RESEARCH CONCERNING THE THEORY OF
NON-LINEAR RESONANCE AND STOCHASTICITY

B.V. Chirikov

Novosibirsk, 1969

Translated at CERN by A.T. Sanders
(Original: Russian)

(CERN Trans. 71-40)

Geneva
October 1971

RESEARCH CONCERNING THE THEORY OF
NON-LINEAR RESONANCE AND STOCHASTICITY

B.V. Chirikov

Novosibirsk, 1969

CONTENTS

INTRODUCTION

The proposed monograph is devoted mainly to many-dimensional non-linear oscillations of a conservative mechanical system studied in a complete way, i.e. in an unlimited period of time and for arbitrary initial conditions. This problem, of which the famous three-body problem in astronomy is a particular example, is probably the most complex and at the same time the most beautiful in classical mechanics. The point is that in the case of finite motion (which is equivalent to oscillations in the broad sense of the term) in the absence of damping, repeated interactions occur in the system, so that very subtle cumulative effects become important (Section 2.12). The complete solution of this problem is still a long way off. Nevertheless, at the present time, particularly as a result of the numerous papers of the last 10-15 years, the general picture of the motion of such a system is already beginning to emerge more and more clearly through the thick fog of innumerable details and the particularity of specific problems.

There are two important reasons for constructing a general theory of non-linear oscillations. On the one hand, in specially interesting cases it is not always possible to remain in the linear oscillation region, i.e. to keep within sufficiently small amplitudes. On the other, the linear region is too narrow and therefore relatively poor in phenomena. Of course, it is difficult to guarantee that qualitatively new processes will not be discovered in this region, particularly if it is remembered that quite recently such interesting and important phenomena as the Kapitsa pendulum[172] or the strong focusing of particles in an accelerator[173] were discovered in this region. Nevertheless, it seems safe to assert that the linear oscillation region has been exhausted to a large extent and for subsequent significant progress, both in understanding and applying oscillatory processes, we shall have to switch to the non-linear region. An attempt to limit investigations to linear oscillations is often very artificial, unduly reduces the possibilities of practical application and resembles the notorious attempt to restrict the search to the area directly in the spotlight. This latter method is certainly a good idea, since the beautifully worked out comprehensive theory of linear oscillations is in sharp contrast to the disconnected descriptions of separate non-linear processes. However, it is becoming increasingly difficult to find anything new "in the spotlight" and the development of a theory of non-linear oscillations can be considered as an attempt to light, albeit a little, the general mass of streets of a large town in addition to the brightly-lit main avenue.

At present, there are two main approaches to the problem. The first is connected with the search for stable periodic or almost periodic motion. This is related to the classical theory of non-linear oscillations (Poincaré, Lyapunov, Mandelstam and others), the basic disadvantage of which -- that the cases of motion considered are too special -- was overcome recently in the famous works of Kolmogorov, Arnold and Moser (KAM theory, Section 2.2). Another approach, the ergodic theory, deals on the contrary with the case of extremely unstable motion, leading up to a statistical description (Birkhoff, Hopf, Anosov, Sinai and others, Sections 2.1, 2.3 and 2.4). Both approaches, in particular recently, have given a series of brilliant results which form a reliable basis for further research in this field. However, on account of the extraordinary mathematical complexity of the problem, they nevertheless remain only special or, rather, limiting cases of motion. It is not even known under what conditions the transition from one approach to the other, i.e. from stable to unstable motion, takes place.

In these circumstances it appears advisable to reject the purely deductive method compulsory in mathematics and adopt the semi-empirical method more usual in physics, which in the present instance means a system of models, analytical estimates and experiments, numerical or "real" (Section 3.1). To a certain extent the Mandelstam school carried out such research with the aim of combining theory and experiment as applied to the special problems of non-linear oscillations. A similar approach to the general problem outlined above was started by Krylov[30], many of whose ideas are used and developed in this paper. The main difference in our approach is that we are interested not so much in the macroscopic molecular systems of statistical physics, the nature of whose motion has in any event been correctly established, as in systems with a few degrees of freedom, where this problem is far from trivial and is not of merely theoretical interest. Bearing in mind the given approach, we shall speak of constructive physics, since the main task here is to construct an approximate system of notions and laws in a region where, in principle (but not in practice!), the exact laws are known. It should be noted that, at present, constructive physics, besides being related to oscillation theory, is connected with such large branches of science as, for instance, statistical physics and chemistry, and in the not too distant future probably also biology. It should be stressed that the centre of gravity of constructive oscillation theory (and this also applies to a certain extent to other regions of constructive physics) does not lie in formulating any new laws of nature, but in applying well-known and firmly established laws of mechanics to the explanation (analysis) and construction (synthesis) of new mechanical systems and processes with the desired characteristics (Section 3.1).

The basis of our analysis of non-linear oscillations is the notion of non-linear resonance (Chapter 1), which first arose apparently in celestial mechanics in connection with the librational motion of the planets (Lagrange) and in a clearer form in accelerator theory in connection with the phase stability mechanism (Veksler, McMillan). The most significant and, as far as we know, new process proves to be the interaction of several resonances, always taking place in a non-linear system.

A large part of the paper (Chapters 2 and 3) is devoted to the study of this interaction.

A system of models was constructed (see diagram on next page) beginning with a one-dimensional non-linear oscillator. The downward-pointing arrows show the simplification of the model down to the elementary one, which is studied in detail analytically (Chapter 2) and by means of numerical experiments (Chapter 3). The results obtained are applied to a series of increasingly complicated models, finishing with a many-dimensional non-linear oscillator (upward-pointing arrows). For the analytical investigation wide use is made of the Krylov-Bogolyubov-Mitropol'sky asymptotic averaging method [KBM theory[3]] on the basis of Hamiltonian formalism. We were naturally obliged to limit ourselves to the case of small (or slow) perturbation (parameter $\varepsilon \ll 1$), assuming that the motion of the unperturbed system is known in one form or another. Since, however, the basic results of the work are estimates in order of magnitude, their range of application can be extended to $\varepsilon \sim 1$.

Let us note here two of the results obtained, in our opinion the most interesting. Firstly, a study was made of stochastic instability, which from a practical point of view is the most dangerous instability of non-linear oscillations (Section 2.5) (and at the same time a peculiar method of particle acceleration, Section 4.1), but from a theoretical point
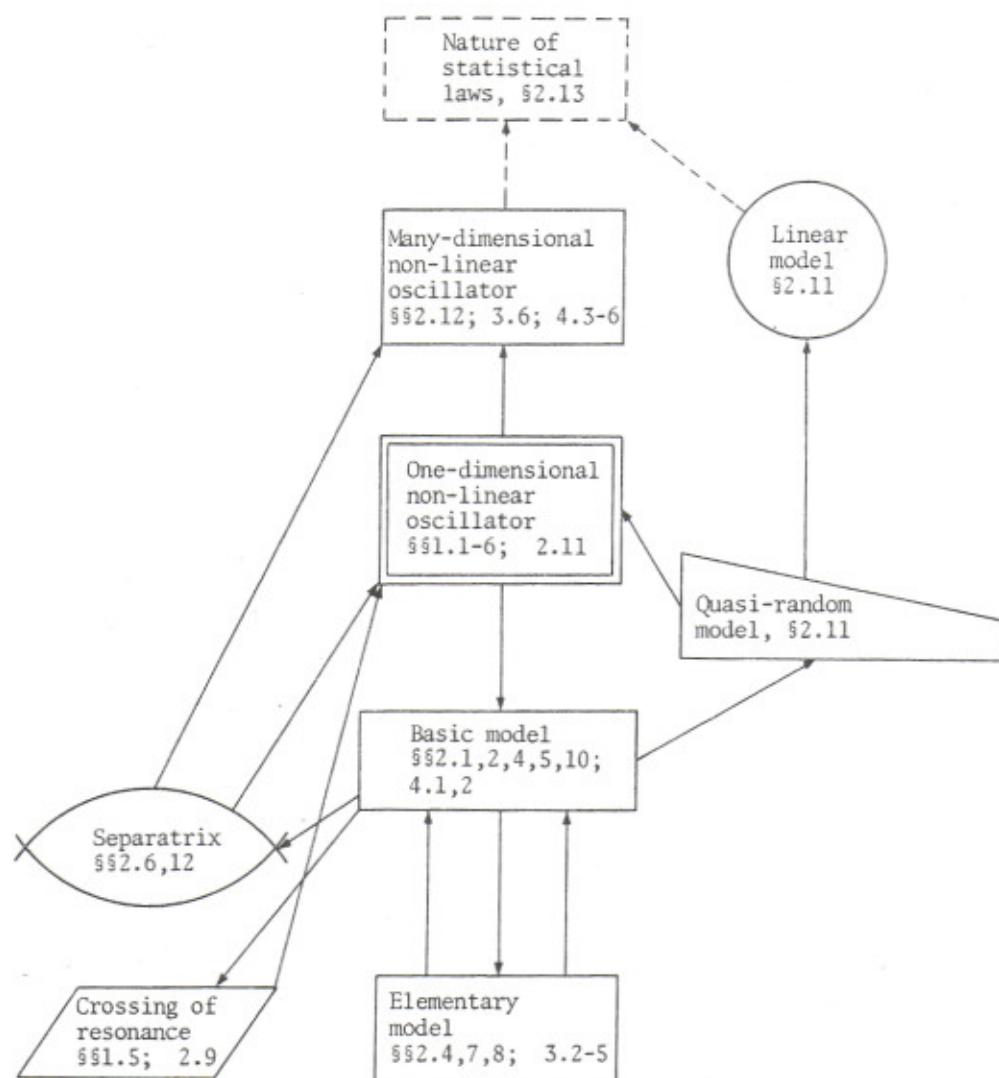
Diagram of models of the interaction of the resonances of non-linear oscillations.

of view gives a model of statistical laws applied, as distinct from the model of contemporary statistical mechanics, to a system with a few degrees of freedom $N \geq 2$ (Section 2.13)[*]. Secondly, a study was made of <u>Arnold diffusion</u>, which proved to be a peculiar universal instability of non-linear oscillations in cases where there was no stochastic instability (Section 2.12).

Furthermore, the studies made seem to us to give a rather detailed general picture of many-dimensional non-linear oscillations, and particularly the rather complicated structure of their phase space. With the above-mentioned limitation on the perturbation strength, the transition from the Kolmogorov region of maximum stability to the region of maximum instability of the ergodic theory can be traced, and it can thus be shown that in the general case both regions interpenetrate deeply in a rather complicated way, forming a system with <u>divided phase space</u>. The latter fact is also an important obstacle to the construction of a rigorous mathematical theory.

In spite of some indistinctness in this picture and some doubt about certain of the details, giving rise to natural dissatisfaction, it can nevertheless serve as a guide line for future research and current applications in this unexplored region. The work can, therefore, be looked upon as a kind of reconnaissance in depth (although perhaps including some superficial observations), intended to facilitate subsequent more accurate investigations.

*   *   *

---

[*] Apparently the first observation of stochastic instability was made by Goward and Hine[114].

LIST OF THE MAIN DIMENSIONLESS PARAMETERS AND SPECIAL SYMBOLS

1. ϵ : small perturbation parameter (Section 1.1).

2. $\alpha \sim |I/\omega \cdot d\omega/dI|$: non-linearity parameter (Section 1.3).

3. $s \sim (\Delta\omega)_H/\Delta$: parameter of the overlapping of resonances;
   $(\Delta\omega)_H$: non-linear width of resonance;
   Δ : distance between resonances (Section 2.1).

4. $s_1 \sim \Omega_\phi/\omega_1 (\sim s)$: parameter of destruction of non-linear resonance separatrix
   (Section 2.6);
   Ω_ϕ : frequency of phase oscillations (Section 1.4);
   ω_1 : perturbation frequency (Section 2.6).

5. s' : stochasticity parameter (s, s_1) for re-normalized resonance
   (Section 2.6).

6. $\mu \sim e^{-c/s}$: exponentially small parameter of destruction of non-linear resonance
   separatrix;
   $c \sim 1$: constant (Section 2.6).

7. $\delta \sim s\mu$: fraction of stochastic component in the region of Kolmogorov sta-
   bility (Section 2.6).

8. N : number of degrees of freedom (Section 2.12).

9. m : multiplicity of interaction (Section 2.12).

10. $\sim$: sign of equivalence in order of magnitude (with correct dimension-
    ality).

*)11. $\propto$: sign of proportionality (dimensionality not maintained).

Note : The above symbols are valid throughout the text, with the exception
       of special cases in which changes in the symbols are specifically
       mentioned.

---

*) Typist's Note: In the handwritten formulae the sign $\propto$ has been used.

).

CHAPTER 1

NON-LINEAR RESONANCE

This short chapter is an introduction. It sets out the basic ideas connected with a single resonance of non-linear oscillations, or, let us say, <u>non-linear resonance</u>. Although, as we shall see later, the difference between resonant and non-resonant motion is not as great for a non-linear system as for a linear one, the main features of the motion are nevertheless determined by the non-linear resonance, which is an "elementary" non-linear oscillation process.

1.1 Formulation of the problem

Let us begin our investigation with a one-dimensional non-linear oscillator, subject to various perturbations. Let us assume that the Hamiltonian of the system is:

$$\mathcal{H} = \mathcal{H}_0(p,q,\lambda) + \varepsilon \mathcal{H}_1(p,q,\Lambda,\vartheta,\varepsilon)$$
$$\lambda = \lambda(\tau); \quad \Lambda = \Lambda(\tau,p,q,\dot{p},\dot{q},\cdots); \quad \frac{d\vartheta}{dt} = \Omega(\tau) \tag{1.1.1}$$

Here $\tau = \varepsilon t$ is the "slow" time, and the parameters define: λ as the adiabatic processes, ϑ as the resonant processes, including those with variable frequency, Λ as the perturbation, depending on the dynamical variables p, q and their derivatives; H_0 is the unperturbed Hamiltonian; εH_1 is the small perturbation ($\varepsilon \ll 1$).

Let us explain the idea of introducing the parameter Λ by the following example. Supposing we want to consider the frictional force -- kq̇. The direct introduction into the Hamiltonian of the term kqp/m "spoils" the second equation: $\dot{q} \neq (\partial H/\partial p) = (p/m) + (kq/m)$. But if we do the same thing through the parameter: $q\Lambda(p)$ where $\Lambda(p) = kp/m$ the equations remain canonical, since differentiation with respect to p, q is carried out with a constant Λ. The dependence of Λ on p should be understood in this case as an explicit dependence on time, so that the Hamiltonian is not conserved. This simple method of taking into account unusual perturbations in the frame of Hamiltonian formalism is equivalent, essentially, to using the generalized Hamilton principle for obtaining Lagrange equations[1] [*]. A similar problem was studied by Volosov[2].

In spite of the apparent limitation of the problem, the Hamiltonian of the form of (1.1.1) covers a fairly wide range of non-linear oscillatory processes, mainly on account of the diversity of the perturbations. In a sense system (1.1.1) may be called an "elementary" non-linear oscillator, which enables us to introduce, investigate and "sound" the basic ideas and regularities of this region. In particular, some many-dimensional problems (see Section 4.5) can be reduced to the form of (1.1.1).

---

[*] However, it should be borne in mind that the said method should be used with caution. Thus, for instance, frictional forces change the phase space volume of the system (violation of Liouville's theorem), while in the case of "real" explicit dependence of the Hamiltonian on time the phase space volume is conserved.

Let us suppose, for instance, that there is a many-dimensional system, which in a zero approximation ($\varepsilon = 0$) splits into independent one-dimensional oscillators. The perturbed Hamiltonian of such a system depends, generally speaking, on variables of all degrees of freedom. However, by calculating these variables in a zero approximation as explicit time functions and substituting them into the perturbation, the system can again be divided (in a first approximation) into separate oscillators of the form of (1.1.1), whose dependence on variables of other degrees of freedom is replaced by an explicit dependence on time. It should nevertheless be stressed that the one-dimensionality of the original model (1.1.1) may sometimes lead to qualitative anomalies (see Section 2.12).

We consider the parameter ε as fairly small, i.e. the perturbation is weak (or slow). This assumption turns out to be correct in a series of cases and is due to the practical need to use a kind of perturbation theory for analytical investigation. Under the condition of small perturbation, resonance, i.e. cumulative perturbation, is the most significant process for the oscillatory system. Thus our problem can be defined as the study of non-linear resonance in a one-dimensional system of the form of (1.1.1).

1.2 Transformation to slow variables

Since the perturbation is small, it is advisable to choose dynamical variables in which the smallness will be expressed explicitly. In other words, it is useful to exclude the "fast" unperturbed motion from the equations. Let the solution of the unperturbed equations take the form:

$$q = q^0(I, \theta, \lambda); \quad \theta = \int \omega(I, \lambda)\, dt + \varphi$$
$$\dot{q} = \omega\, \partial q^0 / \partial \theta; \quad I = \frac{1}{2\pi} \oint p\, dq \qquad (1.2.1)$$

where $2\pi/\omega$ is the period of the motion and I is the action canonically conjugated to the angular variable θ. Although the frequency of the unperturbed motion is constant, it is placed under the integral in order to preserve the functional form of the solution also for the perturbed motion. In this case the constants of the unperturbed motion (I, ϕ) will vary with time, but slowly. We shall choose them as new variables.

In the variables I, θ the Hamiltonian (1.1.1) takes the form:

$$\mathcal{H} = \mathcal{H}_0 + \lambda\, \widetilde{\mathcal{H}}(I, \theta, \lambda) + \varepsilon \mathcal{H}_1(I, \theta, \Lambda, \vartheta, \varepsilon) \qquad (1.2.2)$$

where $\lambda \widetilde{H}$ is an additional term to the unperturbed Hamiltonian because of its explicit dependence on time. In order to find $\widetilde{H}$ we will write the total derivative of I:

$$\dot{I} = \lambda \frac{\partial I(p, q, \lambda)}{\partial \lambda} + \left(\frac{\partial I}{\partial q} \frac{\partial}{\partial p} - \frac{\partial I}{\partial p} \frac{\partial}{\partial q} \right) \left(\mathcal{H}_0 + \varepsilon \mathcal{H}_1 \right) \qquad (1.2.3)$$

and take into account that the operator in brackets depends only on the function $I(p,q,\lambda)$ and thus is equal to $\partial/\partial\theta$. But from (1.2.2) $\dot{I} = -\partial H/\partial\theta$; equating with (1.2.3) we find:

$$\widetilde{\mathcal{H}} = -\int d\theta \left(\frac{\partial I}{\partial\lambda}\right)_{p,q} = \int dI \left(\frac{\partial\theta}{\partial\lambda}\right)_{p,q} \qquad (1.2.4)$$

The latter expression is obtained if a similar procedure is carried out with the function $\theta(p,q,\lambda)$ [*]. When calculating the integral it is necessary to express p,q through I,θ in accordance with (1.2.1).

In slow variables (I,φ) the equations take the form:

$$\dot{I} = -\dot{\lambda}\frac{\partial\widetilde{\mathcal{H}}}{\partial\theta} - \varepsilon\frac{\partial\mathcal{H}_1}{\partial\theta}$$
$$\dot{\varphi} = \dot{\lambda}\frac{\partial\widetilde{\mathcal{H}}}{\partial I} + \varepsilon\frac{\partial\mathcal{H}_1}{\partial I} \qquad (1.2.5)$$

Since the differentiation with respect to both θ,φ is equivalent, system (1.2.5) is canonical.

Let us transform $\partial\widetilde{H}/\partial\theta = (\partial I/\partial\lambda)_{p,q}$ using the relation:

$$\frac{\partial I(W,\lambda)}{\partial W} = \frac{1}{\omega} \; ; \qquad \frac{\partial I(W,\lambda)}{\partial\lambda} = -\frac{1}{\omega}\cdot\overline{\frac{\partial\mathcal{H}_0}{\partial\lambda}} \qquad (1.2.6)$$

where $W = H_0(p,q,\lambda)$, and the bar signifies averaging over the unperturbed motion with constant λ. We have:

$$\frac{\partial I(p,q,\lambda)}{\partial\lambda} = \frac{\partial I(W,\lambda)}{\partial\lambda} + \frac{\partial I(W,\lambda)}{\partial W}\cdot\frac{\partial W(p,q,\lambda)}{\partial\lambda} = \qquad (1.2.7)$$
$$= \frac{1}{\omega}\left(\frac{\partial\mathcal{H}_0}{\partial\lambda} - \overline{\frac{\partial\mathcal{H}_0}{\partial\lambda}}\right)$$

Whence

$$\dot{I} = \frac{\dot{\lambda}}{\omega}\left(\frac{\partial\mathcal{H}_0}{\partial\lambda} - \overline{\frac{\partial\mathcal{H}_0}{\partial\lambda}}\right) - \varepsilon\frac{\partial\mathcal{H}_1}{\partial\theta}. \qquad (1.2.8)$$

This equation clearly shows the adiabatic invariance of the action and is very convenient for constructing various approximate expressions. A similar but approximate equation was

---

[*] (1.2.4) gives the interesting identity:

$$\left(\frac{\partial}{\partial I}\left(\frac{\partial I}{\partial\lambda}\right)_{p,q}\right)_{\theta,\lambda} + \left(\frac{\partial}{\partial\theta}\left(\frac{\partial\theta}{\partial\lambda}\right)_{p,q}\right)_{I,\lambda} = 0 \; ..$$

obtained by Volosov[2]). It should be pointed out in this connection that all the equations in this paragraph are exact.

