Отсюда

$$R^2(x)=\frac{2\mathscr{E}}{N}(x-x_0)^2+R_0^2, \qquad (109,27)$$

где x_0, R_0 — постоянные. Мы видим, что при $\mathscr{E}<0$ на конечном расстоянии вдоль направления распространения достигается полная фокусировка пучка — его радиус R обращается в нуль[1]).

Этот результат, полученный в рамках приближенного уравнения (109,13), не может иметь буквального физического смысла вблизи самого фокуса, где нарушаются предположения, сделанные при выводе уравнения. Достаточно сказать, что при неограниченном увеличении плотности энергии поля при точной фокусировке уже нет оснований ограничиваться низшей степенью нелинейности — кубической. Но существенна уже самая возможность самофокусировки пучка до такой степени, когда нелинейность перестает быть малой. Подчеркнем, что установленный критерий носит достаточный, но не необходимый характер. Пучок с $\mathscr{E}<0$ заведомо фокусируется целиком, но расходимость пучка в среднем при $\mathscr{E}\geqslant 0$ не противоречит тому, что некоторая внутренняя его часть сфокусируется.

§ 110. Генерация второй гармоники

В § 107 были рассмотрены лишь некоторые общие соотношения, относящиеся к характерным для нелинейной оптики процессам преобразования частот. Теперь мы изложим количественную теорию типичного такого процесса — генерации второй гармоники, т. е. возбуждения электромагнитного поля с частотой 2ω полем частоты ω (*Р. В. Хохлов*, 1960; *J. A. Armstrong*, *N. Bloembergen*, *J. Ducuing*, *P. S. Pershan*, 1962).

Генерация второй гармоники — нелинейный эффект второго порядка. Он содержится в тензоре нелинейной восприимчивости

$$\varepsilon_{i,kl}(-2\omega;\ \omega,\ \omega) \qquad (110,1)$$

и потому отсутствует в средах, допускающих пространственную инверсию. Тензор (110,1) симметричен по индексам kl; его свойства симметрии в различных кристаллах — те же, что и у пьезоэлектрического тензора (§ 17). Мы будем считать среду непоглощающей, так что $\varepsilon_{i,kl}$ — вещественные величины.

Задача о генерации второй гармоники может быть поставлена следующим образом. Пусть на плоскую поверхность кристалли-

---

[1]) Отметим, что если распределение $|E_0|^2$ по сечению пучка не меняется вдоль его длины, то $R^2=\mathrm{const}$ и $\mathscr{E}=0$. Обратное, однако, не обязательно: могут существовать решения, в которых $\mathscr{E}=0$, так что $R=\mathrm{const}$, но распределение зависит от x.

ческой среды падает плоская монохроматическая волна с частотой ω. Наряду с отраженной и двумя преломленными (в двупреломляющем кристалле) волнами той же частоты возникают также отраженная и преломленные волны частоты 2ω. Волны этой частоты в кристалле представляют собой решение уравнений (109, 5—6), в которых нелинейный член в индукции, $\mathbf{D}^{(2)}$, должен быть выражен через поле основной волны. Амплитуды всех этих волн выражаются через амплитуду падающей волны с помощью граничных условий, на которых мы не будем останавливаться. Разумеется, амплитуда волн частоты 2ω будет мала в меру малости нелинейной восприимчивости [1]).

Преломленные волны распространяются далее в глубь кристалла как в неограниченной среде. По мере их распространения эффекты нелинейности накапливаются и интенсивность гармоник может достичь больших значений — происходит перекачка энергии из основной частоты в гармонику. Именно этот процесс и будет интересовать нас здесь. Условия же на поверхности кристалла играют при этом лишь роль «начальных» условий, задающих некоторую малую, но отличную от нуля, амплитуду поля второй гармоники. Эти же условия задают (для данного направления падающей волны) волновые векторы первой $\mathbf{k}_1$ и второй $\mathbf{k}_2$ гармоник в кристалле.

