6

Spontaneous and Stimulated Emission

We have already mentioned the term 'spontaneous radiation' in Chap. 5. It was used as a synonym with the notion of the random–phase wave fields summation. In terms of statistics, radiation emitted at random phases is characterized by a spectral–angular distribution of the average power flow. The total field phase is also random in this context. At the same time, it is clear that any realization of the emitter ensemble could be, in a sense, coherent if capable of preserving the fixed correlation between individual emitters during a time interval sufficiently long. For instance, the process of the regular wave scattering by a fixed lattice of charged particles meets these conditions (see Sect. 5.2.2). Surely, the oscillation phase of each particle, prescribed by the wave under scattering, remains correlated with the particle location even if the latter is random. Therefore, a certain degree of coherence is inherent in the total radiation, emitted by this ensemble. Naturally, if the emitters are characterized by a regular spatial distribution, the effects of the radiation coherence are more expressive.

The above–given reasoning relates to fixed ensembles of emitters. As regards the systems which are substantially nonequilibrium (e.g., flows of charged particles), there arises a question: to what extent can regular spatial distributions of emitters keep the coherence in time? Such steady regular distributions might radically alter the basic characteristics of spontaneous radiation, i.e., its spectral–angular distribution and total radiation power. On the one hand, it seems that regular distributions of a large number of independent emitters cannot survive. The matter is that, influenced by a lot of uncontrollable disturbing factors (thermal spread of emitter velocities, collisions, etc.), radiation would rather quickly lose its coherence. On the other hand, symmetry of the processes of emission and absorption of radiation by individual particles indicates that identical emitters cannot be independent of one another: each of them reacts to the radiation fields emitted by other particles. As a result of this, the amplitude and, what is more, the phase of the emitter proper radiation field is subjected to certain changes. There takes place such a reaction even if individual emitters are quasi neutral (e.g., atoms)

and separated from one another by large distances. As regards charged particles in a dense beam, their interaction can also be influenced by comparatively short–range Coulomb forces. The necessity of taking into account these forces essentially complicates the quantitative description of the radiation emission by the corresponding ensembles.

In rarefied ensembles, Coulomb fields are negligible and collective interaction between individual emitters is realized via long–range microwave fields only. The symmetry between the processes of the radiation emission and absorption means that the processes are mutually complementary. Absorption is, of course, a stimulated process meaning that its rate is proportional to the existing electromagnetic power. So is the complementary process that can be called the stimulated emission. The corresponding notions had been for the first time introduced by A. Einstein [2], who used them for the analytical explanation of the black body equilibrium radiation law. In the course of development of quantum theory, the mechanism of the stimulated radiation emission was explained directly.

The quantum genesis of the notion somewhat hampered its application to classical (nonquantum) systems, in particular to intense flows of high–energy electrons. Perhaps, only elaboration of the theory of free electron lasers has revealed the profound correlation between the stimulated emission and the principle of operation of quite traditional devices of microwave electronics (e.g., klystrons or TWT). To emphasize this inner link, we will discuss briefly the quantum and classical interpretations of notions of the spontaneous and stimulated emission. Peculiarity of the mechanisms of realization of these effects in classical ensembles of emitters (flows of charged particles) is also to be discussed below.

6.1 Semiquantum Interpretation

Einstein's semiphenomenological theory is based on the following consideration: in an equilibrium ensemble of emitters the spectral distribution of the radiation energy density is a result of a mutual balance of elementary acts of emitting and absorbing the field energy quanta by individual emitters. For deriving quantitative characteristics of intensity of these processes, the notions of probabilities of the stimulated ('st') and spontaneous ('st') emission of photons as well as the probability of their stimulated absorption were introduced. Transitions between the energy levels E_m and E_n were stimulated by the external microwave radiation of frequency ω_{mn}. The transition probability was supposed to be proportional to the spectral density of the radiation field energy $\varrho\left(\omega_{mn}\right)$:

$$w_{mn}^{\mathrm{st}} = B_{mn}\varrho\left(\omega_{mn}\right). \tag{6.1}$$

Probabilities of emission ($m > n$) and absorption ($m < n$) processes per unit time were supposed to be equal so that $B_{mn} = B_{nm}$.[1] The latter statement corresponded to the supposition that both the processes were symmetric.

Under the equilibrium condition of the 'emitters + field' system, the total numbers of the emission and absorption transitions per unit time have to be equal:

$$w_{mn}^{\mathrm{tot}} N_m = w_{nm}^{\mathrm{tot}} N_n. \tag{6.2}$$

Here N_m and N_n are numbers of the emitters with energies E_m and E_n, respectively. These numbers in equilibrium are correlated via Boltzman's distribution:

$$\frac{N_n}{N_m} = \exp\left(-\frac{\hbar\omega_{mn}}{kT}\right). \tag{6.3}$$

(For simplicity, the energy levels themselves are considered to be nondegenerated.) As it follows from (6.2) and (6.3), the transition probabilities are in the ratio:

$$\frac{w_{mn}^{\mathrm{tot}}}{w_{nm}^{\mathrm{tot}}} = \exp\left(\frac{\hbar\omega_{mn}}{kT}\right). \tag{6.4}$$

Physically it is evident that absorption transitions can be nothing but stimulated ones:

$$w_{nm}^{\mathrm{tot}} = w_{mn}^{\mathrm{st}} = B_{nm}\varrho\left(\omega_{mn}\right); \qquad m > n. \tag{6.5}$$

However, as (6.4) indicates, if the value of $\hbar\omega/kT$ is finite, the total probability of the radiation emission w_{mn}^{tot} is larger than the probability of the stimulated radiation emission w_{mn}^{st}. This conclusion is equivalent to the statement that radiation emission, in contrast to absorption, is possible even at the so–called 'zero' energy of the external microwave field. So, the notion of additional ('spontaneous') radiation, independent of the energy density of the stimulating field $\varrho\left(\omega_{mn}\right)$, has been introduced:

$$w_{mn}^{\mathrm{sp}} = w_{mn}^{\mathrm{tot}} - w_{mn}^{\mathrm{st}}. \tag{6.6}$$

It was the absence of any dependence on the external field that conditioned calling such transitions the spontaneous ones (i.e., those taking place without any external influence).

As it follows from the above-given expressions, the density of the radiation field energy is determined by the relation:

$$\varrho\left(\omega_{mn}\right) = \frac{w_{mn}^{\mathrm{sp}}}{B_{mn}\left[\exp\left(\hbar\omega_{mn}/kT\right) - 1\right]}. \tag{6.7}$$

The relation (6.7) should look like the classical Rayleigh–Jeans spectrum when $\hbar\omega_{mn} \to 0$:

[1] Later on these parameters were called Einstein's coefficients.

$$\varrho\left(\hbar\omega_{mn} \ll kT\right) = \frac{\omega_{mn}^2}{\pi^2 c^3} kT.$$

Making use of this limit, one can relate the coefficients B_{mn} to the probability of the spontaneous emission w_{mn}^{sp}:

$$B_{mn} = \frac{\pi^2 c^3}{\hbar\omega_{mn}^3} w_{mn}^{\mathrm{sp}}. \tag{6.8}$$

Substitution of this expression into the right–hand side of (6.7) finally yields the classical Planck formula:

$$\varrho\left(\omega_{mn}\right) = \frac{\hbar\omega_{mn}^3}{\pi^2 c^3 \left[\exp\left(\hbar\omega_{mn}/kT\right) - 1\right]}\ . \tag{6.9}$$

Within the quantum theory, correctness of the semiphenomenological considerations used as a basis for derivation of (6.8) and (6.9) has been completely proved as well as the equality between the probabilities of the stimulated absorption and emission. Really, matrix elements of these transitions have turned out to be modulo equal and proportional to the amplitude of the radiation field at the frequency corresponding to this transition. As the probabilities of the emission and absorption transitions are proportional to the square of the matrix element, these characteristics turn out to be also proportional to the spectral density of the radiation field energy. In this sense, the notion of spontaneous transitions, postulated by Einstein, looks like being somewhat contradictory to the causality principle. Really, the atom, not influenced by any external force, ought to stay on the corresponding excited level. At present, this apparent paradox is only of historical interest because it has been obviated because of development of quantum electrodynamics. Briefly speaking, the essence of the matter is explicable in the following way.

The spectral density of the free radiation field energy may be presented as the sum of energies $\hbar\omega$ of the photons characterized by the corresponding frequency:

$$\varrho\left(\omega_{mn}\right) = \hbar\omega_{mn} n\left(\omega_{mn}\right).$$

Here $n(\omega)$ is the number of such photons. Consequently, the relation (6.6) of the total probability to the probability of the spontaneous transition may be written as

$$w_{mn}^{\mathrm{tot}} = \left[1 + n\left(\omega_{mn}\right)\right] w_{mn}^{\mathrm{sp}}. \tag{6.10}$$

The unity in the square brackets corresponds to the spontaneous emission, which takes place even if $n(\omega) = 0$.