Let us mention without proof another form of equation for $\dot{\varphi}$ [6]:

$$\dot{\varphi} = -\dot{\lambda}\left(\frac{\partial q^o}{\partial w} \cdot \frac{\partial \mathcal{H}_o}{\partial \lambda} + \frac{\partial q^o}{\partial \lambda}\right) \Big/ \frac{\partial q^o}{\partial \theta} \qquad (1.2.9)$$

At first glance the disagreeable feature of this expression is that the velocity in the denominator may vanish. In practice, however, this fact can be used to check the correctness of the expression for $\dot{\varphi}$, since the numerator, of course, may also vanish along with the denominator.

For solving specific problems one can use any pair of the equally valid equations (1.2.5), (1.2.8) and (1.2.9).

Sometimes it is convenient to use the energy of the unperturbed system W instead of the action. Calculating the total derivative in the same way as in (1.2.3) and using (1.2.1), we find:

$$\dot{W} = \dot{\lambda}\,\frac{\partial \mathcal{H}_o(p,q,\lambda)}{\partial \lambda} - \varepsilon\omega\,\frac{\partial \mathcal{H}_1}{\partial \theta} \qquad (1.2.10)$$

It should be borne in mind, however, that this equation is not canonically conjugated to the equation for $\dot{\varphi}$.

1.3 Single resonance

When the perturbation is small, the most important process for the oscillator is resonance. Resonance generally takes place for a number of values of the oscillator frequency $\omega = \omega_i$. In this chapter we shall consider that the ω_i are rather far apart, so that near one resonance the influence of the others can be completely neglected. Such single resonance is a kind of "elementary" process for a non-autonomous oscillator. The interaction of several resonances will be thoroughly examined in the next chapter.

The time dependence of the unperturbed Hamiltonian is assumed to be slow (but not necessarily small): $\dot{\lambda} \sim \varepsilon$ (1.1.1).

Let us re-specify:

$$\dot{\lambda}\,\tilde{\mathcal{H}} + \varepsilon\mathcal{H}_1 \rightarrow \varepsilon\mathcal{H}_1$$

and use the parameter Λ to describe the losses in the system (for instance, frictional forces). The equations of motion take the form:

$$\dot{I} = -\varepsilon \frac{\partial \mathcal{H}_1(I, \theta, \vartheta)}{\partial \theta} - \varepsilon \Lambda(I, \theta)$$

$$\dot{\varphi} = \varepsilon \frac{\partial \mathcal{H}_1(I, \theta, \vartheta)}{\partial I} \tag{1.3.1}$$

In accordance with (1.1.1) and (1.2.1) H_1 is a periodic function of θ, ϑ of period 2π:

$$\mathcal{H}_1(I, \theta, \vartheta) = \sum_{m,n} \mathcal{H}_{1\,mn}(I) \cdot e^{i(m\theta + n\vartheta)} \tag{1.3.2}$$

The resonance condition takes the form:

$$m\omega \approx n\Omega \tag{1.3.3}$$

Here m,n are any positive integers (we assume that ω, $\Omega > 0$); in contrast to this, in (1.3.2) m,n may be both positive or negative.

All the harmonics that are multiples of the basic ones contribute to the resonance:

$$Nk; \quad N\ell; \quad N = 1, 2, \dots; \quad k/\ell \approx \Omega/\omega$$

Neglecting the non-resonant harmonics in accordance with the averaging method, we obtain from (1.3.1), (1.3.2) the so-called first approximation averaged Hamiltonian[15]:

$$\overline{(\mathcal{H}_1)}_{k\ell} = \sum_{N} \mathcal{H}_{1\,Nk,-N\ell} \cdot e^{iN\psi} - \Lambda(I)\psi = \tag{1.3.4}$$

$$= -\frac{1}{k} U(I, \psi) + \Lambda\psi; \quad \psi = k\theta - \ell\vartheta; \quad \Lambda(I) = \overline{\Lambda}(I, \theta)$$

where U is a periodic function of ψ, of period 2π.

The physical meaning of neglecting the non-resonant harmonics is fully understood; a detailed mathematical proof of the validity of such an approximation and also its accuracy, the limits of its applicability and the construction of the subsequent approximations, form the subject of the Krylov-Bogolyubov-Mitropol'sky theory (KBM theory)[3]. The most important effect of non-resonant harmonics is that new frequencies arise in the system and cause new resonances. For the study of a single resonance this has no significance by definition; as regards the role of higher harmonic resonances for the case of the interaction of several resonances, this question will be discussed in Section 2.7.

From (1.3.4) we obtain the first approximation equations ($\sim \varepsilon$) of the averaging method in the form:

$$\dot{I} = - \varepsilon \frac{\partial U}{\partial \psi} - \varepsilon \Lambda (I)$$
$$\dot{\psi} = k \omega (I) - \ell \Omega + \varepsilon \frac{\partial U}{\partial I} \qquad (1.3.5)$$

Let us recall that the dependence $\Lambda(I)$ is regarded as an explicit dependence on time and therefore is not differentiated when obtaining the equation for $\dot{\psi}$. The parameter Λ is connected with a more usual quantity -- the loss rate (εP) -- by the relation:

$$\Lambda = \frac{P}{\omega} \qquad (1.3.6)$$

The system of equations (1.3.5) is canonical with the <u>resonant Hamiltonian</u>:

$$\mathcal{H}_p = \int (k\omega - \ell\Omega) dI + \varepsilon U(I, \psi) + \varepsilon \Lambda \psi \qquad (1.3.7)$$

For a constant Λ ($d\Lambda/dI = 0$) H_p is the integral of motion and if it can be calculated in explicit form it enables us fully to investigate the behaviour of the oscillator near the resonance. This method is widely used (see for instance Refs. 4 and 5) and is specially suitable (and necessary) when the non-linearity is small ($\alpha \ll 1$). Usually just the case of small non-linearity is studied, often in the hope of simplifying the equations. However, things turn out to be just the opposite[6]. In the case of strong (but not very strong, <u>moderate</u> as we shall call it in what follows) non-linearity

$$\varepsilon \ll \alpha \ll 1/\varepsilon \qquad (1.3.8)$$

the Hamiltonian (1.3.7) is substantially simplified, since the variation of I in this case proves to be always small. Therefore one can neglect the dependence of U on I, having put

$$U(I, \psi) \simeq U(I_\rho, \psi) \to U(\psi) \qquad (1.3.9)$$

and take into account the dependence $\omega(I)$ only in first approximation:

$$k\omega - \ell\Omega \simeq \omega'_k \cdot (I - I_\rho); \quad \omega'_k = k \cdot \frac{\partial \omega(I, \lambda)}{\partial I} \qquad (1.3.10)$$
$$k\omega(I_\rho) = \ell\Omega$$

where $\omega'_k(I_p)$ is the constant characterizing the non-linearity of the oscillator.

In the approximation considered, the conditions of application of which will be discussed in the next paragraph, the resonant Hamiltonian (1.3.7) takes the form:

$$\mathcal{H}_y = \omega'_k \cdot \frac{(I - I_\rho)^2}{2} + \varepsilon U(\psi) \qquad (1.3.11)$$

and the equations of motion:

$$\dot{I} = - \varepsilon \frac{dU}{d\psi}$$
$$\dot{\psi} = \omega'_k \cdot (I - I_\rho)$$

(1.3.12)

Generally speaking, one could also take the losses into account in this same approximation by adding the term $\varepsilon \Lambda(I_p)\psi$ to the Hamiltonian (1.3.11); however, it is more convenient to do this later (Sections 1.5, 1.6).

System (1.3.12) can be reduced to a so-called phase equation[5] after eliminating I:

$$\ddot{\psi} + \varepsilon \omega'_k \cdot \frac{dU}{d\psi} = 0$$

(1.3.13)

The Hamiltonian (1.3.11) describes the oscillations of a certain "particle" with a mass $1/\omega'_k$ in a periodic potential field $\varepsilon U(\psi)$. Thus for moderate non-linearity (1.3.8) the behaviour of the oscillator near the resonance proves in the first approximation to be universal (except for the shape of the "potential well" and consequently the shape of the oscillations). It should be remembered that with weak non-linearity ($\alpha \lesssim \varepsilon$), the behaviour of the system varies qualitatively according to the type of resonance (external, parametric, etc.)[4,14,5].

Since the shape of the oscillations, generally speaking, is not important when studying the general laws of non-linear resonance, it will be specified from time to time in order not to complicate the writing of the formulae unnecessarily. Let us put:

$$U(I, \psi) = U_o(I) \cdot \sin \psi$$

(1.3.14)

Then the original system (1.3.5) takes the form:

$$\dot{I} = - \varepsilon U_o \cos \psi - \varepsilon \Lambda$$
$$\dot{\psi} = k\omega - \ell\Omega - \varepsilon U'_o \sin \psi$$

(1.3.15)

and the universal Hamiltonian becomes:

$$\mathcal{H}_y = \frac{\omega'_k}{2} (I - I_\rho)^2 + \varepsilon U_o \sin \psi$$

(1.3.16)

We studied the periodic dependence of the perturbation on the phase θ. Extension to the case of quasi-periodic perturbation presents no difficulty, but neither does it lead to any new effects. A periodic transient (acting in a finite interval of time) perturbation, is not of much interest from the point of view of resonant processes. There is also steady aperiodic (with a continuous spectrum) perturbation, which leads to a completely different pattern of motion. This case will be discussed later (Section 2.11).

1.4 Phase oscillations

The analogy mentioned in the previous section with the motion of a "particle" in a periodic potential enables us to have a visual picture of non-linear resonance for moderate non-linearity. Let us limit ourselves to the case of the harmonic potential (1.3.14).

System (1.3.12) has two equilibrium states, $I = I_p$, $\psi = \pm\pi/2$, one of which is unstable [depending on the sign U_0, ω'_k see (1.3.13)]. The pattern of the phase plane is periodic in ψ and has a characteristic "bucket" appearance (Fig. 1.4.1). The phase trajectories are
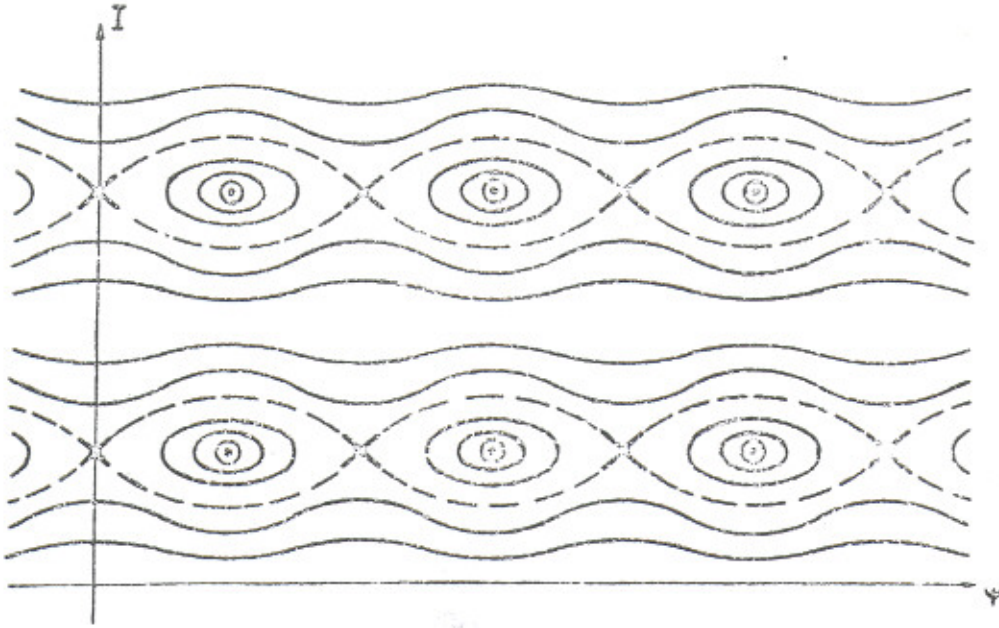
Fig. 1.4.1: Phase trajectories in the vicinity of resonances for moderate non-linearity: ⊙ - stable, or elliptic, points; × - unstable, or hyperbolic points. The dotted lines show the first approximation separatrices; in the subsequent approximations they are destroyed and stochastic layers are formed in their place (Section 2.6).

determined from the condition H_y = const. When $|H_y| < |\varepsilon U_0|$ (inside the "potential well") the phase trajectories are closed, i.e. the phase [and energy [*)]] of the oscillator varies within restricted limits. These oscillations are generally called phase oscillations. This name is fully justified, since the behaviour of the oscillator near the resonance is determined by its phase conditions, namely phase shift law. The frequency of small phase oscillations is equal to (1.3.13):

---

*) From time to time we shall speak of the energy of the oscillator, which depends on the action variable monotonically $dW/dI = \omega > 0$. This is shorter and more usual.

$$\Omega_\varphi^2 = \varepsilon U_o \omega'_k \qquad (1.4.1)$$

When $|H_y| > |\varepsilon U_0|$ (outside the "potential well") the phase shifts to an unlimited extent and the energy oscillations decrease in proportion to their distance from the resonance (to the increase of $|H_y|$). The equation for the separatrix (the upper edge of the "potential well") takes the form: $|H_y| = |\varepsilon K_0|$ or:

$$\left(I - I_\rho\right)^2 = \frac{2\varepsilon U_o}{\omega'_k} \left(\sin \psi \pm 1\right) \qquad (1.4.2)$$

where the sign in brackets coincides with the sign $U_0 \omega'_k$.

The physical meaning of phase oscillations is that the non-linear oscillator deviates from the exact resonance ($k\omega = \ell\Omega$) as a result of the variation of its frequency $\omega(I)$. Alternatively it can be said that the non-linearity stabilizes the resonance, since the unlimited increase of the energy in the case of a linear resonance is replaced by the restricted oscillations. Thus moderate non-linearity always stabilizes the resonance.

The region inside the separatrix is generally called the capture or phase stability region. This means that although the oscillator deviates from the exact resonance as a result of non-linearity, it does not deviate much. Moreover, if, say, the frequency of the external force varies slowly, the energy of the oscillator also varies so that the approximate equality $k\omega \approx \ell\Omega$ is fulfilled all the time.

The size of the capture region is characterized by the width of the separatrix in the direction of I (Fig. 1.4.1):

$$(\Delta I)_H = 4\sqrt{\left|\frac{\varepsilon U_o}{\omega'_k}\right|} \ ; \ k(\Delta\omega)_H = 4\sqrt{|\varepsilon U_o \omega'_k|} = 4\Omega_\varphi \qquad (1.4.3)$$

· These relations determine the non-linear width of the resonance.

From the above-mentioned analysis of the resonance it can be seen that the essential characteristic of a non-linear oscillator is the derivative ω', i.e. the dependence of the frequency on I (or the energy). In what follows, therefore, the term "non-linear oscillator" will be equivalent to the term "oscillator whose frequency depends on the energy" or "non-isochronous oscillator". The oscillations may be of any shape and generally speaking their shape has nothing to do with the non-linearity. Thus the rotation of a relativistic particle in a magnetic field is an example of a non-linear but harmonic oscillator, and an ultra-relativistic particle in a square potential well represents an anharmonic oscillator with constant frequency.

The conditions of applicability of the universal Hamiltonian are connected with the requirement for small variation of its parameters $U_0(I)$, $\omega'_k(I)$ and depend on the specific

form of these functions. In order to obtain a general estimate we shall assume that in the typical case: $U_0' \sim U_0/I$ and $\omega'' \sim \alpha\omega'/I$. It is then sufficient to require small variation of the quantities I, ω:

$$\left(\frac{\Delta I}{I}\right)_H = 4\sqrt{\overline{\frac{\varepsilon}{\alpha} \cdot \frac{U_0}{k\omega \hat{I}}}} \ll 1; \quad \left(\frac{\Delta\omega}{\omega}\right)_H = 4\sqrt{\overline{\varepsilon\alpha \cdot \frac{U_0}{k\omega\hat{I}}}} \ll 1 \qquad (1.4.4)$$

Hence the conditions of (1.3.8) are obtained, if the parameter ε is chosen so that $U_0 \sim \omega I$ [*].

Let us point out that for moderate non-linearity the real expansion parameter is not ε but $\sqrt{\varepsilon}$. The universal equations (1.3.12) prove in this case to be of the first order in $\sqrt{\varepsilon}$ and the original equations (1.3.5) of the second. This also explains the possibility of simplifying the original equations.

Let us note that the behaviour of a non-linear system near to a resonance has been re-investigated many times since the days of Poincaré[36]. A simple picture of phase oscillations and phase stability was set out in the classical papers by Veksler[7] and McMillan[8] which had such a great influence on contemporary accelerator technique. Nevertheless it seems to us that so far due attention has not been paid to the universality of the phase oscillation process and the decisive part it plays for the understanding of non-linear phenomena.

1.5 Crossing the resonance

Let us assume that the value I_p explicitly depends on time, and so the difference $I - I_p$, and thus also $k\omega - \ell\Omega$, change sign. This may occur both as a result of the action of perturbation with variable frequency $\Omega(\tau)$ and as a result of the variation of the frequency of the oscillator ω, if the unperturbed Hamiltonian depends on the parameter λ (1.1.1). Unlike other more usual adiabatic processes, in which one can use the conservation of the adiabatic invariant $J = (1/2\pi)\oint I d\psi$, the crossing of the resonance is a more complex process, since here, generally speaking, J changes considerably independently of the rate of crossing (see Section 1.6).

It is convenient to study the crossing of the resonance graphically, by analogy with the motion of the "particle" in a periodic potential, mentioned at the end of Section 1.3 [6]. Let us first find the variation of the total energy of the "particle" (1.3.11):

$$\frac{d\mathcal{H}_y}{dt} = \frac{\partial\mathcal{H}_y}{\partial t} = -\omega'_k \left(I - \hat{I}_p\right)\cdot \dot{I}_p = -\dot{\psi}\,\dot{I}_p \qquad (1.5.1)$$

When the perturbation is small, the width of the resonance is relatively small $[\sim \sqrt{\varepsilon}$, (1.4.3)]; therefore $\dot{I}_p$ can be treated as a constant, and we obtain:

---

[*] In other words, all the dimensionless parameters of the problem except ε, α are of the order of unity.

$$\mathcal{H}_y = - \dot{I}_\rho \, (\psi - \psi_o) \tag{1.5.2}$$

By inserting this expression in (1.3.16) we find[*]:

$$\frac{\dot{\psi}^2}{2\Omega_\phi^2} = V(\psi - \psi_o) + \sin \psi_o - \sin \psi \tag{1.5.3}$$

where the coefficient V characterizes the rate of crossing the resonance:

$$V = \frac{\dot{\Omega}_1}{\Omega_\phi^2} = \frac{k \frac{\partial \omega}{\partial t} - \ell \frac{\partial \Omega}{\partial t}}{\Omega_\phi^2} \tag{1.5.4}$$

and the phase ψ_0 is taken at the moment of exact resonance ($k\omega = \ell\Omega$).

If we now represent graphically the quantity proportional to the potential energy of the "particle": $\sin \psi$, then analysis of the motion is made in the usual way according to its intersection with the line of the total energy, namely with a horizontal line in the steady case ($\dot{I}_p = 0$, Section 1.4), with a slanting line $V(\psi - \psi_0) + \sin \psi_0$ when $\dot{I}_p$ is constant, and with a curve $f(\psi)$ obtained from (1.5.1) in the general case (Fig. 1.5.1).
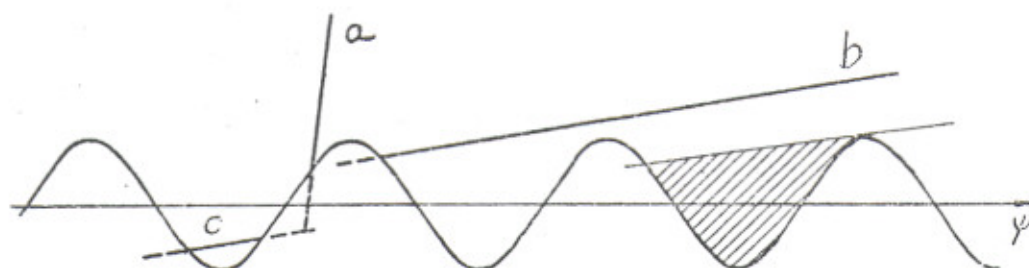
Fig. 1.5.1: Graphical investigation of the crossing of a resonance: a - fast crossing; b - slow crossing; c - phase stability; this region is hatched and limited by a separatrix (thin line).

