon_a} \frac{\partial^2 P^a(x_3, t)}{\partial t^2}.$$

Умножая правую и левую части уравнения на $\sin k_n x_3$ и усредняя по длине резонатора, получаем:

$$\frac{d^2 E_n(t)}{dt^2} + \frac{\sigma}{\epsilon_a} \frac{dE_n(t)}{dt} + v^2 k_n^2 E_n(t) = - \frac{1}{\epsilon_a} \frac{d^2 P_n^a(t)}{dt^2}, \quad (3-37)$$

где

$$P_n^a(t) = \frac{2}{L} \int_0^L P^a(x_3, t) \sin k_n x_3 dx_3.$$

Собственные колебания резонатора при наличии потерь характеризуются комплексными величинами

$$\dot{\omega}_n = \omega_n + j \frac{\omega_n}{2Q_n},$$

где

$$Q_n = \frac{\epsilon_a'(\omega_n)}{\epsilon_a''(\omega_n)} = \frac{\epsilon_a'(\omega_n) \omega_n}{\sigma(\omega_n)} \quad (3-38)$$

— добротность n -го типа колебаний;

$$\omega_n = k_n^2 v^2 \quad (3-39)$$

— резонансная частота n -го типа колебаний;

$$\omega_{nc} = \omega_n \sqrt{1 + \frac{1}{4Q_n^2}} \quad (3-40)$$

— собственная частота n -го типа колебаний.

Подставляя (3-38) и (3-39) в (3-37), получаем уравнение, которое вместе с вторым и третьим уравнениями системы (3-31) образуют систему уравнений для исследования квантового генератора:

$$\left. \begin{aligned} \frac{d^2 E_n(t)}{dt^2} + \frac{\omega_n}{Q_n} \frac{dE_n(t)}{dt} + \omega_n^2 E_n(t) &= -\frac{1}{\varepsilon_a} \frac{d^2 P_n^a(t)}{dt^2}; \\ \frac{\partial^2 P^a}{\partial t^2} + \frac{2}{T_2} \frac{\partial P^a}{\partial t} + \omega_{21}^2 \left(1 + \frac{1}{\omega_{21}^2 T_2^2}\right) P^a &= -\frac{2\omega_{21}}{\hbar} \Delta N E p_e^2; \\ \frac{\partial \Delta N}{\partial t} + \frac{\Delta N - \Delta N_0}{T_1} &= \frac{2}{\hbar \omega_{21}} E \left(\frac{\partial P^a}{\partial t} + \frac{P^a}{T_2} \right), \end{aligned} \right\} \quad (3-41)$$

где

$$E = E(x_3 t) = \sum_n E_n(t) \sin k_n x_3;$$

$$P_n^a(t) = \frac{2}{L} \int_0^L P^a(x_3 t) \sin k_n x_3 dx_3.$$

Ограничимся рассмотрением простейшего случая, когда в генераторе возбуждается только один тип колебаний с линейной поляризацией поля, направление векторов поля и поляризации совпадает, а распределения напряженности поля, поляризации и инверсной населенности равномерны по длине резонатора. В этом случае система (3-41) может быть переписана в скалярной форме:

$$\left. \begin{aligned} \frac{d^2 E}{dt^2} + \frac{\omega_p}{Q_p} \frac{dE}{dt} + \omega_p^2 E &= -\frac{1}{\varepsilon_a} \frac{d^2 P^a}{dt^2}; \\ \frac{d^2 P^a}{dt^2} + \frac{2}{T_2} \frac{dP^a}{dt} + \omega_{21}^2 \left(1 + \frac{1}{\omega_{21}^2 T_2^2}\right) P^a &= -\frac{2\omega_{21}}{\hbar} \Delta N E p_e^2; \\ \frac{d\Delta N}{dt} + \frac{\Delta N - \Delta N_0}{T_1} &= \frac{2}{\hbar \omega_{21}} E \left(\frac{dP^a}{dt} + \frac{P^a}{T_2} \right), \end{aligned} \right\} \quad (3-42)$$

где ω_p и Q_p — резонансная частота и добротность резонатора для выделенного единственного типа колебаний.

Уравнения (3-42) описывают колебания поля в резонаторе с частотой ω_p и колебания поляризации с частотой квантового перехода ω_{21} . Эти колебания можно рассматривать как колебания в двух связанных системах. Поэтому если ω_p и ω_{21} не точно равны друг другу, то в

системе устанавливаются колебания с некоторой промежуточной частотой ω . Это явление называется затягиванием частоты.

Представим электрическое поле E и поляризацию Pa в виде

$$\left. \begin{aligned} E &= \frac{\dot{E}_m(t) e^{j\omega t} + \dot{E}_m^*(t) e^{-j\omega t}}{2}; \\ Pa &= \frac{\dot{P}_m^a(t) e^{j\omega t} + \dot{P}_m^{a*}(t) e^{-j\omega t}}{2}, \end{aligned} \right\} \quad (3-43)$$

где $\dot{E}_m(t)$ и $\dot{P}_m(t)$ — медленно меняющиеся комплексные амплитуды; ω — частота оптического или СВЧ диапазона.

Величина

$$\omega_{21} \sqrt{1 + \frac{1}{\omega_{21}^2 T_2^2}} = \omega_{\text{л}}, \quad (3-44)$$

входящая во второе уравнение системы (3-41), представляет собой собственную частоту линии излучения рабочего перехода. Действительно, величина $2/T_2$ определяет ширину линии рабочего перехода; ω_{21} — резонансная частота и

$$Q_{\text{л}} = \frac{\omega_{21} T_2}{2}$$

— добротность линии рабочего перехода. Отсюда аналогично (3-40)

$$\omega_{\text{л}} = \omega_{21} \sqrt{1 + \frac{1}{4Q_{\text{л}}^2}}.$$

Подставляя (3-43) в уравнение для поля и поляризации системы (3-42), с учетом (3-44) получаем:

$$\left[(\omega^2 - \omega_{\text{л}}^2) (\omega^2 - \omega_{\text{р}}^2) - \frac{2\omega^2 \omega_{\text{р}}}{T_2 Q_{\text{р}}} \right] - j \left[\frac{\omega \omega_{\text{р}}}{Q_{\text{р}}} (\omega^2 - \omega_{\text{л}}^2) + \right. \\ \left. + \frac{2\omega}{T_2} (\omega^2 - \omega_{\text{р}}^2) \right] = - \frac{2\omega_{21} \omega^2}{\hbar \epsilon_a} \Delta N p_e^2.$$

Поскольку правая часть полученного уравнения является действительной величиной, то мнимая часть должна быть равна нулю, откуда получаем:

$$\frac{\omega_{\text{р}}}{Q_{\text{р}}} (\omega^2 - \omega_{\text{л}}^2) + \frac{2}{T_2} (\omega^2 - \omega_{\text{р}}^2) = 0,$$

или, так как $\omega \approx \omega_{\pi} \approx \omega_{21}$ и $\omega \approx \omega_p$, то

$$(\omega - \omega_{21}) \approx \frac{2}{T_2} \frac{Q_p}{\omega_p} (\omega_p - \omega) = \frac{\Delta\omega_{\pi}}{\Delta\omega_p} (\omega_p - \omega). \quad (3-45)$$

Величина $2/T_2 = \Delta\omega_{\pi}$ определяет ширину линии излучения рабочего перехода, величина $\omega_p/Q_p = \Delta\omega_p$ — ширину резонансной кривой резонатора. В зависимости от соотношения полосы резонатора и линии частота ω расположена ближе к частоте перехода ω_{21} или к резонансной частоте ω_p . Для регенеративного усилителя частоту сигнала, а для генератора собственную частоту резонансного пика подстраивают под частоту линии рабочего перехода.

Применяя метод медленно меняющихся амплитуд и фаз, решение системы (3-42) будем искать в виде

$$E = E_m(t) \cos[\omega_{\pi} t + \varphi(t)] = E_m \cos \Phi,$$

$$P^a = P_m^a(t) \cos[\omega_{\pi} t + \psi(t)] = P_m^a \cos \Psi,$$

где $E_m(t)$, $P_m^a(t)$, $\varphi(t)$, $\psi(t)$ — медленно меняющиеся во времени амплитуды и фазы: $\Phi = \omega_{\pi} t + \varphi(t)$, $\Psi = \omega_{\pi} t + \psi(t)$.

Вычислим производные

$$\frac{dE}{dt}, \frac{d^2E}{dt^2}, \frac{dP^a}{dt} \text{ и } \frac{d^2P^a}{dt^2},$$

пренебрегая вследствие малости членами, содержащими вторые производные амплитуд и фаз и произведения их первых производных:

$$\frac{dE}{dt} = \frac{dE_m}{dt} \cos \Phi - \left(\omega_{\pi} + \frac{d\varphi}{dt} \right) E_m \sin \Phi;$$

$$\frac{d^2E}{dt^2} = -2\omega_{\pi} \frac{dE_m}{dt} \sin \Phi - \omega_{\pi} \left(\omega_{\pi} + 2 \frac{d\varphi}{dt} \right) E_m \cos \Phi;$$

$$\frac{dP^a}{dt} = \frac{dP_m^a}{dt} \cos \Psi - \left(\omega_{\pi} + \frac{d\psi}{dt} \right) P_m^a \sin \Psi;$$

$$\frac{d^2P^a}{dt^2} = -2\omega_{\pi} \frac{dP_m^a}{dt} \sin \Psi - \omega_{\pi} \left(\omega_{\pi} + 2 \frac{d\psi}{dt} \right) P_m^a \cos \Psi.$$

Подставляя полученные значения в (3-42) и пренебрегая малыми членами, получаем:

$$\left. \begin{aligned} & \left[(\omega_p^2 - \omega_\pi^2) - 2\omega_\pi \frac{d\varphi}{dt} \right] E_m \cos \Phi - \left(2\omega_\pi \frac{dE_m}{dt} + \right. \\ & \quad \left. + \frac{\omega_\pi \omega_p}{Q_p} E_m \right) \sin \Phi = \frac{\omega_\pi^2 P_m^a}{\varepsilon_a} \cos \Psi; \\ & - 2\omega_\pi \frac{d\Psi}{dt} P_m^a \cos \Psi - 2\omega_\pi \left(\frac{dP_m^a}{dt} + \frac{1}{T^2} P_m^a \right) \sin \Psi = \\ & \quad = - \frac{2\omega_\pi}{\hbar} \Delta N p_e^2 E_m \cos \Phi; \\ & \frac{d\Delta N}{dt} + \frac{\Delta N - \Delta N_0}{T_1} = - \frac{2}{\hbar \omega_{21}} E_m P_m^a \cos \Phi \sin \Psi. \end{aligned} \right\} (3-46)$$

Учитывая в (3-46) тригонометрические соотношения (3-33) и приравнявая коэффициенты при соответствующих косинусах, синусах и свободные члены, получаем:

$$\left. \begin{aligned} & \left[(\omega_p^2 - \omega_\pi^2) - 2\omega_\pi \frac{d\varphi}{dt} \right] E_m = \frac{\omega_\pi^2 P_m^a}{\varepsilon_a} \cos (\psi - \varphi); \\ & 2 \frac{dE_m}{dt} + \frac{\omega_p}{Q_p} E_m = \frac{\omega_\pi P_m^a}{\varepsilon_a} \sin (\psi - \varphi); \\ & P_m^a \frac{d\psi}{dt} = \frac{p_e^2}{\hbar} \Delta N E_m \cos (\psi - \varphi); \\ & \frac{dP_m^a}{dt} + \frac{1}{T_2} P_m^a = \frac{p_e^2}{\hbar} \Delta N E_m \sin (\psi - \varphi); \\ & \frac{d\Delta N}{dt} + \frac{\Delta N - \Delta N_0}{T_1} = - \frac{E_m P_m^a}{\hbar} \sin (\psi - \varphi). \end{aligned} \right\} (3-47)$$

Так как $\omega_\pi \approx \omega_p$, то

$$\frac{\omega_p^2 - \omega_\pi^2}{2\omega_\pi} = \frac{(\omega_p - \omega_\pi)(\omega_p + \omega_\pi)}{2\omega_\pi} \approx \omega_p - \omega_\pi.$$

Учитывая это соотношение в первом уравнении системы (3-47), перепишем систему в виде

$$\left. \begin{aligned}
 \frac{dE_m}{dt} &= -\frac{\omega_p}{2Q_p} E_m + \frac{\omega_l}{2\varepsilon_a} P_m^a \sin(\psi - \varphi); \\
 \frac{dP_m^a}{dt} &= -\frac{1}{T_2} P_m^a + \frac{p_e^2}{\hbar} \Delta N E_m \sin(\psi - \varphi); \\
 \frac{d\Delta N}{dt} + \frac{\Delta N - \Delta N_0}{T_1} &= -\frac{E_m P_m^a}{\hbar} \sin(\psi - \varphi); \\
 \frac{d\varphi}{dt} &= (\omega_p - \omega_l) - \frac{\omega_l}{2\varepsilon_a} \frac{P_m^a}{E_m} \cos(\psi - \varphi); \\
 \frac{d\psi}{dt} &= \frac{p_e^2}{\hbar} \Delta N \frac{E_m}{P_m^a} \cos(\psi - \varphi).
 \end{aligned} \right\} (3-48)$$

Система (3-48) описывает как стационарные, так и нестационарные процессы в генераторе.

3-6. ШУМЫ КВАНТОВЫХ УСИЛИТЕЛЕЙ

Шумы определяют чувствительность усилителя (приемника), т. е. минимальную мощность сигнала на входе приемника, при которой на выходе приемника получается минимально различимый сигнал. Различают внешние и собственные шумы усилителя. Внешние шумы связаны с некогерентным излучением окружающих нагретых объектов и деталей усилителя. Мощность этого излучения в единичном интервале частот при термодинамическом равновесии определяется формулой Планка

$$P_{\text{ш}} = \frac{\hbar\omega}{e^{\frac{\hbar\omega}{kT}} - 1}, \quad (3-49)$$

где T — температура излучающего объекта.

Собственные шумы квантовых усилителей обусловлены флуктуациями числа квантов индуцированного излучения и спонтанным излучением.

На входе в единичном интервале частот мощность шума определяется выражением (3-49). При распространении внутри усилителя бегущей волны изменение шума определяется выражением

$$dP_{\text{ш}} = 2\alpha_{\text{ус}} P_{\text{ш}} dx - 2\alpha_{\text{п}} P_{\text{ш}} dx + \beta_1 P_{\text{ш.бок}} dx + \beta_2 P_{\text{спонт}} dx. \quad (3-50)$$

Первый член правой части этого уравнения определяет усиление входного шума за счет индуцированного излучения в активной среде на элементе длины dx . Согласно (3-2) коэффициент усиления среды по мощности $2\alpha_{yc} = (N_2 - N_1)B$, где $(N_2 - N_1)$ — плотность инверсной населенности,

$$B = \frac{\omega \hbar}{cn} B_{nm} g(\omega).$$

Второй член определяет поглощение шума на элементе длины dx , $2\alpha_n$ — коэффициент поглощения среды по мощности. Коэффициент α_n влияет и на потери усиливаемого сигнала, поэтому выбирают среды с очень малым коэффициентом потерь, улучшая их параметры за счет технологий.

Третий член в выражении (3-50) определяет увеличение шума за счет теплового излучения, проникающего в среду через боковую поверхность. Это может быть, например, тепловое излучение стенок волновода, внутри которого находится активная среда. Здесь β_1 — коэффициент пропорциональности, $P_{ш.бок}$ — мощность тепловых шумов в единичном интервале частот при температуре T . Понижая температуру усилителя за счет охлаждения до температуры жидкого гелия, можно существенно понизить уровень шумов.

Последний член правой части (3-50) определяет увеличение мощности шумов за счет спонтанного излучения. Здесь β_2 — коэффициент пропорциональности, $P_{спонт}$ — мощность спонтанного излучения в единичном интервале частот. Согласно (1-41) мощность спонтанного излучения пропорциональна населенности верхнего уровня рабочего перехода N_2 .

Таким образом, пренебрегая вторым и третьим слагаемыми, выражение (3-50) можно представить в виде

$$dP_{ш} = (N_2 - N_1)BP_{ш}dx + AN_2dx, \quad (3-51)$$

где коэффициент пропорциональности A еще предстоит определить.

Если источник возбуждения отключен, то $dP_{ш} = 0$, все элементы усилителя находятся в состоянии термодинамического равновесия при $T = T_0$, и уравнение (3-51) будет иметь вид:

$$(N_{2равн} - N_{1равн})BP_{ш0}dx + N_{2равн}Adx = 0. \quad (3-52)$$

Здесь $N_{2\text{равн}}$ и $N_{1\text{равн}}$ — населенности рабочих уровней 2 и 1 в состоянии термодинамического равновесия при $T=T_0$, $P_{\text{ш0}}$ определяется (3-49) при $T=T_0$. Согласно распределению Больцмана (П1-3)

$$\frac{N_{2\text{равн}}}{N_{1\text{равн}}} = e^{-\frac{\hbar\omega}{kT_0}}. \quad (3-53)$$

Учитывая (3-53) в (3-52), получаем:

$$A = BP_{\text{ш0}} \left[e^{\frac{\hbar\omega}{kT_0}} - 1 \right],$$

а с учетом (3-49)

$$A = \hbar\omega B. \quad (3-54)$$

Подставляя (3-54) в (3-51), окончательно получаем:

$$dP_{\text{ш}} = (N_2 - N_1) BP_{\text{ш}} dx + \hbar\omega N_2 B dx. \quad (3-55)$$

Обозначая

$$\frac{\hbar\omega N_2}{N_2 - N_1} = a \quad (3-56)$$

и вводя новую переменную

$$P'_{\text{ш}} = P_{\text{ш}} + a,$$

перепишем уравнение (3-55) в виде

$$dP'_{\text{ш}} = (N_2 - N_1) BP'_{\text{ш}} dx.$$

Интегрируя по x , получаем:

$$P'_{\text{ш}}(x) = P'_{\text{ш}}(0) e^{(N_2 - N_1) Bx}$$

или

$$\frac{P'_{\text{ш}}(l)}{P'_{\text{ш}}(0)} = e^{2\gamma_{\text{ус}} l} = G,$$

где G — коэффициент усиления усилителя по мощности; l — длина активной среды усилителя.

Возвращаясь к переменной $P_{\text{ш}}$, получаем:

$$\frac{P_{\text{ш}}(l)}{P_{\text{ш}}(0)} = \frac{P_{\text{ш.вых}}}{P_{\text{ш.вх}}} = G + (G - 1) \frac{a}{P_{\text{ш.вх}}}.$$

Подставляя сюда значение a из выражения (3-56), будем иметь:

$$P_{\text{ш.вых}} = GP_{\text{ш.вх}} + (G - 1) \frac{\hbar\omega}{1 - \frac{N_1}{N_2}}. \quad (3-57)$$

Здесь первое слагаемое определяется внешним шумом, второе — собственным или квантовым шумом усилителя.

При $G \gg 1$ и $N_2 \gg N_1$ выражение (3-57) с учетом (3-49) будет иметь вид:

$$P_{\text{ш.вых}} = G \left[\frac{\hbar\omega}{e^{\frac{\hbar\omega}{kT}} - 1} + \hbar\omega \right], \quad (3-58)$$

где T — температура источников внешних шумов.

Работу усилителя с низким уровнем шумов удобно характеризовать эквивалентной шумовой температурой усилителя T_{yc} . Это температура, до которой должен быть нагрет источник теплового шума на входе такого же, но нешумящего усилителя для получения того же уровня шума на выходе. Для определения эквивалентной шумовой температуры можно воспользоваться выражением

$$P_{\text{ш.вых}} = G \frac{\hbar\omega}{e^{\frac{\hbar\omega}{kT_{\text{yc}}}} - 1}. \quad (3-59)$$

Приравнявая (3-58) и (3-59), получаем:

$$e^{\frac{\hbar\omega}{kT_{\text{yc}}}} = 2 - e^{-\frac{\hbar\omega}{kT}};$$

отсюда

$$T_{\text{yc}} = \frac{\hbar\omega}{k \ln(2 - e^{-\hbar\omega/kT})}. \quad (3-60)$$

В диапазоне СВЧ вплоть до миллиметровых волн $\hbar\omega \gg kT$ и

$$e^{-\frac{\hbar\omega}{kT}} \approx 1 - \frac{\hbar\omega}{kT} + \frac{1}{2} \left(\frac{\hbar\omega}{kT} \right)^2;$$

$$\ln \left[1 + \frac{\hbar\omega}{kT} - \frac{1}{2} \left(\frac{\hbar\omega}{kT} \right)^2 \right] \approx \frac{\hbar\omega}{kT} - \left(\frac{\hbar\omega}{kT} \right)^2,$$

поэтому выражение (3-60) может быть преобразовано к виду

$$T_{\text{yc}} \approx T + \frac{\hbar\omega}{k}. \quad (3-61)$$

Добавка $\hbar\omega/k$ определяет собственные квантовые шумы идеального усилителя. Подстановка $\hbar = \frac{6,62}{2\pi} \cdot 10^{-34}$ Дж·с, $k = 1,38 \cdot 10^{-23}$ Дж/К и $\omega = 2\pi \cdot 10^{10}$ с⁻¹ дает $\hbar\omega/k = 0,5$ К. Таким образом, в СВЧ диапазоне собственные шумы усилителя малы и шумовая эквивалентная температура усилителя определяется шумовой температурой внешних источников, которые имеют тепловой характер.

Поэтому для снижения общего уровня шума квантовые усилители приходится охлаждать до низких температур.

В оптическом диапазоне $\hbar\omega \gg kT$ и согласно (3-60)

$$T_{yc} = \frac{\hbar\omega}{k \ln 2} = 1,44 \frac{\hbar\omega}{k}, \quad (3-62)$$

т. е. основной вклад в шумовую температуру усилителя на оптических частотах дают собственные квантовые шумы усилителя, которые существенно больше внешних шумов.

Действительно, при $\omega = 2\pi \cdot 5 \cdot 10^{14}$ с⁻¹ условие $\hbar\omega \gg kT$ выполняется при любой реально используемой температуре и согласно (3-62) $T_{yc} = 3,6 \cdot 10^4$ К.

Шумы любого усилителя можно также характеризовать коэффициентом шума, определяемым выражением

$$K_{ш} = \frac{P_{с.вх}/P_{ш.вх}}{P_{с.вых}/P_{ш.вых}}, \quad (3-63)$$

где $P_{с.вх}/P_{ш.вх}$ — отношение мощностей сигнала и шума на входе усилителя; $P_{с.вых}/P_{ш.вых}$ — отношение тех же величин на выходе усилителя.

Выражение (3-63) можно представить в виде

$$K_{ш} = \frac{P_{ш.вых}}{GP_{ш.вх}}, \quad (3-64)$$

где $G = P_{с.вых}/P_{с.вх}$ — коэффициент усиления. Согласно (3-64) коэффициент шума определяется как отношение мощности шумов на выходе усилителя к мощности шумов на входе идеального усилителя, собственные шумы в котором отсутствуют. Для идеального усилителя $K_{ш} = 1$.

Поскольку коэффициент усиления зависит от частоты, коэффициент шума $K_{ш}$ также зависит от частоты.

ты. Обычно коэффициент шума минимален в центре полосы пропускания усилителя и возрастает к ее краям.

Подставляя выражение (3-57) в (3-64), получаем:

$$K_{ш} = 1 + \left(1 - \frac{1}{G}\right) \frac{\hbar\omega}{P_{ш.вх} \left(1 - \frac{N_1}{N_2}\right)}.$$

При $G \gg 1$

$$K_{ш} = 1 + \frac{\hbar\omega}{P_{ш.вх} \left(1 - \frac{N_1}{N_2}\right)}$$

и с учетом (3-49)

$$K_{ш} = 1 + \left(e^{\frac{\hbar\omega}{kT}} - 1\right) \frac{1}{\left(1 - \frac{N_1}{N_2}\right)}. \quad (3-65)$$

В диапазоне СВЧ $\hbar\omega \ll kT$ и согласно (3-65)

$$K_{ш} = 1 + \frac{N_2}{N_2 - N_1} \frac{\hbar\omega}{kT},$$

при $N_2 \gg N_1$

$$K_{ш} = 1 + \frac{\hbar\omega}{kT}, \quad (3-66)$$

т. е. близок к единице.

Согласно (3-66) и (3-61) коэффициент шума связан с эквивалентной шумовой температурой усилителя соотношением

$$T_{yc} = K_{ш} T. \quad (3-67)$$

Глава четвертая

ОПТИЧЕСКИЕ КВАНТОВЫЕ ГЕНЕРАТОРЫ И УСИЛИТЕЛИ

4-1. КОГЕРЕНТНОСТЬ ИЗЛУЧЕНИЯ

Важнейшей характеристикой излучения квантовых генераторов и усилителей оптического диапазона является *когерентность*.

Под когерентностью понимают корреляцию (связь) характеристик поля, рассматриваемых в разных точках пространства в разные моменты времени.

Полностью когерентное и полностью некогерентное излучение является абстракцией. Практически любое излучение является частично когерентным.

Эффективность сложения излучений зависит от степени когерентности. В случае полностью когерентных полей складываются напряженности поля. Энергия суммарного поля когерентных источников в пространстве изменяется от точки к точке, что приводит к образованию интерференционных максимумов и минимумов (интерференционные полосы). При одинаковой линейной поляризации источников в точках, где колебания находятся в фазе, суммарная энергия пропорциональна величине $(\sum_{i=1}^n E_i)^2$, и в частном случае, при n одинаковых источниках, — $n^2 E^2$. Когерентность колебаний приводит к перераспределению энергии в пространстве, но не влияет на ее величину.

В случае полностью некогерентных источников складываются мощности. Суммарная энергия полностью некогерентных источников в любой точке поля пропорциональна $\sum_{i=1}^n E_i^2$, и в частном случае, при n одинаковых источниках, пропорциональна $n E^2$.

Теория частичной когерентности в последнее время стала весьма актуальной в связи с разработкой лазеров — источников вынужденного излучения в оптическом диапазоне, степень когерентности которых значительно выше, чем у обычных источников света, но ниже, чем у генераторов радиодиапазона.

Поле излучения в точке $P(\mathbf{r})$ аналитически определяется волновой функцией $\dot{V}(\mathbf{r}, t)$. Мгновенная энергия поля (мгновенная интенсивность) характеризуется величиной

$$I(\mathbf{r}, t) = \dot{V}(\mathbf{r}, t) \dot{V}^*(\mathbf{r}, t). \quad (4-1)$$

Излучение лазеров удовлетворяет условию

$$\frac{\Delta\omega}{\omega_0} \ll 1,$$

где $\Delta\omega$ — ширина спектра излучения; ω_0 — центральная частота спектра, и является квазимонохроматическим

(см. § 3-4). Такое излучение можно характеризовать волновой функцией

$$\dot{V}(\mathbf{r}, t) = \dot{V}_m(\mathbf{r}, t) e^{i\omega_0 t}, \quad (4-2)$$

где

$$\dot{V}_m(\mathbf{r}, t) = V_m(\mathbf{r}, t) e^{i\varphi(\mathbf{r}, t)} \quad (4-3)$$

— комплексная амплитуда, медленно по сравнению с периодом колебания и длиной волны изменяющаяся во времени и пространстве.

В оптическом диапазоне рассматривают усредненные во времени величины, так как вследствие высокой частоты приемники могут регистрировать лишь величины, усредненные по времени наблюдения T , много большему, чем период колебания $1/f$. При этом

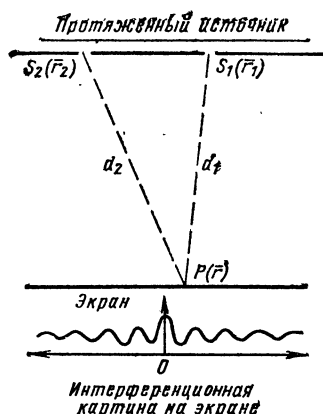


Рис. 4-1. Сложение полей двух источников.

$$\langle \dot{V}(\mathbf{r}, t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \dot{V}(\mathbf{r}, t) dt = 0.$$

Однако среднее значение $\langle \dot{V}(\mathbf{r}, t) \dot{V}^*(\mathbf{r}, t) \rangle$, характеризующее среднюю энергию поля (среднюю интенсивность), и регистрируемое прибором, отлично от нуля, т. е.

$$\begin{aligned} \langle I(\mathbf{r}, t) \rangle &= \langle \dot{V}(\mathbf{r}, t) \dot{V}^*(\mathbf{r}, t) \rangle = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \dot{V}(\mathbf{r}, t) \dot{V}^*(\mathbf{r}, t) dt \neq 0 \end{aligned}$$

(здесь угловыми скобками обозначено усреднение во времени).

Рассмотрим сложение полей двух источников $S_1(\mathbf{r}_1)$ и $S_2(\mathbf{r}_2)$ (рис. 4-1) в точке пространства $P(\mathbf{r})$. Мгновенное значение суммарного колебания в момент времени t определяется волновой функцией

$$\dot{V}(\mathbf{r}, t) = k_1 \dot{V}_1(\mathbf{r}_1, t - t_1) + k_2 \dot{V}_2(\mathbf{r}_2, t - t_2), \quad (4-4)$$

где $t_1 = d_1/v$, $t_2 = d_2/v$ — время распространения излучения от источников S_1 и S_2 до точки сложения $P(\mathbf{r})$; v — скорость распространения; d_1 и d_2 — расстояния от соответствующих источников до точки сложения; $\dot{V}_1(\mathbf{r}_1, t-t_1)$ и $\dot{V}_2(\mathbf{r}_2, t-t_2)$ — волновые функции соответственно на поверхности источников S_1 в момент времени $t-t_1$ и S_2 в момент времени $t-t_2$; \mathbf{r}_1 и \mathbf{r}_2 — координаты источников S_1 и S_2 ; k_1 и k_2 — постоянные коэффициенты, зависящие от размеров и геометрии излучающей поверхности источников S_1 и S_2 , а также расстояний d_1 и d_2 (рис. 4-1).

Средняя интенсивность поля в точке $P(\mathbf{r})$ определяется выражением

$$\begin{aligned} \langle I(\mathbf{r}, t) \rangle \approx \frac{1}{T} \int_0^T [k_1 \dot{V}_1(\mathbf{r}_1, t-t_1) + k_2 \dot{V}_2(\mathbf{r}_2, t-t_2)] \times \\ \times [k_1^* \dot{V}_1^*(\mathbf{r}_1, t-t_1) + k_2^* \dot{V}_2^*(\mathbf{r}_2, t-t_2)] dt. \end{aligned} \quad (4-5)$$

Так как $\dot{V}_1(\mathbf{r}_1, t-t_1)$ и $\dot{V}_2(\mathbf{r}_2, t-t_2)$ — комплексные величины, то

$$\begin{aligned} k_1 k_2^* \dot{V}_1(\mathbf{r}_1, t-t_1) \dot{V}_2^*(\mathbf{r}_2, t-t_2) + \\ + k_1^* k_2 \dot{V}_1^*(\mathbf{r}_1, t-t_1) \dot{V}_2(\mathbf{r}_2, t-t_2) = \\ = 2\text{Re } k_1 k_2^* \dot{V}_1(\mathbf{r}_1, t-t_1) \dot{V}_2^*(\mathbf{r}_2, t-t_2) = \\ = 2\text{Re } k_1^* k_2 \dot{V}_1^*(\mathbf{r}_1, t-t_1) \dot{V}_2(\mathbf{r}_2, t-t_2) \end{aligned}$$

и выражение (4-5) будет иметь вид:

$$\begin{aligned} \langle I(\mathbf{r}, t) \rangle = |k_1|^2 \langle I_1(\mathbf{r}, t-t_1) \rangle + \\ + |k_2|^2 \langle I_2(\mathbf{r}_2, t-t_2) \rangle + 2\text{Re } [k_1^* k_2 \langle \dot{V}_1^*(\mathbf{r}_1, t-t_1) \times \\ \times \dot{V}_2(\mathbf{r}_2, t-t_2) \rangle], \end{aligned} \quad (4-6)$$

где $\langle I_1(\mathbf{r}_1, t-t_1) \rangle$ — средняя интенсивность поля на поверхности источника S_1 ; $I_2(\mathbf{r}_2, t-t_2)$ — средняя интенсивность поля на поверхности источника S_2 , или

$$\begin{aligned} \langle I(\mathbf{r}, t) \rangle = \langle I_1(\mathbf{r}, t) \rangle + \langle I_2(\mathbf{r}, t) \rangle + \\ + 2\text{Re } [k_1^* k_2 \langle \dot{V}_1^*(\mathbf{r}_1, t-t_1) \dot{V}_2(\mathbf{r}_2, t-t_2) \rangle], \end{aligned} \quad (4-7)$$

где $\langle I_1(\mathbf{r}, t) \rangle$ — средняя интенсивность поля в точке $P(\mathbf{r})$ при наличии только источника S_1 ; $\langle I_2(\mathbf{r}, t) \rangle$ — то же при наличии только источника S_2 .

Из выражения (4-7) следует, что усредненная интенсивность в точке сложения излучений источников не равна сумме усредненных интенсивностей, создаваемых в исследуемой точке каждым источником в отдельности.

Если излучение источников стабильно и непрерывно во времени (как, например, излучение газовых лазеров, работающих в непрерывном режиме), то можно изменить начало отсчета, полагая

$$\tau = t_2 - t_1 = \frac{d_2 - d_1}{v}.$$

При этом средняя интенсивность в точке $P(\mathbf{r})$ согласно (4-6) определяется выражением

$$\begin{aligned} \langle I(\mathbf{r}, t) \rangle &= |k_1|^2 \langle I_1(\mathbf{r}_1, t + \tau) \rangle + \\ &+ |k_2|^2 \langle I_2(\mathbf{r}_2, t) \rangle + 2\text{Re } k_1^* k_2 \langle \dot{V}_1^*(\mathbf{r}_1, t + \tau) \dot{V}_2(\mathbf{r}_2, t) \rangle \end{aligned} \quad (4-8)$$

или согласно (4-7) выражением

$$\begin{aligned} \langle I(\mathbf{r}, t) \rangle &= \langle I_1(\mathbf{r}, t) \rangle + \langle I_2(\mathbf{r}, t) \rangle + \\ &+ 2\text{Re } [k_1^* k_2 \Gamma_{12}(\tau)]. \end{aligned} \quad (4-9)$$

Функция

$$\begin{aligned} \Gamma_{12}(\tau) &= \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \langle \dot{V}_1(\mathbf{r}_1, t + \tau) \dot{V}_2^*(\mathbf{r}_2, t) \rangle = \\ &= \langle \dot{V}_1^*(\mathbf{r}_1, t + \tau) \dot{V}_2(\mathbf{r}_2, t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \dot{V}_1(\mathbf{r}_1, t + \tau) \times \\ &\times \dot{V}_2^*(\mathbf{r}_2, t) dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \dot{V}_1^*(\mathbf{r}_1, t + \tau) \dot{V}_2(\mathbf{r}_2, t) dt \end{aligned} \quad (4-10)$$

называется *функцией взаимной когерентности*. Она определяет корреляцию (связь) характеристик поля в двух точках $S_1(\mathbf{r}_1)$ и $S_2(\mathbf{r}_2)$ в моменты времени $t + \tau$ и t соответственно.

Нормируя $\Gamma_{12}(\tau)$, получаем *комплексную степень когерентности* или *коэффициент когерентности*:

$$\begin{aligned} \gamma_{12}(\tau) &= \gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma(\mathbf{r}_1, \mathbf{r}_1, 0)} \sqrt{\Gamma(\mathbf{r}_2, \mathbf{r}_2, 0)}} = \\ &= \frac{\langle \dot{V}_1(\mathbf{r}_1, t + \tau) \dot{V}_2^*(\mathbf{r}_2, t) \rangle}{\sqrt{\langle \dot{V}_1(\mathbf{r}_1, t) \dot{V}_1^*(\mathbf{r}_1, t) \rangle} \sqrt{\langle \dot{V}_2(\mathbf{r}_2, t) \dot{V}_2^*(\mathbf{r}_2, t) \rangle}} = \\ &= \frac{\Gamma_{12}(\tau)}{\sqrt{\langle I_1(\mathbf{r}_1, t) \rangle} \sqrt{\langle I_2(\mathbf{r}_2, t) \rangle}}, \end{aligned} \quad (4-11)$$

т. е.

$$\gamma_{12}(\tau) = \frac{k_1^* k_2 \Gamma_{12}(\tau)}{\sqrt{\langle I_1(\mathbf{r}, t) \rangle} \sqrt{\langle I_2(\mathbf{r}, t) \rangle}}, \quad (4-12)$$

где $\langle I_1(\mathbf{r}_1, t) \rangle$ и $\langle I_2(\mathbf{r}_2, t) \rangle$ — средние интенсивности на поверхности источников S_1 и S_2 соответственно; $\langle I_1(\mathbf{r}, t) \rangle$ и $\langle I_2(\mathbf{r}, t) \rangle$ — средние интенсивности в точке $P(\mathbf{r})$, создаваемые соответственно только источником S_1 или S_2 .

На основании (4-12) перепишем (4-9) в виде

$$\begin{aligned} \langle I(\mathbf{r}, t) \rangle &= \langle I_1(\mathbf{r}, t) \rangle + \langle I_2(\mathbf{r}, t) \rangle + \\ &+ 2 \sqrt{\langle I_1(\mathbf{r}, t) \rangle \langle I_2(\mathbf{r}, t) \rangle} \operatorname{Re} \gamma_{12}(\tau). \end{aligned} \quad (4-13)$$

Если складываемые поля являются квазимонохроматическими полями с одинаковой средней частотой ω_0 , то согласно (4-2) и (4-3) на поверхности источников S_1 и S_2 волновые функции определяются выражениями

$$\begin{aligned} \dot{V}_1(\mathbf{r}_1, t + \tau) &= \dot{V}_{m1}(\mathbf{r}_1, t + \tau) e^{i[\omega_0(t+\tau) + \varphi_1(\mathbf{r}_1, t+\tau)]}, \\ \dot{V}_2(\mathbf{r}_2, t) &= \dot{V}_{m2}(\mathbf{r}_2, t) e^{i[\omega_0 t + \varphi_2(\mathbf{r}_2, t)]}. \end{aligned}$$

Для квазимонохроматических полей с одинаковой средней частотой ω_0

$$\gamma_{12}(\tau) = |\gamma_{12}(\tau)| e^{i[\omega_0 \tau + \varphi(\mathbf{r}_1, \mathbf{r}_2, \tau)]},$$

и согласно (4-13) средняя интенсивность в точке $P(\mathbf{r})$ определяется выражением

$$\begin{aligned} \langle I(\mathbf{r}, t) \rangle &= \langle I_1(\mathbf{r}, t) \rangle + \langle I_2(\mathbf{r}, t) \rangle + \\ &+ 2 \sqrt{\langle I_1(\mathbf{r}, t) \rangle \langle I_2(\mathbf{r}, t) \rangle} |\gamma_{12}(\tau)| \cos[\omega_0 \tau + \\ &+ \varphi(\mathbf{r}_1, \mathbf{r}_2, \tau)]. \end{aligned} \quad (4-14)$$

Здесь средние интенсивности $\langle I_1(\mathbf{r}, t) \rangle$ и $\langle I_2(\mathbf{r}, t) \rangle$, модуль коэффициента когерентности $|\gamma_{12}(\tau)|$ и его аргумент $\varphi(\mathbf{r}_1, \mathbf{r}_2, \tau)$ медленно изменяются при изменении положения точки $P(\mathbf{r})$ в плоскости, параллельной плоскости источников, значение $\omega_0 \tau$ из-за большой величины ω_0 изменяется быстро. Поэтому распределение интенсивности в плоскости, параллельной плоскости источников, будет состоять из однородного фона $\langle I_1(\mathbf{r},$

$t) > +I_2(r, t) >$ и наложенного на него косинусоидального распределения с амплитудой

$$2\sqrt{\langle I_1(r, t) \rangle \langle I_2(r, t) \rangle} |\gamma_{12}(\tau)|$$

(интерференционные полосы). Резкость интерференционных полос зависит от модуля коэффициента взаимной когерентности, изменяющегося в пределах

$$0 \leq |\gamma(r_1, r_2, \tau)| \leq 1.$$

Интенсивности максимума и минимума интерференционных полос определяются выражением (4-14)

$$\langle I(r, t) \rangle_{\text{макс}} = \langle I_1(r, t) \rangle + \langle I_2(r, t) \rangle + 2\sqrt{\langle I_1(r, t) \rangle \langle I_2(r, t) \rangle} |\gamma_{12}(\tau)|; \quad (4-15)$$

$$\langle I(r, t) \rangle_{\text{мин}} = \langle I_1(r, t) \rangle + \langle I_2(r, t) \rangle - 2\sqrt{\langle I_1(r, t) \rangle \langle I_2(r, t) \rangle} |\gamma_{12}(\tau)|. \quad (4-16)$$

Положение максимумов интенсивности определяется из условия

$$\arg \gamma(r_1, r_2, \tau) = \omega_0 \tau + \varphi(r_1, r_2, \tau) = 2m\pi, \quad (4-17)$$

где $m=0, \pm 1, \pm 2, \pm 3 \dots$

Положение минимумов интенсивности определяется из условия

$$\arg \gamma(r_1, r_2, \tau) = \omega_0 \tau + \varphi(r_1, r_2, \tau) = (2m + 1)\pi, \quad (4-18)$$

где $m=0, \pm 1, \pm 2, \pm 3 \dots$

Из выражений (4-17) и (4-18) можно определить аргумент $\gamma_{12}(\tau)$, измеряя расстояние между двумя ближайшими максимумами или минимумами интенсивности.

Резкость интерференционной картины определяется величиной $\vartheta(r)$, называемой «видностью» или «контрастом»:

$$\vartheta(r) = \frac{\langle I(r, t) \rangle_{\text{макс}} - \langle I(r, t) \rangle_{\text{мин}}}{\langle I(r, t) \rangle_{\text{макс}} + \langle I(r, t) \rangle_{\text{мин}}}. \quad (4-19)$$

Подставляя в (4-19), (4-15) и (4-16), получаем:

$$\vartheta(r) = \frac{2\sqrt{\langle I_1(r, t) \rangle \langle I_2(r, t) \rangle} |\gamma_{12}(\tau)|}{\langle I_1(r, t) \rangle + \langle I_2(r, t) \rangle}. \quad (4-20)$$

Если интенсивность обоих источников в плоскости сложения одинакова:

$$\langle I_1(r, t) \rangle = \langle I_2(r, t) \rangle,$$

то

$$\vartheta(r) = |\gamma_{12}(\tau)|$$

и модуль коэффициента взаимной когерентности определяется видностью.

В идеальном случае монохроматических источников одной частоты

$$\dot{V}_1(r_1, t + \tau) = V_{m1} e^{i[\omega_0(t+\tau) + \varphi_1]},$$

$$\dot{V}_2(r_2, t) = V_{m2} e^{i[\omega_0 t + \varphi_2]},$$

подынтегральное выражение функции взаимной когерентности

$$\Gamma_{12}(\tau) \approx \frac{1}{T} \int_0^T \dot{V}_1(r_1, t + \tau) \dot{V}_2^*(r_2, t + \tau) dt$$

не зависит от времени, в результате чего

$$\Gamma_{12}(\tau) = V_{m1} V_{m2} e^{i[\omega_0 \tau + \varphi]},$$

где $\varphi = \varphi_1 - \varphi_2$ и согласно (4-11)

$$\gamma_{12}(\tau) = e^{i(\omega_0 \tau + \varphi)},$$

т. е. $|\gamma_{12}(\tau)| = 1$ и такие источники полностью когерентны.

Если источники монохроматичны, но частоты их разные ($\omega_{01} \neq \omega_{02}$), то аргумент подынтегрального выражения для $\Gamma_{12}(\tau)$ зависит от времени:

$$e^{i[(\omega_{01} - \omega_{02})t + \omega_{01}\tau + \varphi]} = \cos[(\omega_{01} - \omega_{02})t + \omega_{01}\tau + \varphi] + \\ + j \sin[(\omega_{01} - \omega_{02})t + \omega_{01}\tau + \varphi]$$

и оказываются справедливыми соотношения

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \cos[(\omega_{01} - \omega_{02})t + \omega_{01}\tau + \varphi] dt = 0;$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \sin[(\omega_{01} - \omega_{02})t + \omega_{01}\tau + \varphi] dt = 0.$$

При этом $\gamma_{12}(\tau) = 0$, следовательно, $|\gamma_{12}(\tau)| = 0$ и такие источники некогерентны.

Однако, если ω_{01} и ω_{02} не равны, но близки по величине ($\omega_{01} \approx \omega_{02}$), то подынтегральное выражение медлен-

но изменяется во времени и можно подобрать конечное время наблюдения T , при котором

$$\frac{1}{T} \int_0^T \dot{V}_1(\mathbf{r}_1, t + \tau) \dot{V}_2^*(\mathbf{r}_2, t) dt \neq 0$$

и

$$(\omega_{01} - \omega_{02}) T \ll 2\pi;$$

следовательно, такие источники частично когерентны.

Если источники квазимонохроматичны и имеют одну среднюю частоту ω_0 , но работают независимо друг от друга, то фазы их относительно друг друга изменяются случайным образом. В этом случае аргумент подынтегрального выражения $e^{i[(\omega_0\tau + \varphi(t))]}$, где $\varphi(t) = \varphi_1(t) - \varphi_2(t)$, изменяется случайным образом во времени и интегрирование в интервале от 0 до ∞ дает нуль, так как хотя эти изменения и медленные по сравнению с периодом колебаний, но они достаточно быстры по сравнению с временем наблюдения. При этом $\gamma_{12}(\tau) = 0$.

Однако существует такой промежуток времени, называемый *временем когерентности*, в течение которого изменение $\varphi(t)$ много меньше 2π и $\gamma_{12}(\tau) \neq 0$.

Если между источниками есть связь, то изменения $\varphi(t)$ становятся медленнее и время когерентности увеличивается. Если связь между источниками такова, что $\varphi(t) = \text{const}$, то время когерентности стремится к бесконечности, аргумент подынтегрального выражения для $\Gamma_{12}(\tau)$ не зависит от времени и $|\gamma_{12}(\tau)| = 1$. Колебания таких источников полностью когерентны.

Если $\vartheta(\mathbf{r}) = 0$, интерференционные полосы отсутствуют, $\gamma_{12}(\tau) = 0$, т. е. поля полностью некогерентны. Если поля одинаковы и $\vartheta(\mathbf{r}) = 1$, то резкость полос максимальна, $|\gamma_{12}(\tau)| = 1$ и поля полностью когерентны. При $0 \leq |\gamma_{12}(\tau)| \leq 1$ поля частично когерентны.

Когерентные свойства полностью определяются функцией корреляции комплексных амплитуд поля в двух различных точках пространства, определяемых радиус-векторами \mathbf{r}_1 и \mathbf{r}_2 , когда поле в одной точке рассматривается через время τ по сравнению с полем в другой точке

$$\Gamma_{12}(\tau) = \Gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \dot{V}_1(\mathbf{r}_1, t + \tau) \dot{V}_2^*(\mathbf{r}_2, t) dt.$$

Когерентное излучение характеризуется временной и пространственной когерентностью.

Временная когерентность характеризуется постоянством соотношений фаз колебания в одной точке, рассмотренных за одинаковые промежутки времени в любой момент. Временная когерентность характеризуется $\Gamma_{11}(\tau)$, которая равна функции $\Gamma_{12}(\tau)$ при $r_1 = r_2$.

Полностью когерентно во времени лишь монохроматическое поле. Если поле представляет сумму монохро-

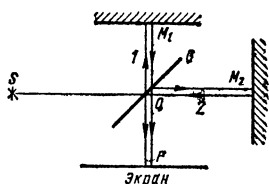


Рис. 4-2. Интерферометр Майкельсона.

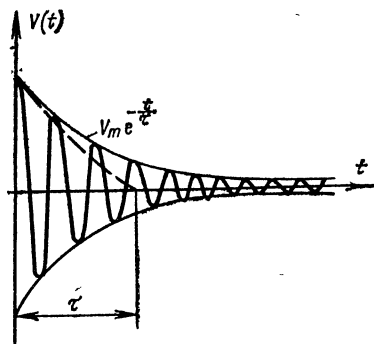


Рис. 4-3. Цуг волны.

матических полей одинаковых частот, то оно также будет полностью когерентным во времени. Если частоты отличаются, то поле не будет монохроматическим и полностью когерентным во времени. Квазимонохроматическое поле характеризуется спектром частот, и поэтому лишь частично когерентно. Время, в течение которого изменение разности фаз любых двух составляющих спектра много меньше 2π и сумма таких колебаний ведет себя как монохроматическое колебание со средней частотой ω_0 , называется временем когерентности $\tau_{\text{ког}}$.

Совмещая в схеме рис. 4-1 отверстия $S_1(r_1)$ и $S_2(r_2)$, но обеспечивая разность хода двух лучей от совмещенного отверстия до точки $P(r)$, переходим от интерференционной картины, характеризующей $\Gamma_{12}(\tau)$, к картине, характеризующей $\Gamma_{11}(\tau)$, т. е. определяющей временную когерентность.

Временная когерентность излучения может быть определена с помощью интерферометра Майкельсона (рис. 4-2). Излучение источника S с помощью полупрозрачной пластины G разделяется на две части 1 и 2, которые проходят различные расстояния OM_1OP и

OM_2OP , складываясь на экране $P(r)$. Монохроматическая волна, представляющая собой гармоническое колебание, существующее бесконечное время — идеализация. Практически волна представляет цуги волн (рис. 4-3) различной частоты ω , длящиеся конечное время τ и имеющие длину $l = v\tau$, где v — скорость распространения. В связи с этим излучение содержит не одну частоту, а спектр частот, ширина которого тем больше, чем короче цуги волн. Интерференционная картина, появляющаяся на экране P , зависит от разности хода лучей OM_1OP — $OM_2OP = L$. Пластиной G каждый цуг разделяется на два. Если разность хода L меньше длины цуга l , то два цуга, полученные от одного, перекроются на экране и получатся интерференционные полосы. Чем меньше разность хода по сравнению с длительностью цуга l , тем резче интерференционные полосы. Если разность хода будет больше длины цуга, то на экране будут встречаться цуги, происходящие от различных первоначальных цугов, между ними не будет корреляции и интерференционных полос не будет. Таким образом, очевидно, что время когерентности $\tau_{\text{ког}}$ связано с длительностью цуга τ .

Чем больше $\tau_{\text{ког}}$, тем на большем временном промежутке наблюдается корреляция колебаний, тем выше степень временной когерентности. В предельном случае полной когерентности $\tau_{\text{ког}} \rightarrow \infty$ (монохроматическое поле).

Отодвигая одно из отражающих зеркал, можно изменять разность хода L . Интерференционная картина наблюдается до тех пор, пока $L < v\tau$. Предельное значение L , найденное по выражению

$$L = v\tau_{\text{ког}},$$

и определяет время когерентности.

Пусть все излучаемые цуги волн одинаковы. Рассмотрим волновой цуг (рис. 4-3), представляющий собой экспоненциально затухающее колебание с частотой ω_0 :

$$\begin{aligned} \dot{V}(t) &= V_m e^{-t/\tau} e^{i\omega_0 t} \quad \text{при } t \geq 0; \\ \dot{V}(t) &= 0 \quad \text{при } t < 0. \end{aligned}$$

Определим длительность цуга τ как время, в течение которого амплитуда колебаний уменьшается в e раз.

Найдем спектр цуга как преобразование Фурье функции $\dot{V}(t)$:

$$\dot{V}(\omega) = \int_0^{\infty} \dot{V}(t) e^{-j\omega t} dt = V_m \int_0^{\infty} e^{-t/\tau} e^{j(\omega_0 - \omega)t} dt = \frac{V_m}{-j(\omega_0 - \omega) + \frac{1}{\tau}}$$

и квадрат его модуля:

$$|\dot{V}(\omega)|^2 = \frac{V_m^2}{(\omega_0 - \omega)^2 + \frac{1}{\tau^2}}.$$

График $|\dot{V}(\omega)|^2$ приведен на рис. 4-4. Ширина полосы спектра по половинному уровню определяется выражением

$$\Delta\omega \approx \frac{2}{\tau},$$

т. е. время когерентности связано со степенью монохроматичности колебания

$$\tau_{\text{кор}} \approx \frac{2}{\Delta\omega}. \quad (4-21)$$

Полученное соотношение определяет среднее время когерентности и ширину спектра, так как при выводе все цуги предполагались одинаковыми.

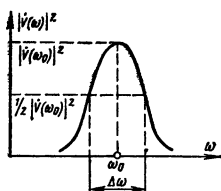


Рис. 4-4. К определению ширины полосы спектра цуга волны.

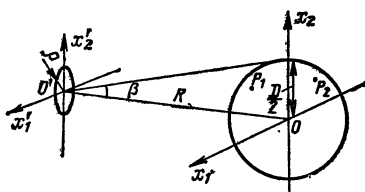


Рис. 4-5. Излучение пространственно когерентного поля от протяженного пространственно некогерентного источника.

Степень монохроматичности определяется отношением $\Delta\omega/\omega_0$. Чем уже спектр колебания, тем меньше отношение $\Delta\omega/\omega_0$, тем ближе колебание к монохроматическому, тем больше $\tau_{\text{кор}}$. Если уменьшить ширину спектра излучения, временная когерентность излучения улучшится.

Поле пространственно когерентно, если в любых двух фиксированных точках разность фаз, определяемая в любой момент времени, постоянна. Пространственная когерентность определяется функцией $\Gamma_{12}(\tau)$ при $\tau=0$, т. е. $\Gamma_{12}(0)$. Если расстояние от источников S_1 и S_2 до точки $P(\mathbf{r})$ на рис. 4-1 одинаково ($t_2 - t_1 = \tau = 0$), то интерференционная картина вблизи этой точки будет характеризовать пространственную когерентность.

Монохроматическое поле полностью пространственно когерентно. Квазимонохроматическое поле можно характеризовать областью когерентности, в пределах которой разность фаз колебаний в двух любых точках изменяется незначительно. Пространственная и временная когерентности взаимосвязаны. Эта связь определяется волновыми уравнениями для функции взаимной когерентности

$$\Delta_i \Gamma_{12}(\tau) = \frac{1}{v^2} \frac{\partial^2 \Gamma_{12}(\tau)}{\partial \tau^2} \quad i = 1, 2, \quad (4-22)$$

где Δ_i — оператор Лапласа относительно координат точек $P_1(\mathbf{r}_1)$ и $P_2(\mathbf{r}_2)$; v — скорость распространения. Каждое из уравнений (4-22) описывает изменение взаимной когерентности, когда одна из точек (P_1 или P_2) фиксирована, а другая точка и параметр τ изменяются.

Из волновых уравнений (4-22) следует, что распространение влияет на когерентность. В случае плоской и круглой поверхности радиуса ρ квазимонохроматического пространственно некогерентного источника (рис. 4-5) решение волнового уравнения для коэффициента когерентности определяется выражением

$$\gamma(\mathbf{r}_1, \mathbf{r}_2, \tau) = \frac{2J_1(u)}{u} e^{j(\omega_0 t - \psi)},$$

где $J_1(u)$ — функция Бесселя;

$$u = \frac{k_0 \rho}{R} \sqrt{(x_1^{(2)} - x_1^{(1)})^2 + (x_2^{(2)} - x_2^{(1)})^2};$$

$$\psi = k_0 \frac{(x_1^{(2)2} + x_2^{(2)2}) - (x_1^{(1)2} + x_2^{(1)2})}{2R};$$

$x_1^{(1)}, x_2^{(2)}$ и $x_1^{(2)}, x_2^{(1)}$ — декартовы координаты точек $P_1(\mathbf{r}_1)$ и $P_2(\mathbf{r}_2)$; $k_0 = 2\pi/\lambda_0$ — волновое число, соответствующее средней частоте полосы излучения; R — расстояние от источника до освещенной поверхности.

Практически время задержки τ прихода луча в одну точку относительно другой мало по сравнению с временем когерентности $\tau_{\text{ког}}$. При этом $|\gamma(r_1, r_2, \tau)|$ незначительно отличается от величины $|\gamma(r_1, r_2, 0)|$, характеризующей пространственную когерентность.

Зависимость $|\gamma(r_1, r_2, 0)|$ от расстояния между точками $P_1(r_1)$ и $P_2(r_2)$ для квазимонохроматического излучения, т. е. зависимость $\left| \frac{2J_1(u)}{u} \right|$ приведена на рис.

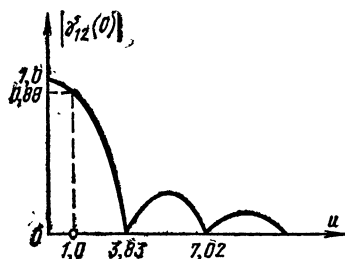


Рис. 4-6. Зависимость $|\gamma_{12}(0)|$ от u .

4-6. При значении $u=0$ получим максимальное значение $|\gamma(r_1, r_2, 0)| = 1$. При возрастании u величина $|\gamma(r_1, r_2, 0)|$ уменьшается и при $u=3,83$ становится равной нулю, что соответствует полностью некогерентному полю. Если считать допустимой для частично когерентного излучения значение

$$|\gamma(r_1, r_2, 0)| > 0,88,$$

то расстояние между точками $P_1(r_1)$ и $P_2(r_2)$ должно лежать в интервале, соответствующем изменению u в пределах

$$0 < u < 1$$

или

$$0 < |r_1 - r_2| < 0,16 \frac{R\lambda_0}{\rho}.$$

Таким образом, квазикогерентное поле от протяженного пространственно некогерентного источника можно получить в круге диаметром

$$D \leq 0,32 \frac{R\lambda_0}{\rho}, \quad (4-23)$$

лежащем в плоскости, параллельной плоскости протяженного источника.

Расходимость излучения определяется углом β , найдем по отношению

$$\frac{D}{2R} = \tan \beta \approx \beta \text{ при } R \gg D,$$

и для пространственно когерентного излучения угол расходимости луча должен согласно (4-23) удовлетворять условию

$$0 \leq \beta \leq 0,16 \frac{\lambda_0}{\rho}.$$

Таким образом, пространственная когерентность связана с направленностью излучения.

4-2. ОТКРЫТЫЕ РЕЗОНАТОРЫ

Основное назначение резонаторов оптических квантовых генераторов и регенеративных усилителей — создание положительной обратной связи, при которой индуцированное излучение многократно проходит через активную среду, каждый раз усиливаясь. Когерентность излучения лазеров также определяется резонатором.

Применение обычных объемных резонаторов, геометрические размеры которых соответствуют настройке на одну собственную частоту, в оптическом диапазоне нецелесообразно, так как технологически трудно создать объемный резонатор с размерами порядка длины волны (доли микрона). Кроме того, при таких малых размерах добротность резонатора резко уменьшается. Объемный же резонатор, размеры которого много больше длины волны, практически теряет свои резонансные свойства.

Действительно, согласно формуле Релея — Джинса число собственных колебаний ρ , возникающих в резонансной трехмерной системе и приходящихся на интервал частот $\Delta\omega$, равно:

$$\rho = \frac{V}{2\pi^2 v^3} \omega^2 \Delta\omega,$$

где V — объем системы; $v = \frac{1}{\sqrt{\epsilon_a \mu_a}}$ — скорость распространения электромагнитной волны. Эта формула асимптотическая, точность ее повышается с увеличением частоты ω . При повышении частоты спектр собственных колебаний сильно сгущается, потери в стенках резонатора возрастают пропорционально $\sqrt{\omega}$, резонансные кривые каждого типа колебания расширяются и перекрываются.

Для одномерной системы .

$$p = \frac{2l}{\pi v} \Delta\omega,$$

где l — длина системы. Согласно этой формуле сгущение спектра для одномерной области не происходит.

В оптическом диапазоне волн используются резонаторы, образуемые системой двух обращенных друг к другу отражающих поверхностей (зеркал). С точки зрения геометрической оптики между системой параллельных зеркал могут существовать пары параллель-

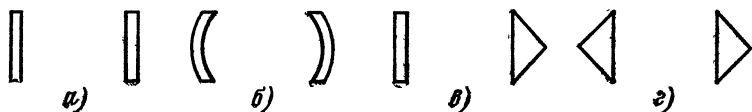


Рис. 4-7. Открытые резонаторы, образованные плоскими зеркалами (а); сферическими зеркалами (б); плоским зеркалом и призмой полного отражения (в); призмами полного отражения (г).

ных, преобразующихся при отражении друг в друга лучей. Эти лучи и определяют собственные колебания резонатора, который эквивалентен одномерной колебательной системе. Под l в этом случае подразумевается расстояние между зеркалами. Так как резонатор ограничен лишь двумя поверхностями и открыт с других сторон, то его называют открытым. Размеры такого резонатора много больше длины волны, а спектр достаточно разрежен. Открытый резонатор длиной l м имеет в 10^{10} раз меньше резонансных частот, чем объемный резонатор такой же длины.

Отражающие поверхности могут представлять собой зеркала (плоские, сферические, параболические), грани призм полного внутреннего отражения или границы сред с различными показателями преломления (рис. 4-7). При этом необходимо иметь поверхности с большим коэффициентом отражения и малыми потерями на поглощение. Такие поверхности создаются с помощью диэлектрических покрытий, которые обеспечивают коэффициент отражения более 0,99 на рабочей длине волны.

Плоский резонатор. Рассмотрим резонатор, образованный параллельными плоскими бесконечнопротяженными зеркалами без потерь (рис. 4-8). В таком резона-

торе возможно существование двух плоских однородных волн, распространяющихся от одной отражающей поверхности к другой вдоль оси резонатора навстречу друг другу. При этом образуются стоячие волны, называемые *продольными или аксиальными типами (модами) основного типа колебаний*, подчиняющиеся условию

$$l = q \frac{\lambda}{2}, \quad (4-24)$$

где q — число полуволн, укладываемых между зеркалами; λ — длина волны в среде, заполняющей резонатор; l — расстояние между зеркалами.

Стоячие волны могут образовываться и при сложении плоских волн, распространяющихся под некоторым углом θ к оси резонатора. Такие стоячие волны называются *продольными типами данного поперечного или углового типа колебаний*. Условие образования таких волн

$$l = q \frac{\lambda_{\perp}}{2} = q \frac{\lambda}{2 \cos \theta}, \quad (4-25)$$

где $\lambda_{\perp} = \lambda / \cos \theta$ определяет вариации поля в направлении оси резонатора.

Собственные частоты для продольных типов основного колебания определяются выражением

$$f_q = q \frac{v}{2l} = q \frac{c}{2nl}; \quad (4-26)$$

для продольных типов поперечных колебаний — выражением

$$f_q = q \frac{v}{2l \cos \theta} = q \frac{c}{2nl \cos \theta}, \quad (4-27)$$

где c — скорость света; n — коэффициент преломления среды, заполняющей резонатор; $v = c/n$.

Для поперечных типов колебаний имеет место выражение, т. е. одной и той же частоте соответствует множество поперечных типов, отличающихся значением q и углом распространения θ . Пусть плоские волны распространяются под углом θ в плоскости x_1x_3 (рис. 4-8). Фиксируя частоту основного типа колебания $f_q(\theta=0)$, найдем соответствующий этой частоте первый поперечный тип колебания. Очевидно, величина

на q для этого типа должна отличаться на единицу и он будет распространяться под фиксированным углом $\theta = \theta_1$, т. е.

$$f_q(\theta = 0) = f_{q-1}(\theta = \theta_1)$$

или согласно (4-27)

$$\cos \theta_1 = \frac{q-1}{q}.$$

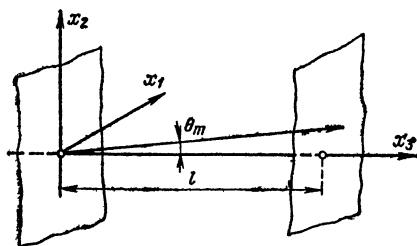


Рис. 4-8. Плоский резонатор.

Аналогично для m -го поперечного типа, распространяющегося под углом θ_m , имеем:

$$f_q(\theta = 0) = f_{q-m}(\theta = \theta_m)$$

и согласно (4-27)

$$\cos \theta_m = \frac{q-m}{q}.$$

При $m \ll q$, что соответствует поперечным колебаниям, имеющим место в резонаторах с зеркалами конечных размеров, получим:

$$\cos \theta_m \approx 1 - \frac{\theta_m^2}{2} \approx 1 - \frac{m}{q}; \quad (4-28)$$

отсюда

$$\theta_m = \sqrt{\frac{2m}{q}}$$

или согласно выражению (4-24)

$$\theta_m = \sqrt{m \frac{\lambda}{l}} \quad (m \ll q).$$

Аналогичное соотношение получаем, рассматривая распространение плоских волн под углом θ_n в плоскости x_2x_3 :

$$\theta_n = \sqrt{n \frac{\lambda}{l}} \quad (n \ll q).$$

При m и $n \ll q$ углы θ_m и θ_n принимают дискретные значения, угловое расстояние между соседними поперечными типами определяется выражениями

$$\Delta\theta_m = \sqrt{\frac{\lambda}{l}} (\sqrt{m+1} - \sqrt{m});$$

$$\Delta\theta_n = \sqrt{\frac{\lambda}{l}} (\sqrt{n+1} - \sqrt{n}),$$

и структура поля на зеркалах неоднородна, имеются периодические вариации поля.

Расстояние по частоте между соседними продольными типами основного и поперечных колебаний одинаково и согласно (4-26) и (4-27) с учетом (4-28) определяется выражением

$$\Delta f = f_q - f_{q-1} = \frac{v}{2l}.$$

Для малых m и n углы θ_m и θ_n малы и поле поперечных типов можно считать имеющим структуру поперечной электромагнитной волны (Т-волны), характеризующуюся определенным числом полуволн q , укладывающихся на длине резонатора, и числами m и n , определяющими направление распространения волн в резонаторе. Отдельный тип колебаний определяется как T_{mnq} .

Открытый резонатор с параллельными плоскими зеркалами конечных размеров можно рассматривать как волновод, образованный параллельными плоскостями, в котором распространяется волна при частоте, лишь немного большей критической. Такая волна, проходя к краям волновода не излучается, а с коэффициентом отражения, близким по модулю к единице, отражается обратно.

Поле пассивного резонатора определяется однородными уравнениями Гельмгольца

$$\Delta \dot{\mathbf{E}} + k^2 \dot{\mathbf{E}} = 0;$$

$$\Delta \dot{\mathbf{H}} + k^2 \dot{\mathbf{H}} = 0.$$

В декартовой системе координат скалярные волновые уравнения для составляющих поля имеют тот же вид:

$$\Delta \dot{u} + k^2 \dot{u} = 0, \quad (4-29)$$

где под функцией \dot{u} подразумевается любая из составляющих \dot{E}_1 , \dot{E}_2 , \dot{H}_1 или \dot{H}_2 . Составляющие поля находят при решении этого уравнения и выполнении граничных условий.

Однако точно решить волновое уравнение из-за дифракции не представляется возможным и от уравнения (4-29) переходят к параболическому уравнению, представляющему собой уравнение второго порядка, в котором отсутствует по крайней мере одна из вторых производных.

Решение волнового уравнения (4-29) приближенно представляют в виде

$$\dot{u} = u_m(x_1, x_2, x_3) e^{-jkx_3}, \quad (4-30)$$

при этом поле в резонаторе можно представить как суперпозицию двух волн вида (4-30), бегущих в противоположные стороны. В выражении (4-30) множитель e^{-jkx_3} быстро изменяется в направлении оси x_3 , а $u_m(x_1, x_2, x_3)$ медленно изменяется в поперечном направлении (поперечная диффузия поля) и еще медленнее в продольном. Функция $u_m(x_1, x_2, x_3)$ характеризует отличие поля в резонаторе от плоской волны: неоднородность поля в поперечном направлении, изменение сечения луча при распространении, кривизну фазового фронта. Подставляя (4-30) в (4-29) и опуская вследствие малости член $\frac{\partial^2 u_m}{\partial x_3^2}$, получаем параболическое уравнение

$$\frac{\partial^2 \dot{u}_m}{\partial x_1^2} + \frac{\partial^2 \dot{u}_m}{\partial x_2^2} - 2jk \frac{\partial \dot{u}_m}{\partial x_3} = 0. \quad (4-31)$$

Возможное решение для основного типа колебания имеет вид:

$$\dot{u}_m = e^{-j \left(p + \frac{k}{2q} r^2 \right)}, \quad (4-32)$$

где $r^2 = x_1^2 + x_2^2$ — расстояние от оси, $p = p(x_3)$ — комплексный параметр, характеризующий изменение ампли-

туды и фазы поля при распространении вдоль оси x_3 , $q = q(x_3)$ — комплексный параметр, характеризующий распределение интенсивности по координате r и кривизну волнового фронта, который вблизи оси является сферическим. Параметр q можно представить в виде

$$\frac{1}{q} = \frac{1}{R} - j \frac{\lambda}{\pi w^2}. \quad (4-33)$$

При подстановке (4-33) в (4-32) видно, что $R = R(x_3)$ — радиус кривизны волнового фронта в точке пересечения с осью x_3 , а $w = w(x_3)$ определяет умень-

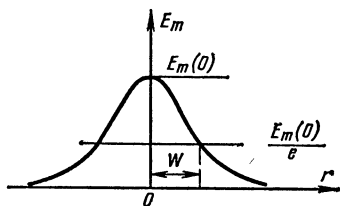


Рис. 4-9. Поперечное распределение амплитуды поля для основного типа колебания.

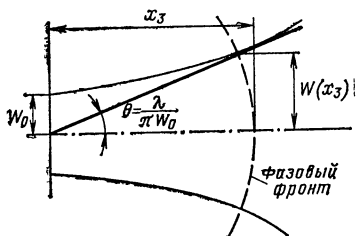


Рис. 4-10. Продольное распределение поля для основного типа колебания.

шение амплитуды поля E с увеличением расстояния от x_3 . Распределение поля в поперечной плоскости подчиняется *закону Гаусса* (рис. 4-9) и w определяется расстоянием, на котором амплитуда поля в e раз меньше, чем на оси. Параметр w обычно называют радиусом пучка, $2w$ — диаметром пучка.

Подставляя (4-32) в (4-31) и приравнивая члены с r в одинаковой степени, получаем:

$$\frac{dq}{dx_3} = 1; \quad (4-34)$$

$$\frac{dp}{dx_3} = - \frac{j}{\frac{dq}{dx_3}}. \quad (4-35)$$

Интегрируя уравнение (4-34), получаем соотношение

$$q_2 = q_1 + \Delta x_3, \quad (4-36)$$

связывающее комплексные параметры q , соответствующие любым поперечным сечениям 1 и 2 и отстоящие друг от друга на расстояние Δx_3 .

Распределение интенсивности в любой поперечной плоскости является гауссовым и изменяется вдоль оси x_3 . Гауссов пучок стягивается к минимальному диаметру $2\omega_0$ в сечении, где фазовый фронт плоский ($R=\infty$, «горловина»). Согласно (4-33) параметр q_0 , соответствующий этому сечению, равен:

$$q_0 = j \frac{\pi \omega_0^2}{\lambda}. \quad (4-37)$$

Если x_3 отсчитывать от этой плоскости, то на расстоянии x_3 согласно (4-36) и (4-37)

$$q(x_3) = q_0 + x_3 = j \frac{\pi \omega_0^2}{\lambda^2} + x_3. \quad (4-38)$$

Согласно (4-33) и (4-38)

$$\frac{1}{x_3 + j \frac{\pi \omega_0^2}{\lambda}} = \frac{1}{R} - j \frac{\lambda}{\pi \omega^2}$$

или

$$\frac{x_3 - j \frac{\pi \omega_0^2}{\lambda}}{x_3^2 + \left(\frac{\pi \omega_0^2}{\lambda} \right)^2} = \frac{1}{R} - j \frac{\lambda}{\pi \omega^2}.$$

Приравнивая действительные и мнимые части правой и левой частей этого уравнения, получаем:

$$R(x_3) = x_3 \left[1 + \left(\frac{\pi \omega_0^2}{\lambda x_3} \right)^2 \right]; \quad (4-39)$$

$$\omega^2(x_3) = \omega_0^2 \left[1 + \left(\frac{\lambda x_3}{\pi \omega_0^2} \right)^2 \right], \quad (4-40)$$

где x_3 отсчитывается от «горловины».

На рис. 4-10 показано расширение пучка, определяемое уравнением (4-40). Образующая пучка представляет собой гиперболу с асимптотами, наклоненными к оси под углом

$$\theta = \frac{\lambda}{\pi \omega_0};$$

этот угол равен дифракционному в дальней зоне для волны основного типа.

Параметр p , определяющий решение (4-32), найдем с помощью уравнения (4-35), подставив в него (4-38).

$$\frac{dp}{dx_3} = - \frac{j}{x_3 + j \left(\frac{\pi w_0^2}{\lambda} \right)}.$$

Интегрирование этого уравнения дает

$$jp(x_3) = \ln \left[1 - j \frac{\lambda x_3}{\pi w_0^2} \right] = \ln \sqrt{1 + \left(\frac{\lambda x_3}{\pi w_0^2} \right)^2} - j \operatorname{arctg} \left(\frac{\lambda x_3}{\pi w_0^2} \right).$$

Действительная часть p

$$\operatorname{Re} p = \varphi = \operatorname{arctg} \frac{\lambda x_3}{\pi w_0^2} \quad (4-41)$$

в соответствии с (4-33) определяет разность фаз между гауссовым пучком и идеальной плоской волной. Мнимая часть p с учетом (4-40)

$$\operatorname{Im} p = \ln \sqrt{1 + \left(\frac{\lambda x_3}{\pi w_0^2} \right)^2} = \ln \frac{w}{w_0}, \quad (4-42)$$

и так как

$$e^{-\ln \frac{w}{w_0}} = \frac{w_0}{w},$$

то (4-42) определяет амплитудный множитель w_0/w , характеризующий уменьшение амплитуды на оси из-за расширения пучка.