On the other hand, as it is known from quantum electrodynamics, the spectral density of the field total energy is determined by the formula:

$$E = \left[\frac{1}{2} + n\left(\omega\right)\right] \hbar\omega\ .$$

The first term on the right-hand side of this equation corresponds to the so-called zero oscillations, i.e., exactly to the case when $n(\omega) = 0$.

An interesting interpretation follows from this expression. The spontaneous radiation emission is not spontaneous in the literal sense: the process is, at least half, stimulated by zero oscillations of the electromagnetic field. A detailed analysis (e.g., see [4]) indicates that this explanation does make sense. And what is more, the second half of the total probability of the spontaneous radiation emission is also stimulated. However, in this case, zero oscillations of the emitter itself play the role of the stimulating factor.

In the experiment, the existence of spontaneous emission transitions has to indicate itself in two ways. First, it is some broadening of spectral lines, which is conditioned by finiteness of the atom life time in the excited state: $\tau_{\rm sp} \propto (w_{\rm sp})^{-1}$. Second, positions of the corresponding energy levels are to be shifted with respect to their locations in the absence of zero oscillations. As a matter of fact, both the effects have been observed experimentally. This fact makes the basic proof of the determining role played by the zero oscillations in the spontaneous radiation emission (even if to say nothing about brilliant conformity of Planck formula with the whole totality of the experimental data).

In general, phases of the spontaneous radiation, directions of the propagation of the corresponding waves, and their polarizations are not correlated because the zero oscillations themselves are of the occasional nature. Physically, it means that the spontaneous radiation emitted by an ensemble is incoherent. Therefore, the above–studied radiation emission by the system of particles, the power of which is equal to the sum of the radiation powers of individual emitters, is also called spontaneous.

There are two suppositions implied in this reasoning. Both of them are inherent in the system of emitters that is in dynamic equilibrium with the isotropic radiation. First, photons are characterized only by the frequency ω without fixation of direction of radiation propagation. Second, the emitter energy distribution is supposed to be thermodynamically equilibrium (Boltzmann distribution). It is easy to generalize the first supposition for the number of photons of a given mode $n(\mathbf{k})$ ($\mathbf{k}$ implies the totality of the indexes, describing the mode).[2] Other suppositions, such as the symmetry of the Einstein coefficients and their link with the spontaneous transition probability, can now be considered as proved by the quantum relations:

$$\begin{aligned} B_{mn} &= B_{nm}; \\ w^{\rm tot}_{mn} &= w^{\rm sp}_{mn}\,[1+N]\,; \\ w^{\rm tot}_{\mathbf{k}\downarrow} &= w^{\rm sp}_{\mathbf{k}\downarrow}\,[1+N]\,. \end{aligned} \tag{6.11}$$

The arrows here indicate the energy variation as a result of the transition.

These relations are not linked with the thermodynamic equilibrium supposition. So, one may write down a kind of a kinetic equation describing a temporal evolution of an average number of photons of the wave vector $\mathbf{k}$

[2] As regards free plane waves, $\mathbf{k}$ is the wave vector.

interacting with emitters of energy $\mathcal{E}$. If the emitters energy distribution is $f(\mathcal{E},t)$, then

$$\frac{\partial n_{\mathbf{k}}}{\partial t} = \int \{f(\mathcal{E},t)w_{\mathbf{k}\downarrow} + n_{\mathbf{k}}f(\mathcal{E},t)w_{\mathbf{k}\downarrow} - n_{\mathbf{k}}f(\mathcal{E}-\hbar\omega,t)w_{\mathbf{k}\uparrow}\}\,\mathrm{d}\mathcal{E}. \qquad (6.12)$$

The first term in the curled brackets describes the probability of the spontaneous emission of a photon $\mathbf{k}$ of energy $\hbar\omega$. The third term corresponds to the absorption process rate, which is proportional to the probability $w_{\mathbf{k}\uparrow}$, to the number of existing photons[3] $n_{\mathbf{k}}$, and to the population of the level $\mathcal{E}-\hbar\omega$ the transition starts from.

The second term describes an inverse process of the induced emission proportional to the number of photons as well. As has been discussed above, this addendum is necessary to provide a stationary distribution with $\partial n/\partial t = 0$ if the emitters energy distribution is Boltzmann one. Really, in the steady state with $f(\mathcal{E}) \propto \exp(-\mathcal{E}/\kappa T)$, the expression in the brackets has to vanish so that

$$n(\mathbf{k}) = \frac{1}{\exp(\hbar\omega/\kappa T)-1}\,. \qquad (6.13)$$

Multiplying (6.13) by the quantum energy $\hbar\omega$ and by the phase volume for the isotropic radiation $4\pi|\mathbf{k}|^2\,\mathrm{d}\mathbf{k}$, one gets the Planck formula (6.9).

It is worth to note here that if the energy distribution is inverse, i.e. if $f(\mathcal{E}) > f(\mathcal{E}-\hbar\omega)$, the second term could be predominant. Under this condition, exponential growing of number of photons takes place and all of them are exact copies of the first "initiating" photon. One can easily recognize this as lasing which is due to the stimulated emission. In what follows we consider this effect of main importance in the classic limit omitting effects specific for quantum emitter (i.e., for quantum lasers).

6.2 Classical Limit

Transferring the concept of stimulated and spontaneous emission to classical (nonquantum) systems meets certain difficulties. Obviously, this is the spontaneous emission to be identified with numerous examples in Part I, where no external electromagnetic waves influencing the particle motion were considered. However, there is a small discrepancy in this approach: the radiation phase is strictly determined for a classical particle in contrast with the quantum spontaneous radiation. It is not surprising: in the quantum description with a determined number of photons the field phase is not defined because of the uncertainty principle, while a number of photons are meaningless in classical electrodynamics.

[3] We keep the same notation $n_{\mathbf{k}}$ for the quantum average of the number of photons.

The physical meaning of the stimulated emission as additional to the spontaneous one is much less obvious in classical electrodynamics. The extra acceleration caused by an external wave does generate some radiation. However, the process is weak and must be regarded as scattering because wave vectors of the corresponding waves do not coincide with that of the incident wave as it should be for stimulated processes.

In spite of that, the concept above can be definitely applied to classical beam systems. Really, the absorption – radiation symmetry for a given wave can be understood as the equality of phase intervals corresponding to the increase and decrease in the particle energy. The temporal variation of the number of quanta is to be interpreted as variations in the field energy spectral density. As regards the quantum discreteness of energy levels, it was used above only for derivation of Planck formula and was not essential for mentioned lasing. Summing, one should expect that the kinetic relation (6.12) is valid even for $\hbar \to 0$ and that it could describe the stimulated effects in classical nonequilibrium systems with inverse populations.

In the classical limit, the photon energy and the distance between energy levels are negligible. So, the energy population can be considered as continuous and can be presented as the expansion:

$$f(\mathcal{E} - \hbar\omega) \approx f(\mathcal{E}) - \hbar\omega \frac{\partial f}{\partial \mathcal{E}} + \cdots . \tag{6.14}$$

Multiplying (6.12) by $\hbar\omega$ and using (6.14), one gets

$$\frac{\partial W_{\mathbf{k}}}{\partial t} = \int \left\{ f(\mathcal{E}) + \frac{\partial f}{\partial \mathcal{E}} W_{\mathbf{k}} \right\} p_{\mathbf{k}}(\mathcal{E}) \mathrm{d}\mathcal{E}, \tag{6.15}$$

where $p_{\mathbf{k}}$ is the intensity of the spontaneous emission of the wave $\mathbf{k}$ by an emitter of energy $\mathcal{E}$. The value $W_{\mathbf{k}} = \hbar\omega n_{\mathbf{k}}$ is, of course, the electromagnetic energy density of the mode. Integrating the second term in right-hand side of (6.15) by parts and putting, for simplicity, $f(\mathcal{E}_{min}) = 0$, we obtain

$$\frac{\partial W_{\mathbf{k}}}{\partial t} = P_{\mathbf{k}} - W_{\mathbf{k}} \int f(\mathcal{E}) \frac{\partial p_{\mathbf{k}}}{\partial \mathcal{E}} \mathrm{d}\mathcal{E}, \tag{6.16}$$

where

$$P_{\mathbf{k}} = \int f(\mathcal{E}) p_{\mathbf{k}}(\mathcal{E}) \mathrm{d}\mathcal{E}$$

is the spontaneous radiation power integrated over all emitters. In particular, for a monoenergetic ensemble[4] of energy $\mathcal{E}_0$

$$\frac{\partial W_{\mathbf{k}}}{\partial t} = P_{\mathbf{k}} - W_{\mathbf{k}} \left(\frac{\partial P_{\mathbf{k}}}{\partial \mathcal{E}} \right)_0 . \tag{6.17}$$

[4] We call an ensemble monoenergetic if its energy distribution is narrow enough but still is much wider than $\hbar\omega$ to ensure the expansion (6.14).