In the Fig. 1.5.1 it can be seen that there are two qualitatively different regimes as for crossing through the resonance. The first is characterized by the existence of two points of intersection ("particle" stops), by restricted phase oscillations and consequently by repeated crossing of the resonance (line c). This regime has been well studied for a special case (charged particle accelerators) and is generally called capture or phase stability[7,8]. Capture is possible only when $|V| < 1$ and for specific initial conditions shown in the Fig. 1.5.1 by hatching. When $|V| \ll 1$, capture takes place for almost any initial phase of the oscillations (when detuning is sufficiently small). Under capture conditions the energy of the oscillator automatically varies in such a way that $k\omega \approx \ell\Omega$. The accuracy of this equality is determined by the depth of the "potential well" and is of the order of Ω_ϕ (1.4.3).

---

[*] Limiting ourselves to the special case of (1.3.14).

Another regime (lines a, b in Fig. 1.5.1) is characterized by a single crossing of the resonance. When $|V| > 1$ crossing is possible for any phase ψ_0, but for $|V| < 1$ only for some ψ_0. It is the last regime that is a real crossing of the resonance, since when $t \to \pm\infty$ the system deviates a long way from the resonance and its energy approaches constant values.

Let us consider two limiting cases in which the solution (1.5.3) in the last regime can be represented analytically[6]. We will suppose that the values $\dot{\Omega}_1$, ω_k' and U_0 are positive. If $\dot{\Omega}_1 < 0$, it is necessary to change the sign of the time in the solution (the resonance is crossed in the opposite direction) and also to make a phase shift ($\psi_0 \to \pi - \psi_0$) on account of the changing of the sign of V, as is easy to see from Fig. 1.5.1. If $U_0\,\omega_k' < 0$, it is necessary to shift the phase by π ($\psi_0 \to \psi_0 + \pi$) (1.5.3). Finally, if both the values $\dot{\Omega}_1$; $U_0\,\omega_k' < 0$, it is necessary to perform both transformations successively, which is equivalent to changing the sign of the time and to the transformation $\psi_0 \to -\psi_0$.

i) **Fast, or linear, crossing of the resonance (V >> 1)**

In this case non-linearity can be neglected in the first approximation and then the phase equation (1.5.3) or (1.3.12) is at once integrated:

$$\psi = \psi_0 + \dot{\Omega}_1 \frac{t^2}{2} \tag{1.5.5}$$

and the equation for I (1.3.12) comes to the Fresnel integral [see for example Ref. 9)]:

$$\Delta(k\omega) = -\sqrt{\frac{2\pi}{V}} \cdot \Omega_\varphi \cdot \cos\left(\psi_0 + \pi/4\right)$$
$$\Delta I = -\sqrt{2\pi} \cdot \frac{\varepsilon U_0}{\sqrt{\dot{\Omega}_1}} \cdot \cos\left(\psi_0 + \frac{\pi}{4}\right) \tag{1.5.6}$$

Let us give the next term of the expansion in powers of the small parameter V^{-1}, characterizing the weak effect of non-linearity for the fast resonance crossing[6]:

$$\Delta(k\omega) = -\sqrt{\frac{2\pi}{V}} \cdot \Omega_\varphi \cdot \left\{ \cos\varphi + \frac{V^{-1}}{2\sqrt{2}}\left[1 + (\sqrt{2}-1)(\sin 2\varphi + \cos 2\varphi)\right] \right\}$$
$$\psi = \psi_0 + \frac{\dot{\Omega}_1}{2}\left(t \mp \frac{\Delta(k\omega)}{2\dot{\Omega}_1}\right)^2 ; \quad \varphi = \psi_0 + \frac{\pi}{4} \tag{1.5.7}$$

The upper sign corresponds to the motion after the resonance, and the lower to that before the resonance. Since the expression in square brackets > 0, the sign of the non-linear contribution to $\Delta(k\omega)$ is the opposite of the sign $\dot{\Omega}_1$. In other words, the non-linear frequency change for the fast resonance crossing is directed to the opposite side with respect to the external change of frequency, as in the capture; it is as if the non-linearity somewhat slowed down the crossing of the resonance.

The variation of the frequency and energy upon crossing the resonance increases with the reduction of the rate of transit, but in contrast to the purely linear case it is limited by the condition $V \gg 1$ and does not exceed: $|\Delta(k\omega)| \lesssim \Omega_\phi \ll \omega$; $|\Delta I| \lesssim \Omega_\phi/\omega'_K \sim \sqrt{\varepsilon U_0 \, \omega'_K} \ll I$ (1.4.4).

ii) Slow, or reversible, crossing of
   the resonance $(V \ll I)$

It can be seen from Fig. 1.5.1 (line b) that in this case the phase at the moment of exact resonance is enclosed in a narrow interval around $\pi/2$:

$$\psi_0 = \frac{\pi}{2} \div \xi ; \qquad -V < \xi < \sqrt{4\pi V} \qquad (1.5.8)$$

The rest of the phases correspond to capture.

An approximate integration of equation (1.5.3) gives[6]:

$$\Delta(k\omega) = 2V\Omega_\phi \cdot \ell n \left(V+\xi\right)\left(4\pi V - \xi^2\right) +$$
$$+ \frac{8}{\pi}\Omega_\phi \sqrt{1 + 2\pi V - \xi^2/4} \qquad (1.5.9)$$

The first term is important only in an exponentially small region on the edges of the interval (1.5.8), where it leads to unlimited variation of ω (and I). The physical meaning of this variation is connected with the very slow motion (almost a halt) of the phase near the value $(\pi/2) - V$ (1.3.12). The sign of $\Delta(k\omega)$ is the opposite of the sign of $\dot{\Omega}_1$ as in capture, i.e. the crossing of the resonance is slowed down. This result is fully understood, since the edges of the interval (1.5.8) are directly adjacent to the separatrix.

The main term in (1.5.9) is the second. In the limit $V \to 0$ it depends neither on the phase ψ_0 (and consequently also on the initial conditions), nor on the rate of crossing the resonance $\dot{\Omega}_1$:

$$\Delta I = \frac{8}{\pi} \sqrt{\frac{\varepsilon U_0}{\omega'_k}} \qquad (1.5.10)$$

Thus under these conditions there is no continuous transition to the steady case $(\dot{\Omega}_1 = 0)$: this transition takes place only in the capture region.

The sign of $\Delta(k\omega)$ for slow crossing agrees with the sign of $\dot{\Omega}_1$, i.e. non-linearity speeds up the crossing of the resonance. Because (1.5.10) is independent of the phase the slow crossing process is reversible. In particular, when there is periodic crossing of the resonance in both directions, the energy of the oscillator is subjected only to the limited [and small (1.4.4)] oscillations in approximation (1.5.10). A more accurate expression

(1.5.9) already depends on ψ_0 and therefore may lead to certain cumulative effects. This question will be dealt with in more detail in Section 2.9.

Let us note that the uniform change (1.5.10) agrees in order of magnitude with the maximum possible fast crossing of the resonance.

Comparatively little is known about the process of slow crossing of a resonance. Apparently it was first mentioned in a paper by Symon and Sessler[11], where it was called the phase displacement mechanism[*] and was proposed as a method of acceleration in addition to the usual phase stability. A qualitative study of slow crossing of a resonance was made by Sturrock[9], but the criterion of slowness in his paper is incorrect:

$$\frac{(\omega - \Omega)^2}{\dot{\Omega}_1} \gg 1 \qquad (1.5.11)$$

In this form it has no sense at all, since it depends on arbitrary detuning $(\omega - \Omega)$. However, as far as can be understood from the text of Ref. 9 the author takes as the width of the resonance the linear expression: $\omega - \Omega \sim \varepsilon U_0'$ [see (1.6.17)], whereas in order to obtain the correct criterion one should take the non-linear one: $\omega - \Omega \sim \Omega_\phi$ (1.5.4).

Let us now consider the effect of losses. In the first approximation to $\sqrt{\varepsilon}$ it is necessary to add to the universal Hamiltonian (1.3.11) the term $\varepsilon\Lambda\psi$, where $\Lambda = \Lambda(I_p) = \text{const.}$ like the other coefficients. The result can be regarded either as a change of the "potential well" $U(\psi)$ (its "slope"), or as some effective change of the speed of crossing through the previous resonance by the value (1.5.1):

$$\Delta\dot{\Omega}_1 = -\varepsilon\Lambda\omega_k' \qquad (1.5.12)$$

In the latter case the parameter of the rate of "crossing" the resonance takes the form (1.3.6):

$$V = \frac{\dot{\Omega}_1 - \varepsilon\Lambda\omega_k'}{\Omega_\phi^2} = \frac{\dot{\Omega}_1}{\Omega_\phi^2} - \frac{P}{\omega U_0} \qquad (1.5.13)$$

In particular, with constant frequencies ($\dot{\Omega}_1 = 0$) capture is possible only under the condition $P < \omega U_0$. In the capture region the energy of the oscillator on the average does not change, since the losses are compensated for by the action of the perturbation. Outside the separatrix the energy of the oscillator decreases, and it goes away from the resonance.

If $\dot{\Omega}_1 = \varepsilon\Lambda\omega_k'$, an interesting "steady" case ($V = 0$) occurs with variable frequencies. Unlike the true steady case ($\dot{\Omega}_1 = \Lambda = 0$), the amplitude of the phase oscillations may vary (Section 1.6).

---

[*] Displacement in phase space.

1.6 Second approximation effects

The effects of the second approximation (namely $\sim \varepsilon$, Section 1.4) in the non-linear resonance are related to the effects due to the variation of the coefficients of the universal Hamiltonian (Section 1.3): $U_0(I)\omega_k'(I)$, $\Lambda(I)$, which in the first approximation to $\sqrt{\varepsilon}$ were taken to be constant. These effects can be divided into two categories: oscillating (at the frequency of the phase oscillations) and cumulative. According to the estimate of Section 1.4 the oscillating effects in the region of moderate non-linearity are always small ($\sim \sqrt{\varepsilon}$), and we shall not write out the corresponding corrections in explicit form. On the other hand, the cumulative effects can be regarded as slow; the simplest way of studying them is to use the adiabatic invariant of the phase oscillations:

$$J = \frac{1}{2\pi} \oint I \, d\psi = \frac{1}{2\pi} \oint_J (I - I_\rho) \, d\psi \qquad (1.6.1)$$

The latter expression is valid for limited phase oscillations, when $\oint d\omega = 0$. Far away from the resonance $J \to I$, i.e. the adiabatic invariant of the phase oscillations changes over to the adiabatic invariant of the oscillator itself.

In order to calculate the variation of J let us return to the resonance Hamiltonian (1.3.7) and use the general formula (1.2.8). The variable parameters here are the frequencies $\Omega_1(t)$ and the loss parameter $\Lambda(t)$. We have:

$$\frac{dJ}{dt} = \frac{T_\Phi}{2\pi}\left(\frac{\partial \mathcal{H}_\rho}{\partial t} - \overline{\frac{\partial \mathcal{H}_\rho}{\partial t}}\right) + \frac{\varepsilon \Lambda' \cdot (\psi - \overline{\psi}) \cdot \dot{I}}{\Omega_\varphi} \qquad (1.6.2)$$

where T_ϕ is the period of the phase oscillations and the explicit dependence on time is due to the frequency variation $(\dot{\Omega}_1)$.

When there is sufficiently slow and smooth frequency variation the first term, as is known (Section 4.4), makes an exponentially small contribution to ΔJ, i.e. J scarcely varies[*], so that it is sufficient to examine only the second term, connected with losses. For the integration of (1.6.2) let us limit ourselves to small phase oscillations:

$$|\varphi| = |\psi - \overline{\psi}| \leqslant \varphi_0 \ll 1 \qquad (1.6.3)$$

where φ_0 is the amplitude of the phase oscillations. In this case one can put

$$\dot{\varphi} \approx \omega_k' (I - I_\rho); \quad I \approx I_\rho + \frac{\ddot{\varphi}}{\omega_k'} = I_\rho - \frac{\varphi \Omega_\varphi^2}{\omega_k'} \qquad (1.4.6)$$

_____

[*] Provided the trajectory does not cross the separatrix of the steady-state phase oscillations, for which $T_\phi = \infty$, and the adiabatic invariant always changes independently of the rate of transit, as is in fact calculated in Section 1.5.

It is easy to verify that the relative accuracy of these expressions $\sim \varphi_0 \sqrt{\varepsilon}$. In order to obtain J with the same accuracy it is sufficient to use the universal Hamiltonian:

$$J = \frac{\Omega_\phi \, \varphi_{10}^2}{2\omega'_k} \tag{1.6.5}$$

By inserting the expressions (1.6.4) and (1.6.5) in (1.6.2) and averaging over the period of the phase oscillations, we find:

$$\frac{dJ}{dt} \approx - \varepsilon \Lambda' J \tag{1.6.6}$$

or

$$J = J_0 \cdot e^{-\varepsilon \int_0^t dt \cdot \Lambda'(I_p)} \tag{1.6.7}$$

In the general case the parameter $\Lambda'(I_p)$ may depend on time because of the variation of I_p. The direction of the variation of J and consequently also of the amplitude of the phase oscillations (damping or increase) depends on the sign of the derivative $\Lambda' = (P/\omega)'$ (1.3.6).

The application of the averaging method to equation (1.6.2) in order to obtain (1.6.6) is permissible under the condition that:

$$\varepsilon \Lambda' \ll \Omega_\phi \tag{1.6.8}$$

In the steady-state case ($\dot{\Omega}_1 = 0$) the only important effect of the second approximation is the damping (or growth) of the amplitude of the phase oscillations with a constant decrement $-\varepsilon\Lambda'(I_p)$ (1.6.7); other effects lead only to small oscillating corrections $\sim \sqrt{\varepsilon}$.

With sufficiently slow crossing of the resonance under capture conditions, the amplitude of the phase oscillations varies adiabatically according to (1.6.5) and (1.6.7). The expression for the adiabatic invariant of the phase oscillations is universal in the same sense as the Hamiltonian (1.3.11), i.e. it does not depend on the type of resonance (except for the shape of the oscillations).

This result, mentioned for the special case of synchrotron oscillations in an accelerator by Kolomensky and Lebedev[5], is completely natural, since the expression for J can be obtained with an accuracy of $\sim \sqrt{\varepsilon}$ from the universal Hamiltonian. In the case of small phase oscillations expression (1.6.5) is entirely universal.

The independence of the adiabatic processes of phase oscillation on the type of perturbation can be considered from another point of view. If the phase of the perturbation θ depends only on time (1.1.1), the phase plane of the resonance (I,ψ) differs from the original phase plane of the oscillator (I,θ) only by a turning of the co-ordinate axes and by the constant transformation of the scale $[\psi = k\theta - \ell\theta(t)]$. In this case the integral (1.6.1) is

proportional to the phase area of the unperturbed oscillator, spanned by the trajectory of the phase oscillations. According to Liouville's theorem this area (defined by the motion of the ensemble of all the points inside the phase trajectory) is always strictly conserved (also when there is perturbation). This corresponds to the approximate conservation of the area spanned by the steady-state trajectory of the phase oscillations, in those cases in which its intersection with the actual trajectories of neighbouring particles can be neglected. The resonance itself determines only the shape of the region, for example, for small phase oscillations: $\Omega_\phi \cdot \Delta\Theta \sim \omega' \cdot \Delta I$ (1.3.12).

In the special case of a harmonic potential (1.3.14), under the condition that $V \ll 1$ (1.5.13) and in the absence of losses we obtain:

$$\varphi_o \sim \left| \frac{\omega'_k}{U_o} \right|^{1/4} \tag{1.6.9}$$

which agrees with the result of the theory of synchrotron oscillations in accelerators[5]. From the last expression it can be seen that damping of the phase oscillations can be ensured both when the energy of the oscillator increases and when it decreases, owing both to the special non-linear characteristic of the oscillator and to the variation of the parameters of the resonance in time. This gives the possibility of using the non-linear resonance for regulating the amplitude of the oscillations, within the limits compatible with Liouville's theorem.

Let us consider the influence of second approximation effects on slow crossing of the resonance[*]. The most important influence is connected with the possibility of changing over from one regime of crossing to another, i.e. with capture (transition to limited phase oscillations) or, on the contrary, with moving out of the resonance. It is evident that moving out of the resonance will necessarily take place sooner or later, if the amplitude of the phase oscillations increases. In the case of damping of the phase oscillations capture is possible (but does not necessarily take place). The point is that with slow crossing of the resonance there is only one phase oscillation intersecting the steady-state separatrix (see Fig. 1.5.1, line b), and therefore the damping may not have time to change the parameters of the phase oscillations so much that capture takes place. However, when $V \to 0$ capture necessarily occurs, because the aforementioned phase oscillation in this case approaches the separatrix and an arbitrarily small change is sufficient for capture. Moreover, as will be shown in Section 2.6, near the separatrix there is always a more or less wide stochastic layer which facilitates the capture process.

Let us estimate the critical value of the rate of crossing. Let us return to the complete equation for $\dot\psi$ (1.3.15), which we will write in the form:

$$\dot\psi = \omega'_k (I - I_\rho) + \frac{(I - I_\rho)^2}{2} \omega''_k - \varepsilon U''_o \sin\psi \tag{1.6.10}$$

---

*) For fast crossing this influence is always small ($\sim \sqrt{\varepsilon}$).

The first term is the main one, the second always being small for moderate non-linearity; capture may take place owing to the last term, if $\dot{\psi}$ vanishes after crossing the resonance near $\sin \psi = 1$, where $\dot{\psi}$ has a minimum value according to the first approximation (Fig. 1.5.1). For capture it is also necessary that $\dot{\psi} \neq 0$ before the crossing of the resonance; in the opposite case all investigation is transferred to the next phase region (one ψ period to the right, Fig. 1.5.1). This happens to be possible, since the first term (1.6.10) changes sign after the crossing of the resonance, and the last one does not change.

The minimum value of $\dot{\psi}$ in a first approximation is of the order of (1.5.3):

$$\dot{\psi}_{min} \sim \Omega_\varphi \sqrt{4\pi V} \tag{1.6.11}$$

Capture is possible under the condition of $\Omega_\varphi \sqrt{4\pi V} \lesssim |\varepsilon U_0'|$:

$$V \lesssim \varepsilon \frac{(U_0')^2}{4\pi U_0 \omega_k'} \sim \frac{\varepsilon}{\alpha} \ ; \quad \dot{\Omega}_1, U_0' > 0 \tag{1.6.12}$$

The last inequality is the condition for the signs of the terms of (1.6.10).

For stable capture it is necessary for the amplitude of the phase oscillations to decrease after capture; in the opposite case only short-term capture is possible. In the absence of an explicit dependence on time and the condition $U_0 > 0$ the oscillations die down, if[*]:

$$\left(\frac{\omega_k'}{U_0} \right)' \cdot \dot{\Omega}_1 > 0 \tag{1.6.13}$$

This is compatible with the capture condition (1.6.12) when

$$\omega_k' < \frac{\omega_k''}{U_0'} \cdot U_0 \tag{1.6.14}$$

In the opposite case stable capture, as a rule, is not possible except for an exponentially small region of resonant phases on the edges of the interval (1.5.8), for which inequality (1.6.12) changes sign (Section 1.5).

Capture is also possible owing to the non-uniform rate of crossing of the resonance ($\ddot{\Omega}_1 \neq 0$), if this leads to the reduction of V by the value $\Delta V \sim V$ (1.5.3), namely under the condition:

---

[*] In approximation (1.6.9), which we use as an example.

$$- \frac{\ddot{\Omega}_1}{\dot{\Omega}_1 \Omega_\phi} \gtrsim 1 \tag{1.6.15}$$

However, for this capture to be stable the damping of the phase oscillations must be sufficiently fast. In fact, under condition (1.6.15) V passes through zero in a time of the order of one phase oscillation and begins to grow again in absolute value, which may lead to motion out of the resonance.

All the estimates of the second approximation effects in this paragraph were made for moderate non-linearity (1.3.8). When there is large non-linearity $\varepsilon\alpha \geq 1$ it is necessary to take into account the subsequent expansion terms of the quantity $(k\omega - \ell\Omega)$ in the equation for $\dot{\psi}$ (1.6.10). In particular, the relation of the second term to the first is of the order of $\sqrt{\varepsilon\alpha} \cdot \varphi_0$. Hence it can be seen that for sufficiently small oscillations

$$\varphi_0 \ll (\varepsilon\alpha)^{-1/2} \tag{1.6.16}$$

all remains as usual. However, the shape of the large oscillations ($\phi_0 \sim 1$), and also the position of the separatrix, may change substantially depending on the specific form of $U(I,\psi)$.