С учетом полученных соотношений (4-41), (4-42), (4-33) и выражения (4-32) решение (4-30) будет иметь вид:

$$\dot{u}_m = \frac{w_0}{w} e^{-r^2 \left(\frac{1}{w^2} + j \frac{k}{2R} \right)} e^{-j(kx_3 - \varphi)}; \quad (4-43)$$

это решение для типов колебаний с гауссовым распределением поля в поперечном сечении.

Другие возможные решения волнового уравнения (4-29) представляют поперечные типы колебаний, при конечных размерах зеркал характеризуемые несколькими вариациями поля в поперечном направлении.

В декартовой системе координат решение уравнения (4-31) для поперечных типов можно представить в виде

$$\dot{u}_m = X_1 \left(\frac{x_1}{w} \right) X_2 \left(\frac{x_2}{w} \right) e^{-j \left[p + \frac{k}{2q} (x_1^2 + x_2^2) \right]}, \quad (4-44)$$

где w — радиус гауссова пучка.

Подставляя это решение в уравнение (4-31), получаем дифференциальные уравнения для определения X_1 и X_2 :

$$\left. \begin{aligned} \frac{d^2 X_1}{dx_1^2} - 2x_1 \frac{dX_1}{dx_1} + 2mX_1 &= 0; \\ \frac{d^2 X_2}{dx_2^2} - 2x_2 \frac{dX_2}{dx_2} + 2nX_2 &= 0, \end{aligned} \right\} \quad (4-45)$$

где m и n — целые числа.

Решениями этих уравнений являются полиномы Эрмита $H_m(x_1)$ и $H_n(x_2)$ соответственно. Полиномы Эрмита низших порядков имеют вид:

$$\begin{aligned} H_0(x) &= 1, \quad H_1(x) = 2x, \quad H_2(x) = 4x^2 - 2; \\ H_3(x) &= 8x^3 - 12x, \quad H_4(x) = 16x^4 - 48x^2 + 12. \end{aligned}$$

Окончательно решение волнового уравнения (4-29) имеет вид:

$$\dot{u}_m = H_m \left(\frac{x_1}{w} \right) H_n \left(\frac{x_2}{w} \right) e^{-r^2 \left(\frac{1}{w^2} + j \frac{k}{2R} \right)} e^{-j(kx_s - \varphi)}, \quad (4-46)$$

где радиус кривизны фронта R одинаков для всех типов и определяется выражением (4-39), радиус пучка w определяется выражением (4-40), фазовый сдвиг φ зависит от индексов типа колебания и для колебания с индексами m, n равен:

$$\varphi_{mn} = \varphi(m, n, x_3) = (m + n + 1) \arctg \left(\frac{\lambda x_3}{\pi w_0^2} \right). \quad (4-47)$$

Итак, собственные типы колебаний открытого резонатора, часто называемые модами, характеризуются определенным распределением амплитуд и фаз по поверхности зеркала. Запаздывание фазы у краев зеркала по сравнению с центром соответствует долям длины волны. В то же время по любому поперечному размеру

прямоугольного зеркала или диаметру круглого зеркала укладывается множество длин волн. Отсюда следует, что кривизна фронта волны в резонаторе мала и колебания в резонаторе являются поперечными. Типы колебаний обычно обозначаются как T_{mnq} , где индексы m и n называются поперечными ($m, n=0, 1, 2, \dots$), а q —

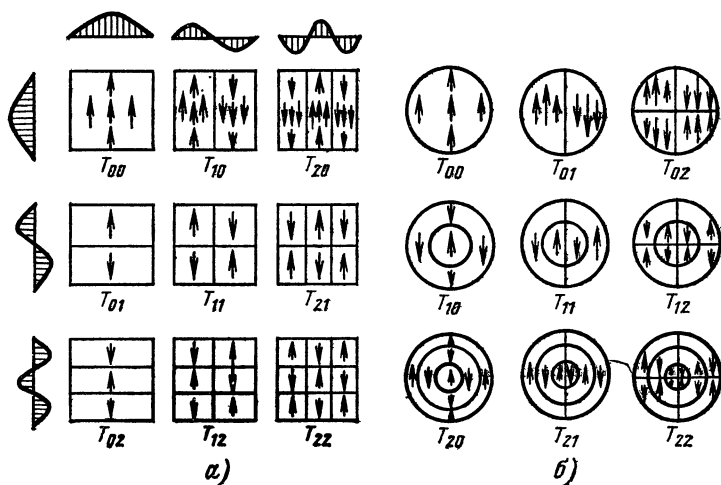


Рис. 4-11. Распределение поля на зеркалах открытого резонатора.

а — с плоскими прямоугольными зеркалами; *б* — с плоскими круглыми зеркалами.

продольным или аксиальным индексом. Поперечные индексы определяют число вариаций знака поля на поверхности прямоугольного зеркала в направлении координат x_1 и x_2 ; для круглого зеркала m означает число изменений знака вдоль радиуса, n — по углу (рис. 4-11). Продольный индекс q равен числу полуволн, укладывающихся на длине резонатора.

Условие резонанса определяется выражением

$$kl - \varphi_{mn} = q\pi, \quad (4-48)$$

где φ_{mn} согласно (4-47) зависит от индексов m и n .

Из равенства (4-48) следует, что резонансные частоты равны:

$$f_{mnq} = \frac{c}{2\pi l n} (\varphi_{mn} + q\pi), \quad (4-49)$$

где n — показатель преломления среды, заполняющей резонатор.

Разность резонансных частот соседних колебаний с одинаковыми поперечными индексами и отличающихся продольными индексами на единицу равна

$$\Delta f = f_q - f_{q-1} = \frac{v}{2l}. \quad (4-50)$$

Так как $q \gg m$ и $q \gg n$, разность частот, определяемая выражением (4-50), мала по сравнению с резонансной частотой, определяемой (4-49), то моды с одинаковыми m и n , но разными q объединяют под названием *поперечной моды*, опуская q и обозначая ее как T_{mn} . Колебания, имеющие одинаковую структуру поля (одинаковые индексы m и n), но отличающиеся индексом q , называют *продольными или аксиальными, относящимися к данной поперечной моде*. Согласно выражению (4-50) расстояние по частоте между соседними продольными колебаниями, относящимися к различным поперечным модам, одинаково. Таким образом, наборы частот, соответствующие различным поперечным модам, сдвинуты друг относительно друга по частоте. Этот сдвиг определяется фазовым сдвигом φ_{mn} , который согласно (4-47) различен для разных поперечных мод. Случай, когда частоты различных поперечных мод совпадают, называется вырождением.

При небольших значениях m , n энергия поля в основном сосредоточена в средней области зеркала и быстро падает до нуля при приближении к краям. Типы колебания с большими m и n образуются в результате сложения плоских волн, распространяющихся под большими углами к оси резонатора. Поле этих типов вблизи краев достаточно велико, поэтому дифракционные потери велики и такие типы практически не могут существовать.

Чем больше индексы m , n , тем больше поле у краев зеркала, тем больше дифракционные потери. Суммарные потери определяются как сумма потерь дифракционных и потерь на зеркалах. Последние одинаковы для всех типов колебаний. Наименьшими дифракционными потерями обладает основной тип T_{00} , для которого распределение амплитуды достигает максимума в центре и довольно круто падает к краям.

Распределение поля на поверхностях зеркал, дифракционные потери и фазовый сдвиг зависят от длины резонатора l , апертурного размера зеркала $2a$, длины волны λ и характеризуются числом Френеля:

$$N_{\Phi} = \frac{a^2}{l\lambda}. \quad (4-51)$$

Чем больше N_{Φ} , тем меньше амплитуда поля на краю зеркал, тем меньше дифракционные потери. Зависимости потерь мощности $\alpha_{\text{диф}}$ и фазового сдвига φ за один проход от числа Френеля N_{Φ} для круглых

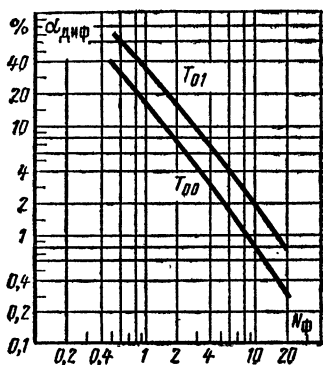


Рис. 4-12. Зависимость относительных потерь мощности $\alpha_{\text{диф}}$ за один проход от числа Френеля N_{Φ} для круглых плоских зеркал.

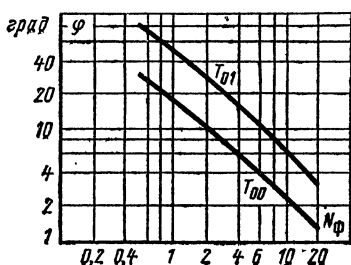


Рис. 4-13. Зависимость фазового сдвига φ за один проход от числа Френеля N_{Φ} для круглых плоских зеркал.

плоских зеркал приведены на рис. 4-12 и 4-13 соответственно. Если $N_{\Phi} > 10$, то дифракционные потери определяются выражением

$$\alpha_{\text{диф}} = 5,23 \cdot 10^{-2} A_{n(m+1)}^2 N_{\Phi}^{-3/2}, \quad (4-52)$$

где $A_{n(m+1)} - (m+1)$ корень уравнения $J_n(x) = 0$; $J_n(x)$ — функция Бесселя n -го порядка.

Рассмотренный метод параболического уравнения широко используется в теории дифракции для получения приближенных решений, когда длина волны много меньше размеров системы. Для исследования открытых резонаторов возможно применить также метод Кирхгофа, основанный на принципе Гюйгенса — Френеля.

Согласно принципу Гюйгенса — Френеля функция u , удовлетворяющая скалярному волновому уравнению (4-29), может быть представлена как суперпозиция

сферических волн, излучаемых элементарными источниками, распределенными на заданной поверхности S . Математическая формулировка этого принципа, данная Кирхгофом, имеет вид:

$$\dot{u} = \frac{1}{4\pi} \int_S \left[\frac{\partial u_S}{\partial n} \left(\frac{e^{-jkr}}{r} \right) - u_S \frac{\partial}{\partial n} \left(\frac{e^{-jkr}}{r} \right) \right] dS. \quad (4-53)$$

Согласно этой формуле можно определить функцию u в любой точке объема по известному значению u_S на граничной поверхности S .

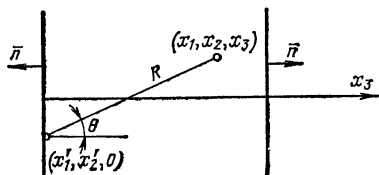


Рис. 4-14. К определению поля внутри резонатора методом Кирхгофа — Френеля.

Метод Кирхгофа является частным случаем более общего метода эквивалентных токов, учитывающего векторный характер электромагнитного поля. По заданным векторам \mathbf{E} и \mathbf{H} или их касательным составля-

ющим на граничной поверхности S можно определить электромагнитное поле в любой точке пространства. При этом поле в исследуемой точке рассматривают как результат векторного сложения полей элементарных источников, распределенных по всей поверхности S .

Поле внутри резонатора (рис. 4-14) образуется в результате отражения волны от зеркал. Пусть при M -м отражении от левого зеркала ($x_3=0$) волна характеризуется скалярной функцией $u_m(x_1, x_2, x_3)$. Под этой функцией можно подразумевать любую составляющую поля \dot{E}_1 , \dot{E}_2 , \dot{H}_1 или \dot{H}_2 . По мере распространения отраженный пучок расширяется. Однако вблизи зеркала можно считать, что

$$\dot{u}_m(x_1, x_2, x_3) \approx u_m(x_1, x_2, 0) e^{-jkx_3} = u_0(x_1, x_2) e^{-jkx_3},$$

т. е. отраженный пучок еще не расширяется. При этом выражение (4-53) принимает вид:

$$\dot{u}_m(x_1, x_2, x_3) = \frac{1}{4\pi} \int_S \left[u_m(x'_1, x'_2, 0) \frac{\partial}{\partial x_3} \left(\frac{e^{-jkR}}{R} \right) - \frac{\partial u_m(x'_1, x'_2, 0)}{\partial x'_3} \left(\frac{e^{-jkR}}{R} \right) \right] dS, \quad (4-54)$$

где $x'_1, x'_2, 0$ — координаты поверхности левого зеркала,

$$\frac{\partial}{\partial n} = -\frac{\partial}{\partial x_3}, \quad R = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + x_3^2}$$

— расстояние от точки на поверхности левого зеркала до исследуемой точки поля.

Так как

$$\begin{aligned} \frac{\partial}{\partial x'_3} \left(\frac{e^{-jkR}}{R} \right) &= \frac{\partial}{\partial R} \left(\frac{e^{-jkR}}{R} \right) \frac{dR}{dx'_3} \bigg|_{x'_3 \rightarrow 0} = \\ &= \left(jk + \frac{1}{R} \right) \frac{e^{-jkR}}{R} \cos \theta; \\ \frac{\partial u_m(x'_1, x'_2, 0)}{\partial x'_3} &= \frac{\partial u_m(x'_1, x'_2, x'_3)}{\partial x'_3} \bigg|_{x'_3 \rightarrow 0} = \\ &= \frac{\partial u_m(x'_1, x'_2, 0)}{\partial x_3} e^{-jkx_3} \bigg|_{x_3 \rightarrow 0} = -jk u_m(x'_1, x'_2, 0) \end{aligned}$$

и

$$k = \frac{2\pi}{\lambda} \gg \frac{1}{R},$$

то выражение (4-54) будет иметь вид:

$$\dot{u}_m(x_1, x_2, x_3) = \frac{jk}{4\pi} \int_S u_m(x'_1, x'_2, 0) \frac{e^{-jkR}}{R} (1 + \cos \theta) dS. \quad (4-55)$$

Поскольку рассматриваемая система симметрична, то в установившемся режиме ($M \gg 1$) распределение поля в плоскости $x_3 = l$, соответствующее $M+1$ отражению и характеризуемое функцией $u_{M+1}(x_1, x_2, l)$ будет с точностью до постоянного комплексного множителя совпадать с распределением поля в плоскости $x_3 = 0$, характеризуемым функцией $u_m(x_1, x_2, 0)$ и соответствующим M -му отражению, т. е.

$$\dot{u}_{M+1}(x_1, x_2, l) = \dot{K} \dot{u}_m(x_1, x_2, 0),$$

где \dot{K} — комплексная постоянная, не зависящая от координат и представляющая коэффициент передачи на один проход

$$\dot{K} = |K| e^{-j\psi}.$$

Модуль \dot{K} характеризует ослабление поля за один проход из-за дифракции и потерь в зеркале на отражение, аргумент $\psi = kl - \varphi$ определяет фазовый сдвиг волны за один проход, φ — добавка к геометрическому фазовому сдвигу.

На идеально проводящем зеркале

$$\dot{u}_{m+1}(x_1, x_2, x_3) = \dot{u}_m(x_1, x_2, x_3),$$

если под u_m подразумеваются составляющие магнитного поля H_1 или H_2 . При этом, используя (4-55), получаем:

$$\dot{K} \dot{u}(x_1, x_2) = \frac{1}{4\pi} \int_S \frac{jk}{R} (1 + \cos \theta) e^{-jkR} u(x'_1, x'_2) ds, \quad (4-56)$$

где

$$\dot{u}(x_1, x_2) = \dot{u}_m(x_1, x_2, 0).$$

Выражение (4-56) представляет собой однородное интегральное уравнение Фредгольма второго рода (см. П 2, пп. 13, 14) с неизвестной функцией $\dot{u}(x_1, x_2)$ и ядром

$$\frac{jk}{R} (1 + \cos \theta) e^{-jkR}.$$

Это уравнение решается методом последовательных приближений. Полученные в результате решения функции $\dot{u}(x_1, x_2)$ являются собственными функциями и определяют распределение поля на поверхности зеркал. Значения K , при которых уравнение имеет нетривиальные решения, являются собственными значениями. Каждому типу колебаний, возникающему в открытом резонаторе, соответствует определенная собственная функция $\dot{u}(x_1, x_2)$, определяющая структуру поля в резонаторе и определенное собственное значение K , определяющее потери и фазовый сдвиг на один проход.

Плоскопараллельный резонатор по сравнению с другими резонаторами обладает максимальной направленностью выходящего излучения, так как в этом случае фронт волны почти плоский.

Конфокальный симметричный резонатор образован одинаковыми сферическими отражателями, расположенными на расстоянии, равном радиусу кривизны зеркал. Так как фокусное расстояние сферических зеркал равно

половине их радиуса, то фокусы обоих зеркал совпадают (рис. 4-15).

Для исследования этих резонаторов также применяют метод Кирхгофа. Нормальные типы колебаний те же, что и для плоского резонатора. Однако поверхности зеркал в этом случае являются поверхностями равных фаз, распределение же амплитуд по поверхности зеркал неоднородно. При этом поле более сконцентрировано в центре зеркал и меньше у края, чем в случае плоских

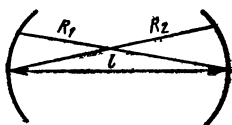
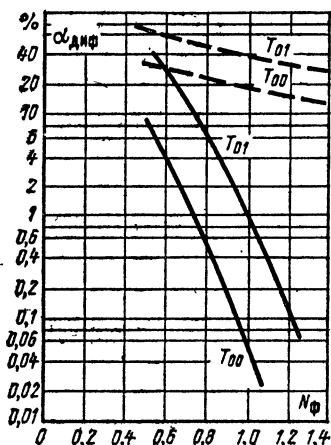


Рис. 4-15. Конфокальный симметричный резонатор ($R_1 = R_2 = l$).

Рис. 4-16. Сравнение потерь мощности за один проход для конфокального (сплошная линия) и плоского (пунктир) резонаторов.



зеркал. Поэтому дифракционные потери конфокального резонатора значительно меньше, чем плоского (рис. 4-16). Отношение потерь низших поперечных типов к потерям основного типа значительно больше, чем в случае плоского резонатора.

Собственные частоты конфокального резонатора с квадратными зеркалами определяются выражением

$$f_{mnq} = \frac{v}{4l} (2q + 1 + m + n), \quad (4-57)$$

где

$v = \frac{1}{\sqrt{\epsilon_a \mu_0}}$ — скорость распространения электромагнитных волн в среде, заполняющей резонатор.

Спектр возможных частот сильно вырожден. Увеличивая $m+n$ на любое целое четное число и уменьшая

q на половину этого числа, получаем то же значение частоты. Расстояние по частоте между соседними типами колебаний, для которых $(m+n)$ различается на единицу, равно:

$$\Delta f = f_{mnq} - f_{(m-1)nq} = \frac{v}{4l}.$$

Расстояние по частоте между продольными типами T_{00} как и в плоском резонаторе

$$\Delta f = f_{00q} - f_{00(q-1)} = \frac{v}{2l}.$$

Таким образом, расстояние по частоте между любыми соседними типами колебаний равно $v/4l$, т. е. в 2 раза меньше, чем в плоском резонаторе.

Для конфокального резонатора с круглыми зеркалами

$$f_{mnq} = \frac{v}{4l} (2q + 1 + 2m + n), \quad (4-58)$$

где m — радиальный индекс.

Здесь также наблюдается вырождение. При увеличении $2m+n$ на целое четное число и уменьшении q на половину этого числа частота не изменяется.

Расстояние по частоте между соседними типами, отличающимися по m на единицу, согласно (4-58), равно:

$$\Delta f = f_{mnq} - f_{(m-1)nq} = \frac{v}{2l}. \quad (4-59)$$

Расстояние по частоте между соседними типами, отличающимися по n на единицу,

$$\Delta f = f_{mnq} - f_{m(n-1)q} = \frac{v}{4l}. \quad (4-60)$$

Расстояние по частоте между продольными типами колебаний T_{00}

$$\Delta f = f_{00q} - f_{00(q-1)} = \frac{v}{2l}. \quad (4-61)$$

Таким образом, расстояние по частоте между любыми соседними типами колебаний также равно $v/4l$.

Для основной волны радиус круга на поверхности зеркала, соответствующий уменьшению напряженности

поля по сравнению с центром в e раз, определяется выражением

$$r_{1/e} = \sqrt{\frac{l\lambda}{\pi}}$$

и не зависит от размеров зеркала. Увеличение размеров отражателя приводит лишь к уменьшению дифракционных потерь.

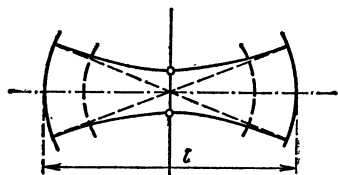


Рис. 4-17. Синфазные поверхности и поверхности равной интенсивности конфокального резонатора.

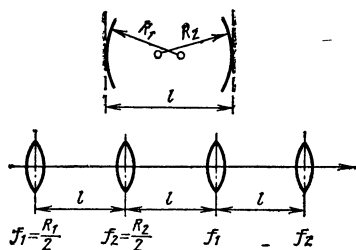


Рис. 4-18. Линзовая эквивалентная схема открытого резонатора.

Диаметр круга на зеркале, в пределах которого мощность падает вдвое, определяется выражением

$$d_{0,5} = 1,18 \sqrt{\frac{l\lambda}{\pi}}.$$

Поле внутри резонатора представляет собой суперпозицию двух волновых пучков, распространяющихся навстречу друг другу. Поверхность постоянной фазы представляет собой сферическую поверхность, радиус кривизны которой изменяется от бесконечности в центре резонатора (фокальная плоскость) до радиуса зеркал на поверхности отражателей. Поверхности равной относительной интенсивности поля имеют вид гиперболоидов вращения, ось которых совпадает с осью резонатора (рис. 4-17). Поперечное сечение пучка для основной моды максимально на поверхности зеркала и минимально в фокальной плоскости, где фазовый фронт плоский. В фокальной плоскости радиус поля, соответствующий ослаблению в e раз, определяется выражением

$$r_{1/e} = \sqrt{\frac{\lambda l}{2\pi}}.$$

Если зеркало частично прозрачно, то энергия излучается в виде луча с максимумом вдоль оси резонатора. Зная распределение поля на поверхности зеркала, можно определить диаграмму направленности. Для поля T_{00} ширина диаграммы по половинному уровню равна:

$$\theta_{0,5} = 0,53 \frac{\lambda}{r_{1/e}} = 0,63 \frac{\lambda}{d_{0,5}}.$$

Вне резонатора кривизна фронта быстро уменьшается и фронт волны стремится к плоскому.

Сферический конфокальный резонатор обладает преимуществами по сравнению с плоскопараллельным. Он обладает меньшими дифракционными потерями и большими различиями этих потерь для низших типов колебаний. Поле конфокального резонатора гораздо больше сконцентрировано у оси. Этот резонатор легче настроить, чем плоскопараллельный. Однако при нарушении конфокальности дифракционные потери резко возрастают.

Матричный метод. В резонаторе со сферическими зеркалами луч периодически фокусируется, отражаясь от этих зеркал. Подобный эффект имеет место в периодической последовательности линз. Поэтому резонатор со сферическими зеркалами радиуса R_1 и R_2 , расположенных на расстоянии l , можно представить как эквивалентную периодическую последовательность линз с фокусными расстояниями $R_1/2$ и $R_2/2$, расположенных на расстоянии l (рис. 4-18).

Оптическую систему такого типа можно рассматривать как четырехполюсник и прохождение параллельного пучка через систему может быть описано матрицей передачи

$$\begin{vmatrix} r_2 \\ r_2' \end{vmatrix} = \begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} \begin{vmatrix} r_1 \\ r_1' \end{vmatrix}, \quad (4-62)$$

где r_1 , r_1' и r_2 , r_2' — расстояние луча от оси системы и тангенс угла наклона луча (предполагается малым) на входе и выходе четырехполюсника соответственно (рис. 4-19)

$$\left. \begin{aligned} r_2 &= A_{11} r_1 + A_{12} r_1'; \\ r_2' &= A_{21} r_1 + A_{22} r_1'. \end{aligned} \right\} \quad (4-63)$$

Значения коэффициентов A_{ij} определяются следующими соотношениями

$$A_{11} = \frac{r_2}{r_1} \text{ при } r_1' = 0;$$

$$A_{21} = \frac{r_2'}{r_1} \text{ при } r_1' = 0;$$

$$A_{12} = \frac{r_2}{r_1'} \text{ при } r_1 = 0;$$

$$A_{22} = \frac{r_2'}{r_1'} \text{ при } r_1 = 0.$$

Здесь A_{11} соответствует коэффициенту передачи по расстоянию луча от оси; A_{22} — коэффициенту передачи по углу. Если система симметрична, то

$$A_{11} = A_{22},$$

если взаимна, то

$$A_{11}A_{22} - A_{12}A_{21} = 1.$$

Оптическая толстая линза характеризуется фокусами, плоскостями входа и выхода и главными плоскостями (рис. 4-20). Фокусы — это точки пересечения лучей,

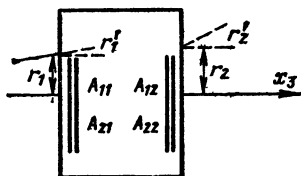


Рис. 4-19. Оптическая система как четырехполюсник.

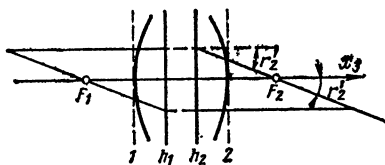


Рис. 4-20. Расположение главных плоскостей, плоскостей входа и выхода для толстой линзы.

падающих на систему параллельно оси в прямом и обратном направлении. Главные плоскости — плоскости, отображающие друг друга в натуральную величину. Главные точки — точки пересечения с осью главных плоскостей. Свет, входящий параллельным пучком в систему с одной стороны, собирается в фокус на расстоянии F_2 от второй главной плоскости. Свет, входящий параллельным пучком с другой стороны, собирается в

фокус, отстоящий от первой главной плоскости на расстоянии F_1 .

Если h_1 и h_2 — расстояния главных плоскостей от плоскостей входа и выхода (рис. 4-20), то уравнения (4-63) можно представить в виде

$$\begin{aligned} r_2 &= A_{11}(r_2 - r'_1 h_1) + A_{12} r'_1; \\ 0 &= A_{21}(r_2 - r'_1 h_1) + A_{22} r'_1 \end{aligned}$$

или

$$\begin{aligned} r_2(1 - A_{11}) &= r'_1(A_{12} - A_{11}h_1); \\ A_{21}r_2 &= (A_{21}h_1 - A_{22})r'_1; \end{aligned}$$

отсюда

$$\frac{1 - A_{11}}{A_{21}} = \frac{A_{12} - A_{11}h_1}{A_{21}h_1 - A_{22}}$$

и

$$h_1 = \frac{A_{12}A_{21} + (1 - A_{11})A_{22}}{A_{21}(1 - A_{11}) + A_{11}A_{21}} = \frac{A_{22} - 1}{A_{21}}.$$

Аналогично,

$$h_2 = \frac{A_{11} - 1}{A_{21}}.$$

При прямом направлении распространения параллельного луча согласно рис. 4-21

$$r'_2 = -\frac{r_1}{F_2}.$$

Сравнивая с (4-63), имеем:

$$F_2 = -\frac{1}{A_{21}}.$$

При обратном направлении распространения параллельного луча согласно рис. 4-21

$$r_2 = F_1 r'_1.$$

Сравнивая с (4-63), имеем:

$$F_1 = A_{12}.$$

Матрицы простейших оптических систем можно найти непосредственным составлением уравнений для этих систем, рассматривая их как четырехполюсники.

Для определения матрицы передачи в свободном пространстве на расстояние l (рис. 4-22) составим уравнения

$$r_2 = r_1 + lr'_1;$$

$$r'_2 = r'_1;$$

отсюда матрица передачи этого пространства на расстояние l согласно (4-63) равна:

$$\|A\| = \begin{vmatrix} 1 & l \\ 0 & 1 \end{vmatrix}. \quad (4-64)$$

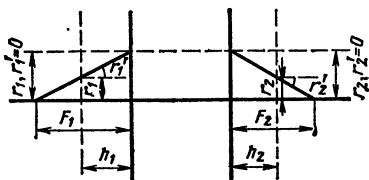


Рис. 4-21. К определению главных плоскостей и фокусных расстояний.

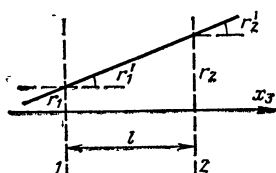


Рис. 4-22. К определению матрицы передачи свободного пространства длиной l .

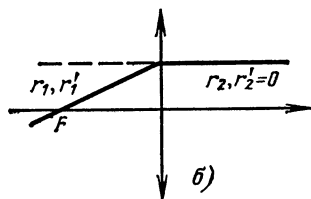
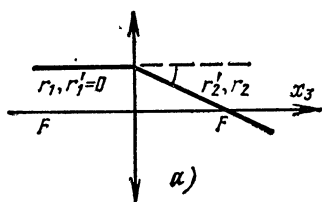


Рис. 4-23. К определению матрицы передачи тонкой симметричной линзы.

Матрица тонкой ($h_1 = h_2 = 0$) симметричной линзы ($F_1 = F_2 = F$) определяется с помощью рис. 4-23 и уравнений (4-63). Если луч на входе параллелен оси линзы ($r'_1 = 0$), то согласно рис. 4-23, а) имеем:

$$r_2 = A_{11}r_1 + A_{12}0 = r_1 \quad (r'_1 = 0);$$

$$r'_2 = A_{21}r_1 + A_{22}0 = -\frac{r_1}{F} \quad (r'_1 = 0);$$

отсюда $A_{11}=1$, $A_{21}=-\frac{1}{F}$, A_{12} и A_{22} не определяются. A_{12} и A_{22} можно определить с помощью рис. 4-23, б, когда луч на входе проходит через фокус ($r_2'=0$). При этом имеем:

$$r_2 = r_1 + A_{12}r_1' = r_1 \left(r_1' = \frac{r_1}{F} \neq 0 \right);$$

$$r_2' = -\frac{1}{F}r_1 + A_{22}r_1' = -\frac{1}{F}r_1 + A_{22}\frac{r_1}{F} = 0;$$

отсюда $A_{12}=0$, $A_{22}=1$ и окончательно матрица передачи тонкой симметричной линзы имеет вид:

$$\|A\| = \begin{vmatrix} 1 & 0 \\ -\frac{1}{F} & 1 \end{vmatrix}. \quad (4-65)$$

Прохождение лучом пути длиной l в свободном пространстве и тонкой симметричной линзы с фокусным расстоянием F определяется матрицей

$$\|A\| = \begin{vmatrix} 1 & l \\ -\frac{1}{F} & 1 - \frac{l}{F} \end{vmatrix}, \quad (4-66)$$

получаемой путем перемножения матриц (4-65) и (4-64).

Прохождение луча через тонкую симметричную линзу с фокусным расстоянием F и пути длиной l в свободном пространстве определяется матрицей

$$\|A\| = \begin{vmatrix} 1 - \frac{l}{F} & l \\ -\frac{1}{F} & 1 \end{vmatrix}, \quad (4-67)$$

получаемой путем перемножения матриц (4-64) и (4-65).

Матрица системы, приведенной на рис. 4-24,

$$\|A\| = \begin{vmatrix} 1 - \frac{l}{F_1} & 2l + \frac{l^2}{F_1} \\ -\frac{1}{F_1} - \frac{1}{F_2} + \frac{l}{F_1 F_2} & 1 - \frac{l}{F_1} - 2\frac{l}{F_2} + \frac{l^2}{F_1 F_2} \end{vmatrix} \quad (4-68)$$

получается путем последовательного перемножения соответствующих матриц.

Распространение луча конечного сечения, напряженность поля в котором убывает от центра к краям, в нелинейной среде приводит к изменению коэффициента преломления среды по закону

$$n = n_0 - \frac{1}{2} n_2 r^2, \quad n_2 \ll n_0,$$

где r — расстояние от оси. При этом происходит фокусировка энергии (см. § 2-4), которая имеет место и в

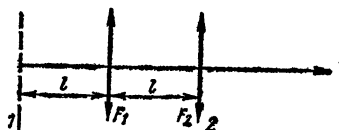


Рис. 4-24. Система свободное пространство — линза — свободное пространство — линза.

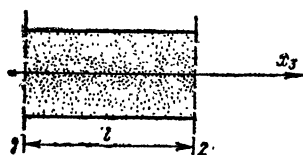


Рис. 4-25. Среда с коэффициентом преломления $n =$

$$= n_0 - \frac{1}{2} n_2 r^2.$$

кристаллах лазеров на твердых диэлектриках. Путь луча в такой среде определяется дифференциальным уравнением

$$\frac{d^2 r}{dx_3^2} = - \frac{n_2}{n_0} r,$$

решение которого имеет вид:

$$r = A e^{j \sqrt{\frac{n_2}{n_0}} x_3} + B e^{-j \sqrt{\frac{n_2}{n_0}} x_3}. \quad (4-69)$$

Рассматривая среду с переменным коэффициентом преломления как четырехполюсник (рис. 4-25), на входе при $x_3 = 0$ согласно (4-69) имеем:

$$r_1 = A + B;$$

$$r'_1 = j \sqrt{\frac{n_2}{n_0}} A - j \sqrt{\frac{n_2}{n_0}} B = j \sqrt{\frac{n_2}{n_0}} (A - B);$$

отсюда

$$\left. \begin{aligned} A &= \frac{1}{2} \left(r_1 + \frac{r'_1}{i \sqrt{\frac{n_2}{n_0}}} \right); \\ B &= \frac{1}{2} \left(r_1 - \frac{r'_1}{i \sqrt{\frac{n_2}{n_0}}} \right); \end{aligned} \right\} \quad (4-70)$$

На выходе четырехполюсника (рис. 4-25) согласно (4-69) с учетом (4-70) получим:

$$\begin{aligned} r_2 &= \frac{1}{2} \left(r_1 + \frac{r'_1}{i \sqrt{\frac{n_2}{n_0}}} \right) e^{i \sqrt{\frac{n_2}{n_0}} l} + \\ &+ \frac{1}{2} \left(r_1 - \frac{r'_1}{i \sqrt{\frac{n_2}{n_0}}} \right) e^{-i \sqrt{\frac{n_2}{n_0}} l}; \\ r'_2 &= \frac{i}{2} \sqrt{\frac{n_2}{n_0}} \left(r_1 + \frac{r'_1}{i \sqrt{\frac{n_2}{n_0}}} \right) e^{i \sqrt{\frac{n_2}{n_0}} l} - \\ &- \frac{i}{2} \sqrt{\frac{n_2}{n_0}} \left(r_1 - \frac{r'_1}{i \sqrt{\frac{n_2}{n_0}}} \right) e^{-i \sqrt{\frac{n_2}{n_0}} l}, \end{aligned}$$

или, преобразуя к виду (4-63),

$$\begin{aligned} r_2 &= \left[\cos \left(\sqrt{\frac{n_2}{n_0}} l \right) \right] r_1 + \left[\sqrt{\frac{n_0}{n_2}} \sin \left(\sqrt{\frac{n_2}{n_0}} l \right) \right] r'_1; \\ r'_2 &= - \left[\sqrt{\frac{n_2}{n_0}} \sin \left(\sqrt{\frac{n_2}{n_0}} l \right) \right] r_1 + \left[\cos \left(\sqrt{\frac{n_2}{n_0}} l \right) \right] r'_1; \end{aligned}$$

отсюда матрица передачи четырехполюсника имеет следующий вид:

$$\|A\| = \begin{vmatrix} \cos \left(\sqrt{\frac{n_0}{n_2}} l \right) & \sqrt{\frac{n_0}{n_2}} \sin \left(\sqrt{\frac{n_2}{n_0}} l \right) \\ - \sqrt{\frac{n_2}{n_0}} \sin \left(\sqrt{\frac{n_2}{n_0}} l \right) & \cos \left(\sqrt{\frac{n_2}{n_0}} l \right) \end{vmatrix}. \quad (4-71)$$

Матрица передачи на расстояние l через диэлектрик с коэффициентом преломления n и коэффициентом преломления внешней среды, равным 1 (рис. 4-26), определяется уравнениями

$$r_2 = r_1 + \frac{l}{n} r_1;$$

$$r'_2 = r'_1;$$

отсюда матрица передачи

$$\|A\| = \begin{vmatrix} 1 & \frac{l}{n} \\ 0 & 1 \end{vmatrix}. \quad (4-72)$$

Передача через двойной слой диэлектрика (рис. 4-27) определяется матрицей

$$\|A\| = \begin{vmatrix} 1 & \frac{l_1}{n_1} + \frac{l_2}{n_2} \\ 0 & 1 \end{vmatrix}. \quad (4-73)$$

Матрицы основных оптических четырехполюсников сведены в таблицу 4-1.

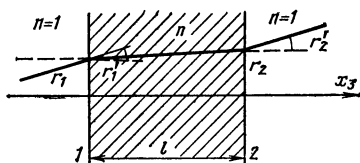


Рис. 4-26. Однослойный диэлектрик.

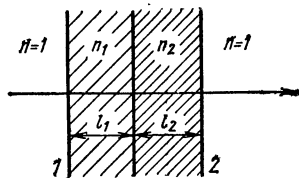


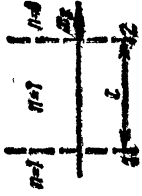
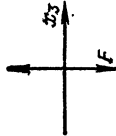
Рис. 4-27. Двуслойный диэлектрик.

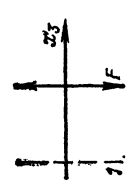
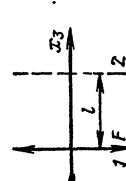
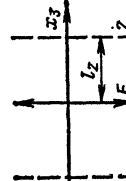
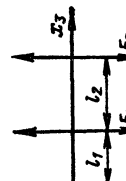
Устойчивые и неустойчивые резонаторы. Открытый резонатор со сферическими зеркалами радиуса R_1 и R_2 , расположенными на расстоянии l друг от друга, как уже упоминалось выше, можно представить в виде эквивалентной оптической схемы (рис. 4-18), представляющей периодическую последовательность линз, учитывающую многократное отражение в резонаторе.

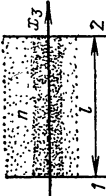
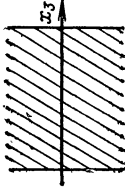
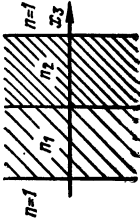
Матрица передачи периодической последовательности, повторяющийся элемент которой характеризуется матрицей передачи

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix},$$

Таблица 4-1

Порядковый номер	Схема четырехполюсника	Наименование четырехполюсника	$\ A\ $
1		Свободное пространство	$\begin{bmatrix} 1 & l \\ 0 & 1 \end{bmatrix}$
2		Тонкая симметричная линза с фокусным расстоянием F .	$\begin{bmatrix} 1 & 0 \\ -\frac{1}{F} & 1 \end{bmatrix}$

3		Система свободное пространство — линза	$\begin{vmatrix} 1 & l \\ -\frac{1}{F} & 1-\frac{l}{F} \end{vmatrix}$
4		Система линза — свободное пространство	$\begin{vmatrix} 1-\frac{l}{F} & l \\ -\frac{1}{F} & 1 \end{vmatrix}$
5		Система свободное пространство — линза — свободное пространство	$\begin{vmatrix} 1-\frac{l_2}{F} & l_1+l_2-\frac{l_1 l_2}{F} \\ -\frac{1}{F} & 1-\frac{l_1}{F} \end{vmatrix}$
6		Система свободное пространство — линза — свободное пространство — линза	$\begin{vmatrix} 1-\frac{l_2}{F_2} & l_1+l_2-\frac{l_1 l_2}{F_1} \\ -\frac{1}{F_1}-\frac{1}{F_2}+\frac{l_2}{F_1 F_2} & 1-\frac{l_1}{F_1}-\frac{l_2}{F_2}+\frac{l_1 l_2}{F_1 F_2} \end{vmatrix}$

Порядковый номер	Схема четырехполюсника	Наименование четырехполюсника	$\ A \ $
7		Линзоподобная среда $n = n_0 - \frac{1}{2} n_2 r^2$ $n_2 \ll n_0$	$\begin{vmatrix} \cos \sqrt{\frac{n_2}{n_0}} l & \sqrt{\frac{n_2}{n_0}} l \\ -\sqrt{\frac{n_2}{n_0}} \sin \sqrt{\frac{n_2}{n_0}} l & \cos \sqrt{\frac{n_2}{n_0}} l \end{vmatrix}$
8		Диэлектрическая пластина	$\begin{vmatrix} 1 & l/n \\ 0 & 1 \end{vmatrix}$
9		Двуслойная диэлектрическая пластина	$\begin{vmatrix} 1 & \frac{l_1}{n_1} + \frac{l_2}{n_2} \\ 0 & 1 \end{vmatrix}$

определяется путем возведения этой матрицы в степень n (n — число периодов системы). При этом согласно формуле возведения матрицы в степень, при условии, что $A_{11}A_{12} - A_{12}A_{21} = 1$ и $A_{11} + A_{22} \neq 2$:

$$\left\| \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \right\|^n = \frac{1}{S_1} \left\| \begin{pmatrix} S_{n+1} - A_{22}S_n & A_{12}S_n \\ A_{21}S_n & S_{n+1} - A_{11}S_n \end{pmatrix} \right\|, \quad (4-74)$$

где $S_n = \text{sh } n\alpha$, а аргумент α определяется соотношением

$$\text{ch } \alpha = \frac{A_{11} + A_{22}}{2}.$$

При этом возможны два случая. Если

$$-1 \leq \frac{A_{11} + A_{22}}{2} \leq 1, \quad (4-75)$$

то гиперболический косинус переходит в тригонометрический и анализ соотношения (4-74) показывает, что в этом случае при прохождении луча через систему он периодически фокусируется. Такая система называется *устойчивой*. Если условие (4-75) не выполняется, то луч по мере прохождения системы все более расширяется. Такая система называется *неустойчивой*.

В рассматриваемом случае (рис. 4-18) элемент периодической последовательности имеет вид, приведенный на рис. 4-24 ($F_1 = R_1/2$, $F_2 = R_2/2$) и характеризуется матрицей передачи (4-68). При этом

$$\frac{A_{11} + A_{22}}{2} = 1 - \frac{2l}{R_1} - \frac{2l}{R_2} - \frac{2l^2}{R_1 R_2} = 2 \left(1 - \frac{l}{R_1}\right) \left(1 - \frac{l}{R_2}\right) - 1.$$

Согласно (4-75) система будет устойчива, если

$$0 \leq \left(1 - \frac{l}{R_1}\right) \left(1 - \frac{l}{R_2}\right) \leq 1. \quad (4-76)$$

Рассмотрим различные виды открытых резонаторов и исследуем их на устойчивость (рис. 4-28). Если оба зеркала резонатора плоские ($R_1 = R_2 = \infty$ — рис. 4-28, а), то согласно (4-76) система устойчива в точке.

Если одно зеркало плоское ($R_1 = \infty$ или $R_2 = \infty$), а другое сферическое вогнутое (рис. 4-28, б), то условие устойчивости (4-76) будет иметь вид:

$$0 \leq \left(1 - \frac{l}{R_{1,2}}\right) \leq 1,$$

т. е. система устойчива в области значений $l/R_{1,2} < 1$. Оптимальный режим соответствует середине области $l = R_{1,2}/2$.

Если оба зеркала сферические вогнутые (рис. 4-28, в), то согласно (4-76) система устойчива, если

$$\frac{l}{R_1} < 1 \text{ и } \frac{l}{R_2} < 1,$$

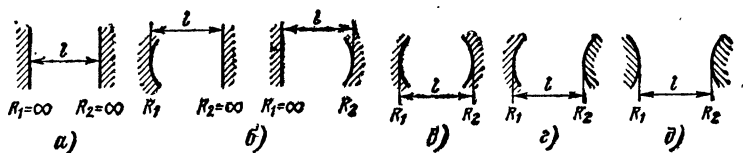


Рис. 4-28. Различные виды открытых резонаторов.

т. е. центры обоих зеркал лежат вне резонатора. Система устойчива и при

$$\frac{l}{R_1} > 1 \text{ и } \frac{l}{R_2} > 1,$$

т. е. когда центры зеркал лежат внутри резонатора, но при выполнении дополнительного условия

$$R_{1,2} \leq l \leq R_1 + R_2.$$

Система (рис. 4-28, в) неустойчива, если согласно (4-76)

$$\frac{l}{R_1} < 1, \text{ но } \frac{l}{R_2} > 1,$$

или

$$\frac{l}{R_1} > 1, \text{ но } \frac{l}{R_2} < 1.$$

Если в резонаторе одно зеркало (R_1) вогнутое, а другое (R_2) — выпуклое (рис. 4-28, з), то согласно условию (4-76), которое в этом случае имеет вид

$$0 \leq \left(1 - \frac{l}{R_1}\right) \left(1 + \frac{l}{R_2}\right) \leq 1, \quad (4-77)$$

для устойчивой работы системы необходимо, чтобы

$$l < R_1 \text{ и } 1 + \frac{l}{R_2} \leq \frac{1}{1 - \frac{l}{R_1}},$$

или

$$R_1 - R_2 \leq l \leq R_1.$$

Система (рис. 4-28, з) неустойчива, если

$$l > R_1,$$

а также если

$$l > R_1 \text{ при условии } l < R_1 - R_2.$$

Если оба зеркала выпуклые (рис. 4-28, д), то условие (4-76) не может быть выполнено и система неустойчива при любых значениях R_1 , R_2 и l .

На рис. 4-29 представлена диаграмма, позволяющая оценить устойчивость открытых резонаторов с различными сферическими зеркалами.

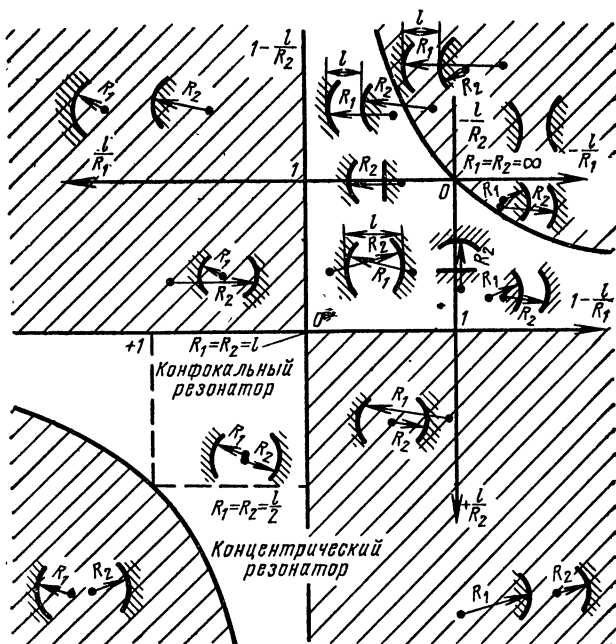


Рис. 4-29. Диаграмма устойчивости резонаторов (заштрихованные области неустойчивы).

Поля устойчивых типов резонаторов сконцентрированы у оси сильнее, чем для неустойчивых, поэтому дифракционные потери неустойчивых резонаторов много больше, чем для устойчивых. Сферические вогнутые зеркала, вогнутое и выпуклое зеркала в зависимости от расстояния между ними могут образовывать как устойчивую, так и неустойчивую систему. Переход от устойчивого состояния к неустойчивому тем резче, чем больше число Френеля. При переходе от устойчивого состояния к неустойчивому резко возрастают дифракционные потери и различие потерь для различных типов.

Добротность открытого резонатора, как и любой резонансной системы, определяется выражением

$$Q = \frac{\omega_p W_p}{P_0}, \quad (4-78)$$

где ω_p — резонансная круговая частота; W_p — энергия, запасенная при резонансе; P_0 — средняя мощность потерь данного типа колебаний.

Потери в резонаторе определяются потерями в зеркалах и потерями на дифракцию. Отражение от зеркала сопровождается частичным рассеянием, поглощением в зеркале и частичным прохождением через него. Все это можно рассматривать как потери в зеркалах независимо от того, используется ли рассеянное и прошедшее излучение в выходном излучении.

При отражении плоской волны от зеркала в результате дифракции отраженная волна представляет собой расходящийся пучок, угол расхождения которого равен $\phi = \lambda/D$ для квадратного зеркала со стороной D , и $\phi = 1,22\lambda/D$ для круглого зеркала с диаметром D . Вследствие конечных размеров зеркал часть энергии, зависящая от амплитуды волны на краю зеркала и угла ϕ , будет теряться при каждом отражении. Для устойчивых резонаторов при размерах зеркал, много больших длины волны, для низших поперечных типов дифракционные потери малы и потери резонатора определяются лишь потерями в зеркалах, которые одинаковы для всех типов колебаний.

Пренебрежем дифракционными потерями и потерями на поглощение и будем при вычислении добротности открытого резонатора учитывать лишь потери за счет частичного прохождения излучения через зеркала.

Энергия электромагнитного поля резонатора W_p распределяется поровну между прямой и обратной волнами. Энергия прямой волны $W_p/2$, заключенная в объеме резонатора, проходит через любое поперечное сечение за время l/v , где l — длина резонатора; v — скорость распространения волны. При этом средний поток мощности через поперечное сечение равен:

$$\int_S \Pi_0 ds = \frac{W_p v}{2l},$$

где Π_0 — среднее значение вектора Пойнтинга.

Средняя мощность потерь на каждом из зеркал при коэффициенте отражения по напряженности электрического поля Γ_E равна:

$$P_{0 \text{ зерк}} = \frac{W_p v}{2l} (1 - |\Gamma_E|^2) = \frac{P_0}{2}$$

и, согласно (4-78), добротность резонатора

$$Q = \frac{\omega_p}{v} \frac{l}{1 - |\Gamma_E|^2} = \frac{2\pi l}{\lambda_p} \frac{1}{1 - |\Gamma_E|^2}. \quad (4-79)$$

При $l/\lambda \gg 1$ и больших значениях $|\Gamma_E|^2$ добротность должна быть очень большой. Например, при $l=1$ м, $\lambda_p=0,63 \cdot 10^{-6}$ м и $|\Gamma_E|^2=0,99$, согласно (4-79),

$$Q = \frac{6,28 \cdot 1}{0,63 \cdot 10^{-6}} \frac{1}{1 - 0,99} \approx 10^9.$$

Ширина резонансного пика определяется выражением

$$\Delta f_p = \frac{f_p}{Q} = \frac{v(1 - |\Gamma_E|^2)}{2\pi l}.$$

Для нашего примера, если резонатор пустой,

$$\Delta f_p = \frac{3 \cdot 10^8 (1 - 0,99)}{6,28 \cdot 1} \approx 0,5 \text{ МГц.}$$

Расстояние по шкале частот между резонансными пиками согласно (4-50) равно:

$$\Delta f = \frac{3 \cdot 10^8}{2} = 150 \text{ МГц.}$$

При учете дифракционных потерь добротности различных типов колебаний будут различаться, так как их

дифракционные потери различны. Наименьшими дифракционными потерями обладает основной тип колебания T_{00} . Для поперечных типов T_{mn} дифракционные потери больше и растут с увеличением значений m и n . В формуле (4-79) величина $(1 - |\Gamma_E|^2)$ характеризует относительные потери на одно отражение от зеркала. Очевидно с учетом дифракционных потерь

$$Q_{mn} = \frac{2\pi l}{\lambda_p} \frac{1}{1 - |\Gamma_E|^2 + \alpha_{\text{диф } mn}}, \quad (4-80)$$

где $\alpha_{\text{диф } mn}$ — дифракционные потери типа T_{mn} на одно отражение.

Добротность тем больше, чем больше длина резонатора. Однако это справедливо лишь до тех пор, пока дифракционные потери меньше потерь в зеркалах. Потери в зеркалах не зависят от длины резонатора, дифракционные потери увеличиваются с увеличением длины. Поэтому после того момента, когда дифракционные потери и потери в зеркалах с увеличением длины сравняются, дальнейшее увеличение длины l согласно (4-80) приводит к уменьшению добротности Q_{mn} .

Активный резонатор. Резонатор, частично или полностью заполненный активной средой, в котором происходит усиление или генерирование электромагнитных колебаний за счёт взаимодействия поля со средой при многократном отражении от зеркал, называется *активным*.

Усиление внешнего электромагнитного поля, поступающего в резонатор, происходит, если прирост мощности за счёт взаимодействия усиливаемого поля с активной средой больше, чем собственные потери резонатора, определяемые потерями в среде, зеркалах (поглощение и рассеяние) и дифракционными потерями, но меньше, чем полные потери резонатора, включающие и излучение через зеркала (§ 3-1). Генерирование электромагнитного поля — когда прирост мощности больше полных потерь.

Усиление среды определяется инверсной населенностью [см. (3-2)], которая в свою очередь зависит от мощности источника возбуждения. Наименьшие значения мощности источника возбуждения и коэффициента квантового усиления, при которых начинается усиление внешнего сигнала или генерирование, называются *поро-*

говыми. При этих значениях происходит компенсация собственных потерь резонатора в случае регенеративного усилителя и полных потерь в случае генератора.

Рассмотрим работу лазера, резонатор которого полностью заполнен активной средой (рис. 4-30), в режиме генерирования. Поле в резонаторе можно рассматривать как результат многократного отражения волны от зеркал 1 и 2. Пусть амплитуда напряженности волны у зеркала 1 равна $\dot{E}_{m0}(0)$. У зеркала 2 после однократного прохождения резонатора и отражения от этого зеркала комплексная амплитуда волны определится выражением

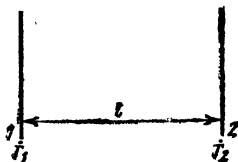


Рис. 4-30. К условию самовозбуждения генератора.

$$\dot{E}_{m0}(l) = \dot{\Gamma}_2 \dot{E}_{m0}(0) e^{(\alpha_{yc} - \alpha_n)l} e^{j\beta l},$$

где $\dot{\Gamma}_2 = \Gamma_2 e^{j\varphi_2}$ — коэффициент отражения от зеркала 2, учитывающий потери на излучение и фазовый сдвиг при отражении; α_n — коэффициент собственных потерь резонатора; α_{yc} — коэффициент квантового усиления активной среды; $\beta = 2\pi/\lambda_p$ — фазовая постоянная.

Вторично к зеркалу 1 волна подойдет с комплексной амплитудой

$$\dot{E}_m(0) = \dot{\Gamma}_2 \dot{E}_{m0}(0) e^{2(\alpha_{yc} - \alpha_n)l} e^{j2\beta l},$$

и, отразившись от этого зеркала, пойдет с амплитудой

$$\dot{E}'_{m0}(0) = \dot{\Gamma}_1 \dot{\Gamma}_2 \dot{E}_{m0}(0) e^{2(\alpha_{yc} - \alpha_n)l} e^{j2\beta l},$$

где $\dot{\Gamma}_1 = \Gamma_1 e^{j\varphi_1}$ — коэффициент отражения от зеркала 1.

Волна будет самоподдерживающейся, если

$$\dot{E}_{m0}(0) = \dot{E}'_{m0}(0),$$

т. е. выполняется условие существования стационарного режима

$$\dot{\Gamma}_1 \dot{\Gamma}_2 e^{2(\alpha_{yc} - \alpha_n)l} e^{j2\beta l} = 1$$

или

$$\Gamma_1 \Gamma_2 e^{2(\alpha_{yc} - \alpha_n)l} e^{j(2\beta l + \varphi_1 + \varphi_2)} = 1.$$

Отсюда получаем два самостоятельных условия существования стационарного режима: *условие баланса фаз*:

$$\frac{2\pi}{\lambda_p} 2l + \varphi_1 + \varphi_2 = 2\pi q \quad (q = 0, 1, 2, \dots)$$

и *условие баланса амплитуд*:

$$\Gamma_1 \Gamma_2 e^{2(\alpha_{yc} - \alpha_n)l} = 1.$$

Условие баланса фаз означает, что в стационарном режиме волна, пройдя в резонаторе путь $2l$ и отразившись от каждого из зеркал, получает фазовый сдвиг, кратный 2π . Согласно формуле (4-48) это условие выполняется для всех резонансных частот и условие стационарного режима сводится к условию баланса амплитуд.

В последнем уравнении экспоненциальный член $e^{2\alpha_{yc}l}$ характеризует индуцированное излучение, остальные члены определяют потери в резонаторе. Отсюда следует, что стационарные колебания будут существовать, если индуцированное излучение компенсирует все потери в резонаторе. Наименьшее или пороговое значение коэффициента усиления определяется выражением

$$\alpha_{yc, \text{пор}} = \alpha_n + \frac{1}{2l} \ln \frac{1}{\Gamma_1 \Gamma_2}.$$

Генератор возбуждётся, если коэффициент усиления активной среды больше или равен пороговому значению коэффициента усиления:

$$\alpha_{yc} \geq \alpha_n + \frac{1}{2l} \ln \frac{1}{\Gamma_1 \Gamma_2}; \quad (4-81)$$

это условие называется условием самовозбуждения оптического квантового генератора.

Пороговое значение $\alpha_{yc, \text{пор}}$ обеспечивается согласно (3-3) при некоторой пороговой инверсной населённости $\Delta N_{\text{пор}}$. Если инверсная населённость превышает пороговое значение, то будет наблюдаться нарастание интенсивности колебаний. При этом за счёт увеличения числа индуцированных переходов инверсная населённость будет уменьшаться до тех пор, пока не достигнет порогового значения $\Delta N_{\text{пор}}$, при котором обеспечивается стационарный режим.

Согласно выражению (4-81), чем больше коэффициент отражения и длина активной среды, тем легче выполняется условие генерирования. Но коэффициенты Γ_1 и Γ_2 принципиально оба не могут равняться 1, по крайней мере одно из зеркал должно быть полупрозрачным для вывода излучения. Длина активной среды не может быть сколь угодно большой не только из-за технологических трудностей, но также из-за увеличения дифракционных потерь.

Количество типов колебаний, которые одновременно возникают в активном резонаторе, ограничивается контуром усиления (рис. 4-31, а), определяемым формулой (3-3). Высота этого контура определяется инверсной населенностью, ширина — шириной линии излучения рабочего перехода (см. П1, п. 20). Ширина линии рабочего перехода (контура усиления) больше, чем расстояние между резонансными частотами продольных типов колебаний,

а так как частоты поперечных типов различны, то на ширине линии рабочего перехода укладывается несколько поперечных типов колебаний с соответствующими наборами продольных типов, сдвинутых относительно друг друга. Для упрощения на рис. 4-31, б приведены резонансные пики только продольных типов основного колебания T_{00} . При малой инверсной населенности и $\alpha_{ус} < \alpha_{ус.пор}$, как на рис. 4-31, в, ни для одной из резонансных частот не выполняется условие баланса амплитуд и генерирование не возникает ни на одном из типов колебаний. С увеличением инверсной населенности усиление активной среды возрастает (рис. 4-31, г) и условие баланса амплитуд выполняется для одной или нескольких

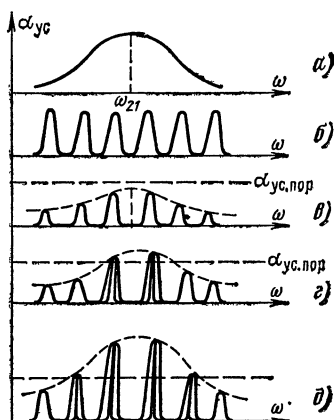


Рис. 4-31. Спектр излучения квантового генератора.

а — контур усиления $\alpha_{ус}$; б — резонансные пики продольных типов; в — малая инверсная населенность; г — условие баланса амплитуд выполняется для середины контура усиления; д — большая инверсная населенность.

частот. При дальнейшем увеличении инверсной населенности число генерируемых частот увеличивается (рис. 4-31, *д*). Начинается генерирование не только продольных типов основных колебания, но и поперечных, добротность резонатора для которых незначительно отличается от добротности для основного колебания. В результате спектр излучения состоит из продольных типов T_{00} и группирующихся около них по частоте продольных ти-

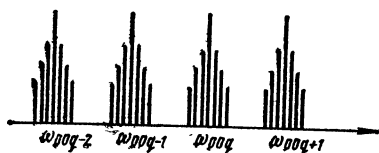
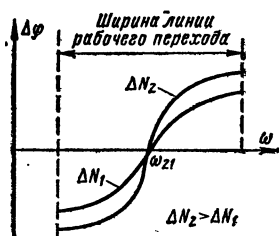


Рис. 4-32. Спектр излучения активного резонатора.

Рис. 4-33. Фазочастотная характеристика активной среды.



пов поперечных колебаний (рис. 4-32). Расстояние между частотами продольных типов составляет сотни мегагерц, ширина диапазона частот, в котором группируются продольные типы различных поперечных типов — единицы мегагерц.

Так как активная среда обладает дисперсией, т. е. скорость распространения волны в среде зависит от частоты, то происходит сдвиг резонансных частот активного резонатора по сравнению с резонансными частотами пассивного резонатора. Резонансные частоты для активного резонатора с плоскими зеркалами определяются согласно (4-49) выражением

$$f_{mnq} = \frac{v}{2\pi l} (\varphi_{mn} + q\pi)$$

или

$$f_{mnq} = \frac{v_0}{2\pi l} (\varphi_{mn} + q\pi - \Delta\varphi),$$

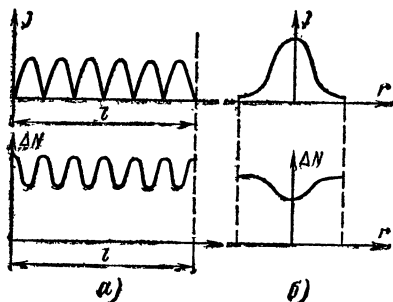
где v_0 — скорость волны для частоты рабочего перехода активной среды ω_{21} , $\Delta\varphi$ — изменение фазы за счет дисперсии активной среды.

Зависимость $\Delta\varphi$ от расстройки относительно центра контура усиления ω_{21} представляет собой фазочастотную характеристику активной среды (рис. 4-33). Если резо-

нансная частота какого-либо типа колебаний в пассивном резонаторе $\omega_{mnq} < \omega_{21}$, то в соответствующем активном резонаторе (при возбуждении среды) эта резонансная частота увеличивается. Если же $\omega_{mnq} > \omega_{21}$, то при возбуждении среды эта резонансная частота уменьшается. Таким образом, все частоты смещаются к центру контура усиления. Это смещение составляет десятки доли мегагерца.

Рис. 4-34. Распределение интенсивности и инверсной населенности в случае колебания T_{00} .

a — по длине резонатора;
б — в поперечном направлении.



С увеличением инверсной населенности крутизна фазочастотной характеристики увеличивается и затягивание проявляется сильнее.

На формирование спектра излучения большое влияние оказывает пространственная модуляция инверсной населенности. Если до начала генерирования инверсная населенность примерно постоянна во всем объеме активной среды, то появление любого типа колебания приводит к модуляции инверсной населенности. Пусть в резонаторе существует, например, один продольный тип колебания T_{00} . Распределение интенсивности поля для этого типа вдоль оси резонатора и в поперечном сечении приведено на рис. 4-34. Чем больше интенсивность, тем больше индуцированных переходов, и там, где интенсивность поля для данного типа больше, инверсная населенность будет меньше. В результате условия возникновения для другого типа колебаний будут лучше, чем для существующего, и он возникает. После возникновения другого типа, распределение поля которого не совпадает с существовавшими ранее, вновь изменяется распределение инверсной населенности и вновь возникает новый тип колебания. Происходит «перескок» типов колебаний. Пространственная модуляция инверсной населенности

создает условия для появления высших типов, несмотря на то, что дифракционные потери для этих типов больше, чем для основного. Действительно, согласно рис. 4-34, б инверсная населенность, а следовательно, и усиление на периферии среды, при наличии основного типа колебания больше, чем в центре, что создает условия для возникновения типа T_{10} .

Теоретически ширина спектральной линии работающего типа колебаний должна быть меньше ширины резонансной кривой того же типа колебаний в неработающем лазере, так как усиление в центре резонансной кривой больше, чем на ее краях. Однако практически ширина спектральной линии излучения из-за нестабильности резонатора вследствие механических и тепловых воздействий оказывается значительно больше теоретической.

Для работы лазера в режиме регенеративного усиления должно выполняться условие

$$\alpha_n < \alpha_{yc} < \alpha_n + \frac{1}{2l} \ln \frac{1}{\Gamma_1 \Gamma_2}; \quad (4-82)$$

в этом случае пороговый уровень возбуждения определяется собственными потерями резонатора α_n . Степень превышения мощности возбуждения над пороговым значением определяет коэффициент усиления. Усиление возможно при мощностях возбуждения, не превышающих пороговое значение генерирования (самовозбуждения).

4-3. СЕЛЕКЦИЯ ТИПОВ КОЛЕБАНИЙ

В оптическом диапазоне ширина линии рабочего перехода значительно больше расстояния между частотами типов колебаний, возбуждаемых в открытом резонаторе. Поэтому обычно генерирование происходит одновременно на большом числе типов колебаний. При этом структура поля оказывается очень сложной, диаграмма направленности сильно изрезана.

Число возбуждающихся типов колебаний и их структуру можно определить по форме и структуре пятна на выходном зеркале, диаграмме направленности излучения и спектру излучаемых частот. Наличие большого числа типов колебаний ухудшает когерентность и монохроматичность излучения, искажает и расширяет диаграмму направленности, мощность излучения распреде-

ляется между всеми типами и в основном типе излучается малая часть общей мощности.

Самый простой метод получения одночастотного режима — это работа при значении коэффициента усиления активной среды, близком к пороговому. Усиление активной среды определяется плотностью инверсной населенности, которая зависит от мощности источника возбуждения. При мощности возбуждения, близкой к пороговой, усиление может превосходить потери лишь для одного вида колебаний (рис. 4-31, *г*). Однако при этом мощность излучения невелика, а к стабильности уровня возбуждения предъявляются жесткие требования.

Поперечные типы колебаний, различающиеся поперечными индексами m и n и, следовательно, распределением амплитуды и фазы на поверхностях зеркал, влияют на пространственную когерентность.

Зависимость пространственной когерентности от числа поперечных типов колебаний определяется выражением

$$|\gamma_{12}(0)|^2 = \frac{\sin^2 \frac{\pi N \Delta r}{4Dn}}{N^2 \sin^2 \frac{\pi \Delta r}{4Dn}},$$

где N — число поперечных типов; Δr — расстояние, на котором определяется когерентность; D — диаметр зеркала; n — показатель преломления.

Подавление поперечных типов колебаний приводит к повышению пространственной когерентности и увеличивает направленность излучения вплоть до дифракционного предела.

Подавление нежелательных высших поперечных типов колебаний осуществляется путем создания в резонаторе условий, при которых дифракционные потери высших типов становятся значительно больше, чем потери основного низшего типа T_{00} . Как видно из рис. 4-12 и 4-16, дифракционные потери растут при уменьшении числа Френеля $N_{\text{ф}}$. При $N_{\text{ф}} \ll 1$ резонатор становится высокоселективной системой. Однако для обычных плоского и конфокального резонаторов число Френеля, меньшее единицы, практически невозможно получить при приемлемых размерах резонаторов. Это возможно для резонатора, образованного плоским и сферическим зеркалами, расположенными на расстоянии, приблизительно равном

радиусу зеркала (полуконцентрический резонатор). При расстоянии между зеркалами, значительно меньшем радиуса, добротность резонатора для всех типов колебаний одинакова, так как потери определяются только потерями в зеркалах и полезным излучением лазера, а они одинаковы для всех типов. Однако с увеличением длины резонатора дифракционные потери быстро растут и становятся различными для разных типов, поэтому добротность, определяемая выражением (4-80), для раз-

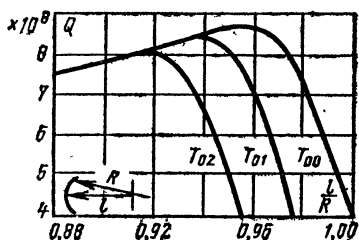


Рис. 4-35. Зависимость добротности плоскосферического резонатора от его длины.

ных типов колебаний уменьшается в различной степени (рис. 4-35).

При увеличении расстояния между зеркалами до значений $l/R \approx 1$ число типов колебаний уменьшается при незначительном уменьшении мощности. При $l/R \approx 0,975$ наблюдается только один основной тип колебания T_{00} (однотиповый или одномодовый режим). При этом мощность излучения составляет 0,8 от максимально излучаемой мощности. При увеличении расстояния ($l > R$) мощность основного типа колебаний падает.

Селекция поперечных типов колебаний может быть осуществлена с помощью специальных диафрагм. Поле основного типа колебаний T_{00} концентрируется ближе к оси резонатора, чем поле поперечных типов T_{11} , T_{20} , T_{21} и т. д. При этом чем больше порядок типа колебаний, тем дальше простирается поле от оси. Обычно диафрагмы представляют собой отверстие в непрозрачной пластинке. Подбирая размеры диафрагмы, можно увеличить дифракционные потери поперечных типов и подавить все типы колебаний, кроме основного. Однако в этом случае и для основного типа вносятся значительные потери. Этот метод может быть усовершенствован путем помещения диафрагмы в совмещенном фокусе двух линз. При этом размер диафрагм меньше, чем в случае отсут-

ствия линз, и селективность системы к поперечным типам увеличивается. Роль, аналогичную диафрагме, играют и зеркала, распределение коэффициента отражения по поверхности которых соответствует конфигурации поля основного типа колебаний. Практически это осуществляется нанесением покрытия лишь на те участки зеркала, где поле этого типа максимально.

Селекция поперечных типов колебаний может быть осуществлена с помощью плоскопараллельной полупрозрачной пластины, устанавливаемой в резонаторе под некоторым углом к его оси. В системе будут возникать

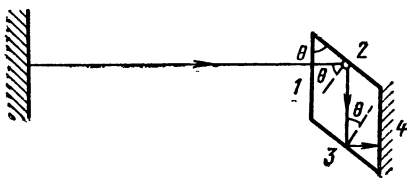


Рис. 4-36. Ромбовидная призма.

только те типы, для которых одновременно будут выполняться условия резонанса и оптимального прохождения через пластину. Коэффициент пропускания такой пластины зависит не только от длины волны, но и от угла падения на эту пластину. Селекция в обеих плоскостях осуществляется двумя пластинами, расположенными под некоторыми углами к оси резонатора в двух взаимно перпендикулярных плоскостях.

Эффективным методом селекции поперечных типов колебаний является такое изменение формы зеркал и расстояния между ними, которое делает резонаторы неустойчивыми. При этом увеличивается разность дифракционных потерь для различных типов колебаний. Работа резонаторов в области неустойчивости повышает эффективность диафрагмы и наклонной плоскопараллельной пластины как селекторов поперечных типов. Наиболее эффективны неустойчивые резонаторы с выпуклыми зеркалами.

Работа призмных селекторов основана на явлении полного внутреннего отражения. Ромбовидная призма, острый угол которой θ равен предельному углу полного внутреннего отражения (для призмы из плавленого

кварца $n=1,4555$, $\theta=43^\circ 24' 10''$) служит отражателем (рис. 4-36). Основной тип колебания падает нормально на грань 1, полностью отражается от грани 2 и 3 и нормально падает на грань 4 с диэлектрическим покрытием. Отразившись от грани 4, луч тем же путем возвращается в резонатор. Поперечным типам соответствуют лучи, наклонные к оси резонатора, которые после отражения от граней призмы покидают резонатор.

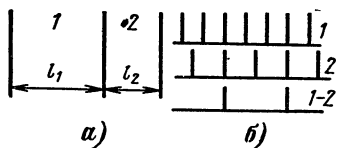


Рис. 4-37. Селекция типов колебаний в связанных резонаторах.

a — связанные резонаторы; b — спектр колебаний (1 — в резонаторе 1; 2 — в резонаторе 2, $1-2$ в связанных резонаторах).

При работе генератора на одном основном типе T_{00} поле на зеркалах описывается функцией Гаусса, диаграмма излучения становится узкой и гладкой. В спектре излучения остаются лишь частоты, соответствующие продольным типам колебаний.

Собственные частоты открытого резонатора, соответствующие различным продольным типам колебаний для данного поперечного типа, образуют спектр частот, расстояние между которыми в зависимости от вида резонатора определяется формулами (4-50) или (4-59) — (4-61). Количество продольных типов, возбуждающихся в активном резонаторе, определяется шириной линии излучения активного вещества.

Генерирование большого числа продольных типов колебаний в активном резонаторе обусловлено пространственной неоднородностью поля. Активные атомы, находящиеся в пучностях поля, взаимодействуют с полем наиболее интенсивно и совершают индуцированные переходы; атомы, находящиеся в минимумах поля данного типа колебаний, не совершают индуцированных переходов и являются источниками возбуждения других типов. Подавление продольных типов колебаний ведет к сужению спектра излучения, повышению временной когерентности.