This equation is of the same structure as (6.12) but does not contain $\hbar$. It shows that the energy spectral density variations take place because of the spontaneous and stimulated emission, the last being proportional to the density itself. The lasing effect is also presented. However, the condition of lasing looks different than that in a quantum system. It says that the derivative $\partial P_{\mathbf{k}}/\partial\mathcal{E}$ has to be negative and large enough. In other words, to get lasing the intensity of spontaneous emission of the given mode should depend sharply on the emitter energy. In quantum systems it is ensured automatically because of the levels discreteness and thus the inverse population is of the main importance. On the contrary, a classical beam system is populated obviously only at large energies so that the inverse population exists for granted and means just the free energy availability. In this case, this is the sharp dependence of the spontaneous emission on energy that must be ensured.

Of course, the relation (6.17) obtained as a limiting case of (6.12) does not explain the mechanism of the stimulated emission in classical electrodynamics. Both of them are based on the energy considerations and contain no information about the field phase. It was completely approved in a steady state when the number of photons was fixed and the phase was random. However, for a temporal evolution of the average number of photons the average phase has to depend on time as well. Really, growing predominance of the emission over absorption (lasing) may take place from the classical point of view only under condition of developing phasing of individual emitters. This phasing is nothing but an appearance of a certain coherence. These arguments lead to the conclusion that the stimulated emission of a given field mode is, in a way, equivalent to the development of its coherence.

Coming to beam systems with the expressed direction of motion along the z-axis a narrow-band character of the spontaneous spectrum is associated with the condition of synchronism which links the phase velocity of the wave and the particle longitudinal velocity. The width of the spectrum is conditioned, first of all, by the finite wave–particle interaction distance. The profile of the spectrum can be obtained from the general considerations. Really, if the particle velocity has a component oscillating with a frequency $\Omega(\gamma)$, the amplitude of a quasi-synchronous spectral harmonic is proportional to

$$k\int_0^L \exp\left[\mathrm{i}\left((\omega\pm\Omega)z/v - kz\right)\right]\mathrm{d}z = \mathrm{i}kL\frac{\exp\left(-\mathrm{i}\mu\right)-1}{\mu}, \tag{6.18}$$

where

$$\mu = \left(1-\frac{\omega\mp\Omega}{kv}\right)kL \tag{6.19}$$

is the phase slippage of the particle with respect to the wave at the total distance of interaction L. The spontaneous emission power as a function of μ is proportional to the absolute value of (6.18) squared:

$$P_k(\mu) = P_k(0)\frac{\sin^2(\mu/2)}{(\mu/2)^2}, \tag{6.20}$$

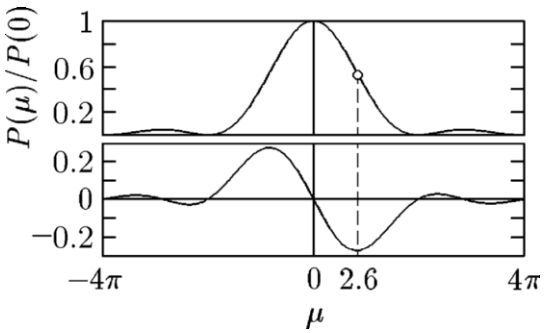

Fig. 6.1. Profile of the spontaneous emission line and its derivative (below)

This universal expression for a profile of the spontaneous radiation spectral line of a single particle at a finite length is shown in Fig. 6.1.

The value of μ depends on the phase velocity and the particle velocity in the fixed combination. So, coming back to (6.17), one can state: for a fixed frequency an equilibrium energy exists ($\mu = 0$) that provides the maximal spontaneous emission. At this energy the induced emission vanishes, but it appears for nonequilibrium energies if $\mu > 0$ where the corresponding derivative in (6.17) is negative. At the opposite side of the resonance ($\mu < 0$), there is the region of the wave absorption. For large μ the particle neither radiates nor absorbs. Treating the particle beam as a medium, one can talk about a band of its optical activity coinciding with the spontaneous emission band. It consists of two symmetric subbands: one of absorbtion and that of stimulated emission (see Fig. 6.1). The latter is possible, of course, only in an active (inversely populated) medium with some intrinsic free energy.

6.3 Stimulated Emission and Beam Phasing

Of course, the arguments above should be considered just as leading ones. First of all, it is unclear to what extent they are valid for nonstationary systems. Second, they do not take into account a degradation of the initial energy distribution, i.e., inevitable saturation effects and beam energy spreading. The last but not the least, the arguments are of a phenomenological character and do not reveal the physics of the correlations developing in the beam of particles. So far as the correlations are related to phasing, we briefly consider below the particles phase dynamics in an external wave.

6.3.1 Phase Dynamics in Quasi-Synchronous Wave

In Chap. 5, we have considered the coherent emission of a single mode by a structure of individual emitters. Now the problem is, in a way, opposite:

this is the development of a spatial structure under action of a monochromatic wave that is of interest. One can call it the second side of the united process of the development of coherence both in the particle motion and in the electromagnetic field.

For the sake of simplicity we consider the electric field of the form

$$\mathbf{E}(z,t) = \mathbf{E}_0 \exp(\mathrm{i}\varphi)\,; \qquad \varphi = kz - \omega t\,; \quad z \geq 0$$

neglecting its dependence on transverse coordinates and treating the amplitude $\mathbf{E}_0$ as constant. Doing this, we ignore the initial stage of the process when the stimulated monochromatic radiation is just appearing against the spontaneous background exactly as we ignored above the prehistory of the emitters lattice. The constancy of the amplitude implies an input signal large enough to be practically unchangeable by additional radiation. Beside that, we neglect Coulomb interaction of the particles. It is clear, by intuition, that all these simplifications are approved only for low-intensity beams.

The particles motion in the absence of the wave is supposed to be a superposition of the longitudinal velocity βc and of transverse oscillations of frequency ω, small enough not to influence the longitudinal velocity (in a particular case of Cherenkov interaction the oscillation amplitude can be zero). We shall accept these conditions for granted because they are quite obvious in many cases of interest (e.g., for an undulator). The constant particle energy means then the constant rate of the particle phase slipping with respect to the wave

$$\frac{\mathrm{d}\varphi}{\mathrm{d}z} = k\left(1 - \frac{\omega \mp \Omega}{k\beta c}\right). \tag{6.21}$$

For a nonzero wave amplitude the energy and the phase slippage undergo variations which can be presented as a superposition of slow (in ω-scale) systematic changes and ripples vanishing in average. Omitting the sign of averaging, one can write down an obvious relation for the systematic part

$$\frac{\mathrm{d}\gamma}{\mathrm{d}z} = gk\cos\varphi, \tag{6.22}$$

where g is a maximal possible increase in the particle energy per a wavelength expressed in mc^2 units. This dimensionless amplitude is a small parameter in the overwhelming majority of cases of interest. Of course, it depends on field and trajectory configurations but is always proportional to the wave amplitude.

The radiation reaction, that is, the proper field of a single particle is not included in the equation. This approximation is valid if the width of the spontaneous radiation spectral line is determined mainly by a finite length of the interaction path rather than by particle acceleration.

For a synchronous particle, by definition, the phase slippage is zero because its velocity $\beta_\mathrm{s} = (\omega \mp \Omega)/kc$. For small energy deviations from the equilibrium, we can present the phase shift per a unit of length as

$$\frac{\mathrm{d}\varphi}{\mathrm{d}z} = \alpha k \left(\gamma - \gamma_\mathrm{s}\right), \tag{6.23}$$

where the index s denotes synchronous values and

$$\alpha = -\beta_\mathrm{ph} \left(\frac{\partial}{\partial \gamma} \frac{\omega \mp \Omega}{\omega \beta} \right)_\mathrm{s} \tag{6.24}$$

is the phase slippage sensibility to energy variations.