When non-linearity decreases ($\omega'_k \to 0$) we finally arrive at a linear resonance. In this case the difference ($k\omega - \ell\Omega$) in the system of equations (1.3.15) is simply constant detuning. The resonance corresponds to the condition $\dot{\psi} = 0$, whence the <u>linear width of the resonance</u> (width of unstable region) is:

$$k \cdot (\Delta\omega)_1 = 2\varepsilon U'_0 \tag{1.6.17}$$

The linear approximation is valid as long as the non-linear frequency variation ($\omega'_k \cdot \Delta I$) is much smaller than the linear width of the resonance (1.6.17). In particular, for $\Delta I \sim I$ we obtain:

$$\frac{\varepsilon U'_0}{I \omega'_k} \sim \frac{\varepsilon}{\alpha} \gg 1 \tag{1.6.18}$$

In the intermediate case of $\varepsilon \sim \alpha$ the motion of the oscillator may be very complex and depends on the type of resonance. The most important feature of this region is the formation under certain conditions of a capture region, or, in other words, stabilization of the resonance by non-linearity. The conditions of such stabilization are usually obtained from the resonance Hamiltonian (1.3.7). An estimate of the order of magnitude can, however, be obtained much more simply from the following considerations. Stabilization occurs in the case when the non-linear frequency variation exceeds the linear width of the resonance and the oscillator thus begins to move out of resonance. On the other hand, the non-linear detuning can be estimated according to the phase oscillation formula (Section 1.4):

$$(k\omega - \ell\Omega)_H \sim \Omega_\phi \sim (\varepsilon U_0 \omega'_k)^{1/2} \qquad (1.6.19)$$

Hence we obtain the stabilization condition in the form:

$$\omega'_k \gtrsim \varepsilon \frac{(U'_0)^2}{U_0} \qquad (1.6.20)$$

As an example, let us consider the resonance for small slightly anharmonic oscillations described by a Hamiltonian[5]:

$$\mathcal{H}_\rho(p,q,t) = \mathcal{H}_0(p,q) + \sum_{k,\ell=-\infty}^{\infty} e^{i(k\omega - \ell\Omega)t} \cdot \sum_{\bar{m}=|k|}^{\infty} U_{k\ell\bar{m}} \cdot I^{\frac{\bar{m}}{2}} \qquad (1.6.21)$$

where H_0 is the linear part and the smallness of the perturbation is ensured by the condition: $I \lesssim 1$. The non-linearity is determined here by the first non-vanishing term U_{00m_1} with $m_1 > 2$ (usually U_{004}):

$$\omega'_k = \frac{k}{4} U_{00\bar{m}_1} \cdot \bar{m}_1(\bar{m}_1 - 2) \cdot I^{\frac{\bar{m}_1}{2} - 2} \qquad (1.6.22)$$

and the value of the perturbation for the resonance of the kth harmonic ($k > 0$) is (see 1.3.4):

$$|U_0| \approx 2k \cdot |U_{k\ell k}| \cdot I^{k/2} \qquad (1.6.23)$$

The stabilization condition (1.6.2) takes the form:

$$\lambda \lesssim I^{\frac{\bar{m}_1 - k}{2}}; \qquad \lambda = \frac{8}{\bar{m}_1(\bar{m}_1 - 2)} \left| \frac{U_{k\ell k}}{U_{00\bar{m}_1}} \right| \qquad (1.6.24)$$

For $I \to 0$ this inequality is always fulfilled when $k > m_1$ (stabilization at small amplitudes) and not fulfilled when $k < m_1$. In the latter case stabilization is possible only for $\lambda \lesssim 1$. and the stabilization boundary is given by the estimate:

$$I_1 \sim \lambda^{\frac{2}{\bar{m}_1 - k}} \qquad (1.6.25)$$

The stable region corresponds to a sufficiently large amplitude: $I > I_1$. Let us note that when $\lambda \gtrsim 1$ this region ($I \gtrsim I_1$) becomes unstable for $k > m_1$. When $k = m_1$, the stabilization condition does not depend on I: $\lambda \lesssim 1$,

For the special case of $m_1 = 4$ the estimates obtained agree with the results of the detailed calculations on a similar problem carried out by Schoch[14] (see also Ref. 5) and Mel'nikov[37].

CHAPTER 2

STOCHASTICITY

This is the main chapter of the paper, in which the interaction of several resonances, due to the non-linearity of the system, will be investigated. The interaction of the resonances is a source of instability of the oscillations, which in turn leads to one or another form of stochasticity, i.e. to the appearance of statistical laws in the dynamical system. At this point, classical oscillation theory merges with statistical mechanics and what interests us mainly is the border zone between the two sciences. In contrast to the more elementary investigations of the previous chapter, we are obliged in what follows to turn to a system of simple models and to make greater use of analytical estimates by order of magnitude. Natural dissatisfaction with such a "non-rigorous" approach may be compensated for to a certain extent by the numerical experiments which will be described in the next chapter.

2.1 The basic model

The central problem of this paper is that of the interaction of several resonances. According to the results of the previous chapter, the size of the region of influence of each resonance (in frequency) is of the order of (Section 1.4)[*]:

$$(\Delta \omega)_H \sim \Omega_\varphi \tag{2.1.1}$$

around the resonance value $\omega = \omega_p$. If there are several resonance values of the frequency (ω_i) (several resonances, as we shall say for the sake of brevity), then it is obvious that the character of the motion will essentially depend, generally speaking, on the ratio:

$$s = \frac{(\Delta \omega)_H}{\Delta} \sim \frac{\Omega_\varphi}{\Delta} \tag{2.1.2}$$

where $\Delta = |\omega_{i+1} - \omega_i|$ is the frequency distance between neighbouring resonances. The case of single resonance, thoroughly studied in the previous chapter, corresponds to the condition

$$s \ll 1 \tag{2.1.3}$$

The asymptotic validity of this condition is fully evident[**]. A more accurate criterion of the applicability of the single resonance approximation will be discussed later (Sections 2.2 and 2.7).

---

[*] For the case of moderate non-linearity (1.3.8), which will always be understood if no special reservation is made.

[**] See Section 2.7 though.

In the opposite case

$$s \gtrsim 1 \tag{2.1.4}$$

it is necessary to take into account the interaction of the resonances, namely the simultaneous effect on a non-linear oscillator of several perturbations with different frequencies.

It is not difficult to extend the universal Hamiltonian (1.3.11) to the case of several resonances. Let us choose one of them as the basic resonance (basis of reference) and designate the values relating to it by a zero index. Let us insert the phases $\psi = \Theta - \varphi_0$; $\psi_i = \varphi_0 - \varphi_i$ (see Section 1.3, $k = \ell = 1$). The universal Hamiltonian can now be written in the form:

$$\mathcal{H}_y = \frac{\omega_o'}{2} \left(I - I_o \right)^2 + \varepsilon \sum_i U_i \, \sin \left(\psi + \psi_i \right) \tag{2.1.5}$$

whence the equations of motion in a first approximation are

$$\dot{I} = \varepsilon \sum_i U_i \, \cos \left(\psi + \psi_i \right)$$
$$\dot{\psi} = \omega_o' \left(I - I_o \right) ; \quad \dot{\psi}_i = \Omega_o - \Omega_i \tag{2.1.6}$$

One can express the following qualitative considerations about the behaviour of this system under conditions of interaction of the resonances (2.1.4). Each term defines its own "centre of attraction" around which the phase oscillations of our "particle" (see Section 1.3) can take place. In other words, in the oscillator phase plane (I,Θ) instead of one "potential well" (or rather one "bucket", Fig. 1.4.1) there are a number of "potential wells" around I_i. Under condition (2.1.4) these "wells" overlap, which makes it possible for the "particle" to cross over from one well to another. The transition conditions depend on the phase relation $\psi + \psi_i$, and generally speaking, vary continuously, since the "wells" shift with respect to each other along Θ on account of the difference of the frequencies Ω_i.

The law governing the migration of the "particle" from one "well" to another depends on the specific form of the perturbation and in particular on the phase relations. Later we shall give examples of the various types of migration (Section 2.4). However, it can be considered that in the limiting case of very large overlapping of the resonant zones

$$s \gg 1 \tag{2.1.7}$$

the law of migration will be almost random. The reason is the very intricate variation of I in this case (2.1.6), especially if one takes into account that the phases $\psi + \psi_i$ determining this variation themselves depend on I by virtue of the non-linearity of the oscillator[*].

---

[*] This conclusion is not trivial, see Section 2.8.

It would seem that the motion cannot be "completely" random since it satisfies the dynamic equation. However, the imitation of all the known properties of a random process is possible and is sometimes so good that the question arises as to whether a "real" random process is only a clever "imitation". Discussion of this question will be postponed until Section 2.13.

Motion of such a quasi-random type will henceforth be called stochastic, on the understanding that this covers all the features of a random process at present known (Section 2.3). The study of the stochastic motion of a mechanical system, begun mainly in connection with the problem of the foundation of statistical mechanics [Section 2.13; see, for instance Ref. 16], has now become a whole new branch of mathematics -- the metric theory of dynamical systems -- which we shall refer to in the rest of the paper by a shorter though less felicitous term, the ergodic theory[*]. Unfortunately this theory, as a rule, is too abstract and is not easy to apply to specific physical problems. It should be stated at once that the most recent and most important results of the theory[17,19,20,31] are better in this sense and will be widely used in this paper.

Our basic task is to validate inequality (2.1.4) as a criterion of stochasticity, namely as the border separating the stable and stochastic regions, for the special case of a mechanical system of the form (1.1.1), and also to calculate the specific parameters of stochastic motion.

The study of the general case of the interaction of resonances (2.1.6) encounters considerable difficulties, the meaning of which will be clear in what follows. Therefore we shall first simplify the model (1.2.5) chosen in the previous chapter, assuming that the perturbation acts on the oscillator periodically (period $T = 2\pi/\Omega$), each time for a very short interval of time $\tau \to 0$ (approximation of short kicks). Equation (1.2.5) in this case takes the form:

$$\dot{I} = -\varepsilon h_\theta (I, \theta)$$
$$\dot{\varphi} = \varepsilon h_I (I, \theta)$$

(2.1.8)

The phase ϑ dependence of the perturbation ($\dot{\vartheta} = \Omega$) reveals itself by the fact that the Hamiltonian $h(I,\Theta)$ is different from zero only at intervals of τ; the indices Θ, I denote partial differentiation with respect to the corresponding argument:

By integrating the system over the interval τ we obtain as a first approximation :

$$\Delta I = -(\varepsilon\tau) \cdot h_\theta (I_\circ, \vartheta_\circ) + O\left(\varepsilon^2, (\tau\omega)^2\right)$$
$$\Delta \varphi = (\varepsilon\tau) \cdot h_I (I_\circ, \theta_\circ) + O\left(\varepsilon^2, (\tau\omega)^2\right)$$

(2.1.9)

---

[*] The present state of the theory is presented rather completely in a paper by Sinai[17]. See also Refs. 41 and 42.

where I_0, Θ_0 are the initial values. In the interval between kicks $I = \text{const.}$, and the phase varies by the value $\Delta\Theta = (T - \tau) \cdot \omega(I_1)$, where I_1 is the value after the kick. The total phase shift during the period is:

$$\Delta\theta = (T-\tau) \cdot \omega(I_1) + \int_0^\tau \omega \, dt + (\varepsilon\tau) \cdot h_I(I_0, \theta_0) =$$
$$= T \cdot \omega(I_1) + (\varepsilon\tau) \cdot h_I(I_0, \theta_0) + O(\varepsilon^2, (\tau\omega)^2) \qquad (2.1.10)$$

We can now describe the motion of the model by means of a system of difference equations:

$$I_{n+1} = I_n - \varepsilon h_\theta(I_n, \theta_n)$$
$$\theta_{n+1} = \theta_n + T\omega(I_{n+1}) + \varepsilon h_I(I_n, \theta_n) \qquad (2.1.11)$$

where $\tau = 1$, and the index "n" denotes the number of the kick (step), the new discrete time of our dynamical system. Let us recall that the Hamiltonian $h(I,\Theta)$ is a periodic function of Θ with a period of 2π.

Equations (2.1.11) are written to first approximation in ε and can be put down more accurately if necessary, using (2.1.8). In particular, let us write the expression for ΔI with an accuracy $\sim \varepsilon^2$, which we shall need later on:

$$\Delta I = -(\varepsilon\tau)h_\theta + \frac{(\varepsilon\tau)^2}{2} \cdot [h_{\theta I} \cdot h_\theta - h_{\theta\theta} \cdot h_I] -$$
$$- \frac{\varepsilon\tau^2}{2} \cdot \omega_0 \cdot h_{\theta\theta} + O(\varepsilon^3, (\tau\omega)^3) \qquad (2.1.12)$$

Since the original system (2.1.8) is canonical, the Jacobian of the transformation (2.1.11) is equal to unity with the corresponding accuracy:

$$\left| \frac{\partial(I_{n+1}, \theta_{n+1})}{\partial(I_n, \theta_n)} \right| = 1 + O(\varepsilon^2) \qquad (2.1.13)$$

which is easy to verify also by direct calculation.

Equations (2.1.11) determine the basic model of the interaction of the resonances. It will sometimes be convenient to simplify it even further. As in the case of a single resonance, the behaviour of the system to a certain degree does not depend on the specific form of the function $h(I,\Theta)$[*], and therefore we shall choose the two most simple cases [(2.1.14) and (2.1.15)]. Further, one can neglect in the first $\sqrt{\varepsilon}$ approximation (see Section 1.4) the last term in the second equation (2.1.11), which represents a linear correction to the frequency [Section 1.6 (1.6.17)]. Finally, instead of the action variable I, one can directly use the frequency of the oscillator ω. As a result we shall obtain the

---

*) See Section 2.7.

following most simple model equations describing the phenomenon of the interaction of the resonances:

$$\omega_{n+1} = \omega_n + \varepsilon \left(\psi_n - 1/2 \right)$$
$$\psi_{n+1} = \left\{ \psi_n + T \cdot \omega_{n+1} \right\} \tag{2.1.14}$$

or

$$\omega_{n+1} = \omega_n + \varepsilon \cdot \cos 2\pi \psi_n$$
$$\psi_{n+1} = \left\{ \psi_n + \frac{T}{2\pi} \omega_{n+1} \right\} \tag{2.1.15}$$

Here the curly brackets represent the fractional part of the argument -- a convenient way of specifying the periodic dependence. The coefficients of the model equations (2.1.14) and (2.1.15) are selected so that the Jacobian $|\partial(\omega_{n+1}, \psi_{n+1})/\partial(\omega_n, \psi_n)| = 1$ exactly. The reasons for the choice of two forms of dependence on ψ will be clear from what follows (see Section 2.4).

We chose for our basic model (2.1.11) a perturbation in the form of short kicks, essentially in the form of a δ-function. This choice is not very special or exceptional; on the contrary, it is typical, since the sum in the right-hand part (2.1.6), when there are a large number of terms, actually represents either a short kick (or series of kicks) or frequency-modulated perturbation. In the latter case periodic crossing of the resonance takes place, which according to the results of Section 1.5 is also equivalent to some kick [(1.5.7) and (1.5.9)]. Thus it can be expected that the properties of model (2.1.11) will be in a sense typical for the problem of the interaction of the resonances and stochasticity.

The transition to the difference equation (2.1.11) or, as they say, to the transformation, means essentially the integration of the original system of differential equations over the period of the perturbation, integration which becomes trivial for the special case considered. We thus obtain some information about the behaviour of the system in a finite, and characteristic, interval of time. This is really a reason for simplifying the original system.

The true significance of the basic model is explained in Section 2.6, where it will be shown that it describes the motion near the non-linear resonance separatrix and in particular the stochastic layer. The latter turns out to be the origin of any instability of non-linear oscillations. Thus it appears possible to study the general case of the interaction of resonances, using the basic model only.

2.2 Kolmogorov stability

Let us return to Eq. (2.1.11). If the perturbation is sufficiently small ($\varepsilon \to 0$) and $T\omega = 2\pi k$ (k is an integer), i.e. if the system is near to the resonance, the difference equations can again be replaced by the differential ones:

$$\dot{I} = -\frac{\varepsilon}{T} h_\theta (I, \theta)$$

$$\dot{\theta} = \left(\omega(I) - \omega_p \right) + \frac{\varepsilon}{T} h_I (J, \theta) \tag{2.2.1}$$

where ω_p is the resonance value of the oscillator frequency ω.

Let us study the nature of the motion in this case. First of all let us note that the Eqs. (2.2.1), of course, are not identical to the original ones (2.1.8), in spite of some resemblance. The derivatives (2.1.8) relate to the interval of time $\ll \tau$ (time of action of the perturbation), whereas the characteristic time for the derivatives (2.2.1) should be $\gg T$ (period of action of the perturbation). This means that both the differential equations (2.2.1) and the difference equations (2.1.11) contain some information about the solution of the original system (2.1.8) during the perturbation period, as noted above.

Let us further point out that Eqs. (2.2.1) agree exactly with the equations (1.3.5) in Section 1.3, describing single resonance. Consequently, in the approximation under consideration there is no interaction of resonances and the motion has the character of limited phase oscillations (Section 1.4).

Let us consider these phase oscillations more thoroughly for model (2.1.15). The differential equations in this case take the form:

$$\dot{\omega} = \frac{\varepsilon}{T} \cos 2\pi \psi$$

$$\dot{\psi} = \frac{\omega}{2\pi} - \frac{k}{T} \tag{2.2.2}$$

where k is an integer. The universal Hamiltonian (see Section 1.3) is equal to:

$$\mathcal{H}_y = \frac{1}{4\pi} \left(\omega - \frac{2\pi k}{T} \right)^2 - \frac{\varepsilon}{2\pi T} \cdot \sin 2\pi \psi \tag{2.2.3}$$

The most important characteristic of a non-linear resonance is the width of the separatrix determining the region of influence of the resonance. In the present case it is (Section 1.4):

$$(\Delta \omega)_H = 4 \sqrt{\frac{\varepsilon}{T}} \tag{2.2.4}$$

The approximate replacement of the difference equations by the differential ones (2.2.1) is thus equivalent to taking into account a single resonance. Let us show this directly. For this let us return to the original equations (2.1.8) which for model (2.1.15) take the form:

$$\dot{\omega} = \varepsilon \sum_{n=-\infty}^{\infty} \delta(t - nT) \cdot \cos 2\pi \theta$$

$$\dot{\theta} = \frac{\omega}{2\pi} \tag{2.2.5}$$

Now by expanding the periodic δ-function into a Fourier series, singling out, as usual, the resonant harmonic, for which $T\omega \approx 2\pi k$ and inserting $\psi = \theta - kt/T$, we obtain exactly system (2.2.2).

Let us consider more accurate conditions under which the difference equations (2.1.11) can be replaced by the differential equations (2.2.1). For this it is evident that the following inequalities must be satisfied:

$$I_{u+1} - I_u = \Delta I \ll I \; ; \qquad \Delta \psi \ll 1 \qquad (2.2.6)$$

In order to satisfy these inequalities it is necessary first of all for the parameter $\varepsilon \ll 1$. This is not, however, an additional limitation, since we always consider the perturbation to be small. Further, the value $T\omega$ must be near to a multiple of 2π:

$$\left| T\omega - 2\pi k \right| \ll 1 \qquad (2.2.7)$$

This condition in its turn can be broken down into two: firstly, the initial detuning must be small:

$$\left| T\omega_0 - 2\pi k \right| \ll 1 \qquad (2.2.7')$$

and secondly, the variation of ω in the process of motion must also be sufficiently small:

$$T \cdot (\Delta \omega)_H \ll 1 \qquad (2.2.8)$$

Let us show in example (2.1.15) that condition (2.2.7') is unimportant. Since it is not connected with non-linearity, let us assume that the system is linear, i.e. that $\omega = \omega_0 = \text{const}$. In this case the second equation (2.1.15) gives $\psi_n = \psi_0 + n\Theta_0/2\pi$: $\Theta_0 = T\omega_0$, whence:

$$\omega_n = \omega_0 + \varepsilon \sum_{k=1}^{u-1} \cos \left(2\pi \psi_0 + k\theta_0 \right) \qquad (2.2.9)$$

The latter sum permits a simple estimate:

$$\left| \omega_n - \omega_0 \right| = \varepsilon \left| \operatorname{Re} \sum_{k=1}^{u-1} e^{2\pi i \psi_0 + i k \theta_0} \right| \leq \varepsilon \left| \frac{1 - e^{i(u-1)\theta_0}}{1 - e^{i\theta_0}} \right| \qquad (2.2.10)$$

Its value is always small except for the resonance regions, where condition (2.2.7') is fulfilled.

In the general case the force $f(\psi)$ in a transformation of type (2.1.15) has all the harmonics: $f(\psi) = \Sigma_q f_q e^{2\pi i q \psi}$ and then the sum (2.2.10) diverges for any rational $\Theta_0/2\pi$. But this simply means that, besides the main resonances $\omega T/2\pi = k$ (integer) in (2.2.2) generally speaking the resonances of the higher harmonics $\omega T/2\pi = r/q$ (rational) should also be taken into account. This question will be discussed in Section 2.7. Going on ahead, let us note that for a sufficiently rapid decrease in the amplitude of the harmonics f_q with the growth of q, the resonances of the higher harmonics can be neglected.

There thus remains one important condition for the validity of the approximation by a single resonance, and precisely condition (2.2.8) which agrees in order of magnitude with inequality (2.1.3), since in the present case the distance between the resonances $\Delta = 2\pi/T$.