Наиболее эффективным методом подавления продольных типов колебаний является использование системы связанных резонаторов. Максимальную добротность

в такой системе будут иметь типы колебаний, являющиеся резонансными для каждого резонатора, входящего в систему. В простейшем случае двух связанных резонаторов, образованных тремя зеркалами (рис. 4-37), подбирая длины этих резонаторов и коэффициент пропускания среднего зеркала, можно добиться работы на тех типах колебаний, резонансные частоты которых совпадают, существенно ослабив или подавив остальные (паразитные) типы. Эффективность этого метода повышается при увеличении числа связанных резонаторов. При настройке расстояний между зеркалами с точностью до долей длины волны и подборе коэффициентов отражения зеркал практически возможно получить одночастотный режим работы.

Свойство связанных резонаторов селективировать продольные типы колебаний используется, когда в качестве одного из отражателей в оптическом генераторе используется стопа из нескольких параллельных пластин. Такие пластины представляют собой резонаторы, в которых собственные частоты продольных колебаний находятся на расстояниях больших, чем у основного резонатора.

Для селекции продольных типов колебаний применяются полупрозрачные параллельные пластины, устанавливаемые под определенными углами к оси резонатора в двух взаимно перпендикулярных плоскостях. Максимальное прохождение излучения с длиной волны λ через пластину толщиной d и наклоненной под углом θ , очевидно, определяется условием

$$\frac{d}{\cos \theta} = n \frac{\lambda}{2}$$

и в системе будут существовать только те продольные типы, которые одновременно удовлетворяют условиям существования в резонаторе и этому условию.

Число продольных и поперечных типов колебаний можно уменьшить, уменьшив пространственную продольную и поперечную модуляцию инверсной населенности. Для этого используют резонатор бегущей волны кольцевого типа (рис. 4-38). Одно из зеркал для уменьшения дифракционных потерь и обеспечения меньшей критичности настройки, имеет слабую вогнутость. Волны, распространяющиеся по часовой и против часовой стрелки в резонаторе, независимы. Помещая в резонатор опти-

ческий вентиль (см. § 4-4), устраняют одну из волн. В таком резонаторе нет продольной модуляции инверсной населенности, так как волна бегущая. Поперечную оптическую неоднородность, а следовательно, и число поперечных типов колебаний можно уменьшить, создав в резонаторе бегущей волны с помощью оптической схемы

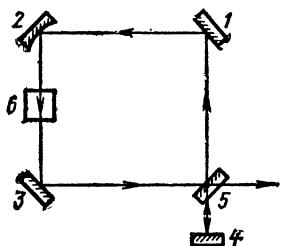


Рис. 4-38. Резонатор бегущей волны.
1, 2, 3, 4 — глухие зеркала; 5 — полупрозрачное выходное зеркало; 6 — оптический вентиль.

вращающееся поле. Продольную модуляцию инверсной населенности можно уменьшить и в обычном резонаторе, перемещая активный элемент в продольном направлении со скоростью примерно 50 м/с.

4-4. УСИЛИТЕЛЬ БЕГУЩЕЙ ВОЛНЫ

Усилитель бегущей волны принципиально содержит два основных элемента: активную среду и источник возбуждения. Если в активной среде создана инверсная населенность, то такой усилитель способен усиливать на частоте квантового перехода.

В режиме бегущей волны усиливаемое поле распространяется в активной среде только в одном направлении. Это всегда можно осуществить, используя оптические вентили, работа которых основана на использовании эффекта Фарадея — повороте плоскости поляризации при прохождении света через среду с неважными свойствами при продольном подмагничивании. Оптический вентиль состоит из двух поляризаторов, между которыми расположен элемент с неважными свойствами (рис. 4-39). Главные плоскости поляризаторов ориентированы под углом 45° друг к другу, неважный элемент поворачивает плоскость поляризации на 45° . При этом линейно поляризованная волна, идущая в

прямом направлении и проходящая через первый поляризатор, пройдет и через второй. Отраженная волна, идущая в обратном направлении, через вентиль не проходит.

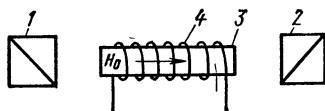
Угол поворота плоскости поляризации ϑ определяется формулой:

$$\vartheta = BH_0 l,$$

где H_0 — напряженность постоянного магнитного поля; l — длина невзаимного элемента; B — постоянная Верде, зависящая от параметров материала и обратно пропорциональная квадрату длины волны.

Рис. 4-39. Оптический вентиль.

1, 2 — поляризаторы; 3 — невзаимный элемент; 4 — соленоид.



Эффективность работы вентилья определяется качеством развязки, т. е. отношением мощности прямой и обратной волны (20—30 дБ), и вносимыми потерями (1—1,4 дБ). В качестве рабочего материала для невзаимных элементов используются различные сорта стекол, преимущественно свинцовые.

Основными уравнениями для исследования усилителя бегущей волны в оптическом диапазоне являются векторные уравнения (3-31) или уравнения в скалярной форме (3-32) или (3-34). Уравнения (3-34) можно упростить, учитывая, что в оптическом диапазоне ширина линии излучения рабочего перехода достаточно велика, т. е.

$$\frac{P_m^a}{T_2} \gg \frac{\partial P_m^a}{\partial t}. \quad (4-83)$$

При этом согласно третьему уравнению системы (3-34)

$$P_m^a = \frac{p_e^2}{\hbar} T_2 \Delta N E_m \sin(\psi - \varphi). \quad (4-84)$$

Подставляя (4-84) в (3-34), получаем:

$$\left. \begin{aligned} \frac{\partial E_m}{\partial t} + v \frac{\partial E_m}{\partial x} + v \alpha_n E_m &= \frac{\omega}{2\varepsilon_a \hbar} p_e^2 T_2 \Delta N \times \\ &\times E_m \sin^2 (\psi - \varphi); \\ \frac{\partial \varphi}{\partial t} + v \frac{\partial \varphi}{\partial x} &= - \frac{\omega}{2\varepsilon_a \hbar} p_e^2 T_2 \Delta N \times \\ &\times \sin (\psi - \varphi) \cos (\psi - \varphi); \\ \frac{\partial \psi}{\partial t} + (\omega - \omega_{21}) &= \frac{1}{T_2} \operatorname{ctg} (\psi - \varphi); \\ \frac{\partial \Delta N}{\partial t} + \frac{\Delta N - \Delta N_0}{T_1} &= - \frac{p_e^2}{\hbar^2} \times \\ &\times T_2 \Delta N E_m^2 \sin^2 (\psi - \varphi). \end{aligned} \right\} \quad (4-85)$$

В стационарном режиме

$$\frac{\partial E_m}{\partial t} = \frac{\partial \varphi}{\partial t} = \frac{\partial \psi}{\partial t} = \frac{\partial \Delta N}{\partial t}$$

и система уравнений (4-85) принимает вид:

$$\left. \begin{aligned} \frac{\partial E_m}{\partial x} + \alpha_n E_m &= \frac{\omega}{2\varepsilon_a \hbar v} p_e^2 T_2 \Delta N E_m \sin^2 (\psi - \varphi); \\ \frac{\partial \varphi}{\partial x} &= - \frac{\omega}{2\varepsilon_a \hbar v} p_e^2 T_2 \Delta N \sin (\psi - \varphi) \cos (\psi - \varphi); \\ \operatorname{ctg} (\psi - \varphi) &= T_2 (\omega - \omega_{21}); \\ \frac{\Delta N - \Delta N_0}{T_1} &= - \frac{p_e^2}{\hbar^2} T_2 \Delta N E_m^2 \sin^2 (\psi - \varphi). \end{aligned} \right\} \quad (4-86)$$

Согласно третьему уравнению системы (4-86)

$$\left. \begin{aligned} \sin (\psi - \varphi) &= \frac{1}{\sqrt{1 + T_2^2 (\omega - \omega_{21})^2}}; \\ \cos (\psi - \varphi) &= \frac{T_2 (\omega - \omega_{21})}{\sqrt{1 + T_2^2 (\omega - \omega_{21})^2}}. \end{aligned} \right\} \quad (4-87)$$

С учетом этих соотношений система уравнений (4-86) принимает вид:

$$\left. \begin{aligned} \frac{\partial E_m}{\partial x} + \alpha_n E_m &= \frac{\omega p_e^2 T_2 \Delta N E_m}{2\varepsilon_a \hbar v [1 + T_2^2 (\omega - \omega_{21})^2]}; \\ \frac{\Delta N - \Delta N_0}{T_1} &= - \frac{p_e^2 T_2 \Delta N E_m^2}{\hbar^2 [1 + T_2^2 (\omega - \omega_{21})^2]}; \\ \frac{\partial \varphi}{\partial x} &= - \frac{\omega p_e^2 T_2 \Delta N (\omega - \omega_{21})}{2\varepsilon_a \hbar v [1 + T_2^2 (\omega - \omega_{21})^2]}. \end{aligned} \right\} \quad (4-88)$$

Первое уравнение системы (4-88) определяет изменение амплитуды поля при распространении в среде

$$E_m = E_{m0} e^{(\alpha_{yc} - \alpha_n)x},$$

где E_{m0} — амплитуда сигнала на входе усилителя бегущей волны;

$$\alpha_{yc} = \frac{\omega p_e^2 T_2 \Delta N}{2\varepsilon_a \hbar v [1 + T_2^2 (\omega - \omega_{21})^2]} \quad (4-89)$$

— коэффициент квантового усиления.

Согласно выражению (4-89) коэффициент квантового усиления возрастает с увеличением дипольного момента p_e и плотности инверсной населенности ΔN , так как при этом возрастает число излучающих частиц. Коэффициент квантового усиления зависит от частоты сигнала ω и имеет резонансный характер (рис. 4-40). Чем ближе частота сигнала ω к частоте квантового перехода ω_{21} , тем больше коэффициент квантового усиления. Ширина резонансной кривой равна $2/T_2$.

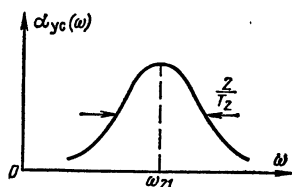


Рис. 4-40. Контур усиления $\alpha_{yc}(\omega)$.

Согласно второму уравнению системы (4-88) стационарная инверсная населенность определяется выражением

$$\Delta N = \frac{\Delta N_0}{1 + \frac{p_e^2 T_1 T_2}{\hbar^2 [1 + T_2^2 (\omega - \omega_{21})^2]} E_m^2}, \quad (4-90)$$

т. е. зависит от начальной инверсной населенности ΔN_0 , которая создается сторонним источником. По мере рас-

пространения сигнала в активной среде E_m^2 увеличивается, а инверсная населенность ΔN уменьшается. При $E_m^2 \rightarrow \infty$, $\Delta N \rightarrow 0$, $\alpha_{yc} \rightarrow 0$ происходит насыщение.

Подставляя (4-90) в (4-89), получаем:

$$\alpha_{yc} = \frac{\omega p_e^2 T_2 \Delta N_0}{2\varepsilon_a v \left\{ \hbar [1 + T_2^2 (\omega - \omega_{21})^2] + \frac{1}{\hbar} p_e^2 T_1 T_2 E_m^2 \right\}}, \quad (4-91)$$

Волна при распространении в активной среде усиливается при условии

$$\alpha_{yc} > \alpha_n. \quad (4-92)$$

Для выполнения этого условия необходимо, чтобы начальная инверсная населенность превышала пороговое значение. Очевидно, это значение определяется из условия, когда

$$\alpha_{yc} = \alpha_n, E_m = E_{m0}$$

и потери в среде компенсируются излучением частиц. Согласно (4-91)

$$\Delta N_{\text{пор}} = 2\varepsilon_a v \alpha_n \left\{ \frac{\hbar}{\omega p_e^2 T_2} [1 + T_2^2 (\omega - \omega_{21})^2] + \frac{T_1}{\omega \hbar} E_{m0}^2 \right\},$$

т. е. пороговое значение начальной плотности инверсной населенности уменьшается с ростом дипольного момента p_e , так как мощность излучения частиц растет с увеличением p_e и уменьшается при уменьшении расстройки сигнала относительно частоты перехода ω_{21} , так как при этом мощность излучения частиц также растет.

Переходя в уравнениях (4-85) с учетом (4-87) к интенсивностям (плотностям мощности излучения), получаем уравнения

$$\left. \begin{aligned} \frac{\partial J}{\partial t} + v \frac{\partial J}{\partial x} &= v [\sigma(\omega) \Delta N - 2\alpha_n] J; \\ \frac{\partial \Delta N}{\partial t} + \frac{\Delta N - \Delta N_0}{T_1} &= -2\sigma(\omega) J \Delta N; \\ \frac{\partial \Phi}{\partial t} + v \frac{\partial \Phi}{\partial x} &= (\omega_{21} - \omega) \frac{T_2}{2} v \sigma(\omega) \Delta N, \end{aligned} \right\} \quad (4-93)$$

где

$$J = \frac{\varepsilon_a v}{2\hbar\omega_{21}} E_m^2$$

— средняя плотность потока мощности излучения, фотонов/(м²·с), интенсивность;

$$\sigma(\omega) = \frac{T_2 \omega_{21} p_e^2}{\epsilon_a \hbar \nu [1 + (\omega - \omega_{21})^2 T_2^2]}$$

— сечение радиационного перехода на частоте ω , определяемое отношением мощности, поглощенной одной частицей, к мощности падающего излучения на единицу площади.

Первое уравнение системы (4-93) описывает изменение интенсивности сигнала при распространении. Первый член правой части уравнения определяет увеличение интенсивности распространяющегося сигнала за счет индуцированного излучения, второй член — уменьшение интенсивности за счет потерь.

Второе уравнение системы (4-93) определяет изменение разности населенностей уровней в единицу времени. Член $\frac{\Delta N - \Delta N_0}{T_1}$ характеризует уменьшение разности населенностей за счет релаксационных процессов и ее увеличение в результате возбуждения внешним источником.

Член $2\sigma(\omega) J \Delta N$ характеризует уменьшение разности населенностей за счет индуцированного излучения. Множитель 2 в правой части уравнения появляется вследствие того, что в двухуровневой системе излучение происходит при переходе одной частицы с верхнего уровня на нижний. При этом населенность верхнего уровня уменьшается на единицу, а нижнего увеличивается на единицу. В результате разность населенностей при таком переходе изменяется на 2.

Третье уравнение определяет изменение фазы в процессе распространения волны.

Часто пользуются только двумя первыми уравнениями системы (4-93)

$$\left. \begin{aligned} \frac{\partial J}{\partial t} + v \frac{\partial J}{\partial x} &= v [\sigma(\omega) \Delta N - 2\alpha_n] J; \\ \frac{\partial \Delta N}{\partial t} + \frac{\Delta N - \Delta N_0}{T_1} &= -2\sigma(\omega) J \Delta N, \end{aligned} \right\} \quad (4-94)$$

которые называются *кинетическими* или *скоростными*. Эти уравнения характеризуют скорость передачи энергии от среды полю по мере распространения волны.

В стационарном режиме, т. е. когда

$$\frac{\partial J}{\partial t} = \frac{\partial \Delta N}{\partial t} = 0,$$

уравнения (4-94) имеют вид:

$$\frac{\partial J}{\partial x} = [\sigma(\omega) \Delta N - 2\alpha_n] J; \quad (4-95)$$

$$\frac{\Delta N - \Delta N_0}{T_1} = -2\sigma(\omega) J \Delta N. \quad (4-96)$$

Решение уравнения (4-95)

$$J = J_0 e^{[\sigma(\omega) \Delta N - 2\alpha_n]x}$$

определяет усиление интенсивности сигнала при распространении в активной среде. Усиление сигнала происходит, если показатель степени этого выражения положительный, т. е.

$$\alpha_n < \frac{\sigma(\omega) \Delta N}{2},$$

что соответствует условию (4-92).

Согласно уравнению (4-96) стационарное значение инверсной населенности определяется выражением

$$\Delta N = \frac{\Delta N_0}{1 + 2\sigma(\omega) T_1 J}, \quad (4-97)$$

т. е. инверсная населенность, создаваемая источником возбуждения, уменьшается при увеличении интенсивности усиливаемого сигнала. При очень большой интенсивности ($J \rightarrow \infty$) инверсная населенность стремится к нулю (насыщение).

Пороговое значение начальной инверсной населенности определяется выражением

$$\Delta N_{\text{пор}} = 2\alpha_n \left[\frac{1}{\sigma(\omega)} + 2T_1 J_0 \right], \quad (4-98)$$

где J_0 — значение интенсивности сигнала на входе усилителя.

Усилитель как четырехполюсник (рис. 4-41) можно характеризовать коэффициентом усиления по интенсивности (мощности):

$$G = \frac{J_{\text{вых}}}{J_{\text{вх}}} = e^{2(\alpha_{\text{ус}} - \alpha_n)l} \quad (4-99)$$

где $J_{\text{вх}}$ и $J_{\text{вых}}$ — интенсивности сигнала на входе и выходе усилителя соответственно; l — длина активного элемента.

При небольших интенсивностях распространяющегося в активной среде сигнала и такой длине активного элемента, при которой усиливаемый сигнал далек от предельного значения при насыщении (см. § 3-1), инверсная населенность ΔN и коэффициент усиления среды $\alpha_{\text{ус}}$ согласно (4-90) и (4-91) практически неизменны. При этом значение коэффициента усиления G также неизменно и выходной сигнал прямо пропорционален входному сигналу (линейный режим).

При больших значениях сигнала на входе или достаточно больших длинах активных элементов усиливаемый сигнал вызывает настолько сильное индуцированное излучение, что инверсная населенность согласно (4-90) уменьшается, при этом уменьшается и $\alpha_{\text{ус}}$ (4-91). Прямой пропорциональности между выходным и входным сигналами уже не наблюдается. Равным приращением входного сигнала соответствуют все меньшие и меньшие приращения сигнала на выходе (нелинейный режим) и при некотором предельном значении входного сигнала $J_{\text{вх.пр}}$ наступает полное насыщение активной среды. При дальнейшем увеличении сигнала на входе сигнал на выходе больше не увеличивается. Зависимость интенсивности сигнала на выходе $J_{\text{вых}}$ от входного называется *амплитудной характеристикой усилителя* (рис. 4-42).

Интенсивность сигнала на выходе ограничивается значением коэффициента потерь $\alpha_{\text{п}}$. В идеальном случае $\alpha_{\text{п}}=0$, интенсивность сигнала на выходе может принимать любые значения вплоть до бесконечно больших. Действительно, усиление сигнала определяется превышением инверсной населенности над пороговым значением (4-98), которое в случае $\alpha_{\text{п}}=0$, также рав-

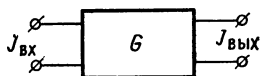


Рис. 4-41. Усилитель как четырехполюсник.

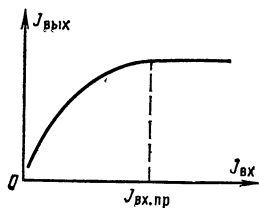


Рис. 4-42. Амплитудная характеристика усилителя.

няется нулю. Поэтому даже при незначительной инверсной населенности увеличение выходной интенсивности сигнала возможно за счет увеличения длины пути в активной среде. При определенной длине это возможно за счет увеличения инверсной населенности.

Динамический диапазон усилителя определяется диапазоном изменения интенсивности сигнала на входе, в пределах которого усилитель обеспечивает требуемый коэффициент усиления.

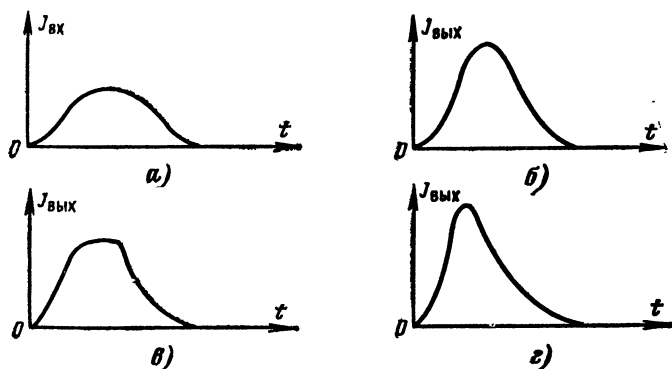


Рис. 4-43. Прохождение импульса через усилитель.

а — импульс на входе усилителя; *б* — линейный режим усилителя; *в* — импульс на выходе при малой входной интенсивности и длине активного элемента $l > l_{кр}$; *г* — при большой входной интенсивности и $l > l_{кр}$

Со стороны слабых сигналов динамический диапазон ограничен собственными шумами усилителя. Сигнал, мощность которого меньше некоторой пороговой величины, не может быть выделен на фоне шумов усилителя. Со стороны мощных входных сигналов динамический диапазон ограничен насыщением среды. Улучшение качества активных сред (уменьшение α_n) приводит к увеличению динамического диапазона и увеличению усиления ($\alpha_{ус} - \alpha_n$) на единицу длины среды.

Прохождение импульсного сигнала через усилитель зависит от его интенсивности на входе (рис. 4-43, *а*). В линейном режиме при малой интенсивности на входе и длине активного элемента, меньшей критического значения, при которой интенсивность распространяющегося сигнала достигает предельной интенсивности, импульс

практически проходит без искажений (рис. 4-43, б). Если длина активного элемента больше критической, то эффекты насыщения проявляются и при малой входной интенсивности. В режиме насыщения среды различные части импульса усиливаются неодинаково. Передняя часть импульса снижает инверсную населенность при своем прохождении и усиление остальных частей становится меньше. Поэтому передний фронт импульса на выходе становится круче, а вершина более

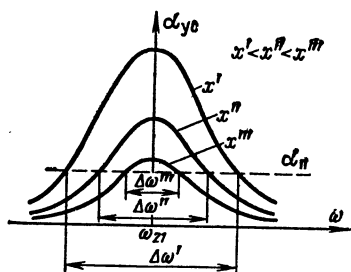


Рис. 4-44. Сужение спектра сигнала при распространении в активной среде с однородно уширенной линией излучения.

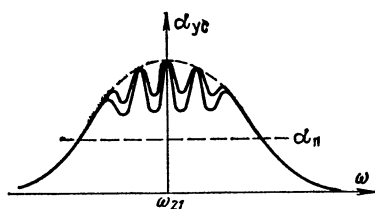


Рис. 4-45. Деформация контура неоднородно уширенной линии при усилении сигнала.

плоской (рис. 4-43, в). При увеличении интенсивности входного импульса эти эффекты проявляются резче, импульс обостряется и максимум у выходного сигнала сдвигается вперед по сравнению с входным (рис. 4-43, з).

При распространении сигнала в активной среде происходит изменение его спектрального состава. По мере распространения сигнала увеличивается его интенсивность, что приводит к увеличению числа индуцированных переходов и уменьшению инверсной населенности. Уменьшение инверсной населенности приводит согласно (4-89) к уменьшению коэффициента усиления α_{yc} . При однородном уширении линии излучения (см. П1, п. 20) увеличение интенсивности распространяющейся волны приводит к пропорциональному уменьшению усиления для всех частот (рис. 4-44). Для волны, прошедшей небольшое расстояние x' , интенсивность будет небольшой, инверсная населенность уменьшится незначительно и коэффициент усиления среды α_{yc} будет

значительным. При этом будут усиливаться спектральные составляющие в интервале $\Delta\omega'$. При увеличении пройденного расстояния до x'' коэффициент усиления уменьшится до α''_{yc} и усиление будет наблюдаться в меньшей полосе $\Delta\omega''$, т. е. происходит монохроматизация излучения.

При неоднородно уширенной линии излучения (см. П1, п. 20) взаимодействие распространяющегося поля происходит не со всеми частицами среды, а лишь с теми из них, в контуры линий излучения которых по-

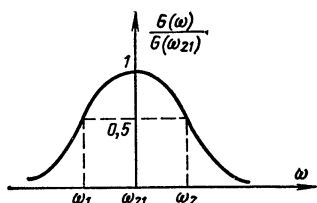


Рис. 4-46. К определению полосы пропускания усилителя.

падают спектральные составляющие сигнала (рис. 4-45). При увеличении пути, проходимого волной в среде, провалы углубляются. При уменьшении возбуждения среды инверсная населенность уменьшается, уменьшается усиление α_{yc} , уменьшается число спектральных составляющих сигнала и излучение монохроматизируется.

Зависимость коэффициента усиления (4-99) от частоты определяется контуром усиления активной среды (рис. 4-40). Усиление в децибелах определяется выражением

$$G_{дБ}(\omega) \approx 8,69\alpha_{yc}(\omega)l$$

или согласно (4-89)

$$G_{дБ}(\omega) = \frac{G_{дБ}(\omega_{21})}{1 + T_2^2(\omega - \omega_{21})^2} = \frac{G_{дБ}(\omega_{21})}{1 + \left[\frac{2(\omega - \omega_{21})}{\Delta\omega_L} \right]^2}, \quad (4-100)$$

где $\Delta\omega_L$ — ширина линии излучения среды; $G_{дБ}(\omega_{21})$ — коэффициент усиления в центре линии.

Полоса пропускания усилителя определяется как полоса частот, на границах которой коэффициент усиления по интенсивности убывает вдвое, т. е. на 3 дБ, относительно усиления в центре полосы (рис. 4-46). Согласно определению

$$G_{дБ}(\omega_1) = G_{дБ}(\omega_2) = G_{дБ}(\omega_{21}) - 3\text{дБ}.$$

С учетом (4-100) получаем:

$$G_{\text{дБ}}(\omega_{21}) - 3\text{дБ} = \frac{G_{\text{дБ}}(\omega_{21})}{1 + \left(\frac{\Delta\omega_{\text{ус}}}{\Delta\omega_{\text{л}}}\right)^2},$$

где $\Delta\omega_{\text{ус}}$ — ширина полосы пропускания усилителя.

Отсюда

$$\Delta\omega_{\text{ус}} = \Delta\omega_{\text{л}} \sqrt{\frac{3}{G_{\text{дБ}}(\omega_{21}) - 3}}, \quad (4-101)$$

т. е. при увеличении усиления на частоте перехода полоса усилителя уменьшается. При $G_{\text{дБ}}(\omega_{21}) \gg 3$ дБ выражение (4-101) переходит в выражение

$$F = \sqrt{G_{\text{дБ}}(\omega_{21})} \Delta\omega_{\text{ус}} \approx \sqrt{3} \Delta\omega_{\text{л}},$$

т. е. произведение корня квадратного из коэффициента усиления в центре полосы, выраженного в децибелах, на полосу усиления для данной среды является величиной постоянной, определяемой шириной линии излучения среды. Эта величина называется *эффективностью* усилителя.

Коэффициент полезного действия усилителя определяется отношением энергии, которую среда отдает распространяющейся электромагнитной волне, к энергии источника возбуждения. В режиме линейного усиления это отношение мало. Более высокий к. п. д. можно получить в режиме, близком к насыщению. Однако при приближении к полному насыщению к. п. д. падает до нуля, так как в этом случае энергия источника возбуждения расходуется только на компенсацию потерь и спонтанное излучение.

Так как вначале при усилении сигнала имеет место линейный режим, при котором эффективность использования инверсной среды мала, а на выходе усилителя энергия ограничивается насыщением, то выгодно использовать несколько каскадов усиления. При этом первые каскады имеют активные элементы с малым поперечным сечением, что уменьшает расход энергии возбуждения. При переходе к следующим каскадам усиления сечение активных элементов увеличивается, сечение светового луча также увеличивается или за счет естественного расхождения луча или с помощью телескопической системы (рис. 4-47). Предельная интенсивность не достигается и выходная энергия ограни-

чивается лишь трудностями возбуждения активных элементов большого поперечного сечения. Каскады разделяются оптическими вентилями. Общий коэффициент усиления такого многокаскадного усилителя равен про-

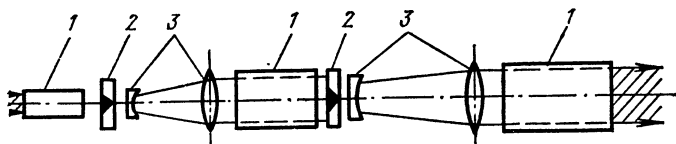


Рис. 4-47. Многокаскадный усилитель бегущей волны.

1 — активный элемент, 2 — оптический вентиль; 3 — телескопическая система.

изведению коэффициентов усиления отдельных каскадов; общая полоса пропускания уже полосы пропускания каждого из каскадов.

4-5. ТЕОРИЯ ОПТИЧЕСКИХ КВАНТОВЫХ ГЕНЕРАТОРОВ (ОКГ)

В основе работы оптических квантовых генераторов лежит принцип усиления вынужденного излучения в резонаторе. Независимо от типа резонатора генерирование возможно при коэффициенте усиления среды, превышающем порог самовозбуждения (4-81). Принципиальная схема генератора приведена на рис. 3-3.

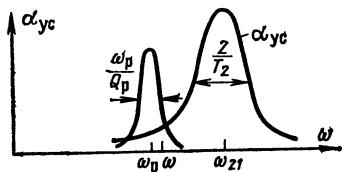


Рис. 4-48. Затягивание частоты.

Будем считать, что генератор работает в однократном режиме, что можно получить, применяя селек-

цию продольных и поперечных типов колебаний (см. § 4-3). Если частота выделенного колебания ω_p и частота квантового перехода ω_{21} не совпадают, происходит затягивание частоты, определяемое выражением (3-45). Так как ширина линии рабочего перехода $\Delta\omega_{\text{л}}$ много больше ширины резонансного пика $\Delta\omega_p$, то согласно (3-45) имеем:

$$(\omega - \omega_{21}) \gg (\omega_p - \omega).$$

Таким образом, частота генерирования значительно ближе к ω_p , чем к ω_{21} , т. е. происходит затягивание частоты более узкой линией (рис. 4-48).

Для исследования генератора при условии настройки резонатора на частоту линии квантового перехода ω_{21} используются векторные уравнения (3-41) и уравнения в скалярной форме (3-42) или (3-48). Хотя инверсная населенность и поле в резонаторе зависят от координат, можно рассматривать усредненные по координатам величины. Учет пространственных вариаций поля и инверсной населенности значительно усложняет решение, результаты же незначительно отличаются от результатов, полученных при пространственном усреднении.

Уравнения (3-48) можно упростить, учитывая (4-83). При этом согласно второму уравнению системы (3-48)

$$P_m^a = \frac{\rho_e^2}{\hbar} T_2 \Delta N E_m \sin(\psi - \varphi)$$

и систему (3-48) можно привести к виду

$$\left. \begin{aligned} \frac{dE_m}{dt} &= -\frac{\omega_p}{2Q_p} E_m + \frac{\omega_p \rho_e^2}{2\hbar \epsilon_a} T_2 \Delta N E_m \times \\ &\quad \times \sin^2(\psi - \varphi); \\ \frac{d\Delta N}{dt} + \frac{\Delta N - \Delta N_0}{T_1} &= -\frac{\rho_e^2}{\hbar^2} T_2 \Delta N E_m^2 \times \\ &\quad \times \sin^2(\psi - \varphi); \\ \frac{d\varphi}{dt} &= (\omega_p - \omega_n) - \frac{\omega_p \rho_e^2}{2\hbar \epsilon_a} T_2 \Delta N \cos(\psi - \varphi) \times \\ &\quad \times \sin(\psi - \varphi); \\ \frac{d\psi}{dt} &= \frac{1}{T_2} \operatorname{ctg}(\psi - \varphi). \end{aligned} \right\} \quad (4-102)$$

Скоростные или кинетические уравнения аналогично (4-94) имеют вид:

$$\left. \begin{aligned} \frac{dJ}{dt} + \frac{\omega_p}{Q_p} J &= v\sigma(\omega) \Delta N J; \\ \frac{d\Delta N}{dt} + \frac{\Delta N - \Delta N_0}{T_1} &= -2\sigma(\omega) J \Delta N, \end{aligned} \right\} \quad (4-103)$$

где

$$J = \frac{\epsilon_a v}{2\hbar \omega_{21}} E_m^2 \text{ — интенсивность излучения;}$$

$$\sigma(\omega) = \frac{T_2 \omega_{21} \rho_e^2}{\epsilon_a \hbar v [1 + (\omega - \omega_{21})^2 T^2]} \text{ — сечение радиационного перехода.}$$

Первое уравнение системы (4-103) описывает изменение интенсивности колебаний в резонаторе во времени. Здесь член $\frac{\omega_p}{Q_p} J$ характеризует уменьшение интенсивности за счет потерь в резонаторе; $v\sigma(\omega)\Delta NJ$ — увеличение интенсивности за счет индуцированного излучения. Влиянием спонтанного излучения из-за его малости пренебрегаем.

Второе уравнение (4-103) описывает изменение разности населенностей рабочих уровней во времени. Здесь член

$$\frac{\Delta N - \Delta N_0}{T_1}$$

характеризует уменьшение инверсной населенности за счет релаксационных процессов, и ее увеличение, создаваемое источником возбуждения. Член $2\sigma(\omega)J\Delta N$ определяет уменьшение инверсной населенности за счет индуцированного излучения. Множитель 2 в этом члене появляется вследствие того, что в двухуровневой системе при переходе одной частицы с верхнего уровня на нижний и излучения одного фотона разность населенностей уровней изменяется на 2.

Кинетические или скоростные уравнения (4-103) в отличие от (4-102) не определяют изменение фазы поля и поэтому не могут быть использованы для описания различных нестационарных процессов в генераторе.

При работе генератора в стационарном режиме выполняются условия

$$\frac{dE_m}{dt} = \frac{d\Delta N}{dt} = \frac{d\varphi}{dt} = \frac{d\psi}{dt} = 0,$$

и система уравнений (4-102) принимает вид:

$$\left. \begin{aligned} -\frac{\omega_p}{2Q_p} E_m + \frac{\omega_n p_e^2}{2\hbar \epsilon_a} T_2 \Delta N E_m \sin^2(\psi - \varphi) &= 0; \\ \frac{\Delta N - \Delta N_0}{T_1} &= -\frac{p_e^2}{\hbar^2} T_2 \Delta N E_m^2 \sin^2(\psi - \varphi); \\ (\omega_p - \omega_n) - \frac{\omega_p p_e^2}{2\hbar \epsilon_a} T_2 \Delta N \cos(\psi - \varphi) \sin(\psi - \varphi) &= 0; \\ \frac{1}{T_2} \operatorname{ctg}(\psi - \varphi) &= 0. \end{aligned} \right\} \quad (4-104)$$

Согласно последнему уравнению

$$\psi - \varphi = \frac{\pi}{2},$$

т. е. поляризация опережает поле по фазе на $\pi/2$ и согласно § 1-3 энергия от среды передается полю. При этом уравнения системы (4-104) будут иметь вид:

$$\left. \begin{aligned} -\frac{\omega_p}{2Q_p} + \frac{\omega_{\pi} p_e^2}{2\hbar \varepsilon_a} T_2 \Delta N &= 0; \\ \frac{\Delta N_0 - \Delta N}{T_1} &= \frac{p_e^2}{\hbar^2} T_2 \Delta N E_m^2; \\ \omega_p - \omega_{\pi} &= 0 \end{aligned} \right\} \quad (4-105)$$

отсюда в стационарном режиме работы

$$\Delta N = \frac{\hbar \varepsilon_a}{Q_p p_e^2 T_2}; \quad (4-106)$$

$$E_m^2 = \frac{\hbar Q_p}{T_1 \varepsilon_a} \left(\Delta N_0 - \frac{\hbar \varepsilon_a}{Q_p p_e^2 T_2} \right). \quad (4-107)$$

При наличии генерирования амплитуда E_m отлична от нуля, что в соответствии с (4-107) возможно при выполнении условия

$$\Delta N_0 \geq \frac{\hbar \varepsilon_a}{Q_p p_e^2 T_2}.$$

Знак равенства в этом выражении соответствует пороговой инверсной населенности. Таким образом, пороговая инверсная населенность определяется выражением

$$\Delta N_{0\text{пор}} = \frac{\hbar \varepsilon_a}{Q_p p_e^2 T_2}. \quad (4-108)$$

Подставляя (4-108) в (4-107), получаем окончательное выражение для амплитуды поля в резонаторе

$$E_m^2 = \frac{\hbar Q_p}{T_1 \varepsilon_a} (\Delta N_0 - \Delta N_{0\text{пор}}). \quad (4-109)$$

Согласно выражению (4-109) чем больше добротность резонатора и превышение инверсной населенности, создаваемой источником возбуждения, над пороговой, тем больше амплитуда поля в резонаторе.

Стационарное значение разности населенностей в резонаторе согласно (4-106) и (4-108) равно пороговому:

$$\Delta N = \Delta N_{\text{пор.}}$$

Пороговая инверсная населенность создается при пороговой мощности источника возбуждения. Казалось бы, что при мощности возбуждения больше пороговой инверсная населенность ΔN должна превышать пороговое значение $\Delta N_{\text{пор.}}$. Однако этого не происходит, так как при увеличении мощности возбуждения увеличивается число частиц в единицу времени, покидающих верхний уровень. При этом инверсная населенность поддерживается на постоянном уровне $\Delta N_{\text{пор.}}$, а энергия, запасаемая резонатором, и мощность на выходе лазера увеличиваются.

Для стационарного режима

$$\frac{dJ}{dt} = \frac{d\Delta N}{dt} = 0,$$

и кинетические уравнения (4-103) преобразуются к виду

$$\left. \begin{aligned} \frac{\omega_p}{Q_p} J &= \nu \sigma(\omega) \Delta N J; \\ \frac{\Delta N - \Delta N_0}{T_1} &= -2\sigma(\omega) J \Delta N. \end{aligned} \right\} \quad (4-110)$$

Отсюда стационарное значение инверсной населенности

$$\Delta N_{\text{стац}} = \frac{\omega_p}{Q_p \nu \sigma(\omega)}, \quad (4-111)$$

и интенсивность излучения

$$J_{\text{стац}} = \frac{1}{2\sigma(\omega)T_1} (\Delta \tilde{N}_0 - 1), \quad (4-112)$$

где $\Delta \tilde{N}_0 = \frac{\Delta N_0}{\Delta N_{\text{стац}}}$ — начальная нормированная инверсная населенность, создаваемая источником возбуждения.

Кинетические уравнения (4-103) удобно представить в нормированных величинах

$$\Delta \tilde{N} = \frac{\Delta N}{\Delta N_{\text{стац}}}, \quad \Delta \tilde{N}_0 = \frac{\Delta N_0}{\Delta N_{\text{стац}}}, \quad \tilde{J} = \frac{J}{J_{\text{стац}}},$$

т. е.

$$\left. \begin{aligned} \frac{d\tilde{J}}{dt} + \frac{\omega_p}{Q_p} \tilde{J} &= \frac{\omega_p}{Q_p} \tilde{J} \Delta\tilde{N}; \\ \frac{d\Delta\tilde{N}}{dt} + \frac{\Delta\tilde{N} - \Delta\tilde{N}_0}{T_1} &= -\frac{\Delta\tilde{N}_0 - 1}{T_1} \tilde{J} \Delta\tilde{N}; \end{aligned} \right\} \quad (4.113)$$

Здесь $\Delta\tilde{N}$ — нормированная инверсная населенность в единице объема; \tilde{J} — нормированная интенсивность излучения; T_1 — время жизни возбужденных состояний при наличии источника возбуждения.

Уравнения (4-113) справедливы, если ширина линии излучения много больше ширины резонансной кривой резонатора, что всегда выполняется для лазеров. Например, для рубинового лазера ширина линии излучения на 3 порядка больше, чем ширина резонансной кривой резонатора. Эти уравнения пригодны для случая однородно уширенной линии, что имеет место для рубинового лазера, работающего при комнатной температуре. В этих уравнениях не учитывается спонтанное излучение, которое сравнимо с индуцированным излучением лишь в начальный момент развития процесса генерирования. В стационарном режиме интенсивность индуцированного излучения на много порядков превышает спонтанное.

Запишем уравнения (4-113) для стационарного режима:

$$\begin{aligned} \frac{\omega_p}{Q_p} \tilde{J} (1 - \Delta\tilde{N}) &= 0; \\ \frac{\Delta\tilde{N} - \Delta\tilde{N}_0}{T_1} + \frac{\Delta\tilde{N}_0 - 1}{T_1} \tilde{J} \Delta\tilde{N} &= 0. \end{aligned}$$

Анализируя эти уравнения, легко убедиться, что они имеют два решения:

$$\Delta\tilde{N} = \tilde{J} = 1, \quad (4.114)$$

$$\Delta\tilde{N} = \Delta\tilde{N}_0, \quad \tilde{J} = 0. \quad (4.115)$$

Решение (4-115) соответствует отсутствию колебаний ($J=0$), поэтому устойчивым стационарным колебаниям в лазере соответствует только решение (4-114). Оба решения соответствуют стационарным состояниям. Практически реализовываться будет то состояние, которое

при данных условиях является устойчивым. Исследуем сначала на устойчивость решение (4-114) (см. П2, пп. 7, 8, 9). Зададим малые отклонения от стационарных значений:

$$\left. \begin{aligned} \tilde{J}(t) &= 1 + \delta i(t) \\ \Delta \tilde{N}(t) &= 1 + \delta n(t), \end{aligned} \right\} \quad (4-116)$$

где $\delta \ll 1$. Подставляя (4-116) в (4-113) и ограничиваясь членами первого порядка малости по δ , получаем:

$$\begin{aligned} \frac{di}{dt} &= \frac{\omega_p}{Q_p} n; \\ \frac{dn}{dt} + \frac{\Delta \tilde{N}_0}{T_1} n &= \frac{1 - \Delta \tilde{N}_0}{T_1} i. \end{aligned}$$

Характеристическое уравнение этой системы имеет вид:

$$\alpha^2 + \frac{\Delta \tilde{N}_0}{T_1} \alpha - \frac{\omega_p}{Q_p T_1} (1 - \Delta \tilde{N}_0) = 0.$$

Корни этого уравнения

$$\alpha_{1,2} = -\frac{\Delta \tilde{N}_0}{2T_1} \pm \sqrt{\left(\frac{\Delta \tilde{N}_0}{2T_1}\right)^2 + \frac{\omega_p}{Q_p T_1} (1 - \Delta \tilde{N}_0)}$$

имеют отрицательные вещественные части при

$$\Delta \tilde{N}_0 > 1; \quad (4-117)$$

следовательно, выражение (4-117) является условием существования стационарных колебаний лазера.

Аналогично можно показать, что решение (4-115), соответствующее отсутствию колебаний, устойчиво при

$$\Delta \tilde{N}_0 < 1.$$

Условие

$$\Delta \tilde{N}_0 = 1$$

является пороговым, так как при $\Delta \tilde{N}_0 > 1$ колебания существуют, а при $\Delta \tilde{N}_0 < 1$ отсутствуют.

Согласно (4-111)

$$\Delta N_{\text{пор}} = \frac{\omega_p}{Q_p \sigma(\omega) v}.$$

Роль источника возбуждения сводится к поддержанию в среде достаточно высокой начальной инверсной населенности, чтобы выполнялось условие (4-117).

Если источник возбуждения перестает работать, то населенность уровней стремится к тепловому равновесному состоянию $\Delta N_{\text{равн}}$ с постоянной времени T_1 . При переходе одной частицы с верхнего уровня на нижний и излучении кванта энергии $\hbar\omega_{21}$ разность населенностей уровней изменяется на 2. Эти переходы происходят и при работе источника возбуждения, действие которого сводится к компенсации убыли частиц, связанной с этими переходами. Плотность мощности источника возбуждения, требуемая для поддержания инверсной населенности ΔN_0 в единице объема активной среды, равна:

$$\rho_{\text{возб}} = \frac{\hbar\omega_{21}}{2} \left(\frac{\Delta N_0 - \Delta N_{\text{равн}}}{T_1} \right);$$

отсюда найдем пороговую плотность мощности возбуждения:

$$\rho_{\text{пор.возб}} = \frac{\hbar\omega_{21}}{2} \left(\frac{\Delta N_{\text{опор}} - \Delta N_{\text{равн}}}{T_1} \right). \quad (4-118)$$

Для вывода излучения из лазера зеркала делаются полупрозрачными, при этом коэффициенты отражения от зеркал Γ_1 и Γ_2 не равны единице. За счет неполного отражения от зеркал происходит уменьшение энергии, запасенной в резонаторе, которое характеризуется выражением

$$w(t) = w(0) e^{-t/\tau_3}, \quad (4-119)$$

где $w(t)$ — плотность энергии, запасенной в резонаторе; $w(0)$ — плотность энергии в момент $t=0$; τ_3 — время релаксации энергии за счет неполного отражения от зеркал.

Условие существования стационарного режима — это условие баланса амплитуд (см. § 4-2):

$$\Gamma_1 \Gamma_2 e^{2(\alpha_{\text{ус}} - \alpha_{\text{п}})l} = 1. \quad (4-120)$$

При однократном отражении электромагнитной волны от каждого из зеркал амплитуда уменьшается, но это уменьшение компенсируется усилением в активной среде при прохождении волной расстояния, равного удвоенной длине резонатора.

За время прохождения волной расстояния, равного двум длинам резонатора,

$$t = \frac{2l}{v},$$

где l — длина резонатора; v — скорость распространения электромагнитной волны; уменьшение запасенной энергии из-за неполного отражения от зеркал согласно (4-119) и (4-120) определится выражением

$$\Gamma_1^2 \Gamma_2^2 = e^{-\frac{2l}{v\tau_3}};$$

отсюда

$$\tau_3 = -\frac{l}{v \ln \Gamma_1 \Gamma_2}. \quad (4-121)$$

Мощность излучения, приходящаяся на единицу объема активной среды, согласно (4-119) равна:

$$p_{\text{изл}} = -\frac{d\omega}{dt} = \frac{\omega}{\tau_3}.$$

В стационарном режиме

$$\omega = \hbar\omega_{21}J_{\text{стац}}/v,$$

или с учетом (4-112)

$$\omega = \frac{\hbar\omega_{21}}{2\sigma(\omega)T_1v}(\Delta\tilde{N}_0 - 1),$$

и мощность излучения, отнесенная к единице объема активной среды, равна:

$$p_{\text{изл}} = \frac{\hbar\omega_{21}}{2\sigma(\omega)\tau_3T_1v}(\Delta\tilde{N}_0 - 1).$$

Регенеративный усилитель по принципиальной схеме (рис. 3-3) ничем не отличается от генератора и имеет значительно больше общего с генератором, чем с усилителем бегущей волны. В основе работы регенеративного усилителя лежит усиление в резонаторе, обеспечивающем положительную обратную связь. Условие работы в режиме регенеративного усиления определяется выражением (4-82).

Введение положительной обратной связи приводит к повышению коэффициента усиления. По способу ввода и вывода усиленного сигнала и типа используемого резонатора регенеративные усилители делятся на про-

ходные, отражательные и кольцевые. В проходных усилителях вход и выход сигнала разделены, в отражательных — совмещены и отделение выходного сигнала от входного осуществляется с помощью невзаимных систем. В кольцевых регенеративных усилителях используются резонаторы кольцевого типа, в которых излучение распространяется в одном направлении. Независимо от типа резонатора усиление сигнала регенеративным усилителем возможно при коэффициенте усиления среды, не превышающем порог самовозбуждения (4-82). Если это условие не выполняется, усилитель переходит в режим генерирования. Наличие положительной обратной связи приводит к значительному увеличению добротности резонатора, поэтому полоса пропускания регенеративного усилителя обычно бывает узкой.

4-6. ЛАЗЕРЫ НА ТВЕРДЫХ ДИЭЛЕКТРИКАХ

Так называют оптические квантовые генераторы и усилители, в которых в качестве активной среды используется кристаллический или аморфный диэлектрик с примесными активными атомами.

Особенностью лазеров на твердых диэлектриках является большая концентрация активных частиц, на несколько порядков превышающая концентрацию частиц в газовых средах. Поэтому твердые среды характеризуются большим коэффициентом усиления и с помощью их можно получать большие мощности при сравнительно небольшой длине активного элемента.

Однако твердые среды обладают меньшей оптической однородностью и в них наблюдаются большие распределенные потери на рассеяние. Это приводит к уменьшению добротности, поэтому нет смысла применять активные элементы большой длины. Для наиболее оптически однородных материалов ограничиваются длиной до 50—60 см. Оптическая неоднородность приводит к тому, что условия генерирования при данном уровне энергии возбуждения удовлетворяются не по всему сечению активного элемента, а в узких каналах. Одновременно генерируется много типов колебаний, монохроматичность получается невысокой, а расходимость луча — значительной, около 10^{-2} рад.

Структурная схема лазера на твердом диэлектрике приведена на рис. 3-3. Она включает активную среду

(кристалл рубина, стекло с неодимом и т. д.); систему возбуждения, создающую в среде инверсную населенность; открытый резонатор, обеспечивающий обратную связь; и систему селекций типов колебаний.

Активная среда лазера на твердом диэлектрике состоит из двух компонентов. Основной материал, так называемая «матрица», не участвует в процессах, связанных с квантовым усилением и генерированием. Матрица представляет собой кристаллическую или аморфную среду, содержащую в качестве примесей ионы или атомы «активатора», число которых невелико и составляет доли процента от полного числа атомов в среде.

Индукцированное излучение лазера определяется энергетическими уровнями атомов активатора. Но эти уровни в твердом диэлектрике (см. П1, п. 3) отличаются от уровней изолированных атомов. Под влиянием электрического поля, создаваемого соседними ионами решетки (внутрикристаллическое поле), уровни расщепляются, образуя полосы. Уширение линии связано и с неоднородностью кристалла, из-за чего атомы активатора оказываются в неодинаковых условиях и поэтому их уровни несколько сдвинуты относительно друг друга («неоднородное» уширение). Уширение линий происходит и из-за тепловых колебаний решетки, амплитуда которых увеличивается при увеличении температуры. Так как в качестве активатора обычно используются парамагнитные ионы, незаполненная оболочка которых обладает магнитным моментом, магнитное взаимодействие между этими ионами приводит к дополнительному расщеплению уровней.

Наличие метастабильного уровня, служащего верхним уровнем рабочего перехода, является непременным условием того, что данная среда может быть использована как активная. Такой уровень позволяет накопить на нем достаточное число активных частиц. Время жизни частиц в метастабильном состоянии должно в основном определяться излучательными, а не безызлучательными переходами (см. П1, п. 7), при которых энергия возбужденных частиц передается кристаллической решетке.

Время жизни τ на метастабильном уровне определяется выражением

$$\frac{1}{\tau} = \frac{1}{\tau_{из}} + \frac{1}{\tau_{бн}},$$

где $\tau_{из}$, $\tau_{би}$ — время излучательных и безызлучательных переходов соответственно. Необходимо, чтобы $\tau_{из} > \tau_{би}$, так как при этом легче выполняется пороговое условие генерирования и уменьшается мощность возбуждения.

Активные среды должны обладать широкими полосами поглощения, так как чем шире полоса поглощения вещества, тем большая часть мощности немонохроматического источника возбуждения используется для возбуждения частиц. Желательно, чтобы широкие линии поглощения лежали достаточно близко к метастабильному уровню, в противном случае при безызлучательных переходах из полосы поглощения на метастабильный уровень будет выделяться энергия, приводящая к нагреванию кристалла.

Активная среда должна обладать малыми потерями на частоте рабочего перехода, не связанными с переходом частиц между рабочими уровнями. Среда должна обладать высокой оптической однородностью, чтобы рассеяние на неоднородностях было минимальным. Потери в среде уменьшают добротность резонатора, увеличивают пороговую мощность возбуждения и уменьшают к. п. д. лазера.

Среда должна обладать большим квантовым выходом на частоте рабочего перехода. *Квантовый выход* определяется отношением числа излученных фотонов к числу фотонов поля возбуждения, поглощенных средой. Квантовый выход всегда меньше единицы, так как на верхнем рабочем уровне часть частиц оказывается после того, как они побывают на промежуточных уровнях. При этом часть частиц переходит на основной уровень, минуя верхний рабочий уровень. Кроме того, часть частиц может переходить с верхнего рабочего уровня без излучения. Чем меньше квантовый выход, тем менее пригодна среда для использования в лазерах.

В широко распространенном рубиновом лазере в качестве матрицы используется кристаллическая решетка Al_2O_3 (корунд), в качестве активатора ионы Cr^{3+} , замещающие в решетке корунда ионы алюминия и окрашивающие корунд в красный цвет (от розового до малинового в зависимости от концентрации ионов Cr^{3+}). Обычно применяемые в лазерах кристаллы рубина содержат около 0,05% ионов Cr^{3+} .

Рубин хорошо поддается обработке, обладает высокой механической прочностью и теплопроводностью, про-

зрачен для излучения на рабочем переходе. Коэффициент преломления рубина равен 1,76.

Рубин относится к трехуровневым системам. Схема уровней приведена на рис. 4-49. Рубин имеет две широкие полосы поглощения в зеленой и синей части спектра. Полосы поглощения выполняют роль третьего уровня. Уровень 2 состоит из двух близко расположенных подуровней, которые находятся в тепловом равновесии и населенность которых

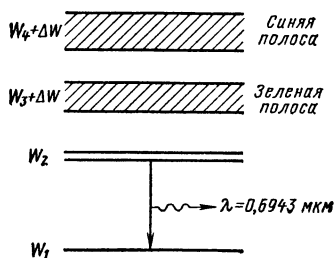


Рис 4-49. Схема энергетических уровней рубина.

подчиняется распределению Больцмана (П1, п. 4). Поэтому населенность верхнего подуровня меньше, чем нижнего. Спонтанное излучение наблюдается при переходах как с верхнего подуровня на основной уровень 1, так и с нижнего подуровня. Вынужденное излучение, если не принимаются специальные меры, происходит при переходах с нижнего подуровня на основной уровень, так как необходимая инверсная населен-

ность для этого перехода достигается раньше, чем для переходов с верхнего подуровня на основной. Уход частиц с нижнего подуровня приводит к восполнению населенности этого подуровня за счет верхнего, так как оба подуровня находятся в тепловом равновесии и верхний «подпитывает» частицами нижний. Поэтому оба подуровня рассматриваются как один уровень со статистическим весом 2 (П1, п. 1). Нижний основной энергетический уровень 1 имеет статистический вес 4.

При комнатной температуре практически все ионы Cr^{3+} находятся на нижнем уровне W_1 . С помощью немонохроматического источника возбуждения они переводятся в полосу $W_3 + \Delta W_3$ и $W_4 + \Delta W_4$.

Уровень W_2 расщеплен слабо и является метастабильным. Время жизни в полосах $W_3 + \Delta W_3$ и $W_4 + \Delta W_4$ около 10^{-8} с, на метастабильном уровне W_2 — 10^{-3} с. Частицы с уровней полос $W_3 + \Delta W_3$ и $W_4 + \Delta W_4$ переходят без излучения на метастабильный уровень и накапливаются на нем. Чем больше время жизни метастабильного уровня по сравнению с временем жизни в полосах

$W_3 + \Delta W_3$ и $W_4 + \Delta W_4$, тем меньше требуется мощность возбуждения.

Чтобы среда стала активной, необходимо

$$N_2 - \frac{g_2}{g_1} N_1 > 0.$$

Учитывая, что $N_1 + N_2 = N_0$ — плотность активных частиц, получаем:

$$N_2 > \frac{N_0}{1 + \frac{g_1}{g_2}} = \frac{N_0}{3},$$

т. е. на втором уровне должно находиться не менее трети всех активных частиц. Так как при комнатной температуре заселен практически только нижний уровень, а верхние уровни пусты, то наименьшая энергия возбуждения, которая требуется для перевода рубина в активное состояние, определяется выражением

$$W_{\text{мин}} = \hbar \omega_{13} \frac{N_0}{3}.$$

Считая, что при концентрации хрома 0,05% плотность активных частиц $N_0 \approx 1,6 \cdot 10^{19} \text{ см}^{-3}$ и $\lambda_{13} = 0,56 \text{ мкм}$, получим $W_{\text{мин}} = 2 \text{ Дж/см}^3$. Практически $W_{\text{мин}}$ оказывается в десятки раз больше, так как к. п. д. лазеров на рубине мал и лишь часть энергии возбуждения поглощается рубином.

Переход с уровня W_2 на уровень W_1 происходит или под действием внешнего излучения (индуцированное излучение), или спонтанно. Этот переход соответствует излучению в красной части видимого спектра ($\lambda = 0,6943 \text{ мкм}$).

Рубиновые образцы изготавливаются обычно в виде цилиндрических стержней, продольная ось которых составляет с оптической осью угол в 90 или 60°. Излучение, генерируемое в стержне с такой ориентацией оптической оси, линейно поляризовано, и вектор электрического поля перпендикулярен плоскости, проходящей через ось стержня и оптическую ось. Если оптическая ось и ось стержня параллельны, то излучение не поляризовано.

Рубин нашел широкое применение благодаря тому, что он генерирует излучение в видимой части спектра, может работать при комнатной температуре, имеет вы-

сокий порог разрушения. Однако кристаллы рубина обладают значительной оптической неоднородностью из-за дефектов кристаллической решетки (инородные включения, неравномерное распределение хрома, дислокации и т. п.). Из-за неравномерного распределения хрома, концентрация которого меньше в центре кристалла и увеличивается к краям, коэффициент преломления изменяется от центра к краям, что приводит к тому, что образец с плоскими параллельными торцами эквивалентен рассеивающей линзе. Дислокации и другие дефекты решетки приводят к возникновению напряжений внутри кристалла. Все это приводит к повышению расходимости лазерного луча и неоднородному распределению энергии в нем, что приводит к деформации волнового фронта.

В лазерах на стекле, активированном неодимом, матрицей является аморфное стекло. Преимуществом этих активных сред является их технологичность, возможность изготовления стержней больших размеров, сравнительная дешевизна и высокая оптическая однородность. Однако стекло не обладает высокой термостойкостью и теплопроводностью, поэтому при работе в режиме, когда интервал между импульсами излучения мал, стекло разрушается. Большое выделение тепла приводит к ухудшению свойств излучения. Неоднородное распределение энергии возбуждения по сечению стержня приводит к перепаду температур и неоднородному распределению инверсной населенности по сечению, что в свою очередь создает анизотропию, которая влияет на распределение энергии по сечению луча лазерного излучения и на его поляризацию. Влияние термических искажений существенно зависит от формы активного стержня и уменьшается при использовании стержней прямоугольного сечения.

Стекло с ионами неодима Nd^{3+} относится к четырехуровневым системам. Схема уровней приведена на рис. 3-9.

При возбуждении ионы Nd^{3+} переходят в полосу $W_4 + \Delta W_4$, откуда они без излучения переходят на метастабильный уровень (время жизни около 10^{-3} с). Индуцированные переходы происходят между уровнями W_3 и W_2 (длина волны излучения $\lambda = 1,06$ мкм). Спектр излучения гораздо шире, чем в случае рубина. Излучение неполяризовано.

Создать инверсную населенность между уровнями W_3 и W_2 гораздо легче, чем в случае трехуровневой среды, так как уровень W_2 мало заселен. Поэтому для создания инверсной населенности в четырехуровневых средах требуется на порядок меньшая мощность возбуждения, чем для создания инверсной населенности в трехуровневых системах. Уровень W_2 должен лежать как можно выше, что обеспечивает его малую населенность.

Практически пороговая мощность возбуждения больше, чем мощность, определяемая выражением (4-118), так как часть мощности возбуждения идет на компенсацию потерь, связанных с переходами частиц из полос поглощения на верхний уровень рабочего перехода. При этом активная среда нагревается.

Если нижний уровень рабочего перехода является основным, то при тепловом равновесии все активные частицы будут находиться на этом уровне и, следовательно, $\Delta N_{\text{равн}}$ будет определяться концентрацией активных частиц в единице объема N_0 .

Для рубина

$$\begin{aligned}\Delta N_{\text{опор}} &\approx 10^{18} \text{ см}^{-3}; \\ N_0 &= 2 \cdot 10^{19} \text{ см}^{-3}; \\ T_1 &\approx 4 \cdot 10^{-3} \text{ с}; \\ \hbar\omega_{21} &\approx 2,7 \cdot 10^{-19} \text{ Дж}.\end{aligned}$$

Если резонатор лазера на рубине характеризуется постоянной затухания $\omega_p/Q_p \approx 2 \cdot 10^9 \text{ с}^{-1}$, то плотность пороговой мощности возбуждения согласно (4-118) составляет 600 Вт/см^3 .

Для лазера на стекле с неодимом, работающего по четырехуровневой схеме, требуется меньшая плотность пороговой мощности возбуждения, так как нижний уровень рабочего перехода лежит на значительном расстоянии от основного уровня и при тепловом равновесии населенность этого уровня близка к нулю. При этом согласно формуле (4-118)

$$P_{\text{пор.возб}} = \frac{\hbar\omega_{21}}{2} \frac{\Delta N_{\text{опор}}}{T_1}.$$

Для стекла с неодимом

$$\begin{aligned}\Delta N_{\text{опор}} &\approx 2,2 \cdot 10^{16} \text{ см}^{-3}; \\ T_1 &\approx 10^{-4} \text{ с}; \\ \hbar\omega_{21} &\approx 1,7 \cdot 10^{-19} \text{ Дж},\end{aligned}$$

и плотность пороговой мощности при резонаторе лазера, характеризуемом постоянной затухания $\omega_p/Q_p \approx 2 \times 10^9 \text{ с}^{-1}$, составляет 19 Вт/см^2 .

В последнее время широкое применение в качестве матриц находит иттрий-алюминиевый гранат $\text{Y}_3\text{Al}_5\text{O}_{12}$, имеющий кристаллическую структуру, который сокращенно обозначается YAG. В качестве активаторов используются Cr^{3+} и Nd^{3+} . В первом случае схема рабочих уровней совпадает с трехуровневой схемой рубина, во втором случае — с четырехуровневой схемой стекла с неодимом. Соответственно излучение происходит при $\lambda = 0,6943 \text{ мкм}$ и $\lambda = 1,06 \text{ мкм}$.

Кристаллы YAG имеют низкую пороговую энергию возбуждения при комнатной температуре, высокую механическую прочность и теплопроводность. Поэтому такие кристаллы могут быть использованы в лазерах с непрерывным излучением и лазерах, работающих в режиме повторяющихся импульсов. Иттрий-алюминиевый гранат не может быть применен в мощных генераторах из-за низкого порога разрушения материала.

В настоящее время лазерный эффект получен на большом числе редкоземельных ионов во многих кристаллических и аморфных средах.

Методы возбуждения. В основе работы лазеров лежит усиление света с помощью вынужденного излучения. В равновесных системах вынужденное поглощение преобладает над вынужденным излучением. Для вынужденного усиления или генерирования необходимо создать в среде состояние с инверсной населенностью. Существуют различные методы получения инверсной населенности. Основным является метод оптического возбуждения. Он эффективен, если большая часть энергии, излучаемая источником, соответствует спектру полос поглощения активной среды. Так как для возбуждения, особенно трехуровневых систем, требуются большие интенсивности облучения, то твердотельные лазеры работают, как правило, в импульсном режиме.

Система возбуждения лазера состоит из лампы и арматуры, концентрирующей лучистый поток на активном элементе, накопителя энергии, устройства для управления режимом заряда накопителя и формирования разрядных импульсов с необходимой периодичностью. Эффективность системы определяется отношением энергии, поглощаемой активным элементом на соответствующих

квантовых переходах, к потребляемой энергии. Согласно закону Вина максимум излучения абсолютно черного тела, зависит от абсолютной температуры T и соответствует длине волны $\lambda_{\text{макс}}$, мкм, определяемой формулой

$$\lambda_{\text{макс}} = \frac{3 \cdot 10^3}{T}.$$

Чтобы $\lambda_{\text{макс}}$ соответствовала оптическому диапазону, необходима температура излучения 5000—10 000 К. Такая температура реализуется в газоразрядных лампах.

Газоразрядные лампы характеризуются высоким к. п. д. преобразования энергии. Однако только 25% из общего количества излучаемой энергии соответствует спектру полос поглощения активной среды.

Для импульсного возбуждения используются обычно импульсные ксеноновые лампы, так как ксенон обеспечивает более высокую по сравнению с другими инертными газами светоотдачу и плотность излучения, что связано с низким потенциалом ионизации этого газа и его высокой атомной массой. Газ помещается в кварцевую трубку, оба конца которой вакуумплотно соединяются с электродами. Оптимальное давление ксенона 40—80 кПа.

Спектр излучения импульсных газоразрядных ламп имеет сплошной характер. Положение максимума излучения на шкале частот зависит от напряжения, подаваемого на лампу, и при повышении напряжения смещается в сторону коротких волн.

Для возбуждения ОКГ, работающих в непрерывном режиме, применяются дуговые лампы постоянного горения. Активные среды, применяемые для работы в непрерывном режиме, имеют низкий порог возбуждения генерирования, поэтому интенсивность излучения дуговых ламп может быть значительно ниже, чем у импульсных ламп. При этом снижается требуемая плотность тока, что облегчает режим работы лампы.

Спектр дуговых ламп непрерывного действия имеет сплошной характер, на фоне которого выделяются интенсивные линии излучения атомов инертного газа. Поэтому определяющим при выборе дуговых ламп является совпадение интенсивных линий излучения инертного газа с полосами поглощения активной среды. Для возбуждения ОКГ на иттрий-алюминиевом гранате с неодимом применяют криптоновые лампы высокого давления.

Применяются также ксеноновые лампы высокого давления (около 1 МПа при комнатной температуре), ртутные лампы сверхвысокого давления (около 10 МПа) и йодные лампы накаливания. Однако при одной и той же подводимой к лампам электрической мощности эффективность криптоновых ламп выше.

Важным электрическим параметром лампы является напряжение поджига, при котором в промежутке между электродами лампы образуется плазма, обладающая определенной проводимостью, обеспечивающей развитие разряда. Для импульсных ламп остро встает вопрос о долговечности, определяемой числом вспышек, которое выдерживает лампа до выхода из строя. Долговечность зависит от отношения рабочей энергии вспышки $W_{\text{раб}}$ к предельной энергии $W_{\text{пред}}$, которую лампа может выдержать. При $W_{\text{раб}}/W_{\text{пред}}=0,3$ долговечность лампы составляет 10^4 — 10^5 вспышек.

Для питания импульсных ламп применяют накопители энергии, обеспечивающие высокую импульсную мощность. Обычно накопитель представляет собой импульсный конденсатор или систему импульсных конденсаторов. Энергия $W_э$, запасаемая конденсатором, определяется емкостью конденсатора C и напряжением U :

$$W_э = \frac{CU^2}{2}.$$

Напряжение U целесообразно выбирать равным номинальному напряжению импульсной лампы. При этом емкость, а следовательно, и размеры конденсатора будут минимальными.

Энергия, излученная лампой, концентрируется на активных образцах с помощью отражателей, имеющих форму эллиптических цилиндров, вдоль фокусных линий которого расположены образец и лампа. Как известно, лучи, исходящие из одного фокуса эллипса, собираются в другом. Поэтому излученная лампой энергия концентрируется на активном элементе (рис. 4-50).

С целью увеличения энергии возбуждения используют отражатель, образованный несколькими эллиптическими цилиндрами (рис. 4-51). Образец помещается в общем фокусе, лампы — в других фокусах.

Оптическое возбуждение можно осуществить за счет света, выделяемого при взрыве тонких проволок при пропускании через них электрического тока большой си-

лы. Спектральный состав излучения зависит от материала проволоки и условий взрыва.

Для возбуждения возможно использовать солнечное излучение, концентрируемое на активном элементе с помощью зеркальной оптической системы.

Для оптического возбуждения вместо ламп можно использовать излучение полупроводниковых диодов (например, из арсенида галлия). Спектр излучения дио-

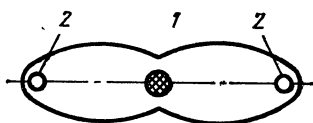


Рис. 4-50. Двухэллипсный отражатель.

1 — активный элемент; 2 — газоразрядные лампы.

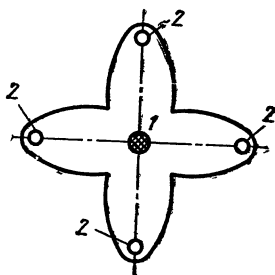


Рис. 4-51. Многоэллипсный отражатель.

1 — активный элемент; 2 — газоразрядные лампы.

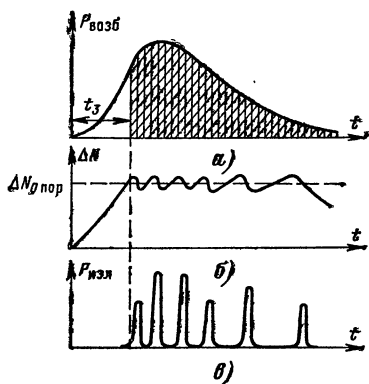


Рис. 4-52. Режим свободного генерирования.

а — импульс возбуждения; б — изменение инверсной населенности во времени; в — импульсы излучения (пикки).

дов достаточно узок, а к. п. д. высокий. В настоящее время ведется работа по исследованию возбуждения с помощью полупроводниковых диодов и по мере их совершенствования они смогут конкурировать с газоразрядными лампами.

Возможны также химические методы оптического возбуждения, в которых для возбуждения активной среды используется свет, выделяемый при химической реакции.

Режим свободного генерирования. Обычно лазеры на твердых диэлектриках (рубин, стекло с неодимом) работают в импульсном режиме, что определяется импульсным режимом работы источника возбуждения.

Большую мощность источника, определяемую высоким значением пороговой мощности возбуждения, возможно обеспечить лишь в импульсе. Длительность импульса возбуждения составляет 0,1—1 мс. Форма импульса имеет вид, приведенный на рис. 4-52, а.

Для большинства лазеров на твердых диэлектриках, работающих в импульсном режиме, стационарный режим за время импульса не достигается.

Под действием возбуждения населенность верхнего уровня увеличивается. Но вначале мощность импульса возбуждения недостаточна, чтобы создать инверсную населенность, превышающую пороговую. По мере увеличения мощности возбуждения инверсная населенность возрастает (рис. 4-52, б), и, когда она превышает пороговую величину $\Delta N_{\text{пор}}$, начинается генерирование. Начало генерирования запаздывает относительно момента включения источника возбуждения на время $t_3 \approx 10$ —100 мкс.

Индукцированное излучение при генерировании возникает вследствие усиления слабого спонтанного излучения. В начале генерирования энергия индуцированного излучения мала и мало число переходов сверху вниз, источник возбуждения успевает компенсировать эти переходы и населенность верхнего уровня возрастает. Но постепенно индуцированное излучение возрастает, число переходов сверху вниз увеличивается, и населенность становится меньше пороговой, генерирование прекращается — формируется первый импульс излучения.

При отсутствии генерирования поле в резонаторе уменьшается, уменьшается число индуцированных переходов и источник возбуждения восстанавливает инверсную населенность, необходимую для генерирования — формируется следующий импульс излучения и т. д. (рис. 4-52, в). В результате излучение лазера имеет вид последовательности нерегулярных по амплитуде и временному положению «пиков», длительность которых составляет доли микросекунд, временной интервал между ними — микросекунды. Этот режим работы называется *режимом свободного генерирования*.

Как видно из рис. 4-52, б, во время генерирования «пиков» разность населенностей уровней ΔN колеблется около стационарного значения и можно приближенно считать, что в течение всего импульса излучения нормированная инверсная населенность $\Delta \tilde{N} \approx 1$. В этом при-

ближении, деля первое уравнение системы (4-113) на второе и исключая время, получаем:

$$\frac{d\tilde{J}}{d\Delta\tilde{N}} = \frac{T_{1\omega p} (1 - \Delta\tilde{N}) \tilde{J}}{Q_p (1 - \Delta\tilde{N}_0) (1 - \tilde{J})}. \quad (4-122)$$

Приравнивая производную $\frac{d\tilde{J}}{d\Delta\tilde{N}}$ постоянной величине (см. П2, пп. 10—12)

$$\frac{d\tilde{J}}{d\Delta\tilde{N}} = C$$

и придавая ей значения $C_1, C_2, C_3 \dots$, строим изоклины.

Значению $\frac{d\tilde{J}}{d\Delta\tilde{N}} = 0$ соответствует изоклина $\Delta\tilde{N} = 1$, значению

$\frac{d\tilde{J}}{d\Delta\tilde{N}} = \infty$ — изоклина $\tilde{J} = 1$.

В общем случае уравнение изоклины имеет вид:

$$\tilde{J} = \frac{1}{1 + \frac{T_{1\omega p}}{CQ_p} (1 - \Delta\tilde{N})}.$$

Фазовую плоскость $(\tilde{J}, \Delta\tilde{N})$ можно разделить на четыре области, границы которых определяются прямыми $\Delta\tilde{N} = 1$ и $\tilde{J} = 1$ (рис. 4-53). Производная $\frac{d\tilde{J}}{d\Delta\tilde{N}}$ в областях *I* и *II* положительна, в областях *III* и *IV* — отрицательна. Стационарный режим соответствует точке $\tilde{J} = \Delta\tilde{N} = 1$, начальная точка генерирования $\tilde{J} = 0, \Delta\tilde{N} = 1$. Вычерчивая из этой точки интегральную кривую, получаем спираль, свертывающуюся в точку $\tilde{J} = \Delta\tilde{N} = 1$.