Как будет видно из дальнейшего, эффективная перекачка энергии происходит, только если выполняется *условие синхронизма* основной волны и гармоники [2]):

$$\mathbf{k}_2 \approx 2\mathbf{k}_1. \tag{110,2}$$

Подчеркнем, что для самой постановки задачи о генерации одной лишь второй гармоники принципиальное значение имеет наличие дисперсии. В отсутствие дисперсии условие (110,2) выполняется при преломлении автоматически вместе с аналогичными условиями и для более высоких гармоник ($\mathbf{k}_3 \approx 3\mathbf{k}_1$ и т. п.). При наличии дисперсии это уже не так и можно считать, что условие синхронизма, выполняясь для второй гармоники, не выполняется для других. Подчеркнем, что уже само условие (110,2) фактически может быть выполнено, только если основная волна и гармоника относятся к различным поляризационным типам и, соответственно, имеют различные законы дисперсии.

---

[1]) Расчет условий отражения и преломления на границе нелинейной среды для некоторых частных случаев — см. статью *Bloembergen N., Pershan P. S.* — Phys. Rev., 1962, v. 128, p. 606 (русский перевод в книге *Бломберген Н.* Нелинейная оптика. — М.: Мир, 1966).

[2]) Природа этого условия особенно ясна с квантовой точки зрения, когда генерация второй гармоники рассматривается как «слияние» двух фотонов в один. Равенство $\hbar\mathbf{k}_2 = 2\hbar\mathbf{k}_1$ выражает сохранение импульса в этом процессе.

Запишем поле в среде как суперпозицию двух волн:

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 = \mathrm{Re}\left[\mathbf{e}_1 E_{10} e^{i(\mathbf{k}_1\mathbf{r} - \omega t)} + \mathbf{e}_2 E_{20} e^{i(\mathbf{k}_2\mathbf{r} - 2\omega t)}\right], \quad (110,3)$$

причем, в силу условия (110,2),

$$\mathbf{k}_2 = 2\mathbf{k}_1 + \mathbf{q}, \quad (110,4)$$

где $q \ll k_1$. Амплитуды волн представлены в виде произведений $\mathbf{E}_0 = \mathbf{e}E_0$, где $\mathbf{e}$ — единичный ($\mathbf{ee}^* = 1$) вектор поляризации. В линейном приближении эти амплитуды были бы постоянны, а с учетом нелинейности они — медленно (т. е. мало на расстояниях $\sim 1/k_1$) меняющиеся функции координат.

Уравнения для амплитуд обеих волн получаются путем подстановки (110,3) в уравнения Максвелла (109,5—6) и отделения в них членов с одинаковой зависимостью от времени. Мы не станем подробно приводить здесь эти простые, но громоздкие вычисления и ограничимся лишь некоторыми принципиальными указаниями.

Мы будем искать решения уравнений, описывающие стационарную генерацию второй гармоники в кристалле бегущей основной волной; такие решения не зависят от времени. При точном синхронизме ($\mathbf{q} = 0$) уравнения для амплитуд не содержали бы вовсе координат в явном виде; при неточном синхронизме координаты входят в уравнение только в комбинации $\mathbf{qr}$ (в множителе $\exp(-i\mathbf{qr})$). Выбрав направление вектора $\mathbf{q}$ в качестве оси z, можно поэтому искать решения уравнений, зависящие только от z. Если исходить из указанной выше постановки вопроса о падении волны ω на поверхность кристалла, то задача однородна в плоскостях, параллельных этой поверхности. Поэтому ось z ей перпендикулярна (перпендикулярность вектора $\mathbf{q}$ поверхности кристалла выполняется автоматически в силу граничных условий).

В линейном приближении волны в анизотропной (негиротропной) среде линейно поляризованы (см. § 97); для них $\mathbf{e}$ может быть определен как вещественный вектор, и именно эти значения для обеих волн будут подразумеваться ниже под $\mathbf{e}_1$ и $\mathbf{e}_2$. Если разложить амплитуду $\mathbf{E}_0$ каждой волны по направлениям $\mathbf{e}$, $\mathbf{k}$, $[\mathbf{ek}]$, то составляющие вдоль двух последних направлений будут малы в меру малости производных dE_0/dz, появляющихся как эффект нелинейности. Составляющие же вдоль $\mathbf{e}$ приближенно совпадают с абсолютными величинами E_0 векторов $\mathbf{E}_0$. Уравнения для них получаются умножением уравнения (109,5) на вектор $\mathbf{e}_1$ или $\mathbf{e}_2$. Поскольку волны с $E_0 = \mathrm{const}$ предполагаются точными решениями уравнений Максвелла в линейном приближении, то все линейные члены в уравнениях, не содержащие производных по z, взаимно сокращаются. Члены же с составляющими $\mathbf{E}_0$ по направлениям $\mathbf{k}$ и $[\mathbf{ek}]$ (которые могли бы быть того же порядка величины, что и члены с производными dE_0/dz), как оказывается, выпадают при указанном умножении; это обстоятельство связано с ортогональностью индукции **D** направлениям как **k**, так и [**ek**] (см. (97,3)).