So far we have restricted ourselves to the first approximation only, taking into account some rough effects of the second approximation. Naturally the question arises as to whether some fine effects of the higher approximations qualitatively change the solution after a sufficiently long time; in other words, are there not some kind of cumulative corrections of the higher approximations?

The KBM theory enables us to construct a solution in the form of an asymptotic series in powers of the small parameter ε, the residual term of which is of the order of $R_N \sim \varepsilon^{N+1} \cdot t$ [3]. Such series, as is known, diverge and therefore there is no guarantee against exponentially small error, say $\sim t \cdot e^{-A/\varepsilon}$. It is true that if the system has finite damping the asymptotic solution remains valid for any t when there is a sufficiently small fixed ε [3]. However, for conservative systems the question remains open[*].

The practical construction of asymptotic series is a highly laborious task. Apparently the best technique for such construction was devised by Kruskal[18].

Only relatively recently, in papers by Kolmogorov[19], Arnold[20] and Moser[28], a new technique for constructing convergent series was developed, which makes it possible in some cases to solve the problem of the stability of the motion of a conservative system in an infinite interval of time[**]. This progress was possible because the problem was formulated in a different way. The perturbed trajectory is generally calculated for given initial conditions. In the averaging method[3] the calculation of the variation of the frequencies of the motion in each successive approximation plays an important part; this makes it possible to avoid trivial secular terms[3]. Instead of this in the KAM theory the perturbed trajectory, or rather the invariant surface (torus), is calculated for given frequencies and the torus shifts a little and becomes deformed in the phase space in each successive approximation. In other words, in the KAM theory a different principle of splitting up the phase space into trajectories is applied. It turns out that in order to conserve such tori in the presence of perturbation it is necessary, firstly, for the system to be non-linear and, secondly, so that the frequencies of the motion on the torus shall have some special arithmetical properties, roughly speaking, it is necessary for their quotients not to be too close to rational numbers (see Section 2.1.2). The change in the formulation of the problem and the success in solving it are connected with precisely this latter condition. However, this condition is of a rather specific nature, it is not physical. Although the invariant surface of the unperturbed system has "good" frequencies with a probability of unity, arbitrarily near to it are surfaces with "bad" frequencies which are destroyed by the perturbation. In a real system it is not possible to distinguish between these two kinds of invariant tori. Thus real conclusions on the stability of the motion can be drawn only for a two-dimensional autonomous or a one-dimensional non-autonomous system. In this

---

[*] In the case of small damping some effects may also be missed. See Section 2.10.

[**] From now on we will refer to these papers as the KAM theory.

case the invariant tori are inserted one inside the other and thus the "bad" tori are confined between the "good" ones, which ensures general stability of the motion independently of the mythical arithmetical properties of the frequencies[*]. For the many-dimensional case the question remains open for the present; there is only an example of instability constructed by Arnold[21]. This question will be more thoroughly discussed in Section 2.12.

Thus, in the limiting case of $s \to 0$ (2.1.3) the motion of a system of the form (1.1.1) actually has the character of limited stable phase oscillations. However, in its present state the KAM theory does not make it possible effectively to estimate the critical value ε_{cr}. The existing estimates[20] are clearly too low by many orders of magnitude. The numerical experiments (Chapter III) show that ε_{cr} is of the same order of magnitude as the border of stochasticity $s \sim 1$.

2.3 An elementary example of stochasticity

Let us go over to the solution of the system of difference equations (2.1.11) in the case when condition (2.2.8), or inequality (2.1.3) which is equivalent to it, is violated.

Let us begin with an elementary example. Let us consider model (2.1.14), rewriting the equations in the form:

$$\omega_{n+1} = \omega_n + \varepsilon \left(\psi_n - 1/2 \right)$$
$$\psi_{n+1} = \left\{ \psi_n + T\omega_n + \varepsilon T (\psi_n - 1/2) \right\} \tag{2.3.1}$$

Condition (2.2.8) in the present case may be written in the form [see (2.2.4)]:

$$\varepsilon T \gg 1 \tag{2.3.2}$$

The second of the equations takes on essentially the character of phase extension with a coefficient εT. Thus it can be replaced in its turn by a model transformation of the form:

$$\psi_{n+1} = \left\{ k \psi_n \right\} \tag{2.3.3}$$

It is difficult to imagine a simpler (and rougher) model of a dynamical system. Nevertheless, it enables us to trace the most important features of the phenomenon of stochasticity. Moreover this is the only model whose properties are completely known and furthermore in the form of rigorous mathematical theorems with all the necessary conditions and reservations[**]. It can therefore serve as a safe point of departure, from which we will endeavour to progress further by means of less rigorous methods of qualitative estimates, physical (model) considerations and numerical experiments.

---

[*] We shall call this case one-dimensional.

[**] The main results are in the papers by Rokhlin[22] and Postnikov[33].

When $k > 1$ the motion of the system (2.3.3) has all the attributes of a random process so far known -- ergodicity, mixing and positive K-entropy[22] (see below). As mentioned above, we shall call such systems stochastic[*].

The <u>ergodicity</u> of system (2.3.3) means the uniform distribution of the sequence ψ_n in the segment $(0,1)$. The mixing is closely connected with the correlations in the system. Let us consider several different trajectories with initial conditions: $\psi_0^{(1)}$, $\psi_0^{(2)}$, ..., $\psi_0^{(r)}$. Let us combine them in one trajectory of an r-dimensional point $(\psi_n^{(1)}, ..., \psi_n^{(r)})$. We will speak of the absence of r-fold correlations in the original system (2.3.3), if the combined r-dimensional system possesses ergodicity, i.e. if the trajectory of the point $(\psi_n^{(1)}, ..., \psi_n^{(r)})$ uniformly fills the r-dimensional hypercube when $n \to \infty$.

What is known as <u>weak mixing</u> means the absence of pair (twofold) correlations[**]. The term "weak" shows that this property is not sufficient for obtaining stochasticity. It turns out[26] that with weak mixing only, the continuous distribution function (of the ensemble of the systems) in the phase space even in the steady state undergoes strong, although also infrequent, variations; this is unsatisfactory from the point of view of statistical mechanics. Let us recall for purposes of comparison that when only ergodicity is present there is no steady state at all, but the distribution function varies almost periodically[26].

Infrequent but strong oscillations of the distribution function when there is weak mixing are apparently due to the higher correlations $(r > 2)$. If the distribution function relaxes to a steady-state function (constant), i.e. if the oscillations of the distribution function decrease infinitely when $t \to \infty$, one talks of <u>strong mixing</u> or simply <u>mixing</u>. It is natural to assume that (strong) mixing is equivalent to the absence of correlations of any multiplicity[†]. In order to give a full picture let us mention, going on a little further ahead, that in the special but very important case when the relaxation process goes according to an exponential law, one speaks of the positive K-entropy of the system.

By virtue of the ergodicity, the correlations of several trajectories are equivalent to the correlations of several points taken successively in the same trajectory: $(\psi_n^{(1)}, ..., \psi_n^{(r)}) \to (\psi_{n+k_1}, ..., \psi_{n+k_{r-1}})$. However, in this case all shifts in time between the points $(|k_i - k_j|)$ must increase infinitely with the growth of n. Correlations with constant shifts are called <u>autocorrelations</u>. These always exist in a mechanical system, since its motion is unambiguously determined by reversible dynamical equations. Thus mixing means asymptotic (i.e. with $|k_i - k_j| \to \infty$) dying down of the autocorrelations.

The notion of mixing is also connected with the notion of the <u>completely uniformly distributed sequence</u> introduced by Korobov[23] (see also Ref. 24). This last term means the absence of autocorrelations of any multiplicity with arbitrary non-zero shifts $(k_i \neq k_j \neq 0)$. This sequence obviously cannot be given by dynamical equations. However,

---

[*] Another term used in ergodic theory is <u>K-systems</u>, in honour of Kolmogorov who discovered them.

[**] For other definitions of mixing and the connection between them see the book by Halmos[26].

[†] This assumption still remains only a more or less plausible hypothesis.

the dynamical sequence is asymptotically completely uniformly distributed when there is mixing.

Let us consider pair autocorrelation more thoroughly. We determine the correlation coefficient by

$$\rho_n^{(2)}(q) = \left\langle e^{2\pi i (\psi_n + q\psi_o)} \right\rangle \qquad (2.3.4)$$

where averaging can be carried out over ψ_0 by virtue of the ergodicity, and q is an integer. An advantage of this definition of the correlation coefficient for a system of type (2.3.3) as compared to the standard

$$\rho_n = \frac{\left\langle (\psi_n - \langle \psi_n \rangle)(\psi_o - \langle \psi_o \rangle) \right\rangle}{\left\langle (\psi_o - \langle \psi_o \rangle)^2 \right\rangle} \qquad (2.3.5)$$

is that the integer part of ψ is automatically excluded, which considerably simplifies the calculation. At the same time $|\rho_n^{(2)}(q)|$ has the properties of a standard correlation coefficient[*].

From (2.3.3) and (2.3.4) we have:

$$\rho_n^{(2)}(q) = \frac{e^{2\pi i (k^n + q)} - 1}{2\pi i (k^n + q)} \qquad (2.3.6)$$

For integer k the correlation coefficient vanishes, because of the nature of its definition[**], for all q except $q = -k^n$; in the other remaining cases it is of the order of:

$$|\rho_n^{(2)}(q)| \sim \frac{1}{|k^n + q|} \rightarrow e^{-n \ln k} \qquad (2.3.7)$$

and asymptotically decreases exponentially.

From this estimate one can also draw interesting conclusions on the space structure of the mixing, which is characterized by the parameter q. In fact, expression (2.3.4) represents the qth Fourier component of the correlation, i.e. it characterizes the correlations in the region of scale 1/q. From estimate (2.3.7) it follows that the correlation coefficient for a given q does not decrease immediately, but only after some time (number of steps), when $k^n > q$. In other words, the mixing process spreads gradually into increasingly small regions. Assuming that $k^n \sim q$, one can obtain an estimate of the size of the region up to which the mixing extends in time:

---

[*] The idea of such a definition arises from Weyl's[12] criterion for uniform distribution of a sequence (see also Ref. 24). Let us note that (2.3.4) is the standard correlation coefficient of the quantities $e^{2\pi i \psi_n}$ and $e^{2\pi i q\psi_0}$.

[**] See also Section 2.11.

$$\left(\frac{1}{q}\right)_n \sim e^{-n\,\ell n\,k} \tag{2.3.8}$$

Thus the size of the region not yet affected by mixing also decreases exponentially with time.

Let us consider, finally, yet another, probably the most important, property of a stochastic system -- local instability of motion. This means that trajectories that are close together at first rapidly diverge. For model (2.3.3) we obtain directly:

$$(\delta \psi)_n = (\delta \psi)_o \cdot k^n = (\delta \psi)_o \cdot e^{n\,\ell n\,k} \tag{2.3.9}$$

i.e. the instability also develops exponentially at the same rate as the correlations (2.3.7) decrease and the correlation length (2.3.8) is reduced.

Local instability of motion is the specific mechanism which ensures mixing and decrease of the correlations in the mechanical system.

The connection between local instability and stochasticity was first noted, apparently, by Hopf[29] and Hedlund[35], analysed in detail as applied to mechanical systems by Krylov[30], and rigorously proved for a rather general case in recent papers by Anosov[31] and Sinai[34,17]. Local instability appears to be a very convenient practical criterion of stochasticity, since it needs only the investigation of linearized equations. It is also not out of the question that local instability plays a decisive part in understanding the nature of the statistical laws (see Section 2.13).

Using relation (2.3.9) the whole mixing process for our model can be visually traced. At first $\left[n \leq \ln\,(1/\delta\,\psi_0)/\ln\,k\right]$ the segment $\delta\psi_0$ simply extends until it reaches the size of the whole region $(\delta\psi_n \sim 1)$. After this begins the mixing of the trajectories emerging from $(\delta\psi)_0$ throughout the whole region (0,1). At the moment when $(\delta\psi)_n \sim 1$ the correlation length $\sim (\delta\psi)_0$, since the trajectories of this segment $(\delta\psi_0)$ just begin to mix. This condition leads, of course, to the previous estimate (2.3.8).

The distinctive feature of a stochastic system is just the exponential development of local instability and the development of the resulting process of mixing and decrease of the correlations. Exponential law ensures fast transition to "random" motion with a high degree of accuracy. It will therefore be understood that the characteristic rate of this exponential process is of the greatest importance for the stochastic system. It was introduced in Ref. 25 and is generally called entropy. In our case:

$$h = \ell n\,k \tag{2.3.10}$$

Sinai[34,17] established that this definition of entropy was equivalent to the original one in Ref. 25, which was more complex. This quantity had already actually been widely used by Krylov[30] and may therefore be called Krylov-Kolmogorov entropy, or K-entropy.

The term entropy for the quantity (2.3.10) [25] cannot be regarded as felicitous, because there is confusion with the usual thermodynamic entropy. In fact these quantities are completely different even in dimension.

Thermodynamic entropy characterizes the statistical state of the system and depends only on the distribution function[27]:

$$H = - \int f \, \ln f \, d\mu \tag{2.3.11}$$

where μ is the invariant measure of the region (the volume of the phase space for Hamiltonian systems). In particular, thermodynamic entropy is constant in a steady state (f = const.). For the classical system it is defined except for the constant, whose value is connected with phase space quantization[27] namely with the fact that the quantum system cannot occupy a region in the phase space less than some $\Delta\mu_{kb}$. This condition leads to the standard expression[27]:

$$H = - \int f \, \ln \left(f \cdot \Delta\mu_{\kappa e} \right) d\mu \tag{2.3.12}$$

For a purely classical system one can also introduce some minimum permissible phase volume ($\Delta\mu_{k\Lambda}$) from the following considerations. In its physical meaning entropy characterizes just the stochastic motion of the system. On the other hand, although in principle a classical system may also occupy an arbitrarily small volume, its motion will not be stochastic in the regions smaller than correlation volume $\Delta\mu_c(t)$, similar to the correlation length $1/q$ (2.3.8) for model (2.3.3). It is therefore natural to choose $\Delta\mu_c$ as the minimum permissible volume when calculating the entropy: $\Delta\mu_{k\Lambda} \sim \Delta\mu_c$. As a result we obtain the relation:

$$H(t) = - \int f \cdot \ln \left(f \cdot \Delta\mu_c(t) \right) d\mu \tag{2.3.13}$$

defining a new entropy which now depends not only on the statistical state (f) but also on the dynamics of the mixing $\left[\Delta\mu_c(t) \right]$. In this form it is difficult to use, both in statistical and dynamical theory. However, it is easy to obtain from it the quantity characterizing mixing dynamics only. For this let us choose any specific statistical state, for example, steady state (f = const.). Defined, in such a way, the dynamical entropy permanently increases with time (in a state of statistical equilibrium!) for any system with mixing. In the case of systems with an exponentially decreasing correlation length the entropy (2.3.13) proves to be asymptotically proportional to time. It is therefore natural to introduce as a characteristic its mean rate of change:

$$h = \lim_{t \to \infty} \left(- \frac{\int \ln \left(\Delta\mu_c(t) \right) d\mu}{t} \right) \tag{2.3.14}$$

This is precisely K-entropy; it is of the same dimension as the frequency, and is therefore sometimes called entropy per time unit (or per one step). We will call it K-entropy.

2.4 Stochasticity of the basic model

Let us turn to the more real non-linear resonance model given by the difference equations (2.1.11). In order to study stochasticity it is most convenient to investigate the local stability of the solution.

For this let us write the linearized equations in the variations[*]:

$$\delta I_{n+1} = (1 - \varepsilon h_{\theta I}) \cdot \delta I_n - \varepsilon h_{\theta\theta} \cdot \delta\theta_n$$

$$\delta\theta_{n+1} = [T\omega'(1 - \varepsilon h_{\theta I}) + \varepsilon h_{II}] \cdot \delta I_n + (1 - \varepsilon h_{\theta\theta} T\omega' + \varepsilon h_{I\theta})\delta\theta_n \qquad (2.4.1)$$

By equating the right-hand sides of the equations obtained to $\lambda \cdot \delta I_n$ and $\lambda \cdot \delta\theta_n$, respectively, we shall find the characteristic equation for λ:

$$\lambda^2 - (2 + K)\lambda + 1 = 0 \qquad (2.4.2)$$

where the coefficient

$$K \approx -\varepsilon T\omega' h_{\theta\theta} \qquad (2.4.3)$$

We put the last term of the characteristic equation as unity, since it is equal to the Jacobian of transformation (2.1.11), which in its turn is obtained from the exact Hamiltonian equations (2.1.8). In fact for this it is necessary to take into account the subsequent terms of the expansion in ε (Section 2.1). In the expression for K there also appear additional terms of the order of ε^2, $\varepsilon^2 T\omega'$, $\varepsilon^3 T\omega'$, ... It is, however, essential for the factor $T\omega'$, which may become large, always to participate only in the first power and therefore the additional terms mentioned are small in comparison to the main one (2.4.3).

The only coefficient of the characteristic equation (K) is closely connected with the extension of the phase:

$$\frac{d\theta_{n+1}}{d\theta_n} \approx 1 + K + \varepsilon h_{I\theta} \qquad (2.4.4)$$

The validity condition for the approximation of the single resonance (2.2.8) or (2.1.3), for system (2.1.11) takes the form:

$$K \ll 1 \qquad (2.4.5)$$

In this paragraph we shall deal with the opposite case of $K \gtrsim 1$. Thus the last term in (2.4.4) can be neglected.

The roots of the characteristic equation are given by the expression:

$$\lambda = 1 + \frac{K}{2} \pm \sqrt{K\left(1 + \frac{K}{4}\right)} \qquad (2.4.6)$$

Depending on the value of λ the solution may be of two qualitatively different types. The first corresponds to the complex conjugate roots and takes place under the condition

---

[*] Another name for (2.4.1) common in the ergodic theory is tangent transformation.

$$- 4 < K < 0 \qquad (2.4.7)$$

It is easy to verify that in this case $|\lambda_1| = |\lambda_2| = 1$, and consequently the linearized transformation (2.4.1) represents rotation by an angle of ϕ_0:

$$tg\,\varphi_o = \pm \frac{\sqrt{-K\,(1 + K/4)}}{1 + K/2}. \qquad (2.4.8)$$

which corresponds to the oscillatory nature of the solution of (2.4.1) with a frequency of ϕ_0/T. This is the case of local stability of motion.

The quantity K is a periodic function of Θ; let us introduce the amplitude:

$$K_o = \max |K| \qquad (2.4.9)$$

When $K_0 \ll 1$ the solution of (2.4.8) gives well-known phase oscillations (Section 2.2) near to the stable equilibrium state ($h_\Theta = 0$; $h_{\Theta\Theta} \cdot \omega' > 0$). It is important, however, that such oscillatory solutions are possible, generally speaking, for any K_0, including $K_0 \gg 1$, near to the points of weak phase extension ($h_{\Theta\Theta} \approx 0$; $K \approx 0$; $d\Theta_{n+1}/d\Theta_n \approx 1$). Let us discuss this a little more thoroughly.

First of all let us estimate the size of the region of local stability in the phase space of the system (ΔI, $\Delta\Theta$). This can be done using the condition that the parameter K does not go outside the limit of the interval (2.4.7). We obtain the size of the stable phase region immediately (2.4.7)

$$\Delta\theta \lesssim 4 \left| \frac{\partial K}{\partial \theta} \right|^{-1} \sim K_o^{-1} \qquad (2.4.10)$$

The latter estimate is valid provided that the function $K(\Theta)$ is sufficiently smooth. The permissible value of ΔI is obtained from the second equation (2.1.11), with the requirement that the phase variation $\Theta_{n+1} - \Theta_n$ shall not exceed the value (2.4.10):

$$\Delta I \lesssim \frac{\varepsilon}{K_o^2} \qquad (2.4.11)$$

Thus the phase volume of the region of local stability turns out to be $\leq \varepsilon K_0^{-3}$.

Further, let us note that when $K_0 \gg 1$ the stability ($h_{\Theta\Theta} \approx 0$) and constancy of I [$h_\Theta \approx 0$, see (2.1.11)] are mutually exclusive as a rule. Therefore even in the stable region I varies. This means that the system leaves the stable region after one step because of the variation of the term $T\omega$ in (2.1.11) and thus the actual size of the stable region proves to be considerably smaller (see Sections 2.8 and 3.5).

However, a special case is possible, when the variation of $T\omega$ is equal to an integral multiple of 2π:

$$\varepsilon T\omega' h_\theta = \pm 2\pi k \ ; \quad k \geqslant 1 \qquad (2.4.12)$$

Here a stable process of increasing or decreasing the energy of the oscillator may take place. Such a process is used to accelerate charged particles in a microtron and historically this was the first proposal for the use of the phenomenon of phase stability in accelerators[7].