Скручивающаяся спираль пересекает все четыре области. Направление возрастания времени определяется обходом против часовой стрелки. В области *I* происходит нарастание интенсивности от $\tilde{J} = 0$ до $\tilde{J} = 1$. При этом $\Delta\tilde{N}$ изменяется от значения, равного 1, до максимального значения $\Delta\tilde{N}_+$, незначительно превышающего 1. В области *II* $\left(\frac{d\tilde{J}}{dt} > 0\right)$ продолжается увеличение интенсивно-

сти от $\tilde{J}=1$ до $\tilde{J}=\tilde{J}_{\text{макс}}$ при изменении $\Delta\tilde{N}$ от $\Delta\tilde{N}_+$ до $\Delta\tilde{N}=1$. В области III $\left(\frac{d\tilde{J}}{d\Delta\tilde{N}} < 0\right)$ интенсивность убывает от значения $\tilde{J}_{\text{макс}}$ до значения $\tilde{J}=1$, $\Delta\tilde{N}$ от $\Delta\tilde{N}=1$ — до минимального значения $\Delta\tilde{N}_-$, незначительно отличающегося от 1. В области IV уменьшение интенсивности продолжается. При этом $\Delta\tilde{N}$ возрастает и когда достигает значения $\Delta\tilde{N}=1$, восстанавливаются усло-

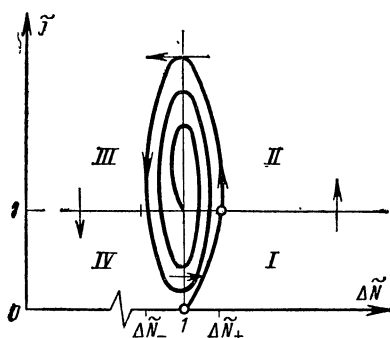


Рис. 4-53. Фазовая плоскость \tilde{J} , $\Delta\tilde{N}$.

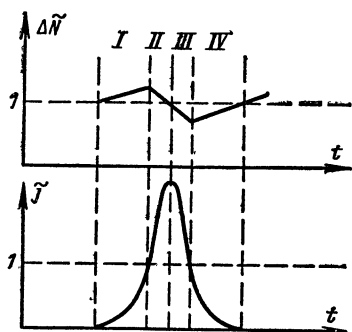


Рис. 4-54. Цикл формирования пучка.

вия, имевшие место в области I, и цикл повторяется. На рис. 4-54 приведен полный цикл изменения нормированной населенности и интенсивности во времени. Если рассмотреть изменение интенсивности за несколько циклов, то получим последовательность «пиков» уменьшающейся высоты при приближении к стационарному значению.

Интегральную кривую можно получить и из уравнения (4-122). Разделяя переменные и интегрируя, получаем:

$$\frac{T_1\omega_p}{2Q_p(\Delta\tilde{N}_0 - 1)} (\Delta\tilde{N} - 1)^2 = \ln \frac{\tilde{J}}{\tilde{J}_0} - (\tilde{J} - \tilde{J}_0), \quad (4-123)$$

где \tilde{J}_0 — начальное значение нормированной интенсивности при $\Delta\tilde{N}=1$, соответствующее спонтанному излучению.

Максимальное значение интенсивности в пичке $\tilde{J}_{\text{макс}} \gg \tilde{J}_0$ наблюдается при $\Delta\tilde{N}=1$ и согласно (4-123) определяется выражением

$$\tilde{J}_{\text{макс}} = \ln \frac{\tilde{J}_{\text{макс}}}{\tilde{J}_0}. \quad (4-124)$$

Для одномодового режима $\tilde{J}_0 = 10^{-10}$ и $\tilde{J}_{\text{макс}} = 26$, т. е. интенсивность в пичке в 26 раз превышает интенсивность в стационарном режиме. Изменение \tilde{J}_0 незначительно влияет на интенсивность в пичке. Например, увеличение \tilde{J}_0 в 10^5 раз уменьшает интенсивность в пичке $\tilde{J}_{\text{макс}}$ только вдвое.

Инверсная населенность в течение периода формирования пичка изменяется от $\Delta\tilde{N}_-$ до $\Delta\tilde{N}_+$. При этих значениях инверсной населенности $\tilde{J}=1$ и, учитывая, что $\tilde{J}_0 \ll 1$, согласно (4-123) получим:

$$\Delta\tilde{N}_{\pm} = 1 \pm \sqrt{\frac{2Q_p (\Delta\tilde{N}_0 - 1) (-1 - \ln \tilde{J}_0)}{\omega_p T_1}}.$$

Вычисление этих величин показывает, что они отличаются от единицы в третьем знаке, т. е. менее чем на 1%.

Плотность энергии, излучаемая единицей объема активной среды в течение времени длительности пичка, определяется выражением

$$\omega = \hbar \omega_{21} (\Delta\tilde{N}_+ - \Delta\tilde{N}_-) \Delta\tilde{N}_{\text{стац}},$$

где $\Delta N_{\text{стац}}$ можно найти с помощью формулы (4-111). Так как примерно половина энергии идет на компенсацию потерь в резонаторе, то энергия излучения в пичке приблизительно равна

$$\omega_{\text{изл}} \approx \frac{\hbar \omega_{21}}{2} (\Delta\tilde{N}_+ - \Delta\tilde{N}_-) \Delta\tilde{N}_{\text{стац}}. \quad (4-125)$$

Для рубинового лазера, работающего в режиме свободного генерирования, энергия в пичке, полученная с 1 см^3 активной среды, составляет около 10^{-3} Дж, мощность — около 1 кВт.

Режим гигантских импульсов или режим модуляции добротности обеспечивает импульсы длительностью 10—100 нс и пиковую мощность примерно 1 ГВт.

Методы модуляции добротности. Для получения импульсов большой мощности добротность резонатора тем или иным способом изменяется за время меньшее, чем время релаксации возбужденных уровней. Этот метод применяется для таких активных сред, как рубин и стекло с неодимом.

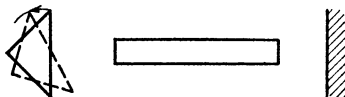
Чем больше число частиц активатора на возбужденном уровне W_2 , тем быстрее они переходят на уровень W_1 , тем короче длительность импульса излучения и тем больше мощность излучения. Под действием лампы-вспышки ионы активатора легко переходят в возбужденное состояние на уровень W_2 . Это происходит до тех пор, пока не возникнет инверсия населенностей, т. е. превышение числа возбужденных ионов над невозбужденными. При этом возникает генерирование и возбужденные частицы интенсивно переходят вниз. С таким интенсивным переходом лампа-вспышка справиться не может. Подсчитано, что если не половина, а большая часть ионов Cr^{3+} в рубине оказалась бы на уровне W_2 , то генерирование могло бы развиваться за время 10^{-8} с, что повысило бы мощность в импульсе до нескольких МВт. Перекачку всех ионов Cr^{3+} на возбужденный уровень W_2 можно осуществить с помощью лампы-вспышки, если на время выключить обратную связь, т. е. отключить зеркала, устранив тем самым возможность генерирования. За это время нужно перевести ионы на уровень W_2 , включить зеркала (добротность) и получить мощный и короткий импульс. Время отключения не должно быть больше времени жизни частиц на метастабильном уровне W_2 ($2 \cdot 10^{-3}$ с).

Основные требования, предъявляемые к модуляторам добротности — это быстрота включения добротности и малые потери в открытом состоянии.

Механические методы наиболее просты и дешевы. Широкое распространение получил модулятор с вращающейся призмой. В любом положении, кроме одного-единственного, обратной связи нет; лучи, отражаясь от граней призмы, выходят из системы, генерирования нет (рис. 4-55). В это время происходит возбуждение рубина, переход частиц с уровня W_1 в полосу W_3 и затем на метастабильный уровень W_2 . И только в единственном

положении лучи, отражаясь от грани призмы, попадают в активный стержень. Призма продолжает вращаться, но генерирование развивается столь быстро, что успевает закончиться, пока призма существенно не изменит своего положения. Оптимальным положением призмы,

Рис. 4-55. Модуляция добротности резонатора путем вращения призмы полного отражения.



обеспечивающем максимальную добротность резонатора и максимальную выходную мощность, является такое, при котором вершина призмы делит диаметр стержня пополам. Смещение вершины призмы на 10% диаметра активного стержня уменьшает выходную мощность на 20%.

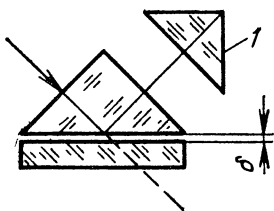


Рис. 4-56. Контактный затвор (1—призма).

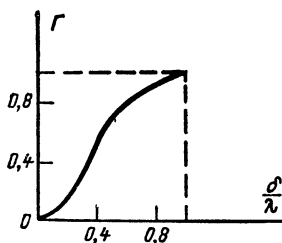


Рис. 4-57. Зависимость коэффициента отражения от толщины воздушного зазора в контактном затворе.

Модуляцию добротности можно осуществить и с помощью колебаний одного из зеркал резонатора около некоторого положения на частоте собственного механического резонанса под действием переменного магнитного поля.

Для управления добротностью резонатора можно применить контактный затвор (рис. 4-56), использующий зависимость полного внутреннего отражения от размера зазора между двумя диэлектриками. В зависимости от толщины зазора δ коэффициент отражения изменяется от 0 до 1 (рис. 4-57). Изменение толщины зазора осуществляется перемещением диэлектрической

пластины с помощью пьезокерамики, изменяющей свою толщину под действием приложенного напряжения.

Даже при мгновенном срабатывании затвора длительность импульса излучения не может быть меньше 10^{-8} — 10^{-9} с, но если обратная связь включается не мгновенно, то длительность импульса определяется уже временем включения обратной связи.

Электрооптические методы обеспечивают более быстрое включение обратной связи. Электрооптические затворы основаны на электрооптических эффектах в твердых и жидких средах и состоят из двух скрещенных поляризаторов, между которыми расположен электрооптический элемент, управляемый электрическим полем (рис. 4-58). Для

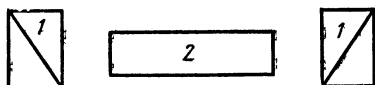


Рис. 4-58. Схема электрооптического затвора.

1 — поляризаторы; 2 — электрооптический элемент.

изготовления электрооптических элементов применяют кристаллы дигидрофосфата калия (КДР), дигидрофосфата аммония (АДР) и другие, а также жидкости — нитробензол и сероуглерод. Принцип действия обычно применяемых электрооптических затворов основан на эффекте Поккельса или эффекте Керра.

Линейный эффект Поккельса наблюдается только в твердых средах. В затворах, работающих на этом эффекте, преимущественно используются кристаллы КДР. Сущность эффекта Поккельса состоит в следующем.

Если направление распространения волны и управляющего электрического поля совпадают, то в кристалле возникают две ортогонально поляризованные волны, распространяющиеся с разными скоростями. При распространении эти волны приобретают разность фаз, которая зависит от длины кристалла, его характеристик и управляющего электрического поля или напряжения, приложенного к торцам кристалла,

$$\Delta\varphi = 2\pi B l E_0, \quad (4-126)$$

где B — постоянная, см/В; l — длина кристалла; E_0 — приложенное постоянное поле.

При этом плоскость поляризации суммарного электромагнитного поля поворачивается. Этот поворот в за-

творах составляет обычно 90° , что соответствует сдвигу фаз между ортогонально поляризованными волнами 180° (рис. 4-59). Формулу (4-126) удобно представить в виде

$$\Delta\varphi = \pi \frac{U}{U_{\lambda/2}}, \quad (4-127)$$

где $U_{\lambda/2}$ — полуволновое напряжение, соответствующее сдвигу фаз $\Delta\varphi$, зависящее от параметров кристалла.

Квадратичный эффект Керра наблюдается в кристаллах группы перовскитов и жидкостях — нитробензоле и сероуглероде. При отсутствии электрического поля материал представляет собой изотропную среду. При на-

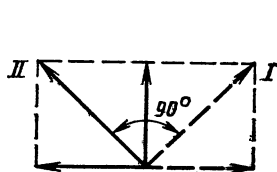


Рис. 4-59. Поворот плоскости поляризации в электрооптическом кристалле под действием полуволнового напряжения (положение I — напряжение отсутствует, положение II — напряжение $U = U_{\lambda/2}$).

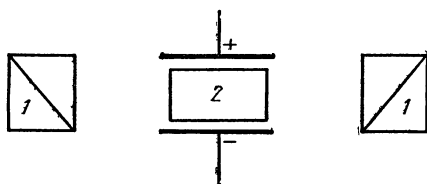


Рис. 4-60. Схема электрооптического затвора на основе ячейки Керра.

1 — поляризаторы; 2 — ячейка Керра.

ложении поля, перпендикулярного распространяющемуся излучению, среда приобретает анизотропные свойства, в результате чего возникает двойное лучепреломление. При этом диэлектрическая проницаемость среды для обыкновенной и необыкновенной волны отличаются и эти волны, пройдя расстояние l , приобретают разность фаз

$$\Delta\varphi = 2\pi B F_0^2 l = \pi \frac{U}{U_{\lambda/2}}.$$

Использование затворов на эффекте Керра предпочтительнее из-за квадратичной зависимости сдвига фаз от поля [ср. с формулой (4-126)]. Схема затвора приведена на рис. 4-60. Электрооптический элемент представляет собой ячейку Керра — кювету, заполненную нитробензолом и помещенную между обкладками конденсатора. До и после ячейки помещают поляризаторы,

пропускающие световые лучи, поляризация которых взаимно перпендикулярна и составляет угол в 45° с направлением постоянного поля конденсатора. При отсутствии электрического поля после поляризатора 1 получается линейно-поляризованный свет, а поляризатор 2 его не пропускает. При наложении поля среда в ячейке Керра становится анизотропной и в результате двойного лучепреломления излучение, проходящее через ячейку, будет иметь эллиптическую поляризацию и часть излучения будет проходить через оба поляризатора.

Управляющее напряжение имеет форму прямоугольных электрических импульсов. При амплитуде импульса, равной полуволновому напряжению, затвор полностью открывается, при нулевом напряжении — закрыт.

Если излучение активного элемента поляризовано (например, при использовании рубина с 90° -ной ориентацией), то схему, приведенную на рис. 4-58 или 4-60, можно упростить, удалив первый поляризатор.

Электровзрываемые пленочные затворы представляют собой тонкую металлическую пленку, которая мгновенно испаряется при пропускании через нее мощного электрического импульса. Это затвор одноразового действия. Обычно такие затворы представляют собой лавсановую пленку-подложку, на которую наносится алюминиевое покрытие. Эта пленка наматывается на две катушки, которые после каждого срабатывания затвора перематываются для обновления рабочей зоны пленки.

Фототропные затворы основаны на использовании сред, прозрачность которых возрастает с увеличением интенсивности света, проходящего через них. Это прозрачные среды, в которые введены резонансно-поглощающие вещества. Просветление затворов связано с насыщением поглощения на резонансной длине волны, совпадающей с длиной волны излучения активного вещества.

Для управления добротностью рубиновых лазеров применяют растворы фталоцианина лития, натрия, калия и других металлов в органических растворах. Могут быть приготовлены как жидкие, так и твердые растворы. Для управления добротностью лазеров на стекле с неодимом используют различные типы полиметиновых красителей.

Энергетические уровни органического красителя должны совпадать с рабочими уровнями активного веще-

ства. Тогда излучение активного вещества поглощается в пассивном затворе тем сильнее, чем больше разность числа частиц между нижним и верхним уровнем. Пассивный затвор работает как поглотитель, и добротность мала. Однако поглощение приводит к увеличению числа частиц на верхнем уровне, разность населенностей между верхним и нижним уровнем уменьшается, затвор просветляется. Полностью затвор открывается тогда, когда населенности уровней сравниваются и поглощение прекращается (добротность полностью включена). После прекращения импульса излучения затвор восстанавливает свои поглощающие свойства.

Наибольшее применение нашли жидкостные и пленочные фототропные затворы. Жидкостные фототропные затворы представляют собой стеклянные или кварцевые кюветы, заполненные органическим красителем. Пленочные затворы из-за разрушения фототропной пленки излучением лазера требуют замены рабочей зоны пленки после каждого импульса излучения. Такая замена обеспечивается вращающимся механизмом с двумя катушками, между которыми перематывается фототропная пленка.

Анализ режима модуляции добротности. Уже в начальный момент возникновения импульса излучения $\tilde{J} \gg 1$. В этом приближении, деля первое уравнение системы (4-113) на второе и исключая время, получаем:

$$\frac{d\tilde{J}}{d\Delta\tilde{N}} = - \frac{T_1\omega_p}{Q_p(\Delta\tilde{N}_0 - 1)} \frac{\Delta\tilde{N} - 1}{\Delta\tilde{N}}.$$

Разделяя переменные и интегрируя с учетом начальных условий $\tilde{J}=0$ при $\Delta\tilde{N}=\Delta\tilde{N}_0$, получаем:

$$\tilde{J} = \frac{T_1\omega_p}{Q_p(\Delta\tilde{N}_0 - 1)} \left(\Delta\tilde{N}_0 + \ln \frac{\Delta\tilde{N}}{\Delta\tilde{N}_0} - \Delta\tilde{N} \right). \quad (4-128)$$

Максимальная интенсивность получается при $\Delta\tilde{N}=1$ и согласно выражению (4-128)

$$\tilde{J}_{\text{макс}} = \frac{T_1\omega_p}{Q_p(\Delta\tilde{N}_0 - 1)} (\Delta\tilde{N}_0 - \ln \Delta\tilde{N}_0 - 1) \quad (4-129)$$

Зависимость $\tilde{J}=f(\Delta\tilde{N})$ для различных $\Delta\tilde{N}_0$ приведена на рис. 4-61. При $\Delta\tilde{N}_0=2$ максимальная нормирован-

ная интенсивность $\tilde{J}=0,62 \cdot 10^6$, т. е. более чем в $6 \cdot 10^5$ раз превышает интенсивность в стационарном режиме. Мощность в импульсе составляет около 10 МВт при объеме рубиновой активной среды 1 см³.

Согласно формуле (4-125) плотность энергии излучения примерно равна 0,1 Дж/см³ и длительность импульса составляет около 10^{-8} с.

В регенеративных усилителях с модуляцией добротности затвор включается в момент начала усиления.

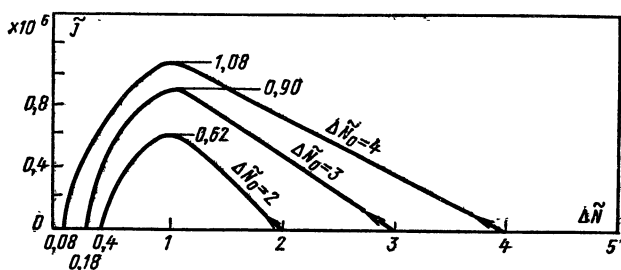


Рис. 4-61. Зависимость нормированной интенсивности от нормированной инверсной населенности в случае модуляции добротности и $T_1\omega_p/Q_p=2 \cdot 10^6$. Стрелки указывают направление возрастания времени.

При этом усиление происходит при инверсных населенностях, превышающих порог самовозбуждения, но время усиления должно быть меньше, чем время развития процесса генерирования.

Преобразование частоты излучения лазера. Лазеры на твердых диэлектриках работают на ограниченном числе рабочих волн (0,6943 и 1,06 мкм), которое можно увеличить с помощью преобразования частот в нелинейных кристаллах, например КДР. При этом необходимо, чтобы выполнялось условие волнового синхронизма (см. § 2-1) или закон сохранения энергии и импульса.

Действительно, при преобразовании волны частоты ω_1 в волну второй гармоники

$$\omega_2 = 2\omega_1;$$

закон сохранения энергии получим, умножая обе части этого равенства на \hbar (см. П1, п. 8)

$$\hbar\omega_2 = 2\hbar\omega_1;$$

отсюда следует, что на образование каждого кванта $\hbar\omega_2$ тратится два кванта $\hbar\omega_1$.

Условие волнового синхронизма для второй гармоники имеет вид:

$$2k_1(\omega_1) = k_2(2\omega_1).$$

Умножая обе части этого равенства на \hbar (см. П1, п. 8), получаем закон сохранения импульса квантов при образовании второй гармоники

$$2p_1 = p_2.$$

При падении на нелинейный кристалл излучения двух лазеров с частотами ω_1 и ω_2 в кристалле возникает волна частоты ω_3 с волновым вектором k_3 . Мощность этой волны максимальна при условии выполнения закона сохранения энергии и импульса для всех трех волн, т. е.

$$\omega_1 + \omega_{2,3} = \omega_{3,2};$$

$$k_1 + k_{2,3} = k_{3,2}$$

для случая сложения волн в нелинейном кристалле, и

$$\omega_1 - \omega_{2,3} = \omega_{3,2};$$

$$k_1 - k_{2,3} = k_{3,2}$$

для случая вычитания волн.

Возможна плавная перестройка частоты. Кристалл КДР помещается в резонатор, собственные частоты которого совпадают с ω_2 и ω_3 и могут плавно перестраиваться. Через кристалл пропускается мощное излучение лазера на частоте ω_1 . При этом за счет собственного спонтанного излучения в нелинейном кристалле возбуждаются слабые волны на частотах ω_2 и ω_3 . Если для этих волн выполняется условие волнового синхронизма, что достигается поворотом кристалла в резонаторе, то при фиксированной частоте излучения лазера ω_1 возможно получить излучение на плавно перестраиваемых частотах ω_2 и ω_3 .

Лазеры на жидких диэлектриках (жидкостные лазеры). Активной средой в таких лазерах служат жидкие диэлектрики. Плотность активных частиц в таких случаях достаточно велика и они обладают высокой оптической однородностью. Использование жидких диэлектриков позволяет иметь активный элемент любых разме-

ров и формы. Проблема охлаждения при работе в непрерывном режиме легко решается путем циркуляции жидкого диэлектрика.

Лазеры на жидких диэлектриках объединяют в себе преимущества газовых лазеров (высокая когерентность, расходимость луча 10^{-3} рад) и лазеров на твердых диэлектриках (большая мощность излучения).

Весьма перспективны лазеры на неорганических красителях с растворенным в них неодимом. Структура энергетических уровней сходна со структурой уровней для стекла с неодимом. Спектр излучения такого лазера гораздо уже, т. е. временная когерентность выше, чем у лазера на стекле с неодимом. Энергетические характеристики такие же, как у лазера на стекле с неодимом.

Лазеры на органических красителях могут генерировать на любых частотах в видимой и инфракрасной частях спектра. Частота их может плавно перестраиваться в широких пределах. Эти лазеры работают в импульсном (одиночные и периодические импульсы) и непрерывном режимах.

Для возбуждения лазеров на жидких диэлектриках используется когерентное излучение газовых лазеров и лазеров на твердых диэлектриках, а также некогерентное излучение специальных импульсных ламп. Если для возбуждения используется импульсная лампа, то она вместе с кюветой, в которой находится рабочий раствор, помещается в отражатель. Если для возбуждения применяется лазер, то его излучение пропускается через кювету либо вдоль оси резонатора, либо в поперечном направлении.

Лазеры с некогерентным возбуждением обеспечивают более высокие к.п.д. (около 1%) и более высокие уровни энергии излучения (около 10 Дж). Это определяется возможностью применения источников возбуждения с высокими энергетическими параметрами и отсутствием многоступенчатости в процессе преобразования энергии возбуждения, характерной для когерентного возбуждения: излучение импульсных ламп превращается в излучение лазера на твердом диэлектрике, используемого для возбуждения лазера на жидком диэлектрике, а затем в его излучение.

Наиболее часто применяется краситель родамин 6Ж, который по своим энергетическим параметрам близок к рубину и стеклу с неодимом. Раствор красителя поме-

щается в кварцевую кювету с двойными стенками, предусматривающими охлаждение, и окнами, наклоненными под углом Брюстера. Кюветы, заполненные различными красителями, помещаются в барабан, поворот которого обеспечивает быструю смену активных растворов внутри резонатора и перекрытие широкой спектральной полосы излучения лазера. Спектры поглощения используемых растворов должны совпадать с длинами волн излучения возбуждающих лазеров или их гармониками. Система возбуждения представляет генератор на рубине или стекле с неодимом с модуляцией добротности. Если для возбуждения используется вторая гармоника, то для удвоения частоты применяются нелинейный кристалл КДР и фильтр, поглощающий излучение основной частоты и пропускающий вторую гармонику.

4.7. ГАЗОВЫЕ ЛАЗЕРЫ

Активной средой в таких лазерах является газ или смесь газов. Энергетический спектр газа отличается от спектра твердого тела тем, что он весьма точно соответствует энергетическим уровням отдельных атомов и молекул, что позволяет довольно точно предсказать возможную схему переходов. Так как взаимодействие между частицами газовой среды значительно меньше, чем в твердых средах, ширина линии излучения узкая. Излучение газовых лазеров характеризуется высокой монохроматичностью, время когерентности примерно 10^{-3} с. При соответствующем выборе активной среды излучение может быть получено в любой части спектра, от ультрафиолетовой до субмиллиметровой области.

Газы отличаются высокой оптической однородностью, потери на рассеяние и дифракцию малы. Расходимость излучения характеризуется углами около 10^{-3} рад, т. е. близка к дифракционному пределу.

Однако плотность газа значительно меньше плотности твердых тел, поэтому нельзя получить такую большую плотность активных частиц и мощность излучения с единицы объема вещества, как в лазерах на твердых диэлектриках. Условие генерирования (4-81) для среды с малой плотностью активных частиц может быть выполнено за счет увеличения протяженности. Поэтому газовые лазеры имеют большие размеры, чем лазеры на твердых диэлектриках.

Рабочий газ помещается в кварцевую или стеклянную трубку. Чтобы уменьшить потери на торцах трубки, торцевые пластины помещаются не перпендикулярно оси трубки, а так, чтобы нормаль к торцевой пластине составляла с осью трубки угол Брюстера (рис. 4-62). Спонтанное излучение не поляризовано и его можно представить в виде суммы двух линейных поляризаций — горизонтальной и вертикальной, одинаковых по амплитуде с случайно изменяющимися фазами. Доброт-

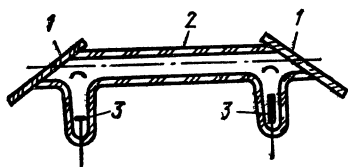


Рис. 4-62. Газоразрядная трубка газового лазера.

1 — торцевые пластины, наклоненные к оси трубки под углом Брюстера; 2 — стенки трубки; 3 — электроды.

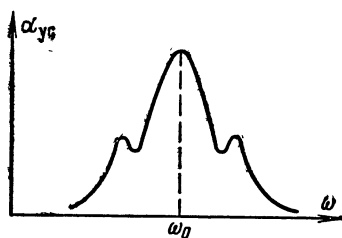


Рис. 4-63. Уменьшение коэффициента усиления вблизи частоты генерирования при неоднородно уширенной линии.

ность резонатора будет больше для вертикальной поляризации, индуцированное излучение повторяет поляризацию индуцирующего и в конце концов устанавливается излучение вертикальной поляризации.

Трубка помещается в резонатор, образованный плоскими, вогнутыми или выпуклыми зеркалами. Наиболее часто применяют конфокальный резонатор, обладающий малыми потерями, так как поле в нем сконцентрировано у оси (см. § 4-2). Конфокальный резонатор менее критичен к настройке.

Газовые лазеры классифицируются по виду активного вещества на три группы: газовые атомные, ионные и молекулярные газовые лазеры.

Газовые лазеры отличаются от лазеров на твердых диэлектриках формой линии излучения. Активные частицы в газовом лазере имеют различные частоты переходов вследствие своего движения. Если частица движется в направлении распространения волны со скоростью v_i , то частота перехода определится выражением

$$\omega_i = \omega_0 \left(1 + \frac{v_i}{c} \right),$$

где ω_0 — частота перехода неподвижной частицы; c — скорость света. При противоположном направлении движения знак «+» в этой формуле заменяется на «-». В результате такого смещения линия излучения превращается в совокупность лоренцевых линий, имеющих различные центральные частоты ω_i (см. П1, п. 20). Такая линия называется неоднородно уширенной линией. В газах она имеет гауссову форму. Зависимость инверсной населенности от частоты определяется формой линии излучения.

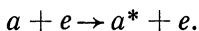
При однородном уширении линии инверсная населенность и связанный с ней коэффициент квантового усиления при взаимодействии среды с распространяющимся полем уменьшаются по всей ширине линии. При неоднородном уширении линии распространяющееся поле частоты ω будет взаимодействовать с частицами, частота квантового перехода которых с учетом доплеровского сдвига совпадает с частотой поля. В результате этого взаимодействия инверсная населенность и усиление будут уменьшаться лишь в пределах лоренцевой линии, соответствующей частоте ω . Этот эффект называется эффектом прожигания «дырки». Так как поле в резонаторе можно представить как сумму двух волн, бегущих навстречу друг другу, то на неоднородно уширенной гауссовой линии появятся две симметричные «дырки» (рис. 4-63). Если частота поля совпадает с частотой ω_0 , то провал инверсной населенности и усиления возникнет в максимуме неоднородной линии.

Инверсная населенность в газовых средах может быть получена с помощью электрических разрядов, химических реакций и оптического возбуждения.

При электрическом разряде возбуждение атомов или молекул газа происходит за счет соударений с быстрыми свободными электронами, образующимися при разряде. Давление в газоразрядных лазерах выбирается от единиц до нескольких сотен паскалей. При меньших давлениях уменьшается число соударений электронов с частицами, и возбуждение частиц протекает недостаточно интенсивно, при больших давлениях число столкновений увеличивается, благодаря чему электроны не успевают приобрести достаточную скорость и соударения становятся мало эффективными, частицы не возбуждаются.

Столкновения первого рода используются для возбуждения Ne, Ar, Kr, Xe и других однокомпонентных

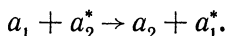
сред. Процесс заключается в неупругом столкновении между свободными электронами (e) и атомами газовой смеси (a), находящимися в основном невозбужденном состоянии, и определяется формулой



Здесь a^* — атом в возбужденном состоянии. Этот механизм используется не только для возбуждения нейтральных атомов, но и для возбуждения однократно ионизированных атомов, а также паров металлов (Pb, Cu, Au, Ag и др.).

Столкновения второго рода (неупругие) используются для возбуждения двухкомпонентной газовой смеси, в которой энергия метастабильного уровня атомов одного газа (a) близка к энергии возбужденного уровня атомов другого газа (a^*). При этом наблюдается эффективная передача энергии.

Процесс определяется формулой



Эффективность тем больше, чем ближе энергии рассматриваемых уровней. Столкновения второго рода используются для возбуждения газовых смесей (гелий — неон, азот — углекислый газ).

Для возбуждения газовых лазеров используется импульсный и стационарный (тлеющий и дуговой) разряды. Импульсный разряд характеризуется короткими импульсами тока большой плотности (10^3 А/см²). Тлеющий разряд характеризуется небольшими плотностями тока (10^{-5} — 10^{-1} А/см²). Образующаяся при этом разряде плазма имеет малую степень ионизации (10^{-4} — $10^{-2}\%$). Дуговой разряд характеризуется большой плотностью тока (10^3 А/см²), степень ионизации около 10%.

Химический метод возбуждения приводит к созданию инверсной населенности в результате химических реакций, при которых образуются атомы или радикалы в возбужденном состоянии. Так как время жизни возбужденных уровней невелико, то возбуждение возможно только с помощью быстротекущих реакций, таких как фотодиссоциация молекул (распад молекулы на несколько частей при вспышке света), взрыв. Этот метод позволяет создать лазеры с большим к. п. д. и мощностью.

Оптический метод возбуждения малоэффективен для газовых лазеров. Обычный оптический источник имеет широкую полосу излучения. Газы же обладают узкими линиями поглощения и используют лишь малую часть энергии источника с широким спектром. Применение импульсных газоразрядных ламп, обладающих мощным линейчатым спектром, требует точного совпадения линии излучения лампы и линии поглощения рабочего газа. Оптический метод возбуждения более эффективен

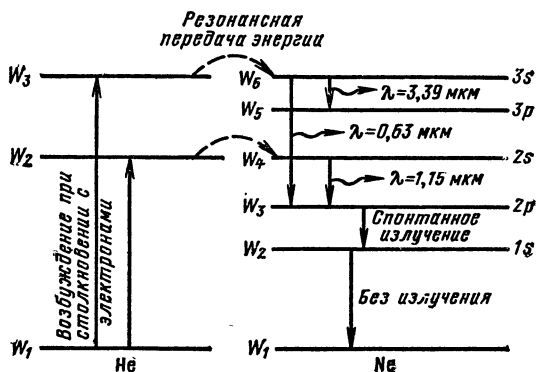


Рис. 4-64. Схемы энергетических уровней гелия и неона.

при использовании широких полос поглощения в спектрах газов, состоящих из двухатомных и многоатомных молекул.

Атомные лазеры. В качестве рабочей среды применяются чистые инертные газы (неон, криптон, ксенон, аргон); смеси газов (гелий — неон и гелий — ксенон). Наибольшее распространение получил гелий-неоновый лазер, рабочими атомами в котором являются атомы неона. Атомы гелия введены для увеличения инверсной населенности неона и соответственно выходной мощности излучения. По сравнению с излучением чистого неона интенсивность излучения лазера на смеси гелия и неона примерно в 50 раз больше.

Схема уровней гелия и неона приведена рис. 4-64. Атомы гелия вследствие соударений со свободными электронами переходят из основного состояния W_1 в возбужденные состояния, соответствующие метастабильным

уровням (время жизни 10^{-3} с) W_2 и W_3 . Верхние уровни неона также заселяются за счет соударений с электронами. В дополнение к этому заселение уровней W_4 и W_6 неона происходит за счет соударений второго рода атомов неона, находящихся в невозбужденном состоянии, с возбужденными атомами гелия, находящихся на уровнях W_2 и W_3 , совпадающих с уровнями W_4 и W_6 неона. При столкновениях возбужденные атомы гелия переходят в основное состояние, а атомы неона в возбужденное W_4 и W_6 . Одновременно происходит и обратный процесс передачи энергии от возбужденных атомов неона, находящихся на уровнях W_4 и W_6 , к невозбужденным атомам гелия. Чтобы преобладал прямой процесс, концентрация гелия должна превышать концентрацию неона примерно в отношении 10 : 1. При этом можно добиться большой заселенности уровней W_4 и W_6 и получить инверсную населенность между уровнями W_6 и W_5 , W_6 и W_3 , W_4 и W_3 . Опустошение уровня W_3 происходит за счет спонтанных переходов на уровень W_2 , который является метастабильным и опустошается за счет соударений атомов неона со стенками. Чтобы эти соударения интенсивно опустошали уровень W_2 , необходимо подобрать оптимальный диаметр трубки. Обычно трубка имеет диаметр 7—9 мм. При больших диаметрах мощность падает, несмотря на увеличение объема рабочего газа. Падение мощности обусловлено тем, что эффективно опустошаются уровни W_2 тех атомов, которые находятся вблизи стенок; атомы, находящиеся вблизи оси трубки, практически выключаются из процесса генерирования. При меньших диаметрах трубки мощность также уменьшается, так как уменьшается объем газовой смеси.

Наиболее эффективным является переход $3s \rightarrow 2p$, соответствующий длине волны 3,39 мкм. Усиление на этом переходе достигает 20 дБ/м. На переходах $2s \rightarrow 2p$ усиление составляет 10—12% на метр и $\lambda = 1,15$ мкм. На переходе $3s \rightarrow 2p$ усиление равно 4—6% на метр, но излучение соответствует видимому диапазону $\lambda = 0,63$ мкм и этот переход, несмотря на малое усиление, используется особенно часто.

Возбуждение газовой среды осуществляется или высокочастотным генератором с частотой от единиц до десятков мегагерц или источником постоянного напряжения, вызывающим тлеющий разряд в газоразрядной трубке.

Гелий-неоновый лазер обладает относительной нестабильностью частоты около 10^{-8} . Спектральная линия излучения неона значительно шире резонансных пиков, соответствующих собственным частотам резонатора, и стабильность определяется стабильностью длины резонатора. Под действием различного рода механических вибраций зеркал собственные частоты резонатора смещаются. Чтобы повысить стабильность, применяют различные схемы автоподстройки, использующие обратную связь. В одночастотном режиме информация о положении линии излучения лазера относительно центра доплеровского контура используется для стабилизации длины резонатора с помощью пьезокерамики, на которой укреплены зеркала. Размеры пьезокерамики зависят от приложенного поля. Изменение частоты преобразуется в изменение электрического поля, зеркала перемещаются, резонатор подстраивается и частота стабилизируется. Принципиально относительная нестабильность частоты может быть снижена до 10^{-14} .

Выходная мощность гелий-неонового лазера составляет сотни милливатт, коэффициент полезного действия — сотые доли процента, расходимость излучения — $3 \cdot 10^{-4}$ рад (дифракционный предел 10^{-4} рад).

Гелий-неоновый лазер может работать как в непрерывном режиме, так и в импульсном. Импульсный режим создается при импульсном питании газоразрядной трубки.

Ионные лазеры. В качестве рабочей среды в ионных лазерах применяются ионизированные инертные газы — ксенон, криптон, неон, аргон, а также ионы фосфора, серы и хлора. Ионизация рабочего вещества достигается в результате дугового разряда с плотностью тока до нескольких тысяч ампер на квадратный сантиметр. Поэтому ионные лазеры обычно работают в импульсном режиме.

Наиболее широкое применение нашли аргоновые ионные лазеры. Схема энергетических уровней аргона приведена на рис. 4-65. Инверсная населенность создается между уровнями $W_4(3p^4p)$ и $W_3(3p^4s)$. Уровень W_2 соответствует ионизации аргона (Ar^+). С уровня W_2 на уровни $W_4(3p^4p)$ и $W_5(3p^4d)$ ионы аргона переходят в результате столкновения с электронами. Уровень W_4 заселяется не только за счет соударений ионов с электронами в разряде, но и за счет переходов с W_5 . Уровень

W_4 является метастабильным, обладает большим временем жизни по сравнению с уровнем W_5 и частицы накапливаются на нем. Уровень W_3 обладает малым временем жизни, что быстро приводит к его опустошению. Таким образом, обеспечивается инверсия заселенности уровня W_4 по отношению к уровню W_3 . Так как уровни W_4 и W_3 состоят из групп подуровней, то генерирование может одновременно происходить в диапазоне от 0,45 до 0,53 мкм. В выходном излучении преобладают линии 0,488 (голубая) и 0,515 мкм (зеленая), в которых содержится до 50% общей мощности излучения.

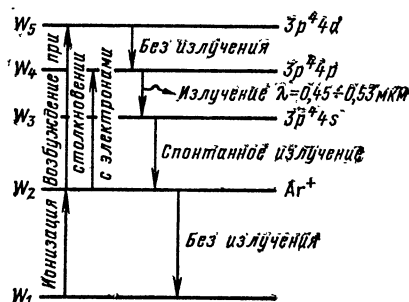


Рис. 4-65. Схема энергетических уровней аргона.

Прежде чем возбудить уровни, необходимо сначала газ ионизировать, что требует пропускания через газ тока большой плотности до нескольких тысяч ампер на 1 см^2 . Ток пропускают через тонкий (1—3 мм) капилляр, что позволяет получить большую плотность тока в разряде. С помощью продольного магнитного поля, создаваемого соленоидом, пучок электронов сжимается. Электроны вылетают из катода и летят через капилляр к аноду. При этом газ также перемещается в этом направлении. Чтобы компенсировать перекачку газа, анодную и катодную части трубки соединяют между собой обводной трубкой, что обеспечивает циркуляцию газа (рис. 4-66).

Главная трудность при конструировании — это борьба с разрушением капилляра вследствие бомбардировки его стенок ионами и электронами.

Выходная мощность ионных аргонных лазеров в непрерывном режиме составляет единицы ватт, к. п. д. де-

сятые доли процента. В импульсном режиме получена мощность в сотни киловатт.

Молекулярные лазеры. Газовые лазеры, использующие атомные и ионные переходы, имеют незначительный коэффициент полезного действия. Это связано с тем, что верхний рабочий уровень находится на значительном расстоянии от основного и поэтому не все электроны принимают участие в возбуждении. Только очень быстрые электроны способны возбудить в этом случае атом или ион.

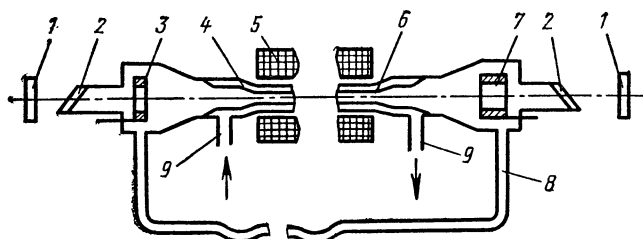


Рис. 4-66. Аргоновый лазер.

1 — зеркала; 2 — выходные окна; 3 — катод; 4, 9 — жидкостное охлаждение; 5 — катушка намагничивания; 6 — капилляр; 7 — анод, 8 — обводная трубка.

Более перспективным в отношении к. п. д. использования системы возбуждения является рабочее вещество, имеющее близко расположенные энергетические уровни. Такими являются, например, колебательные уровни молекул. Возможно создание молекулярных генераторов на углекислом газе CO_2 , азоте N_2 , кислороде O_2 и воде H_2O .

Наиболее перспективен молекулярный генератор на CO_2 . В молекуле CO_2 возможны три вида колебательных движений: симметричные, антисимметричные и деформационные (рис. 4-67). Соответственно и энергетический спектр состоит из трех групп состояний. Колебательные энергетические уровни расположены значительно ближе друг к другу, чем электронные, поэтому переходы между ними соответствуют излучению в более длинноволновом спектре. Схема нижних колебательных уровней молекулы CO_2 приведена на рис. 4-68. Инверсия населенностей может быть создана между уровнями W_5 и W_4 , W_5 и W_3 .

Верхний уровень W_5 отстоит от нижнего W_1 всего на 0,35 эВ.

Обычно в лазере углекислый газ используется в смеси с азотом. Схема уровней азота N_2 имеет более простой вид (рис. 4-68). Уровень W_2 азота совпадает с уровнем W_5 для углекислого газа. Верхние уровни азота W_2 и углекислого газа W_5 заселяются за счет соударений частиц с электронами, образующимися при разряде. Уровень W_2 азота заселяется очень сильно (до 30% об-

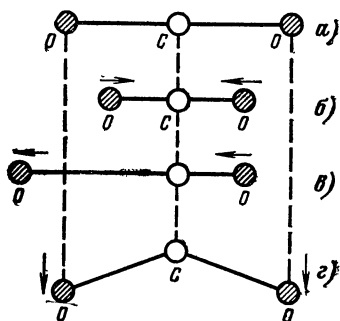


Рис. 4-67. Виды колебаний молекулы CO_2 .

a — положение равновесия; b — симметричные колебания; c — антисимметричные; d — деформационные.

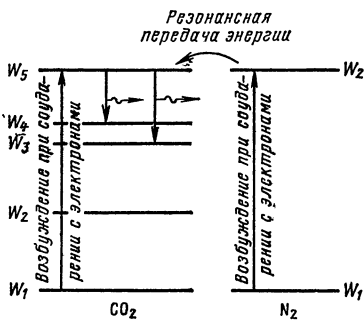


Рис. 4-68. Схемы энергетических уровней молекул CO_2 и N_2 .

щего числа молекул). Так как уровень W_2 азота и уровень W_5 совпадают, то заселение уровня W_5 происходит не только за счет электронного возбуждения, но и за счет передачи энергии возбужденными молекулами азота молекулам CO_2 .

Часто в рабочую смесь CO_2 и N_2 добавляют гелий, который также увеличивает заселенность уровня W_5 за счет переходов молекул CO_2 с уровней более высоких, чем W_5 при соударениях с атомами гелия. За счет этих столкновений также обедняются нижние уровни рабочих переходов W_4 и W_3 . Благодаря высокой теплопроводности гелия уменьшается температура газовой смеси, что уменьшает скорость релаксаций верхнего уровня рабочих переходов W_5 и увеличивает выходную мощность.

Каждый из колебательных уровней W_5, W_4 и W_3 представляет собой полосу близко расположенных

уровней. Переходы соответствуют излучению в диапазоне от 9 до 18 мкм. Наиболее эффективным является излучение на длине волны $\lambda = 10,6$ мкм. Путем перестройки резонатора возможна плавная перестройка частоты в довольно широких пределах (от 10,15 до 10,75 мкм), что связано с наличием множества колебательно-вращательных уровней (см. П1, пп. 3 и 7) молекулы CO_2 .

Лазеры на CO_2 характеризуются высоким уровнем генерируемой мощности. В непрерывном режиме получена мощность десятки киловатт при к. п. д. 20—30%.

Так как время жизни молекулы CO_2 в возбужденном состоянии велико (около 10^{-3} с), то возможна работа лазера в режиме модуляции добротности. Когда добротность отключается, возбужденные частицы накапливаются на верхнем уровне, при включении добротности все частицы разом переходят на более низкие уровни, излучая энергию в виде импульса длительностью 10—100 нс и пиковой мощностью до 10 МВт. Модуляция добротности осуществляется вращающейся призмой из кристалла поваренной соли NaCl , прозрачной для излучения в этом диапазоне.

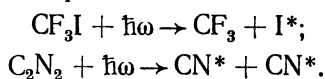
Химические лазеры. Инверсная населенность в этих лазерах создается в результате химических реакций, при которых происходит перестройка химических связей и образуются частицы в возбужденном состоянии. Это возбуждение осуществляется путем химической реакции между частицами рабочего вещества или путем химической реакции между добавками к рабочему веществу с последующей передачей энергии молекулам, участвующим в излучении. Реакции должны протекать настолько быстро, чтобы обеспечить достаточное количество возбужденных частиц за время, сравнимое с временем их жизни. К таким реакциям относятся фотодиссоциация молекул, диссоциация молекул при электрическом разряде. Инициированные внешним источником (вспышка света, электрический разряд и т. д.) эти реакции протекают без потребления внешней энергии. Возбуждение осуществляется за счет энергии, освобождающейся при реакции, создающей инверсную населенность. Реакции цепные, квантовый выход очень большой (до 10^5). Таким образом, химические лазеры обладают свойством самовозбуждения и высоким к. п. д.

Наибольшее распространение получили фотодиссоциационные лазеры. Явление фотодиссоциации молекул

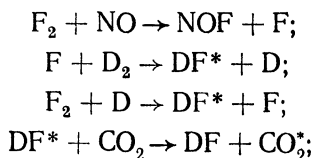
заключается в распаде молекулы под воздействием излучения на отдельные атомы или радикалы. При этом один из атомов или радикалов переходит в возбужденное состояние. Процесс протекает в том случае, если энергия кванта, поглощаемого молекулой, больше энергии связи частиц в молекуле.

Лазеры, основанные на фотодиссоциации, представляют собой кювету с газом, помещенную в резонатор и облучаемую импульсной газоразрядной лампой. Газоразрядная лампа характеризуется излучением, спектральный состав которого соответствует спектральной характеристике фотодиссоциации молекулы. Спектр фотодиссоциации гораздо шире (в 10^4 раз), чем ширина линии рабочего перехода в обычных газовых лазерах. Поэтому коэффициент использования излучения лампы достаточно высок. Для эффективного использования излучения газоразрядная лампа и кювета окружены отражателем. Так как энергия фотонов излучения лампы должна превышать энергию связи частиц в молекуле, которая обычно больше 3 эВ, то излучение лампы соответствует ультрафиолетовому диапазону.

В фотодиссоциационных лазерах в качестве активных сред используются молекулы TeBr , CH_3I , CH_3F , CF_3I и др. Реакция фотодиссоциации имеет вид:



В чисто химических лазерах реакции протекают без внешнего инициирующего воздействия на рабочую смесь. Примером может служить лазер на CO_2 с передачей энергии от возбужденных молекул DF^* . Рабочая среда представляет смесь газов CO_2 , NO , F_2 и D_2 . Возбуждение CO_2 происходит в результате следующих реакций:

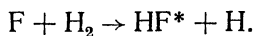


реакция протекает без затрат внешней энергии.

Химические лазеры работают в импульсном и непрерывном режиме. При работе в непрерывном режиме

требуется достаточно быстрая прокачка рабочей смеси через реакционную камеру.

Высокие мощности и к. п. д. получены в непрерывном режиме в быстроточных системах — газодинамических лазерах. В лазерах на H_2 и F реакция молекулярного водорода с атомарным фтором происходит при диффузии водорода в поток фтора, продуваемого через сопло со сверхзвуковой скоростью,



Основное достоинство химических лазеров — это возможность прямого преобразования химической энергии в энергию когерентного излучения большой мощности, независимость от внешних источников. Энергия излучения во много раз превышает внешнюю, подводимую к лазеру энергию.

Коэффициент преобразования химической энергии достаточно высок (более 10%). В этом состоит преимущество химических лазеров перед газоразрядными. Большой к. п. д. определяется тем, что в активной среде химических лазеров преобладает основной процесс (химическая реакция), ведущий к селективному заселению верхнего уровня рабочего перехода. В газоразрядных лазерах энергия свободных электронов, образовавшихся при разряде, расходуется на заселение всех состояний, в том числе и нижнего рабочего уровня. Добавки буферных газов, уменьшающих населенность нижнего рабочего уровня, для химических лазеров более эффективны. В газоразрядных лазерах эти добавки возмущают плазму, уменьшают концентрацию и энергию свободных электронов, а следовательно, уменьшают и заселенность верхнего уровня. Отсутствие плазмы в химических лазерах определяет стабильность характеристик.

4.8. ПОЛУПРОВОДНИКОВЫЕ ЛАЗЕРЫ

Полупроводник как рабочий материал (см. П1, п. 5) позволяет создать лазеры, работающие в широком диапазоне волн (от длинноволновой инфракрасной области до ультрафиолетовой), так как индуцированное излучение в полупроводниках связано с переходами через запрещенную зону, величина которой для различных полупроводников меняется в широких пределах от тысячных долей до единиц электронвольт. Полупроводни-

ковые лазеры имеют высокий к. п. д. (70—80%). Кроме того, полупроводниковые лазеры имеют малые габариты, простую конструкцию и могут перестраиваться по частоте. Однако излучение этих лазеров характеризуется более широким спектром частот, сравнимым со спектром тепловых источников света и большей расходимостью ($5 \cdot 10^{-2}$ рад), чем излучение газовых лазеров и лазеров на твердых диэлектриках.

Полупроводниковый лазер состоит из трех основных элементов: среды, резонатора и источника возбуждения.

В качестве рабочих сред наибольшее применение нашли антимониды индия и галлия ($\lambda = 4,9$ мкм и $1,53$ мкм), арсениды индия и галлия ($\lambda = 3,0$ мкм и $0,82—0,9$ мкм) и фосфид индия ($\lambda = 0,9$ мкм).

Излучение в оптическом диапазоне происходит в результате рекомбинации электронов и дырок (П1, п. 5 и § 1-4). Разработанные в настоящее время лазеры используют прямые переходы (П1, п. 7), хотя в принципе можно использовать среды и с непрямыми переходами.

Полупроводниковая среда активна при условии (1-52)

$$f_{\text{эп}} > f_{\text{зв}},$$

когда вероятность нахождения электрона в зоне проводимости больше, чем в валентной. Инверсная населенность обеспечивается при возбуждении полупроводника, когда в зоне проводимости и в валентной зоне создается избыточная, неравновесная концентрация носителей.

Неравновесные носители скапливаются у границ зон: электроны у дна зоны проводимости, дырки у потолка валентной зоны. Ширина полос энергетических уровней, занятых неравновесными носителями приблизительно равна kT и излучение полупроводниковых лазеров соответствует сравнительно узкой полосе частот

$$\Delta\omega = \frac{2kT}{\hbar},$$

сконцентрированных около центральной частоты

$$\omega_0 = \frac{\Delta W}{\hbar},$$

где ΔW — ширина запрещенной зоны.

Для создания инверсной населенности можно использовать следующие методы возбуждения: 1) инжекция (вспрыскивание) извне электронов и дырок в полупроводник. Этот метод позволяет получить инверсию населенностей в тонком слое, соответствующем p - n переходу; 2) бомбардировка полупроводника пучком электронов большой энергии. Электроны пучка, двигаясь сквозь кристаллическую решетку, передают часть своей кинетической энергии валентным электронам, которые переходят в зону проводимости; 3) оптическое возбуждение; 4) прямое электрическое возбуждение (пробой полупроводниковой структуры). Наибольшее распространение получили первый и второй методы.

Резонатор образуется между двумя плоскими параллельными поверхностями кристалла. Так как коэффициент преломления полупроводников достаточно высок (для арсенида галлия 3,5), то получаемый при этом коэффициент отражения от границы полупроводник — воздух вполне достаточен для выполнения условия самовозбуждения (4-81), если учесть, что коэффициент квантового усиления $\alpha_{\text{ус}}$ может превышать 100 см^{-1} . Остальные четыре плоскости кристалла делают шероховатыми, чтобы исключить отражения в нерабочих направлениях.

Условию $f_{\text{эл}} > f_{\text{эв}}$ соответствует условие вырождения (1-53):

$$W_{F_3} - W_{F_d} > \Delta W. \quad (4-130)$$

В чистых полупроводниках этому условию удовлетворить трудно, но ему легко удовлетворить в системе, состоящей из n и p полупроводников, в которых соответственно вырождены электроны и дырки. При соединении этих полупроводников в месте соединения (p - n переходе) будет выполняться условие инверсии населенности (4-130) (рис. 4-69, а). Часть электронов проводимости из n -области очень быстро перейдет в p -область, а часть дырок из p -области в n -область. При этом будет наблюдаться рекомбинация электронов и дырок, сопровождающаяся излучением. Излучение будет наблюдаться до тех пор, пока будет выполняться условие (4-130).

Однако это условие будет выполнено в области перехода только в начальный момент времени. Через короткий промежуток времени движение электронов и

дырок прекратится. Они перераспределятся по энергетическим уровням и придут в равновесие друг с другом. Уровни Ферми в p и n -областях совместятся (рис. 4-69, б), в области p - n перехода исчезнет одновременное вырождение электронов и дырок, а следовательно, и инверсия населенностей. Для того чтобы снова создать инверсию населенности в p - n переходе, надо добиться, чтобы уровни Ферми разошлись на расстояние, большее запрещенной зоны. Для этого надо приложить к системе электрическое поле прямого на-

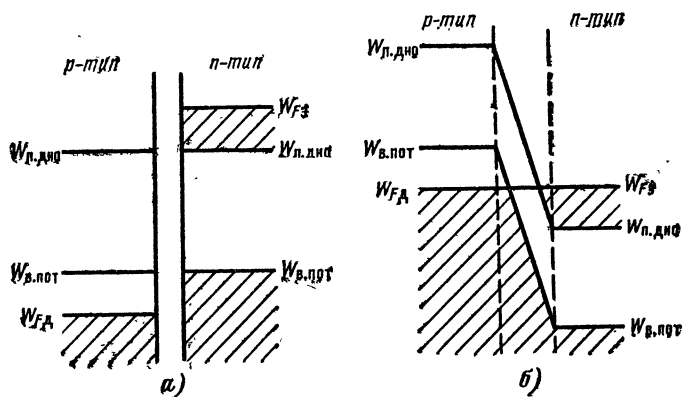


Рис. 4-69. Переход p - n .

а — в момент соединения полупроводников p и n -типа; б — установление равновесия.

правления. При этом через p - n переход потечет ток, состоящий из электронов и дырок. Эти два потока встречаются и, рекомбинируя, создают излучение. Встречные потоки будут поддерживать в переходном слое концентрацию, достаточную для вырождения электронов и дырок, и условие (4-130) выполняется (рис. 4-69, а).

Так как инверсная населенность обеспечивается лишь в сравнительно узкой области p - n перехода, а в n и p полупроводниках инверсия отсутствует, то при прохождении электромагнитной волны через p - n переход она усиливается, а в n и p -областях поглощается. Инверсная населенность тем больше, чем больше приложенное поле, т. е. чем больший ток протекает через p - n переход. Минимальная плотность тока $J_{\text{пор}}$, при

которой усиление в активной части среды компенсирует потери, называется *пороговой*:

$$J_{\text{пор}} = \frac{i_{\text{пор}}}{S},$$

где $i_{\text{пор}}$ — пороговый ток; S — площадь p - n перехода.

При токе, меньшем порогового, наблюдается некогерентное спонтанное излучение. Если ток больше порогового, то при облучении p - n перехода светом он усиливается. Усиленное вынужденное излучение когерентно. Для генерирования излучения необходима обратная связь (резонатор).

Качество инжекционного лазера в значительной степени определяется пороговой плотностью тока, так как эта характеристика определяет плотность потока энергии, вводимой в полупроводник. Величина $J_{\text{пор}}$ определяется качеством полупроводника, технологией его изготовления, температурой и размерами кристалла. Пороговое значение плотности тока можно найти из условия самовозбуждения (4-81), учитывая, что показатель усиления пропорционален пороговой плотности тока $\alpha_{\text{ус.пор}} = \beta J_{\text{пор}}$, где β — коэффициент, зависящий от температуры:

$$J_{\text{пор}} = \frac{1}{\beta} \left[\alpha_n + \frac{1}{l} \ln \frac{1}{\Gamma_1 \Gamma_2} \right].$$

Согласно этому выражению пороговое значение плотности тока уменьшается при увеличении длины кристалла. Однако при увеличении длины возрастают технологические трудности — граница раздела p - n перехода должна быть плоской и параллельной отражающим поверхностям кристалла.

При повышении температуры плотность порогового тока увеличивается. Это связано с уменьшением вырождения, а следовательно, и инверсной населенности и усиления полупроводниковой среды. Потери (коэффициент α_n) с повышением температуры также возрастают.

Инжекционные лазеры, в которых в качестве активной среды используется арсенид галлия, получили наибольшее распространение. Электронно-дырочный переход образуется путем диффузии цинка (акцептора) в пластину сильнолегированного арсенида галлия (полупроводник n -типа с концентрацией доноров около 10^{18} см^{-3}).

Конструкция лазера достаточно проста (рис. 4-70). Внешняя поверхность p -слоя металлизирована, и к ней, как и к металлической пластине — основанию, припаяны контактные выводы. Кристалл имеет форму усеченной пирамиды. Две боковые грани, образующие резонатор, строго параллельны, а две другие во избежание резонанса скошены под углом. Размеры кристалла составляют доли миллиметра.

В импульсном режиме возбуждение осуществляется импульсами тока длительностью около 10^{-6} с. Порого-

вое значение тока зависит от температуры активной среды и возрастает с повышением этой температуры. При температуре жидкого гелия пороговый ток имеет плотность около 10^2 А/см², при температуре жидкого азота плотность порогового тока возрастает до 10^5 А/см², при комнатной температуре эта величина в несколько раз больше. Максимальная достигнутая мощность полупроводниковых лазеров составляет около 100 Вт при средней мощности десятки милливатт.

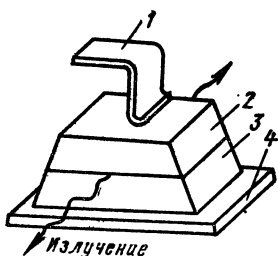


Рис. 4-70. Полупроводниковый инжекционный лазер.

1, 4 — контакты; 2 — p -полупроводник; 3 — n -полупроводник.

В непрерывном режиме при температуре жидкого гелия требуется плотность тока до 10^2 А/см², при комнатной температуре требуемая плотность тока ведет к недопустимому перегреву кристалла. В непрерывном режиме при гелиевых температурах мощность достигает 10 Вт.

Значительное снижение пороговой плотности тока получено в гетероструктурах. В инжекционных лазерах, представляющих собой двойную гетероструктуру, пороговая плотность тока при $T=300$ К составляет лишь $1600\text{--}3500$ А/см², что обеспечивает работу в непрерывном режиме при комнатной температуре. Двойная гетероструктура на основе арсенида галлия состоит из слоя GaAs p -типа толщиной около 1 мкм, заключенного между двумя областями n и p -типа смешанного кристалла (Ga+Al) As, обладающего более широкой за-

прещенной зоной. Контактные поля гетеропереходов удерживают инжектированные электроны и дырки в центральном слое из GaAs, где значительно увеличивается концентрация электронов в зоне проводимости и дырок в валентной зоне, а следовательно, увеличивается инверсия населенностей и коэффициент усиления $\alpha_{\text{ус}}$. При этом порог генерирования достигается при меньшей плотности тока. Снижению порога способствует уменьшение коэффициента поглощения $\alpha_{\text{п}}$ из-за отсутствия межзонного поглощения в областях (Ga+Al) As, так как энергия квантов генерируемого излучения существенно меньше, чем ширина запрещенной зоны в этих областях.

Высокий коэффициент полезного действия (до 70—80%) обеспечивается непосредственным преобразованием электрической энергии в излучение.

Расходимость излучения составляет несколько градусов, что намного больше, чем у газового лазера. Это объясняется малыми размерами резонатора, что приводит к ухудшению его добротности, и малой площадью поверхности излучения активного элемента.

Для увеличения мощности излучения используются решетки из несинхронизированных между собой лазеров. Такая решетка компактна, если она работает при комнатной температуре.

Для охлаждения полупроводниковых лазеров в лабораторных условиях используются сосуды Дьюара с жидким азотом или гелием, в промышленных установках — охладители, использующие эффект понижения температуры газа при его резком расширении и термоэлектрический эффект.

Полупроводниковые лазеры с электронным возбуждением имеют более сложную конструкцию, чем инжекционные. Поверхность полупроводника бомбардируется потоком электронов, формируемым электронной пушкой, включающей систему фокусировки и модуляции луча. Это определяет габариты лазера. Энергия электронов в возбуждающем электронном потоке составляет 20—50 кэВ. Проникая в полупроводник на глубину 0,01—0,1 мм, электроны потока возбуждают связанные электроны полупроводника и переводят их из валентной зоны в зону проводимости. Вновь образовавшиеся электроны также участвуют в возбуждении. Процесс нарастает лавинно. Один бомбардирующий электрон

энергией в 20—50 кэВ может образовать несколько тысяч свободных электронов.

В результате возбуждения электроны переходят на уровни, далеко отстоящие от дна зоны проводимости, но вследствие взаимодействия с кристаллической решеткой они переходят на более низкие уровни вблизи дна зоны проводимости. Если плотность потока электронов достаточна, то условие (4-130) выполняется и полупроводниковая среда становится активной.

Лазеры с электронной бомбардировкой работают в импульсном режиме при температуре жидкого азота или гелия.

Глава пятая

КВАНТОВЫЕ ГЕНЕРАТОРЫ И УСИЛИТЕЛИ СВЧ ДИАПАЗОНА

5-1. УСИЛЕНИЕ ЭЛЕКТРОМАГНИТНОЙ ВОЛНЫ В ПАРАМАГНИТНОЙ СРЕДЕ

Квантовые парамагнитные генераторы не нашли практического применения, так как не отличаются стабильностью частоты, существенно уступая в этом отношении молекулярным генераторам, имеют значительно более низкий уровень генерируемой мощности и к. п. д. по сравнению с электронными СВЧ генераторами.

Широкое применение нашли квантовые парамагнитные усилители, которые обладают по сравнению с другими типами усилителей СВЧ диапазона чрезвычайно низким уровнем собственных шумов.

Усиление распространяющейся электромагнитной волны в парамагнитной среде с инверсной населенностью происходит за счет индуцированного излучения при электронном парамагнитном резонансе (см. § 2-3).

Парамагнетизм сред связан с наличием магнитного момента у атомов или ионов. В результате вращения вокруг собственной оси (см. П1, п. 1) электрон, входящий в состав атома или молекулы, обладает механическим моментом количества движения (спином) и магнитным моментом, направленным противоположно механическому и равным магнетону Бора. Вращаясь по

орбите вокруг ядра, электрон обладает еще и орбитальным магнитным моментом. Полный магнитный момент атома определяется суммой орбитальных и спиновых магнитных моментов электронов.

Электроны полностью заполненных электронных оболочек (см. П1, п. 2) составляют пары, имеющие скомпенсированные, т. е. противоположно направленные магнитные (орбитальные и спиновые) моменты. Исключение составляют внешние электроны, образующие незаполненные оболочки. Вступая в химическую связь, атомы или обмениваются электронами, как, например, при образовании молекулы NaCl , или внешние электроны образуют общую оболочку, охватывающую оба атома, как в молекулах H_2 . В обоих случаях парамагнитные свойства атомов исчезают, так как образуются пары электронов со скомпенсированными магнитными моментами. Обычно наблюдается тенденция к образованию пар со скомпенсированными магнитными моментами, поэтому большинство сред диамагнитно. Однако существуют целые группы элементов с недостроенными внутренними оболочками, например, группа железа с недостроенной $3d$ -оболочкой и группа редкоземельных элементов с недостроенной $4f$ -оболочкой. Атомы этих элементов обладают парамагнитными свойствами, определяемыми нескомпенсированными магнитными спиновыми моментами электронов недостроенных оболочек. В химических соединениях эти атомы, превращаясь в ионы, в большинстве случаев сохраняют парамагнитные свойства.

Для квантовых парамагнитных усилителей используются диамагнитные кристаллы (матрица) с примесью парамагнитных ионов (активатор). Концентрация парамагнитных ионов мала (0,01—1%) и взаимодействие магнитных моментов соседних ионов мало. Но эти ионы находятся под действием достаточно сильного внутрикристаллического электрического поля диамагнитного кристалла. Под действием этого электрического поля энергетические уровни ионов расщепляются (эффект Штарка). Энергетические уровни ионов с нечетным числом электронов расщепляются на подуровни, различающиеся абсолютной величиной проекции магнитного момента на направление поля. Величина расщепления зависит от величины и конфигурации электрического поля кристалла.

Электрическое поле кристалла не полностью снимает вырождение. Магнитные подуровни парамагнитных ионов остаются двукратно вырожденными в электрическом поле. Снять это вырождение, т. е. расщепить до конца все магнитные подуровни, можно только с помощью внешнего магнитного поля (эффект Зеемана). При этом у «многоэлектронных» парамагнитных частиц возникает целый набор энергетических уровней $W_1, W_2, W_3...$, расстояние между которыми определяется величиной внешнего магнитного поля и его направлением относительно оси кристалла. Изменяя величину и направление магнитного поля, можно изменять расстояние между уровнями.

Под действием внешнего переменного электромагнитного поля, частота которого соответствует разностям энергий соседних подуровней, осуществляются вынужденные переходы ионов с одного уровня на другой. Эти переходы подчиняются правилам отбора (см. П1, п. 7):

$$\Delta J = 0, \Delta M_J = \pm 1$$

и при умеренных значениях постоянного магнитного поля (порядка тысячи гаусс) соответствуют диапазону СВЧ. Вероятность перехода максимальна при круговой поляризации переменного магнитного поля, направление вращения которого совпадает с направлением прецессии спинов.

Вероятность спонтанных переходов между магнитными подуровнями в СВЧ диапазоне мала. И спонтанное излучение, определяющее шумы усилителей, примерно на 15 порядков меньше, чем в оптическом.

При тепловом равновесии согласно распределению Больцмана (П1-3)

$$\frac{N_n}{N_m} = e^{-\frac{\hbar\omega}{kT}}$$

нижние магнитные подуровни населены больше, чем верхние. Поэтому вынужденные переходы приводят к резонансному поглощению (электронный парамагнитный резонанс).

Соотношением типа больцмановского

$$\frac{N_2}{N_1} = e^{-\frac{\hbar\omega}{kT_s}},$$

где T_s — так называемая *спиновая температура*, можно характеризовать любое состояние парамагнитной среды. При тепловом равновесии спиновая температура равна температуре окружающей среды T . При $T_s \rightarrow +0$, все частицы находятся на нижнем уровне. С повышением температуры T_s все большее число частиц переходит на верхний уровень и при $T_s \rightarrow +\infty$ населенность обоих уровней выравнивается.

Состояние с инверсной населенностью, очевидно, характеризуется отрицательной спиновой температурой ($T_s < 0$). В этом случае спиновая температура совпадает с термодинамической температурой изолированной спиновой системы. Наибольшая инверсная населенность будет получена при $T_s \rightarrow -0$, когда все частицы будут находиться на верхнем уровне. Вынужденные переходы в среде с инверсной населенностью приводят к резонансному излучению.

Ширина линии парамагнитного резонанса определяется в основном временем поперечной или спин-спиновой релаксации [см. (2-46)], а также временем спин-решеточной релаксации, неоднородностями кристаллического и внешнего магнитного полей. Для руфина и рутила с железом, применяемым в парамагнитных усилителях, ширина линий приблизительно 60 МГц. Ширина линий для рутила с хромом 15—20 МГц.

Форма линии в рубине зависит от концентрации хрома. При концентрациях Cr^{3+} , меньших 0,1%, она близка к гауссовой, при увеличении концентрации приближается к лоренцевой (см. П1, п. 20). Форма линий для рутила с железом при концентрациях, меньших 0,1% — лоренцева, для рутила с концентрацией хрома 0,01—0,3% — промежуточная между гауссовой и лоренцевой.

Распространение электромагнитной волны в среде характеризуется изменением амплитуды волны по экспоненциальному закону:

$$E_m(x) = E_{m0} e^{-\alpha x}.$$

Парамагнитная среда обладает малыми диэлектрическими и магнитными потерями и коэффициент затухания α определяется выражением

$$\alpha = \frac{\omega \sqrt{\epsilon'_a \mu'_a}}{2} (\text{tg } \delta_s + \text{tg } \delta_m),$$

где δ_ϵ и δ_μ — углы диэлектрических и магнитных потерь, или

$$\alpha = \frac{\omega}{2v} \left(\frac{\epsilon''}{\epsilon'} + \frac{\mu''}{\mu'} \right), \quad (5-1)$$

где $v = \frac{1}{\sqrt{\epsilon'_a \mu'_a}}$ — скорость распространения волны в неограниченной парамагнитной среде.

В парамагнитной среде с инверсной населенностью $\epsilon'' > 0$, а $\mu'' < 0$. При этом первое слагаемое в выражении (5-1) положительно и определяет коэффициент затухания

$$\alpha_n = \frac{\omega}{2v} \frac{\epsilon''}{\epsilon'} > 0,$$

а второе слагаемое отрицательно и определяет коэффициент квантового усиления

$$\alpha_{yc} = \frac{\omega}{2v} \frac{\mu''}{\mu'} < 0. \quad (5-2)$$

Отношение μ''/μ' в формуле (5-2) определяется выражениями (2-45), в которых статическая магнитная восприимчивость χ_{m0} определяется для случая парамагнитной среды с инверсной населенностью. Намагниченность среды M_0 определяется проекциями магнитных моментов электронов на направление постоянного магнитного поля H_0 . Для простейшего случая среды с двумя магнитными подуровнями возможны лишь две ориентации магнитных моментов: по полю и против поля. Ориентации по полю соответствует нижний уровень, против поля — верхний. Так как магнитный момент электрона равен магнетону Бора $\mu_B = \frac{\hbar}{2} \frac{e}{m}$, то намагниченность, как магнитный момент единицы объема, равна:

$$M_0 = \mu_B (N_2 - N_1),$$

где N_2 и N_1 — населенности верхнего и нижнего магнитного подуровней соответственно.

Для среды с инверсной населенностью

$$\Delta N = N_2 - N_1 > 0.$$

Таким образом,

$$\chi_{m0} = \frac{M_0}{H_0} = \frac{\mu_B \Delta N}{H_0}. \quad (5-3)$$

Учитывая, что расстояние между подуровнями

$$\Delta W = \hbar \omega_{21} = 2\mu_0 \mu_B H_0,$$

выражение (5-3) можно представить в виде

$$\chi_{m0} = \frac{2\mu_0 \mu_B^2 \Delta N}{\hbar \omega_L}. \quad (5-4)$$

При резонансе ($\omega = \omega_{21} = \omega_L$) коэффициент квантового усиления максимален и согласно выражению (5-2) с учетом (2-45) и (5-4) получим:

$$\alpha_{\text{ус. макс}} = \frac{\omega_L}{2\nu} \mu_{\text{макс}} = \frac{\omega_L \mu_0 \mu_B^2 \Delta N T_2}{2\nu \hbar}. \quad (5-5)$$

Таким образом, коэффициент квантового усиления возрастает при возрастании частоты парамагнитного резонанса, так как при этом возрастает энергия излучаемого кванта. Увеличение $\alpha_{\text{ус. макс}}$ при возрастании инверсной населенности определяется увеличением числа активных частиц, отдающих энергию электромагнитному полю. Возрастание $\alpha_{\text{ус. макс}}$ с увеличением времени спин-спиновой релаксации T_2 определяется увеличением времени взаимодействия поля со спиновой системой. При уменьшении скорости распространения волны $v = c/\sqrt{\epsilon\mu}$ величина $\alpha_{\text{ус. макс}}$ возрастает, так как при этом также возрастает время взаимодействия поля со средой.

Если коэффициент квантового усиления $\alpha_{\text{ус}}$ больше коэффициента затухания $\alpha_{\text{п}}$, то волна усиливается и ее амплитуда при прохождении участка активной среды длиной l возрастает по закону

$$E_m = E_{m0} e^{\alpha l},$$

где $\alpha = \alpha_{\text{ус}} - \alpha_{\text{п}}$.

5.2. ПАРАМАГНИТНЫЙ УСИЛИТЕЛЬ БЕГУЩЕЙ ВОЛНЫ

Квантовый парамагнитный усилитель бегущей волны представляет собой волновод, заполненный парамагнитной средой полностью или частично. Это может быть прямоугольный волновод с волной H_{01} , перпендикулярно

узкой стенке которого приложено постоянное магнитное поле. Магнитное поле должно быть однородным в объеме, занимаемом рабочим веществом. Электромагниты со сверхпроводящими обмотками из ниобиевой проволоки помещают вместе с парамагнитным кристаллом в криостат с жидким гелием. Такие магниты работают без постоянных источников тока. После возбуждения тока обмотки электромагнита закорачиваются сверхпроводящей перемычкой. Масса таких магнитов невелика.