Equations (6.22) and (6.23) describe the phase stability mechanism well known in the theory of accelerators (see, e.g., [33]). For $g = \mathrm{const}$ they have an integral of motion (Hamiltonian):

$$\mathcal{H} = (\gamma - \gamma_\mathrm{s})^2 - \frac{2g}{\alpha} \sin\varphi, \tag{6.25}$$

which predicts stable "synchrotron" oscillations[5] around equilibrium values $\varphi_\mathrm{s} = \mathrm{sign}(\alpha)\pi/2$ and $\gamma = \gamma_\mathrm{s}$. If deviations from the equilibrium values are small, the Hamiltonian (6.25) can be presented as

$$\mathcal{H} = (\gamma - \gamma_\mathrm{s})^2 + \frac{g}{|\alpha|} (\varphi - \varphi_\mathrm{s})^2 \ .$$

This positive quadratic form corresponds to harmonic oscillations of period $2\pi\sqrt{|\alpha|/g}$ in space. Note that this period is expressed in units of the wavelength and usually exceeds the latter.

The synchrotron oscillations are nonlinear, their period increasing with the amplitude. For $\mathcal{H} = 4\pi/|\alpha|$, the period becomes infinitely large. The corresponding phase trajectory

$$(\gamma - \gamma_\mathrm{s})^2 = \frac{4g}{|\alpha|} + \frac{2g}{\alpha} (\sin\varphi - \sin\varphi_\mathrm{s})$$

is called a separatrix, dividing trapped particles oscillating around the equilibrium from nontrapped or librating ones. The latter ones slip in phase with respect to the wave in the positive or negative direction, depending on the sign of $\alpha(\gamma - \gamma_\mathrm{s})$. The separatrix passes through the points $\gamma = \gamma_\mathrm{s}$ and $\varphi = \varphi_\mathrm{s} \mp \pi$ with a maximum deviation from the φ-axis

$$\gamma_\mathrm{max} - \gamma_\mathrm{s} \pm \sqrt{4g/|\alpha|} \tag{6.26}$$

taking place at $\varphi = \varphi_\mathrm{s}$. The qualitative structure of other phase trajectories (supplied with arrows) is shown in the Fig. 6.2.

6.3.2 Phase Bunching by External Wave (Low-Gain Regime)

Let us consider now evolution of particles initially distributed uniformly over phases and having the same initial energy $\gamma_\mathrm{i} > \gamma_\mathrm{s}$. In the plane (φ, γ), this

[5] Also known in electronics as bounce oscillations.

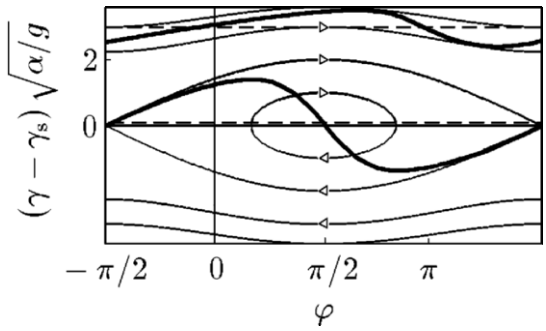

Fig. 6.2. Phase trajectories and beam bunching at $zk < \pi/2\sqrt{\alpha/g}$. Dashed lines show initial distributions

distribution is presented by a straight line $\gamma = \gamma_\mathrm{i}$ (see Fig. 6.2). Moving along bent phase trajectories, half of the particles are accelerated and half decelerated (i.e., are absorbing or radiating). At the beginning these processes are quite symmetric what corresponds to the symmetry of Einstein coefficients. But accelerated particles move faster to the right in Fig. 6.2 and overtake the decelerated ones. As a result, the particles begin to bunch in the phase region of deceleration, disturbing the initial symmetry between absorption and radiation in the latter's favor. If the initial energy is lesser than γ_s, the bunching process goes in opposite direction resulting in absorption. One can easily see that this scenario describes the amplification and absorption bands discussed above. If the initial energy is outside the optical activity band, the phase trajectories are almost straight lines, bunching vanishes, and the beam stays transparent.

The exact solution of (6.22) and (6.23) can be obtained in terms of elliptical functions. However, it is rather cumbersome and is not really necessary if we limit ourselves by the initial stage of bunching described above. Because of the smallness of the amplitude g, one can exploit a perturbation theory if the interaction distance is not very large.

In the zeroth approximation ($g = 0$) just kinematic slipping takes place:

$$\varphi = \varphi_\mathrm{i} + \zeta; \quad \zeta = \alpha k \delta_\mathrm{i} z; \quad \delta = \gamma - \gamma_\mathrm{s}.$$

Substituting this in (6.22) gives the induced energy modulation of the first-order

$$\delta - \delta_\mathrm{i} = \frac{g}{\alpha k \delta_\mathrm{i}} \left[\sin(\varphi_\mathrm{i} + \zeta) - \sin\varphi_\mathrm{i}\right]$$

and the corresponding dynamical phase slippage:

$$\varphi = \varphi_\mathrm{i} + \zeta + \frac{g}{\alpha \delta_\mathrm{i}^2} \left[\cos\varphi_\mathrm{i} - \cos(\varphi_\mathrm{i} + \zeta) - \zeta \sin\varphi_\mathrm{i}\right]. \tag{6.27}$$

Note that both values vanish after averaging over initial phases. This means, in particular, that the first-order radiation losses are zero because the numbers of absorbing and radiating particles are equal (one can remind again the

symmetry of Einstein coefficients). However, the induced phase shift (6.27) does disturb the symmetry. In the next approximation

$$\frac{\mathrm{d}\gamma}{\mathrm{d}z} = kg\cos(\varphi_\mathrm{i}+\zeta) - \frac{g^2k}{\alpha\delta_\mathrm{i}^2}\sin(\varphi_\mathrm{i}+\zeta)\left[\cos\varphi_\mathrm{i} - \cos(\varphi_\mathrm{i}+\zeta) - \zeta\sin\varphi_\mathrm{i}\right]. \quad (6.28)$$

Averaging over initial phases gives now

$$\left\langle\frac{\mathrm{d}\gamma}{\mathrm{d}z}\right\rangle = -\frac{g^2k}{2\alpha\left(\gamma_\mathrm{i}-\gamma_\mathrm{s}\right)^2}\left[\sin\zeta - \zeta\cos\zeta\right], \quad (6.29)$$

i.e., an additional induced energy change proportional to the field amplitude squared. The average energy loss for this induced or stimulated radiation can be calculated by integrating (6.29) once more and noting that the kinematic phase shift at the length L is $\mu = \alpha kL\delta_\mathrm{i}$:

$$\langle\gamma(L)-\gamma_\mathrm{s}\rangle = g^2\frac{\alpha k^3L^3}{\mu^3}\left[\cos\mu - 1 + \frac{\mu}{2}\sin\mu\right] = g^2\frac{\alpha k^3L^3}{4}\frac{\mathrm{d}}{\mathrm{d}\mu}\left(\frac{\sin\mu/2}{\mu/2}\right)^2. \quad (6.30)$$

To calculate the radiation power emitted by the beam as a whole, one should multiply (6.30) by the number of particles passing the region per unit time, i.e., by I/q where I is the beam current. This power obviously contains the spontaneous radiation spectral line profile (6.20), exactly in the same way as the phenomenological expression (6.16). Beside, the average radiation losses are proportional to g^2, i.e., to the external wave power. So, the stimulated emission in a classical system really can be interpreted as coherence self-organization due to the autophasing mechanism with a consequent increase in the radiation spectral brightness within the optical activity region.

The dependence of the radiation power on the phase slippage parameter deserves a special comment. For a fixed interaction length the power is maximized by $\mu \approx 2.6$, that is, the wave that overtakes the beam almost by a wavelength is the most prosperious.[6] Note that the exactly synchronous wave is not amplified at all while all slow waves are attenuated.

These calculations predict the evolution of the initially monoenergetic beam as well. The method of successive approximations used above works only if the induced phase shift is small enough, or under qualitative conditions:

$$\delta_\mathrm{i} \gg \sqrt{g/\alpha}\,\min\{\mu,1\}, \quad (6.31)$$

$$kL \ll \frac{1}{\sqrt{\alpha g}}\min\{\mu,1\}. \quad (6.32)$$

[6] This is true for $g = \mathrm{const}$ only, i.e., for the low-gain regime.

So, the case $\mu > 1$ corresponds to particles situated mainly or totally outside the separatrix. Note that this is the condition to realize the optimizing value $\mu = 2.6$ (see the footnote on page 109). In the case of $\mu < 1$ almost all particles can be initially trapped.