Since the left-hand side of the last equality $\sim K_0$ and the size of the stable region shrinks abruptly with the growth of K_0, in practice microtron conditions are essential in the region $K_0 \sim 1$. More complex periodic conditions are also possible, under which the quantities I, θ pass through several different values before returning to the original ones. Such conditions are thoroughly studied by Moroz[182], where they are called generalized microtron regimes. The role of all such stable regions and the related estimates will be discussed in Section 2.8.

Let us now go over to the solution of (2.4.1) in the case of the real roots of the characteristic equation, that takes place for values of K outside the interval (2.4.7). At the edges of this interval $\lambda = \pm 1$. Excluding this trivial case one of the real roots is always greater than unity in absolute value, and the other smaller, because of the condition $\lambda_1 \cdot \lambda_2 = 1$.

Let us first consider the simplest case, when the roots $\lambda_{1,2}$ and the eigenvectors of the transformation $\vec{e}_\xi, \vec{e}_\eta$ are constant (do not depend on θ). Then the solution of (2.4.1) can be written in the form:

$$\xi_n = \xi_0 \cdot \lambda^{-n} ; \quad \eta_n = \eta_0 \cdot \lambda^n \qquad (2.4.13)$$

where ξ, η are the coordinates along the eigenvectors: $\lambda = \lambda_1 = 1/\lambda_2 > 1$. Model (2.1.14) has just this property and is an exception in this sense (see below).

The description of the motion by the variables ξ, η, namely the description of the relative shifting of the points of the phase plane (not necessarily close), may be called the transverse flux[34,17]. In the simplest case which we are discussing the structure of the transverse flux is very simple: all the trajectories asymptotically approach the η axis when $n \to \infty$, and the ξ axis when $n \to -\infty$. The flux of such a structure will be called asymptotic. Let us note that the two special trajectories of the transverse flux, along which either continuous extension ($\xi = 0$) or continuous contraction ($\eta = 0$) takes place, are asymptotes. According to the aforementioned results of the papers by Anosov[31] and Sinai[34,17], the stochasticity of a Hamiltonian system is equivalent to the existence of an asymptotic transverse flux in the vicinity of any point of the phase space, or in other words, to the splitting up of all the phase space into asymptotic trajectories.

The regular nature of the transverse flux necessarily leads to residual autocorrelations, vanishing only when $n \to \infty$. Any initial region of the phase plane extends exponentially in the direction of η and contracts along ξ. The mixing process begins after the length of the region extended in the direction of η reaches the maximum size permissible for this system[*]. The initial region is then transformed into a set of increasingly thin

---

[*] This limitation is always fulfilled for systems of the oscillator type, which are being studied in the present paper, at least for some of the variables (phases). Let us note that the extension in such systems occurs mainly just in phase (see below).

(along ξ) layers crossing the phase plane along the axis η and uniformly filling it like "flaky pastry". The initial stage of the mixing process for model (2.1.16) is shown schematically in Fig. 2.4.1.
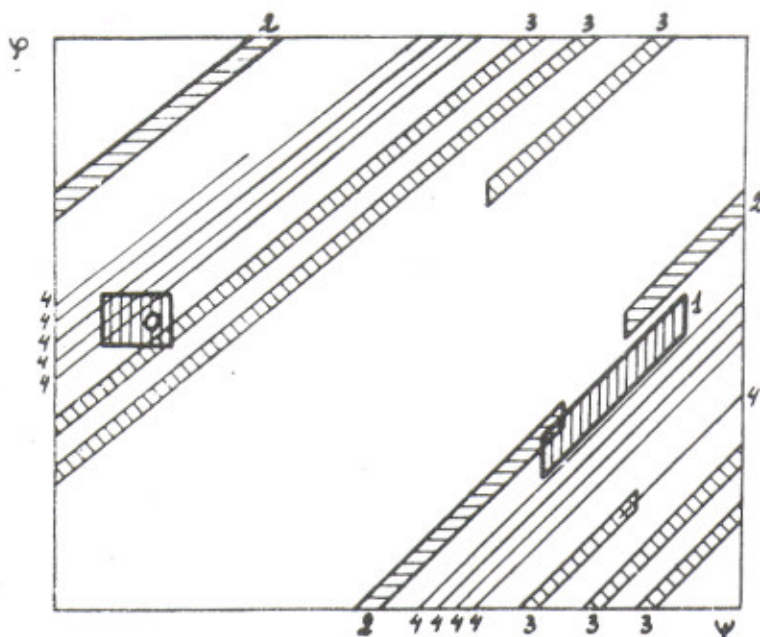
Fig. 2.4.1: Schematic picture of the mixing process for model (2.4.16) with $f(\psi) = \psi - \frac{1}{2}$. The initial region is represented by a square; the figures indicate the number of the step. The direction of the extension coincides approximately with the diagonal of the phase square, and the direction of contraction with the axis ϕ.

The general character of the mixing process here is the same as for the model of Section 2.3. However, there are also significant differences due to the fact that we are now considering a Hamiltonian system, the motion of which is reversible in time, in contrast to the model of the previous paragraph.

The first difference is due to the fact that the mixing time (n_1) depends now only on the width of the initial region along $\eta(\Delta\eta_0)$ and increases indefinitely when $\Delta\eta_0 \to 0$, while the region along ξ may be any size: $n_1 \sim -\ln (\Delta\eta_0)/\ln \lambda \to \infty$. Let us note, however, that the area of the initial region also tends towards zero together with $\Delta\eta_0$ for any finite $\Delta\xi_0$.

Another more important difference is that the initial region, which generally speaking is of a very complex structure and has a large quantity of thin layers uniformly covering the phase plane, can always be chosen such that in the process of motion it will shrink to a region of simple form. In other words, a process which is the reverse of the mixing process will take place. For this it is sufficient merely to change the time sign and

trace the reverse process. The possibility of such a process appears to be inconsistent with statistical irreversibility, and this leads to the so-called Loschmidt paradox[32]. This will be discussed (for the nth time!) in Section 2.13. Let us only note that for any initial region consisting of layers of <u>finite</u> thickness, or what amounts to the same, for any non-singular initial distribution function, the shrinking process lasts only for a finite time, inevitably changing subsequently to a process of extension [along the other eigenvector (2.4.13)] and mixing[*]. This is easy to verify by again tracing the reverse process (in time).

Thus in a stochastic system mixing always takes place asymptotically ($t \to \pm\infty$) for any direction in time! However, on the other hand one can always so choose the initial state that the reverse process takes place during any finite interval of time.

Let us return to the general case, when the value of λ and also the eigenvectors change from step to step. The direction of the eigenvectors can be obtained from (2.4.1) in the form of the ratio:

$$\frac{\delta I}{\delta \theta} = \frac{\varepsilon h_{\theta\theta}}{1 - \lambda - \varepsilon h_{\theta I}} \tag{2.4.14}$$

For small K_0 ($\lambda \approx 1 + \sqrt{K}$) the eigenvectors can always be orthogonalized and instability occurs only for $K > 0$, i.e. roughly speaking for half the phase region. It is easy to see that this region corresponds to an unstable equilibrium state.

For large K_0 ($\lambda \approx K$) the vectors $\vec{e}_\xi$, $\vec{e}_\eta$, generally speaking, are non-orthogonal (for $K < 0$). The direction of the vector $\vec{e}_\eta$ (extension) is almost constant ($\delta I/\delta\theta \approx 1/T\omega'$) and forms a small angle with the axis θ. This shows that the extension, and consequently also the development of instability (and mixing), goes mainly along the phase.

The direction of the vector $\vec{e}_\xi$ (contraction) generally speaking varies considerably as a result of the dependence of $h_{\theta\theta}$ on θ. This can lead, in principle, to the solution entering sometimes into the extension and sometimes into the contraction region, which leads to limited oscillations instead of instability. Such a situation may arise when $\vec{e}_\xi$, $\vec{e}_\eta$ are almost parallel, which corresponds to the phase values of θ near to the stable region ($h_{\theta\theta} \approx 0$). The size of this region decreases by at least $1/K_0$. Therefore, when $K_0 \gg 1$ entry into this region will take place rather seldom, so that even if the system enters the contraction region it succeeds in going over to the extension region again. When $K_0 \gg 1$ this transition takes place relatively fast (2.4.13) except for an exponentially small region near the vector $\vec{e}_\xi$. It may be considered that frequent entry into the latter region is possible only for a very special dynamical system or very special initial conditions[**].

Another important question concerns the possibility of capture in stable regions, or regions adjoining them, where the vectors $\vec{e}_\xi$, $\vec{e}_\eta$ are almost parallel. This is not possible for an autonomous system by virtue of Poincaré's recurrence theorem, valid for any

---

[*] This process is essentially development and relaxation of a big fluctuation. A bright demonstration of such a process was given by Orban and Bellemans (see Ref. 180).

[**] This question will be more closely studied in Section 2.8.

Hamiltonian system with limited motion. For a non-autonomous system, which we are now investigating, in principle capture is possible under microtron conditions with unlimited energy variation. However, the size of the region of the initial conditions corresponding to such capture is always small, since according to Liouville's theorem it cannot exceed the size of the stable region $[\sim \epsilon K_0^{-3}$ (2.4.10) and (2.4.11)$]$.

Thus we reach the conclusion that when $K_0 \gg 1$ our model system is locally unstable almost everywhere. The term "almost" signifies here the exclusion of a region that is small but of finite measure, in contrast to the ergodic theory, where it relates to sets of zero measure. This "negligible" difference unfortunately prevents the rigorous application of the latest results of the ergodic theory[17,31] to the problem under consideration. Systems of type (2.1.14) are an exception, since they have no stable regions, because of the special dependence $K(\theta)$.

For model (2.1.14) with its constant λ, $\vec{e}_\xi$, $\vec{e}_\eta$ the basic theorems[17,31] lead directly to the conclusion of stochastic motion, provided the parameter K lies in the unstable region. In fact, the demonstration of stochasticity can be extended also to the more general case of the variable λ, $\vec{e}_\xi$, $\vec{e}_\eta$, with, however, the necessary condition that $K(\theta)$ lies entirely (for all θ) outside the stable interval (2.4.7). This was recently shown by Oseledets and Sinai[150] (see Section 2.8). The proof was based on the existence of an asymptotic transverse flux (see above). However, the direction of the asymptotes now disagrees, generally speaking, with the local direction of the eigenvectors $\vec{e}_\xi$, $\vec{e}_\eta$ [compare (2.4.13)].

Since $K(\theta)$ is a periodic function, the absence of stable regions necessary for the proof of stochasticity is possible only in the case of discontinuity of $K(\theta)$ or its derivative. If this condition is not fulfilled then, according to Sinai's paper[17], in order to prove stochasticity an independent proof of ergodicity is required, or at least the existence of an ergodic component. Thus in the general case the question of the stochasticity of the basic model is still open in the sense of rigorous mathematical proof.

Another difficulty in using the results of the ergodic theory lies in the different formulation of the problem. Generally, mixing is considered in the whole region accessible for the dynamical system (for instance on the full energy surface for autonomous systems). The result of such mixing from the point of view of statistical description is a steady (statistical) state. In this paper, however, we wish to go further, and in particular to obtain the kinetic equation enabling us to trace the evolution of the statistical state of the system (Section 2.10). For this it is necessary to split the motion into "fast", which represents the mixing process, and "slow", described by the kinetic equation. For our model system the motion is fast in the phase θ. Accordingly, we need the mixing only in phase.

Let us show how the latter difficulty can be overcome by means of a new model, which will be called elementary. We will base ourselves on model (2.1.15), in which we replace $\cos 2\pi\psi$ by arbitrary function $f(\psi)$. Further, let us multiply the first equation of (2.1.15) by $T/2\pi$ and introduce a new variable:

$$\varphi_n = \left\{ \frac{T\omega_n}{2\pi} \right\} \tag{2.4.15}$$

The idea of this variable is that the variation ω_n interests us only in so far as it leads to the variation of the phase ψ_n. As a result we obtain a new transformation, describing the elementary model:

$$\varphi_{n+1} = \{\varphi_n + k \cdot f(\varphi_n)\}$$
$$\psi_{n+1} = \{\psi_n + \varphi_{n+1}\} \tag{2.4.16}$$

with a single parameter

$$k = \frac{\varepsilon T}{2\pi} \tag{2.4.17}$$

The essential difference between the new and the old model is that both the variables (φ, ψ) are now periodic, the phase plane is limited (system in a square or on a torus) and for $k \geq 1$ all the φ region, as well as the ψ region is passed through in one step, i.e. one can consider the classical problem of mixing in the whole accessible region of the phase space of the system.

Model (2.4.16) is the most simple non-trivial model of stochasticity in a Hamiltonian system. With its aid it is possible to go more or less straight over to the real physical problems. Therefore in the next paragraphs the behaviour of the elementary model will be thoroughly studied analytically (Sections 2.7 and 2.8) and by means of numerical experiments (Chapter 3).

Turning to the question of the K-entropy of the Hamiltonian system, let us first consider again the simplest case (2.1.14). Since K-entropy (2.3.14) is asymptotic (in time) like all the other quantities of the ergodic theory, only the asymptotic behaviour of the transverse flux is essential, i.e. actually, only its asymptote with extension, towards which all the other trajectories tend when $t \to \infty$. For model (2.1.14) according to (2.4.13) the asymptote is characterized by constant extension with a coefficient λ^+, where the index $+$ shows that the eigenvalue is chosen > 1, corresponding to the extension. The asymptotic motion, in this case, thus coincides with the motion of the model of Section 2.3 and this means that the K-entropy will also be the same as (2.3.10):

$$h = \ln \lambda^+ \tag{2.4.18}$$

In the general case of the variable λ, $\vec{e}_\xi$, $\vec{e}_\eta$, the K-entropy depends on the extension coefficient on the asymptote of the transverse flux λ_a^+. This coefficient, generally speaking, will be variable, since the position of the asymptote in relation to the vectors $\vec{e}_\xi$, $\vec{e}_\eta$, changes. According to Sinai[17] the entropy is equal in this case to

$$h = \langle \ln \lambda_a^+ \rangle \tag{2.4.19}$$

where the averaging is done either along the trajectory of the system, or by virtue of the ergodicity over the phase space, or to be more precise, over the ergodic component in the presence of stable regions.

For our basic model (2.1.11) the expression for the K-entropy can be simplified when $K_0 \gg 1$. According to (2.4.6) in this case:

$$\lambda^+ \approx \mathcal{K} = K_0 \cdot F(\theta) \qquad (2.4.20)$$

Further, let us note that λ_a^+ differs from λ^+ only by the factor depending on the angles between the direction of the asymptote and the vectors $\vec{e}_\xi$, $\vec{e}_\eta$. From (2.4.19) in this case we obtain:

$$h = \ln K_0 + C \rightarrow \ln K_0 \qquad (2.4.21)$$

where C is a constant ~ 1, depending on the specific form of the system. The latter expression becomes valid when $\ln K_0 \gg 1$. More accurate estimates of the K-entropy for some cases will be given in Section 3.4.

2.5 The border of stochasticity

In the previous paragraphs we have thoroughly discussed two limiting cases of very small (as compared to unity) and very large values of the parameter of stochasticity K_0. With some reservations, in the first case the motion is stable and in the second it is stochastic. The question of the position of the border of stochasticity separating the two cases naturally arises. In other words, it is a question of deciding under what conditions stochasticity arises in the system, or, on the contrary, the motion becomes stable.

Let us point out that stochasticity is the most dangerous instability of a non-linear oscillator. In fact, stochasticity means a diffusion process which makes the energy of the oscillations change, roughly speaking $\propto \sqrt{t}$ (see Section 2.10). The proportionality factor is in a sense maximal for a given perturbation (Section 2.12). The only faster process is linear resonance, in which the energy varies proportionally to t. However, for a non-linear oscillator such resonance is not possible, because the frequency of the oscillations changes with the energy. Microtron regimes, considered in the previous paragraph, are an exception, in which also the energy may vary $\propto t$. But such regimes require very special initial conditions, at least when $K_0 \gg 1$. But stochasticity takes place in a wide parameter region ($K_0 \gtrsim 1$). Hence, the border of stochasticity is at the same time a criterion for the occurrence of the most dangerous instability of non-linear oscillations.

It is not usual in ergodic theory to formulate the problems of the border of stochasticity. Although there is also in the theory the term "ergodic component" referring to the situation concerned, the approach generally adopted is to ask whether the system under consideration is stochastic. In our case, namely for systems of the non-linear oscillator type (1.1.1), the approach should be different and we should ask what the stochasticity region of the system concerned is like. We are first of all concerned with a region of values of the parameters of the system, such as the parameters of perturbation ε, non-linearity α (Section 1.3), etc. The border of stochasticity defines the critical values of these parameters, corresponding to the transition from the stable to the stochastic region. If these parameters are constants, i.e. do not depend on dynamic variables, the problem can be formulated in the classical way: is the system stochastic (in all the phase space accessible to it) for given values of the parameters?

Let us note that for an autonomous system it is sufficient to require the parameters to be constant on the energy surface, and the energy of the system can also be considered as one of the parameters.

In the general case the parameters determining the border of stochasticity or, as we shall say for the sake of brevity, the parameters of stochasticity, depend also on the dynamical variables. For instance, the parameter K depends on the phase (2.4.3). This means that the border of stochasticity divides up the phase space of the system. For autonomous systems this implies the dividing up of the energy surface, but for the sake of brevity we shall simply speak of systems with divided phase space.

In the example given above (2.4.3) the stable regions are small when $K_0 \gg 1$, and in a certain sense they can simply be ignored. However, cases are possible (see for example Section 4.1) in which the border of stochasticity divides up the phase space into regions of the same order of magnitude, so that neither of them can be neglected. This situation is similar to the action of some weak additional conservation law; unlike the standard one it does not single out a subspace of smaller dimension, but part of the phase space of the same dimension. From the physical point of view it seems completely unsound to renounce a statistical description in such cases. Consequently the problem arises of extending the ergodic theory to systems with divided phase space. The difficulties of this problem can be seen from the following rather plausible hypothesis of Sinai[*]: for systems of the type concerned the stable regions of the phase space form an everywhere dense set, which as it were penetrates (saturates) the ergodic component. Thus in a rigorous formulation of the problem the shape and even the topology of the border of stochasticity can be very complex. From the physical point of view, however, such impregnations of the ergodic component by the stable region are not of essential significance, provided that their dimensions and over-all volume are sufficiently small. Therefore, the border of stochasticity can be defined (of necessity approximately) as some intermediate zone of the phase space having a finite thickness, approximately separating the region of quasi-stability, namely stability for the majority of the initial conditions, from the region of quasi-stochasticity. Such a border can probably also be introduced in a rigorous mathematical way, i.e. with all the necessary conditions and reservations. An example of quasi-stability is discussed in Section 2.2. This is so-called "Kolmogorov stability", with the region being penetrated throughout by an everywhere dense system of unstable domains of small but finite measure (see Ref. 35); the structure of this region will be discussed more completely in Sections 2.6 and 2.7. The term "quasi" here again signifies the exclusion of regions of small but finite measure in contrast to the classical ergodic theory.

Returning to the basic model (2.1.11), it can be asserted that the border of stochasticity lies somewhere in the neighbourhood

$$K_0 \sim 1 \tag{2.5.1}$$

---

[*] For a more thorough discussion of this hypothesis, see Section 2.8.

This border, obtained for transformation (2.1.11) corresponds in order of magnitude to the criterion of stochasticity (2.1.4) for the differential equations of motion, i.e. for continuous time. Indeed, according to (2.4.8), when $K_0 \leq 1$ the quantity $\Omega_\phi^2 \sim K_0/T^2$; on the other hand it is clear that $\Delta \sim 1/T$, whence:

$$K_0 \sim s^2 \tag{2.5.2}$$

Thus the criterion of stochasticity (2.1.4) is confirmed for the special case of model (2.1.11). It will be extended to the general case of the original model (1.1.1) in the next section.

The physical border of stochasticity, defined above, in fact represents an intermediate zone of highly complex structure, as was thoroughly described above. According to the initial conditions, very different kinds of motion are possible in it: stable limited oscillations (Section 2.2), isolated ergodic components unconnected, generally speaking, with the main quasi-stochastic region (Section 2.6), and even systematic variations of the energy of the oscillations similar to linear resonance (microtron regimes, Section 2.4). The intermediate zone penetrates deeply on both sides, into the stochastic region in the form of narrow stable regions (Section 2.8) and into the region of Kolmogorov stability in the form of thin stochastic layers (Section 2.6). Nevertheless, it can be asserted that estimate (2.5.1) defines some real physical border, the border of strong stochastic instability of non-linear oscillations.

This is the main conclusion of this paper. It is completely confirmed by numerical experiments, i.e. by numerical integration of the equations of motion of very different systems (Chapters 3 and 4).

2.6 The stochastic layer in the vicinity of the separatrix

This section will be devoted to a closer study of the structure of the region of Kolmogorov stability (Section 2.2) and at the same time to extending the criterion of stochasticity for the basic model (2.5.1) to the general case of overlapping of the resonances. As noted above, the KAM theory (Section 2.2) establishes the stability only of "good" invariant tori. A "good" torus means one that is non-resonant and located "far away" from all "bad", i.e. resonant[*] tori. The term "far away" may be bewildering, since the system of resonant tori, generally speaking, forms an everywhere dense set throughout the phase space. The answer is that the term "far away" relates to the width of the resonance. The fundamental result of the KAM theory is precisely just that, roughly speaking, it shows that the total width of all the resonances becomes arbitrarily small when the perturbation is sufficiently small.