Ввиду предполагаемой медленности зависимости амплитуд от координат, в уравнениях можно пренебречь вторыми производными от $\mathbf{E}_0$ по z. Поэтому, например, имеющееся в уравнении для $\mathbf{E}_2$ (умноженном на $\mathbf{e}_2$) выражение

$$e_{2i}\left\{(\operatorname{rot}\operatorname{rot})_{ik}-\frac{4\omega^2}{c^2}\,\varepsilon_{ik}\,(2\omega)\right\}e_{2k}E_{20}\,e^{i\mathbf{k}_2\mathbf{r}}$$

с учетом всего сказанного приближенно сводится к

$$2i\mathbf{e}_2\,[\mathbf{k}_2[\mathbf{l}\mathbf{e}_2]]\,\frac{dE_{20}}{dz}$$

(где **l** — орт оси z), и аналогично для $\mathbf{E}_1$.

Окончательные уравнения, получающиеся в результате описанных действий, гласят:

$$\begin{aligned}\alpha_2\frac{dE_{20}}{dz}&=-i\eta e^{-iqz}E_{10}^2,\\ \alpha_1\frac{dE_{10}^*}{dz}&=i\eta e^{-iqz}E_{10}E_{20}^*,\end{aligned}\tag{110,5}$$

где введены обозначения[1]):

$$\eta=\frac{\omega^2}{2c^2}\,\varepsilon_{i,kl}(\omega,\ \omega)\,e_{2i}\,e_{1k}\,e_{1l},\tag{110,6}$$

$$\alpha_1=\mathbf{l}\,[\mathbf{e}_1[\mathbf{k}_1\mathbf{e}_1]],\quad \alpha_2={}^1/_2\,\mathbf{l}\,[\mathbf{e}_2[\mathbf{k}_2\mathbf{e}_2]]\approx\mathbf{l}\,[\mathbf{e}_2[\mathbf{k}_1\mathbf{e}_1]].\tag{110,7}$$

Умножив первое уравнение (110,5) на E_{20}^*, а второе — на E_{10} и сложив их, найдем первый интеграл этой системы:

$$\alpha_1|E_{10}|^2+\alpha_2|E_{20}|^2=\mathrm{const}\equiv P.\tag{110,8}$$

Он выражает собой постоянство суммарного (в обеих волнах) потока энергии вдоль оси z[2]).

От комплексных величин удобно перейти к вещественным — абсолютным значениям и фазам величин E_{10} и E_{20}. Для максимального упрощения уравнений введем новые неизвестные ρ_1, ρ_2,

---

[1]) Первое из уравнений (110,5) получается из членов с $e^{-2i\omega t}$, а второе — из членов с $e^{i\omega t}$. В правых сторонах этих уравнений использованы соотношения (108,13).

[2]) Усредненная по времени плотность потока энергии в волне $\mathbf{E}_1$:

$$\overline{S}_{1z}=\frac{c}{8\pi}\,\mathrm{Re}\,\mathbf{l}\,[\mathbf{E}_{10}^*\mathbf{H}_{10}]=\frac{c^2}{8\pi\omega}\,\mathbf{l}\,[\mathbf{E}_{10}^*\,[\mathbf{k}_1\,\mathbf{E}_{10}]]=\frac{c^2\alpha_1}{8\pi\omega}\,|E_{10}|^2$$

и аналогично для волны $\mathbf{E}_2$.

φ_1, φ_2 как безразмерные величины, определив их согласно

$$E_{10}=\sqrt{\frac{P}{\alpha_1}}\rho_1 e^{i\varphi_1}, \quad E_{20}=\sqrt{\frac{P}{\alpha_2}}\rho_2 e^{i\varphi_2}. \tag{110,9}$$