If the interaction length exceeds the value (6.32), one can predict the faster particles outrunning the slower ones, the distribution "overturning" and filamentation taking place. At the final stage of this mixing, the phase distribution would be symmetric again with a corresponding increase in the energy spread. The larger is the initial μ value the later this filamentation occurs. From the viewpoint of physics the process can be interpreted as a nonlinear saturation of the stimulated emission (or absorbtion) accompanied by beam heating. The corresponding length (by order of magnitude)

$$L_{\text{sat}} = 1/k\sqrt{\alpha g}$$

can be called the distance of saturation.

Although the results of this section are restricted by the fixed field approximation, i.e., do not take into account possible amplitude and phase variations, they can be directly used in some cases of interest. In particular, if the field is "locked" in a cavity of a finite Q-value and the amplification exceeds certain threshold, one may foresee a steady state with a time independent established amplitude. For large Q, this amplitude could be large enough although the transient process takes a long time. This low-gain regime is typical for generators of coherent radiation with low current beams, where the feedback necessary for self-excitation is provided by a cavity.[7] These problems will be considered in detail in Part III.

6.3.3 Spatial Amplification in Particles Flow (High-Gain Regime)

Nevertheless, high gain systems are also important from the general viewpoint as well as for applications. Suppose that there is no feedback and the steady state self-consistent field depends essentially on the longitudinal coordinate being determined by an input signal and by emitted radiation. This regime may be called a high-gain spatial amplification of the input signal by the beam. Naturally, only those input waves could be amplified that are inside the beam optical activity domain.

Of course, the spatial amplification depends on electrodynamic properties of the system as a whole. But, basing on general arguments, one should expect that a quasi-synchronous mode would be mainly amplified if, of course, it is presented in the input signal. This is the mode that, according to the previous considerations, pumps the energy out of the beam most effectively. So, we restrict ourselves below by this one-mode approximation, bearing in mind that modes are independent in a linear system. Other suppositions are the same as above including the near-zone interaction being neglected.

[7] Or by a negative group velocity typical for backward wave tubes.

So, we consider now the field amplitude growing with distance but being independent of time. Note, by the way, that the beam represents both active and reactive loads for the wave changing its phase $\psi(z)$ as well as the amplitude $g(z)$. Taking this into account, the equations of phase dynamics (6.22), (6.23) are to be rewritten as

$$\frac{\mathrm{d}\delta}{\mathrm{d}z} = kg(z)\cos(\phi+\psi)\ , \qquad \frac{\mathrm{d}\phi}{\mathrm{d}z} = k\alpha\delta. \tag{6.33}$$

In the same way as in the previous section we get in the first approximation with respect to g:

$$\begin{aligned}\delta &= \delta_\mathrm{i} + k\int_0^z g(z')\cos\left(\varphi_\mathrm{i} + k\alpha\delta_\mathrm{i}z' + \psi(z')\right)\,\mathrm{d}z';\\ \varphi &= \varphi_\mathrm{i} + k\alpha\delta_\mathrm{i}z\\ &\quad + k^2\alpha\int_0^z \mathrm{d}z'\int_0^{z'} g(z'')\cos\left(\varphi_\mathrm{i} + k\alpha\delta_\mathrm{i}z'' + \psi(z'')\right)\mathrm{d}z''\end{aligned}$$

or, with the same precision,

$$\begin{aligned}&\frac{\mathrm{d}\varphi_\mathrm{i}}{\mathrm{d}\varphi} - 1\\ &\quad = k^2\alpha\int_0^z \mathrm{d}z'\int_0^{z'} g(z'')\sin\left[\varphi - k\alpha\delta_\mathrm{i}\left(z-z''\right) + \psi(z'')\right].\end{aligned} \tag{6.34}$$

Note that this expression describes particles bunching in the φ space.

Now we need the second equation relating field variations to beam bunching. Restricting ourselves by plane motion with the amplitude of the oscillation velocity $\tilde{v}$, we note that a single particle at the phase φ creates a resonant harmonic of the transverse current equal to

$$\frac{q\tilde{v}}{4\pi\beta c}\exp\left[\mathrm{i}\left(kz - \varphi(z)\right)\right].$$

Averaging it over all initial phases with the help of (6.34) yields the driving transverse current

$$\begin{aligned}j_\perp &= -j_0\frac{\mathrm{i}k^2\alpha\tilde{v}}{8\pi\beta c}\exp\left(\mathrm{i}kz\right)\\ &\quad\times\int_0^z \mathrm{d}z'\int_0^{z'} g(z'')\exp\left[-\mathrm{i}k\alpha\delta_\mathrm{i}\left(z-z''\right) + \mathrm{i}\psi(z'')\right]\mathrm{d}z''.\end{aligned} \tag{6.35}$$

Looking for a solution of the wave equation for the corresponding transverse component of the electric field

$$\frac{\mathrm{d}^2E}{\mathrm{d}z^2} + k^2E = -\mathrm{i}\frac{4\pi k}{c}j_\perp,$$

we present it as $E_0(z)\exp(\mathrm{i}kz)$, where the complex amplitude $E_0(z)$ is a slow function of the argument. Neglecting its second derivative, we get

$$\frac{\mathrm{d}E_0}{\mathrm{d}z} = \mathrm{i}\frac{k^2 j_0 \tilde{v}\alpha}{4c^2\beta} \times \int_0^z \mathrm{d}z' \int_0^{z'} g(z'')\exp\left[\mathrm{i}\left(-k\alpha\delta_\mathrm{i}\left(z-z''\right)+\psi(z'')\right)\right]. \tag{6.36}$$

Note now that, by definition,

$$g = \frac{q\tilde{v}}{2k\beta mc^3}\left|E_0\right| \quad \text{and} \quad g\exp(\mathrm{i}\psi) = \frac{q\tilde{v}}{2k\beta mc^3}E_0.$$

Substituting that into (6.36) yells the self-consistent equation for the electric field amplitude:

$$\frac{\mathrm{d}E_0}{\mathrm{d}z} = \frac{\mathrm{i}}{L_\mathrm{r}^3}\int_0^z \mathrm{d}z' \int_0^{z'} E_0(z'')\exp\left[-\mathrm{i}k\alpha\delta_\mathrm{i}\left(z-z''\right)\right]\mathrm{d}z'', \tag{6.37}$$

where

$$L_\mathrm{r} = 2\left(\frac{\beta^2 c^2 I_0}{kj_0\alpha\tilde{v}^2}\right)^{1/3} \tag{6.38}$$

with $I_0 = mc^3/q \approx 17$ kA. For reasons explained below, L_r can be called a radiation length.

Equation (6.35) can be easily reduced to the third-order differential linear equation:

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{\mathrm{d}}{\mathrm{d}z}+\mathrm{i}k\alpha\delta_\mathrm{i}\right)^2 E_0 = \frac{\mathrm{i}}{L_\mathrm{r}^3}E_0. \tag{6.39}$$

The solution of this equation under the initial condition

$$E_0 = 1\,; \quad \mathrm{d}E_0/\mathrm{d}z = 0\,; \quad \mathrm{d}^2E_0/\mathrm{d}z^2 = 0 \quad \text{for } t = 0 \tag{6.40}$$

represents a complex amplification coefficient describing amplitude and phase characteristics of the process. [8]

Looking for a solution in the form $\exp(\mathrm{i}\nu z)$, one gets the characteristic equation

$$\nu\left(\nu+k\alpha\delta_\mathrm{i}\right)^2 = -L_\mathrm{r}^{-3}. \tag{6.41}$$

An equation of this type will be investigated in details in Chap. 10. Here we just note that the coefficients in (6.41) are real. So it has either three real roots or one real and two complex conjugated ones. In the first case, all three linearly independent partial solutions are of an oscillatory type. Their linear combination cannot exceed essentially the initial field, meaning that

[8] We suppose that the final signal amplitude is small enough to exclude nonlinear processes.

amplification does not occur. In the second case, which takes place under condition

$$\mu_{\rm r} > -2^{-2/3}3\,; \qquad \mu_{\rm r} = k\alpha\delta_{\rm i}L_{\rm r} \tag{6.42}$$

one of the complex conjugated roots has a negative imaginary part and the corresponding partial solution grows exponentially with an e-fold length $L_{\rm r}$. Note that $\mu_{\rm r}$ has a meaning of the kinematic phase shift at the length $L_{\rm r}$ related to the corresponding detuning in initial energy.

If the distance z is essentially smaller than $L_{\rm r}$, as it happens for small currents, the amplification coefficient slightly exceeds unity. One can easily see that this leads to the result discussed above: at a fixed length, maximal amplification occurs for the wave with the optimal kinematic phase shift $\mu \approx 2.6$, but the gain itself is small and linearly proportional to the beam current.