Парамагнитный кристалл является наиболее важным элементом парамагнитного усилителя, определяющим его характеристики. Кристаллы должны обладать соответствующей системой энергетических уровней, позволяющей использовать имеющиеся источники возбуждения. Частота рабочего перехода должна соответствовать нужному диапазону частот при приемлемых напряженностях внешнего магнитного поля. Время релаксации должно быть около 10^{-2} — 10^{-4} с, что позволяет получить инверсную населенность при небольшой мощности источника возбуждения. Кристаллы должны иметь высокую теплопроводность при гелиевых температурах и обладать малыми диэлектрическими потерями.

Этим требованиям наиболее полно удовлетворяют кристаллы рубина и кристаллы рутила с примесью ионов Cr^{3+} и Fe^{3+} .

Рубин применяется в усилителях сантиметрового и дециметрового диапазона. Рубин представляет собой диамагнитный кристалл корунда (Al_2O_3), в котором часть ионов алюминия замещена трехвалентными ионами хрома Cr^{3+} . Парамагнетизм хрома с электронной конфигурацией $1s^2 2s^2 2p^6 3s^2 3p^6 3d^5 4s^1$ обусловлен электронами внутренней незаполненной $3d$ -оболочки. В химических соединениях валентные связи обеспечиваются одним $4s$ и двумя $3d$ электронами и атом превращается в парамагнитный ион с электронной конфигурацией $1s^2 2s^2 \times \times 2p^6 3s^2 3p^6 3d^3$, парамагнитные свойства которого определяются нескомпенсированными спинами трех электронов теперь уже внешней $3d$ -оболочки. Результирующий магнитный момент иона равен утроенному магнитному моменту электрона. Проекция этого момента на выбранное направление может принимать четыре значения ($-3/2 \hbar$, $-1/2 \hbar$, $+1/2 \hbar$, $+3/2 \hbar$) и энергетические уровни изолированного иона хрома имеют четырехкратное вырождение. Ионы хрома в кристаллической решетке Al_2O_3 испы-

тывают действие сильного внутрикристаллического электрического поля, которое приводит к расщеплению основного энергетического уровня на два, отличающихся друг от друга только абсолютной величиной проекции магнитного момента на ось кристалла (рис. 5-1). Каждый из этих уровней двукратно вырожден ($\pm 1/2$, $\pm 3/2$). Внешнее магнитное поле снимает это вырождение, расщепляя каждый уровень на два простых уровня, различающихся знаком проекции магнитного момента. Таким образом, основной уровень расщепляется на четыре подуровня. Величина расщепления зависит как от величины постоянного магнитного поля, так и от ориентации этого поля относительно кристаллографической оси кристалла. Обычно ориентируют поле под углом $\theta = 90^\circ$ или $54,7^\circ$. На длинных волнах ($\lambda > 5$ см) используется 90° -ная ориентация поля (рис. 5-2) и трехуровневая схема (см. рис. 3-6). В слабых магнитных полях уровни W_1 и W_2 практически сливаются, возбуждение осуществляется на частоте ω_{14} (или ω_{24}), рабочий переход соответствует частоте ω_{43} . При сильном внешнем магнитном поле уровни W_1 и W_2 расщепляются и получается трехуровневая схема с частотой возбуждения ω_{14} и частотой рабочего перехода ω_{21} . Инверсная населенность в этом случае больше, чем в предыдущем.

При длине волны $\lambda < 5$ см используется ориентация поля под углом $\theta = 54,7^\circ$ и четырехуровневая схема (рис. 3-7) при двойном возбуждении (см. § 3-3). Возбуждение осуществляется одновременно на частотах ω_{13} и ω_{24} ($\omega_{13} = \omega_{24}$), частота рабочего перехода ω_{32} . Для возбуждения используются обычно клистронные генераторы соответствующего диапазона. Энергия возбуждения вводится в волновод с помощью специального ввода коаксиального типа в дециметровом диапазоне и длинноволновой части сантиметрового диапазона и волноводного типа в коротковолновой части сантиметрового и миллиметрового диапазона.

Охлаждение парамагнитного кристалла до температуры жидкого гелия позволяет резко увеличить инверсную населенность рабочего перехода, т. е. получить в усилителе большой коэффициент усиления; уменьшить тепловые процессы в кристалле, т. е. увеличить время спин-решеточной релаксации и, следовательно, уменьшить мощность источника возбуждения, необходимую для создания определенной инверсной населенности.

Кроме того, как уже говорилось, охлаждение понижает уровень собственных шумов усилителя, имеющих в основном тепловой характер (см. § 3-6).

В миллиметровом диапазоне волн в качестве активной среды используется рутил (окись титана TiO_2) с примесями ионов хрома Cr^{3+} или железа Fe^{3+} . Парамагнетизм железа с электронной конфигурацией $1s^2 2s^2 2p^6 3s^2 3p^6 \times$

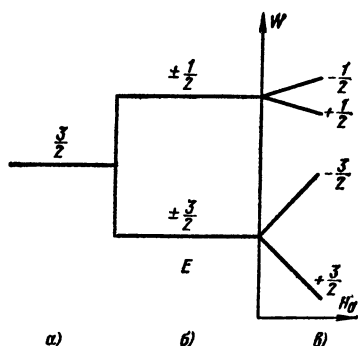


Рис. 5-1. Расщепление основного энергетического уровня хрома в кристалле рутина.

a — основной уровень; *б* — расщепление в электрическом поле кристалла; *в* — расщепление во внешнем магнитном поле.

$\times 3d^6 4s^2$ определяется незаполненной $3d$ -оболочкой. В химических соединениях валентные связи обеспечиваются двумя $4s$ и одним $3d$ электронами. При этом атом железа превращается в парамагнитный ион Fe^{3+} с электронной конфигурацией $1s^2 2s^2 2p^6 3s^2 3p^6 3d^5$.

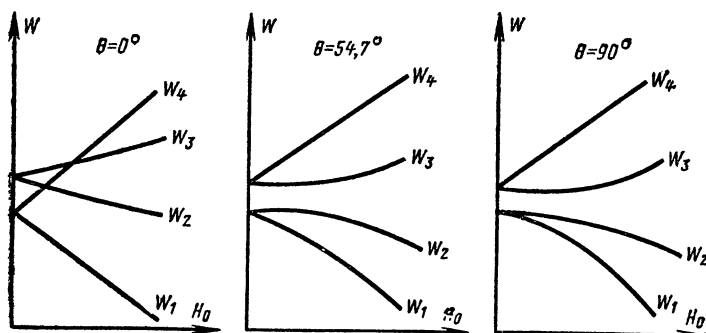


Рис. 5-2. Расщепление уровней хрома в рутине в зависимости от величины и ориентации постоянного магнитного поля.

В миллиметровом диапазоне, как это следует из графиков на рис. 5-2, постоянное магнитное поле должно быть больше, чем в сантиметровом диапазоне.

Электромагнитное поле в волноводе представляет собой суперпозицию плоских волн, распространяющихся в волноводе от одной стенки к другой (рис. 5-3). В волноводе, полностью заполненном парамагнитной средой с инверсной населенностью, распространение плоских волн будет характеризоваться коэффициентом $\alpha = \alpha_{\text{ус}} - \alpha_{\text{п}}$, определяемом выражением (5-1) при $\epsilon'' > 0$ и $\mu'' < 0$. Изменение поля вдоль оси волновода, очевидно, будет определяться коэффициентом

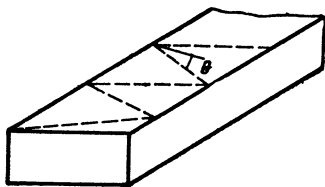


Рис. 5-3. Распространение плоской волны в волноводе.

$$\alpha' = \frac{\alpha}{\sin \theta},$$

где θ — угол падения плоской волны на стенку волновода. Этот угол определяется выражением

$$\sin \theta = \sqrt{1 - (\lambda/\lambda_{\text{кр}})^2},$$

где $\lambda_{\text{кр}}$ — критическая длина волны, определяемая размерами волновода.

Таким образом,

$$\alpha' = \alpha'_{\text{ус}} - \alpha'_{\text{п}} = \frac{\omega}{2v_{\text{гр}}} \left(\frac{\epsilon''}{\epsilon'} + \frac{\mu''}{\mu'} \right), \quad (5-6)$$

где $v_{\text{гр}}$ — групповая скорость в волноводе:

$$v_{\text{гр}} = v \sqrt{1 - \left(\frac{\lambda}{\lambda_{\text{кр}}} \right)^2}.$$

Групповая скорость характеризует скорость распространения энергии поля. Уменьшение групповой скорости приводит к увеличению времени взаимодействия распространяющегося поля с активной средой, т. е. к увеличению числа активных частиц, которые отдают свою энергию полю, и, следовательно, к увеличению коэффициента усиления.

Если коэффициент квантового усиления в волноводе больше коэффициента затухания ($\alpha'_{\text{ус}} > \alpha'_{\text{п}}$), то распро-

страняющееся поле усиливается и амплитуда его на выходе волновода длиной l равна:

$$E_m = E_{m0} e^{\alpha' l},$$

где $\alpha' = \alpha'_{\text{ус}} - \alpha'_n$; E_{m0} — амплитуда на входе волновода.

Коэффициент усиления по мощности в децибелах без учета потерь в стенках волновода и диэлектрических потерь в парамагнетике, заполняющем волновод, определится согласно (5-6) выражением

$$G \text{ (дБ)} = 27,3 \frac{c}{v_{\text{гр}}} \frac{l}{\lambda} \frac{\mu''}{\mu'}. \quad (5-7)$$

Из формулы (5-7) следует, что для получения усиления $G = 30$ дБ при типичном значении $\mu''/\mu' = 5 \cdot 10^{-3}$ и распространении волны со скоростью света ($c/v_{\text{гр}} = 1$) волноводная система должна иметь длину $l \approx 200 \lambda$, т. е. при $\lambda = 5$ см $l = 10$ м. Обычно парамагнитные кристаллы заполняют лишь часть поперечного сечения волновода и длина системы должна быть еще больше.

Для получения необходимого усиления при приемлемой длине усилителя ($l \approx 2 \lambda$) используют замедляющие системы, увеличивающие отношение $c/v_{\text{гр}}$ [см. (5-7)]. Практически используется замедление $c/v_{\text{гр}}$ около сотни.

Выражения (2-45) и (5-4) при инверсной населенности среды ($\Delta N = N_2 - N_1 > 0$) определяют активную среду. При резонансе ($\omega = \omega_n$) коэффициент усиления по мощности максимален и с учетом этих соотношений согласно (5-7) равен:

$$G_{\text{макс}} \text{ (дБ)} = 27,3 \frac{c}{v_{\text{гр}}} \frac{l}{\lambda} \frac{\mu_0 \mu_B^2 \Delta N T_2}{\hbar}. \quad (5-8)$$

Из полученного выражения видно, что коэффициент $G_{\text{макс}}$ (дБ) аналогично коэффициенту $\alpha_{\text{ус. макс}}$ [см. (5-5)] растет с увеличением частоты парамагнитного резонанса, инверсной населенности, времени спин-спиновой релаксации T_2 и замедления $c/v_{\text{гр}}$. При сильных распространяющихся полях инверсная населенность уменьшается, происходит насыщение. При этом уменьшается и коэффициент усиления $G_{\text{макс}}$.

При лоренцевой форме линии излучения электрического парамагнитного резонанса согласно формуле (5-7)

с учетом (2-45), (2-46) и (5-8) получим аналогично (4-100)

$$G(\text{дБ}) = \frac{G_{\text{макс}}(\text{дБ})}{1 + \left[\frac{2(\omega - \omega_{\text{Л}})}{\Delta\omega_{\text{Л}}} \right]^2}, \quad (5-9)$$

где $\omega - \omega_{\text{Л}}$ — расстройка; $\Delta\omega_{\text{Л}}$ — ширина линии парамагнитного резонанса, определяемая выражением (2-46).

Ширина полосы пропускания усилителя $\Delta\omega_{\text{ус}}$ определяется согласно формуле (5-9) аналогично (4-101) выражением

$$\Delta\omega_{\text{ус}} = \omega_2 - \omega_1 = \Delta\omega_{\text{Л}} \sqrt{\frac{3}{G_{\text{макс}}(\text{дБ}) - 3}}. \quad (5-10)$$

Обычно $\Delta\omega_{\text{Л}} \approx 60$ МГц и при усилении $G_{\text{макс}}(\text{дБ}) = 30$ дБ полоса пропускания составляет 20 МГц.

Согласно формуле (5-10) полоса пропускания усилителя всегда меньше ширины линии рабочего перехода. Увеличение ширины линии $\Delta\omega_{\text{Л}} = 2/T_2$ приводит к уменьшению времени T_2 и согласно выражению (5-8) к уменьшению коэффициента усиления. Таким образом, полоса пропускания усилителя и коэффициент усиления взаимосвязаны. Возможности усилителя наиболее полно характеризуются эффективностью

$$F = (\omega_2 - \omega_1) \sqrt{G_{\text{макс}}} \quad (5-11)$$

(практически достигнута эффективность $F \approx 1000$ МГц).

Динамический диапазон усилителя снизу ограничен шумами, сверху — насыщением рабочего вещества. В режиме насыщения инверсная населенность резко уменьшается и коэффициент усиления падает. Обычно динамический диапазон сверху ограничивается уменьшением коэффициента усиления по мощности вдвое (рис. 5-4). Максимальная мощность входного сигнала для квантового парамагнитного усилителя составляет примерно 10^{-6} Вт.

Динамический диапазон парамагнитного усилителя может быть увеличен, с одной стороны, за счет уменьшения собственных

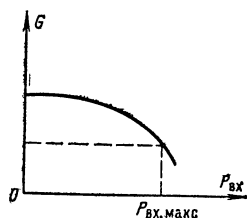


Рис. 5-4. Зависимость коэффициента усиления от мощности входного сигнала.

шумов усилителя, с другой, за счет увеличения объема активного вещества и концентрации активных частиц.

Как уже говорилось, для получения требуемого усиления при приемлемых размерах волноводной системы необходимо замедление волны. В миллиметровом диапазоне при использовании рутила с диэлектрической проницаемостью $\epsilon = 100 \div 200$ замедление ($c/v_{гр} \approx 10 \div 14$) осуществляется самим парамагнетиком. Но в сантиметровом

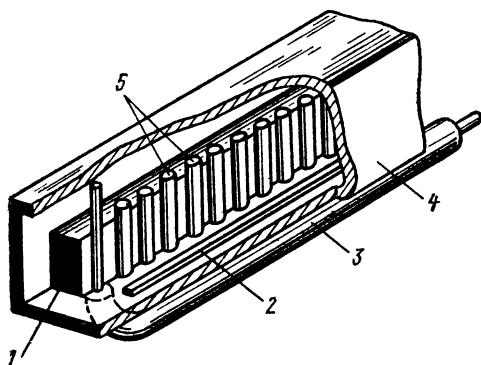


Рис. 5-5. Гребенчатая замедляющая система.

1 — парамагнетик; 2 — ферритовый вентиль; 3 — коаксиальный ввод; 4 — волновод; 5 — замедляющие штыри.

и дециметровом диапазоне используются периодические замедляющие структуры, дающие значительно большее замедление, чем замедление за счет диэлектрической проницаемости парамагнетика (для рубина $\epsilon \approx 10$, $c/v_{гр} \approx 3$), заполняющего волновод.

Наибольшее применение нашли гребенчатые замедляющие системы, представляющие собой гребенку параллельных штырей, находящихся в плоскости, параллельной широкой стенке волновода на одинаковом расстоянии друг от друга (рис. 5-5). Замедление $c/v_{гр}$ увеличивается при уменьшении периода системы. Периодические системы позволяют получить замедление $c/v_{гр} = 100 \div 1000$. Структура поля бегущей волны определяется периодичностью замедляющей системы. В точках, удаленных друг от друга на один период системы, поля имеют одинаковую форму и величину и отличаются толь-

ко фазам. Магнитное поле волны поляризовано в плоскости, перпендикулярной штырям, и сконцентрировано около их основания, что позволяет уменьшить сечение парамагнитного стержня. Плоскость гребенки разбивает пространство волновода на две области, в которых поляризация магнитного поля имеет противоположные направления. Парамагнитный кристалл расположен в той части, где магнитное поле имеет круговую поляризацию и направление вращения поля совпадает с направлением прецессии спинов электронов. Такой кристалл при постоянном подмагничивании и инверсной населенности ($\mu'' < 0$) усиливает распространяющееся поле. С полем обратной волны кристалл не взаимодействует.

Для устранения обратной волны по другую сторону от гребенки симметрично парамагнитному усиливающему кристаллу помещают ферритовый вентиль, основанный на ферромагнитном резонансном поглощении (см. § 2-3). Поглощение при ферромагнитном резонансе значительно превышает усиление в парамагнетике, поэтому при сравнительно небольшом объеме можно получить значительное поглощение обратной волны (около 30 дБ). Величина поля подмагничивания внутри феррита зависит от его геометрии и размеров. Размеры подбираются таким образом, чтобы ферромагнитный резонанс происходил при том же внешнем постоянном магнитном поле, что и парамагнитный резонанс в усиливающем парамагнитном кристалле.

Гребенчатая структура имеет конечную полосу пропускания, расположенную несколько ниже резонансной частоты отдельных штырей. Замедление возрастает к краям полосы пропускания до $c/v_{гр} = \infty$ на границах полосы и имеет минимальную, но все же большую величину на значительном центральном участке. Увеличение замедления на этом центральном участке достигается путем уменьшения полосы пропускания. Полоса уменьшается при увеличении расстояния от конца штырей до узкой стенки. Практически на волне 6 см можно получить замедление $c/v_{гр} = 200$ на участке $\Delta\lambda/\lambda = 5\%$, на волне 21 см — замедление $c/v_{гр} = 400$ на участке $\Delta\lambda/\lambda = 10\%$.

Замедляющая система должна быть согласована с входной и выходной передающими линиями. В качестве линий передачи в дециметровом диапазоне и длинноволновой части сантиметрового диапазона используются коаксиальные линии, в коротковолновой части сантимет-

рового и миллиметровом диапазоне — волноводы. Согласование замедляющей системы с коаксиальной линией осуществляется подбором формы и положения штыря связи, являющегося продолжением центральной жилы коаксиальной линии по отношению к первому штырю замедляющей системы. Согласование волноводов с замедляющей системой осуществляется с помощью плавного волноводного перехода.

Полоса пропускания парамагнитного усилителя бегущей волны определяется полосой пропускания замедляющей системы и полосой электронного парамагнитного резонанса. Так как полоса электронного парамагнитного резонанса меньше полосы замедляющей системы, то возможна перестройка усилителя в пределах полосы замедляющей системы за счет изменения величины и направления постоянного магнитного поля.

Эффективный способ расширения полосы пропускания состоит в расширении резонансной линии парамагнитного резонанса $\Delta\omega_{\text{л}}$ [см. формулу (5-10)] за счет распроделенной расстройки парамагнитного резонанса, при которой центральные частоты линии парамагнитного резонанса в различных участках замедляющей системы сдвинуты относительно друг друга из-за неоднородности постоянного магнитного поля вдоль или поперек замедляющей системы.

Целесообразность применения парамагнитных усилителей, несмотря на сложность их конструкции и эксплуатации, определяется низким уровнем их собственных шумов. Источниками этих шумов являются спонтанное излучение активной среды и тепловое излучение элементов усилителя (см. § 3-6). Спонтанное излучение представляет собой принципиально неустраняемый источник шума и, таким образом, определяется предельно достигаемый уровень шумов.

Обычно шумы усилителей оценивают коэффициентом шума $K_{\text{ш}}$, определяемым выражением (3-64). Но малошумящие усилители удобнее характеризовать шумовой температурой в градусах Кельвина, определяемой формулой (3-61) и связанной с коэффициентом шума соотношением (3-67).

В диапазоне СВЧ спонтанное излучение мало и обусловленная им эквивалентная шумовая температура составляет 0,5—5 К. Тепловое излучение элементов усилителя обуславливает основную долю шумов квантового

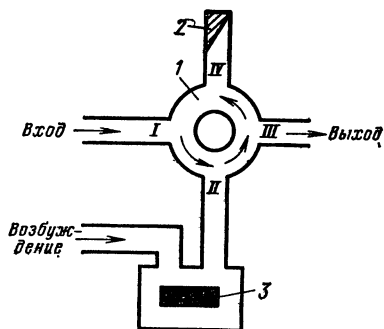
усилителя, но при глубоком охлаждении (до температуры жидкого гелия) ее можно резко уменьшить. Шумовая температура усилителей бегущей волны обычно менее 15 К, что на два порядка меньше уровня шумов наиболее чувствительных электронных усилителей СВЧ диапазона.

5-3. ПАРАМАГНИТНЫЕ РЕГЕНЕРАТИВНЫЕ УСИЛИТЕЛИ

Время взаимодействия электромагнитной волны с активной средой можно увеличить, применяя резонатор. Отражаясь от стенок резонатора, волна многократно проходит через активную среду и время взаимодействия поля со средой возрастает во много раз по сравнению с распространением в свободном пространстве.

Рис. 5-6. Квантовый парамагнитный регенеративный усилитель отражательного типа.

1 — циркулятор; 2 — поглощающая нагрузка; 3 — парамагнетик.



Существует два типа резонаторных усилителей: отражательный и проходной. Основные элементы этих усилителей одинаковы и они отличаются лишь способами разделения входного и выходного сигналов.

В усилителе отражательного типа входной и выходной сигналы поступают в один и тот же волновод (рис. 5-6) и для разделения их применяют ферритовый циркулятор.

Применяют четырех- и трехплечие циркуляторы. По усилительным свойствам усилители с такими циркуляторами одинаковы, но уровень шумов в схемах с трехплечим циркулятором выше. Сигнал входит в плечо I циркулятора и благодаря невзаимным свойствам подмагниченного феррита направляется в плечо II и в резонатор, в котором размещен парамагнетик. Если под действием источника возбуждения парамагнетик

находится в состоянии с инверсной населенностью, то сигнал усиливается и через циркулятор попадает в плечо *III*. Благодаря своим невзаимным свойствам циркулятор разделяет вход и выход резонатора, превращая его из двухполюсника в четырехполюсник. К выходному плечу *III* подсоединяется последующий каскад усиления. При этом шумы последующего каскада не влияют на первый, так как они через циркулятор попадают

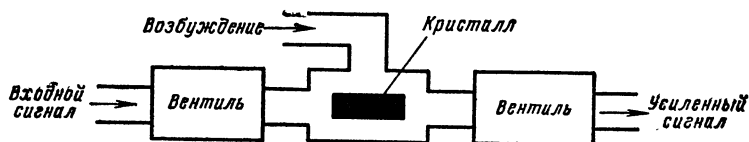


Рис. 5-7. Квантовый парамагнитный регенеративный усилитель проходного типа.

в балластную нагрузку *IV* и там поглощаются. Однако сама балластная нагрузка является источником шумов, которые идут в антенну и, частично отражаясь от нее, попадают в резонатор первого каскада и усиливаются вместе с сигналом. Но уровень шумов нагрузки невелик и за счет охлаждения нагрузки его можно существенно понизить. В схеме с трехплечим циркулятором шумы последующего каскада поступают в антенну и, отражаясь от нее, попадают в резонатор первого каскада и усиливаются. Мощность же шума последующего каскада значительно больше, чем мощность шума нагрузки, поэтому уровень шума усилителя с трехплечим циркулятором больше, чем с четырехплечим.

Усилитель проходного типа (рис. 5-7) имеет разделенные вход и выход. Во избежание самовозбуждения на входе и выходе включены ферритовые вентили. Усилитель этого типа имеет меньшие коэффициент усиления и полосу пропускания и больший уровень шумов, чем усилитель отражательного типа, на практике он применяется редко.

Основным элементом усилителя является резонатор с парамагнитной средой. Резонатор осуществляет обратную связь через активную среду. Входной сигнал вызывает индуцированные переходы, сигнал, возникаю-

щий в результате этих переходов, вызывает последующие индуцированные переходы. Во избежание самовозбуждения обратная связь, создаваемая резонатором, должна быть меньше пороговой.

Интенсивность вынужденных переходов, определяющих коэффициент усиления, зависит от добротности резонатора. Чем больше добротность резонатора, тем больше напряженность поля в резонаторе, больше интенсивность переходов, а следовательно, и коэффициент усиления. Но при увеличении добротности порога пропускания уменьшается, поэтому резонаторные усилители с большим коэффициентом усиления узкополосны.

Интенсивность вынужденных переходов возрастает не только при увеличении добротности, но особенно резко при парамагнитном резонансе. Поэтому резонатор одновременно должен обеспечить необходимую добротность и выполнение условий парамагнитного резонанса. Кроме того, поскольку частота рабочего перехода и частота возбуждения не совпадают, резонатор должен одновременно работать на этих частотах. Добротность резонатора на частоте возбуждения определяет требуемую мощность возбуждения, которая тем меньше, чем больше добротность резонатора на этой частоте.

Выбор типа резонатора определяется рабочим диапазоном волн и свойствами парамагнитного вещества. В дециметровом диапазоне и длинноволновой части сантиметрового диапазона используют резонаторы коаксиального или полоскового типа. В сантиметровом диапазоне применяют объемные прямоугольные резонаторы. Обычно это прямоугольные кристаллы рубина, покрытые тонким слоем серебра. Возбуждение и усиление в объемных резонаторах осуществляется на двух различных типах колебаний. Часто применяют комбинированные резонаторы (рис. 5-8), состоящие из прямо-

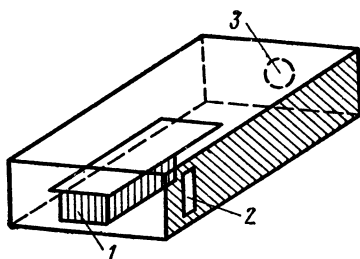


Рис. 5-8. Комбинированный резонатор.

1 — парамагнетик; 2 — ввод сигнала; 3 — ввод энергии возбуждения.

угольного объемного резонатора, настроенного на частоту возбуждения и полоскового резонатора, являющегося собственно резонатором усилителя и настроенного на частоту рабочего перехода. Сигнал и энергия возбуждения подводятся через отдельные элементы связи.

В миллиметровом диапазоне применяют диэлектрические резонаторы, так как используемый в этом диапазоне рутил обладает большой диэлектрической проницаемостью. Небольшой кристалл рутила обладает свойствами резонатора и без металлической оболочки.

Добротность резонатора, как известно, определяется выражением

$$Q_p = \frac{\omega_p W_p}{P_0}, \quad (5-12)$$

где ω_p — резонансная частота; W_p — энергия, запасенная при резонансе; P_0 — мощность потерь.

Мощность потерь удобно представить в виде суммы

$$P_0 = P_{\text{собр}} + P_m + P_{\text{изл}}, \quad (5-13)$$

где $P_{\text{собр}}$ — мощность, определяемая потерями в стенках резонатора и диэлектрическими потерями в среде, заполняющей резонатор (собственные потери); P_m — мощность магнитных потерь; $P_{\text{изл}}$ — мощность излучения из резонатора.

Подставляя (5-13) в (5-12), получаем:

$$\frac{1}{Q_p} = \frac{1}{Q_n} = \frac{1}{Q_0} + \frac{1}{Q_m} + \frac{1}{Q_{\text{вн}}}, \quad (5-14)$$

где Q_n — нагруженная добротность; Q_0 , Q_m , $Q_{\text{вн}}$ — частичные добротности;

$$Q_0 = \frac{\omega_p W_p}{P_{\text{собр}}}$$

— добротность, определяемая потерями в стенках и диэлектрическими потерями;

$$Q_{\text{вн}} = \frac{\omega_p W_p}{P_{\text{изл}}}$$

— внешняя добротность;

$$Q_m = \frac{\omega_p W_p}{P_m} \quad (5-15)$$

— магнитная добротность.

Магнитные потери в парамагнетике имеют резонансный характер. Если среда пассивная, то при парамагнитном резонансе наблюдается поглощение

$$P_m > 0 \text{ и } Q_m > 0.$$

Если в среде создана инверсная населенность, то наблюдается излучение,

$$P_m < 0 \text{ и } Q_m < 0.$$

Очевидно, условие генерирования (3-7) для парамагнитного генератора будет иметь вид:

$$-P_m > P_0 + P_{\text{изл}}$$

или

$$-\frac{1}{Q_m} > \frac{1}{Q_0} + \frac{1}{Q_{\text{вн}}},$$

а условие усиления в регенеративном режиме

$$P_0 < -P_m < P_0 + P_{\text{вн}}$$

или

$$\frac{1}{Q_0} < -\frac{1}{Q_m} < \frac{1}{Q_0} + \frac{1}{Q_{\text{вн}}}. \quad (5-16)$$

Определим магнитную добротность согласно формуле (5-15). В этой формуле W_p и P_m определяются выражениями

$$W_p = \int_{V_p} \frac{\mu'_a H_m^2}{2} dV; \quad (5-17)$$

$$P_m = \omega_p \int_{V_{\text{кр}}} \frac{\mu''_a H_m^2}{2} dV, \quad (5-18)$$

где интегрирование производится по объему резонатора V_p [формула (5-17)] и объему кристалла $V_{\text{кр}}$ [формула (5-18)]; μ'_a и μ''_a определяются выражениями (2-45) при резонансе ($\omega = \omega_p = \omega_l$).

Мощность магнитных потерь при парамагнитном резонансе согласно формуле (5-18) с учетом (2-45) и (5-4)

$$P_m = \frac{\omega_p \mu_0^2 \mu_B^2 \Delta N T_2}{\hbar} \int_{V_{\text{кр}}} H_m^2 dV. \quad (5-19)$$

Подставляя выражения (5-19) и (5-17) в формулу (5-15), получаем:

$$Q_m = \frac{\hbar}{\mu_0 \mu_B^2 \Delta N T_2 \xi_p}, \quad (5-20)$$

где

$$\xi_p = \frac{\int_{V_{кр}} H_m^2 dV}{\int_V H_m^2 dV}.$$

— коэффициент заполнения резонатора ($\xi_p \leq 1$).

Чтобы среда обладала усилительными свойствами, из условия (5-16) следует, что магнитная добротность должна быть как можно меньше. Согласно выражению (5-20) уменьшение магнитной добротности может быть получено за счет увеличения инверсной населенности, коэффициента заполнения резонатора и времени спин-спиновой релаксации.

Магнитная добротность, определяющая работу усилителя, характеризуется в основном свойствами активного вещества, а не резонатора. Единственная характеристика резонатора, существенно влияющая на работу усилителя — это коэффициент заполнения ξ_p . Желательно, чтобы коэффициент заполнения был как можно ближе к единице, т. е. активное вещество по возможности полностью заполняло резонатор. Типичное значение магнитной добротности $Q_m \approx 100$.

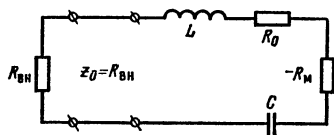


Рис. 5-9. Эквивалентная схема квантового парамагнитного регенеративного усилителя отражательного типа.

Так как магнитная добротность является основным параметром, характеризующим работу усилителя, то можно воспользоваться теорией электрических цепей. Эквивалентная схема усилителя приведена на рис. 5-9. Она состоит из внешней нагрузки с сопротивлением $R_{вн}$ согласованной передающей линии с волновым сопротивлением $Z_0 = R_{вн}$ и резонансного контура с отрицательным сопротивлением R_m , связанным с излучением

вещества, и сопротивлением R_0 , связанным с диэлектрическими потерями в веществе и потерями в стенках резонатора.

Коэффициент усиления по мощности отражательно-однорезонаторного усилителя определяется отношением

$$G = \frac{P_{\text{отр}}}{P_{\text{пад}}} = |\Gamma_u|^2, \quad (5-21)$$

где $P_{\text{отр}}$ — мощность, отраженная от резонатора; $P_{\text{пад}}$ — мощность, поступающая на вход резонатора; Γ_u — коэффициент отражения по напряжению.

Коэффициент отражения по напряжению согласно теории длинных линий определяется выражением

$$\Gamma_u = \frac{Z_H - Z_0}{Z_H + Z_0}.$$

В рассматриваемом случае

$$\begin{aligned} Z_H &= -R_M + R_0 + j\left(\omega L - \frac{1}{\omega C}\right) = \\ &= -R_M + R_0 + 2j\omega_p L \frac{\Delta\omega}{\omega_p}, \end{aligned}$$

где

$$\omega_p = \frac{1}{\sqrt{LC}}, \quad \Delta\omega = \omega - \omega_p.$$

Таким образом,

$$\Gamma_u = \frac{R_{\text{вн}} + R_M - R_0 - 2j\omega_p L \frac{\Delta\omega}{\omega_p}}{R_{\text{вн}} - R_M + R_0 + 2j\omega_p L \frac{\Delta\omega}{\omega_p}},$$

или

$$\Gamma_u = \frac{\frac{1}{Q_{\text{вн}}} - \frac{1}{Q_M} - \frac{1}{Q_0} - 2j \frac{\Delta\omega}{\omega_p}}{\frac{1}{Q_{\text{вн}}} + \frac{1}{Q_M} + \frac{1}{Q_0} + 2j \frac{\Delta\omega}{\omega_p}},$$

где

$$Q_{\text{вн}} = \frac{\omega_p L}{R_{\text{вн}}}, \quad Q_0 = \frac{\omega_p L}{R_0}, \quad Q_M = \frac{\omega_p L}{-R_M}.$$

Согласно (5-21) коэффициент усиления по мощности определяется выражением

$$G = \frac{\left(\frac{1}{Q_{\text{вн}}} - \frac{1}{Q_{\text{м}}} - \frac{1}{Q_0}\right)^2 + 4\left(\frac{\Delta\omega}{\omega_p}\right)^2}{\left(\frac{1}{Q_{\text{вн}}} + \frac{1}{Q_{\text{м}}} + \frac{1}{Q_0}\right)^2 + 4\left(\frac{\Delta\omega}{\omega_p}\right)^2}. \quad (5-22)$$

При резонансе ($\omega = \omega_p$) коэффициент усиления максимален

$$G_{\text{макс}} = \frac{\left(\frac{1}{Q_{\text{вн}}} - \frac{1}{Q_{\text{м}}} - \frac{1}{Q_0}\right)^2}{\left(\frac{1}{Q_{\text{вн}}} + \frac{1}{Q_{\text{м}}} + \frac{1}{Q_0}\right)^2}. \quad (5-23)$$

Для усиления необходимо, чтобы $G_{\text{макс}} > 1$, что выполняется, если

$$-\frac{1}{Q_{\text{м}}} > \frac{1}{Q_0}.$$

Это условие легко выполняется, так как даже без применения специальных мер добротность резонатора при температуре жидкого гелия $Q_0 > 1000$, а у хорошо выполненных резонаторов $Q_0 \approx 20\,000$.

Чтобы получить значительный коэффициент усиления (около тысячи), необходимо, чтобы усилитель работал вблизи порога самовозбуждения, и

$$-\frac{1}{Q_{\text{м}}} \approx \frac{1}{Q_{\text{вн}}} + \frac{1}{Q_0}.$$

При

$$-\frac{1}{Q_{\text{м}}} = \frac{1}{Q_{\text{вн}}} + \frac{1}{Q_0}$$

коэффициент усиления $G_{\text{макс}} \rightarrow \infty$ и начинается самовозбуждение.

При большом $G_{\text{макс}}$ вторым слагаемым в числителе формулы (5-22) можно пренебречь и представить эту формулу в виде

$$G = G_{\text{макс}} \frac{1}{1 + a^2},$$

где $a = 2Q_{\text{н}} \frac{\Delta\omega}{\omega_p}$ — обобщенная настройка; $Q_{\text{н}}$ — нагруженная добротность, определяемая выражением (5-14).

Зависимость коэффициента усиления от частоты совпадает по форме с частотной характеристикой обычного контура. Полоса пропускания усилителя определяется нагруженной добротностью резонатора, так как ширина линии рабочего перехода гораздо шире. Как и для обычного контура, полоса пропускания определяется по половинному уровню мощности

$$G(\omega_1) = G(\omega_2) = \frac{G_{\text{макс}}}{1 + a_{1,2}^2} = \frac{G_{\text{макс}}}{2};$$

отсюда

$$a_{2,1} = \pm 1$$

и ширина полосы пропускания определяется выражением

$$\omega_2 - \omega_1 = \frac{\omega_p}{Q_n} = \omega_p \left(\frac{1}{Q_{\text{вн}}} + \frac{1}{Q_m} + \frac{1}{Q_0} \right). \quad (5-24)$$

Поскольку величины $1/Q_m$ и $1/Q_{\text{вн}} + 1/Q_0$ почти полностью компенсируют друг друга, полоса пропускания усилителя очень узкая. Полоса пропускания связана с коэффициентом усиления. При $-\frac{1}{Q_m} \rightarrow \frac{1}{Q_{\text{вн}}} + \frac{1}{Q_0}$ полоса пропускания уменьшается, а коэффициент усиления увеличивается. Согласно выражениям (5-23) и (5-24) эффективность усилителя

$$F = \sqrt{G_{\text{макс}}}(\omega_2 - \omega_1) = \omega_p \left(\frac{1}{Q_{\text{вн}}} - \frac{1}{Q_m} - \frac{1}{Q_0} \right) = \text{const.}$$

Для данной конструкции эффективность усилителя не изменяется при изменении полосы пропускания $\omega_2 - \omega_1$. Практически получены значения эффективности около 200 МГц. С повышением температуры эффективность уменьшается.

Увеличение полосы пропускания при сохранении постоянным коэффициента усиления достигается в многорезонаторных схемах усилителей. Активное вещество обычно располагают в одном из резонаторов. В противном случае резко возрастают трудности, связанные с необходимостью одновременного возбуждения нескольких резонаторов с парамагнетиком, точной настройкой их резонансных частот, регулировкой связи и т. п.

Практически в многорезонаторной системе парамагнетиком заполняют последний резонатор, а цепочка пустых резонаторов представляет собой фильтр, расширяющий полосу пропускания. Такие усилители имеют полосу пропускания, близкую к полосе пропускания усилителя бегущей волны, но габариты их меньше и конструкция проще.

5.4. ГЕНЕРАТОР НА ПУЧКЕ МОЛЕКУЛ АММИАКА

В молекулярных генераторах инверсная населенность создается посредством сортировки частиц, находящихся в различных энергетических состояниях, с помощью неоднородных статических электрических или магнитных полей.

В качестве рабочего вещества используются пучки молекул аммиака, атомарного водорода, формальдегида и синильной кислоты.

Молекула аммиака NH_3 имеет вид тетраэдра. Три атома водорода расположены в вершинах равностороннего треугольника, служащего основанием тетраэдра, а атом азота — в вершине тетраэдра. В спектре аммиака

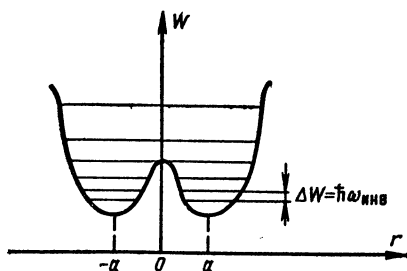


Рис. 5-10. Зависимость потенциальной энергии молекулы U от расстояния r атома азота от плоскости атомов водорода.

ка в диапазоне СВЧ наибольшая мощность излучения наблюдается для спектральной линии $f=23\,870$ МГц ($\lambda=1,25$ см). Излучение на этой частоте связано с колебательными движениями молекулы, при которых атом азота колеблется относительно плоскости атомов водорода. Эти колебания называются инверсионными.

Зависимость потенциальной энергии молекулы U от расстояния r атома азота до плоскости атомов водорода приведена на рис. 5-10 и имеет вид симметричной кривой с двумя минимумами ($r=\pm a$), которые соответствуют двум крайним положениям атома азота по разные стороны от плоскости атомов водорода. Максимум кривой при $r=0$ соответствует положению атома азота в плоскости атомов водорода. Потенци-

альный барьер, разделяющий минимумы, достаточно узок; и вероятность обнаружения частицы по ту или другую сторону барьера, т. е. переход атома азота через плоскость атомов водорода, имеет конечную величину. Поэтому каждый энергетический колебательный уровень расщепляется на два инверсионных подуровня, расстояние между которыми возрастает по мере перехода к более высоким колебательным энергетическим состояниям.

Молекула аммиака имеет электрический дипольный момент, так как вследствие различия электронных оболочек атомов водорода и азота «центр тяжести» отрицательных зарядов электронов в молекуле не совпадает с «центром тяжести» положительных зарядов ядер. Ориентация дипольного момента меняется на противоположную при инверсии атома азота относительно плоскости атомов водорода.

При термодинамическом равновесии распределение частиц по энергетическим уровням подчиняется распределению Больцмана (П1, п. 4).

Основное назначение молекулярных генераторов — служить источниками высокостабильных колебаний. Поэтому принимаются все меры для устранения причин, приводящих к уширению спектральной линии (см. П1, п. 20). Это достигается созданием направленного молекулярного пучка и понижением давления в нем до нескольких миллиметров ртутного столба. При этом устраняется влияние эффекта Допплера и уменьшается число столкновений между частицами и со стенками генератора.

Устройство молекулярного генератора схематически представлено на рис. 5-11. Его основные элементы: источник пучка молекул, сортирующая система и резонатор. Все эти элементы расположены в кожухе, охлаждаемом жидким гелием. В кожухе создается вакуум около 1 мПа.

Молекулы из источника через специальную решетку выходят в рабочее пространство. Решетка формирует пучек невзаимодействующих между собой молекул за счет подбора размеров отверстий решетки и расстояний между ними, а также давления газа в источнике. Расход источника — около 10^{18} молекул в секунду. Для компенсации убыли газа источник подсоединен к внешнему резервуару.

Инверсная населенность в пучке создается посредством пространственной сортировки частиц, находящихся в различных энергетических состояниях, с помощью неоднородного электрического поля. Сортирующая система представляет собой квадрупольный конденсатор, состоящий из четырех стержневых электродов (рис. 5-11, *а*), расположенных симметрично относительно общей оси на равном расстоянии друг от друга. Одна пара электродов (через один) подсоединена к плюсу

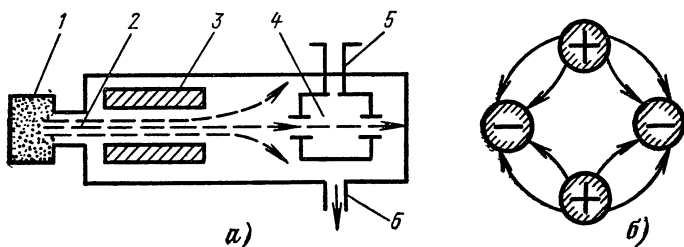


Рис. 5-11. Молекулярный генератор.

а — схема (1 — источник; 2 — пучок молекул; 3 — квадрупольный конденсатор; 4 — резонатор; 5 — вывод энергии; 6 — патрубок вакуумного насоса); *б* — поле в квадрупольном конденсаторе.

постоянного источника напряжения, другая пара — к минусу. Поле в конденсаторе аксиально-симметричное (рис. 5-11, *б*). На оси системы оно минимально и возрастает по мере удаления от оси конденсатора.

По выходе из источника большая часть молекул пучка попадает в конденсатор. Молекула аммиака обладает электрическим дипольным моментом p_e и в электрическом поле E приобретает дополнительную потенциальную энергию

$$U = -p_e E.$$

В неоднородном поле E на частицу действует сила

$$F = -\text{grad } U = \text{grad } p_e E,$$

под действием которой частица смещается в положение с минимальной энергией и отклоняется от первоначальной траектории движения. Частицы, у которых дипольный момент ориентирован параллельно вектору E , находятся на нижнем энергетическом уровне и по мере прохождения квадрупольного конденсатора отклоняются

к периферии, где поле максимально, а энергия минимальна. Частицы, находящиеся на верхнем энергетическом уровне, имеют антипараллельную ориентацию дипольного момента полю и смещаются к центру системы. Таким образом, вдоль оси сортирующей системы образуется пучок молекул с инверсной населенностью, так как в этом пучке больше молекул, находящихся на верхнем энергетическом уровне, чем на нижнем. Одновременно с сортировкой происходит и фокусировка пучка частиц, находящихся на верхнем энергетическом уровне, так как неоднородное поле прижимает эти частицы к оси системы. Большая часть частиц, находящихся на верхнем энергетическом уровне, через диафрагму попадает в резонатор.

Требования к генератору определяются получением наибольшей мощности, монохроматичности и стабильности излучения. Этим требованиям наиболее полно удовлетворяют резонаторы с равномерным распределением поля вдоль оси. К ним относится цилиндрический резонатор, в котором возбуждается основной тип колебания E_{010} . Электрическое поле этого типа имеет только одну составляющую по оси резонатора, совпадающую с направлением движения молекул, что обеспечивает минимальное влияние эффекта Допплера на уширение линии излучения.

Основными характеристиками резонатора являются собственная частота и добротность. Резонансная частота цилиндрического резонатора радиуса a и длиной l для колебаний типа E_{nmp} определяется выражением

$$\omega_{nmp} = c \sqrt{\frac{A_{nm}^2}{a^2} + \frac{\pi^2 p^2}{l^2}},$$

где A_{nm} — m -й корень уравнения $J_n(x) = 0$; $J_n(x)$ — функция Бесселя n -го порядка; p — число полуволн, укладывающихся вдоль длины резонатора.

Для колебания E_{010} резонансная частота не зависит от длины резонатора, а определяется лишь его радиусом:

$$\omega_{010} = \frac{c A_{01}}{a} = \frac{2,405c}{a}.$$

Для эффективной работы молекулярного генератора необходимо, чтобы собственная частота резонатора сов-

падала с частотой рабочего перехода молекул. Точная настройка резонатора на частоту рабочего перехода осуществляется специальным подстроечным винтом.

Собственная добротность резонатора для колебаний E_{010} , определяемая потерями в стенках, равна:

$$Q_{010} = \frac{al}{d(a+l)},$$

где $d = \sqrt{\frac{2}{\omega \mu_0 \sigma}}$ — глубина проникновения тока в стенки резонатора.

Для входа и выхода молекулярного пучка на торцах резонатора имеются отверстия. Чтобы уменьшить излучения через эти отверстия и не исказить поле в резонаторе, отверстия выполняются в виде предельных волноводов, диаметр которых меньше 0,2 диаметра резонатора. Чтобы обеспечить достаточный приток молекул в резонатор, в торцах резонатора делают большое количество таких отверстий. Добротность таких резонаторов около 10^4 .

Применим общую теорию, изложенную в § 3-5, к молекулярному генератору. Теория квантовых генераторов СВЧ более проста и точна, чем теория лазеров, так как генераторы СВЧ работают на одном типе колебаний и распределение поля однородно вдоль оси резонатора.

Если частота резонатора ω_p и частота рабочего перехода ω_{21} не совпадают, то происходит затягивание частоты, определяемое выражением (3-45). Так как ширина линии рабочего перехода $\Delta\omega_{\text{л}}$ в диапазоне СВЧ много меньше ширины резонансной кривой резонатора $\Delta\omega_p$, то

$$(\omega - \omega_{21}) \ll (\omega_p - \omega),$$

т. е. происходит затягивание частоты генерирования контуром линии, и

$$\omega \approx \omega_{21} = \omega_{\text{л}}. \quad (5-25)$$

Основные уравнения для исследования молекулярного генератора — это векторные уравнения (3-41) и уравнения в скалярной форме (3-42) или (3-48). Для молекулярного генератора продольное и поперечное времена релаксации T_1 и T_2 определяются средним временем нахождения молекул в резонаторе

$$\tau = \frac{l}{v_{\text{пр}}} = T_1 = T_2,$$

где l — длина резонатора; $v_{\text{пр}}$ — наиболее вероятная скорость молекулы в продольном направлении.

Для молекулярного генератора типичен стационарный режим и в этом случае в уравнениях (3-48) можно положить:

$$\frac{dE_m}{dt} = \frac{dP_m^a}{dt} = \frac{d\Delta N}{dt} = \frac{d\varphi}{dt} = \frac{d\psi}{dt} = 0;$$

при этом система уравнений (3-48) принимает вид

$$\left. \begin{aligned} -\frac{\omega_p}{Q_p} E_m + \frac{\omega_L P_m^a}{\varepsilon_a} \sin \Phi &= 0; \\ -\frac{1}{\tau} P_m^a + \frac{\rho_e^2}{\hbar} \Delta N E_m \sin \Phi &= 0; \\ \frac{\Delta N - \Delta N_0}{\tau} + \frac{E_m P_m^a}{\hbar} \sin \Phi &= 0; \\ (\omega_L - \omega_p) + \left(\frac{\omega_L P_m^a}{2\varepsilon_a E_m} + \frac{\rho_e^2 \Delta N E_m}{\hbar P_m^a} \right) \cos \Phi &= 0, \end{aligned} \right\} \quad (5-26)$$

где $\Phi = \varphi - \psi$.

Согласно второму уравнению системы (5-26)

$$P_m^a = \frac{\rho_e^2}{\hbar} \tau \Delta N E_m \sin \Phi,$$

и с учетом этого соотношения система (5-26) будет иметь вид:

$$\left. \begin{aligned} -\frac{\omega_p}{Q_p} + \frac{\omega_L \rho_e^2}{\hbar \varepsilon_a} \tau \Delta N \sin^2 \Phi &= 0; \\ \frac{\Delta N - \Delta N_0}{\tau} + \frac{\rho_e^2}{\hbar^2} \tau \Delta N E_m^2 \sin^2 \Phi &= 0; \\ (\omega_L - \omega_p) + \left(\frac{\omega_L \rho_e^2}{2\hbar \varepsilon_a} \tau \Delta N \sin \Phi + \frac{1}{\tau \sin \Phi} \right) \times \\ &\times \cos \Phi = 0; \end{aligned} \right\} \quad (5-27)$$

отсюда

$$\operatorname{ctg} \Phi = \frac{\omega_p - \omega_L}{\frac{1}{\tau} + \frac{\omega_p}{2Q_p}} = \delta \tau,$$

где $\delta = \frac{\omega_p - \omega_l}{1 + \frac{\omega_p}{2Q_p} \tau}$ — величина, связанная с расстройкой

резонатора относительно линии рабочего перехода.

В результате получим выражение для стационарного значения инверсной населенности:

$$\Delta N = \frac{\hbar \epsilon_a}{Q_p \rho_e^2 \tau} [1 + (\delta \tau)^2], \quad (5-28)$$

где учтено соотношение (5-25), а также выражение для напряженности поля:

$$E_m^2 = \frac{\hbar Q_p}{\tau \epsilon_a} \left\{ \Delta N_0 - \frac{\hbar \epsilon_a}{Q_p \rho_e^2 \tau} [1 + (\delta \tau)^2] \right\}. \quad (5-29)$$

Генерирование ($E_m^2 \geq 0$) наблюдается при

$$\Delta N_0 \geq \frac{\hbar \epsilon_a}{Q_p \rho_e^2 \tau} [1 + (\delta \tau)^2].$$

Отсюда условие самовозбуждения, определяющее пороговую инверсную населенность, имеет вид:

$$\Delta N_{\text{опор}} = \frac{\hbar \epsilon_a}{Q_p \rho_e^2 \tau} [1 + (\delta \tau)^2]. \quad (5-30)$$

Пороговая инверсная населенность уменьшается при увеличении добротности резонатора, дипольного момента молекул, длины резонатора и уменьшении расстройки. При выполнении условия (5-30) мощность излучения частиц компенсирует мощность потерь в резонаторе. Мощность излучения согласно (5-29) возрастает при увеличении дипольного момента и времени взаимодействия молекулы с полем, которое пропорционально длине резонатора. Эффективность взаимодействия возрастает при уменьшении расстройки δ и максимальна при $\delta=0$. Мощность потерь определяется добротностью резонатора и уменьшается при увеличении добротности.

Пороговая инверсная населенность $\Delta N_{\text{опор}}$ для типичного молекулярного генератора на аммиаке составляет примерно $2,5 \cdot 10^7 \text{ см}^{-3}$.

Амплитуда поля в резонаторе может быть выражена через пороговую инверсную населенность. Согласно формулам (5-29) и (5-30) имеем:

$$E_m^2 = \frac{\hbar Q_p}{\tau \epsilon_a} (\Delta N_0 - \Delta N_{\text{опор}}). \quad (5-31)$$

Таким образом, амплитуда излучаемого поля определяется превышением инверсной населенности пучка молекул, влетающих в резонатор, над пороговой.

Стационарное значение инверсной населенности согласно выражениям (5-28) и (5-30) равно пороговому. Увеличение инверсной населенности пучка молекул, влетающих в резонатор, приводит к увеличению мощности излучения согласно формуле (5-31), стационарное же значение инверсной населенности постоянно. Мощность излучения генераторов на аммиаке составляет от 10^{-10} до 10^{-7} Вт.

При инверсной населенности меньше пороговой наблюдается переход из режима генерирования в режим усиления. Однако такой усилитель по своим характеристикам значительно уступает квантовым парамагнитным усилителям.

Более точно, чем выражением (5-25), частота генерирования определяется согласно (3-45) выражением

$$\omega \approx \left(\omega_{\text{л}} + \frac{\Delta\omega_{\text{л}}}{\Delta\omega_{\text{р}}} \omega_{\text{р}} \right) \frac{1}{1 + \frac{\Delta\omega_{\text{л}}}{\Delta\omega_{\text{р}}}}.$$

Так как $\frac{\Delta\omega_{\text{л}}}{\Delta\omega_{\text{р}}} \ll 1$, то

$$\omega \approx \left(\omega_{\text{л}} + \frac{\Delta\omega_{\text{л}}}{\Delta\omega_{\text{р}}} \omega_{\text{р}} \right) \left(1 - \frac{\Delta\omega_{\text{л}}}{\Delta\omega_{\text{р}}} \right)$$

или

$$\omega \approx \omega_{\text{л}} \left(1 - \frac{\Delta\omega_{\text{л}}}{\Delta\omega_{\text{р}}} \frac{\omega_{\text{л}} - \omega_{\text{р}}}{\omega_{\text{л}}} \right),$$

т. е. частота генерирования тем ближе к частоте линии перехода, чем точнее настроен резонатор на частоту линии перехода (чем меньше разность $\omega_{\text{л}} - \omega_{\text{р}}$). Влияние расстройки уменьшается при увеличении длины резонатора, так как

$$\Delta\omega_{\text{л}} = \frac{2}{\tau} = \frac{2}{l} v_{\text{пр}}.$$

Практически настройку частоты генерирования на частоту линии перехода возможно осуществить с точностью $10^{-10} - 10^{-11}$, что и обусловило применение молекулярного генератора как высокоточного стандарта частоты.

ЭЛЕКТРОННЫЕ ГЕНЕРАТОРЫ И УСИЛИТЕЛИ СВЧ ДИАПАЗОНА

6-1. УСИЛЕНИЕ ЭЛЕКТРОМАГНИТНОЙ ВОЛНЫ В ЭЛЕКТРОННЫХ ПОТОКАХ

Как уже было показано в § 1-4, при взаимодействии электромагнитного поля с электронным потоком энергия поля увеличивается, если поток тормозится полем, и уменьшается, если поток ускоряется полем. В первом случае энергия поля увеличивается за счет убыли кинетической энергии движущихся электронов, во втором случае энергия электронов возрастает, а энергия поля убывает.

В генераторах и усилителях СВЧ диапазона энергия постоянного тока (источника питания) преобразуется в энергию высокочастотных колебаний. При конструировании этих приборов большое внимание уделяется созданию условий, при которых достигается наиболее полное преобразование энергии.

Кинетическая энергия электронов, с помощью которых энергия постоянного тока источника преобразуется в энергию СВЧ колебаний, определяется выражением $W_{\text{кин}} = \frac{mv_0^2}{2}$, где m — масса электрона; v_0 — скорость движения.

В общем случае масса движущегося электрона зависит от его скорости:

$$m = \frac{m_0}{\sqrt{1 - \frac{v_0^2}{c^2}}},$$

но для большинства электронных приборов СВЧ

$$\left(\frac{v_0}{c}\right)^2 \ll 1$$

и релятивистской поправкой можно пренебречь.

Скорость электрона зависит от постоянного электрического поля, создаваемого источником питания, и определяется уравнением

$$eU_0 = \frac{mv_0^2}{2},$$

где e — заряд электрона; U_0 — разность потенциалов между точкой, в которой находится электрон, и точкой, где его скорость равна нулю; отсюда

$$v_0 = \sqrt{\frac{2e}{m} U_0} \approx 5,9 \cdot 10^5 \sqrt{U_0}. \quad (6-1)$$

Как известно, мощность излучения электрона при торможении определяется выражением

$$P_{\text{изл}} = \left(\frac{e\mu_0 a}{4\pi} \right)^2 \frac{1}{60} \int_0^\pi \frac{\sin^3 \vartheta}{\left(1 - \frac{v_0}{c} \cos \vartheta \right)^6} d\vartheta,$$

где a — отрицательное ускорение; ϑ — угол, образованный направлением излучения и вектором скорости движения электрона v_0 .

Наибольшее замедление, а следовательно, и наибольшая мощность излучения достигаются при параллельности направлений тормозящего электрического поля и движения электронов.

Скорость электрона v_0 по мере движения в тормозящем поле уменьшается, в результате чего при выходе электрона из пространства взаимодействия с тормозящим полем кинетическая энергия оказывается меньше, чем при входе в это пространство. Разность между этими значениями равна энергии излучения.

Торможение движущегося электрона наблюдается, если вблизи его траектории движения находится проводник. В последнем возникает индуцированное движение собственных зарядов, т. е. возникает наведенный переменный ток с собственной частотой колебаний, определяющейся геометрической формой и размерами проводящего тела. Это явление и лежит в основе работы генераторов и усилителей СВЧ.

Высокочастотное переменное поле будет тормозить электрон только в течение той половины периода, когда направление электрического поля совпадает с направлением движения электрона. При этом поле усиливается за счет излучения электрона. В следующую часть периода электрон будет ускоряться и энергия от поля будет переходить к электрону.

Если имеется непрерывный поток электронов, то обязательно часть электронов будет тормозиться высоко-

частотным полем, а часть ускоряться. При этом поток разбивается на группы электронов (рис. 6-1).

Процесс группировки электронов зависит от соотношения скорости электронов v_0 и фазовой скорости электромагнитной волны v_ϕ . Если $v_0 = v_\phi$ (рис. 6-1, а) и электроны не взаимодействуют с полем, то электроны будут перемещаться синхронно с волной и все время будут находиться в точках с одной и той же напряженно-

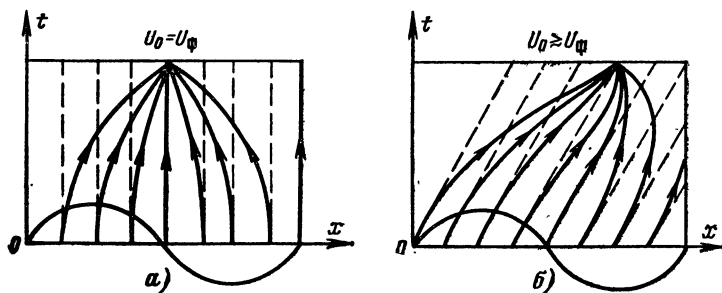


Рис. 6-1. Группировка электронов в сгустки в поле бегущей волны.

а — группировка в области $E=0$, где не происходит торможения электронов;
б — группировка в области максимального тормозящего поля E .

стью поля бегущей волны (пунктирные прямые на графике). При взаимодействии с полем электроны, находящиеся в области тормозящего поля, будут тормозиться и отставать от бегущей волны. Электроны, находящиеся в области ускоряющего поля, будут ускоряться и опережать волну. В результате электроны образуют сгустки в области, где продольное поле E равно нулю. При дальнейшем движении сгустков электронов скорость их остается неизменной, электроны и поле больше не взаимодействуют (см. § 1-4). Затраты энергии поля на группировку электронов малы, так как плотность потока электронов на входе в пространство взаимодействия с полем постоянна и в среднем поле замедляет и ускоряет одинаковое число электронов.

Как уже было показано в § 1-4, усиление электромагнитного поля при взаимодействии с потоком электронов происходит при условии $v_0 > v_\phi$. Действительно, если $v_0 > v_\phi$ (рис. 6-1, б), то при отсутствии взаимодействия электронов с полем они опережают волну (пунктирные прямые). При взаимодействии с волной электро-

ны, находящиеся в ускоряющем поле, будут ускоряться и перемещаться в область тормозящего поля. Таким образом, электроны группируются в области тормозящего поля и отдают свою энергию распространяющемуся полю.

Так как скорость электронов в процессе взаимодействия с полем будет уменьшаться, то сгусток по мере движения будет отставать от волны. Различие скоростей v_0 и v_ϕ должно быть таким, чтобы за время движения в пространстве взаимодействия электронов с полем сгусток не вышел из тормозящего поля.

Для эффективного взаимодействия сгустка электронов с полем необходимо, чтобы фаза распространяющегося поля в месте нахождения сгустка соответствовала максимальному значению тормозящего поля и перемещалась в пространстве синхронно со сгустком. Условие постоянства фазы высокочастотного поля по отношению к электронным сгусткам означает, что фазовая скорость v_ϕ волны должна быть приблизительно равна скорости электронов v_0 . Условие

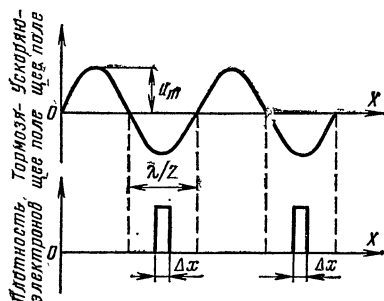


Рис. 6-2. Распределение высокочастотного поля и электронов в пространстве.

$$v_0 \gtrsim v_\phi \quad (6-2)$$

называют *условием синхронизма электронов и бегущей волны*. Но скорость электронов в устройствах СВЧ много меньше скорости света, поэтому практические устройства должны использовать замедляющие системы.

Электронные сгустки должны занимать как можно меньше места в пространстве и частота их следования должна равняться частоте поля. Однако на практике не удастся получить сгустки идеальной формы, изображенной на рис. 6-2. В интервалах между сгустками плотность потока электронов хотя и меньше, чем в сгустках, но не равна нулю. Получение возможно более коротких, близких к прямоугольной форме сгустков — одна из проблем электроники СВЧ. Требование, чтобы протя-

женность каждого сгустка электронов Δx в направлении движения электронов была много меньше половины длины волны, ограничивает возможности создания генераторов и усилителей СВЧ в коротковолновой части диапазона СВЧ.

Для обеспечения свободного движения электронного потока в этих приборах создается вакуум с давлением менее 10^{-4} Па, вследствие чего они относятся к электровакуумным приборам.

Анализ процессов в усилителях и генераторах СВЧ можно производить с помощью волнового уравнения. С учетом движущегося потока электронов согласно (1-6) волновое уравнение можно записать в виде

$$\operatorname{rot} \operatorname{rot} \mathbf{E} + \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = -\mu_0 \frac{\partial \mathbf{J}}{\partial t}, \quad (6-3)$$

где $\mathbf{J} = \mathbf{J}_{\text{конв}} = \rho \mathbf{v}_0$ — конвекционный ток (ток переноса); ρ — объемная плотность зарядов.

Правая часть уравнения (6-3) определяется уравнением движения электронов

$$\mathbf{F}_L = m \frac{d\mathbf{v}}{dt} = e \{ \mathbf{E} + \mu_0 [\mathbf{v}_0 \mathbf{H}] \}, \quad (6-4)$$

где \mathbf{F}_L — сила Лоренца.

Напряженности электрического и магнитного полей в выражении (6-4) могут иметь как постоянные, так и переменные составляющие. Однако практически учитывают переменную составляющую только электрического поля. Действительно, составляющие силы Лоренца

$$\mathbf{F}_L = \mathbf{F}_E + \mathbf{F}_H$$

находятся в отношении

$$\frac{|\mathbf{F}_H|}{|\mathbf{F}_E|} \leq \mu_0 v_0 \frac{|\mathbf{H}|}{|\mathbf{E}|},$$

но отношение

$$\frac{|\mathbf{H}|}{|\mathbf{E}|} = \frac{1}{Z_0} = \sqrt{\frac{\epsilon_0}{\mu_0}},$$

где Z_0 — волновое сопротивление свободного пространства, и

$$\frac{|\mathbf{F}_H|}{|\mathbf{F}_E|} \leq \frac{v_0}{c}.$$

Так как $v_0 \ll c$, то

$$|\mathbf{F}_H| \ll |\mathbf{F}_E|. \quad (6-5)$$

В системах, где происходит взаимодействие поля с потоком электронов, возбуждаются поля вида E , т. е. имеющие продольную составляющую электрического поля и обеспечивающие поэтому наибольшее взаимодействие с электронным потоком. Пренебрегая действием магнитного поля, в линейном приближении уравнение (6-4) можем переписать в виде

$$\frac{dv}{dt} = \frac{e}{m} E_z, \quad (6-6)$$

где E_z — продольная составляющая электрического поля, совпадающая с направлением движения электронов.

Система уравнений (6-3) и (6-4) должна быть дополнена граничными условиями для электрического и магнитного полей и начальными условиями — значениями начальных скоростей электронов на фиксированных поверхностях.

Кроме основных уравнений (6-3) и (6-4) важную роль при анализе работы электронных приборов СВЧ играет уравнение непрерывности

$$\operatorname{div} \mathbf{J} = \operatorname{div} (\rho \mathbf{v}) = - \frac{\partial \rho}{\partial t}. \quad (6-7)$$

6-2. ЛАМПЫ БЕГУЩЕЙ ВОЛНЫ ТИПА О

В генераторных и усилительных лампах бегущей волны (ЛБВ) взаимодействие высокочастотного поля бегущей волны с электронами происходит на участке протяженностью в несколько длин волн и носит непрерывный характер. Поэтому ЛБВ относят к классу приборов СВЧ с длительным взаимодействием электронов с электромагнитным полем. Для эффективного взаимодействия электронов с полем на участке взаимодействия необходимо соблюдение условия волнового синхронизма (6-2), т. е. приблизительного совпадения скорости движения электронов v_0 с фазовой скоростью волны v_{ϕ} . При этом предполагается, что направление движения электронов совпадает с направлением фазовой скорости волны.

Так как $v_0 \ll c$, а v_ϕ в свободном пространстве равна скорости света, то необходимо использование замедляющих систем. В отличие от парамагнитных усилителей бегущей волны (см. § 5-2), где замедляющие системы используются для уменьшения групповой скорости (увеличения замедления $c/v_{гр}$), в ЛБВ выполнение условия волнового синхронизма означает приблизительное совпадение скорости электронов с фазовой, а не групповой скоростью. Мало того, направление движения энергии в замедляющей системе ЛБВ, т. е. направление вектора групповой скорости, может быть противоположно движению электронов.

Лампы, в которых электронный поток взаимодействует с волной, распространяющейся в замедляющей системе в прямом направлении, называются *лампами прямой бегущей волны*. Обычно их называют просто *лампами бегущей волны* (ЛБВ). Так как направление движения энергии и электронов совпадает, вывод энергии у этих ламп расположен со стороны коллектора, собирающего отработанные электроны.

Лампы, в которых электронный поток взаимодействует с волной, энергия которой движется навстречу потоку электронов, называются *лампами обратной волны* (ЛОВ). В таких лампах вывод энергии расположен на конце замедляющей системы, обращенном к источнику электронного потока — электронной пушке.

Лампы как прямой, так и обратной волны подразделяются на типы О и М, отличающиеся направлением и назначением применяемого в лампах магнитного поля. В лампах типа О магнитное поле продольное, в лампах типа М — поперечное.

Замедляющие системы, используемые в ЛБВ и ЛОВ, представляют собой периодические структуры типа гребенки, одиночной или двойной спирали, диафрагмированного волновода и т. д.

Требуемое замедление c/v_ϕ определяется скоростью движения электронов и зависит от ускоряющего напряжения U_0 . Обычно $c/v_\phi \approx 10$.

Поле в периодической замедляющей структуре с периодом h должно удовлетворять периодическим граничным условиям. В системах без потерь это означает, что поле в соседних ячейках должно лишь отличаться сдвигом по фазе, т. е. множителем $e^{-j\beta_0(z+h)}$, где β_0 — постоянная распространения волны. Очевидно, поле $E(z)$

представляется периодической в пространстве функций, которую можно разложить в ряд Фурье

$$\dot{E}_m(z) = \left(\sum_{q=-\infty}^{\infty} \dot{E}_{mq} e^{-\frac{2\pi q}{h} z} \right) e^{-i\beta_0 z},$$

где

$$\dot{E}_{mq} = \frac{1}{h} \int_0^h \dot{E}_m(z) e^{i \frac{2\pi q}{h} z} dz,$$

т. е. представить в виде суммы бесконечного числа бегущих волн с амплитудами \dot{E}_{mq} и постоянными распространения

$$\beta_q = \beta_0 + \frac{2\pi q}{h}, \quad q = 0, \pm 1, \pm 2, \pm 3 \dots$$

Эти волны называются *пространственными гармониками*. Фазовые скорости гармоник определяются выражениями

$$v_{\phi q} = \frac{\omega}{\beta_q} = \frac{\omega}{\beta_0 + \frac{2\pi q}{h}}. \quad (6-8)$$

Чем больше номер гармоники, тем меньше ее скорость. Наибольшую фазовую скорость имеет нулевая пространственная гармоника, ее называют основной. Так как q принимает как положительные, так и отрицательные значения, то фазовая скорость в зависимости от знака q может быть как положительной, так и отрицательной. Фазовая скорость основной гармоники положительна.

Групповая скорость определяется выражением

$$v_{гр} = \frac{d\omega}{d\beta}$$

и не зависит от номера гармоники, поскольку

$$\frac{d\omega}{d\beta_1} = \frac{d\omega}{d\beta_2} = \dots = \frac{d\omega}{d\beta_0}.$$

Гармоники с одинаковыми направлениями фазовой и групповой скорости называются *прямыми*; если направления этих скоростей противоположны, гармоники называются *обратными*.

Фазовая скорость прямых гармоник может как уменьшаться с увеличением частоты (нормальная дисперсия), так и увеличиваться (аномальная дисперсия). Все обратные гармоники обладают аномальной дисперсией.

Важной характеристикой замедляющей системы является величина, эквивалентная волновому сопротивлению, имеющая размерность сопротивления и называемая *сопротивлением связи*:

$$R_{\text{св}(q)} = \frac{E_{mq}^2}{2\beta_q^2 P_0}, \quad (6-9)$$

где E_{mq} — амплитуда напряженности продольной составляющей электрического поля на оси системы; P_0 — поток мощности в системе.

Чем больше сопротивление связи, тем больше продольная составляющая напряженности электрического поля на оси системы при данном потоке мощности, тем эффективнее взаимодействие электронного потока с полем.

Условие синхронизма (6-2) можно выполнить для любой пространственной гармоники, регулируя скорость электронов изменением постоянного напряжения [см. формулу (6-1)]. Согласно (6-8) наибольшую фазовую скорость имеет основная гармоника, высшие гармоники имеют меньшую фазовую скорость. Поэтому для выполнения условия синхронизма для низших гармоник требуется большее напряжение U_0 . Однако обычно в ЛБВ и ЛОВ используются низшие гармоники (основная и первые прямая и обратная), так как сопротивление связи, определяемое выражением (6-9), для этих гармоник велико и взаимодействие электронов с полем этих гармоник наиболее эффективно.

Полоса пропускания ЛБВ и ЛОВ определяется дисперсионной характеристикой замедляющей системы. Чем слабее дисперсия, тем в большей полосе частот при фиксированном постоянном напряжении можно обеспечить синхронизм электронов и пространственной гармоники.

Наиболее часто применяются замедляющие системы типа спирали, имеющие полосу пропускания, составляющую 50—100% средней частоты и обеспечивающие широполосное согласование с входным и выходным

трактами. Шаг спирали на концах системы для обеспечения согласования плавно увеличивается.

Замедление волны осуществляется на однородном участке спирали. Электромагнитное поле распространяется вдоль витков спирали, и фазовая скорость волны в продольном направлении, очевидно, определяется выражением

$$v_{\phi} = \frac{ch}{\sqrt{(2\pi a)^2 + h^2}} \approx c \frac{h}{2\pi a}, \quad (6-10)$$

где a — радиус спирали; h — шаг спирали.

Поле замедленной волны быстро убывает по мере удаления от поверхности спирали и на расстоянии половины длины замедленной волны мало. Поэтому обычно

$$a \approx (0,2 \div 0,25) \lambda_{\text{зам}},$$

где $\lambda_{\text{зам}}$ — длина замедленной волны.

Как и для любой направляющей системы, в случае спирали

$$\lambda_{\text{зам}} = \lambda \frac{v_{\phi}}{c}$$

и

$$a \approx (0,2 \div 0,25) \lambda \frac{v_{\phi}}{c}.$$

Чем короче рабочая длина волны λ , на которую рассчитана ЛБВ или ЛОВ, тем меньше при заданном замедлении c/v_{ϕ} должен быть радиус спирали. В миллиметровом диапазоне из-за конструктивных трудностей, возникающих при уменьшении диаметра спирали, и уменьшения тока электронного пучка, который можно пропустить через спираль, использование спирали в качестве замедляющей системы нецелесообразно. Обычно в этом диапазоне применяют замедляющие системы типа гребенки.