Система уравнений (110,5) инвариантна относительно преобразования

$$\varphi_1 \to \varphi_1 + c, \quad \varphi_2 \to \varphi_2 + 2c.$$

Поэтому фактически в ней отделяются уравнения для функций ρ_1, ρ_2 и для инвариантной комбинации $2\varphi_1-\varphi_2$, образующие вместе замкнутую систему уравнений. Они гласят:

$$\frac{d\rho_1}{d\zeta}=-\rho_1\rho_2\sin\theta, \quad \frac{d\rho_2}{d\zeta}=\rho_1^2\sin\theta, \tag{110,10}$$

$$\frac{d\theta}{d\zeta}=-s-\left(2\rho_2-\frac{\rho_1^2}{\rho_2}\right)\cos\theta, \tag{110,11}$$

где

$$\theta=2\varphi_1-\varphi_2-s\zeta \tag{110,12}$$

и введены безразмерная переменная

$$\zeta=z\eta\sqrt{\frac{P}{\alpha_1^2\alpha_2}} \tag{110,13}$$

и безразмерный параметр

$$s=\frac{q}{\eta}\sqrt{\frac{\alpha_1^2\alpha_2}{P}}. \tag{110,14}$$

Первый интеграл (110,8) в этих переменных принимает вид

$$\rho_1^2+\rho_2^2=1. \tag{110,15}$$

Рассмотрим случай точного синхронизма: $q=0$, т. е. $s=0$. Тогда уравнения (110,10—11) имеют еще один первый интеграл:

$$\rho_1^2\rho_2\cos\theta=\text{const}\equiv\delta \tag{110,16}$$

(причем постоянная $\delta^2\leqslant {}^4/_{27}$; это следует, как легко убедиться, из равенств (110,15—16) в силу условия $|\cos\theta|\leqslant 1$). С использованием обоих этих интегралов решение уравнений (110,10) приводится к квадратурам — к эллиптическому интегралу

$$\zeta=\pm\frac{1}{2}\int\limits_{\rho_2^2(0)}^{\rho_2^2(\zeta)}\frac{du}{[u(1-u)^2-\delta^2]^{1/2}}; \tag{110,17}$$

выбор знака перед интегралом определяется знаком начального (при $\zeta=0$) значения $\sin\theta$. Кубическое уравнение

$$u(1-u)^2-\delta^2=0 \tag{110,18}$$

имеет (при $\delta^2 \leqslant {}^4/_{27}$) три вещественных положительных корня, из которых два меньше единицы; обозначим последние как ρ_a^2 и ρ_b^2, причем $\rho_a^2 < \rho_b^2$[1]). Определяемая формулой (110,17) функция $\rho_2^2(\zeta)$ периодически меняется в пределах между этими значениями с периодом, равным интегралу

$$\int\limits_{\rho_a^2}^{\rho_b^2} \frac{du}{[u(1-u)^2-\delta^2]^{1/2}}. \qquad (110,19)$$

Аналогичным образом меняется и функция $\rho_1^2(\zeta) = 1 - \rho_2^2(\zeta)$, причем, когда одна из них максимальна, другая минимальна.

Величину ρ_a^2 надо отождествить со значением интенсивности второй гармоники, $\rho_2^2(0)$, задаваемой граничными условиями на поверхности кристалла ($z=0$). Мы видим, что в направлении оси z в глубь кристалла происходит периодическая перекачка энергии из основной частоты во вторую гармонику и обратно. При уменьшении $\rho_2(0)$ период этого процесса возрастает, и при $\rho_2(0) \to 0$ стремится к бесконечности по логарифмическому закону. Предельному значению $\rho_2(0) = \rho_a = 0$ отвечает решение

$$\rho_1 = 1/\operatorname{ch}\zeta, \qquad \rho_2 = \operatorname{th}\zeta, \qquad (110,20)$$

получающееся из (110,17) при $\delta=0$ элементарным интегрированием; в нем амплитуда второй гармоники монотонно возрастает и асимптотически при $\zeta \to \infty$ вся энергия переходит из основной частоты в гармонику.

Рассмотрим теперь обратный случай, когда амплитуда ρ_2 остается везде малой по сравнению с ρ_1. Мы увидим, что этот случай отвечает значительной расстройке синхронизма волн.