Unlike the case of an almost constant amplitude, the high–gain amplification takes place not only at positive μ values but also at small negative ones if the condition (6.42) is fulfilled. The maximal gain is reached for the exactly synchronous wave[9] when the total solution under the initial condition (6.40) is relatively compact:

$$E_0 = \frac{1}{3}\left[\exp\left({\rm i}z/L_{\rm r}\right) + 2\exp\left(-{\rm i}z/2L_{\rm r}\right)\cosh\left(\sqrt{3}z/2L_{\rm r}\right)\right]; \tag{6.43}$$

$$|E_0| = \frac{1}{3}\sqrt{1 + 4\cosh\left(\sqrt{3}z/2L_{\rm r}\right)\cos\left(3z/2L_{\rm r}\right) + 4\cosh^2\left(\sqrt{3}z/2L_{\rm r}\right)}\,. \tag{6.44}$$

It is worth to note that the increment is proportional to $j_0^{1/3}$.

For $\mu = 0$ the characteristic equation (6.41) has the roots

$$\nu_n = L_{\rm r}^{-1}\exp\left[{\rm i}\pi n/3\right], \qquad n = 0, 1, 2.$$

The asymptotic behavior of the electric field amplitude is determined by the root with maximal imaginary part, i.e., $\nu_0 = L_{\rm r}^{-1}\exp\left({\rm i}\pi/3\right)$.

The mechanism of the spatial amplification is a basic one for a variety of high–power amplifiers, using high–current electron beams. Besides, it can be used in coherent radiation sources where spontaneous radiation plays the role of the input signal. The selective mechanism of the spatial amplification shares out a narrow spectral line from the spontaneous radiation spectrum. In the theory of free electron lasers, such regime is called SASE (Self-Amplification of Spontaneous Emission) and appears as a direct analog of optical superradiance [35]. The notion of "spontaneous coherent radiation" used sometimes is intrinsically contradictory on our opinion. We will return to these problems in Sect. 10.

Considerations above are related to a flow of harmonic oscillators with amplitude determined by $\tilde{v}$. Of course, the increase in the wave power comes

[9] We mean the asymptotic behavior.

from the total energy of particles which must be accompanied by a change of the oscillations amplitude. As it has been proved in Part I, a resonant interaction under conditions of a normal Doppler effect damps the amplitude and the amplification is limited because of oscillation energy exhaustion. Prolongation of interaction would lead to the inverse process of wave absorption. However, under conditions of the anomalous Doppler effect, this saturation does not exist and the oscillations amplitude grows at the account of an additional decrease in the particle longitudinal momentum.

6.4 Dynamic Chaos

All material above was based on the assumption that a particle interacts with a single harmonic wave under synchronous or resonance conditions. Being quite productive for explanation of the induced radiation mechanism as a result of particles self-bunching in the wave, this assumption still needs additional discussion of its applicability, especially in the case of large amplitude fields.

We do not mean here negligible changes of a particle trajectory still governed by external fields while radiation fields can be treated as perturbations. The dimensionless electric field amplitude g introduced by (6.22) remains small in practically all cases of interest. Thereby, all our arguments were based on consideration of resonances (Cherenkov type, Doppler-shifted oscillator resonances, cyclotron ones, etc.,). For small g factors only resonant conditions and a long-term wave–particle interaction can provide a large energy transfer from particles to the wave (amplifiers and oscillators) or vice versa (accelerators). In the case of large g, a particle could get a relativistic velocity during one period of the wave and the resonant conditions would lose their paramount importance. Electrodynamics of so large fields is still in the developmental stage.

Nevertheless, even within the frames of the resonant perturbation theory, one cannot exclude a simultaneous action on the particle of two waves of different frequencies satisfying approximate resonant conditions for two different degrees of freedom, for example, of Cherenkov and cyclotron type. For very small wave amplitudes, when these resonances are reliably separated, provide two well-separated stability regions one can consider them independently. But an increase in field amplitude which is desirable for high-power devices leads to an increase in the resonances width. In a sense they can act together, so that the particle motion in the phase plane becomes unpredictable and close to stochastic one. They call this phenomenon as a dynamic chaos limiting, naturally, the power increase and broadening the spectrum of oscillations. The last plays, of course, an essential role.

Remaining within the frames of the self-consisting theory and neglecting the radiation damping, one can consider fields as external ones and formulate the problem as canonical one. This permits to use the well-developed powerful formalism of Hamiltonian mechanics, especially the resonant perturbation

theory. This implies that the Hamiltonian of the dynamical system under consideration (in our case of a particle in the external field) can be presented as a sum of two terms: $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$. Here $\mathcal{H}_0$ is a nonperturbed Hamiltonian and $\mathcal{H}_1$ is a small periodic perturbation. Note that the canonical formulation of the problem implies the applicability of the Liouville theorem which simplifies essentially the description of the energy exchange between the particle and the field.

So, the processes of interest can be investigated using the theory of small perturbations acting over a long period of time. It will be shown below that the particle motion can be reduced to dynamics of either one nonlinear pendulum or that of a system of interacting pendulums. In the first case, the particle motion remains regular, and in the second, it can be chaotic. The conditions of this change will be discussed below, but before we discuss we need the basic notions of the resonant perturbation theory applied to the problems of microwave electronics.

6.4.1 Resonant Perturbation Theory

The main ideas and methods of the resonant perturbation theory will be considered as applied to motion of a particle under action of two waves, denoted below by indices 1 and 2. Let the system be described by a Hamiltonian in action–phase variables

$$\mathcal{H} = \mathcal{H}_0 \left(J_1, J_2\right) + \mathcal{H}_1 \left(J_1, J_2, \theta_1, \theta_2\right), \tag{6.45}$$

where $\mathcal{H}_0$ is a nonperturbed Hamiltonian and $\mathcal{H}_1$ is a perturbation supposed to be periodic over θ_1, θ_2. So, it can be presented as a Fourier series:

$$\mathcal{H}_1 = \sum_{l,n} H_{l,n} \exp\left[\mathrm{i}\left(l\theta_1 + n\theta_2\right)\right], \tag{6.46}$$

where l, n are integers. Suppose a resonance condition is fulfilled, that is, there exists a relation between the proper frequencies of the nonperturbed system ω_1, ω_2:

$$r\omega_1 - s\omega_2 \approx 0\,, \tag{6.47}$$

where

$$\omega_1 = \frac{\partial \mathcal{H}_0}{\partial J_1}\,, \quad \omega_2 = \frac{\partial \mathcal{H}_0}{\partial J_2}\,, \quad \dot{\theta}_1 = \omega_1\,, \quad \dot{\theta}_2 = \omega_2\,, \quad r, s \qquad \text{integers.}$$

It is necessary to describe slow (in the proper frequencies scale) but systematic variations of the values $J_{1,2}$, which are integrals of motion in the absence of perturbations.

Let us come to the variables denoted by a bar by means of a generating function

$$F_2 = \left(r\theta_1 - s\theta_2\right)\bar{J}_! + \theta_2 \bar{J}_2\,. \tag{6.48}$$

According to general rules [31], we get

$$J_1 = \frac{\partial F_2}{\partial \theta_1} = r\bar{J}_1, \quad J_2 = \frac{\partial F_2}{\partial \theta_2} = -s\bar{J}_1 + \bar{J}_2,$$
$$\bar{\theta}_1 = \frac{\partial F_2}{\partial \bar{J}_1} = r\theta_1 - s\theta_2, \quad \bar{\theta}_2 = \frac{\partial F_2}{\partial \bar{J}_2} = \theta_2. \tag{6.49}$$

The third equation of the system (6.49) yields

$$\dot{\bar{\theta}}_1 = r\dot{\theta}_1 - s\dot{\theta}_2 = r\omega_1 - s\omega_2 = 0.$$

So, the first new angular variable turns out to be a slow varying one while the second coincides with the original phase ($\bar{\theta}_2 = \theta_2$). Taking into account (6.46) and the resonant conditions (6.47), we obtain the new Hamiltonian:

$$\bar{\mathcal{H}} = \mathcal{H}_0\left(r\bar{J}_1, -s\bar{J}_1 + \bar{J}_2, r\bar{\theta}_1 - s\bar{\theta}_2, \bar{\theta}_2\right) + \sum_{l,n} H_{l,n}\left(\bar{J}\right) \exp\left[\frac{\mathrm{i}}{r}\left(l\bar{\theta}_1 + (ls + nr)\,\bar{\theta}_2\right)\right]. \tag{6.50}$$

Note that fast motion in (6.50) is represented by $\bar{\theta}_2$, while $\bar{\theta}_1$ is slow because of the resonant conditions (6.47). So, one can expect that the application of the perturbation theory would not give rise to appearance of new essential resonances. Really, as far as $\bar{\theta}_1$ is a slow variable while $\bar{\theta}_2$ is a fast one, a resonance can occur only at large values r. These are so-called secondary resonances developing under certain peculiar circumstances only. Otherwise resonant terms are not presented in (6.50). Thus, the procedure above reduces the problem to investigation of particle motion in a vicinity of a single chosen resonance.