Усилительная лампа бегущей волны типа О со спиральной замедляющей системой схематически изображена на рис. 6-3. Ускоряющее напряжение U_0 , приложенное между катодом и анодом, обеспечивает синхронизм между электронами и замедленной волной. Скорость электронов на входе замедляющей системы согласно выражению (6-1)

$$v_0 = \sqrt{\frac{2e}{m} U_0} \quad (6-11)$$

и с учетом условия синхронизма $v_{\phi} \approx v_0$ и выражения (6-10) приближенно (без учета дисперсии замедляющей системы) оптимальное ускоряющее напряжение определяется выражением

$$U_{0 \text{ опт}} = \frac{mc^2}{8\pi^2 e} \left(\frac{h}{a} \right)^2 \approx 0,65 \cdot 10^4 \left(\frac{h}{a} \right)^2.$$

При отклонении ускоряющего напряжения от оптимального значения выходная мощность ЛБВ резко падает.

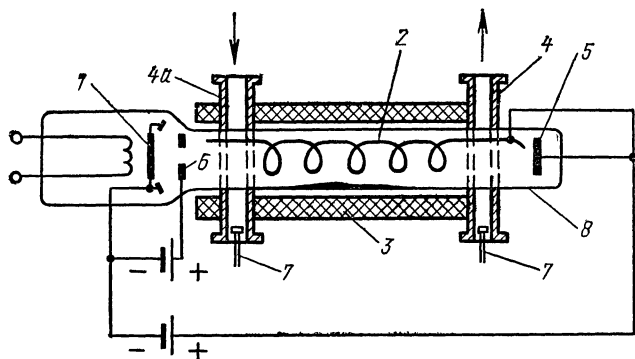


Рис. 6-3. Схема конструкции усилительной ЛБВ типа О.

1 — катод; 2 — спираль; 3 — соленоид; 4 — выход энергии СВЧ; 4а — вход усиливаемых колебаний СВЧ; 5 — коллектор; 6 — анод; 7 — согласующие устройства; 8 — оболочка лампы.

Фокусировка электронного потока осуществляется с помощью постоянного магнитного поля, направленного вдоль оси ЛБВ. Под действием этого поля электроны, движущиеся в продольном направлении, сжимаются в шнур, внутри которого расталкивающая сила преодолевается силой Лоренца. При этом траектория движения электронов имеет вид спирали.

В замедляющей системе ЛБВ происходит модуляция электронов по скорости и как следствие этого образование сгустков (групп) электронов. Электронные сгустки наводят в замедляющей спирали ток и создают тормозящее поле. Взаимодействуя с этим полем, электронные сгустки отдают энергию усиливаемому распространяющемуся полю.

Анализ работы ЛБВ в режиме усиления высокочастотного поля малой амплитуды (линейный режим) про-

ведем в предположении, что электроны движутся только в продольном направлении, высокочастотное поле и плотность электронов в пределах поперечного сечения пучка однородны.

Электромагнитное поле, распространяющееся внутри электронного потока, в символической форме согласно (6-3) описывается уравнением

$$\text{rot rot } \dot{\mathbf{E}} - k^2 \dot{\mathbf{E}} = -j\omega\mu_0 \mathbf{j}$$

или

$$\text{grad div } \dot{\mathbf{E}} - \Delta \dot{\mathbf{E}} - k^2 \dot{\mathbf{E}} = -jkZ_0 \mathbf{j}, \quad (6-12)$$

где $k = \omega \sqrt{\epsilon_0 \mu_0}$ и $Z_0 = \sqrt{\mu_0 / \epsilon_0}$ — постоянная распространения и волновое сопротивление плоской волны в вакууме.

Входящую в уравнение (6-12) величину $\text{grad div } \dot{\mathbf{E}}$ можно найти, если учесть, что

$$\text{div } \dot{\mathbf{E}} = \frac{\dot{\rho}}{\epsilon_0}$$

и согласно уравнению непрерывности (6-7)

$$\text{div } \mathbf{j} = -j\omega \dot{\rho}.$$

В цилиндрической системе координат, полагая, что поток однороден в поперечном сечении, движется и изменяется в направлении оси Z , получим:

$$\text{grad div } \dot{\mathbf{E}} = \frac{1}{\epsilon_0} \text{grad } \frac{\text{div } \mathbf{j}}{-j\omega} = \mathbf{e}_z \frac{jZ_0}{k} \frac{\partial^2 \mathbf{j}}{\partial z^2}. \quad (6-13)$$

Продольное электрическое поле бегущей волны в потоке электронов может быть представлено в виде

$$\dot{E}_z = E_{mz} e^{j(\omega t - \dot{k}z)}, \quad (6-14)$$

где \dot{k} — постоянная распространения с учетом электронного потока.

Подставляя выражения (6-13) и (6-14) в уравнение (6-12), для продольной составляющей напряженности электрического поля получаем:

$$(\dot{k}^2 - k^2) \dot{E}_{mz} = -jkZ_0 j_m - j \frac{Z_0}{k} \frac{\partial^2 j_m}{\partial z^2}. \quad (6-15)$$

В ЛБВ электронный пучок, движущийся по оси спиральной замедляющей системы, достаточно тонок и поле

в пределах поперечного сечения пучка можно считать однородным. При отсутствии потерь в замедляющей системе влияние ее учитывается заменой постоянной распространения в вакууме k на постоянную распространения в замедляющей системе без потерь $\beta = \omega/v_\phi$ и волнового сопротивления Z_0 на сопротивление связи $R_{св}$. Таким образом, уравнение (6-15) для ЛБВ будет иметь вид:

$$(\dot{k}^2 - \beta^2) \dot{E}_{mz} = -j\beta R_{св} j_m - j \frac{R_{св}}{\beta} \frac{\partial^2 j_m}{\partial z^2}. \quad (6-16)$$

Уравнение (6-16) связывает продольную составляющую поля в пучке \dot{E}_{mz} с переменной составляющей плотности тока, которая сама зависит от \dot{E}_{mz} , так как электроны движутся под действием этого поля. Эта связь определяется уравнением движения электронов (6-6)

$$\frac{dv}{dt} = \frac{e}{m} E_z.$$

Здесь пренебрегают разбросом электронов по скоростям и направлениям и учитывают взаимодействие электрона только с электрической составляющей поля, так как сила, обусловленная магнитной составляющей, согласно (6-5) мала.

Взаимодействие поля бегущей волны с электронным потоком и модуляция плотности этого потока рассмотрены в § 1-4. Согласно проведенному анализу плотность тока в электронном пучке определяется выражением

$$j_m = -j \frac{k_3 \frac{\rho_0}{v_0} \frac{e}{m}}{(\dot{k}_3 - \dot{k})^2} \dot{E}_{mz},$$

где величина $k_3 = \omega/v_0$.

С учетом выражения (6-11)

$$j_m = -j \frac{\rho_0 \omega}{2U_0 (\dot{k}_3 - \dot{k})^2} \dot{E}_{mz}. \quad (6-17)$$

Постоянная составляющая плотности тока в пучке

$$J_0 = -\rho_0 v_0.$$

С учетом этого соотношения выражение (6-17) можно представить в виде

$$j_m = j \frac{J_0 k_3}{2U_0 (\dot{k}_3 - \dot{k})^2} \dot{E}_{mz}. \quad (6-18)$$

При выполнении условия синхронизма

$$v_{\phi} \approx v_0, \quad k \approx k_z,$$

амплитуда волны плотности тока максимальна и значительно превосходит постоянную составляющую J_0 :

$$J_m \gg J_0,$$

т. е. происходит образование сгустков электронов большой плотности.

Согласно выражению (6-18) переменная составляющая плотности тока и продольная составляющая электрического поля сдвинуты по фазе во времени на $\pi/2$, т. е. сгустки образуются в тех частях замедляющей системы, где продольная составляющая электрического поля равна нулю. При приблизительном равенстве скоростей электронов и фазовой скорости волны образовавшиеся сгустки будут все время находиться в минимуме поля и усиления распространяющегося поля не будет. Чтобы поле усиливалось, скорость электронов должна несколько превышать скорость волны $v_0 \gtrsim v_{\phi}$.

При этом согласно выражению (6-18) образование сгустков будет продолжаться и они, обгоняя волну поля, будут перемещаться в область тормозящего поля. Здесь, теряя свою кинетическую энергию, электроны будут усиливать поле. Превышение скорости электронов над скоростью волны должно быть таким, чтобы сгустки не проскакивали участки с тормозящим полем и не оказывались на участках с ускоряющим полем, где электроны отбирают энергию от поля. Это осуществляется подбором оптимального ускоряющего напряжения $U_0 \text{ опт}$, при котором усиление ЛБВ максимально.

Выражение (6-18) определяет свойства электронного пучка под действием известного электрического поля. Кроме слабого усиливаемого поля на электроны действует электрическое поле, которое они сами создают. Сгустки электронов, образующиеся в начале замедляющей системы идвигающиеся в ней, наводят в спирали токи, которые в свою очередь являются источником, изменяющим поле усиливаемой волны. Другими словами, распространяющееся поле должно быть самосогласованным, т. е. при взаимодействии этого поля с электронным пучком последний изменяется таким образом, что наводимые им в спирали токи создают то же самое электромагнитное поле. Такое самосогласованное решение най-

дем, подставляя выражение (6-18) в уравнение (6-16). При этом получим

$$(\dot{k}^2 - \beta^2) \left(1 + \frac{J_0 k_3 R_{св}}{2U_0 (k_3 - \dot{k})^2 \beta} \right) \dot{E}_{mz} = 0,$$

или, поскольку $\dot{E}_{mz} \neq 0$,

$$(\dot{k}^2 - \beta^2) \left(1 + \frac{J_0 k_3 R_{св}}{2U_0 (k_3 - \dot{k})^2 \beta} \right) = 0. \quad (6-19)$$

Уравнение (6-19) называется *характеристическим или дисперсионным уравнением ЛБВ*. Корни этого уравнения определяют постоянную распространения \dot{k} . Поскольку это уравнение четвертого порядка, то \dot{k} должно иметь четыре различных значения

$$\begin{aligned} \dot{k}_{1,2} &= \pm \beta, \\ \dot{k}_{3,4} &= k_3 \pm j \sqrt{\frac{J_0 k_3 R_{св}}{2U_0 \beta}}, \end{aligned} \quad (6-20)$$

соответствующие четырем различным волнам.

Первые две волны, соответствующие \dot{k}_1 и \dot{k}_2 , представляют собой прямую и обратную волны в замедляющей системе, на распространение которых в принятом приближении электронный поток не оказывает влияния. Наличие обратной волны является одной из причин самовозбуждения ЛБВ. Третья и четвертая волны распространяются в прямом направлении, причем третья волна, постоянная распространения которой \dot{k}_3 имеет положительную мнимую часть, при распространении усиливается. Четвертая волна, постоянная распространения которой \dot{k}_4 имеет отрицательную мнимую часть, в процессе распространения поглощается.

Для работы ЛБВ в усилительном режиме важна третья волна, амплитуда которой при распространении растет по экспоненциальному закону

$$E_m(x) = E_{m0} e^{\alpha_{yc} z},$$

где согласно выражению (6-20),

$$\alpha_{yc} = \sqrt{\frac{J_0 k_3 R_{св}}{2U_0 \beta}}. \quad (6-21)$$

Коэффициент усиления по мощности на участке l определяется выражением

$$G(\text{дБ}) = 10 \lg \frac{1}{3} \frac{P(z=l)}{P(z=0)}.$$

Здесь множитель $1/3$ введен с учетом того, что мощность на входе примерно поровну распределяется между всеми тремя бегущими волнами. Таким образом,

$$G(\text{дБ}) \approx 8,6 \alpha_{yc} l - 9,5, \quad (6-22)$$

где α_{yc} определяется выражением (6-21).

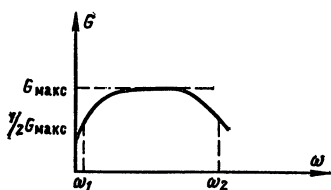


Рис. 6-4. Амплитудно-частотная характеристика ЛБВ.

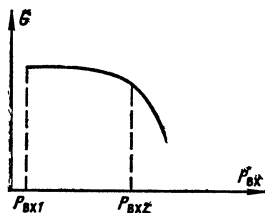


Рис. 6-5. Зависимость коэффициента усиления от входной мощности.

Согласно формуле (6-22) коэффициент усиления увеличивается при увеличении длины l пространства взаимодействия волны с электронным потоком, т. е. при увеличении длины замедляющей системы. Однако формула (6-22) справедлива только для линейного режима. В действительности коэффициент усиления ограничен возможностью самовозбуждения ЛБВ и тем, что по мере продвижения в замедляющей системе и взаимодействия с полем электроны теряют свою кинетическую энергию и условие работы ЛБВ (небольшое превышение скорости электронов над фазовой скоростью волны в замедляющей системе) нарушается и дальнейшее увеличение длины системы приводит не к увеличению, а к уменьшению усиления. Практически коэффициент усиления ЛБВ равен 30—40 дБ, т. е. достаточно высок.

Амплитудно-частотная характеристика ЛБВ приведена на рис. 6-4. Полоса пропускания ($\omega_2 - \omega_1$), как и для любого усилителя, определяется уменьшением коэффициента усиления по мощности вдвое [на 3 дБ]. Сни-

жение коэффициента усиления со стороны нижних частот обусловлено увеличением длины волны в замедляющей системе. При этом уменьшается отношение $1/\lambda_{\text{зам}}$, т. е. число волн, укладываемых на длине замедляющей системы, а следовательно, и число сгустков электронов, с которыми происходит взаимодействие распространяющегося поля. Снижение коэффициента усиления на верхних частотах обусловлено тем, что при повышении частоты увеличивается скорость спада амплитуды продольного поля, которое как бы прижимается к виткам спирали, и поле в месте нахождения электронного пучка уменьшается. При этом взаимодействие электронов с распространяющимся полем становится менее эффективным. Обычно полоса пропускания ЛБВ бывает очень широкой и соответствует приблизительно двукратному перекрытию рабочего диапазона по частоте.

Зависимость коэффициента усиления от уровня входного сигнала приведена на рис. 6-5. При малом уровне мощности на входе коэффициент усиления максимален и в пределах динамического диапазона остается практически постоянным. При дальнейшем увеличении входной мощности коэффициент усиления монотонно убывает. Ширина динамического диапазона ЛБВ достигает 80—90 дБ.

Коэффициент полезного действия ЛБВ составляет единицы процентов для ламп миллиметрового диапазона и 20—35% для ламп сантиметрового и дециметрового диапазонов.

В практике ЛБВ широко используются в диапазоне СВЧ. ЛБВ непрерывного действия малой мощности (единицы милливатт) применяются в радиоприемниках; ЛБВ средней мощности (от сотен милливатт до единиц ватт) и мощные ЛБВ (сотни ватт) — в радиопередатчиках.

Мощность импульсных ЛБВ составляет от нескольких сотен киловатт до нескольких мегаватт.

При разработке маломощных ЛБВ основное внимание уделяется снижению уровня шумов, ЛБВ средней мощности — увеличению коэффициента усиления, мощных ЛБВ — повышению уровня выходной мощности и к. п. д.

Шумы в ЛБВ усилителях характеризуются коэффициентом шума или эквивалентной шумовой температурой усилителя (см. § 3-6). Коэффициент шума на прак-

тике часто выражают не только в абсолютных величинах, но и в децибелах:

$$K_{ш} \text{ (дБ)} = 10 \lg K_{ш}.$$

Коэффициент шума ЛБВ невелик, около 3—6 дБ и менее. Это позволяет использовать ЛБВ в высокочувствительной приемной аппаратуре.

Основными источниками шумов являются флуктуации плотности конвекционного тока и скоростей электронов, эмитированных катодом, флуктуации тока в замедляющей системе за счет оседания электронов на замедляющей системе и ионные токи, возникающие за счет остаточных газов. На уровень шумов ЛБВ оказывают влияние также шумы, возникающие в замедляющей системе, определяемые температурой системы и высокочастотными потерями в ней.

ЛБВ могут самовозбуждаться, что недопустимо для усилительного режима. Положительная обратная связь, необходимая для самовозбуждения, возникает за счет отражения энергии от выхода ЛБВ и движения ее в обратном направлении. Условие самовозбуждения сводится к условию баланса фаз и условию баланса амплитуд (см. § 3-2). Условие баланса амплитуд выполняется, когда мощность, отдаваемая потоком электронов, компенсирует потери в замедляющей системе и в нагрузке. Условие баланса фаз

$$\frac{2\pi}{\lambda_{зам}} 2l + \varphi_1 + \varphi_2 = 2\pi q$$

определяет частоту генерирования. Для простоты положим фазы отражения на обоих концах замедляющей системы равными нулю. Тогда получим:

$$l = q \frac{\lambda_{зам}}{2}. \quad (6-23)$$

Число q , равное числу замедленных полуволн, укладываемых на длине замедляющей системы, определяет тип колебаний ЛБВ генератора. Частота этих колебаний равна:

$$f = \frac{v_{\phi}}{\lambda_{зам}} = \frac{c}{\lambda}. \quad (6-24)$$

Согласно условию синхронизма и выражению (6-11)

$$v_{\phi} \approx v_0 = \sqrt{2 \frac{e}{m} U_0}. \quad (6-25)$$

Подставляя (6-25) и (6-23) в выражение (6-24), получаем:

$$f = \frac{q}{l} \sqrt{\frac{eU_0}{2m}}. \quad (6-26)$$

Таким образом, ЛБВ может самовозбуждаться на ряде дискретных частот, определяемых выражением (6-26). Диапазон возможных частот генерирования ограничивается полосой пропускания усилителя. Со стороны высоких частот сильное влияние на возможность самовозбуждения лампы оказывают потери в замедляющей системе, которые возрастают с увеличением частоты. Поэтому чаще всего ЛБВ возбуждается на низких частотах.

Чтобы избежать самовозбуждения, в ЛБВ усилитель вводят распределенный или сосредоточенный поглотитель. При этом коэффициент усиления по мощности, определяемый выражением (6-22), равен:

$$G(\text{дБ}) = 8,6 \alpha_{yc} l - 9,5 - L(\text{дБ}),$$

где $L(\text{дБ})$ — коэффициент, характеризующий ослабление сигнала поглотителем.

Лампа бегущей волны как генератор. Если не принимать мер к устранению самовозбуждения, то ЛБВ усилитель может превратиться в генератор. Однако вследствие большой длины замедляющей системы частоты, определяемые выражением (6-26), лежат близко друг к другу и при работе генератора наблюдаются перескоки с одной частоты на другую.

Разрешение видов колебаний возможно путем сокращения длины l и введения внешней обратной связи в виде коаксиальной линии или волновода, соединяющих выход лампы со входом. В этом случае условие баланса фаз имеет вид:

$$2\pi \frac{l}{\lambda_{\text{зам}}} + 2\pi \frac{l_{\text{ос}}}{\lambda_{\text{ос}}} = 2\pi q,$$

где $\lambda_{\text{ос}}$ — длина волны в линии внешней обратной связи; $l_{\text{ос}}$ — длина линии обратной связи, или

$$\frac{\omega l}{v_{\text{ф}}} + \frac{\omega l_{\text{ос}}}{v_{\text{ф ос}}} = 2\pi q. \quad (6-27)$$

Так как фазовая скорость волны в ЛБВ зависит от частоты и скорости электронов v_0 , а следовательно, и от

напряжения U_0 , то согласно (6-27), изменяя величину U_0 , можно перестраивать частоту ЛБВ генератора (электронная перестройка). Диапазон электронной перестройки мал и составляет от долей процента до нескольких процентов от средней частоты.

Из-за малого диапазона перестройки ЛБВ генераторы практически не применяются.

Лампа обратной волны как генератор имеет широкий диапазон перестройки частоты и поэтому находит широкое применение. В ЛОВ электронный поток взаимодействует с полем первой обратной пространственной гармоники, фазовая скорость которой совпадает с направлением движения электронного потока и синхронизована с ним. Поглотитель, расположенный в замедляющей системе, препятствует возбуждению лампы на прямых волнах. Энергия в соответствии с направлением групповой скорости распространяется навстречу электронному потоку, что обеспечивает положительную обратную связь между полем волны и потоком электронов. Взаимодействуя с полем обратной гармоники, электроны группируются и отдают свою энергию полю волны. Амплитуда волны нарастает и поле еще более эффективно группирует электроны. При этом возрастает передача энергии волне и т. д. Так как сгустки электронов образуются вблизи нулей поля обратной пространственной гармоники, то это означает, что ток в пучке отстает от напряженности поля, образующего сгустки, на $\pi/2$. В свою очередь поле СВЧ, создаваемое сгустками электронов, отстает от тока пучка на $\pi/2$.

Стационарные колебания устанавливаются при условии выполнения условий баланса фаз и амплитуд. Баланс амплитуд сводится к тому, что энергия, получаемая полем от электронного потока, компенсирует потери в замедляющей системе и нагрузке. Фазовый сдвиг по петле, охватывающей электронный поток и замедляющую систему, должен составлять при балансе фаз целое число 2π . Геометрические длины замедляющей системы и электронного пучка одинаковы. Таким образом, условие баланса фаз можно представить в виде

$$\frac{\omega l}{v_{\Phi(-1)}} + \frac{\omega l}{v_0} - \pi = 2\pi q, \quad (6-28)$$

где $q=0, 1, 2, 3, \dots$; $v_{\Phi(-1)}$ — фазовая скорость первой обратной гармоники.

В выражении (6-28) первое слагаемое определяет набег фазы волны при ее движении вдоль замедляющей системы длиной l , второе слагаемое — набег фазы за время движения электронов вдоль замедляющей системы. Угол π характеризует смещение сгустка электронов относительно волны в замедляющей системе.

Для создания ЛОВ генератора с широким диапазоном электронной перестройки необходимо, чтобы условие (6-28) при фиксированном q выполнялось для любой частоты, т. е.

$$\frac{\partial}{\partial \omega} \left(\frac{\omega}{v_{\Phi(-1)}} + \frac{\omega}{v_0} \right) = 0$$

или

$$v_{\text{гр}(-1)} = -v_{\text{гр.эл}},$$

т. е. групповые скорости обратной гармоник и волны, возникающей вследствие группировки электронов и бегущей в противоположном направлении, должны быть равны по величине и противоположны по направлению. Это условие выполняется для ЛОВ генератора в широком диапазоне частот. Существуют ЛОВ, перестраиваемые в диапазоне от 2 до 8 ГГц. Мощность генерирования ЛОВ составляет от нескольких милливатт до единиц ватт, к. п. д. — несколько процентов.

6-3. ПРИБОРЫ МАГНЕТРОННОГО ТИПА

В основе работы приборов магнетронного типа лежит взаимодействие электронного потока с электромагнитной волной в скрещенных (перпендикулярных друг дру-

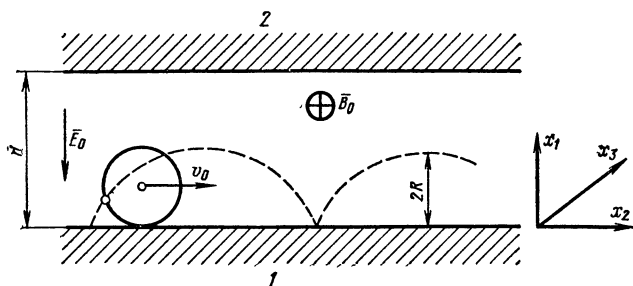


Рис. 6-6. Траектория электрона в скрещенных полях при плоских параллельных аноде и катоде.

1 — катод; 2 — анод.

гу) постоянных электрическом поле \mathbf{E}_0 и магнитном поле индукции \mathbf{B}_0 .

Рассмотрим плоскую конструкцию с заданной ориентацией постоянных полей $\mathbf{E}_0 = -e_1 E_0$ и $\mathbf{B}_0 = e_3 B_0$ (рис. 6-6). В статическом режиме, т. е. при отсутствии поля СВЧ, движение электрона определяется уравнением движения, которое согласно (6-4) имеет вид:

$$m \frac{d\mathbf{v}}{dt} = e (\mathbf{E}_0 + [\mathbf{v} \mathbf{B}_0])$$

или в проекциях на оси декартовой системы координат при заданных направлениях постоянных полей

$$\left. \begin{aligned} \frac{d^2 x_1}{dt^2} &= -\frac{e}{m} B_0 - \omega_{\text{ц}} \frac{dx_2}{dt}; \\ \frac{d^2 x_2}{dt^2} &= \omega_{\text{ц}} \frac{dx_1}{dt}; \quad \frac{d^2 x_3}{dt^2} = 0, \end{aligned} \right\} \quad (6-29)$$

где $\omega_{\text{ц}} = \frac{|e|}{m} B_0$ — циклотронная частота.

Пусть в начальный момент $t=0$ электрон находится в точке x_1^0, x_2^0, x_3^0 и имеет скорость $v_i^0 = (v_1^0, v_2^0, v_3^0)$. Интегрируя последнее уравнение системы (6-29), получаем:

$$v_3 = v_3^0, \quad x_3 = x_3^0 + v_3^0 t, \quad (6-30)$$

т. е. скорость электрона в направлении оси x_3 при заданном направлении полей не изменяется.

Далее, интегрируя второе уравнение системы (6-29), с учетом начальных условий получаем:

$$\frac{dx_2}{dt} = \omega_{\text{ц}} (x_1 - x_1^0) + v_2^0. \quad (6-31)$$

Подставляя полученное значение dx_2/dt в первое уравнение системы (6-29), получаем неоднородное уравнение

$$\frac{d^2 (x_1 - x_1^0)}{dt^2} + \omega_{\text{ц}}^2 (x_1 - x_1^0) = \omega_{\text{ц}} \left(\frac{E_0}{B_0} - v_2^0 \right),$$

решение которого можно представить в виде

$$x_1 - x_1^0 = A e^{j\omega_{\text{ц}} t} + B e^{-j\omega_{\text{ц}} t} + \frac{\frac{E_0}{B_0} - v_2^0}{\omega_{\text{ц}}}, \quad (6-32)$$

где $Ae^{j\omega_{\text{ц}}t} + Be^{-j\omega_{\text{ц}}t}$ — общее решение однородного уравнения; $\frac{\frac{E_0}{B_0} - v_2^0}{\omega_{\text{ц}}}$ — частное решение неоднородного уравнения.

Удовлетворяя начальным условиям $x_1 = x_1^0$ и $dx_1/dt = v_1^0$ при $t=0$, найдем постоянные интегрирования A и B . В окончательном виде решение (6-32) можно записать как

$$x_1 = x_1^0 + \frac{v_2^0 - \frac{E_0}{B_0}}{\omega_{\text{ц}}} \cos \omega_{\text{ц}} t + \frac{v_1^0}{\omega_{\text{ц}}} \sin \omega_{\text{ц}} t + \frac{\frac{E_0}{B_0} - v_2^0}{\omega_{\text{ц}}}. \quad (6-33)$$

Подставляя выражение (6-33) в уравнение (6-31), интегрируя его и находя постоянную интегрирования из начального условия $x_2 = x_2^0$ при $t=0$, получаем:

$$x_2 = x_2^0 - \frac{v_1^0}{\omega_{\text{ц}}} \cos \omega_{\text{ц}} t + \frac{v_2^0 - \frac{E_0}{B_0}}{\omega_{\text{ц}}} \sin \omega_{\text{ц}} t + \frac{E_0}{B_0} t + \frac{v_1^0}{\omega_{\text{ц}}}. \quad (6-34)$$

Выражения (6-30), (6-33) и (6-34) определяют траекторию движения электрона в общем случае. Практический интерес представляют два частных случая.

В первом случае электрон вылетает из катода ($x_1^0 = 0$) с нулевой начальной скоростью ($v_1^0 = v_2^0 = v_3^0 = 0$). При этом траектория движения электрона согласно выражениям (6-33), (6-34) и (6-30) определяется выражениями

$$\left. \begin{aligned} x_1 - R &= -R \cos \omega_{\text{ц}} t; \\ x_2 - (x_2^0 + v_{\text{центр}} t) &= -R \sin \omega_{\text{ц}} t; \\ x_3 &= x_3^0, \end{aligned} \right\} \quad (6-35)$$

где

$$R = \frac{E_0}{B_0 \omega_{\text{ц}}} = \frac{m}{|e|} \frac{E_0}{B_0^2}; \quad (6-36)$$

$$v_{\text{центр}} = \frac{E_0}{B_0}. \quad (6-37)$$

Уравнения (6-35) описывают окружность радиуса R , центр которой перемещается вдоль оси x_2 со скоростью

$v_{\text{центр}}$. Эта окружность катится без скольжения в плоскости $x_3 = \text{const}$ вдоль оси x_2 . Точка, лежащая на этой окружности, вращается с частотой $\omega_{\text{ц}}$ и при качении окружности описывает кривую, называемую циклоидой. Таким образом, траектория движения электрона в данном частном случае представляет собой *циклоиду*.

Скорость электрона, находящегося в верхней точке арки циклоиды, максимальна. Она складывается из линейной скорости точки, вращающейся по окружности радиуса R с угловой скоростью $\omega_{\text{ц}}$, и скорости перемещения центра окружности, образующей циклоиду. С учетом формул (6-36) и (6-37) получим:

$$v_{\text{макс}} = v_{\text{центр}} + \omega_{\text{ц}} R = 2v_{\text{центр}} = 2 \frac{E_0}{B_0}. \quad (6-38)$$

В точках на поверхности катода эти скорости вычитаются и скорость электрона равна

$$v_{\text{мин}} = v_{\text{центр}} - \omega_{\text{ц}} R. \quad (6-39)$$

Во втором случае электрон влетает в пространство между анодом и катодом со скоростью, составляющие которой $v_1^0 = v_3^0 = 0$ и $v_2^0 = v_{\text{центр}} = E_0/B_0$. Исходя из выражений (6-33), (6-34) и (6-30), получим уравнение траектории электрона:

$$\left. \begin{aligned} x_1 &= x_1^0; \\ x_2 &= \frac{E_0}{B_0} t; \\ x_3 &= x_3^0. \end{aligned} \right\} \quad (6-40)$$

Легко видеть, что это уравнение прямой, параллельной оси x_2 . В этом случае электрон движется по прямой со скоростью $v = v_2 = v_2^0$.

Из рис. 6-6 видно, что электроны при движении по циклоиде попадают на анод при условии, если диаметр $2R$ катящейся окружности, характеризующей движение электрона, больше расстояния d между анодом и катодом. Если $2R < d$, то ток отсутствует.

Режим, соответствующий равенству $d = 2R$ или

$$d = 2 \frac{v_{\text{центр}}}{\omega_{\text{ц}}} = 2 \frac{m}{e} \frac{E_0}{B_0},$$

называется *критическим*. Значения магнитного и электрических полей, соответствующих этому режиму, называются *критическими*.

Критическое значение магнитной индукции равно

$$B_{0\text{кр}} = \sqrt{2 \frac{mE_0}{ed}} = \frac{\sqrt{2 \frac{m}{e} U_0}}{d}, \quad (6-41)$$

так как $E_0 = U_0/d$, где U_0 — напряжение, приложенное между анодом и катодом.

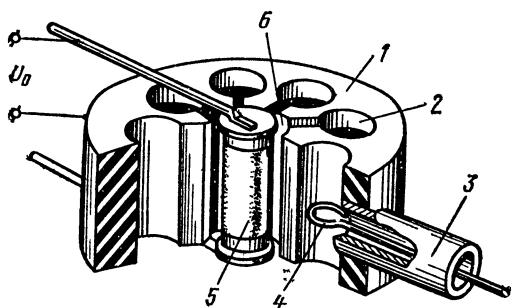


Рис. 6-7. Многорезонаторный магнетрон.

1 — анод; 2 — резонатор; 3 — нагрузка; 4 — петля связи;
5 — катод; 6 — пространство взаимодействия.

Группировка электронов и взаимодействие электронов с полем СВЧ (динамический режим) происходит в пространстве взаимодействия между замедляющей системой, находящейся под положительным потенциалом, и отрицательным электродом. Замедляющая система может иметь резонансный характер и состоять из резонаторов, связанных с пространством взаимодействия с помощью щелей, а может быть нерезонансной широкополосной периодической структурой.

Многорезонаторный магнетрон (рис. 6-7) представляет собой замкнутую в кольцо резонансную замедляющую систему, состоящую из резонаторов со стоячими волнами, поле СВЧ которых проникает в пространство взаимодействия с электронами, вылетающими из катода. Под действием постоянного электрического поля $E_0 = e_r E_0$, создаваемого источником питания, и перпендикулярного ему магнитного поля $B_0 = e_z B_0$, создаваемого специальным магнитом, электроны в пространстве взаимодействия вращаются вокруг катода. При выполнении

определенных условий синхронизма между вращающимися электронами и высокочастотным полем электроны группируются в сгустки, которые в тормозящем поле СВЧ отдают энергию полю (активная среда).

В статическом режиме задача определения траектории движения электрона принципиально не отличается от аналогичной задачи для плоской конструкции. В цилиндрической системе координат траектория движения электрона совпадает с траекторией точки, находящейся на окружности, которая катится по цилиндрическому катоду без скольжения. Кривая, описываемая этой точкой, называется *эпициклоидой*.

На рис. 6-8 показаны эпициклоидальные траектории для коаксиальной конструкции при отсутствии поля СВЧ (статический режим) при различных значениях постоянного магнитного поля. На этом же рисунке приведена зависимость тока от постоянного магнитного поля.

Критическое значение магнитной индукции согласно формуле (6-41) при $d = a_2 - a_1$ и $E_0 = \frac{U_0}{a_2 - a_1}$ (рис. 6-8) определяется выражением

$$B_{\text{кр}} = \frac{\sqrt{2 \frac{m}{e} U_0}}{(a_2 - a_1)}. \quad (6-42)$$

Как видно из соотношения (6-42), $B_{\text{кр}}$ возрастает с увеличением потенциала U_0 .

При возникновении электромагнитных колебаний в пространстве взаимодействия между анодом и катодом помимо постоянных скрещенных полей \mathbf{E}_0 и \mathbf{B}_0 на электроны действует переменное поле СВЧ, имеющее радиальные и азимутальные составляющие (рис. 6-9). В результате этого происходит движение электронов от

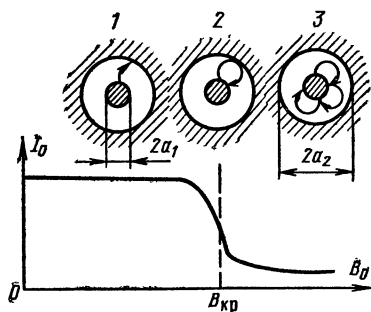


Рис. 6-8. Траектории движения электронов и ток в магнетроне.
1 — $B_0 < B_{\text{кр}}$; 2 — $B_0 = B_{\text{кр}}$; 3 — $B_0 > B_{\text{кр}}$.

катода к аноду по сложным траекториям. При этом происходит группировка электронов в сгустки, образующие спицеобразную конфигурацию, вращающуюся синхронно с изменением фазы переменного поля СВЧ в пространстве взаимодействия (рис. 6-10).

Механизм группировки электронов иллюстрируется рис. 6-11, на котором изображено пространство взаимодействия между плоскими катодом и анодом. Электроны «неправильной фазы», т. е. те, которые, выходя с катода,

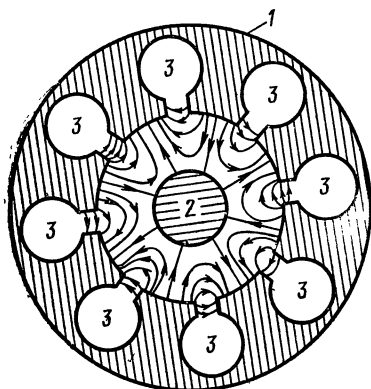


Рис. 6-9. Структура поля СВЧ в магнетроне.

1 — анод; 2 — катод; 3 — резонаторы.

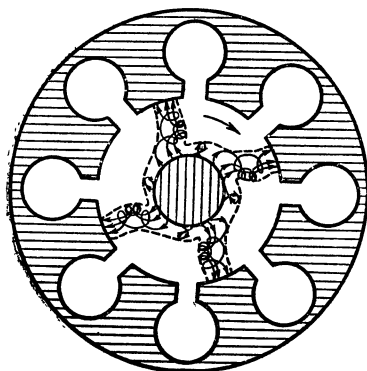


Рис. 6-10. Структура «спиц» в резонаторе.

попадают в ускоряющее поле азимутальной составляющей СВЧ, двигаясь по петле циклоиды, возвращаются на катод (траектория T_1) и тем самым автоматически выходят из пространства взаимодействия. Электроны «правильной фазы», т. е. электроны, при выходе из катода попадающие в тормозящее поле азимутальной составляющей поля СВЧ, двигаясь по петле циклоиды, теряют свою скорость и останавливаются, не доходя до катода. Если выполняется условие синхронизации, эти электроны во время следующего периода СВЧ колебания снова начинают двигаться по петле циклоиды, приближаясь к аноду (траектория T_2). Электроны «правильной фазы» образуют активную среду в пространстве взаимодействия. Потенциальная энергия n электронов «правильной

фазы», равная neU_0 , создаваемая источником питания, преобразуется в энергию СВЧ.

Процесс разделения электронов на электроны «неправильной» и «правильной» фазы может происходить при постоянном магнитном поле, превышающем его критическое значение, определяемое формулой (6-41), причем диаметр окружности, определяющей циклоиду (6-36), должен быть меньше расстояния между соседними резонаторами.

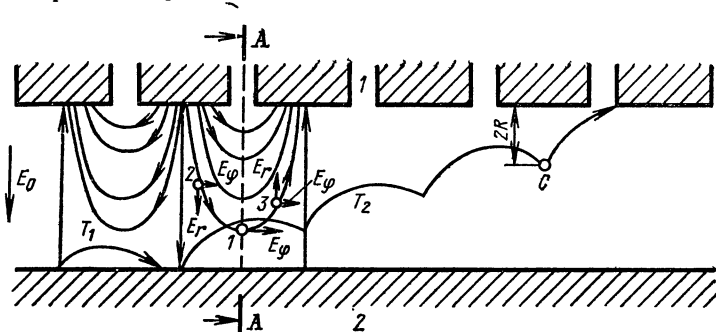


Рис. 6-11. Траектории электронов в многорезонаторном магнетроне.
1 — щели резонаторов в аноде; 2 — катод.

Максимальные значения азимутальных составляющих поля СВЧ наблюдаются в радиальных плоскостях, проходящих через середину резонаторных щелей (плоскость $A-A$, электрон 1 на рис. 6-11). Если электрон 2 будет отставать от такой плоскости, то под действием радиального суммарного поля ($E_0 + E_r$) согласно (6-37) его скорость возрастет и он догонит эту плоскость. Скорость электрона 3, находящегося по другую сторону плоскости $A-A$, под действием суммарного радиального поля ($E_0 - E_r$) согласно той же формуле (6-37) уменьшится. В результате этого образуются сгустки электронов и автоматически поддерживается синхронное вращение «спиц» и поля СВЧ; при этом энергия электронов, запасенная от источника питания, создающего напряжение U_0 , преобразуется в энергию СВЧ, когда концы спиц проходят через тормозящее переменное электрическое поле около щелей резонаторов. При небольших изменениях величины U_0 синхронизация автоматически восстанавливается.

Вращение концов «спиц» происходит со скоростью, при которой время перехода их от одной щели до соседней равно полупериоду колебаний СВЧ (колебания основного вида π). Вследствие этого для основного вида колебаний число спиц равно половине числа резонаторов (щелей).

Колебания этого вида являются наиболее устойчивыми и обычно используются на практике. Разумеется, возможны и колебания высших типов, однако работа магнетрона при этом менее устойчива.

Все электроны «правильной фазы», пройдя пространство взаимодействия, оседают на аноде, теряя при этом кинетическую энергию на его разогрев. Эта кинетическая энергия составляет на практике сравнительно небольшую часть полной энергии, затрачиваемой источником питания; основная часть последней переходит в энергию поля СВЧ. Вследствие этого к. п. д. многорезонаторного магнетрона значительно больше, чем у ЛБВ и клистрона. Для определения к. п. д. можно полагать, что электрон, попадая на анод, обладает кинетической энергией, равной его потенциальной энергии $W_{\text{пот}}$, когда он находился на расстоянии $2R$ от поверхности анода (см. рис. 6-11), т. е.

$$W_{\text{пот}} \approx W_0 \frac{2R}{(a_2 - a_1)},$$

где $W_0 = eU_0$.

Электронный к. п. д. определяется выражением

$$\eta_e = \frac{W_0 - W_{\text{пот}}}{W_0}$$

и с учетом (6-36) и (6-41)

$$\eta_e = 1 - \left(\frac{B_{\text{окр}}}{B} \right)^2.$$

Из этой формулы следует, что при увеличении магнитной индукции электронный к. п. д. возрастает. Полный к. п. д. магнетрона, учитывающий также потери в резонаторной системе, достигает 80%.

На рис. 6-12 приведены так называемые «рабочие характеристики» магнетрона, определяющие зависимость U_0 от анодного тока I_0 при различных значениях мощности СВЧ и магнитной индукции B_0 .

При малых значениях U_0 колебания СВЧ не возникают, так как при этом все электроны под действием B_0

возвращаются на катод. Максимальное значение анодного тока I_0 ограничивается возможностью разрушения поверхности катода большим числом возвращающихся на него электронов «неправильной фазы».

Кроме рассмотренного выше многорезонаторного магнетрона для генерирования и усиления колебаний

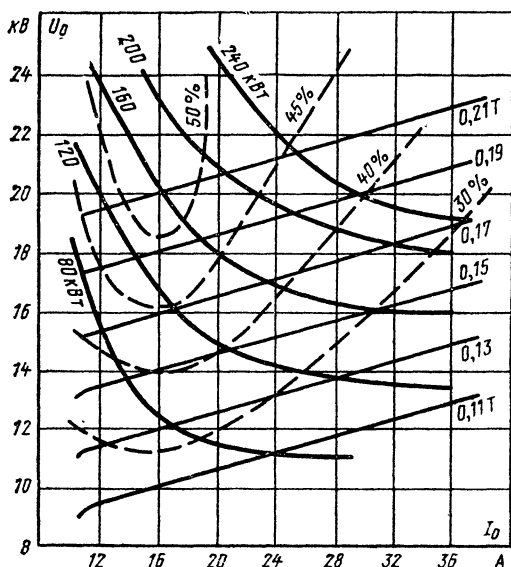


Рис. 6-12. Рабочие характеристики многорезонаторного магнетрона.

СВЧ диапазона применяются другие электронные приборы магнетронного типа: амплитроны, лампы бегущей и обратной волны типа М и некоторые другие.

Амплитрон—это магнетрон с независимым возбуждением, т. е. усилитель СВЧ колебаний, отличающийся от рассмотренного выше магнетрона наличием разомкнутой замедляющей системы. При этом один конец ее возбуждается слабым полем СВЧ, а другой конец является выходом усиленных колебаний СВЧ. Амплитрон, как усилитель, в отличие от ЛБВ и клистрона обладает большим к. п. д., но небольшим коэффициентом усиления — около 10 дБ.

Лампа бегущей волны типа М как усилитель схематически показана на рисунке 6-13. Замедляющая си-

стема представляет собой широкополосную нерезонансную периодическую структуру. Пространство взаимодействия образовано замедляющей системой и отрицательным электродом (ложный катод). В пространстве взаимодействия из плоской системы анод — катод электрон входит с вершины первого витка циклоиды. Скорость та-

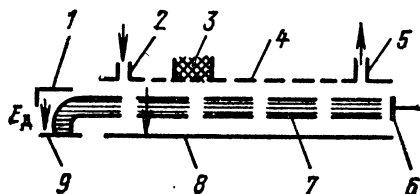


Рис. 6-13. ЛБВ типа М плоской конструкции.

1 — анод; 2 — вход энергии; 3 — поглотитель; 4 — замедляющая система; 5 — выход энергии; 6 — коллектор; 7 — электронный пучок; 8 — ложный катод; 9 — катод.

кого электрона согласно формуле (6-38) определяется выражением

$$v_0 = 2 \frac{E_{0ак}}{B_0}, \quad (6-43)$$

где $E_{0ак}$ — напряженность постоянного электрического поля в пространстве анод — катод.

В пространстве взаимодействия электрон движется прямолинейно. Для этого, как уже было показано, необходимо, чтобы

$$v_0 = \frac{E_0}{B_0}, \quad (6-44)$$

где E_0 — напряженность постоянного электрического поля в пространстве взаимодействия. Постоянное магнитное поле одинаково и в системе анод — катод, и в пространстве взаимодействия.

Из сравнения соотношений (6-43) и (6-44) видно, что в пространстве взаимодействия напряженность постоянного электрического поля должна быть в 2 раза больше, чем в пространстве анод — катод.

Процессы группировки электронов и взаимодействия их с полем электромагнитной волны аналогичны соответствующим процессам в магнетроне. При движении электронов синхронно с одной из пространственных гар-

моник (практически всегда основной) электронный пучок распадается на сгустки, которые в отличие от сгустков в магнетроне имеют форму тонких узких лент, а не спиц. Ленты ориентированы перпендикулярно плоскости чертежа на рис. 6-13. При достижении коллектора лента прекращает свое существование в отличие от спиц в магнетроне, которые непрерывно вращаются вокруг катода. В тормозящем поле СВЧ сгустки отдают свою энергию полю и усиливают его.

Усилители на ЛБВ типа М работают в сантиметровом и дециметровом диапазонах. В сантиметровом диапазоне конструкция усилителя обычно плоская, в дециметровом — цилиндрическая для уменьшения габаритов. Коэффициент усиления до 20 дБ, относительная полоса пропускания до 25%. Импульсные ЛБВ усилители типа М, работающие в сантиметровом диапазоне, имеют мощность на выходе до сотен киловатт, в дециметровом — до десятка мегаватт при к. п. д. до 40%.

Лампа обратной волны типа М как генератор работает на обратной пространственной гармонике. Внутренняя положительная обратная связь создается в результате движения энергии навстречу потоку электронов. Конструкция лампы преимущественно цилиндрическая. ЛОВ генераторы типа М работают в дециметровом и сантиметровом диапазонах с выходными мощностями от единиц киловатт до единиц ватт, соответственно, и к. п. д. до 40%. Диапазон электронной перестройки может достигать 30—40%.

6-4. КЛИСТРОНЫ

Пролетный двухрезонаторный клистрон схематически представлен на рисунке 6-14. Электроны эмиттируются подогревным катодом 2. Под действием постоянного напряжения источника питания U_0 , приложенного между катодом и ускоряющим электродом 8, электроны ускоряются и попадают в зазор 3а. Этот зазор образован двумя сетками, расположенными на близком расстоянии друг от друга, и представляет часть первого резонатора, имеющего тороидальную форму. В зазоре концентрируется электрическое поле, параллельное направлению движения электронов. Усиливаемые СВЧ колебания через петлю связи подводятся к первому резонатору, называемому обычно модулятором. Второй выходной резона-

тор можно назвать активным. Настройка первого и второго резонатора на частоту усиливаемых колебаний производится либо с помощью прогиба стенок резонаторов, либо подстроечными винтами. Оба резонатора находятся под одним и тем же постоянным потенциалом, равным потенциалу U_0 ускоряющего электрода.

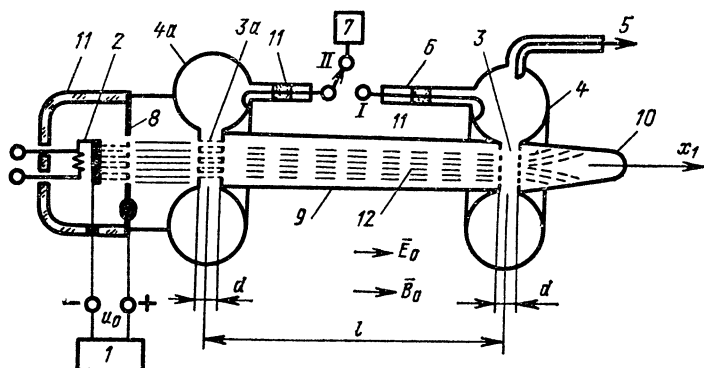


Рис. 6-14. Пролетный двухрезонаторный клистрон.

1 — источник питания; 2 — катод; 3 — активный зазор; 3a — модулирующий зазор; 4, 4a — резонаторы; 5 — выход; 6 — линия обратной связи; 7 — задающий генератор; 8 — ускоряющий электрод; 9 — пространство дрейфа; 10 — коллектор; 11 — изолятор; 12 — группа электронов.

Подводимое к первому резонатору напряжение возбуждает в нем колебания, в результате чего между сетками возникает переменное напряжение

$$\left. \begin{aligned} u_1 &= U_{m1} \sin \omega t; \\ U_{m1} &\ll U_0. \end{aligned} \right\} \quad (6-45)$$

Переменное электрическое поле, возникающее в зазоре, экранируется сетками и не влияет на поле вне резонатора.

Скорость электронов, влетающих в зазор, $v_0 = \sqrt{2 \frac{e}{m} U_0}$. При пролете зазора модулирующего резонатора скорость электронов изменяется в зависимости от мгновенной разности потенциалов между сетками. Значение скорости определяется выражением

$$\frac{mv^2}{2} = eU_0 + eU_{m1} \sin \omega t;$$

отсюда скорость электрона, выходящего из зазора,

$$v = \sqrt{2 \frac{e}{m} (U_0 + U_{m1} \sin \omega t)},$$

или, так как $U_{m1} \ll U_0$,

$$v \approx v_0 \left(1 + \frac{1}{2} \frac{U_{m1}}{U_0} \sin \omega t \right), \quad (6-46)$$

т.е. модуляция скорости происходит по гармоническому закону.

Полученное выражение (6-46) приближенно, так как в нем не учтено изменение напряжения на сетках за время пролета электроном зазора. Обычно ширина зазора мала и время пролета также мало. Пренебрегая малым изменением скорости электронов внутри зазора, определим время пролета

$$\tau = \frac{d}{v_0},$$

где d — ширина зазора.

В зазоре на каждый электрон, находящийся в момент времени t в середине зазора, действует среднее значение напряжения

$$u_{10} = \frac{1}{\tau} \int_{t-\tau/2}^{t+\tau/2} U_{m1} \sin \omega t \, dt = \gamma U_{m1} \sin \omega t, \quad (6-47)$$

где

$$\gamma = \frac{\sin \frac{\theta}{2}}{\frac{\theta}{2}} \quad (6-48)$$

— коэффициент эффективности взаимодействия электронов с полем СВЧ, а

$$\theta = \omega \tau = \omega \frac{d}{v_0} \quad (6-49)$$

— угол пролета в зазоре.

С учетом конечного значения времени пролета согласно (6-47) выражение (6-46) можно переписать как

$$v \approx v_0 (1 + M \sin \omega t), \quad (6-50)$$

где

$$M = \gamma \frac{U_{mi}}{2U_0} \quad (6-51)$$

— коэффициент модуляции скорости электронов.

Модуляция электронов по скорости максимальна, когда максимален коэффициент эффективности взаимодействия электронов с полем. Это наблюдается согласно формулам (6-48) и (6-49) при сравнительно малой ширине зазора d и больших значениях постоянного напряжения U_0 . При этом угол пролета в зазоре θ мал ($\theta \ll \pi$) и $\gamma \approx 1$. При θ , равном четному числу π , коэффициент эффективности $\gamma = 0$ и модуляции по скорости в среднем не происходит, так как ускорение и торможение электронов в зазоре компенсируют друг друга. В клистронах угол пролета θ лежит в пределах от $\pi/2$ до π . При этом коэффициент $\gamma \approx 0,9 \div 0,6$.

После выхода из модулирующего зазора часть электронов ускорена, часть — замедлена. Пространство между зазорами первого и второго резонаторов 9 называется пространством дрейфа (трубка дрейфа). В этом пространстве электроны движутся по инерции, так как оба резонатора находятся под одним и тем же потенциалом. Группировка электронов в пространстве дрейфа иллюстрируется рис. 6-15. По вертикальной оси отложены расстояния от центра модулирующего зазора и мгновенные значения напряжения на сетках этого зазора, по горизонтальной оси — время. График движения каждого электрона представляет собой прямую, наклон которой к горизонтальной оси определяется скоростью электрона. Точки пересечения прямых с осью времени определяют моменты входа электронов в пространство дрейфа. Электроны, пролетающие модулирующий зазор, при $u_1 = 0$, не изменяют своей скорости. Графики движения этих электронов представлены сплошными прямыми линиями. Для тех электронов, которые проходят модулирующий зазор, когда высокочастотное поле является тормозящим, наклон прямых увеличивается (пунктирные прямые); для электронов, ускоряемых полем, уменьшается (жирные прямые). Точки пересечения этих прямых соответствуют положению сгустков электронов. Группируются электроны, прошедшие модулирующий зазор в ту половину периода, когда поле в зазоре изменяется от максимально тормозящего до максимально ускоряю-

щего (часть синусоиды с положительным наклоном). Электроны, прошедшие модулирующий зазор в другую половину периода, разгруппировываются. На каждый период колебания высокочастотного поля приходится один сгусток, центром которого являются электроны, прошедшие модулирующий зазор без изменения скорости.

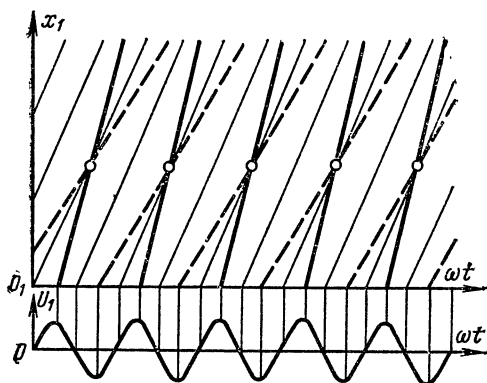


Рис. 6-15. Процесс группировки электронов в сгустки.

Из графика (рис. 6-15) видно, что процесс образования сгустка в пространстве дрейфа начинается сразу около модулирующего зазора. По мере удаления от зазора плотность электронов возрастает и на некотором определенном расстоянии (при заданных U_0 и U_{m1}) достигает максимального значения. При дальнейшем удалении от зазора сгусток расплывается, а затем процесс образования начинается снова и второй сгусток образуется уже на большем расстоянии от зазора и т. д.

Образовавшиеся сгустки после прохождения пространства дрейфа попадают в зазор второго резонатора, который можно назвать активным, так как в него сгустки влетают в тот момент, когда направление поля в зазоре совпадает с направлением движения электронов. При этом электроны тормозятся и отдают свою энергию полю.

Для работы клистрона необходимо, чтобы сгусток максимальной плотности образовывался в зазоре второ-

го резонатора. Рассмотрим условия образования такого сгустка.

Пусть t_1 — момент прохождения электроном середины модулирующего зазора. Электрон достигает середины зазора второго резонатора в момент

$$t_2 = t_1 + \frac{l}{v},$$

или в угловых единицах

$$\omega t_2 = \omega t_1 + \omega \frac{l}{v}, \quad (6-52)$$

где l — длина пространства дрейфа.

Считая, что в пространстве дрейфа электроны движутся по инерции (силами рассталкивания между электронами пренебрегаем), подставляя (6-50) в выражение (6-52), получаем:

$$\omega t_2 = \omega t_1 + \frac{\omega l}{v_0 (1 + M \sin \omega t_1)}.$$

Так как $U_{m1} \ll U_0$, то согласно (6-51) $M \ll 1$ и

$$\omega t_2 = \omega t_1 + \theta_0 - A \sin \omega t_1, \quad (6-53)$$

где $\theta_0 = \omega \frac{l}{v_0}$ — угол пролета пространства дрейфа невозмущенными электронами (при прохождении модулирующего зазора не изменяющими своей скорости), $A = \theta_0 M = \theta_0 \gamma \frac{U_{m1}}{2U_0}$ — амплитуда переменной составляющей угла пролета пространства дрейфа, называемая параметром группировки.

Действительно, величина A связана с группировкой. Если $A = 0$, то модуляция по скорости отсутствует и согласно (6-53)

$$\omega t_2 = \omega t_1 + \theta_0,$$

т.е. угол пролета через пространство дрейфа для всех электронов одинаков. Если $A \neq 0$, то происходит группировка электронов. Рассмотрим этот процесс с помощью графика (рис. 6-15) для трех электронов: «среднего», скорость которого не изменяется при прохождении модулирующего зазора, и двух «крайних», наиболее замедленного и наиболее ускоренного в зазоре. Остальные электроны во внимание не принимаем. Как видно из гра-

фика, фазы влета «крайних» электронов в пространство дрейфа отличаются от фазы «среднего» электрона на $\pm\pi/2$. При этом согласно выражению (6-53) получим для «среднего» электрона

$$\omega t_2 = \omega t_1 + \theta_0,$$

для «крайних» электронов

$$\omega t_2 = \omega t_1 + \theta_0 \pm \frac{\pi}{2} \mp A.$$

Отсюда следует, что все три электрона влетят в зазор второго резонатора одновременно при $A = \theta_0 \gamma \frac{U_{m1}}{2U_0} = 1,57$. Это значение соответствует оптимальной группировке трех рассмотренных электронов.

Вычислим конвекционный ток в зазоре второго резонатора. При достаточно сильном продольном постоянном магнитном поле, фокусирующем пучок электронов, можно считать, что все электроны вылетевшие из катода, достигнут второго зазора. Применяя закон сохранения заряда, запишем:

$$dq_1 = dq_2, \quad (6-54)$$

где dq_1 — заряд, пролетевший через плоскость середины первого зазора за время dt_1 , dq_2 — заряд, пролетевший через плоскость середины второго зазора за время dt_2 .

Так как

$$i = \frac{dq}{dt},$$

то равенство (6-54) можно переписать в виде

$$i_1 dt_1 = i_2 dt_2,$$

где i_1 и i_2 — мгновенные значения конвекционного тока в середине первого и второго зазоров.

Отсюда конвекционный ток в зазоре второго резонатора определяется выражением

$$i_2 = i_1 \frac{dt_1}{dt_2}.$$

Так как модуляция пучка по плотности в середине первого зазора практически отсутствует, то ток i_1 в пер-

вом зазоре равен постоянному току I_0 , поступающему из катода, т. е.

$$i_2 = I_0 \frac{dt_1}{dt_2}. \quad (6-55)$$

Определяя производную $\frac{dt_1}{dt_2}$ из уравнения (6-53) и подставляя полученное значение в (6-55), получаем:

$$i_2 = \frac{I_0}{1 - A \cos \omega t_1}. \quad (6-56)$$

Из соотношения (6-56) видно, что ток в зазоре второго резонатора определяется параметром группировки A . При $A=0$ ток постоянный, при малом A ток близок к гармоническому, при $A=1$ и $\cos \omega t_1=1$ ток обращается в бесконечность, что соответствует идеальной группировке при сделанном допущении, что рассталкивающих сил между электронами нет.

В реальных условиях конвекционный ток во втором зазоре является периодическим пульсирующим с частотой ω током и его можно представить в виде ряда Фурье для четной функции

$$i_2(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos n(\omega t_2 - \theta_0), \quad (6-57)$$

где

$$a_n = \frac{2}{\pi} \int_0^{\pi} i_2(t) \cos n(\omega t_2 - \theta_0) d(\omega t_2) \quad n = 0, 1, 2, \dots \quad (6-58)$$

— амплитуды гармонических составляющих.

Подставляя в выражения (6-58) значения $i_2(t)$, $\omega t_2 - \theta_0$ и $d(\omega t_2)$, определяемые (6-55) и (6-53), получаем:

$$a_0 = 2I_0,$$

$$a_n = 2I_0 \frac{1}{\pi} \int_0^{\pi} \cos n(\omega t_1 - A \sin \omega t_1) d(\omega t_1), \quad n = 1, 2, \dots$$

Интеграл

$$\frac{1}{\pi} \int_0^{\pi} \cos n(\omega t_1 - A \sin \omega t_1) d(\omega t_1) = J_n(nA)$$

является интегральным представлением функции Бесселя первого рода n -го порядка.

Таким образом, ряд (6-57) можно записать в виде

$$i_2 = I_0 + \sum_{n=1}^{\infty} 2I_0 J_n(nA) \cos n(\omega t_2 - \theta_0). \quad (6-59)$$

Тороидальные резонаторы клистронов характеризуются высокой добротностью и могут возбуждаться лишь одной из гармонических составляющих тока, на частоту которой они настроены. Поэтому зазор второго резонатора помещается на таком расстоянии от зазора первого резонатора, чтобы в месте расположения второго зазора был максимален не весь ток, а соответствующая гармоника.

Наибольший интерес представляет первая гармоника, определяемая функцией Бесселя $J_1(A)$, максимальное значение которой равно 0,58 при $A=1,84$. Оптимальным значением параметра группировки для всего тока, как уже было показано, является значение $A=1$.

Так как длина пространства дрейфа l определяется конструкцией клистрона, то оптимальный параметр группировки A можно получить, меняя постоянное напряжение U_0 и амплитуду модулирующего колебания U_{m1} .

Максимальная амплитуда первой гармоники конвекционного тока во втором зазоре при оптимальной группировке равна:

$$I_m \approx 1,16I_0. \quad (6-60)$$

Если второй резонатор настроен на частоту этой гармоники, то она возбуждает в нем поле СВЧ с частотой ω . При этом электрическое поле, сконцентрированное в зазоре и тормозящее электроны, будет практически синусоидальным, как и напряжение на сетках резонатора. Чтобы все электроны проходили через активный зазор, а не возвращались обратно в пространство дрейфа, амплитуда напряжения между сетками U_m не должна превышать постоянного напряжения U_0 :

$$U_m \leq U_0. \quad (6-61)$$

На основании этого и учитывая коэффициент эффективности взаимодействия электронов с полем γ , определяемый выражением (6-48), находим максимальную

мощность в выходном резонаторе и максимально возможный электронный к. п. д. двухрезонаторного клистрона:

$$P_0 \leq \frac{I_m U_m}{2} \gamma \text{ и } \eta \leq \gamma \frac{P_0}{U_0 I_0}.$$

Подставляя сюда (6-60) и (6-61), получаем:

$$P_0 \leq 0,58 U_0 I_0 \text{ и } \eta \leq 0,58 \gamma.$$

С учетом неполной группировки всех электронов из-за оседания части их на сетках, ограничивающих зазоры, а также с учетом потерь в стенках резонатора на практике удастся получить к. п. д. $\eta \approx 20 \div 30\%$.

Важной характеристикой клистрона является коэффициент усиления по мощности G , который увеличивается при удлинении трубки дрейфа, так как на большой длине может образоваться большее число сгустков. Однако при большой длине дрейфа будет возрастать частичная потеря электронов, не достигающих активного зазора, за счет расталкивающих сил, действующих в пространственном заряде. На практике у двухрезонаторного клистрона $G \leq 20$ дБ.

Другой важной характеристикой клистрона как усилителя является пропускаемая полоса частот $\Delta\omega$. Относительное значение ее определяется добротностью Q нагруженного резонатора, т. е.

$$\frac{\Delta\omega}{\omega} = \frac{1}{Q}$$

и в диапазоне СВЧ для двухрезонаторного клистрона равна примерно 10^{-3} .

Для увеличения коэффициента усиления и получения при этом большей выходной мощности применяются многорезонаторные клистроны с числом резонаторов от трех до пяти. Между первым входным резонатором и последним выходным располагаются промежуточные резонаторы, которые играют роль активного — для предыдущего резонатора и модулирующего — для последующего. В пространстве дрейфа между каждой парой резонаторов осуществляется группировка электронов.

Для многорезонаторных клистронов к. п. д. около 35%, коэффициент усиления 40—60 дБ и относительная полоса пропускания $\Delta\omega/\omega \approx 10^{-2}$. Выходная мощность мощных клистронов, работающих в непрерывном режи-

ме, около сотни киловатт, клистронов, работающих в импульсном режиме, около десятка мегаватт.

Для умножения частоты используются двухрезонаторные клистроны. Входной резонатор клистрона настраивается на частоту умножаемого напряжения, а выходной — на частоту нужной (n -й) гармоники. Обычно

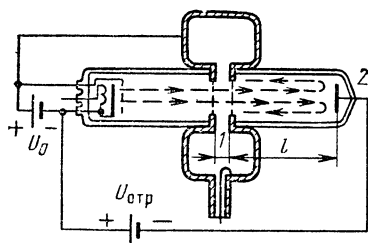


Рис. 6-16. Отражательный клистрон.

1 — зазор резонатора; 2 — отражатель.

$n=3\div 10$. Возможность умножения частоты обусловлена наличием гармоник в конвекционном токе второго резонатора [см. формулу (6-59)], амплитуда которых зависит от параметра группировки. Коэффициент полезного действия умножителей мал и составляет единицы процентов. Фокусировка электронов в пролетном клистроне, как и в ЛБВ, осуществляется продольным постоянным магнитным полем.

Отражательные клистроны (рис. 6-16) являются генераторами малой мощности. Такой клистрон имеет только один резонатор, который находится под тем же потенциалом U_0 , что и ускоряющий электрод. За резонатором находится отражатель, к которому приложено отрицательное относительно катода напряжение.

Начальное поле СВЧ в резонаторе возникает из-за флуктуаций в электронном потоке, создаваемом катодом. В зазоре концентрируется электрическое поле, которое модулирует скорость движения электронов в потоке. Эта модуляция определяется выражением (6-50). Группировка электронов происходит в пространстве между зазором резонатора и отражателем. В отличие от пролетного клистрона, где в пространстве дрейфа поля нет и электроны движутся по инерции, в отражательном клистроне создается статическое тормозящее поле,

так как отражатель имеет отрицательный потенциал относительно катода, а резонатор — положительный. Если считать, что тормозящее поле между сеткой резонатора и отражателем однородно, то напряженность этого поля

$$E_0 = \frac{U_0 - U_{\text{отр}}}{l},$$

где $U_{\text{отр}}$ — напряжение на отражателе; l — расстояние от сетки резонатора до отражателя.

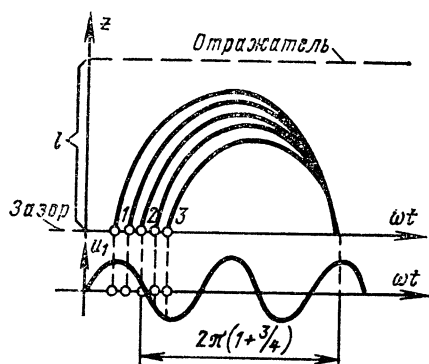


Рис. 6-17. Процесс группировки электронов в отражательном клистроне.

Процесс группировки электронов рассмотрим с помощью графиков, приведенных на рис. 6-17. Здесь на нижнем графике показано изменение напряжения на сетках зазора резонатора, на верхнем — движение электронов в тормозящем поле (от второй сетки резонатора в сторону отражателя и обратно). На графике представлено движение только тех электронов, которые проходят зазор резонатора в период, когда напряжение на сетках изменяется от максимального ускоряющего до максимального тормозящего (часть синусоиды с отрицательным наклоном).

Рассмотрим движение трех электронов, обозначенных на графике цифрами 1, 2 и 3. Электрон 2 пролетает зазор в момент, когда поле СВЧ равно нулю и скорость электрона не изменяется. Электрон 1, пролетевший зазор на четверть периода раньше, вылетает из зазора с большей кинетической энергией и потому удаляется от

зазора на большее расстояние, затрачивая на путь до возвращения в зазор большее время. Электрон 3 вылетает из зазора на четверть периода позже, чем электрон 2, но скорость его меньше, чем у электронов 2 и 1; в тормозящем поле он проделывает меньший путь, затрачивая на него меньше времени. При определенных условиях моменты возвращения всех трех электронов могут совпасть. Электроны, вылетающие из зазора в тот период времени, когда напряжение между сетками зазора изменяется от максимального замедляющего до максимального ускоряющего (часть синусоиды с положительным наклоном), сгруппировать невозможно, так как электроны, вылетевшие раньше, будут находиться в тормозящем поле меньше время, чем вылетевшие позже.

Образовавшиеся сгустки должны влетать в зазор в тот период, когда поле в нем будет тормозящим. При этом электроны теряют свою кинетическую энергию, усиливая поле СВЧ в резонаторе. Для электронов, вылетающих из зазора в этот период, поле СВЧ будет ускоряющим. Центром сгустка является электрон 2. Этот электрон попадает в максимально тормозящее поле СВЧ, как видно из рис. 6-17, при условии, что угол пролета его от сетки зазора и обратно

$$\theta_0 = 2\pi(n + 3/4). \quad (6-62)$$

Соотношение (6-62) является условием самовозбуждения отражательного клистрона. Если угол пролета не соответствует этому условию, то сгусток будет влетать в зазор не в момент максимального тормозящего поля, а раньше или позже. При этом сгустки будут либо отдавать полю СВЧ меньшую энергию, либо даже поглощать энергию поля. В последнем случае самовозбуждение вообще невозможно.

Время пролета зависит от напряжения на отражателе $U_{отр}$ и ускоряющего напряжения U_0 . Чем больше значение $U_{отр}$ и меньше U_0 , тем на меньшее расстояние от зазора будут удаляться электроны, тем меньше время пролета, и наоборот. Меняя значения $U_{отр}$ и U_0 , можно в широких пределах изменять угол пролета и удовлетворять условию (6-62) при различных n , соответствующих различным зонам генерирования. Значения напряжения $U_{отр}$ и U_0 , при которых удовлетворяется условие (6-62), соответствуют центрам зон. Значения $U_{отр}$ и U_0 между

зонами генерирования соответствуют поглощению энергии СВЧ сгустками, влетающими в зазор.

Конвекционный пульсирующий ток в резонаторе определяется выражением (6-59), в котором знак плюс заменяется на минус из-за сдвига процесса группировки электронов в тормозящем поле отражательного клистро-на по фазе на π по сравнению с пролетным клистроном. Этот сдвиг очевиден из сравнения графиков, приведенных на рис. 6-15 и 6-17. Для работы может быть использована только одна основная гармоника, так как для модуляции скорости электронов и отбора у них энергии используется один и тот же резонатор.

Отражательные клистроны имеют низкий к. п. д. (не более 2—3%), выходную мощность — единицы ватт и относительную полосу электронной перестройки менее 1%.

ФИЗИЧЕСКИЕ ДОПОЛНЕНИЯ (ЭЛЕМЕНТЫ КВАНТОВОЙ МЕХАНИКИ)

1. **Уровни энергии атома водорода и водородоподобных ионов.** Согласно теории Бора атом или молекула не могут находиться в состояниях с произвольной энергией, а могут находиться лишь в некоторых дискретных состояниях, называемых *устойчивыми* или *стационарными*.

Уровни энергии атома водорода определяются выражением

$$W_n = -\frac{R}{n^2}, \quad (\text{П1-1})$$

где $n=1, 2, 3 \dots$ называется *главным квантовым числом*; $R = \frac{mc^4}{8h^2e_0}$ — постоянная Ридберга; m — масса электрона; e — его заряд; $h = 6,6 \cdot 10^{-34}$ Дж·с — постоянная Планка; e_0 — электрическая постоянная.

В этом случае энергия зависит только от главного квантового числа и отрицательна для всех устойчивых состояний. Состояние с самой низкой энергией соответствует $n=1$ и движению электрона по самой близкой к ядру орбите. Это состояние называется *основным* или *нормальным*.

С увеличением n энергия растет и приближается к нулю, а уровни сближаются (рис. П1-1). При увеличении n орбиты движения электрона удаляются от ядра. Когда электрон удален на бесконечность ($n=\infty$) и находится в покое, энергия атома предполагается равной нулю.

Электрон, удаленный от ядра, может находиться в движении, приближаться или удаляться от ядра. Кинетическая энергия двух частиц, приближающихся друг к другу или удаляющихся друг от друга, положительна и может принимать любые значения (не квантуется).

Так как электрон обладает некоторой массой и движется вокруг ядра, то это движение можно характеризовать *орбитальным моментом количества движения* электрона в атоме. Орбитальный момент квантуется и определяется формулой

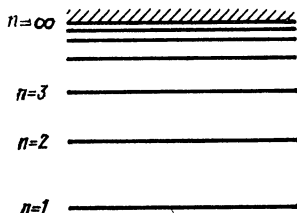
$$||| = \sqrt{l(l+1)} \hbar,$$

где $\hbar = \frac{h}{2\pi}$, l — *орбитальное*, или *азимутальное*, *квантовое число*.

При заданном значении n орбитальное квантовое число может принимать только следующие значения

$$l = 0, 1, 2, \dots, n - 1.$$

Так как электрон кроме массы обладает и зарядом, то при вращении его вокруг ядра возникает магнитный орбитальный момент, противоположный механическому моменту по направлению. Отношение магнитного момента к механическому определяется величиной e/m , где e — заряд электрона; m — его масса. В магнитном поле маг-



нитный момент и связанный с ним механический момент будут прецессировать вокруг направления магнитного поля. При этом угол прецессии (рис. П1-2) может иметь только определенные (квантованные) значения. Возможны лишь те углы, при которых проекция орбитального момента l на направление поля имеет значение

$$m_l \hbar,$$

Рис. П1-1. Уровни энергии атома водорода и водородоподобных ионов.

где m_l — магнитное квантовое число, которое при заданном значении l равно:

$$m_l = 0, \pm 1, \pm 2, \pm 3, \pm \dots, \pm l.$$

Электрон вращается вокруг собственной оси и обладает собственным магнитным моментом и моментом количества движения s , называемым спином. Этот момент квантуется и равен:

$$|s| = \sqrt{s(s+1)} \hbar,$$

где s — спиновое магнитное квантовое число. Отношение спинового магнитного момента к механическому равно $\frac{e}{2m}$.