Although the KAM theory does not deal with the behaviour of the system in the vicinity of the resonant tori (it simply excludes these regions), it enables us to conclude that the motion in these regions is unstable. This conclusion can be drawn by comparing its results with Poincaré's theorem[36] [**]. The latter maintains that under very general conditions[†]

---

*) With the resonance relation of the oscillation frequencies: $\Sigma_k n_k \omega_k = 0$, n_k are integers (see Section 2.12).

**) See also Ref. 49 (Chapter 14, Section 2).

†) For example, systems with separable variables are an exception.

a Hamiltonian system has no other <u>analytical</u> integrals of motion but the energy (or more precisely all the additive integrals: energy, momentum and angular momentum). Comparison shows that non-analytical (in the dynamical variables) integrals of motion may exist, which are destroyed in the vicinity of the resonant tori.

This conclusion may appear strange, since we saw (Sections 1.4 and 2.2) that just near the resonance there are stable phase oscillations. In fact, the region near the resonance can be studied by means of the same KAM theory[*], applying it to the phase oscillation equations. It turns out that a large part of this region is stable. Then where are the unstable regions? Poincaré already noted[**] that a likely place was the neighbourhood of the separatrix (Section 1.4). Apparently the first detailed investigation of the neighbourhood of the separatrix was made by Mel'nikov[37] who, however, was not able to estimate the width of the unstable region. Such an estimate was made for the first time for a special dynamical system by Zaslavsky, Sagdeev and Filonenko[38].

Below, an estimate is given of the width of the stochastic layer in the vicinity of a non-linear resonance separatrix under very general conditions. In fact, the only essential condition is that the separatrix must pass through the hyperbolic fixed point, i.e. the point of unstable equilibrium, at which both the velocity and the acceleration vanish. This condition can be violated only for a singular phase oscillation potential $U(\psi)$. When the above condition is fulfilled, the frequency of the phase oscillations $\Omega \to 0$ as it approaches the separatrix, and the oscillations become anharmonic and non-linear. In particular, the velocity of the motion during a great part of the period is near to zero (the system is almost motionless near the point of unstable equilibrium) and substantially increases only in an interval of time $\sim \Omega_\phi^{-1}$, where Ω_ϕ is the frequency of the small phase oscillations. This means that the effective action of the perturbation on such oscillations will also be limited by the interval $\sim \Omega_\phi^{-1}$ and consequently when $\Omega \to 0$ the perturbation may be represented as a δ-function. Thus the oscillations near the separatrix are described by our basic model (2.1.11), with the sole difference that it is now necessary to take the variable half-period of the oscillations π/Ω as one step of the transformation. This means that we can directly use the criterion of stochasticity (2.5.1) or, more conveniently, its equivalent (2.1.4).

Let us assume that the perturbation is characterized by a force μf_ω with a frequency ω. The system of resonances will now be determined by the spectrum of the oscillations themselves. This contains frequencies $k\Omega$ and has the form of a δ-function's spectrum up to frequencies $\sim \Omega_\phi$, and then decreases exponentially[†]. This follows directly from the uncertainty relation for frequency and time: $\Delta\omega \cdot \Delta t \sim 1$. The amplitude of the velocity harmonics can most easily be estimated from the normalization condition (Parseval's equation):

$$v_k \sim v_\phi \cdot \frac{\Omega}{\Omega_\phi} \cdot e^{-C_o \frac{k\Omega}{\Omega_\phi}} \tag{2.6.1}$$

---

*) With some modification[77,78,176].

**) Ref. 36, see also footnote on p. 179 of Arnold's review[20].

†) For analytical $U(\psi)$, see, for example, Ref. 20.

where $\dot{v}_\phi$ is the maximum value of the velocity and is approximately constant in the vicinity of the separatrix; the parameter $c_0 \sim 1$ depends on the form of the separatrix: for example, for a harmonic potential (1.3.16) we have:

$$U_k = \left(4v_\phi \cdot \frac{\Omega}{\Omega_\phi} \right) e^{-\frac{\pi k \Omega}{2\Omega_\phi}} \qquad \text{(see Ref. 21)}$$

Let us first consider the case when the perturbation frequency lies in the main part of the spectrum (2.6.1), i.e. $\omega \sim \Omega_\phi$. The resonance condition has the form:

$$\Omega_k \approx \frac{\omega}{k} \qquad (2.6.2)$$

The essential difference between this system of resonances and that considered earlier, for example, for the basic model (2.1.11) is that now the distance between the frequencies of the spectrum (Ω) is not equal to the distance between the resonant values of the frequency $(\Delta = \Omega_k - \Omega_{k-1} \approx d\Omega/dk)$, which enters into the criterion of stochasticity and which as usual we will call the distance between resonances [see (2.1.2)]. The latter is considerably smaller:

$$\Delta \approx \frac{\omega}{k^2} \approx \frac{\Omega^2}{\omega} \qquad (2.6.3)$$

It remains for us to consider the second order phase oscillations, which arise owing to the action of the perturbation μf_ω on the main phase oscillations (of first order), and to estimate the size of their separatrix. The most simple is to use relation (1.4.3), where εU_0 is now $\sim \mu f_\omega k v_k / \Omega$, which follows from the definition of U (1.3.4) and from the fact that the perturbation of the Hamiltonian in our case is equal to: $H_{1k} \sim \mu f_\omega v_k / \Omega$. As a result we obtain an estimate of the border of stochasticity $(\Omega \sim \Omega_\phi \sim \omega)$:

$$s^2 \sim \left| \frac{\mu f_\omega v_\phi \Omega' \omega^2}{\Omega_\phi \Omega^4} \right| \sim 1 \qquad (2.6.4)$$

where the derivative Ω' is taken with respect to the action. It can be seen that there is always a stochastic region near the separatrix, since $\Omega \to 0$, and $\Omega' \to \infty$ (see below). This region is situated practically symmetrically on both sides of the separatrix, since by virtue of the periodicity of the potential $U(\psi)$ the "external" and "internal" phase oscillations are almost identical near the separatrix. By virtue of the aforementioned approximate symmetry, the second order resonances lie not only inside the first order resonances, i.e. inside their separatrix, but also outside it, in the immediate vicinity of the resonance. For the sake of brevity we will henceforth keep to the term "inside" when referring to the above situation. Similarly, the third order resonances lie inside the second order resonances and so on. We thus obtain a hierarchy of resonances, also described by Greene[47].

The formation of a stochastic layer in the vicinity of the separatrix is thus due to the overlapping of second order resonances, although the parameters of only first order resonances appear in the final equations (2.6.12) to (2.6.17). The parameters of resonances of higher orders may prove to be important when calculating the diffusion rate (see Section 2.10).

In order to obtain a more definite estimate let us make the natural assumption that the potential energy of the phase oscillations near the hyperbolic point ($\psi = 0$) has the form $U(\psi) = - m\Omega_1^2\psi^2/2$, where m is a mass, and Ω_1 is a constant representing the inverse time of the exponential drift of the system away from the point of unstable equilibrium. It is easy to obtain the asymptotic expression ($W \to 0$): $\Omega \approx \pi\Omega_1/\ln|A/W|$ [*)]. It is convenient to choose the constant A so that $\Omega(W_\phi) = \Omega_\phi$:

$$\Omega \approx \frac{\pi\Omega_1}{\ln\left|\frac{W_\phi}{W}\right| + \frac{\pi\Omega_1}{\Omega_\phi}}$$

(2.6.5)

Here W is the energy of the oscillations near the separatrix, and W_ϕ is the energy of the small oscillations, and both energies are measured from the separatrix. For the non-linearity of the oscillations we find:

$$\Omega' \approx \frac{\Omega^3}{\pi\Omega_1 W_\phi} \cdot e^{\pi\Omega_1\left(\frac{1}{\Omega} - \frac{1}{\Omega_\phi}\right)}$$

(2.6.6)

Let us now fix the small perturbation parameter μ so that $f_\omega v_\phi \sim \Omega_\phi W_\phi$ [**)]. This means that when $\mu \sim 1$ the energy of the oscillations changes considerably after one period ($\Omega \sim \Omega_\phi$). Inserting expression (2.6.6) in (2.6.4) we obtain the following estimate of the width of the stochastic layer along the separatrix in units of phase oscillation frequency ($\mu \ll 1$) [†)]

$$\Omega_c \approx \frac{\pi\Omega_1}{\ln\frac{1}{\mu} + \frac{\pi\Omega_1}{\Omega_\phi}}$$

(2.6.7)

Here we have preserved the sign for approximate equality (instead of the one for a rough estimate), since the indeterminate factor (~ 1) in the criterion of stochasticity (2.6.4) is found in (2.6.7) under the logarithm.

It is more natural perhaps to take the energy width of the stochastic layer which is equal to (2.6.7):

$$\left|\frac{W_c}{W_\phi}\right| \sim \mu$$

(2.6.8)

The width of the stochastic layer in units of the action I is also of the same order. Here we neglect the disparity of the frequencies $\Omega \sim \Omega_\phi \sim \Omega_1 \sim \omega$. For the case when $\omega \ll \Omega_\phi$ see below.

---

*) We assume that there are two identical hyperbolic points. In the opposite case one should put: $1/\Omega_1 \to (1/\Omega_1 + 1/\Omega_2)/2$; the non-hyperbolic stop point corresponds to $\Omega_2 = \infty$.

**) Here and below we assume that the amplitude of the phase oscillations $\psi_0 \sim 1$.

†) A comparison with Ref. 38 and with the results of the numerical computation is given in Section 4.2.

It is clear that the estimates obtained remain valid for any oscillator having a separatrix. Now, however, we are interested in the non-linear resonance separatrix, for which we can even further specify the above-mentioned estimates.

If there is a single resonance, the only perturbation will be the non-resonant harmonic (Section 1.3) with a frequency of $\omega \sim \Omega_\phi/\sqrt{\epsilon}$ (Section 1.4). This frequency lies far away in the "tail" of the spectrum (2.6.1) and therefore there appears in (2.6.4) an additional factor $\sim e^{-c/\sqrt{\epsilon}}$ ($c \sim 1$), which can be included in the parameter μ, by putting:

$$\mu \sim e^{-c/\sqrt{\epsilon}} \tag{2.6.9}$$

The pre-exponential factor here is ~ 1, since the parameter is determined by means of (see above): $\mu \sim f_\omega v_\phi/\Omega_\phi W_\phi \sim f_\omega/\Omega_\phi^2 \sim 1$; $f_\omega \sim \epsilon$; $\Omega_\phi \sim \sqrt{\epsilon}$. As a result we obtain the following estimate of the size of the stochastic layer caused, if one may so express it, by a self-perturbation, i.e. by the same perturbation that is responsible for the formation of the separatrix ($\epsilon \ll 1$):

$$\delta_\Omega = \frac{\Omega_c}{\Omega_\phi} \approx \frac{\pi \Omega_1/\Omega_\phi}{\frac{c}{\sqrt{\epsilon}} + \frac{\pi \Omega_1}{\Omega_\phi}} \approx \frac{\pi \Omega_1}{\Omega_\phi} \cdot \frac{\sqrt{\epsilon}}{c} ; \quad \frac{\Omega_c}{\omega} \sim \epsilon \tag{2.6.10}$$

expressed in terms of frequency or:

$$\delta_w = \left| \frac{W_c}{W_\phi} \right| \sim e^{-\frac{c}{\sqrt{\epsilon}}} \tag{2.6.11}$$

expressed in terms of energy.

This width is very small and agrees in order of magnitude with the splitting up of the separatrix (far away from the hyperbolic points) obtained by Mel'nikov[37] (see also Ref. 21). Hence it follows that the tongues of the split-up separatrix, the length of which increase infinitely as they approach the hyperbolic point, spread along the unperturbed separatrix and the stochastic region splits up into increasingly thin layers. This is a typical mixing process, similar in structure to that described in Section 2.4 for the elementary model (see Fig. 2.4.1).

Our result (2.6.11) agrees with Ref. 37 in the sense that it can be concluded from the latter that the width of the stochastic layer in any case is not smaller than (2.6.11). From our estimates it can be concluded that it is also not greater.

Let us now turn to the more interesting case when there are several resonances. First let the system of resonances be determined by the perturbation, the spacing $\Delta\omega$ between resonances and their width being of the same order. From the general expansion (1.3.2) it can be seen that the nearest non-resonant perturbation in this case has a frequency $\omega_1 = \Delta\omega$. Since in estimate (2.6.11) $\sqrt{\epsilon} \sim \Omega_\phi/\omega$, we now obtain a new estimate [see (2.6.1)]:

$$\delta_w \sim \mu \sim e^{-c_o \frac{\omega_1}{\Omega_\phi}} \tag{2.6.12}$$

But the stochasticity criterion (2.1.4) follows directly from this. It is essential, however, to have two different criteria. The estimate (2.6.12) shows that when the condition $\Omega_\phi \sim \omega_1$ is fulfilled the resonant region is almost completely destroyed, both inside and outside the separatrix, i.e. the width of the stochastic layer becomes of the order of the width of the resonance. The criterion (2.1.4) characterizes the overlapping of neighbouring resonances. When both criteria are simultaneously fulfilled this ensures the formation of a wide stochastic region, determined by all the resonances.

Now let the system of resonances be determined by the oscillator itself as in the motion near the separatrix which has just been considered. Taking into account that $\omega_1 = \Omega$ (1.3.2); $\Delta \approx \Omega/k$ (2.6.3) and $(\Delta\omega)_H \sim \Omega_\phi/k$ [(2.1.2), Section 1.4], we find that both criteria [(2.1.4) and (2.6.12)] again agree: $(\Delta\omega)_H/\Delta \sim \Omega_\phi/\omega_1$.

The two limiting cases considered above are characterized by the presence of a single perturbation or oscillation harmonic. It is clear that in itself a harmonic (sinusoidal) form is within some limits unimportant (for further detail see Sections 2.7 and 2.8). What is important is the structure of the resonance spectrum, which in both cases can be called locally equidistant. The essential property of this structure is the finite (non-zero) distance between resonances. The general case of a discrete spectrum of resonances that is everywhere dense will be considered in the next section.

Thus, on the basis of the properties of the special model (2.1.11), we validated the stochasticity criterion (2.1.4) for a system with a locally equidistant spectrum of resonances. The most simple case of such a spectrum is a pair of resonances of the same order of width. According to the criterion (2.6.12) this is already sufficient for obtaining the stochastic layer inside the resonant region, i.e. of a width of $\sim \sqrt{\varepsilon}$ [*)].

Now we can estimate the relative fraction (δ) of the stochastic component in the region of Kolmogorov stability. Since (2.6.12) gives the width of the stochastic layer in relation to the width of the resonance, in order to obtain the required estimate it is sufficient to multiply (2.6.12) by $s = (\Delta\omega)_H/\Delta$; in both the limiting cases considered above we obtain:

$$\delta \sim s \cdot e^{-c/s} \qquad (2.6.13)$$

where the stochasticity parameter (2.1.2) $s < 1$ in the region of Kolmogorov stability.

Let us have a closer look at the simplest case of two resonances, mentioned above. Let us first of all ascertain how the mutual destruction of the resonances changes if their width is substantially different. The perturbation parameter μ will in this case contain an additional factor (see p. 52) $f_p/f_q \sim (\Omega_p/\Omega_q)^2$, where the index p relates to the destroying resonance and the index q to the one that is destroyed. The frequency of the phase oscillations in estimate (2.6.12) characterizes the destroyed resonance: $\Omega_\phi \to \Omega_q$, and the minimum perturbation frequency

---

*) This conclusion was recently verified by means of a numerical experiment[48] and by "real" experiments[181].

$$\omega_1 = \bar{m}ax\left(\Omega_p, \Omega_q\right) \tag{2.6.14}$$

on the border of overlapping. The special case when a weaker resonance is entirely inside a stronger one, so that ω_1 is substantially less than (2.6.14), will be considered below. Estimate (2.6.12) now takes the form:

$$\delta_{pq} \sim \left(\frac{\Omega_p}{\Omega_q}\right)^2 \cdot e^{-c \cdot \frac{\omega_1}{\Omega_q}} \tag{2.6.15}$$

Taking into account (2.6.14) it will be seen that the most stable is the weak resonance, for which the (absolute) width of the stochastic layer is exponentially small: $\Omega_q \cdot \delta_W \sim (\Omega_p^2/\Omega_q) \cdot e^{-c\Omega_p/\Omega_q}$; $\Omega_p \gg \Omega_q$, while for the strong resonance ($\Omega_p \ll \Omega_q$): $(\Omega_q \delta_W) \sim \Omega_p^2/\Omega_q$. It is essential, however, for the destruction of the strong resonance also to be only negligible[*]. Therefore, in the case under consideration the relative fraction of the stochastic component proves to be small ($\sim \Omega_p/\Omega_q$) even under conditions when the resonances overlap. Nevertheless, owing to the overlapping of the stochastic layers of neighbouring resonances some diffusion from one resonance to another is possible, although its rate may be very small (Sections 2.7 and 3.3).

This example shows the difference between the two criteria of stochasticity particularly clearly: the criterion of the overlapping of the resonances (2.1.4), which determines the possibility of some diffusion for part of the initial conditions, and the criterion of the destruction of the resonances (2.6.15), which determines the formation of a continuous or, more precisely, almost continuous (Section 2.8) stochastic region with a maximum diffusion rate (Section 2.10).

Now let a few neighbouring resonances almost coincide: $\omega_1 \ll \Omega_p \sim \Omega_q$. Then we can consider them as one resonance with slowly changing parameters: $U(\psi, \lambda)$ and the characteristic time of variation of λ is $\sim 1/\omega_1$. The effectiveness of such perturbation is determined by the accuracy of conservation of the adiabatic invariant. The latter always breaks down near the separatrix where the phase oscillation frequency passes through zero. The width of the stochastic layer in this case may be shown[183] to be of the order

$$\delta_W \sim \frac{\omega_1}{\Omega_q} . \tag{2.6.16}$$

We shall call the formation of a single resonance from a group of almost coincident resonances renormalization of the resonances. It is seen from (2.6.16) that a continuous limit transition takes place when $\omega_1 \to 0$. Stochastic destruction of a narrow group of resonances as a function of the perturbation reaches the maximum (full destruction) when $s \sim 1$ [**]. At the same time, when there is strong overlapping of a wide group of resonances (much wider than the renormalized resonance) a system of renormalized resonances forms, for which the condition: $s' \sim 1$ is automatically fulfilled.

---

[*] This effect can be used for stabilization of stochastic instability by an additional strong resonance. The stable region appears inside the separatrix of this resonance.

[**] This and other aspects of stochastic destruction of non-linear resonances have been investigated in detail experimentally by Kulipanov, Mishnev and Skrinsky[181].

The application of the simple estimates of the width of the stochastic layer obtained above [(2.6.12), (2.6.13), and (2.6.15)] requires some caution. In fact they are based on estimates for μ of the type of (2.6.9), which takes into account only the frequency spectrum of the perturbation. This is certainly true if there is only one perturbation harmonic (two resonances). In the case of several harmonics it is necessary to take into account their phase relations which, in particular, may considerably reduce the value of μ as compared to the above-mentioned estimates. The simplest example is the basic model (2.1.11) when $T \to \infty$. In spite of the strong overlapping of the resonances in this case (s >> 1) the motion will be stable during each of the intervals T between kicks. This occurs precisely owing to the special phase relation of the resonances. A more complex example of the effect of phase relations will be considered in Section 2.9.

2.7 Full set of resonances

So far we have considered the interaction of an approximately equidistant set of reso-nances, formed owing to the anharmonicity of either the perturbation alone (basic model) or the oscillations themselves (separatrix, Section 2.6). In both cases the stochasticity criterion had a fully defined sense, since the mean distance between resonances Δ remained finite (2.1.2).

In the general case a complete set of resonances is dense in frequency, so that formally $\Delta = 0$. Physically it is clear that the amplitudes of the high harmonics, generally speaking, rapidly decrease with the increase of the harmonic number (for an analytical func-tion -- exponentially). Therefore a finite number of harmonics actually works and this means also a finite number of resonances. The more accurate result is that the total width of all the resonances is finite and small (for sufficiently small perturbation). As already noted above, this is also the main result of the KAM theory. However, the technical dif-ficulties of constructing convergent series in this theory lead to excessive requirements for smoothness of the functions entering into the equations of motion (smoothness of force, as we shall say from now on for the sake of brevity). Originally the analyticity of the force was even assumed[19,20], although it was perfectly clear that this was simply a technical requirement[39].