It follows from (6.50) that there is only one slow varying term $\exp\left[\mathrm{i}\bar{\theta}_1 l/r\right]$ in the sum representing the perturbation. In so far as the variable $\bar{\theta}_1$ is slow, one may average (6.50) over the fast variable $\bar{\theta}_2$ (for justification of this method see, for example, [31]):

$$\bar{\mathcal{H}} = \bar{\mathcal{H}}_0(J) + \langle \mathcal{H}_1 \rangle \tag{6.51}$$

where

$$\langle \mathcal{H}_1 \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathcal{H}_1 \mathrm{d}\bar{\theta}_2.$$

Taking into account that the remaining term corresponds to $ls + nr = 0$, the perturbed Hamiltonian takes the form:

$$\langle \mathcal{H}_1 \rangle = \sum_{p=-\infty}^{\infty} H_{-rp,sp}(\bar{J}) \exp\left[-\mathrm{i}p\bar{\theta}_1\right] \quad \text{where} \quad p = -\frac{l}{r} \quad s = \frac{n}{l}r\,. \tag{6.52}$$

The new averaged Hamiltonian is independent of $\bar{\theta}_2$. So, the new canonical action is an integral of motion:

$$\bar{J}_2 = J_2 + s\bar{J}_1 = J_2 + \frac{s}{r}J_1 = \text{const}. \tag{6.53}$$

Hence, the original system (6.50) of two degrees of freedom is reduced to the system (6.51) with one degree of freedom. The new canonical momentum $\bar{J}_2$=const is an integral of motion and can be considered in what follows as a parameter.

Let us consider now the system described by the Hamiltonian (6.51). Traditionally, the first step is determination of stationary points. According to general rules, the equation for the stationary points has the form:

$$\dot{\bar{J}}_1 = -\frac{\partial\bar{\mathcal{H}}}{\partial\bar{\theta}_1} = 0; \quad \dot{\bar{\theta}}_1 = \frac{\partial\bar{\mathcal{H}}}{\partial\bar{J}_1} = 0. \tag{6.54}$$

For the overwhelming majority of systems of interest, members of the series (6.52) decrease with increasing p. Bearing this in mind, one can leave only three terms with numbers $p = 0, \pm 1$. It should be noted also that the perturbation is a real value, that is, $H_{-l,m} = H_{m,-l}$. In result the Hamiltonian (6.51) can be presented in the form:

$$\bar{\mathcal{H}} = \bar{\mathcal{H}}_0(\bar{J}) + H_{0,0}(\bar{J}) + 2H_{r,-s}(\bar{J})\cos\bar{\theta}_1 \ . \tag{6.55}$$

It follows from (6.54) and (6.55) that the stationary points are disposed at $2H_{r,-s}(\bar{J})\sin\bar{\theta}_1 = 0\,, \quad \bar{\theta}_{10} = 0\,, \quad \bar{\theta}_{11} = \pi$. The value of the new canonical action at these points is determined by the equation:

$$\frac{\partial\bar{\mathcal{H}}_0}{\partial\bar{J}_1} + \frac{\partial H_{0,0}}{\partial\bar{J}_1} + 2\frac{\partial H_{r,-s}}{\partial\bar{J}_1}\cos\bar{\theta}_1 = 0. \tag{6.56}$$

It should be noted also that as far as

$$\frac{\partial\bar{\mathcal{H}}_0}{\partial\bar{J}_1} = \frac{\partial\bar{\mathcal{H}}_0}{\partial J_1}\frac{\partial J_1}{\partial\bar{J}_1} + \frac{\partial\bar{\mathcal{H}}_0}{\partial J_2}\frac{\partial J_2}{\partial\bar{J}_1} = \omega_1 r - \omega_2 s \approx 0$$

the second term in (6.56) vanishes and the equation for stationary points takes the form:

$$\frac{\partial H_{0,0}}{\partial\bar{J}_1} \pm 2\frac{\partial H_{r,-s}}{\partial\bar{J}_1}\cos\bar{\theta}_1 = 0, \tag{6.57}$$

where the sign plus is to be chosen for $\bar{\theta}_1 = 0$ and the sign minus for $\bar{\theta}_1 = \pi$.

It should be noted that variations in the canonical action $\bar{\theta}_1$ can be large while changes in the variable $\bar{J}_1$ are small (proportional to perturbations). Hence, to describe the system near the resonance, the Hamiltonian (6.55) can be expanded into a Taylor series over powers of a small deviation from the stationary value of the canonical action $\bar{J}_{10}$:

$$\bar{\mathcal{H}} = \bar{\mathcal{H}}_0\left(\bar{J}_{10}\right) + \frac{\partial \mathcal{H}_0}{\partial \bar{J}_1}\left(\bar{J}_1 - \bar{J}_{10}\right) + \frac{1}{2}\frac{\partial^2 \mathcal{H}_0}{\partial \bar{J}_1^2}\left(\bar{J}_1 - \bar{J}_{10}\right)^2 + 2H_{r,-s}\left(\bar{J}_{10}\right)\cos\bar{\theta}_1 + \cdots \tag{6.58}$$

Then

$$\Delta\mathcal{H} \equiv \bar{\mathcal{H}} - \bar{\mathcal{H}}_0(\bar{J}_{10}) = G(\Delta\bar{J})^2 - F\cos\bar{\theta}_1, \tag{6.59}$$

where

$$G = \frac{1}{2}\frac{\partial^2 \mathcal{H}_0}{\partial \bar{J}_1^2} \quad \text{and} \quad F = 2H_{r,-s}\left(\bar{J}_{10}\right).$$

The Hamiltonian (6.59) known as a standard one describes the nonlinear system dynamics in the vicinity of a resonance. The most interesting cases of resonant interaction, including the particle phasing in a single external wave considered above, can be reduced to its analysis.

Note that exactly the same arguments could be applied to the variable θ_1. In other words, the approximation under consideration describes two independent resonances and yields the phase trajectory schematic pattern shown in Fig. 6.3a.

6.4.2 Randomization of Motion

Strictly speaking, the concept of a single isolated resonance is adequate to 1-D conservative systems only. There are two exact integrals of motion in such cases, namely an action and a phase of oscillation, which provide totally determined motion. In the overwhelming majority of cases, such conditions cannot be supported. Particularly, this relates to short-wave radiation where the electrodynamic system permits a coexistence of a large number of proper modes interacting with various degrees of particles freedom. Even for two degrees of freedom, the resonances interact, in a sense. This coupling leads to a principally new phenomena – to development of a dynamic chaos.[10] It should be emphasized that in microwave electronics this phenomenon mainly relates to a nonlinear character of a particle motion in a wave. However, sometimes a nonlinearity of interacting waves plays its role as well.

The possibility of chaotic regimes is of essential interest for both general theory and applications. Really, a chaotic electrodynamic system should generate wide spectrum radiation. This can be used for the development of powerful generators of electromagnetic noise. On the other hand, such noisy regime is definitely deleterious for narrow spectrum highly coherent radiation. In both cases, the region of parameters corresponding to the dynamic chaos is of a principal importance.

[10] This relates as well to open one-dimensional systems where an external force is in a certain resonant ratio with proper oscillations. Such systems can be described in a 3-D phase space and are usually used for simple illustration of the chaos development. However, such examples are hardly peculiar for our problems.

The germs of the dynamic chaos are those points of the phase space where two phase trajectories cross each over (homocline points). In our case they are the saddle points which belong to a separatrix. The reason is rather obvious: particles spend long time in the vicinity of these points and are subjects of small but long-acting perturbations. Then a so-called local instability appears when two points originally close together go away rapidly their separation increasing exponentially with time. This is exactly what happens when two points initially close but coming to a saddle point along different sides of a separatrix have quite different destinations.

Under action of a small perturbation provided by other resonances, the outgoing branches of a separatrix split and oscillate with increasing amplitude. As a result, they cross each over creating new homocline points where the process is repeated behaving like an avalanche. Then a kind of a particle trajectories web appears and the motion becomes indistinguishable from the stochastic one. Practically at every point phase trajectories run away from each other.