В магнитном поле спин электрона может ориентироваться лишь так, что его проекция на направление поля (рис. П1-3) равна $s\hbar$, т. е. спиновое магнитное квантовое число имеет только два значения:

$$s = \pm \frac{1}{2}.$$

Квантовые числа n , l , m_l , s определяют устойчивые состояния атома. Но энергия (П1-1) зависит только от главного квантового числа n . Следовательно, одному и тому же уровню энергии соответствует несколько различных состояний атома. Состояния с одинаковой энергией называются *вырожденными*. Число состояний с одинаковой энергией называется *степенью* (кратностью) *вырождения* или *статистическим весом*.

Для атома водорода каждому значению энергии, определяемому числом n , соответствует

$$g = 2 \sum_{l=0}^{n-1} (2l+1) = 2n^2$$

устойчивых состояний.

Обычно вырождение связано с симметрией. У атома водорода благодаря сферической симметрии внутриатомного электрического поля энергия не зависит от ориентации орбитального и спинового моментов, что и приводит к вырождению уровней. Во внешнем поле симметрия нарушается, атом приобретает добавочную энергию

$$\Delta W = Hm \cos \theta,$$

где θ — угол между направлениями магнитного момента \mathbf{m} и внешнего магнитного поля \mathbf{H} . В этом случае вырождение снимается. Уровни, соответствующие одному и тому же n , расщепляются. Состояния

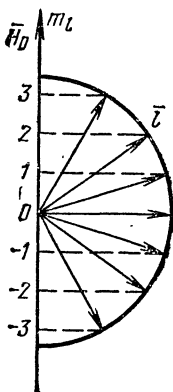


Рис. П1-2. Пространственное квантование вектора \vec{l} в магнитном поле.

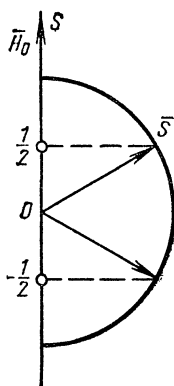


Рис. П1-3. Пространственное квантование вектора \vec{s} в магнитном поле.

с разными магнитными квантовыми числами m_l будут соответствовать различной энергии.

Аналогичной схемой уровней энергии обладают водородоподобные ионы He^+ , Li^{2+} , Be^{3+} , ..., состоящие из ядра с зарядом Ze (Z — порядковый номер элемента в периодической таблице Менделеева) и одного электрона. При этом уровни энергии будут определяться выражением

$$W_n = -\frac{RZ^2}{n^2}.$$

2. Энергетические уровни атомов. В многоэлектронном атоме каждому отдельному электрону можно приписать те же самые квантовые числа, что и одному электрону, движущемуся в поле атомного ядра n , l , m_l , s . Это справедливо, так как действие всех других электронов можно заменить приближенно некоторым средним полем с центральной симметрией. Но так как симметрия в этом случае все же нарушается, то энергия будет зависеть не только от n , но и от l :

$$W_{n,l} = -\frac{RZ^2}{(n + a)^2},$$

где $a=a(l) \leftarrow$ поправка Ридберга, зависящая от орбитального квантового числа l и быстро стремящаяся к нулю с уменьшением l . Таким образом, при заданном n энергия увеличивается с ростом l .

Степень вырождения при заданных n и l равна:

$$g_l = 2(2l + 1),$$

т. е. энергии $W_{n,l}$ соответствует $2(2l+1)$ состояний.

Согласно принципу Паули не может быть двух электронов в одном и том же состоянии, т. е. никакие два электрона не могут характеризоваться одинаковым рядом квантовых чисел n, l, m_l, s . Электроны, имеющие данные значения n и l , образуют электронную оболочку, которую заполняют $2(2l+1)$ электронов.

Электроны с $l=0, 1, 2, 3 \dots$ называются $s^-, p^-, d^-, f^- \dots$ — электронами, а уровни соответственно обозначаются $1s, 2s, 2p, 3s, 3p, 3d$ и т. д., где цифра соответствует значению квантового числа n . Уровни с одинаковым n, l , но разными m_l, s можно считать вырожденными, если пренебречь взаимодействием электронов друг с другом и с магнитным полем, создаваемым орбитальным движением электрона и его спиновым магнитным моментом.

При последовательном заполнении электронами оболочек атома получается *основная* или *нормальная электронная конфигурация*:

водород ($Z=1$) $1s$;

гелий ($Z=2$) $1s^2$;

аргон ($Z=18$) $1s^2 2s^2 2p^6 3s^2 3p^6$

(степень определяет число электронов в оболочке).

В противоречие с классической электродинамикой согласно постулатам *Бора* энергия электрона при движении в пределах электронной оболочки не изменяется. Изменение энергии происходит при переходе электрона с одной электронной оболочки, соответствующей, например, уровню энергии W_1 , на другую с энергией W_2 .

Расстояние между электронными уровнями атома составляет $1-2$ эВ.

При поглощении энергии один или несколько электронов переходят на более удаленные электронные оболочки, соответствующие более высоким уровням энергии. При этом получают *возбужденные электронные конфигурации*. Например, для Не конфигурация $1s 2p$ соответствует переходу одного электрона с уровня $1s$ на уровень $2p$.

Каждый электрон в атоме обладает орбитальным моментом количества движения l_k и спином s_k . Векторная сумма орбитальных моментов отдельных электронов в атоме называется орбитальным моментом атома L ; векторная сумма спинов всех электронов в атоме называется спиновым моментом атома S :

$$L = \sum_{k=1}^Z l_k, \quad S = \sum_{k=1}^Z s_k,$$

где Z — порядковый номер элемента.

Так как векторная сумма моментов электронов заполненных оболочек атома равна нулю, то суммарные моменты L и S определяются соответственно суммой l_k и s_k электронов незаполненных оболочек атома.

Так как l_k и s_k отдельных электронов квантованы, то и результирующие моменты атомов L и S также квантованы. Величина $|L|$ может принимать дискретные значения, равные $\sqrt{L(L+1)} \hbar$, где

L — целое число, значения которого могут лежать в пределах от $|\sum_k l_k|_{\min}$ до $|\sum_k l_k|_{\max}$. Аналогично $|S|$ может принимать значения $\sqrt{S(S+1)}\hbar$, где S — целое для четного числа электронов и полуцелое для нечетного: $|\sum_k s_k|_{\min} \leq S \leq |\sum_k s_k|_{\max}$.

Полный момент количества движения атома

$$\mathbf{J} = \mathbf{L} + \mathbf{S}$$

также принимает только квантованные значения:

$$|J| = \sqrt{J(J+1)}\hbar,$$

где J — целое число при четном числе электронов и полуцелое при нечетном; $J = L+S, L+S-1 \dots |L-S|$. При полном заполнении всех оболочек $J=0$. Степень вырождения уровня энергии при заданном J равна $2(J+1)$.

Если атом с полным моментом \mathbf{J} поместить в поле, то происходит пространственное квантование и проекция \mathbf{J} на направление поля может принимать значения

$$M_J \hbar,$$

где

$$M_J = 0, \pm 1, \pm 2 \dots$$

В магнитном поле состояния с различными M_J имеют различную энергию (эффект Зеемана), в электрическом поле различную энергию имеют состояния с различными $|M_J|$ (эффект Штарка).

3. Энергетические уровни молекулы. Энергетические уровни молекулы имеют более сложную схему, чем атомов. В молекуле наряду с движением электронов относительно ядер возможно колебательное движение атомов относительно друг друга и вращательное движение молекулы как целого.

Полная энергия невозмущенной молекулы равна:

$$W = W_{эл} + W_{кол} + W_{вр},$$

где $W_{эл}$, $W_{кол}$, $W_{вр}$ — соответственно электронная, колебательная и вращательная энергии.

Энергия внутримолекулярных колебаний и вращения квантована. При этом $W_{эл} \gg W_{кол} \gg W_{вр}$. Поэтому к резким электронным уровням атомов добавляется последовательность колебательных уровней, заканчивающихся сверху границей, за которой начинается сплошной спектр, соответствующий распаду молекулы на части (диссоциация). К каждому колебательному уровню примыкает последовательность вращательных уровней.

Расстояние между электронными уровнями молекулы, как и в атоме, 1—2 эВ. Электронные уровни молекулы расщепляются, так как любой атом молекулы находится в электрическом поле ее остальных атомов. Расщепление определяется эффектом Штарка.

У двухатомной молекулы схема энергетических уровней наиболее проста. Колебательное движение в этом случае соответствует одной степени свободы. При малых энергиях молекулу можно рас-

смагивать как гармонический осциллятор и ее уровни определяются выражением

$$W_{\text{кол}} = \omega_0 \left(v + \frac{1}{2} \right) \hbar,$$

где v — колебательное квантовое число, $v=0, 1, 2, 3 \dots$; ω_0 — частота собственных колебаний. Расстояние между уровнями — десятые и сотые доли электронвольт. При малой энергии расстояния между уровнями одинаковые. При больших энергиях уровни сближаются.

Вращение двухатомной молекулы происходит вокруг оси симметрии молекулы. Энергия вращения определяется выражением

$$W_{\text{вр}} = BJ(J+1),$$

где $J=0, 1, 2, 3, 4 \dots$ — вращательное квантовое число; $B = \frac{\hbar}{2I_{\text{н}}}$ — постоянная; $I_{\text{н}}$ — момент инерции молекулы.

Расстояние между вращательными уровнями — тысячные и десятитысячные доли электронвольт; оно увеличивается с увеличением вращательной энергии.

Вращательный момент количества движения молекулы J квантован и может принимать значения $|J| = \sqrt{J(J+1)} \hbar$.

Проекция вращательного момента на выделенное направление может принимать лишь определенные значения

$$M_J \hbar,$$

где $M_J=0, \pm 1, \pm 2, \pm 3, \pm \dots, \pm J$.

У многоатомных молекул схема уровней еще сложнее. Молекула, состоящая из N атомов, имеет $r=3N-6$ колебательных степеней свободы. Каждой степени свободы соответствуют свои колебательные уровни со своей частотой нормальных колебаний ω_{0i} :

$$W_{\text{кол}} = \sum_{i=1}^r W_{\text{кол}i} = \sum_{i=1}^r \omega_{0i} \left(v_i + \frac{1}{2} \right) \hbar.$$

У молекул, состоящих из большого числа атомов, имеется большое число колебательных и вращательных энергетических уровней. В действительности уровни не являются бесконечно узкими, а имеют конечную ширину. Если расстояние между соседними уровнями меньше их ширины, то уровни сливаются в энергетические полосы.

В средах наблюдается взаимодействие частиц друг с другом. Газы из-за слабого взаимодействия частиц друг с другом имеют узкие энергетические полосы. В твердых и жидких средах отдельные атомы и молекулы могут находиться настолько близко друг к другу, что их внешние электронные оболочки могут перекрываться и взаимодействовать друг с другом. Это приводит к тому, что в твердых и жидких средах вместо узких полос образуются широкие энергетические полосы.

4. Распределение Больцмана. Населенность энергетического уровня определяется числом частиц в единице объема вещества, находящихся на данном уровне, деленном на статистический вес (степень вырождения). Если вырождение отсутствует, то населенность

определяется числом частиц в единице объема вещества, находящихся на данном энергетическом уровне.

При условии термодинамического равновесия населенность состояний атома или молекулы определяется выражением

$$\frac{N_i}{g_i} = \frac{N_0 e^{-\frac{W_i}{kT}}}{\sum_i g_i e^{-W_i/kT}},$$

которое называется распределением Больцмана. Здесь N_i — число частиц на i -м уровне в единице объема (если одному уровню энергии соответствует несколько состояний, то населенность их одинакова); g_i — статистический вес состояния i , W_i — энергия, соответствующая уровню i , k — постоянная Больцмана, равная $1,38 \times 10^{-23}$ Дж/К; N_0 — полное число частиц на всех уровнях в единице объема; T — абсолютная температура. Суммирование производится по всем уровням.

Согласно распределению Больцмана населенность состояний возрастает с ростом температуры T . Однако при любой температуре населенность верхних состояний меньше нижних. При абсолютном нуле ($T=0$ К) все частицы находятся на самом нижнем уровне (основное состояние). Система, подчиняющаяся распределению Больцмана, может только поглощать энергию, т. е. является *пассивной*.

В частности, распределение Больцмана описывает и соотношение между населенностями различных уровней атомов или молекул:

$$\frac{N_n}{N_m} = \frac{g_n}{g_m} e^{-\frac{W_n - W_m}{kT}}. \quad (\text{П1-2})$$

Если вырождение уровней отсутствует, то распределение Больцмана имеет вид:

$$N_i = \frac{N_0 e^{-\frac{W_i}{kT}}}{\sum_i e^{-W_i/kT}},$$

а соотношение между населенностями различных уровней определяется формулой

$$\frac{N_n}{N_m} = e^{-\frac{W_n - W_m}{kT}}. \quad (\text{П1-3})$$

5. Энергетические зоны твердых кристаллических полупроводниковых сред. В отличие от атомов и молекул, где движение электронов локализовано, в кристаллических средах происходит движение валентных электронов от атома к атому, которые находятся в узлах кристаллической решетки. Так как энергетические уровни электронов зависят от положения их относительно кристаллической решетки, а число электронов велико, то число энергетических уровней чрезвычайно увеличивается. Кроме того, в твердых средах атомы и моле-

кулы находятся так близко друг к другу, что их внешние электронные оболочки могут перекрываться и взаимодействовать друг с другом. Эти взаимодействия вызывают сдвиги и расщепление уровней. В результате уровни перекрываются, образуя *разрешенные энергетические зоны*, разделенные областями запрещенных значений — *запрещенными зонами*.

Внутренние электроны атомов подвергаются меньшему влиянию соседних атомов, чем внешние валентные электроны. А так как расширение дискретного уровня в зону происходит из-за влияния соседних атомов, то зоны с низкими энергиями, соответствующие внутренним электронам, узкие (рис. П1-4).

Согласно принципу Паули на N уровнях каждой разрешенной зоны может находиться не более $2N$ электронов (по два электрона, различающихся спинами, на каждом уровне). Верхняя, целиком заполненная валентными электронами зона называется *валентной*, последующая, пустая при $T=0$ К — *зоной проводимости*. Электроны, находящиеся в валентной зоне, не дают вклада в проводимость среды, так как под влиянием внешнего поля не происходит их ускорения (увеличения энергии) из-за отсутствия свободных состояний, и электрический ток не возникает. В зоне проводимости электроны ускоряются внешним полем (переходят на свободные уровни) и возникает электрический ток. Такие электроны называются *электронами проводимости*.

При температуре $T=0$ К валентная зона диэлектриков и полупроводников полностью заполнена, а зона проводимости пуста. Поэтому при $T=0$ К эти среды являются идеальными диэлектриками. При $T \neq 0$ К происходит перескок электронов из валентной зоны в зону проводимости, но у диэлектриков ширина запрещенной зоны $\Delta W \gg kT$ (≥ 3 эВ) поэтому проводимость мала и среда остается диэлектриком при всех реальных температурах. У полупроводников ширина запрещенной зоны менее 3 эВ и проводимость заметна уже при комнатной температуре. У проводников при $T=0$ К валентная зона и зона проводимости перекрываются (рис. П1-5).

Если в полупроводниковой среде электрон при температуре $T \neq 0$ К переходит в зону проводимости, в валентной зоне образуется незанятое место — *дырка*. При отсутствии электрического поля дырка перемещается беспорядочно за счет перескока электронов. Под действием электрического поля дырка перемещается в направлении электрического поля, т. е. ведет себя как положительный заряд.

Электроны проводимости и дырки являются носителями тока в полупроводниках. Процесс образования пары дырка — электрон проводимости может происходить при нагревании. С повышением температуры возрастает амплитуда колебаний атомов решетки и электрон валентной зоны может получить энергию от решетки, достаточную для перехода в зону проводимости. Это тепловая *генерация электронно-дырочных пар*.

Одновременно с возникновением электронно-дырочных пар происходит процесс *рекомбинации*. Спонтанно электроны из зоны проводимости переходят на свободные уровни валентной зоны. При этом электрон проводимости и дырка исчезают, рекомбинируют.

Вероятность генерации электронно-дырочных пар мала, так как средняя энергия, получаемая электроном от решетки, порядка kT и при комнатной температуре составляет около 0,025 эВ, ширина же запрещенной зоны $\Delta W \gg kT$. Чем выше температура, тем больше

электронов переходит в зону проводимости и тем больше проводимость полупроводника.

При тепловом равновесии число электронов проводимости определяется выражением

$$n = AT^{3/2} e^{-\frac{\Delta W}{2kT}},$$

где A — постоянная.

Носители тока (электроны и дырки), возникающие при тепловом равновесии, называются *равновесными*.

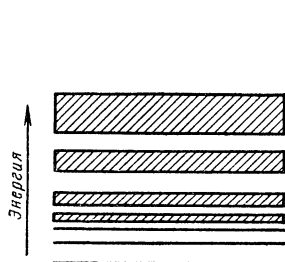


Рис. П1-4. Энергетические зоны кристаллического полупроводника.

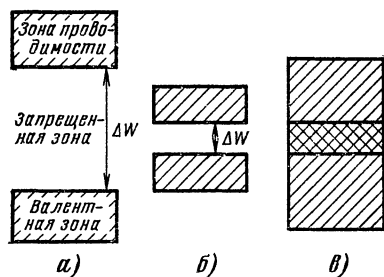


Рис. П1-5. Зона проводимости и валентная зона.

a — диэлектрика; $б$ — полупроводника; $в$ — проводника.

Чистый полупроводник, не содержащий примесей, называется *собственным полупроводником*. В собственном полупроводнике концентрации равновесных электронов n и дырок p равны:

$$n = p = n_i$$

(величина n_i называется *концентрацией носителей* в беспримесном полупроводнике).

Атомы чистого полупроводника располагаются в кристаллической решетке на таких расстояниях, что внешние электронные оболочки атомов перекрываются. У электронов соседних атомов появляются общие орбиты, на которых в соответствии с принципом Паули находится по два электрона с противоположными спинами. Общие орбиты связывают атомы, образуя так называемые ковалентные или парноэлектронные связи. При этом каждый атом дополняет свою внешнюю оболочку до восьми электронов. Наибольшее применение нашли такие чистые полупроводники, как германий и кремний, относящиеся к IV группе периодической системы Менделеева. Атомы этих полупроводников на внешней оболочке имеют по четыре электрона.

Добавление примесей из соседних групп таблицы Менделеева приводит к образованию примесных полупроводников донорного или акцепторного типа. Примесные атомы вводятся в решетку как заместители основных атомов в небольших концентрациях.

Если основные атомы относятся, например, к IV группе, а вводятся атомы V группы с пятью валентными электронами, то четыре

из них образуют ковалентную связь с основными атомами, а пятый электрон становится избыточным. Такой электрон слабее связан с атомом и для превращения его в свободный носитель заряда (перевод в зону проводимости) требуется меньше энергии, чем для освобождения электрона из ковалентной связи. Таким образом, уровни, на которых находятся эти избыточные электроны, должны располагаться в запрещенной зоне вблизи от дна зоны проводимости. Такой примесный полупроводник называется *донорным*.

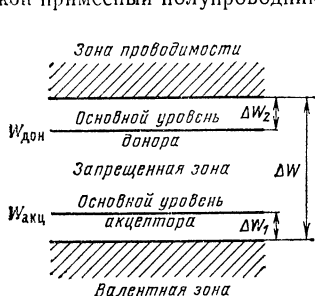


Рис III-6. Энергетическая схема примесных полупроводников

Добавление примеси из III группы, атомы которой имеют три валентных электрона, к полупроводнику IV группы приводит к образованию трех ковалентных связей. Четвертая связь остается незаполненной. При небольшом тепловом возбуждении, значительно меньшем, чем требуется для перевода электрона основного атома в зону проводимости и образования в валентной зоне дырки, электрон одного из основных атомов переходит в незаполненную связь. В той связи, откуда электрон ушел, образуется дырка. Таким образом, наличие примеси приводит к появлению дополнительного уровня в запрещенной зоне вблизи

от потолка валентной зоны. Такой примесный полупроводник называется *акцепторным*.

При температуре абсолютного нуля валентная зона и основной уровень донора будут полностью заполнены электронами, зона проводимости и основной уровень акцептора — пусты (рис. П1-6).

При температуре $T \neq 0$ К, если примесь только донорная, электроны с уровня донора будут переходить в зону проводимости. Так как ширина энергетического барьера ΔW_2 значительно меньше ширины запрещенной зоны ΔW , то число донорных электронов, перешедших в зону проводимости, будет гораздо больше, чем число электронов, перешедших из валентной зоны. Основными носителями тока в этом случае будут электроны в зоне проводимости. Такие полупроводники называются *электронными* или *n-типа*.

Если примесь только акцепторная и $T \neq 0$ К, то наиболее вероятным является переход электронов из валентной зоны на уровни акцептора. При этом эти уровни полностью заполняются и, электроны, находящиеся на уровнях акцепторов, в создании тока не участвуют. В валентной зоне образуются дырки, которые и являются основными носителями тока. Такие полупроводники называются *дырочными* или *p-типа*.

В примесных полупроводниках произведение концентраций равновесных электронов n и дырок p при данной температуре T не зависит от концентрации примесей:

$$n_i^2 = np,$$

где n_i — концентрация носителей в беспримесном полупроводнике при той же температуре.

Практически полупроводники не являются ни абсолютно чистыми, ни идеальными n или p -типа. Реальный полупроводник рассматривается как полупроводник n -типа, если число электронов в зоне проводимости значительно больше, чем дырок в валентной зоне; в противном случае полупроводник относят к p -типу.

Особый интерес представляет система полупроводников n и p -типа, имеющих общую границу раздела, которая называется *электронно-дырочным* или *p - n переходом*. Так как концентрация электронов и дырок по обе стороны p - n перехода различна, то происходит диффузия дырок из полупроводника p -типа в полупроводник n -типа и электронов из полупроводника n -типа в полупроводник p -типа. При этом область p -типа вблизи границы раздела заряжается отрицательно, n -типа — положительно. Возникающее при этом контактное электрическое поле задерживает диффузию и перемещает электроны и дырки от границы раздела в глубь полупроводников; граничный слой толщиной в несколько десятков микрон, обедненный носителями тока, называется запирающим.

6. Распределение Ферми. Это распределение характеризует населенность электронами энергетических уровней квантовой системы (атома, молекулы, твердой среды). При температуре $T=0$ К в квантовой системе с одним электроном электрон займет уровень, соответствующий наименьшей энергии. Увеличение числа электронов в системе приводит к последовательному заполнению электронами уровней все с большей энергией. Энергия последнего заполненного уровня при $T=0$ К называется *уровнем Ферми*. Чем больше число электронов в системе, тем выше находится уровень Ферми.

Распределение Ферми определяет вероятность заполнения электронами уровня с энергией W при любой температуре и имеет вид:

$$f_s(W) = \frac{1}{\frac{W - W_F}{e k T} + 1}, \quad (\text{П1-4})$$

где W_F — химический потенциал, определяемый энергией, на которую возрастает энергия системы при добавлении одного электрона.

Среднее число электронов на данном уровне с энергией W определяется выражением

$$N(W) = f_s(W) g(W),$$

где $g(W)$ — степень вырождения уровня с энергией W .

Согласно (П1-4) при температуре $T=0$ К вероятность заполнения электронами уровня с энергией $W \leq W_F$ равна 1, а с энергией $W > W_F$ равна нулю. Распределение Ферми имеет вид ступеньки (рис. П1-7), когда все уровни ниже уровня W_F заполнены, а выше — пусты. При $T=0$ К уровень W_F является уровнем Ферми.

Если температура $T \neq 0$ К, то электроны с энергией, меньшей W_F , за счет теплового движения могут перебрасываться на уровни, энергия которых больше W_F . Поэтому вероятность пребывания электронов на уровнях с энергией $W > W_F$ становится отличной от нуля, а на уровнях с энергией $W < W_F$ — меньшей единицы (рис. П1-7). При $W = W_F$ и $T \neq 0$ К $f_s(W) = 1/2$, т. е. при $T \neq 0$ К химический потенциал W_F соответствует уровню, вероятность заполнения которого электронами равна $1/2$. Этот уровень в физике полупроводников называют *уровнем Ферми*.

Вероятность заполнения электроном любого уровня с энергией W при любой температуре T , как в валентной зоне, так и зоне проводимости, определяется выражением (П1-4).

Так как вероятность обнаружить на одном уровне электрон или дырку равна единице, т. е.

$$f_v(W) + f_d(W) = 1,$$

то распределение Ферми для дырок будет иметь вид:

$$f_d(W) = \frac{1}{e^{\frac{W_F - W}{kT}} + 1}. \quad (\text{П1-5})$$

Для полупроводников ширина запрещенной зоны $\Delta W \gg kT$ и для собственных полупроводников энергия электрона в зоне проводимости удовлетворяет условию

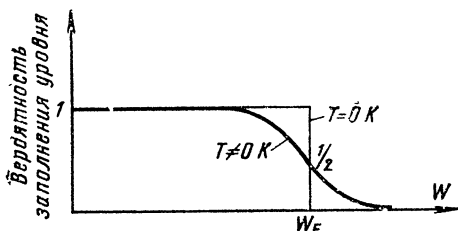


Рис. П1-7. Распределение Ферми при температурах $T=0$ К и $T \neq 0$ К.

$$W - W_F \gg kT.$$

При этом распределение Ферми (П1-4) переходит в распределение Больцмана

$$f_v(W) = e^{-\frac{W - W_F}{kT}}$$

[сравните с формулой (П1-3)], если энергия отсчитывается вверх от уровня Ферми.

При $W_F - W \gg kT$ распределение Ферми (П1-5) для дырок также переходит в распределение Больцмана

$$f_d(W) = e^{-\frac{W_F - W}{kT}};$$

в этой формуле энергия дырок W отсчитывается от уровня Ферми вниз

Полупроводники, к которым применимо распределение Больцмана, называются *невырожденными*.

Уровень Ферми в собственных полупроводниках лежит посредине запрещенной зоны. Действительно, при больцмановском рас-

пределении тепловое возбуждение электронов и дырок происходит так, как если бы уровень Ферми был для них основным. Если электрон переходит с верхнего уровня валентной зоны на нижний уровень зоны проводимости, то на это затрачивается энергия ΔW , равная ширине запрещенной зоны. Но одновременно с появлением электрона в зоне проводимости появляется дырка в валентной зоне. Следовательно, энергия ΔW делится пополам: половина на образование дырки, половина на образование электрона проводимости, и начало отсчета энергии этих процессов (уровень Ферми) лежит посередине запрещенной зоны. При температуре $T=0$ К считают, что уровень Ферми также совпадает с серединой запрещенной зоны.

Для примесных полупроводников положение уровня Ферми зависит от типа полупроводника, концентрации примеси и температуры. При низкой концентрации полупроводники не вырождены и уровень Ферми представляет собой энергию начала отсчета при образовании пары электрон — дырка.

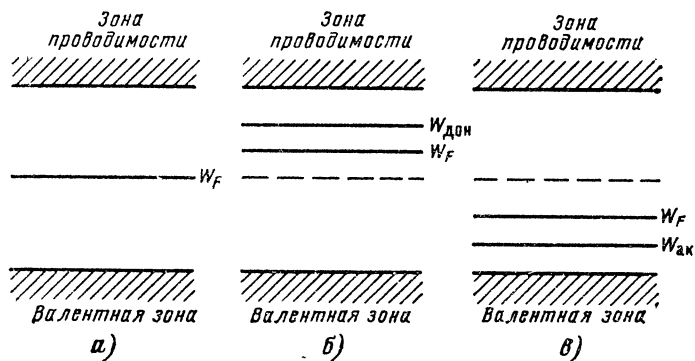


Рис. П1-8 Положение уровня Ферми при температуре $T \neq 0$ К для полупроводников.

а — собственного; б — донорного; в — акцепторного.

При температуре $T=0$ К уровень Ферми полупроводника с донорной примесью лежит в середине зоны ΔW_1 между нижним уровнем (дном) зоны проводимости и уровнем донора, а для полупроводников с акцепторной примесью — в середине зоны ΔW_2 между верхним уровнем (потолком) валентной зоны и уровнем акцептора. При низких температурах в основном происходят переходы между соответствующими зонами и уровнями примеси. Число переходов зона — зона незначительно. При повышении температуры число переходов зона — зона увеличивается и при достаточно высокой температуре эти переходы будут преобладать над переходами зона — уровень примеси. При этом в случае донорной примеси уровень Ферми опускается вниз к середине запрещенной зоны ΔW (рис. П1-8, б), т. е. стремится к положению уровня Ферми для собственного полупроводника (рис. П1-8, а). В случае проводника с акцепторной примесью уровень Ферми поднимается вверх, приближаясь к середине зоны ΔW (рис. П1-8, в).

При данной температуре с ростом концентрации донорной примеси уровень Ферми поднимается вверх, приближаясь к зоне проводимости. С ростом концентрации акцепторной примеси уровень Ферми опускается вниз к валентной зоне. При достаточно большой концентрации донорной примеси уровни примесных атомов образуют полосу, сливающуюся с зоной проводимости. При этом уровень Ферми оказывается в зоне проводимости (рис. П1-9, а). Такой полупроводник называется *вырожденным полупроводником n-типа*. При температуре $T=0$ К электроны заполняют не только все уровни валентной зоны, но и все уровни зоны проводимости, лежащие

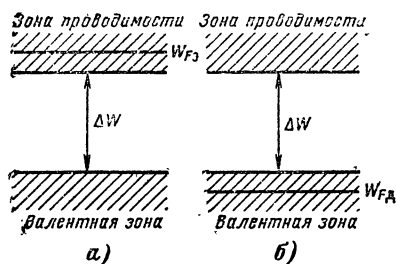


Рис. П1-9. Положение уровня Ферми для вырожденных полупроводников.

а — n-типа; б — p-типа.

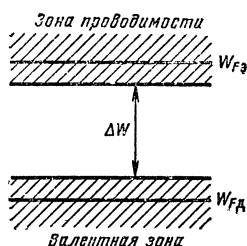


Рис. П1-10. Электронный и дырочный квазиуровни Ферми для вырожденного полупроводника.

ниже уровня Ферми (вырождение электронов). При повышении температуры электроны переходят на уровни с большей энергией, уровень Ферми опускается вниз и при достаточно большой температуре может оказаться в запрещенной зоне, т. е. вырождение исчезает. Температура, при которой исчезает вырождение, тем выше, чем больше концентрация донорной примеси.

Аналогично при достаточно большой концентрации акцепторной примеси уровень Ферми оказывается в валентной зоне (рис. П1-9, б). Такой полупроводник называется *вырожденным полупроводником p-типа* (вырождение дырок). При температуре $T=0$ К уровни валентной зоны, лежащие выше уровня Ферми, будут пусты (заполнены дырками). При повышении температуры электроны заполняют верхнюю часть валентной зоны, уровень Ферми поднимается вверх. При достаточно большой температуре уровень Ферми переходит в запрещенную зону, вырождение исчезает.

К вырожденным полупроводникам распределение Больцмана не применимо, в этом случае имеет место распределение Ферми.

При тепловом равновесии в вырожденных полупроводниках n-типа почти нет дырок, а в вырожденных полупроводниках p-типа очень мало электронов проводимости. Однако под влиянием внешнего воздействия электроны переходят из валентной зоны в зону проводимости. Генерируемые под влиянием внешнего воздействия электронно-дырочные пары рекомбинируют не сразу. Дополнительные электроны и дырки называются *неравновесными носителями тока*. При этом

$$np > n_i^2,$$

где n — концентрация электронов; p — концентрация дырок; n_i — концентрация носителей в чистом полупроводнике.

Время жизни неравновесных носителей много больше времени релаксации и неравновесные носители до рекомбинации распределяются по энергиям, определяемым распределением Ферми (П1-4) и (П1-5), но уровни Ферми для электронов и дырок различны. В этом случае их называют электронным W_{Fe} и дырочным W_{Fd} квазиуровнями Ферми (рис. П1-9, а, б). Если электроны и дырки вырождены одновременно, то расстояние между квазиуровнями Ферми превышает ширину запрещенной зоны (рис. П1-10).

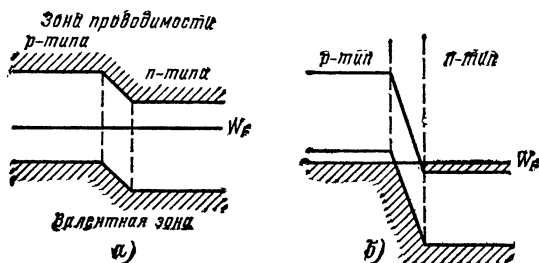


Рис. П1-11. Уровни Ферми для p - n перехода.

а — при малой концентрации (вырождения нет); б — при большой концентрации (вырождение).

В случае, когда полупроводник n -типа находится в контакте с полупроводником p -типа, уровень Ферми один и тот же для обоих типов. При малых концентрациях примеси уровень Ферми лежит в запрещенной зоне (рис. П1-11, а), при большой концентрации уровень Ферми находится в зоне проводимости для полупроводника n -типа и в валентной зоне для полупроводника p -типа (рис. П1-11, б).

7. Квантовые переходы. Переход системы (атома, молекулы) из одного энергетического состояния, соответствующего определенному уровню энергии, в другое, соответствующее другому уровню, называется *квантовым переходом*. При переходе с более низкого на более высокий уровень система поглощает энергию, при переходе с более высокого на более низкий — отдаёт.

Различают *излучательные* и *безызлучательные* квантовые переходы. При безызлучательных переходах изменение энергии не связано с излучением, т. е. испусканием или поглощением фотонов; энергия системы изменяется при взаимодействии с другими атомами или окружающей средой (взаимодействие с кристаллической решеткой, соударение атомов газа друг с другом и стенками сосуда и т. д.). При излучательных переходах система излучает или поглощает электромагнитное излучение в виде квантов, энергия которых равна

$$\hbar\omega_{ik} = W_i - W_k.$$

Для атома или иона могут иметь место переходы, удовлетворяющие правилам отбора:

$$\Delta L = \pm 1, \Delta J = 0, \pm 1 \text{ (кроме } 0 \rightarrow 0);$$

$$\Delta M_J = 0, \pm 1 \text{ (кроме } 0 \rightarrow 0 \text{ для } \Delta J = 0).$$

Совокупность излучательных квантовых переходов с нижних уровней на верхние образует *спектр поглощения*, совокупность переходов с верхних уровней на нижние образует *спектр излучения*.

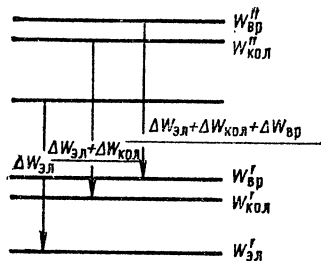


Рис П1-12. Электронно-колебательно-вращательные переходы.

бательными уровнями $\Delta W_{\text{кол}}$ и еще меньшие при переходах между вращательными уровнями $\Delta W_{\text{вр}}$ (рис. П1-12). Боровское условие частот в этом случае имеет вид:

$$\omega = \frac{\Delta W_{\text{эл}} + \Delta W_{\text{кол}} + \Delta W_{\text{вр}}}{\hbar} = \omega_{\text{эл}} + \omega_{\text{кол}} + \omega_{\text{вр}}.$$

Переходы, связанные с изменением $W_{\text{эл}}$, $W_{\text{кол}}$ и $W_{\text{вр}}$, образуют электронно-колебательно-вращательные спектры, лежащие в видимой и ультрафиолетовой областях. При $\Delta W_{\text{эл}} = 0$, $\Delta W_{\text{кол}} \neq 0$ и $\Delta W_{\text{вр}} \neq 0$ получаются колебательно-вращательные спектры, лежащие в инфракрасной области. При $\Delta W_{\text{эл}} = 0$, $\Delta W_{\text{кол}} = 0$, $\Delta W_{\text{вр}} \neq 0$ — вращательные спектры, лежащие в далекой инфракрасной области и в области сантиметровых и миллиметровых волн (диапазон СВЧ). Диапазону СВЧ соответствуют переходы между подуровнями, образующимися при расщеплении уровней энергии в электрических (эффект Штарка) и в магнитных (эффект Зеемана) полях. Спектры молекул имеют *полосатую структуру*.

Спектры атомов и молекул могут быть и сплошными, если квантовые переходы осуществляются между дискретными уровнями энергии и непрерывными полосами энергии, лежащими выше энергии ионизации для атома и энергии диссоциации для молекул.

В полупроводниках возможны следующие переходы электронов: зона — зона, зона — уровень примеси и переходы между уровнями примесей. Переходы на более высокие уровни связаны с увеличением энергии электрона и происходят с поглощением энергии извне, а переходы на более низкие уровни сопровождаются выделением

энергии. Выделяющаяся энергия передается кристаллической решетке или излучается в виде электромагнитных колебаний.

При переходах между зонами различают прямые и непрямые переходы, связанные с зависимостью энергии электрона и дырки от импульса. Электроны, находящиеся вблизи дна зоны проводимости и дырки вблизи потолка валентной зоны, ведут себя как свободные

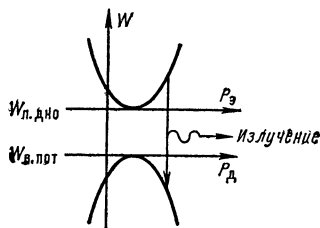


Рис. П1-13. Зависимость энергии электронов проводимости и дырок от их импульсов. Прямой переход.

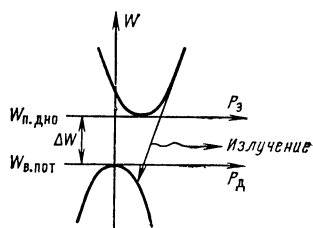


Рис. П1-14. Зависимость энергии электронов проводимости и дырок от их импульсов. Непрямой переход.

движущиеся частицы. Энергия электронов вблизи дна зоны проводимости $W_{п\text{ дно}}$ растет с ростом импульса:

$$W = W_{п\text{ дно}} + \frac{p^2}{2m^*};$$

энергия дырки вблизи потолка валентной зоны $W_{в\text{ пот}}$ падает:

$$W = W_{в\text{ пот}} - \frac{p^2}{2m'^*}.$$

Здесь m^* и m'^* — эффективные массы электрона и дырки, отличающиеся от массы свободного электрона и не равные друг другу. Эффективные массы электрона и дырки зависят от структуры кристаллической решетки и могут быть как больше, так и меньше массы свободного электрона.

Зависимость энергии электрона и дырки от импульса имеет вид двух парабол, расположенных друг над другом. Расстояние между вершинами парабол равно ширине запрещенной зоны ΔW (рис. П1-13).

Переход электрона из зоны проводимости на вакантное место в валентной зоне называется *прямым* (вертикальным), если он не сопровождается изменением импульса частицы (рис. П1-13). Переход называется *непрямым* (рис. П1-14), если он сопровождается изменением импульса электрона, за счет испускания или поглощения кванта звука (фонона).

В случае прямого перехода

$$p_1 = p_2 \pm p_{\text{фот}},$$

где p_1 и p_2 — соответственно начальный и конечный импульсы электрона; $p_{\text{фот}}$ — импульс фотона. Так как

$$p_{\text{фот}} = \frac{\Delta W}{c},$$

где ΔW — ширина запрещенной зоны (менее 3 эВ); c — скорость света, то $p_{\text{фот}}$ мало, и

$$p_1 \approx p_2.$$

Прямой переход происходит в том случае, когда импульсы встречающихся электрона и дырки одинаковы, т. е. в полупроводниках, для которых вершины парабол (рис. П1-13) расположены одна под другой, что свойственно лишь определенным кристаллическим структурам. В случае непрямого перехода

$$p_1 = p_2 \pm p_{\text{фон}} \pm p_{\text{фон}},$$

где $p_{\text{фон}}$ — импульс фонона, передаваемый или отбираемый у кристаллической решетки, за счет чего колебательное состояние решетки изменяется. Так как

$$p_{\text{фон}} = \frac{\Delta W}{v_{\text{зв}}},$$

где $v_{\text{зв}}$ — скорость распространения звука в среде, а $v_{\text{зв}} \ll c$, то

$$p_{\text{фон}} \gg p_{\text{фот}}$$

и

$$p_1 \neq p_2.$$

В этом случае импульсы встречающихся дырок и электрона не равны и вершины парабол сдвинуты (рис. П1-14).

8. Уравнение Шредингера. Согласно представлениям квантовой механики изменение энергии поля происходит не непрерывно, поле квантовано. Оно излучается и поглощается в виде квантов. Энергия кванта определяется выражением

$$W = \hbar \omega,$$

где $\hbar = h/2\pi$; $h = 6,6 \cdot 10^{-34}$ Дж·с — постоянная Планка; ω — круговая частота.

Импульс кванта

$$p = \frac{W}{v} = \hbar \frac{\omega}{v} = \hbar k,$$

где k — волновой вектор; v — скорость распространения.

С другой стороны, волновыми свойствами обладают частицы. Движение внутри атома также является волновым и характеризуется волновой функцией, изменяющейся по периодическому закону в каждой точке пространства

$$\psi(\mathbf{r}, t) = \psi_0 e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})} = \psi_0 e^{i\left(\frac{W}{\hbar} t - \frac{\mathbf{p} \cdot \mathbf{r}}{\hbar}\right)} = \psi_0 e^{i \frac{W}{\hbar} t}, \quad (\text{П1-6})$$

где ψ_0 — амплитуда, которая в случае движения материальной точки массы m зависит только от координат пространства.

Волны, определяемые этой функцией, называются *волнами де Бройля*. Длина волны определяется выражением

$$\lambda = \frac{2\pi\hbar}{p},$$

частота

$$\omega = \frac{W}{\hbar}.$$

Из вида волновой функции (П1-6) очевидно, что она является решением волнового уравнения

$$\Delta\psi + k^2\psi = 0.$$

Учитывая, что для свободной частицы

$$k^2 = \frac{p^2}{\hbar^2} = \frac{2mW_{\text{кин}}}{\hbar^2},$$

где m — масса частицы; $W_{\text{кин}}$ — ее кинетическая энергия, получаем волновое уравнение частицы

$$\Delta\psi + \frac{2m}{\hbar^2} W_{\text{кин}}\psi = 0. \quad (\text{П1-7})$$

Если частица находится под воздействием внешних сил, то ее кинетическая энергия равна:

$$W_{\text{кин}} = W - U, \quad (\text{П1-8})$$

где W — полная энергия; $U = U(r)$ — потенциальная энергия частицы.

Подставляя [П1-8] в [П1-7], получаем уравнение Шредингера

$$\Delta\psi + \frac{2m}{\hbar^2} (W - U)\psi = 0, \quad (\text{П1-9})$$

решения которого (П1-6) являются *собственными волновыми функциями*, характеризующими стационарные состояния системы (частицы или совокупности частиц).

Дифференцируя (П1-6) по времени, получаем:

$$\frac{\partial\psi}{\partial t} = j \frac{W}{\hbar} \psi;$$

отсюда следует, что

$$W\psi = -j\hbar \frac{\partial\psi}{\partial t}. \quad (\text{П1-10})$$

Подставляя (П1-10) в (П1-9), будем иметь уравнение Шредингера в более общем виде:

$$j\hbar \frac{\partial\psi}{\partial t} = \frac{\hbar^2}{2m} \Delta\psi - U\psi. \quad (\text{П1-11})$$

Уравнение Шредингера линейно и его решения удовлетворяют принципу суперпозиции, т. е. если волновые функции ψ_1 и ψ_2 являются решениями и характеризуют два состояния системы, то функция

$$\psi = a_1\psi_1 + a_2\psi_2, \quad (\text{П1-12})$$

где a_1 и a_2 — некоторые постоянные, также является решением.

В общем случае

$$\psi = \sum_n a_n \psi_n. \quad (\text{П1-13})$$

Из решения уравнения Шредингера с учетом граничных условий следует, что полная энергия W может принимать лишь определенные отрицательные собственные значения W_n , которым соответствуют собственные волновые функции ψ_n . Эти значения W_n и ψ_n характеризуют стационарные состояния, соответствующие боровским орбитам. При переходе из одного стационарного состояния в другое энергия меняется скачком.

9. **Волновая функция.** Физический смысл волновой функции (П1-6) определяется выражением

$$\psi\psi^* = |\psi|^2 = |\psi_0|^2.$$

Функция ψ_0 зависит только от координат пространства, и величина $|\psi|^2$ представляет плотность вероятности нахождения частицы в момент времени t в определенной точке пространства. Распределение плотности вероятности нахождения электрона или электронов в атоме — это все, что мы можем знать о движении электрона в атоме.

Электрон может находиться с некоторой вероятностью в любой точке пространства. Он как бы размазан по всему пространству. Функция ψ уменьшается экспоненциально с увеличением расстояния от ядра, так что вероятность нахождения электрона вне области боровских орбит чрезвычайно мала.

В состояниях с положительной энергией собственная функция представляет собой выходящую или входящую сферическую волну, соответствующую удалению или захвату электрона. Здесь ψ уменьшается обратно пропорционально квадрату расстояния от центра, а не экспоненциально.

Собственные функции, соответствующие различным собственным значениям волнового уравнения, взаимно ортогональны, т. е.

$$\int_V \psi_i \psi_k^* dV = 0 \text{ при } i \neq k,$$

где i и k — индексы, определяющие два различных состояния. Интегрирование распространяется на все пространство.

Если собственные функции нормированы, то

$$\int_V \psi_i \psi_k^* dV = \delta_{ik}, \quad (\text{П1-14})$$

где

$$\delta_{ik} = \begin{cases} 1 & \text{при } k = i, \\ 0 & \text{при } k \neq i. \end{cases}$$

10. **Принцип неопределенности.** В волновой механике понятие орбиты теряет свой смысл и положение электрона и его импульс (скорость) не могут быть одновременно измерены с желаемой точностью.

Из вероятностного характера волновой функции следует принцип неопределенности Гейзенберга, согласно которому нельзя одно-

временно точно определить координаты и импульс частицы, координаты и скорость, т. е.

$$\Delta x \Delta p \approx \hbar ;$$

$$\Delta x \Delta v \approx \frac{\hbar}{m} ,$$

а также

$$\Delta W \Delta t \approx \hbar ,$$

где Δx , Δp , Δv , ΔW — соответственно неопределенность определения координаты, импульса, скорости и энергии частицы, Δt — неопределенность определения времени, в течение которого частица может иметь энергию $W_0 \pm \Delta W$.

Из последнего соотношения следует:

$$\Delta t \approx \frac{\hbar}{\Delta W} ,$$

или

$$\Delta \omega \Delta t \approx 1. \quad (\text{П1-15})$$

Чтобы получить сведения о системе, необходимо провести измерения, причем измерения могут дать величины, характеризующие лишь одну сторону процесса. Или мы получаем набор величин, характеризующих положение частиц, или набор величин, характеризующих их импульсы. Создать условия, при которых одновременно можно провести точные измерения тех и других величин, невозможно.

Для квантовых измерений характерна неоднозначность результатов. Ряд измерений какой-либо величины при одних и тех же внешних условиях дает ряд значений, соответствующих вероятности получения этих величин. В процессе измерения измерительный прибор вступает во взаимодействие с квантовой системой, возмущая ее.

11. Квантовые ансамбли. Совокупность одинаковых частиц, находящихся в одинаковых макроскопических условиях (температура, давление, внешнее поле и т. д.), называется *квантовым ансамблем*. Среднее значение любой величины, характеризующее состояние квантового ансамбля, называется средним по ансамблю. Эти средние значения характеризуют и поведение частиц, входящих в ансамбль. С другой стороны, состояние всей макроскопической системы в целом характеризуется средними значениями величин для ансамблей, входящих в ее состав.

Если состояние всех частиц, входящих в ансамбль, до измерения каких-либо величин описывается одной и той же волновой функцией ψ , то ансамбль называется *чистым*. При этом измерение какой-либо величины L может дать целый ряд значений L_1, L_2, L_3 и т. д., и среднее по ансамблю значение равно:

$$\bar{L} = \sum_i \frac{n_i L_i}{n} ,$$

где n — число всех измерений; n_i — число измерений, дающих одинаковое значение L_i .

Каждому значению L_i соответствует собственная функция ψ_i , так как каждое измерение приводит систему к новому состоянию,

и чтобы повторить измерение, необходимо вернуть систему в исходное состояние. При этом согласно принципу суперпозиции (П1-13)

$$\psi = \sum_n a_n \psi_n,$$

комплексно-сопряженная функция

$$\psi^* = \sum_m a_m^* \psi_m^*,$$

плотность вероятности

$$\psi\psi^* = |\psi|^2 = \sum_n |a_n \psi_n|^2 + \sum_{m \neq n} \sum_n a_m^* a_n \psi_m^* \psi_n.$$

Второй член в последнем выражении называется интерференционным и в случае чистого ансамбля он отличен от нуля.

Смешанным называется ансамбль состоящий из групп, состоящие частиц в каждой из которых характеризуется своей собственной функцией ψ_i , т. е. смешанный ансамбль представляет собой совокупность чистых ансамблей.

Плотность вероятности в случае смешанного ансамбля представляется суммой плотностей вероятности всех чистых ансамблей

$$\sum_i |\psi_i|^2 \quad (i = 1, 2, 3, \dots, n),$$

где n — число чистых ансамблей, входящих в сложный смешанный. В этом случае интерференционный член равен нулю.

Состояние частиц, входящих в смешанный ансамбль, нельзя описать одной волновой функцией. Поэтому для описания состояния смешанного ансамбля используется сложная функция, называемая матрицей плотности.

12. Операторы и их свойства. *Оператором* называется математическое понятие, обозначающее совокупность математических действий, устанавливающих соответствие между функциями.

В квантовой механике применяются линейные операторы, удовлетворяющие условию

$$\hat{L}[u_1(x) + u_2(x)] = \hat{L}u_1(x) + \hat{L}u_2(x), \quad (\text{П1-16})$$

где \hat{L} — оператор; $u_1(x)$ и $u_2(x)$ — функции, на которые действует оператор.

Действие оператора на сумму функций эквивалентно сумме результатов действия этого оператора на каждую функцию.

Действие алгебраической суммы операторов \hat{L}_1 и \hat{L}_2 на функцию $u(x)$ эквивалентно алгебраической сумме результатов действий операторов на функцию:

$$(\hat{L}_1 \pm \hat{L}_2)u(x) = \hat{L}_1 u(x) + \hat{L}_2 u(x). \quad (\text{П1-17})$$

Линейные операторы коммутативны

$$\hat{L}_1 + \hat{L}_2 = \hat{L}_2 + \hat{L}_1$$

и ассоциативны

$$\hat{L}_1 + (\hat{L}_2 + \hat{L}_3) = (\hat{L}_1 + \hat{L}_2) + \hat{L}_3.$$

Произведение двух операторов удовлетворяет свойству ассоциативности

$$\left(\hat{L}_1 \hat{L}_2\right) u(x) = \hat{L}_1 \left(\hat{L}_2 u(x)\right)$$

и свойству дистрибутивности

$$\hat{L}_1 \left(\hat{L}_2 + \hat{L}_3\right) = \hat{L}_1 \hat{L}_2 + \hat{L}_1 \hat{L}_3.$$

Свойство коммутативности для произведения в общем случае не имеет места:

$$\hat{L}_1 \hat{L}_2 \neq \hat{L}_2 \hat{L}_1.$$

Перестановочное соотношение для операторов \hat{L}_1 и \hat{L}_2 обозначается как

$$\left[\hat{L}_1 \hat{L}_2\right] = \hat{L}_1 \hat{L}_2 - \hat{L}_2 \hat{L}_1;$$

очевидно,

$$\left[\hat{L}_1 \hat{L}_2\right] = -\left[\hat{L}_2 \hat{L}_1\right].$$

Линейный оператор, удовлетворяющий соотношению

$$\int u_1^*(x) \hat{L} u_2(x) dx = \int u_2(x) \hat{L}^* u_1^*(x) dx, \quad (\text{П1-18})$$

называется *самосопряженным или эрмитовым*. Здесь $u^*(x)$ — функция, комплексно-сопряженная с $u(x)$; \hat{L}^* — оператор, комплексно-сопряженный с оператором \hat{L} .

В квантовой механике применяются только эрмитовы операторы, так как только такие операторы могут изображать вещественные физические величины.

В результате действия оператора на некоторую функцию в общем случае получается новая функция

$$\hat{L} u(x) = \varphi(x).$$

Однако для некоторых функций $u(x)$ в результате действия оператора получается та же функция, умноженная на некоторое постоянное число L :

$$\hat{L} u(x) = L u(x). \quad (\text{П1-19})$$

Число L называется *собственным значением оператора \hat{L}* , а функция $u(x)$, удовлетворяющая уравнению (П1-19), называется *собственной функцией оператора*.

Если оператор является линейным дифференциальным оператором, то уравнение (П1-19) будет линейным дифференциальным уравнением, имеющим при заданных граничных условиях ненулевые решения лишь при определенных значениях L , являющихся собственными значениями. Обычно имеется множество собственных значений и собственных функций. Совокупность всех собственных значений

образует спектр, который может быть как дискретным, так и сплошным.

Собственные значения самосопряженного оператора всегда вещественны. Собственные функции линейного самосопряженного оператора удовлетворяют условию ортогональности

$$\int_V u_m^* u_n dV = 0 \text{ при } m \neq n.$$

Совокупность собственных функций $u_n (n=0, 1, 2 \dots)$ образует полную систему, т. е. любая непрерывная функция, заданная на том же интервале, что и u_n , удовлетворяющая таким же граничным условиям, может быть представлена в виде ряда

$$\Phi = \sum_n a_n u_n, \quad (\text{П1-20})$$

где

$$a_n = \int_V \Phi u_n^* dV.$$

Ряд (П1-20) сходится, если интеграл $\int_V |\Phi|^2 dV$ существует.

В квантовой механике операторы служат для аналитического описания физических величин. При этом значения физической величины, определяемой опытным путем, должны совпадать с собственными значениями оператора. Так как физическая величина представляется вещественным числом, то операторы, соответствующие им, должны быть самосопряженными.

13. **Оператор энергии и импульса.** Уравнение Шредингера (П1-9) можно представить в виде

$$\left[-\frac{\hbar^2}{2m} \Delta + U(x_i) \right] \psi = W\psi,$$

т. е. воздействие оператора

$$\left[-\frac{\hbar^2}{2m} \Delta + U(x_i) \right] \quad (\text{П1-21})$$

на функцию ψ равносильно умножению ее на величину W — полную энергию системы. Таким образом, оператор (П1-21) является *оператором полной энергии*, его называют также *гамильтонианом* и обо-

значают символом \hat{H} . В этом операторе удобно заменить функцию $U(x_i)$ на оператор потенциальной энергии $\hat{U}(x_i)$:

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta + \hat{U}(x_i);$$

при этом уравнение Шредингера для стационарных состояний можно записать в виде

$$\hat{H} \psi = W\psi, \quad (\text{П1-22})$$

или согласно (П1-10) в виде

$$-i\hbar \frac{\partial \psi}{\partial t} = \hat{H} \psi. \quad (\text{П1-23})$$

Так как полная энергия системы равна сумме кинетической и потенциальной энергий, то согласно (П1-23) оператор кинетической энергии определяется выражением

$$\hat{W}_{\text{кин}} = -\frac{\hbar^2}{2m} \Delta. \quad (\text{П1-24})$$

С другой стороны,

$$W_{\text{кин}} = \frac{mv^2}{2} = \frac{p^2}{2m},$$

и, следовательно,

$$\hat{W}_{\text{кин}} = -\frac{\hat{p}^2}{2m}, \quad (\text{П1-25})$$

где \hat{p} — оператор импульса.

Сравнивая (П1-24) и (П1-25), получаем:

$$\hat{p}^2 = \hbar^2 \Delta$$

или

$$\hat{p} = i\hbar \nabla.$$

Проекция оператора \hat{p} на оси координат записываются в виде

$$\hat{p}_i = i\hbar \frac{\partial}{\partial x_i}.$$

14. Среднее значение. Матрицы. Среднее значение (или математическое ожидание) физической величины x определяется выражением

$$\tilde{x} = \int_{-\infty}^{\infty} x f(x) dx, \quad (\text{П1-26})$$

где $f(x)$ — плотность вероятности случайной непрерывной величины.

Если x — координата, то ее среднее значение

$$\tilde{x} = \int_{-\infty}^{\infty} x \psi^*(x) \psi(x) dx = \int_{-\infty}^{\infty} \psi^*(x) x \psi(x) dx,$$

где

$$f(x) = |\psi(x)|^2 = \psi^* \psi.$$

Среднее значение функции $F(x)$

$$\tilde{F}(x) = \int_{-\infty}^{\infty} \psi^* F(x) \psi dx. \quad (\text{П1-27})$$

Среднее значение оператора \hat{L}

$$L = \int_V \psi^* \hat{L} \psi dV, \quad (\text{П1-28})$$

где интегрирование производится по конфигурационному пространству, т. е. по совокупности координат, определяющих положение частицы и ее элементарных частей в пространстве.

Так как L есть действительная измеряемая величина, то

$$L = L^*,$$

а

$$L^* = \int_V \psi \hat{L}^* \psi^* dV, \quad (\text{П1-29})$$

т. е. оператор должен удовлетворять условию (П1-18).

Так как согласно (П1-13)

$$\begin{aligned} \psi &= \sum_n a_n \psi_n; \\ \psi^* &= \sum_m a_m^* \psi_m^*, \end{aligned}$$

то, комбинируя попарно волновые функции ψ_n и ψ_m , получаем согласно (П1-29) последовательность средних значений, в общем случае зависящих от времени,

$$L_{nm}(t) = \int_V \psi_m^* \hat{L} \psi_n dV = L_{mn}^*(t), \quad (\text{П1-30})$$

которые можно представить в виде эрмитовой матрицы, рассматривая L_{nm} как элемент матрицы, где первый индекс относится к номеру строки, а второй к номеру столбца:

$$L(t) = \begin{vmatrix} L_{11} & L_{12} & \dots & \dots & \dots & \dots & L_{1n} \\ L_{21} & L_{22} & \dots & \dots & \dots & \dots & L_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{vmatrix}.$$

Алгебра операторов соответствует алгебре матриц. Переход от операторов к матрицам соответствует переходу от дифференциальных уравнений к алгебраическим.

15. Взаимодействие поля с веществом. Рассмотрим взаимодействие внешнего электрического поля с частицей (атомом или молекулой), представляющей систему зарядов. Пусть начало декартовой системы координат находится внутри системы, а координаты точек, в которых расположены заряды e_i , определяются величинами Δx_1 , Δx_2 , Δx_3 .

Разложим функцию зависимости потенциала внешнего поля от координат $\varphi(x_1, x_2, x_3)$ в ряд Тейлора:

$$\begin{aligned} \varphi_i &= \varphi + \left(\Delta x_1 \frac{\partial \varphi}{\partial x_1} + \Delta x_2 \frac{\partial \varphi}{\partial x_2} + \Delta x_3 \frac{\partial \varphi}{\partial x_3} \right) + \\ &+ \frac{1}{2} \left(\Delta x_1^2 \frac{\partial^2 \varphi}{\partial x_1^2} + \Delta x_2^2 \frac{\partial^2 \varphi}{\partial x_2^2} + \Delta x_3^2 \frac{\partial^2 \varphi}{\partial x_3^2} + 2\Delta x_1 \Delta x_2 \frac{\partial^2 \varphi}{\partial x_1 \partial x_2} + \right. \\ &\left. + 2\Delta x_1 \Delta x_3 \frac{\partial^2 \varphi}{\partial x_1 \partial x_3} + 2\Delta x_2 \Delta x_3 \frac{\partial^2 \varphi}{\partial x_2 \partial x_3} \right) + \dots, \end{aligned}$$

где φ — значение потенциала в начале координат.

Энергия взаимодействия поля со всей системой зарядов определится выражением

$$\begin{aligned} \sum_i e_i \varphi_i = \varphi \sum_i e_i + \sum_i e_i \left(\Delta x_1 \frac{\partial \varphi}{\partial x_1} + \Delta x_2 \frac{\partial \varphi}{\partial x_2} + \Delta x_3 \frac{\partial \varphi}{\partial x_3} \right) + \\ + \sum_i \frac{1}{2} e_i \left(\Delta x_1^2 \frac{\partial^2 \varphi}{\partial x_1^2} + \Delta x_2^2 \frac{\partial^2 \varphi}{\partial x_2^2} + \Delta x_3^2 \frac{\partial^2 \varphi}{\partial x_3^2} + 2 \Delta x_1 \Delta x_2 \frac{\partial^2 \varphi}{\partial x_1 \partial x_2} + \right. \\ \left. + 2 \Delta x_1 \Delta x_3 \frac{\partial^2 \varphi}{\partial x_1 \partial x_3} + 2 \Delta x_2 \Delta x_3 \frac{\partial^2 \varphi}{\partial x_2 \partial x_3} \right) + \dots \quad (\text{П1-31}) \end{aligned}$$

Сумма $\sum_i e_i$ определяет полный заряд системы, не равный нулю для иона; для незаряженной молекулы или атома эта величина равна нулю.

Так как напряженность электрического поля определяется выражением

$$\mathbf{E} = -\text{grad } \varphi = \left(-\frac{\partial \varphi}{\partial x_1}, -\frac{\partial \varphi}{\partial x_2}, -\frac{\partial \varphi}{\partial x_3} \right),$$

а дипольный момент

$$\mathbf{p}_e = (\sum_i e_i \Delta x_1, \sum_i e_i \Delta x_2, \sum_i e_i \Delta x_3),$$

то второй член в выражении для энергии взаимодействия (П1-31) можно представить в виде скалярного произведения

$$-\mathbf{p}_e \mathbf{E}.$$

Молекулы, обладающие конечным дипольным моментом, называются полярными.

Третий член в выражении (П1-31) не равен нулю, только в неоднородном поле. Сама система будет характеризоваться величинами

$$\begin{aligned} \sum_i e_i \Delta x_1^2, \sum_i e_i \Delta x_2^2, \sum_i e_i \Delta x_3^2, \sum_i e_i \Delta x_1 \Delta x_2, \\ \sum_i e_i \Delta x_1 \Delta x_3, \sum_i e_i \Delta x_2 \Delta x_3, \end{aligned} \quad (\text{П1-32})$$

определяющими квадрупольный момент молекулы или атома. Соответствующим выбором направления осей системы координат величины $\sum_i e_i \Delta x_1 \Delta x_2$, $\sum_i e_i \Delta x_1 \Delta x_3$, $\sum_i e_i \Delta x_2 \Delta x_3$ можно свести к нулю, при этом квадрупольный момент системы будет определяться величинами $\sum_i e_i \Delta x_1^2$, $\sum_i e_i \Delta x_2^2$, $\sum_i e_i \Delta x_3^2$. При высокой степени симметрии возможен случай, когда $\sum_i e_i \Delta x_1^2 = \sum_i e_i \Delta x_2^2 = \sum_i e_i \Delta x_3^2$. При этом третий член в выражении (П1-31) будет равен нулю, так как φ удовлетворяет уравнению Лапласа $\Delta \varphi = 0$.

В то время как дипольный момент является вектором, величины (П1-32) определяют тензор второго ранга. Мультипольные моменты более высоких порядков, определенные из соответствующих членов разложения (П1-31), являются тензорами более высокого ранга.

Четвертый член выражения (П1-31) характеризует взаимодействие поля с октупольным моментом системы и т. д.

Обычно дипольное взаимодействие значительно сильнее, чем квадрупольное, квадрупольное взаимодействие значительно сильнее, чем октупольное и т. д. Поэтому если дипольный момент системы имеет конечное значение, то рассматривают взаимодействие системы с полем в дипольном приближении, пренебрегая другими моментами системы. Если дипольный момент системы практически равен нулю, то учитывают следующие по порядку величины моменты, а именно квадрупольный и магнитный.

В классической электродинамике излучение связано с изменением электрического или магнитного момента. Внутриатомное и внутримолекулярное движение приводит к излучению только в том случае, если имеет место переменный момент. Молекула, состоящая из двух различных атомов представляет собой электрический диполь, так как центры тяжести положительных и отрицательных зарядов в этом случае не совпадают. При колебаниях молекулы расстояние между центрами тяжести положительных и отрицательных зарядов изменяется, меняется и значение момента электрического диполя. Определенному квантовому переходу между колебательными уровнями соответствует и определенное изменение дипольного момента. При вращении молекулы меняется проекция дипольного момента на заданное направление с частотой вращения и наблюдается излучение с этой частотой. Определенному квантовому переходу между вращательными уровнями соответствует и определенное изменение этой проекции.

Разрешены лишь те квантовые переходы, которые связаны с изменением дипольного, квадрупольного или магнитного моментов. Если изменения равны нулю, то такие переходы запрещены. Разрешенные переходы определяются правилами отбора.

Постоянный дипольный момент имеют лишь асимметричные молекулы, состоящие из неоднородных атомов. Симметричные молекулы дипольного момента не имеют. Однако в неоднородном внешнем поле их симметрия искажается и появляются электрические мультипольные моменты: квадрупольные, октупольные и т. д. Такие молекулы обладают магнитным моментом.

Дипольный момент отсутствует у молекул H_2 , O_2 , N_2 и т. д. У молекулы CO_2 дипольный момент отсутствует из-за симметрии — все три атома расположены по одной прямой симметрично.

В молекуле H_2O атомы расположены по углам треугольника и она является полярной. Молекула аммиака NH_3 имеет пирамидальную структуру с атомом N в вершине и атомами H в углах треугольника основания. Вследствие асимметрии она также обладает дипольным моментом.

16. Матрица электрического дипольного момента. Элемент этой матрицы согласно (П1-30) определяется выражением

$$p_{em}(t) = \int_V \psi_m^* \hat{p}_e \psi_n dV, \quad (\text{П1-33})$$

где \hat{p}_e — оператор дипольного момента, и интегрирование производится по конфигурационному пространству частицы (атома или молекулы).

Движение электронов можно охарактеризовать вероятностью нахождения их в той или иной точке пространства. Совокупность этих

гочек можно рассматривать как электронное облако. Вероятность нахождения электрона в объеме dV при стационарном состоянии, характеризуемом волновой функцией ψ_n , определяется выражением (см. П1, п. 9)

$$\psi_n^* \psi_n dV,$$

плотность заряда в данном объеме — выражением

$$e\psi_n^* \psi_n,$$

а среднее значение дипольного момента, соответствующее диагональному элементу матрицы (П1-33),

$$\mathbf{p}_{enn} = e \int_V \psi_n^* \psi_n \mathbf{r} dV = e \int_V \psi_n^* \mathbf{r} \psi_n dV, \quad (\text{П1-34})$$

где \mathbf{r} — радиус-вектор, проведенный из начала координат, где находится ядро атома или центр тяжести зарядов ядер молекулы, к произвольной точке электронного облака; интегрирование производится по конфигурационному пространству частицы. При симметричном распределении электронного облака выражение (П1-34) равно нулю, при несимметричном равно постоянной величине. Действительно, в последнем случае, если \mathbf{r} от времени не зависит, то согласно (П1-6) подынтегральное выражение (П1-34) от времени также не зависит.

Переходы между двумя уровнями m и n происходят в обе стороны и с учетом (П1-6) можно записать:

$$\left. \begin{aligned} \mathbf{p}_{enm} &= e \int_V \psi_m^* \mathbf{r} \psi_n dV = e \int_V \psi_{0m}^* \mathbf{r} \psi_{0n} e^{j\omega_{nm}t} dV; \\ \mathbf{p}_{emn} &= e \int_V \psi_n^* \mathbf{r} \psi_m dV = e \int_V \psi_{0n}^* \mathbf{r} \psi_{0m} e^{-j\omega_{nm}t} dV, \end{aligned} \right\} \quad (\text{П1-35})$$

где

$$\omega_{nm} = \frac{W_n - W_m}{\hbar}.$$

Недиагональные элементы матрицы (П1-33) зависят от времени и определяют поглощение или излучение энергии частицей при соответствующих переходах.

Обозначим

$$\begin{aligned} \mathbf{r}_{mn} &= \int_V \psi_{0m}^* \mathbf{r} \psi_{0n} dV; \\ \mathbf{r}_{nm} &= \mathbf{r}_{mn}^* = \int_V \psi_{0n}^* \mathbf{r} \psi_{0m} dV. \end{aligned}$$

Так как это действительные величины, то

$$r_{mn} = r_{nm} = \int_V \psi_m^* r \psi_n dV,$$

и в общем случае электрический дипольный момент может быть представлен в виде эрмитовой матрицы

$$\rho_e^{(t)} = \left\| \begin{array}{cccc} \rho_{e11} & \rho_{e12} e^{j\omega_{12}t} & \dots & \rho_{e1n} e^{j\omega_{1n}t} \\ \rho_{e21} e^{j\omega_{21}t} & \rho_{e22} & \dots & \rho_{e2n} e^{j\omega_{2n}t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{en1} e^{j\omega_{n1}t} & \rho_{en2} e^{j\omega_{n2}t} & \dots & \rho_{enn} \end{array} \right\|, \quad (\text{П1-36})$$

где $\rho_{eik} = \rho_{eki}$, $\omega_{ik} = -\omega_{ki}$.

В случае квантовомеханической системы с двумя энергетическими уровнями, считая, что другие уровни достаточно далеко удалены от этих двух и не взаимодействуют с ними, матрица дипольного момента имеет вид:

$$\rho_e = \left\| \begin{array}{cc} \rho_{e11} & \rho_e e^{j\omega_{12}t} \\ \rho_e e^{j\omega_{21}t} & \rho_{e22} \end{array} \right\|.$$

В большинстве случаев из-за симметричности электронного облака дипольные моменты, соответствующие стационарным состояниям, равны нулю и

$$\rho_e = \left\| \begin{array}{cc} 0 & \rho_e e^{j\omega_{12}t} \\ \rho_e e^{j\omega_{21}t} & 0 \end{array} \right\|. \quad (\text{П1-37})$$

17. Матрица энергии. Согласно (П1-30) элементы этой матрицы определяются выражением

$$H_{nm} = \int_V \psi_m^* \hat{H} \psi_n dV, \quad (\text{П1-38})$$

где \hat{H} — оператор энергии и интегрирование производится по конфигурационному пространству.

Если \hat{H} от времени не зависит, а согласно (П1-6)

$$\begin{aligned} \psi_m^* &= \psi_{0m}^* e^{-j \frac{W_m t}{\hbar}}; \\ \psi_n &= \psi_{0n} e^{j \frac{W_n t}{\hbar}}, \end{aligned}$$

то подынтегральное выражение (П1-38) имеет вид:

$$\psi_m^* \hat{H} \psi_n = \psi_{0m}^* \hat{H} \psi_{0n} e^{j\omega_{nm}t},$$

где

$$\omega_{nm} = \frac{W_n - W_m}{\hbar}.$$

При $m=n$

$$\psi_n^* \hat{H} \psi_n = \psi_{0n}^* \hat{H} \psi_{0n}.$$

Таким образом, диагональные элементы матрицы не зависят от времени и соответствуют стационарным состояниям, недиагональные зависят от времени и соответствуют переходам системы из одного стационарного состояния в другое.

В случае квантовомеханической системы с двумя энергетическими уровнями W_1 и W_2 ($W_1 < W_2$) стационарные состояния описываются

волновыми функциями вида (П1-6) $\psi_{01} e^{j \frac{W_1 t}{\hbar}}$ и $\psi_{02} e^{j \frac{W_2 t}{\hbar}}$. Подставляя эти волновые функции в уравнение (П1-22), получаем:

$$\hat{H}_0 \psi_{01} = W_1 \psi_{01}; \quad \hat{H}_0 \psi_{02} = W_2 \psi_{02},$$

где \hat{H}_0 — невозмущенный гамильтониан системы (частицы), который согласно (П1-30) можно представить в виде матрицы. С учетом условия ортогональности (П1-14) получим:

$$H_{011} = \int_V \psi_{01}^* \hat{H}_0 \psi_{01} dV = W_1;$$

$$H_{012} = \int_V \psi_{01}^* \hat{H}_0 \psi_{02} dV = 0;$$

$$H_{021} = \int_V \psi_{02}^* \hat{H}_0 \psi_{01} dV = 0;$$

$$H_{022} = \int_V \psi_{02}^* \hat{H}_0 \psi_{02} dV = W_2,$$

или

$$\hat{H}_0 = \begin{vmatrix} W_1 & 0 \\ 0 & W_2 \end{vmatrix} = \hbar \begin{vmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{vmatrix}. \quad (\text{П1-39})$$

Под воздействием поля излучения частиц E система уже не будет находиться в стационарном состоянии. Состояние системы будет определяться уравнением (П1-23), которое будет иметь вид:

$$-j\hbar \frac{\partial \psi}{\partial t} = [\hat{H}_0 + \hat{U}(t)] \psi, \quad (\text{П1-40})$$

где $\hat{U}(t)$ — оператор возмущения.

При дипольном взаимодействии

$$U(t) = -pE,$$

действием мультиполей можно пренебречь, так как оно много меньше дипольного взаимодействия (см. П1, п. 15) и уравнение (П1-40) можно представить в виде

$$-j\hbar \frac{\partial \psi}{\partial t} = [\hat{H}_0 - \hat{p} \hat{E}] \psi,$$

где $\hat{H}_0 - \hat{p} \hat{E}$ — гамильтониан возмущенной частицы.