При $\rho_2 \ll \rho_1$ можно, в первом приближении, считать ρ_1 постоянным ($\rho_1 = \rho_1(0)$), а уравнения для ρ_2 и θ записать в виде

$$\frac{d\rho_2}{d\zeta} = \rho_1^2(0)\sin\theta, \qquad \frac{d\theta}{d\zeta} = -s + \frac{\rho_1^2(0)}{\rho_2}\cos\theta.$$

Решение этих уравнений, равное нулю в начальной точке $\zeta=0$:

$$\rho_2(\zeta) = \frac{2}{s}\rho_1^2(0)\sin\frac{s\zeta}{2}, \qquad \theta = \frac{\pi}{2} - \frac{s\zeta}{2}. \qquad (110,21)$$

Эти формулы определяют изменение поля на протяжении одного интервала $0 \leqslant \zeta \leqslant 2\pi/s$ (т. е. $0 \leqslant x \leqslant 2\pi/q$), после чего процесс повторяется периодически [2]). Условие $\rho_2 \ll \rho_1$ означает, что должно

---

[1]) Между этими корнями полином в левой стороне (110,18) имеет максимум в точке $u={}^1/_3$, равный ${}^4/_{27}-\delta^2$; при $\delta^2 = {}^4/_{27}$ этот максимум обращается в нуль, два вещественных корня сливаются, а затем исчезают.

[2]) В каждом последующем периоде к постоянному члену в фазовой переменной θ надо добавлять по π. В точке, где $\rho_2=0$, фаза φ_2 теряет смысл и разность фаз θ может испытывать скачок.

быть $\rho_1^2(0)/s \ll \rho_1(0)$, т. е. $s \gg \rho_1(0)$ или

$$qz_0 \gg 1, \quad z_0 \sim 1/\eta\rho_1(0)\sqrt{P}.$$

Это — условие сравнительно большой расстройки синхронизма. Вообще, величина параметра q определяет, какой эффект ограничивает генерацию гармоники (т. е. рост амплитуды ρ_2) — линейный эффект нарушения синхронизма при $qz_0 \gg 1$ или нелинейные эффекты при $qz_0 \ll 1$ [1]).

Выше в этом параграфе мы говорили везде о генерации второй гармоники за счет основной частоты. Но рассмотренные уравнения описывают и обратное явление: усиление (его называют *параметрическим*) слабого сигнала частоты ω в поле интенсивного излучения частоты 2ω; этот процесс, рассмотрением которого мы здесь ограничимся, представляет собой простейший случай более общего явления — усиления сигналов с различными частотами ω_2 и $\omega_1 - \omega_2$ в поле интенсивной волны с частотой ω_1 (*С. А. Ахманов, Р. В. Хохлов*, 1962; *R. H. Kingston*, 1962).

Прежде всего, подчеркнем следующее отличие этого процесса от генерации второй гармоники. Последняя могла начинаться с нулевой интенсивности гармоники. Возможность же усиления основной частоты во всяком случае требует хотя бы малой ее начальной интенсивности: если в начальной точке $\rho_1(0) = 0$, то так останется везде; из уравнений (110,10) видно, что вместе с ρ_1 обращаются в нуль также и производные всех порядков от ρ_1 и ρ_2.

Снова рассмотрим случай точного синхронизма и, сверх того, пусть начальное значение фазовой переменной $\theta(0) = -\pi/2$; при точном синхронизме это значение будет сохраняться. Тогда параметр $\delta = 0$ в силу равенства $\cos\theta = 0$, хотя начальные значения ρ_1 и ρ_2 отличны от нуля. Решение уравнений (110,10) в этом случае есть

$$\rho_1 = 1/\mathrm{ch}(\zeta - \zeta_0), \quad \rho_2 = -\mathrm{th}(\zeta - \zeta_0), \tag{110,22}$$

где $\zeta_0 > 0$ — постоянная. При большом значении этой постоянной начальное значение $\rho_1(0) = 1/\mathrm{ch}\,\zeta_0$ мало. Мы видим, что вдоль оси z в глубь кристалла происходит усиление волны основной частоты за счет интенсивности гармоники. Последняя убывает до нуля (при $\zeta = \zeta_0$), а затем снова возрастает, пока асимптотически вся интенсивность не окажется сосредоточенной в ней [2]).

---

[1]) Напомним, что все проведенное рассмотрение основано на предположении $q \ll k_1$. Это условие уже полностью использовано при выводе уравнений (110,5) и малого параметра q/k_1 в них не осталось. Говоря о случае $qz_0 \gg 1$, мы предполагаем, конечно, что он совместим с условием $q \ll k_1$.

[2]) При $\zeta > \zeta_0$ фазовой переменной надо приписать значение $\theta = \pi/2$ и изменить знак перед th в $\rho_2(\zeta)$.