This scenario looks rather apocalyptically and can open to question the possibility of dynamical description of nonlinear systems in general. Fortunately, the reality is not so bad. There is a remarkable theorem by Kolmogorov–Arnold–Moser (KAM theorem) [36] telling that the dynamic chaos takes place only in a close vicinity of a homocline trajectory if the perturbation is small enough. Only a few phase trajectories leave this domain. So, the concept of almost independent resonances stated above still has a right for existence. But in the case of large enough perturbations, numerous computer simulations really show practically stochastic wandering of representing points over all phase space in the absence of external stochastic forces. This remarkable and comparatively new result of classical mechanics actually opens a way to the understanding of irreversible character of real physical processes in nondissipating systems. Moreover, if the dissipation does exist, a lot of remarkably new effects are predictable including so-called strange attractors, a fractal structure of phase portraits etc. This theory is intensively developing now and can be found in a row of specialized monographies [37]. Here we have to return to our main question – to the criterion of the dynamic chaos regime and to its consequences. Unfortunately, KAM theorem itself does not answer the question.

6.4.3 Criteria of Dynamic Chaos

The main problem in this context is a definition of the parameters domain where the system becomes chaotic under action of several resonances. There are several more or less formal criteria. In what follows we stay only on two simplest ones.

Lyapunov's Criterion

Consider a dynamic system described by the ordinary differential equations:

$$\dot{x}_i = f_i(\mathbf{x}), \qquad i = \{1...N\}, \tag{6.60}$$

where $\mathbf{x}(t)$ is an arbitrary trajectory of the system (6.60). Let $\mathbf{x}_1(t) = \mathbf{x}(t) + \delta\mathbf{x}(t)$ be another trajectory being in a close vicinity of $\mathbf{x}(t)$. Then for a small deviation $\delta\mathbf{x}(t)$, one gets the following system of ordinary differential equations:

$$\delta\dot{\mathbf{x}} = \mathsf{M}\delta\mathbf{x}, \tag{6.61}$$

where

$$\mathsf{M} = \{a_{ik}\}; \qquad a_{ik} \equiv \frac{\partial f_i}{\partial x_k}.$$

Generally, the matrix M depends on time, but we restrict ourselves at the moment by the case when it has constant elements. A validity of this approximation will be considered below.

For a constant matrix M, the general solution of (6.61) has the form:

$$\delta\mathbf{x} = \sum C_j \mathbf{e}_j \exp[\lambda_j t], \tag{6.62}$$

where $\mathbf{e}_j$ are the proper vectors of the matrix M, λ_j are the corresponding eigenvalues, and C_j are constants. It follows from (6.62) that for negative $\mathrm{Re}\,\lambda_j < 0$ the small deviations under consideration damp. They say that the system is locally stable in this case. However, if just one of the eigenvalues has a positive real part, the small deviations rise exponentially as well as deviations of the phase trajectories. In these cases, the system is locally unstable.

Using this criterion, we identify the local instability with the possibility of the dynamic chaos. The eigenvalues λ_j depend on the system (6.61) and those parameters which give $\mathrm{Re}\,\lambda_j > 0$ just for one eigenvalue determine the dynamic chaos domain.

Mathematically, this criterion is immaculate but difficult for applications if the matrix M depends on time. Then every point of the phase trajectory requires a separate investigation of the eigenvalues.

Chirikov's Criterion

Now we stay with another criterion based on simple physical arguments. In spite of its semi-intuitional character, it describes qualitatively a great variety of systems. One should not, of course, expect an exact quantitative result all the more that dynamic chaos appearance is sometimes rather sensitive to governing parameters.

Bearing in mind our particular interests, we shall consider now the motion of a particle in the fields of two waves satisfying close resonance conditions.

Let us start with one isolated resonance (e.g., Cherenkov one) for certain ω and k. As it was stated above, the equations of motion in this case look like the equations of a nonlinear pendulum:

$$\frac{\mathrm{d}\gamma}{\mathrm{d}z} = g\cos\varphi; \qquad \frac{\mathrm{d}\varphi}{\mathrm{d}z} = \alpha k(\gamma - \gamma_\mathrm{s}), \tag{6.63}$$

where g is a maximal increase in the particle energy (in units of mc^2) per unit of length which can be obtained from the wave. Note that we use the energy deviation instead of the action in the general theory 6.4.1. In a case of neighboring resonances, it does not matter but is more convenient for physical interpretations.

The analysis of Sect. 6.3.1 shows that the regions of vibration and librations (phase slippage) are separated by a which passes via points $\gamma = \gamma_s$ and $\varphi - \varphi_s \mp \pi$ and has a spread along the energy axis

$$\gamma_{max} - \gamma_s = \pm\sqrt{4g/k|\alpha|}$$

realized at $\gamma = \gamma_s$ and $\varphi - \varphi_s \mp \pi$. This spread should be identified as the nonlinear resonance width. The phase trajectory pattern was shown in Fig. 6.2.

Suppose now that the particle at the same time is in the vicinity of another resonance that produces the same picture by itself. If their interaction is weak enough, the resulting picture would look like that presented in Fig. 6.3

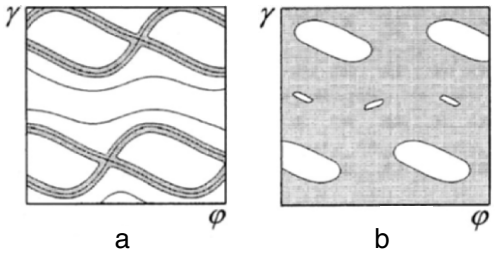

Fig. 6.3. A phase portrait of a nonlinear system under action of two resonances. (**a**) Almost independent resonances. (**b**) Overlapping resonances. The stochastic regions are shadowed

Note that according to the previous section the separatrix of an almost independent resonance has a narrow stochastic layer also shown in the figure.

According to the arguments above Fig. 6.3a describes satisfactory the motion if the waves phase velocities are essentially different. Otherwise, the resonances are close to each other, and the procedure of averaging above is not justified anymore. One can foresee a situation when the stochastic layers of both resonances are partially overlapped and a lot of additional homocline points appear. The corresponding set of parameters can be identified as a criterion of the dynamical chaos development[11] (see Fig. 6.3b). This condition suggested by B. Chirikov [38] is sometimes quoted as the resonances overlapping criterion. Numerous analytic and computer investigations show that it is in a good qualitative (and sometimes quantitative) agreement with reality for a wide variety of systems.

[11] The small additional islands of stability appear because of secondary resonances mentioned above.

Using this criterion to find the dynamic chaos condition, one has to find a nonlinear resonance width and a distance from the nearest adjacent resonance. In our particular case the latter is supposed to be a cyclotron one.

The distance between two resonances is determined by a difference between the corresponding equilibrium particle velocities or by a difference of equilibrium energies. If this distance turns out to be lesser than a sum of semi-widths of the resonances (including their stochastic layers), the particle dynamics becomes stochastic. Labeling parameters of the Cherenkov and cyclotron resonances by indices 1 and 2, one has for the equilibrium values

$$\omega_1 - k_1 v_1 = 0; \qquad \omega_2 - k_2 v_2 = \pm \Omega_0/\gamma_2, \tag{6.64}$$

where $\Omega_0 = qB/mc$ is Larmor frequency in the magnetic field B. Assume for simplicity that the resonances are close to each other. Then the difference of the equilibrium energies is

$$|\Delta\gamma| = \left|\gamma^3 \beta \Delta\beta\right| = \left| \frac{\gamma^3 \beta \left(\beta - \beta_{\rm g}\right)}{1 \pm \gamma \beta^2 \Omega_0/\omega} \, \frac{\Delta k}{k} \right|, \tag{6.65}$$

where $\beta_{\rm g} c = \partial\omega/\partial k$ is the group velocity in the vicinity of the resonances. For semiwidths of the resonances, one can take the expression (6.26). If they are equal to each other, the condition of a stochastic regime takes a simple form:

$$\sqrt{g} > \left| \frac{\sqrt{\alpha}\gamma^3 \beta \left(\beta - \beta_{\rm g}\right)}{4 \left(1 \pm \gamma \beta^2 \Omega_0/\omega\right)} \, \frac{\Delta k}{k} \right|. \tag{6.66}$$

Note that if the group velocity of the wave is close to the phase one, the limiting amplitude can be rather small. For nonequal widths, one should take into account that for a cyclotron resonance the phase slippage coefficient is equal to

$$\alpha = -\beta \frac{\partial}{\partial\gamma} \frac{\omega \mp \Omega_0/\gamma}{\omega\beta} = \frac{1 \mp \Omega_0 \gamma/\omega}{\gamma^3 \beta^2}. \tag{6.67}$$

In more detail, this effect will be considered in use to cyclotron resonance masers (Sect.9) and to free electron lasers (Sect.10).