Согласно (П1-39) и (П1-37)

$$\left[\hat{H}_0 - \hat{p} \hat{E} \right] = \begin{vmatrix} W_1 & -pE \\ -pE & W_2 \end{vmatrix} = \hbar \begin{vmatrix} \omega_1 & -\frac{pE}{\hbar} \\ -\frac{pE}{\hbar} & \omega_2 \end{vmatrix}. \quad (\text{П1-41})$$

Частоты $\omega_{nm} = \frac{W_n - W_m}{\hbar}$ также образуют матрицу

$$\omega = \begin{vmatrix} 0 & \omega_{12} & \dots & \dots & \dots & \omega_{1n} \\ \omega_{21} & 0 & \dots & \dots & \dots & \omega_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \omega_{n1} & \omega_{n2} & \dots & \dots & \dots & 0 \end{vmatrix}.$$

18. Матрица плотности. Систему частиц можно подразделить на группы, в каждой из которых частицы находятся в определенных энергетических состояниях с энергией W_i , характеризующимся волновыми функциями ψ_i ($i=1, 2, 3, 4 \dots$). Статистический вес групп определяется выражениями

$$g_i = \frac{N_i}{N},$$

где N_i — число частиц i -й группы; N — полное число частиц. Очевидно,

$$\sum_i g_i = 1.$$

Среднее значение некоторой величины p для всей системы равно:

$$\tilde{p} = \sum_i g_i \tilde{p}_i, \quad (\text{П1-42})$$

где \tilde{p}_i — среднее значение этой величины для i -й группы согласно (П1-27) определяется выражением

$$\tilde{p}_i = \int_V \psi_i^* p_i \psi_i dV. \quad (\text{П1-43})$$

Функции ψ_i и ψ_i^* согласно (П1-13) можно представить в виде

$$\left. \begin{aligned} \psi_i &= \sum_n a_n^{(i)} \psi_n; \\ \psi_i^* &= \sum_m a_m^{(i)*} \psi_m^*. \end{aligned} \right\} \quad (\text{П1-44})$$

Подставляя (П1-43) и (П1-44) в (П1-42), получаем:

$$\tilde{p} = \sum_i g_i \sum_m \sum_n a_m^{(i)*} a_n^{(i)} p_{mn},$$

где $p_{mn} = \int_V \psi_m^* p \psi_n dV$ — матричный элемент, или

$$\tilde{p} = \sum_{m,n} p_{mn} p_{mn}, \quad (\text{П1-45})$$

где

$$\rho_{mn} = \sum_i g_i a_m^{(i)*} a_n^{(i)} \quad (\text{П1-46})$$

— матричный элемент матрицы ρ , называемой *статистической* или *матрицей плотности*.

Диагональные элементы матрицы плотности ρ_{mn} определяют вероятность нахождения системы в данном состоянии ψ_m , поэтому эти элементы не могут быть отрицательными и сумма всех диагональных элементов равна 1. Недиагональные элементы ρ_{mn} характеризуют квантовые переходы из состояния ψ_m в состояние ψ_n (с m -уровня энергии на n -уровень). Так как среднее значение физической величины должно быть вещественно, то матрица плотности является *эрмитовой*.

Выражение (П1-45) можно представить также в виде

$$\tilde{\rho} = \sum_i (\rho\rho)_{kk}, \quad (\text{П1-47})$$

т. е. средняя величина равна сумме диагональных элементов матрицы, представляющей собой произведение матрицы ρ и ρ .

Если система изменяется во времени, то

$$\rho_{mn} = \sum_i q_i a_m^{(i)*}(t) a_n^{(i)}(t);$$

$$\frac{\partial \rho_{mn}}{\partial t} = \sum_i q_i \left[\frac{\partial a_m^{(i)*}}{\partial t} a_n^{(i)} + a_m^{(i)*} \frac{\partial a_n^{(i)}}{\partial t} \right].$$

Величину $\frac{\partial a_n^{(i)}}{\partial t}$ определим с помощью уравнения Шредингера (П1-23)

$$\hat{H}\psi_i = -j\hbar \frac{\partial \psi_i}{\partial t}, \quad (\text{П1-48})$$

представляя ψ_i в виде

$$\psi_i = \sum_n a_n^{(i)} \psi_n,$$

где ψ_n — собственные функции оператора \hat{H} . Подставляя это разложение в (П1-48), умножая обе части на ψ_m^* и интегрируя, получаем

$$-j\hbar \frac{\partial a_m^{(i)}}{\partial t} = \sum_n H_{mn} a_n^{(i)}, \quad (\text{П1-49})$$

где

$$H_{mn} = \int \psi_m^* \hat{H} \psi_n dV.$$

С учетом (П1-49) в (П1-48) и эрмитовости матрицы энергии, т. е. $H_{mi} = H_{im}^*$ получаем

$$-j\hbar \frac{\partial \rho_{mn}}{\partial t} = \sum_i (H_{mi} \rho_{in} - \rho_{mi} H_{in})$$

или

$$\frac{\partial \rho}{\partial t} = \frac{i}{\hbar} (H\rho - \rho H) = [H\rho] \quad (\text{П1-50})$$

— закон изменения матрицы плотности во времени. Здесь $[H\rho] = \frac{i}{\hbar} (H\rho - \rho H)$ — квантовые скобки Пуассона.

19. Вероятность перехода. Частица может находиться только в определенных энергетических состояниях. Под влиянием внешних воздействий она может переходить из одного состояния в другое.

Пусть в момент времени $t < t_1$ частица находится в стационарном состоянии, определяемом волновой функцией $\psi_{0m}(\mathbf{r}) e^{i \frac{W_m t}{\hbar}}$. В промежутке времени $t_1 \leq t \leq t_2$ частица подвергается внешнему воздействию (возмущению) и переходит в новое стационарное состояние, характеризуемое волновой функцией $\psi_{0n}(\mathbf{r}) e^{i \frac{W_n t}{\hbar}}$. Переход частицы из одного состояния в другое — вероятностный процесс. Уравнение Шредингера для данного случая будет иметь вид (П1-23):

$$-i\hbar \frac{\partial \psi}{\partial t} = [\hat{H}_0 + \hat{U}(t)] \psi, \quad (\text{П1-51})$$

где \hat{H}_0 — гамильтониан невозмущенной системы, а

$$\hat{U}(t) = \begin{cases} U_B(t) & \text{при } t_1 \leq t \leq t_2, \\ 0 & \text{при } t < t_1 \text{ и } t > t_2 \end{cases}$$

— оператор возмущения, характеризующий внешнее воздействие.

Уравнение Шредингера для невозмущенной системы имеет вид:

$$-i\hbar \frac{\partial \psi}{\partial t} = \hat{H}_0 \psi;$$

его решение

$$\psi_l = \psi_{0l} e^{i \frac{W_l t}{\hbar}}. \quad (\text{П1-52})$$

Решение уравнения (П1-51) будем искать в виде линейной комбинации волновых функций (П1-52):

$$\psi = \sum_l a_l(t) \psi_l = \sum_l a_l(t) \psi_{0l} e^{i \frac{W_l t}{\hbar}}, \quad (\text{П1-53})$$

где $a_l(t)$ — неизвестные функции времени.

Пусть $t_1 = 0$, $t_2 = \tau$. Так как при $t < 0$ частица находилась в стационарном состоянии m , т. е.

$$\psi|_{t < 0} = \psi_{0m} e^{i \frac{W_m t}{\hbar}},$$

$$a_l(t) = \begin{cases} 1 & \text{при } l = m, \\ 0 & \text{при } l \neq m. \end{cases}$$

При $t > \tau$ система вновь переходит в стационарное состояние. При этом согласно (П1-53)

$$\psi|_{t>\tau} = \sum_l a_l(\tau) \psi_{0l} e^{i \frac{W_l t}{\hbar}}, \quad (\text{П1-54})$$

где коэффициенты $a_l(t) = a_l(\tau)$ постоянны при $t > \tau$.

Величина $|a_l(\tau)|^2$ определяет вероятность нахождения системы в состоянии l . Вероятность нахождения системы в состоянии m при $t < 0$ по условию равна 1, т. е.

$$|a_m(t)|^2 = |a_m(0)|^2 = 1.$$

Вероятность нахождения системы в состоянии n определяется величиной $|a_n(\tau)|^2$. Таким образом, вероятность перехода системы из состояния m в состояние n

$$P_{mn} = |a_n(\tau)|^2.$$

Подставляя выражение (П1-54) в уравнение (П1-51), получаем:

$$\begin{aligned} -j\hbar \sum_l \frac{da_l}{dt} \psi_{0l} e^{i \frac{W_l t}{\hbar}} + \sum_l a_l W_l \psi_{0l} e^{i \frac{W_l t}{\hbar}} &= \hat{H}_0 \sum_l a_l \psi_{0l} e^{i \frac{W_l t}{\hbar}} + \\ &+ \hat{U}_B(t) \sum_l a_l \psi_{0l} e^{i \frac{W_l t}{\hbar}}. \end{aligned} \quad (\text{П1-55})$$

Так как согласно (П1-22)

$$\hat{H}_0 \psi_{0l} = W_l \psi_{0l},$$

то

$$\sum_l a_l W_l \psi_{0l} e^{i \frac{W_l t}{\hbar}} = \sum_l a_l (\hat{H}_0 \psi_{0l}) e^{i \frac{W_l t}{\hbar}}$$

и выражение (П1-55) переходит в равенство

$$j\hbar \sum_l \frac{da_l}{dt} \psi_{0l} e^{i \frac{W_l t}{\hbar}} = \hat{U}_B \sum_l a_l \psi_{0l} e^{i \frac{W_l t}{\hbar}}.$$

Умножая обе части последнего равенства на $\psi_{0n}^* e^{-j \frac{W_n t}{\hbar}}$ и интегрируя по конфигурационному пространству частицы, с учетом условия ортогональности (П1-14) получаем:

$$j\hbar \frac{da_n}{dt} = \sum_l a_l U_{nl} e^{-j\omega_{nl} t}, \quad (\text{П1-56})$$

где

$$\omega_{nl} = \frac{W_n - W_l}{\hbar};$$

$$U_{nl} = \int_V \psi_{0n}^* \hat{U}_B \psi_{0l} dV \quad (\text{П1-57})$$

— матричный элемент энергии возмущения.

Начальные условия можно представить в виде

$$a_l(0) = \delta_{lm} = \begin{cases} 1 & \text{при } l = m, \\ 0 & \text{при } l \neq m. \end{cases}$$

Уравнение (П1-56) можно решить методом последовательных приближений (П2, п. 3), полагая возмущение малым ($U_B \rightarrow 0$). В качестве нулевого приближения определим коэффициенты $a_l(t)$ из начальных условий

$$a_l^{(0)}(t) = \delta_{lm}.$$

Первое приближение найдем, используя это решение. Подставляя его в (П1-56), получаем:

$$i\hbar \frac{da_n^{(1)}}{dt} = \sum_l a_l^{(0)} U_{nl}(t) e^{-i\omega_{nl}t} = U_{nm}(t) e^{-i\omega_{mn}t};$$

отсюда

$$a_n^{(1)}(t) = -\frac{1}{i\hbar} \int_0^t U_{nm}(t') e^{-i\omega_{nm}t'} dt' + \delta_{nm}.$$

Ограничиваясь первым приближением, получаем:

$$P_{mn}(\tau) = |a_n^{(1)}(\tau)|^2 = \frac{1}{\hbar^2} \left| \int_0^\tau U_{nm}(t) e^{-i\omega_{nm}t} dt \right|^2. \quad (\text{П1-58})$$

Если при расчете окажется, что $U_{nm}=0$, то соответствующий переход невозможен. Такой переход называется запрещенным.

Рассмотрим возмущение системы (частицы) электрическим полем:

$$\mathbf{E} = \mathbf{E}_m \cos(\omega t - \mathbf{k}\mathbf{r}).$$

Если эквивалентный электрический дипольный момент частицы равен \mathbf{p}_e , то энергия ее взаимодействия с полем равна:

$$U_B = -\mathbf{p}_e \mathbf{E}.$$

Пусть в момент $t < 0$ частица находилась в состоянии с энергией W_m , а под влиянием возмущения переходит в состояние с энергией W_n :

$$\hat{U}_B = -\hat{\mathbf{p}}_e \hat{\mathbf{E}} = -\hat{\mathbf{p}}_e \mathbf{E}_m \cos(\omega t - \mathbf{k}\mathbf{r}), \quad 0 \leq t \leq \tau,$$

где τ — время взаимодействия.

Согласно (П1-57)

$$U_{nm} = \int_V \psi_{0n}^* \hat{p}_e \hat{E}_m \cos(\omega t - \mathbf{k}\mathbf{r}) \psi_{0m} dV.$$

Так как длина волны электромагнитного поля много больше размеров конфигурационного объема частицы, то поле в пределах этого объема можно считать постоянным и

$$U_{nm} = - \left(\int_V \psi_{0n}^* \hat{p}_e \psi_{0m} dV \right) E_m \cos \omega t$$

или

$$U_{nm} = - p_{enm} E_m \cos \omega t,$$

где $p_{enm} = \int_V \psi_{0n}^* \hat{p}_e \psi_{0m} dV$ — матричный элемент электрического момента диполя.

Согласно (П1-58)

$$P_{mn}(\tau) = \left| \frac{p_{enm} E_m}{\hbar} \right|^2 \left| \frac{e^{j(\omega - \omega_{nm})\tau} - 1}{\omega - \omega_{nm}} - \frac{e^{-j(\omega - \omega_{nm})\tau} - 1}{\omega + \omega_{nm}} \right|^2. \quad (\text{П1-59})$$

Первое слагаемое в этом выражении быстро увеличивается при приближении частоты возмущенного поля ω к частоте перехода ω_{nm} . Практически переход возможен, если

$$\omega \approx \omega_{nm} = \frac{W_n - W_m}{\hbar}$$

или $W_n = W_m + \hbar\omega$. Таким образом, первый член связан с переходом в состояние n , которое выше состояния m на величину энергии $\hbar\omega = \hbar\omega_{nm}$.

При таком переходе энергия частицы увеличивается за счет поглощения энергии электромагнитного поля (резонансное поглощение). Второе слагаемое (П1-59) быстро увеличивается при приближении частоты поля ω к частоте $-\omega_{nm}$. Это слагаемое велико, если

$$\omega \approx -\omega_{nm} = \frac{W_m - W_n}{\hbar},$$

т. е. $W_n = W_m - \hbar\omega$. Следовательно, второй член связан с переходом частицы вниз, при котором энергия частицы уменьшается на величину $\hbar\omega$, отдаваемую частицей полю (индуцированное излучение).

Вероятность, связанная с поглощением, согласно (П1-59) определяется выражением

$$P_{mn}^{\text{погл}}(\tau) = \left| \frac{p_{enm} E_m}{\hbar} \right|^2 \frac{4 \sin^2 \frac{\omega - \omega_{nm}}{2} \tau}{(\omega - \omega_{nm})^2}; \quad (\text{П1-60})$$

вероятность, связанная с излучением,

$$P_{mn}^{\text{изл}}(\tau) = \left| \frac{p_{enm} E_m}{\hbar} \right|^2 \frac{4 \sin^2 \frac{\omega + \omega_{nm}}{2} \tau}{(\omega + \omega_{nm})^2}. \quad (\text{П1-61})$$

Очевидно, вероятность переходов вверх и вниз одинакова:

$$P_{mn}^{\text{погл}}(\tau) = P_{nm}^{\text{изл}}(\tau). \quad (\text{П1-62})$$

Рассмотрим переходы, связанные с поглощением. При достаточно большом значении τ множитель

$$\frac{4 \sin^2 \frac{\omega - \omega_{nm}}{2} \tau}{(\omega - \omega_{nm})^2}$$

обладает свойствами δ -функции и его можно представить в виде

$$\frac{4 \sin^2 \frac{\omega - \omega_{nm}}{2} \tau}{(\omega - \omega_{nm})^2} = 2\pi\tau\delta(\omega - \omega_{nm}).$$

Подставляя этот результат в (П1-61), получаем окончательное выражение

$$P_{mn}^{\text{изл}}(\tau) = \left| \frac{p_{enm} E_m}{\hbar} \right|^2 2\pi\tau\delta(\omega - \omega_{nm}), \quad (\text{П1-63})$$

т. е. вероятность перехода обращается в бесконечность при резонансе $\omega = \omega_{nm}$ и равна нулю при $\omega \neq \omega_{nm}$. В рассмотренном приближении энергетический уровень считается бесконечно узким. При учете ширины энергетического уровня $P_{mn}^{\text{изл}}(\tau)$ будет максимальна при $\omega = \omega_{nm}$, но конечна. Согласно формуле (П1-63) вероятность перехода пропорциональна времени взаимодействия поля с частицей.

Вероятность перехода в единицу времени получим, разделив выражение (П1-63) на τ :

$$p_{mn} = \frac{P_{mn}(\tau)}{\tau} = \left| \frac{p_{enm} E_m}{\hbar} \right|^2 2\pi\delta(\omega - \omega_{nm}).$$

Если уровень m соответствует одному состоянию, а уровень n представляет собой полосу состояний, разделенных по энергии непрерывно или дискретно с очень малым расщеплением, то переходы с уровня m в полосу состояний n будут соответствовать энергиям, заключенным между W и $W+dW$. Если $N(W)$ — плотность распределения состояний в зависимости от энергии, то $N(W)dW = \frac{1}{\hbar} N(W)\hbar d\omega$ — число состояний в полосе n . Вероятность перехода из состояния m в полосу состояний n равна:

$$P_{mn}(\tau) = \int \left| \frac{p_{enm} E_m}{\hbar} \right|^2 2\pi\tau\delta(\omega - \omega_{nm}) N(W) \hbar d\omega.$$

Интегрируя по полосе состояний уровня n , т. е. по узкому спектру частот в области $\omega \approx \omega_{nm}$, получаем:

$$P_{mn}(\tau) = \frac{2\pi}{\hbar} |p_{enm} E_m|^2 N(W) \tau; \quad (\text{П1-64})$$

при этом вероятность перехода в единицу времени определится выражением

$$p_{mn}(\tau) = \frac{P_{mn}(\tau)}{\tau} = \frac{2\pi}{\hbar} |p_{enm} E_m|^2 N(W). \quad (\text{П1-65})$$

20. **Ширина спектральной линии.** Ранее предполагалось, что энергетические уровни бесконечно узки и переходы между ними связаны с излучением или поглощением в бесконечно узком спектре частот. На самом деле уровни имеют конечную ширину и излучение или поглощение происходит в конечном интервале частот, т. е. спектральная линия излучения или поглощения имеет некоторую ширину, а форма ее количественно характеризуется функцией, называемой фактором.

Конечная ширина энергетических уровней связана с конечностью времени жизни частицы на энергетическом уровне Δt . Соответствующая ширина спектральной линии, определяемая как полоса частот $\Delta\omega$, на границах которой интенсивность излучения или поглощения уменьшается в 2 раза по сравнению с максимальной, называется «естественной» шириной линии.

Согласно квантовой механике «естественная» ширина спектральной линии $\Delta\omega$ и время жизни частицы на энергетическом уровне Δt связаны соотношением неопределенностей (П1-15)

$$\Delta\omega\Delta t \approx 1. \quad (\text{П1-66})$$

Конечность времени жизни частицы на энергетическом уровне учитывается затуханием волновой функции. В этом случае m -состояние частицы характеризуется волновой функцией

$$\psi_m = \psi_{m0} e^{-\frac{\gamma_m t}{2}} e^{j\frac{W_m}{\hbar} t}.$$

Вероятность перехода с уровня m на уровень n определится выражением, аналогичным (П1-58), где $e^{-j\omega_{mn}t}$ заменяется на $e^{-j\omega'_{mn}t}$ и

$$\omega'_{nm} = \frac{W_n - W_m}{\hbar} + j\frac{\gamma_n + \gamma_m}{2} = \omega_{nm} + j\frac{\gamma_n + \gamma_m}{2}$$

и аналогично (П1-59)

$$P_{nm}(\tau) = \left| \frac{p_{enm} E_m}{\hbar} \right|^2 \left| \frac{e^{j(\omega - \omega'_{nm})\tau} - 1}{\omega - \omega'_{nm}} - \frac{e^{-j(\omega - \omega'_{nm})\tau} - 1}{\omega + \omega'_{nm}} \right|^2.$$

Рассмотрим первое слагаемое, характеризующее поглощение, полагая $\tau \gg \frac{1}{\gamma_n + \gamma_m}$ (установившийся режим). При этом

$$P_{mn} = \left| \frac{p_{enm} E_m}{\hbar} \right|^2 \frac{1}{(\omega - \omega_{nm})^2 + \left(\frac{\gamma_n + \gamma_m}{2} \right)^2}.$$

Множитель

$$\frac{1}{(\omega - \omega_{nm})^2 + \left(\frac{\gamma_n + \gamma_m}{2} \right)^2}$$

определяет форму линии поглощения (рис. П1-15), связанную с конечностью времени жизни частицы на энергетических уровнях и на-

зываемую лоренцевой. Тот же результат будет и для линии излучения. Обозначая

$$\gamma = \frac{\gamma_n + \gamma_m}{2}$$

и учитывая условие нормировки

$$\int_{-\infty}^{\infty} g(\omega) d\omega = 1,$$

получаем форм-фактор лоренцевой формы линии в виде

$$g(\omega) = \frac{1}{2\pi} \frac{1}{(\omega - \omega_{nm})^2 + \frac{\gamma^2}{4}} \quad (\text{П1-67})$$

или

$$g(\omega) = \frac{1}{2\pi} \frac{\gamma}{(\omega - \omega_{nm})^2 + \left(\frac{\Delta\omega}{2}\right)^2}, \quad (\text{П1-68})$$

где $\Delta\omega = \gamma$ — «естественная» ширина линии. Форма линии симметрична относительно ω_{nm} .

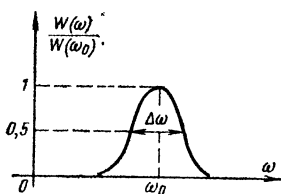


Рис. П1-15. Лоренцева форма линии излучения.

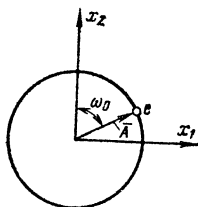


Рис. П1-16. Движение электрона по круговой орбите.

«Естественная» ширина линии определяет тот предел, меньше которого ширина спектральной линии быть не может. В диапазоне СВЧ естественная ширина линии составляет 10^{-3} — 10^{-4} Гц, в оптическом — около 10^9 Гц.

В классическом представлении поглощение и излучение, соответствующее квантовым переходам, связано с колебаниями электронов, квазиупругосвязанных с положительным зарядом атома.

В атомах электроны движутся по эллиптическим орбитам. Это движение происходит с неравномерной скоростью, что эквивалентно неравномерности распределения заряда, и, следовательно, каждую эллиптическую орбиту можно характеризовать средним электрическим дипольным моментом

$$p_e = e l,$$

где e — заряд электрона; l — плечо дипольного момента, совпадающее по направлению с большой осью эллипса.

Для упрощения рассмотрения можно считать, что электрон движется по круговой орбите относительно ядра, совпадающего с центром декартовой системы координат (рис. П1-16) с угловой скоростью ω_0 . Такое движение эквивалентно движению двух электронов во взаимно перпендикулярных направлениях со сдвигом по фазе на $\pi/2$ и описывается уравнениями

$$x_1 = A \cos \omega_0 t;$$

$$x_2 = A \sin \omega_0 t.$$

Соответствующие дипольные моменты равны:

$$p_{e1} = eA \cos \omega_0 t;$$

$$p_{e2} = eA \sin \omega_0 t.$$

Для простоты рассмотрим лишь один диполь. Переменный диполь непрерывно излучает. Поле излучения оказывает обратное влияние на диполь, что представляет реакцию излучения, малую по сравнению с силой упругости, действующей на электрон со стороны ядра. Уравнение движения имеет вид:

$$m \frac{d^2 x}{dt^2} + \beta \frac{dx}{dt} + kx = 0, \quad (\text{П1-69})$$

где kx — сила упругости; $\beta \frac{dx}{dt}$ — реакция излучения.

Уравнение (П1-69) можно представить в виде

$$\frac{d^2 x}{dt^2} + 2\gamma \frac{dx}{dt} + \omega_0^2 x = 0,$$

где

$$\omega_0^2 = \frac{k}{m}, \quad 2\gamma = \frac{\beta}{m}.$$

Решая это уравнение, получаем выражение

$$x(t) = Ae^{-\gamma t} \cos(\omega_0 t + \varphi),$$

определяющее затухающие колебания, энергия которых равна:

$$W(t) \sim x^2(t)$$

или

$$W(t) = W_0 e^{-t/\Delta t},$$

где $\Delta t = \frac{1}{2\gamma}$ — время затухания, в течение которого энергия убывает в e раз. Это время называется «естественным» временем затухания и определяется только излучением колеблющегося электрона.

Так как γ мало, то это почти незатухающие колебания. Используя преобразование Фурье, получаем:

$$x(\omega) = A \int_0^{\infty} e^{-\gamma t} \cos(\omega_0 t + \varphi) e^{-j\omega t} dt = \\ = \frac{A}{2} \left[\frac{e^{j\varphi}}{\gamma/2 + j(\omega - \omega_0)} + \frac{e^{-j\varphi}}{\gamma/2 + j(\omega + \omega_0)} \right].$$

Спектральная плотность энергии

$$W(\omega) d\omega \sim \frac{1}{2\pi} |x^2(\omega)| d\omega \approx \frac{\gamma}{2\pi} \frac{d\omega}{(\omega - \omega_0)^2 + \gamma^2}$$

или

$$W(\omega) d\omega \approx \frac{\gamma}{2\pi} \frac{d\omega}{(\omega - \omega_0)^2 + \left(\frac{\Delta\omega}{2}\right)^2}, \quad (\text{П1-70})$$

где $\Delta\omega = 2\gamma$ — ширина линии.

Наблюдаемые спектральные линии могут являться суперпозицией нескольких неразрешенных спектральных линий. Это *неоднородно уширенная линия*. Если такой суперпозиции нет и частица излучает или поглощает в пределах всей спектральной линии, то линию называют *однородно уширенной*. Рассмотренная выше естественная ширина линии является примером однородно уширенной линии.

Примером неоднородно уширенной линии являются спектральные линии газа, уширенные за счет эффекта Доплера. Частицы газа находятся в хаотическом тепловом движении. Излучение частицы, движущейся со скоростью v , из-за эффекта Доплера смещено по частоте на величину $\omega_0 \frac{v}{c}$, где ω_0 — частота неподвижного излучателя; c — скорость света. Так как частицы движутся с различными скоростями, частотные сдвиги излучения разных частиц будут различны и форм-фактор линии будет определяться распределением скоростей $f(v)$ излучающих частиц:

$$g(\omega) d\omega = f(v) dv.$$

Так как частота излучения частицы

$$\omega = \omega_0 + \frac{v}{c} \omega_0,$$

то

$$v = c \frac{\omega - \omega_0}{\omega_0};$$

$$dv = c \frac{d\omega}{\omega_0}$$

и

$$g(\omega) d\omega = f\left(c \frac{\omega - \omega_0}{\omega_0}\right) c \frac{d\omega}{\omega_0}.$$

В случае Максвелловского распределения скоростей

$$\left. \begin{aligned} f(v) &= \frac{1}{\sqrt{\pi}} \frac{e^{-\left(\frac{v}{v_0}\right)^2}}{v_0} dv; \\ g(\omega) &= \frac{1}{\sqrt{\pi} \Delta\omega_D} e^{-\left(\frac{\omega - \omega_0}{\Delta\omega_D}\right)^2}, \end{aligned} \right\} \quad (\text{П1-71})$$

где

$$\Delta\omega_D = \omega_0 \frac{v_0}{c}, \quad v_0 = \sqrt{\frac{2kT}{m}}.$$

Линия, форм-фактор которой определяется выражением (П1-71), называется *доплеровски уширенной линией*. Форма линии симметрична относительно ω_0 . Ширина ее равна

$$\Delta\omega = 2 \sqrt{\ln 2} \Delta\omega_D.$$

Так как доплеровский сдвиг частоты пропорционален частоте перехода ω_0 , уширение за счет эффекта Доплера значительно в оптическом диапазоне волн. Для молекулы аммиака доплеровское уширение в диапазоне СВЧ составляет примерно 100 кГц, в оптическом диапазоне — на пять порядков больше.

В действительности спектральная линия гораздо шире естественной ширины. В газах на ширину линии влияют столкновения частиц друг с другом и со стенками сосуда, при которых энергия частиц изменяется, что эквивалентно уменьшению времени жизни на данном уровне или уширению линии. Из-за соударений ширина спектральной линии аммиака при комнатной температуре и давлении около 1 Па увеличивается примерно на 300 кГц, что значительно превышает естественную ширину линии. Уширение из-за соударений со стенками при не слишком низких давлениях значительно меньше, чем из-за соударений частиц друг с другом. Для аммиака при комнатной температуре и стандартных размерах сосуда это уширение составляет около 20 кГц. При увеличении давления ширина линии увеличивается. При малых давлениях столкновения между частицами реже и основную роль в уширениях начинают играть столкновения со стенками сосуда.

Уширение линии происходит также из-за влияния магнитных и электрических полей (эффект Зеемана и Штарка) как внешних, так и полей соседних атомов и молекул.

В твердых телах уширение линии связано с неоднородностью кристаллической решетки, что приводит к неоднородности ее поля и сдвигу уровней атомов на неодинаковую величину. Тепловые колебания решетки приводят к изменению внутрикристаллического электрического поля и колебаниям его относительно некоторого постоянного значения. Спектральная линия при этом смещается, что эквивалентно ее уширению. При охлаждении кристалла тепловые колебания решетки и вызываемые ими уширения линии уменьшаются.

Спектральные линии уширяются за счет насыщения при увеличении мощности внешнего электромагнитного поля, поглощаемого веществом. При этом скорость перехода с нижнего уровня на верхний за счет поглощения поля больше, чем скорость восстановления термодинамического равновесия. Это приводит к уменьшению поглощения, причем в центре линии уменьшение более значительно, чем на краях, что приводит к уширению линии. В диапазоне СВЧ уширение за счет насыщения составляет около 100 кГц.

ПРИЛОЖЕНИЕ 2

МАТЕМАТИЧЕСКИЕ ДОПОЛНЕНИЯ (НЕЛИНЕЙНЫЕ УРАВНЕНИЯ)

1. Физические процессы в нелинейных устройствах или системах описываются нелинейными дифференциальными уравнениями.

Способы решения нелинейных уравнений более трудоемки, чем линейных. Только для небольшого числа уравнений могут быть найдены точные решения. Чаще возможно найти лишь приближенные решения, удовлетворяющие определенным условиям.

Если влияние нелинейности невелико, то для получения приближенных решений могут быть применены аналитические методы.

При значительной нелинейности решения находят численными или графическими методами. При этом используют численные значения параметров и начальных условий. Полученное решение справедливо лишь для одной определенной системы условий.

При решении практических задач целесообразно бывает отбросить второстепенные эффекты и сосредоточить внимание на основных. Это иногда позволяет перейти к упрощенным уравнениям, решения которых правильно описывают особенности изучаемого явления.

2. Аналитические методы. Решение, полученное в аналитической форме, удобно для исследования в широких пределах изменения параметров. Аналитические методы применяют к уравнениям с малой нелинейностью. Наиболее результативными являются метод последовательных приближений (метод итераций), метод возмущений, метод медленно меняющихся амплитуд и фаз и метод вариации параметров.

Однако ни один из этих методов не может дать исчерпывающих сведений о поведении среды. Например, метод возмущений применяется для исследования установившегося режима, но его нецелесообразно использовать для исследования переходного режима; метод вариации параметров позволяет исследовать переходный режим и дает лишь грубое представление об установившемся режиме. Метод медленно меняющихся амплитуд и фаз применяется для исследования процессов, характеризующихся периодическими или близкими к периодическим колебаниями.

3. Метод последовательных приближений (итераций) заключается в нахождении по известному приближенному решению уравнения следующего более точного приближения. Этот метод может быть применен, если последовательность полученных приближенных решений сходится.

Электромагнитный процесс в слабонелинейных средах может описываться уравнением вида

$$\frac{d^2x}{dt^2} + \omega_0^2 x + f(x) = F(t), \quad (\text{П2-1})$$

где $f(x)$ — малая по величине нелинейная функция переменной x . В линейном приближении это уравнение имеет вид:

$$\frac{d^2x}{dt^2} + \omega_0^2 x = F(t). \quad (\text{П2-2})$$

Допустим, что линейное уравнение (П2-2) имеет решение x_0 . Это решение называется порождающим и может рассматриваться как приближенное решение нелинейного уравнения (П2-1).

Подставляя решение x_0 в выражение для функции $f(x)$ в уравнении (П2-1), получаем:

$$\frac{d^2x}{dt^2} + \omega_0^2 x = F(t) - f(x_0).$$

Интегрируя последнее уравнение, получаем решение x_1 , которое является решением нелинейного уравнения (П2-1) в первом приближении:

$$x_1 = x_0 + \text{поправка}.$$

Следующее приближение получаем, подставляя в уравнение (П2-1) первое приближение решения x_1 :

$$\frac{d^2x}{dt^2} + \omega_0^2 x = F(t) - f(x_1).$$

Полученное решение

$$x_2 = x_1 + \text{поправка}$$

является решением уравнения (П2-1) во втором приближении.

Аналогично можно найти последующие приближения $x_3, x_4 \dots$

4. Метод возмущений (метод малого параметра) применим к уравнениям, близким к линейным, когда нелинейные члены малы и их можно связать с малым параметром.

Уравнение вида

$$\frac{d^2x}{dt^2} + \omega_0^2 x + f(x) = F(t), \quad (\text{П2-3})$$

где $f(x)$ — нелинейная функция, для слабонелинейных сред можно представить в виде

$$\frac{d^2x}{dt^2} + \omega_0^2 x + \mu f(x) = F(t), \quad (\text{П2-4})$$

где μ — малый параметр, означающий малость нелинейной функции $f(x)$.

Решение уравнения (П2-4) ищется в виде ряда, члены которого содержат возрастающие степени малого параметра

$$x = x_0(t) + \mu x_1(t) + \mu^2 x_2(t) + \dots, \quad (\text{П2-5})$$

где x_0, x_1, x_2 — пока неизвестные функции.

Достаточно малые значения μ обеспечивают быструю сходимость и два или три члена ряда (П2-5) дают достаточную точность. При $\mu \rightarrow 0$ решение

$$x = x_0(t)$$

является решением линейного уравнения

$$\frac{d^2x}{dt^2} + \omega_0^2 x = F(t).$$

Поскольку известно, что решение уравнения (П2-3) носит колебательный характер, то естественно предположить, что основная частота ω в данной нелинейной среде будет зависеть от амплитуды колебаний, а так как в уравнение (П2-3) входит ω_0^2 , то частоту колебаний удобно представить в виде

$$\omega^2 = \omega_0^2 + \mu b_1(A) + \mu^2 b_2(A) + \dots, \quad (\text{П2-6})$$

где b_1, b_2 — функции, зависящие от амплитуды A .

Подставляя выражения (П2-5) и (П2-6) в уравнение (П2-4), получаем:

$$\begin{aligned} & \frac{d^2x_0}{dt^2} + \mu \frac{d^2x_1}{dt^2} + \mu^2 \frac{d^2x_2}{dt^2} + \dots + \omega^2 (x_0 + \mu x_1 + \mu^2 x_2 + \dots) - \\ & - \mu b_1(x_0 + \mu x_1 + \mu^2 x_2 + \dots) - \mu^2 b_2(x_0 + \mu x_1 + \mu^2 x_2 + \dots) - \\ & - \dots + \mu f(x_0 + \mu x_1 + \mu^2 x_2 + \dots) = F(t). \end{aligned}$$

Приравнявая члены, содержащие одинаковую степень параметра μ , получаем систему взаимосвязанных уравнений:

$$\frac{d^2x_0}{dt^2} + \omega^2 x_0 = F(t);$$

$$\frac{d^2x_1}{dt^2} + \omega^2 x_1 = b_1 x_0 - f(x_0);$$

$$\frac{d^2x_2}{dt^2} + \omega^2 x_2 = b_1 x_1 + b_2 x_0 - f(x_0, x_1);$$

.....

Решая последовательно эту систему уравнений, получаем решение вида (П2-5), представляющее собой ряд поправок к нулевому приближению (порождающему решению).

Начальные условия используются для вычисления постоянных интегрирования, появляющихся на каждом этапе решения.

Если внешнее воздействие $F(t)$ имеет гармонический характер, т. е. $F(t) = F_m \cos \omega_1 t$, то частота колебаний ω или равна частоте ω_1 , или связана с частотой ω_1 определенным соотношением. В этом случае нет необходимости пользоваться разложением (П2-6).

5. Метод медленно меняющихся амплитуд и фаз применяется в тех случаях, когда заранее известно, что форма колебаний близка к синусоидальной. Рассмотрим нелинейное уравнение вида

$$\frac{d^2x}{dt^2} + \omega_0^2 x + \mu f\left(x, \frac{dx}{dt}\right) = 0, \quad (\text{П2-7})$$

где $f\left(x, \frac{dx}{dt}\right)$ — функция, определяющая нелинейность и потери; μ — безразмерный параметр, указывающий на малость нелинейности и потерь.

Решение уравнения при $\mu=0$ представляет собой гармоническое колебание. Очевидно, при малом μ решение близко к гармоническому и его можно представить в виде

$$x = A_m(t) \cos[\omega_0 t + \varphi(t)] = A_m(t) \cos \Phi(t), \quad (\text{П2-8})$$

где $A_m(t)$ и $\varphi(t)$ — соответственно амплитуда и фаза, медленно изменяющиеся во времени $\Phi(t) = \omega_0 t + \varphi(t)$.

Условие медленного изменения амплитуды и фазы заключается в том, что они мало изменяются за период колебания $T = 2\pi/\omega_0$, т. е.

$$\left| \frac{1}{A_m} \frac{dA_m}{dt} \right| \ll \omega_0, \quad \left| \frac{d\varphi}{dt} \right| \ll \omega_0. \quad (\text{П2-9})$$

Производные величины x определяются выражениями

$$\frac{dx}{dt} = \frac{dA_m}{dt} \cos \Phi - \left(\omega_0 + \frac{d\varphi}{dt} \right) A_m \sin \Phi; \quad (\text{П2-10})$$

$$\begin{aligned} \frac{d^2x}{dt^2} = & \left[\frac{d^2A_m}{dt^2} - \left(\omega_0 + \frac{d\varphi}{dt} \right)^2 A_m \right] \cos \Phi - \\ & - \left[2 \left(\omega_0 + \frac{d\varphi}{dt} \right) \frac{dA_m}{dt} + \frac{d^2\varphi}{dt^2} A_m \right] \sin \Phi. \end{aligned} \quad (\text{П2-11})$$

Нелинейная функция $f\left(x, \frac{dx}{dt}\right)$ мала, что подчеркивается множителем μ в уравнении (П2-7). Поэтому среди слагаемых производной $\frac{dx}{dt}$, определяемой выражением (П2-10), имеет значение лишь наибольшее $\omega_0 A_m \sin \Phi$ и

$$f\left(x, \frac{dx}{dt}\right) = f(A_m \cos \Phi, -\omega_0 A_m \sin \Phi)$$

или

$$f\left(x, \frac{dx}{dt}\right) = P_m(t) \cos(\omega_0 t + \psi) = P_m(t) \cos \Psi, \quad (\text{П2-12})$$

где $\Psi = \Phi + (\psi - \varphi)$; $(\psi - \varphi)$ — фазовый сдвиг функции $f\left(x, \frac{dx}{dt}\right)$ относительно функции $x(t)$, определяемой формулой (П2-8).

Подставляя выражения (П2-8), (П2-11) и (П2-12) в уравнение (П2-7), пренебрегая вследствие малости согласно условию (П2-9) вторыми производными амплитуд и фаз, а также произведениями этих производных, учитывая соотношение

$$\cos \Psi = \cos(\psi - \varphi) \cos \Phi - \sin(\psi - \varphi) \sin \Phi,$$

получаем:

$$2\omega_0 \frac{d\varphi}{dt} A_m \cos \Phi + 2\omega_0 \frac{dA_m}{dt} \sin \Phi = \\ = P_m [\cos(\psi - \varphi) \cos \Phi - \sin(\psi - \varphi) \sin \Phi]. \quad (\text{П2-13})$$

Приравнивая коэффициенты при $\cos \Phi$ и $\sin \Phi$ в правой и левой частях уравнения (П2-13), получаем:

$$\frac{d\varphi}{dt} = \frac{P_m \cos(\psi - \varphi)}{2\omega_0 A_m}; \\ \frac{dA_m}{dt} = - \frac{P_m \sin(\psi - \varphi)}{2\omega_0}.$$

Интегрируя эти выражения по времени, определим амплитуду $A_m(t)$ и фазу $\varphi(t)$, подставляя которые в выражение (П2-8), найдем решение уравнения (П2-7).

6. **Метод вариации параметров** целесообразно применять в том случае, когда из физических соображений известен вид искомого решения.

Рассмотрим нелинейное уравнение второго порядка, описывающее поведение нелинейной системы при отсутствии внешней возбуждающей силы

$$\frac{d^2x}{dt^2} + \omega_0^2 x + \mu f\left(x, \frac{dx}{dt}\right) = 0, \quad (\text{П2-14})$$

где $f\left(x, \frac{dx}{dt}\right)$ — нелинейная функция, определяющая потери и нелинейность системы. Амплитуда и фаза колебаний в такой системе переменны.

Уравнение (П2-14) представим в виде системы уравнений первого порядка, вводя переменную $y = \frac{dx}{dt}$. При этом получим:

$$\left. \begin{aligned} \frac{dx}{dt} &= y; \\ \frac{dy}{dt} &= -\omega_0^2 x - \mu f(x, y). \end{aligned} \right\} \quad (\text{П2-15})$$

Порождающее решение находится из решения системы (П2-15) при $\mu=0$ и имеет вид:

$$\left. \begin{aligned} x &= A_m \cos(\omega_0 t + \varphi) = A_m \cos \Phi; \\ y &= -\omega_0 A_m \sin(\omega_0 t + \varphi) = -\omega_0 A_m \sin \Phi; \end{aligned} \right\} \quad (\text{П2-16})$$

где A_m и φ — постоянные, определяемые начальными условиями, $\Phi = \omega_0 t + \varphi$.

Решение нелинейной системы (П2-15) имеет вид (П2-16), где A_m и φ являются функциями времени:

$$\left. \begin{aligned} x &= A_m(t) \cos[\omega_0 t + \varphi(t)] = A_m(t) \cos \Phi; \\ y &= -\omega_0 A_m(t) \sin[\omega_0 t + \varphi(t)] = -\omega_0 A_m(t) \sin \Phi. \end{aligned} \right\} \quad (\text{П2-17})$$

Стационарные состояния определяются из условия

$$\frac{dx_1}{dt} = \frac{dx_2}{dt} = \dots = \frac{dx_n}{dt} = 0. \quad (\Pi 2-19)$$

Это условие определяет особые точки, в которых система находится в положении равновесия. Линейная система имеет одно единственное положение равновесия, нелинейная может иметь несколько положений.

Решение нелинейной системы уравнений (П2-18) при условии (П2-19) дает группы значений $x_{1s}, x_{2s}, \dots, x_{ns}$, соответствующие положениям равновесия. Чтобы выяснить, является ли какое-либо конкретное состояние равновесия устойчивым, систему, находящуюся в этом состоянии, возмущают, отклоняя из состояния равновесия на величину δx_i . При этом необходимо следить, чтобы величина возмущения δx_i была достаточно малой, с тем чтобы система осталась в области, близкой к рассматриваемому положению. Если начальные возмущения велики, то система может оказаться в области, соответствующей какой-либо другой точке равновесия.

Если при $t \rightarrow \infty$ переменные x_i возвращаются к исходным равновесным значениям $x_{i,0}$, то систему называют *асимптотически устойчивой*, если при $t \rightarrow \infty$ значения x_i удаляются от исходных равновесных значений, то систему называют *неустойчивой*. Если при возмущении переменные x_i не удаляются и не возвращаются к исходным значениям, то систему называют *безразлично устойчивой*. *Неустойчивой* называют систему, в которой возмущения нарастают даже по истечении довольно долгого времени.

Итак, исследуем стационарные решения x_{is} на устойчивость, задавая малые отклонения δx_i . Подставляя $x_{is} + \delta x_i$ в систему уравнений (П2-18) при условии (П2-19) получаем:

[illegible]

Разлагая нелинейные функции в ряд Тейлора в окрестности точек равновесия и удерживая первый линейный член, получаем линейную систему

$$\left. \begin{aligned} \frac{d\delta x_1}{dt} &= C_{11}\delta x_1 + C_{12}\delta x_2 + \dots + C_{1n}\delta x_n; \\ \frac{d\delta x_2}{dt} &= C_{21}\delta x_1 + C_{22}\delta x_2 + \dots + C_{2n}\delta x_n; \\ &\vdots \\ \frac{d\delta x_n}{dt} &= C_{n1}\delta x_1 + C_{n2}\delta x_2 + \dots + C_{nn}\delta x_n, \end{aligned} \right\} \quad (\Pi 2-20)$$

где C_{ni} — постоянные, значения которых зависят от значений x_{is} .

Естественно предположить для δx_1 экспоненциальную зависимость от времени вида $e^{\alpha t}$, т. е.

$$\delta x_i = A_i e^{\alpha t}, \quad (\Pi 2-21)$$

где A_i — постоянные, определяемые начальными условиями.

теристического уравнения (П2-23) была отрицательной, необходимо и достаточно, чтобы все определители были положительны, т. е.

$$\Delta_1 > 0, \Delta_2 > 0, \Delta_3 > 0, \dots, \Delta_n > 0.$$

Поскольку последняя строка определителя Δ_n , за исключением последнего члена C_n , заполнена нулями, то $\Delta_n = C_n \Delta_{n-1}$ и для устойчивости системы необходимо, чтобы

$$C_n > 0 \text{ и } \Delta_{n-1} > 0.$$

Условие, что все коэффициенты C_r характеристического уравнения должны быть положительны, не является достаточным, так как характеристическое уравнение (П2-23), все коэффициенты ко-

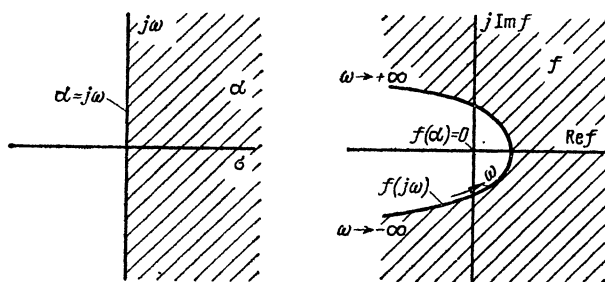


Рис. П2-1. α - и f -плоскости. $f(\alpha) = C_0 \alpha^2 + C_1 \alpha + C_2$.

торого положительны, при определенных условиях может иметь корни с положительной вещественной частью.

9. **Критерий Найквиста** основан на применении конформного отображения. Корни характеристического уравнения (П2-23) комплексны и могут быть представлены на комплексной плоскости $\alpha = \sigma + j\omega$ (рис. П2-1). Значения α , связанные с неустойчивостью состояния, расположены в правой половине плоскости (заштрихованная область), поскольку их вещественные части положительны. Граница этой области — мнимая ось.

Левую часть характеристического уравнения (П2-23) можно рассматривать как функцию $f(\alpha)$. Значения этой функции, соответствующие значениям α , можно представить на комплексной плоскости f . Задаваясь значениями $\alpha = j\omega$, пробегающими мнимую ось от $-j\infty$ до $j\infty$, строят кривую $f(j\omega)$ в плоскости f . Заштрихованные области плоскостей α и f соответствуют друг другу. Так как степень полинома $f(\alpha)$ равна n , то каждой точке плоскости f соответствует n различных точек плоскости α . С другой стороны, заданной точке плоскости α соответствует только одна точка плоскости f . Каждой точке заштрихованной области плоскости α должна соответствовать точка заштрихованной же области плоскости f . Эти точки соответствуют тем значениям α , которые приводят к неустойчивым решениям.

Характеристическому уравнению $f(\alpha) = 0$ соответствует начало координат плоскости f , корни уравнения выражаются точками

плоскости α , соответствующими началу координат плоскости f . Если начало координат плоскости оказывается в заштрихованной области, то вещественная часть по меньшей мере одного из корней положительна, что указывает на неустойчивость системы. Если начало координат оказывается в незаштрихованной области, то вещественная часть у всех корней отрицательна и система может иметь лишь устойчивые решения.

10. **Графические методы** решения в основном дают качественную информацию о решениях нелинейных уравнений. Эти методы могут быть применены для исследования нелинейных уравнений, когда нелинейность становится значительной и приведенные выше аналитические методы исследования дают совершенно неправильные результаты.

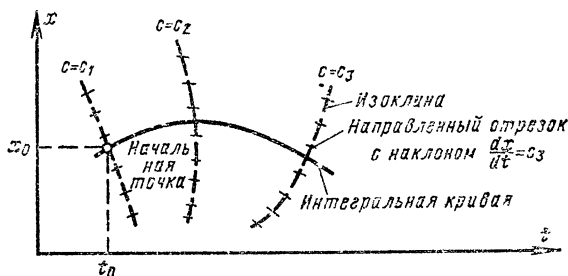


Рис. П2-2. Построение интегральной кривой методом изоклин.

11. **Метод изоклин** применяется для решения уравнений первого порядка вида

$$\frac{dx}{dt} = f(x, t), \quad (\text{П2-24})$$

где $f(x, t)$ — непрерывна и однозначна за исключением особых точек, где $f(x, t)$ становится неопределенной.

Изоклина — геометрическое место точек, в которых производная $\frac{dx}{dt}$ имеет одно и то же значение. Графические построения выполняются в плоскости с ортогональной системой координат x и t . Задавая значения

$$\frac{dx}{dt} = C_1, C_2, \dots,$$

вычерчивают изоклины $f(x, t) = C$ и заполняют всю плоскость x, t (рис. П2-2).

Каждому решению уравнения (П2-24) соответствует кривая на плоскости x, t (**интегральная кривая**), касательная к которой в каждой точке перпендикулярна проходящей через эту точку изоклине. Конкретная интегральная кривая, соответствующая

определенным постоянным интегрирования, задается начальными условиями

$$x = x_0, \quad t = t_0.$$

От этой точки проводится кривая, перпендикулярная изоклинам. Если изоклины на плоскости x, t расположены достаточно близко друг к другу, то решение можно получить с достаточной точностью. Различные начальные условия в общем случае приводят к различным интегральным кривым, соответствующим разным постоянным интегрирования. Так как в регулярных точках крутизна однозначно определяется значением $\frac{dx}{dt}$, то интегральные кривые не могут пересекаться друг с другом. Из особой точки может выходить или входить в нее бесконечное число интегральных кривых.

12. Метод фазовой плоскости представляет собой метод изоклин в применении к автономным уравнениям второго порядка. Автономным называется уравнение, в котором отсутствует явная зависимость от независимой переменной. Уравнение вида

$$\frac{d^2x}{dt^2} + f\left(\frac{dx}{dt}, x\right) = 0 \quad (\text{П-25})$$

является автономным, так как в нем отсутствует явная зависимость от времени t . Это уравнение можно свести к уравнению первого порядка путем замены

$$\frac{dx}{dt} = v.$$

Тогда

$$\frac{d^2x}{dt^2} = \frac{dv}{dt} = \frac{dv}{dx} \frac{dx}{dt} = v \frac{dv}{dx}$$

и уравнение (П2-25) будет иметь вид

$$\frac{dv}{dx} = -\frac{f(v, x)}{v},$$

аналогичный (П2-24), и его можно решить методом изоклин.

Графические построения выполняются на плоскости с ортогональными координатами $v = \frac{dx}{dt}$ и x (рис. П2-3). Такая плоскость называется фазовой. Величину x можно рассматривать как перемещение точки, $\frac{dx}{dt} = v$ — как скорость перемещения. Графическое изображение скорости v как функции перемещения x называется *фазовым портретом*. Кривые в плоскости x, v называются *фазовыми траекториями*. Мгновенные значения x и v являются точками фазовой траектории и называются *изображающими точками*. Ско-

рость движения изображающей точки вдоль траектории называется фазовой скоростью и определяется выражением

$$\frac{ds}{dt} = \left[\left(\frac{dx}{dt} \right)^2 + \left(\frac{dv}{dt} \right)^2 \right]^{1/2},$$

где s — длина дуги, измеренная вдоль траектории.

Фазовая скорость изменяется по мере движения изображающей точки вдоль траектории. Она нигде не равна нулю, за исключением особых точек, где $\frac{dx}{dt}$ и $\frac{d^2x}{dt^2}$ равны нулю.

Если фазовая траектория представляет собой замкнутую интегральную кривую, к которой независимо от начальных условий приближаются другие интегральные кривые, эта интегральная кривая называется предельным циклом. Она изображает устойчивые периодические колебания, определяемые лишь свойствами самого уравнения и не зависящие от начальных условий. Это характерно для нелинейных систем и не имеет места в системах линейных. Если начальная точка находится вблизи предельного цикла, то фазовая траектория представляет собой развертывающуюся спираль, приближающуюся к предельному циклу. Если начальная точка находится вне предельного цикла, то интегральная кривая свертывается внутрь, приближаясь к предельному циклу.

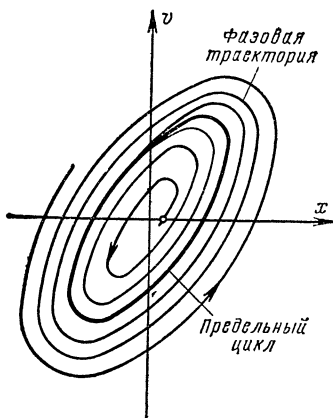


Рис. П2-3. Фазовая плоскость.

Уравнение второго порядка (П2-25) можно свести к системе уравнений первого порядка вида

$$\left. \begin{aligned} \frac{dx_1}{dt} &= P(x_1, x_2); \\ \frac{dx_2}{dt} &= Q(x_1, x_2). \end{aligned} \right\} \quad (\text{П2-26})$$

Исключая время t , получаем:

$$\frac{dx_2}{dx_1} = \frac{Q(x_1, x_2)}{P(x_1, x_2)}.$$

Полагая $\frac{dx_2}{dx_1} = C_1$, определяем на фазовой плоскости x_1, x_2 изоклину

$$F(x_1, x_2) = \frac{Q(x_1, x_2)}{P(x_1, x_2)} = C_1,$$

вдоль которой наклон интегральных кривых остается постоянным:

$$\frac{dx_2}{dx_1} = C_1.$$

Придавая постоянной другие значения C_2, C_3, \dots , изоклинами покрывают всю фазовую плоскость. На изоклины наносится направление производной $\frac{dx_2}{dx_1} = C_1, C_2, C_3 \dots$. Имея начальную точку x_{01}, x_{02} и следуя направлениям линий поля, чертят интегральную кривую.

13. Интегральные уравнения наиболее полно описывают физическое содержание задачи, так как они включают в себя и граничные условия. *Интегральным уравнением* называется уравнение, содержащее искомую функцию под знаком интеграла.

Интегральным уравнением Фредгольма второго рода называется уравнение вида

$$u(x_1) - \mu \int_a^b G(x_1, x_2) u(x_2) dx_2 = f(x_1).$$

Это уравнение написано для случая, когда искомая функция должна быть определена на оси x_1 . Здесь μ — параметр, $u(x_1)$ — искомая функция, $f(x_1)$ и $G(x_1, x_2)$ — заданные непрерывные функции. Функция $G(x_1, x_2)$ называется ядром уравнения.

Если функция $f(x_1) = 0$, то получаем однородное уравнение Фредгольма второго рода:

$$u(x_1) = \mu \int_a^b G(x_1, x_2) u(x_2) dx_2. \quad (\text{П2-27})$$

Если искомая функция входит только под знак интеграла, то уравнение имеет вид

$$\mu \int_a^b G(x_1, x_2) u(x_2) dx_2 = f(x_1)$$

и называется *уравнением Фредгольма первого рода*.

Функции $G(x_1, x_2)$, $u(x_1)$, $f(x_1)$ и параметр μ могут принимать как действительные, так и комплексные значения.

Функция $u(x_1)$, не равная тождественно нулю и удовлетворяющая однородному интегральному уравнению Фредгольма второго рода (П2-27) при определенном значении параметра μ , называется *собственной функцией уравнения или ядра*. Соответствующие значения параметра μ называются *собственными значениями ядра* или соответствующего ему интегрального уравнения.

14. Метод последовательных приближений может быть применен к решению уравнения (П2-27). Будем искать решение этого уравнения в виде ряда

$$u(x_1) = u_0(x_1) + \mu u_1(x_1) + \mu^2 u_2(x_1) + \dots \quad (\text{П2-28})$$

Если этот ряд равномерно сходится в промежутке (a, b) , то, подставляя его в уравнение (П2-27) и приравнявая коэффициенты при одинаковых степенях μ в обеих частях полученного ра-

венства, получим формулы для последовательного определения $u_n(x_1)$:

$$\begin{aligned} u_0(x_1) &= f(x_1); \\ u_1(x_1) &= \int_a^b G(x_1, x_2) u_0(x_2) dx_2; \\ u_2(x_1) &= \int_a^b G(x_1, x_2) u_1(x_2) dx_2; \\ &\dots \dots \dots \\ u_n(x_1) &= \int_a^b G(x_1, x_2) u_{n-1}(x_2) dx_2. \end{aligned}$$

Если μ по модулю достаточно мало, то ряд (П2-28) сходится. Функцию $u_n(x_1)$ можно выразить непосредственно через свободный член $f(x_1)$. Для удобства записи заменим x_2 на t . Тогда

$$\begin{aligned} u_1(x_1) &= \int_a^b G(x_1, t) f(t) dt; \\ u_2(x_1) &= \int_a^b G(x_1, t) u_1(t) dt = \int_a^b \int_a^b G(x_1, t) G(t, t_1) f(t) dt_1 dt = \\ &= \int_a^b G_2(x_1, t_1) f(t_1) dt_1 \end{aligned}$$

и вообще

$$u_n(x_1) = \int_a^b G_n(x_1, t) f(t) dt,$$

где

$$\begin{aligned} G_n(x_1, t) &= \int_a^b \int_a^b \dots \int_a^b G(x_1, t_{n-1}) G(t_{n-1}, t_{n-2}) \dots \\ &\dots G(t_2, t_1) G(t_1, t) dt_1 dt_2 \dots dt_{n-1}. \end{aligned}$$

Подставляя это в (П2-28), получаем:

$$u(x_1) = f(x_1) + \mu \sum_{n=0}^{\infty} \int_a^b G_{n+1}(x_1, t) \mu^n f(t) dt$$

или

$$u(x_1) = f(x_1) + \mu \int_a^b R(x_1, t, \mu) f(t) dt,$$

где функция

$$R(x_1, t, \mu) = \sum_{n=0}^{\infty} G_{n+1}(x_1, t) \mu^n$$

называется *резольвентой ядра* $G(x_1, x_2)$ или уравнения (П2-27).

СПИСОК ЛИТЕРАТУРЫ

- 1 Кугушев А. М., Голубева Н. С. Основы радиоэлектроники. М., «Энергия», 1969. 880 с. с ил.
2. Бломберген Н. Нелинейная оптика. М., «Мир», 1966. 860 с. с ил.
- 3 Шуберт М., Вильгельми Б. Введение в нелинейную оптику. М., «Мир», 1973. 244 с. с ил.
4. Ярив А. Квантовая электроника и нелинейная оптика. М., «Советское радио», 1973. 454 с. с ил.
5. Никольский В. В. Электродинамика и распространение радиоволн. М., «Наука», 1973. 606 с. с ил.
6. Ахманов С. А., Хохлов Р. В. Проблемы нелинейной оптики. М., ВИНТИ, 1965. 296 с. с ил.
7. Гуревич А. В., Шварцбург А. Б. Нелинейная теория распространения радиоволн в ионосфере. М., «Наука», 1973. 272 с. с ил.
8. Квантовая электроника. М., «Советская энциклопедия», 1969. 432 с. с ил.
- 9 Гинзбург В. Л., Рухадзе А. А. Волны в магнитоактивной плазме. М., «Наука», 1975. 255 с. с ил.
10. Цытович В. Н. Нелинейные эффекты в плазме. М., «Наука», 1967. 288 с. с ил.
11. Ферриты в нелинейных сверхвысокочастотных устройствах. М., Изд-во иностр. лит., 1961. 634 с. с ил.
12. Колесников П. М. Введение в нелинейную электродинамику. Минск, «Наука и техника», 1971. 380 с. с ил.
13. Борн М., Вольф Э. Основы оптики. М., «Наука», 1960. 856 с. с ил.
14. Гольдштейн Л. Д., Зернов Н. В. Электромагнитные поля и волны. М., «Советское радио», 1971. 664 с. с ил.
15. Квантовая оптика и квантовая радиофизика. М., «Мир», 1966. 452 с. с ил.
16. Ферромагнитный резонанс. Под ред. С. В. Вонсовского М., «Физматгиз», 1964. 344 с. с ил.
17. Герцберг Г. Спектры и строение двухатомных молекул. М., Изд-во иностр. лит. 1949. 404 с. с ил.
18. Страховский Г. М., Успенский А. В. Основы квантовой электроники. М., «Высшая школа», 1973. 312 с. с ил.
19. Ораевский А. Н. Молекулярные генераторы. М., «Наука», 1964. 296 с. с ил.
20. Пантел Р., Путхоф Г. Основы квантовой электроники. М., «Мир», 1972. 384 с. с ил.
21. Васильев В. Н. Электронные и квантовые приборы СВЧ. М., «Связь», 1972. 256 с. с ил.
22. Батраков А. С. Квантовые приборы. Л., «Энергия», 1972. 176 с. с ил.
23. Франсон М., Сланский С. Когерентность в оптике. М., «Наука», 1967. 80 с. с ил.
24. Дьяков В. А., Тарасов Л. В. Оптическое когерентное излучение. М., «Советское радио», 1974. 168 с. с ил.
25. Дьяков В. А. Введение в квантовую электронику. М., «Энергия», 1969. 264 с. с ил.
26. Дулин В. Н. Электронные и квантовые приборы СВЧ. М., «Энергия», 1972. 224 с. с ил.

27. Ван дер Зил А. Шум. Источники, описание, измерение. М., «Советское радио», 1973. 230 с. с ил.
28. Льюиселл У. Излучение и шумы в квантовой электронике. М., «Наука», 1972. 400 с. с ил.
29. Ищенко Е. Ф., Климов Ю. М. Оптические квантовые генераторы. М., «Советское радио», 1968. 470 с. с ил.
30. Рябцев Н. Г. Материалы квантовой электроники. М., «Советское радио», 1972. 382 с. с ил.
31. Вайнштейн Л. А. Открытые резонаторы и открытые волноводы. М., «Советское радио», 1966. 476 с. с ил.
32. Микаэлян А. Л., Тер-Микаэлян М. Л., Турков Ю. Г. Оптические генераторы на твердом теле. М., «Советское радио», 1967. 384 с. с ил.
33. Бирнбаум Д. Оптические квантовые генераторы. М., «Советское радио», 1967. 360 с. с ил.
34. Лазеры. Сборник статей. Пер. с англ. под ред. М. Е. Жаботинского и Т. А. Шмаонова. М., Изд-во иностр. лит., 1963. 470 с с ил.
35. Вейлстеке А. Основы теории квантовых усилителей и генераторов. М., Изд-во иностр. лит., 1963. 410 с. с ил.
36. Оптические квантовые генераторы. Сборник статей. Пер. с англ. под ред. Ф. В. Бункина. М., «Мир», 1966. 376 с. с ил.
37. Бертен Ф. Основы квантовой электроники. М., «Мир», 1971. 630 с. с ил.
38. Белостоцкий Б. Р., Любавский Ю. В., Овчинников В. М. Основы лазерной техники. М., «Советское радио», 1972. 408 с. с ил.
39. Штейншлейгер В. Б., Мисежников Г. С., Лифанов П. С. Квантовые усилители СВЧ. М., «Советское радио», 1971. 432 с с ил.
40. Зингер Дж. Мазеры. М., Изд-во иностр. лит., 1961, 206 с. с ил.
41. Троуп Г. Квантовые усилители и генераторы. М., Изд-во иностр. лит., 1961. 170 с. с ил.
42. Файн В. Н., Ханин Э. И. Квантовая радиофизика. М., «Советское радио», 1965. 608 с. с ил.
43. Уэрт Ч., Томсон Р. Физика твердого тела. М., «Мир», 1969.
44. Сигмен А. Мазеры. М., «Мир», 1966. 520 с. с ил.
45. Займан Дж. Принципы теории твердого тела. М., «Мир», 1966. 416 с. с ил.
46. Базаров В. К. Полупроводниковые лазеры и их применение. М., «Энергия», 1969. 56 с. с ил.
47. Гвоздовер С. Д. Теория электронных приборов сверхвысоких частот. М., Гостехиздат, 1956. 528 с. с ил.
48. Лебедев И. В. Техника и приборы сверхвысоких частот. Т. 2. М., «Энергия», 1964. 616 с. с ил.
49. Шевчик В. Н. Основы электроники сверхвысоких частот. М., «Советское радио», 1959. 308 с. с ил.
50. Андрушко Л. М., Бурмистенко В. М. Электронные и квантовые приборы СВЧ. М., «Связь», 1974. 192 с. с ил.
51. Каннингхэм В. Введение в теорию нелинейных систем. М.—Л., Госэнергоиздат, 1962. 456 с. с ил.
52. Блохинцев Д. И. Основы квантовой механики. М., «Высшая школа», 1963. 620 с. с ил.
53. Боголюбов Н. Н., Митропольский Ю. А. Асимптотические методы в теории нелинейных колебаний. М., Физматгиз, 1963. 410 с. с ил.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Амплитрон 321

Баланс амплитуд и фаз 186
Больцмана распределение 39, 343

Возмущения метод 381
Волновая функция 356
Волнового синхронизма условие 58, 59, 199
Волновое уравнение 16, 18, 19, 25
— — для диэлектрика 46
— — — магнетика 68
— — — плазмы 63

Генератор квантовый 107
— — молекулярный 284, 285
Граничные условия 26
Гюйгенса — Френеля принцип 161

Зеемана эффект 352
Зона валентная 43
— запрещенная 43
— проводимости 43

Изоклин метод 389
Инверсная населенность 40
— — двухуровневых сред 108
— — трехуровневых сред 109—111, 113
— — четырехуровневых сред 112—114
Индукированное излучение 37, 44

Квантового усиления коэффициент 101—103
— — — пороговый 184, 186
Квантовые переходы 351
— числа 337, 338
Квантовый выход 219
Керра эффект 235
Кирхгофа метод 162
Клистрон отражательный 333
— пролетный 323
Когерентность 136
— временная 144
— пространственная 147
Коэффициент когерентности 139
— шума 134, 135

Лазер газовый 241
— — атомный 245
— — ионный 247
— — молекулярный 249
— жидкостный 239—241
— на иттрийалюминиевых гранатах 224
— — рубине 218—221
— — — стекле с неодимом 222
— полупроводниковый 253
— — инжекционный 257
— — с электронным возбуждением 259
Лампа бегущей волны, баланс фаз и амплитуд 309, 310
— обратной волны 311

Магнетрон 316
Максвелла уравнения 12, 16, 24
Метод медленно меняющихся амплитуд и фаз 382

Намагниченность среды 23, 24
Насыщение 43, 78
Неопределенности принцип 356, 357

Парамагнетизм 260
Парамагнитный резонанс 77
Паули принцип 346
Плазма 63—68
Плотности матрица 369
Поглощение резонансное 37
Поккельса эффект 234
Поляризация среды 19—22, 24
— — ионная 60
— — ориентационная 61
— — электронная 46—60
Последовательных приближений метод 380
Процесс линейный 13
— нелинейный 13—15
— параметрический 14, 15

Режим модуляции добротности 232
— — свободного генерирования 228
Резонатор открытый 150
— — добротность 182—184
— — конфокальный 164
— — плоский, дифракционные потери 161

Резонатор открытый плоский
типы колебаний 151, 158, 160
— — плотность типов колеба-
ний 150

Релаксации время 42

— процесс 41, 42, 71—73

Релаксация спин-решеточная
71, 72

— спин-спиновая 72, 73

Самоканализация 87—89

Самофокусировка 86, 87

Сегнетоэлектрики 62, 63

Снелиуса законы 92

— — модифицированные 95

Сопротивление связи 300

Спонтанные излучения 35, 37

Среда активная 31, 32, 35, 40,
41, 45

— диэлектрическая 46—63

— квантовая двухуровневая 35

— линейная 13

— нелинейная 13, 14

— парамагнитная 69—78

— параметрическая 14

— пассивная 31, 32, 35, 40, 41,
45

— ферромагнитная 78—86

Ток намагниченности 17, 18

— поляризации 17

Умова — Пойнтинга теорема
26, 27

Уравнение состояния среды 12,
16

Усилитель квантовый 107

Усилитель квантовый бегущей
волны 119, 120, 124

— — — условия усиления 190

— — оптического диапазона бе-
гущей волны 196, 197

— — — регенеративный 216,
217

— — регенеративный 104, 105

Усилитель парамагнитный 261

— — бегущей волны 265

— — регенеративный 275

Условия синхронизма электро-
нов и бегущей волны 295, 300

Фазовой плоскости метод 390

Ферми распределение 44, 45,
347, 348

— уровень 349—351

Ферромагнитный резонанс 81

Фредгольма уравнения 164

Френеля формулы модифициро-
ванные 97, 98

— число 161

Ширина спектральной линии
естественная 375, 376

— — — доплеровски уширен-
ная 379

Шредингера уравнение 355

Штарка эффект 342, 352

Эйнштейна коэффициент для
индуцированного излучения
37, 38

— — — поглощения 38

— — — спонтанного излучения
32—38

ОГЛАВЛЕНИЕ

Предисловие : : :	3
Основные обозначения	6
Глава первая. Основные уравнения и характеристики электромагнитного поля и среды	12
1-1. Нелинейные и параметрические электромагнитные процессы в средах	12
1-2. Основные уравнения электромагнитного поля	15
1-3. Энергия электромагнитного поля	26
1-4. Пассивные и активные среды	31
Глава вторая. Нелинейные процессы в пассивных средах	46
2-1. Распространение электромагнитного поля в безграничных диэлектрических средах	46
2-2. Распространение электромагнитного поля в неограниченной плазме	63
2-3. Распространение электромагнитного поля в неограниченной магнитной среде	68
2-4. Самофокусировка и самоканализация энергии электромагнитного поля : : :	86
2-5. Наклонное падение на границу изотропных линейной и нелинейной диэлектрических сред без потерь	90
Глава третья. Усиление и генерирование в активных квантовых средах	100
3-1. Квантовое усиление и генерирование	100
3-2. Квантовые генераторы и усилители	105
3-3. Создание инверсной населенности в средах	108
3-4. Усиление бегущей волны. Общая теория	114
3-5. Усиление в резонаторе. Общая теория	124
3-6. Шумы квантовых усилителей	130
Глава четвертая. Оптические квантовые генераторы и усилители	135
4-1. Когерентность излучения	135
4-2. Открытые резонаторы	149
4-3. Селекция типов колебаний :	190
4-4. Усилитель бегущей волны : :	196
4-5. Теория оптических квантовых генераторов (ОКГ)	208
4-6. Лазеры на твердых диэлектриках	217
4-7. Газовые лазеры	241
4-8. Полупроводниковые лазеры	253
Глава пятая. Квантовые генераторы и усилители СВЧ диапазона	260
5-1. Усиление электромагнитной волны в парамагнитной среде	260
5-2. Парамагнитный усилитель бегущей волны	265

5-3. Парамагнитные регенеративные усилители	275
5-4. Генератор на пучке молекул аммиака	284
Глава шестая. Электронные генераторы и усилители СВЧ	
диапазона	292
6-1. Усиление электромагнитной волны в электронных по- токах	292
6-2. Лампы бегущей волны типа О	297
6-3. Приборы магнетронного типа	312
6-4. Клистроны	323
<i>Приложение 1. Физические дополнения</i>	<i>337</i>
<i>Приложение 2. Математические дополнения</i>	<i>380</i>
Список литературы	394
Предметный указатель	396

Александр Михайлович Кугушев
Нина Сергеевна Голубева

**Основы радиоэлектроники
(нелинейные электромагнитные процессы)**

Редактор *В. Г. Карташев*
Редактор издательства *И. Н. Суслова*
Переплет художника *Е. В. Никитина*
Художественный редактор *Д. И. Чернышев*
Технический редактор *Н. Н. Хотулева*
Корректор *И. А. Володяева*

ИБ № 705

Сдано в набор 12/I 1977 г. Подписано к печати 26/IV 1977 г.
Т-08433. Формат 84×108¹/₃₂. Бумага типографская № 3.
Усл. печ. л. 21. Уч.-изд. л. 20,85. Тираж 58 000 экз. Зак. 13
Цена 99 коп.

Издательство «Энергия», Москва, М-114, Шлюзовая наб., 10

Владимирская типография Союзполиграфпрома
при Государственном комитете Совета Министров СССР
по делам издательств, полиграфии и книжной торговли
600610, г. Владимир, ул. Победы, д. 18-б.