Moser has recently developed a technique for "smoothing" non-analytical functions, namely approximating them by the sequence of analytical ones[28]; as a result it turned out to be sufficient to require the existence of a number (L_c) of continuous derivatives of the force. The minimum value of L_c obtained by Moser is[28]:

$$L_c > 2N + 2 \tag{2.7.1}$$

where N is the number of degrees of freedom of the autonomous Hamiltonian system. For a non-autonomous system and also for transformations, there are no estimates; as far as can be understood from Ref. 28, in this case one should put: $N \to N + 1$. Moser's result gives essentially the upper (sufficient) limit of smoothness of the force, since it is determined by smoothing technique.

In this paragraph we will try to give some estimates of the lower limit of smoothness of the force necessary for Kolmogorov stability. It is assumed that an effective estimate

of the stability conditions, taking into account the complete set of resonances, can be given in first approximation (see below). Going on ahead, let us say that this assumption is made not only to simplify the task, but also on the basis of the result of numerical experiments (see Section 3.3).

The value of this hypothesis for obtaining practical estimates is evident, it simply eliminates the need to calculate higher approximations, not to mention questions of convergence. Let us note that the estimates of this paragraph are not equivalent to the first approximation of the KAM theory[20], since the size and the other characteristics of the resonances are taken into account and not simply excluded.

We will limit ourselves in the main to the elementary model (2.4.16), which will be written in the form:

$$\omega_{n+1} = \left\{ \omega_n + \varepsilon \cdot f(\psi_n) \right\}$$
$$\psi_{n+1} = \left\{ \psi_n + \omega_{n+1} \right\} \tag{2.7.2}$$

where $f(\psi)$ is a certain function ("force") which we will define more precisely later, and ε is the small parameter.

The main resonances, to which we have so far restricted ourselves, lie at $\omega = r$ (r is an integer) and correspond to the fixed point of transformation (2.7.2). It is not difficult to see that in the general case the resonance takes place for any rational value of ω:

$$\omega = \frac{r}{q} \tag{2.7.3}$$

Indeed, under this condition the phase ψ changes by exactly r periods after q steps. These higher harmonic resonances ($q > 1$), as they will be called, thus correspond to the periodic motion of system (2.7.2) with a period q.

The resonance condition (2.7.3) becomes especially clear if one changes over from transformation (2.7.2) to the differential equations, i.e. to continuous time:

$$q\omega = r\Omega \tag{2.7.4}$$

where Ω is the basic frequency of the perturbation. Then the resonance (r,q) is the resonance of the rth harmonic of the perturbation with the qth harmonic of the oscillations.

For what follows it is important to understand that the high harmonics occur for two completely different reasons. First of all, owing to the anharmonicity of the force as a function of the coordinate:

$$f(\psi) = \sum_q' f_q \cdot e^{2\pi i q \psi} \tag{2.7.5}$$

The resonances thus arising will be called first approximation resonances or higher harmonic resonances. Their width is determined by the coefficients f_q, which can be obtained without any fundamental difficulties.

However, there is also another reason for the occurrence of higher resonances, even for $f(\psi) = \sin 2\pi\psi$, when $f_q = 0$ ($q > 1$). This is as follows. The resonant frequencies (2.6.4) are obtained from a Fourier expansion in time, whereas only the Fourier expansion according to the coordinate (2.6.5) is easy to obtain. But the phase ψ does not vary strictly in proportion to time, since the frequency ω in its turn varies under the action of the perturbation, particularly under the action of the non-resonant harmonics (Section 1.3). It is easy to see that modulation frequencies $\omega + r\Omega$ appear in the first approximation. This leads to second approximation resonances of the form: $2\omega = r\Omega$, and the amplitude of such resonant terms is $\sim \varepsilon^2$. Similarly it can be shown[3] that the resonances $q\omega = r\Omega$ are defined by terms $\sim \varepsilon^q$ decreasing exponentially with q. This gives grounds for hoping that the influence of the higher approximation resonances will be unimportant.

In reality, however, the question is a highly complex one. A more accurate investigation[20] shows that the terms of the qth approximation $\sim q^{\ln q} \cdot \varepsilon^q$. In the case of a non-analytical force with a power-law spectrum of the (2.7.6) type, this may lead to divergence for very high harmonics. In fact, however, divergence does not occur, as was shown by Moser[28].

It will be assumed that somehow or other the total (actual) width of the resonance can be estimated in first approximation. In any event we can rely on thus obtaining the lower limit of smoothness of the force necessary for Kolmogorov stability.

Bearing in mind the comparison with Moser's result[28], let us choose as $f(\psi)$ a function whose $(\ell + 1)^{th}$ derivative undergoes discontinuity of the order of unity. It is easy to see that the asymptotic (when $q \gg 1$) spectrum of such a function is given by the expression:

$$f_q \sim q^{-(\ell+2)} \tag{2.7.6}$$

Let us consider some resonance (r,q). Ignoring the non-resonant terms the transformation (2.7.2) near the resonance (r,q) can be written in the form:

$$\omega'_{n+1} = \left\{ \omega'_n + \varepsilon q^{-(\ell+1)} \cdot e^{2\pi i \psi'_n} \right\} \tag{2.7.7}$$

$$\psi'_{n+1} = \left\{ \psi'_n + \omega'_{n+1} \right\}$$

where we changed the variables: $\omega' = q\omega$; $\psi' = q\psi$. In the first approximation transformation (2.7.7) has a single resonance (basic, $q' = 1$), the width of which according to (2.2.4) is $(\Delta\omega')_H \sim \sqrt{\varepsilon} \cdot q^{-(\ell+1)/2}$ or in the variable ω:

$$(\Delta\omega)_q \sim \sqrt{\varepsilon} \cdot q^{-\frac{\ell+3}{2}} \tag{2.7.8}$$

For a given value of q there are q different resonances, corresponding to r = 1, 2, ..., q (2.7.3). Simple summation of the width of all the resonances gives:

$$(\Delta\omega)_{\Sigma} \sim \sqrt{\varepsilon} \cdot \sum_{q=1}^{\infty} q^{-\frac{\ell+1}{2}} \tag{2.7.9}$$

This sum converges if $\ell > 1$ [*].

However, even in the frame of a first approximation simple summation of the width of the resonances is not really justified. The point is that many resonances coincide or almost coincide, or rather fall inside one another. The total (renormalized, see Section 2.6) width of such coincident resonances will, generally speaking, be smaller than the sum of the width of the separate resonances. The summation rule (renormalization) depends on the phase relations. If all the resonances are "in phase" the $(\Delta\omega)_q^2$ proportional to the amplitude of the perturbation harmonic accumulate; for "random" phases it is necessary to sum up the $(\Delta\omega)_q^4$. Apparently the latter case is nearer to reality, since the majority of resonances do not coincide exactly and the phase relations vary with time. It turns out that the convergence of the resonance sum does not depend on the power of renormalization, which is denoted by n and left arbitrary for the present (see below).

Thus we estimate the sum width of all the resonances with $q' > q$ coinciding with one of the q resonances, i.e. falling inside the resonant region $(\Delta\omega)_q$ (2.7.8). It is clear that out of q' resonances of the q'^{th} harmonic only $q_1' = (\Delta\omega)_q \cdot q'$ will coincide, on the average, with the resonance q. We have:

$$(\Delta\omega)_{q\Sigma}^n = (\Delta\omega)_q^n + \sum_{q'=q+1}^{\infty} q_1' \cdot (\Delta\omega)_{q'}^n \approx$$
$$\approx (\Delta\omega)_q^n \left(1 + \frac{\sqrt{\varepsilon} \cdot q^{(1-\ell)/2}}{n\ell + 3n - 4} \right) \tag{2.7.10}$$

where $(\Delta\omega)_{q\Sigma}$ is the renormalized width. The sum converges if:

$$\ell > \frac{4}{n} - 3 \tag{2.7.11}$$

Renormalization is unimportant when $\varepsilon \to 0$, $q \to \infty$, if:

$$\ell \geqslant 1 \tag{2.7.12}$$

In the case of $\ell < 1$, $(\Delta\omega)_{q\Sigma} \gg (\Delta\omega)_q$, so that it is natural when determining q_1' to take $(\Delta\omega)_{q\Sigma}$ instead of $(\Delta\omega)_q$ and ignore the value $(\Delta\omega)_q^n$ in the right-hand side of (2.7.10). We obtain:

$$(\Delta\omega)_{q\Sigma} \approx \left(\sum_{q'=q+1}^{\infty} q' (\Delta\omega)_{q'}^n \right)^{\frac{1}{n-1}} \approx$$
$$\approx (\varepsilon q)^{\frac{n}{2(n-1)}} \cdot q^{-2} \cdot \left(\frac{2}{n\ell+3n-4} \right)^{\frac{1}{n-1}} \tag{2.7.13}$$

---

[*] A similar estimate was obtained in Ref. 46.

When summing up expression (2.7.13) for all the resonances, the coincident ones should be excluded. For this let us introduce the value G(q) -- the total width of all the gaps (in frequency) between the resonances with q' < q. Assuming that the reciprocal distribution of gaps and resonances is an uncorrelated ("random") one, we can write the equation:

$$G(q+1) = G(q) \cdot \left(1 - q \cdot (\Delta\omega)_{q\varepsilon}\right) \qquad (2.7.14)$$

When $\varepsilon \to 0$ the solution takes the form:

$$G(q) = \exp\left\{ -\sum_{q'=1}^{q-1} q' \cdot (\Delta\omega)_{q'\varepsilon} \right\} \qquad (2.7.15)$$

where we put G(1) = 1. Thus the conditions of the overlapping of the resonances are nevertheless determined by the direct sum of (2.7.13) for all the resonances. This converges when:

$$\ell > 1 \qquad (2.7.16)$$

This is also a necessary condition for the existence of Kolmogorov stability or the lower limit of the required smoothness of force. Taking into account condition (2.7.12) this limit is obtained from the most simple sum of (2.7.9).

The value of (2.7.16) is considerably smaller than the upper limit (2.7.1) which in the present case (N = 2) is: $L_c > 6$. The numerical experiments seem to testify in favour of a lower limit (see Section 3.3).

Let us investigate the case of $\ell \leq 1$ when the sum in (2.7.15) diverges and $G \to 0$ when $q \to \infty$. Nevertheless, exponentially small gaps remain for any finite q. There is doubt as to whether they really exist, for two reasons. Firstly, for this there must be a very sharp edge to the resonant region [the destroyed separatrix (Section 2.6)]. According to the KAM theory there exists in fact a border of absolute stability. However, in the neighbourhood of this border (on the separatrix side) in the general case there is a very complex transitional layer, characterized in particular by very slow development of instability (Section 3.3). Secondly, solution (2.7.15) is essentially connected with the assumption made above concerning the "randomness" of the gap distribution. This assumption is admittedly violated in two cases: if $\varepsilon \sim 1$, so that exponentially small gaps appear already when $q \sim 1$, or if resonances of one harmonic overlap. In the first case, the total overlapping of a small number of lower resonances is possible, the conditions of which can easily be obtained (when $\varepsilon \sim 1$) from (2.7.13). In the second case, total overlapping is possible with any ε, if $q \cdot (\Delta\omega)_{q\varepsilon} \to \infty$ when $q \to \infty$, i.e. if:

$$\ell < -1 + \frac{2}{n} = -\frac{1}{2} \qquad (2.7.16)$$

This condition now depends on n, which was taken here as four (see above). Assuming that $q \cdot (\Delta\omega)_{q\Sigma} \sim 1$, the minimum harmonic number that already ensures overlapping and determines the diffusion rate can be estimated; we have (n = 4):

$$q_1 \sim \varepsilon^{\frac{2}{1+2\ell}} \tag{2.7.17}$$

We used the renormalized width of the resonances (2.7.13). From expression (2.7.10) it can be seen that this is valid only for sufficiently high harmonics $q > q_\Sigma$, where

$$q_\Sigma \sim \varepsilon^{\frac{1}{\ell-1}} \tag{2.7.18}$$

It is easy to see that indeed $q_1 > q_\Sigma$ in the region of applicability (2.7.16) of expression (2.7.17).

Excluding the two special cases considered above, it can nevertheless be expected that the gap distribution G(q) will be nearly "random". This is mainly due to the fact that asymptotically (q → ∞) the position of the gaps depends essentially on the width of the resonances determining the gap, and as a rule these will be resonances of different harmonics having a different width. It is essential also for the distribution of the resonances (2.7.3) to be asymptotically uniform (see below).

To sum up, we reach the conclusion that it is apparently not possible completely to exclude the existence in the phase plane of gaps for any q, if

$$-\frac{1}{2} < \ell \leq 1 \tag{2.7.19}$$

These gaps, in principle, can completely stop diffusion, in spite of the absence of Kolmogorov stability. The results of the corresponding numerical experiments and the subsequent discussion are given in Section 3.3.

Let us now verify the criterion of destruction of the resonant region, which in the general case of resonances of a different width can be obtained from estimate (2.6.15). First of all let us make this estimate a little more accurate, taking into consideration the fact that for resonances of different harmonics p and q we have the relation $f_p/f_q \sim (q/p) \, (\Omega_p/\Omega_q)^2$, whence:

$$\delta_{pq} \sim \frac{q}{p} \left(\frac{\Omega_p}{\Omega_q} \right)^2 \cdot e^{-\frac{\omega_1}{\Omega_q}} \tag{2.7.20}$$

Let us note that when $\ell > -3/2$, namely in practice in all cases of interest to us, this more accurate expression does not change the character of the estimate (2.6.15). Let us recall that the index p relates to the destroying resonance and q to the one destroyed; Ω_q is the width of the resonance, equal to $q \cdot (\Delta\omega)_q$, and as $(\Delta\omega)$ it is necessary to use expression (2.7.8) or (2.7.13) depending on the value of ℓ.

From estimate (2.7.20) it can be seen that mutual destruction is possible only for resonances of close harmonics as was thoroughly demonstrated in Section 2.6. In particular, for the power-law spectrum (2.7.6) the following condition should be fulfilled:

$$|p - q| \ll q \tag{2.7.21}$$

When this condition is fulfilled the criterion of mutual destruction of the resonances can be written in the form[*]:

$$S_1 = \frac{\Omega_q}{\omega_1} = \frac{\Omega_q}{|p\omega - k|} \geq 1 \tag{2.7.22}$$

Here ω is the frequency of the system near the destroyed resonance (r,q), but not necessarily exactly equal to r/q, since what interests us is strong destruction of the resonance and stochasticity; (k,p) is the destroying resonance complying with condition (2.7.21).

Let us estimate the denominator (2.7.22). The lower estimate may be taken from Moser's paper[28]:

$$|p\omega - k| > \frac{C_1}{p^\tau} ; \quad \tau > 1 \tag{2.7.23}$$

Noting further that the minimum value of interest to us $\omega_1 = p\omega - k_1 = \{p\omega\}$, when $p = 1$, $2, \ldots$, forms a sequence which is ergodic for any irrational ω [33], we obtain the upper estimate:

$$|p\omega - k_1| < \frac{C_2}{p^\tau} ; \quad \tau < 1 \tag{2.7.24}$$

Comparing (2.7.23) and (2.7.24) we see that there is an effective estimate:

$$|p\omega - k| \sim \frac{1}{p} \tag{2.7.25}$$

and the numbers of p,q can be chosen close together, if they are large enough; this closeness enters into the constant $c_2 \sim p/|p - q|$.

Fulfilment of the criterion (2.7.22) depends now on the asymptotic behaviour of Ω_q when $q \to \infty$. For both cases (2.7.8) and (2.7.13) $s_1 \sim q^2 \cdot (\Delta\omega)_q \to \infty$, if:

$$\ell < 1 \tag{2.7.26}$$

Thus for a complete set of resonances the criterion of destruction of the resonant region also agrees in order of magnitude with the criterion of their overlapping, if one does not consider the possible formation of the gaps mentioned above. For the criterion (2.7.22) such gaps are completely unimportant, since the value of ω_1 is determined, roughly speaking, by the distance between the centres of the resonances and not between their separatrices.

The only case in which there is a considerable difference between conditions (2.7.26) and (2.7.19) corresponds to $\ell = 1$. In this border case the resonances overlap, but the

---

[*] In accordance with the observation in the previous paragraph (Section 2.7) for the re-normalized set of resonances $s_1' \leq 1$ always.

width of the stochastic layers is exponentially small. Consequently the total area of the stochastic component, and also the diffusion rate, are negligibly small when $\varepsilon \to 0$.

Assuming that $q^2 \cdot (\Delta\omega)_{q\Sigma} \sim 1$, we can find the boundary of destruction of the resonances in q, which, it turns out, coincides with the renormalization boundary q_Σ (2.7.18). From estimate (2.7.20) it follows that the separatrices of the lower resonances $(q < q_\Sigma)$ are negligibly destroyed. Since the total width of the undestroyed resonances $(q < q_\Sigma)$ is just ~ 1, the stable regions occupy a considerable part of the phase plane. However, they are separated from each other by a thick network of interwoven stochastic layers. The scale of the mesh of this network is determined by the mean distance between the destroyed resonances and is:

$$ (\delta\omega) \sim q_\Sigma^{-2} \sim \varepsilon^{\frac{2}{1-\ell}} \tag{2.7.27} $$

The estimates obtained in this paragraph are also important for the analytical force $f(\psi)$ of a special form with sharp (in the section $\Delta\psi \ll 1$) variation of the $(\ell + 1)^{\text{th}}$ derivative. In this case the spectrum of $f(\psi)$ is a power-law one (2.7.6) up to $q_m \sim 1/\Delta\psi$. When $\ell < 1$ the previous stochasticity criterion $\varepsilon \sim 1$ (2.5.1) changes by

$$ \varepsilon \sim q_m^{\ell-1} \tag{2.7.28} $$

which is obtained from the condition: $q_\Sigma \sim q_m$.

Let us turn in conclusion to continuous time, i.e. to the differential equations instead of the transformation. The amplitude of the resonance harmonic will depend in this case not only on q but also on r (2.7.3). Let us put [compare with (2.7.6)]:

$$ f_{qr} \sim q^{-(\ell+2)} \cdot r^{-(\ell_t + 2)} \tag{2.7.29} $$

when $q, r \gg 1$. The resonance sum (2.7.9) now takes the form:

$$ (\Delta\omega)_\Sigma \sim \sqrt{\varepsilon} \sum_{q=1}^{\infty} \sum_{r=1}^{q} q^{-\frac{\ell+3}{2}} \cdot r^{-\frac{\ell_t+2}{2}} \tag{2.7.30} $$

Convergence, and hence also Kolmogorov stability, takes place under the condition:

$$ \ell > -1 \; ; \quad \ell + \ell_t > -1 \tag{2.7.31} $$

In particular, for analytical dependence on $t(\ell_t \to \infty)$ the only essential condition is the first, which is considerably weaker than the previous one (2.7.16) for the transformation. The latter is obtained from the second condition (2.7.31) if one assumes: $\ell_t = -2$ (δ-function).

2.8 Quasi-resonances

Let us now make a more detailed study of the stochastic region. A troublesome feature of this region for systems of the type of our basic model (2.1.11), a feature which puts in doubt the possibility of "real" stochasticity occurring, is the presence of "islets"

of stability, which do not vanish even for values of the stochasticity parameter $K_0 \to \infty$ (2.5.1). For the reasons given below we shall call these "islets" quasi-resonances. Our task, therefore, is to estimate the size and over-all area of these "islets". Let us again restrict ourselves to the elementary model (2.4.16):

$$\varphi_{n+1} = \left\{ \varphi_n + k \cdot f(\psi_n) \right\}$$
$$\psi_{n+1} = \left\{ \psi_n + \varphi_{n+1} \right\}$$

(2.8.1)

which was also used for the numerical experiments (Section 3.5).

The stable regions are situated near the periodic trajectories of the system. The simplest periodic solution of transformation (2.8.1) -- fixed point (period T = 1 step) -- can be stable only for special values of k (see below and Section 2.4). However, generally speaking, there exists an innumerable set of other periodic solutions with $T \to \infty$. More precisely Sinai[40] showed that a stochastic system has an everywhere dense set of periodic trajectories in the phase space. Of course, the measure of this set is equal to zero and all the periodic trajectories are unstable. The following estimate follows from Ref. 40:

$$\nu(T) \sim e^{h(T-1)}$$

(2.8.2)

where $\nu(T)$ is the number of periodic trajectories with a period $\leq T$; h is the K-entropy[*].

Our system is not stochastic in the full (classical) sense of this word because of the presence of regions of stability around part of the periodic trajectories (2.8.2). However, it can be assumed that estimate (2.8.2) does not change essentially, at least if the fraction of stable regions is sufficiently small.

Before proceeding with the estimates, let us explain the stability mechanism near the periodic solutions. As was thoroughly described in Section 2.4, for stochasticity the existence of a so-called asymptotic transverse flux is required. This means that in the vicinity of every point of the phase plane the trajectories of the transverse flux should tend asymptotically towards a particular trajectory (asymptote) any segment of which will expand exponentially in the process of motion, at least on the average. It is not diffi-cult to see that the easily proved property of local instability of motion is not sufficient to fulfil this condition. Indeed, by virtue of the conservation of the phase volume, the transverse transformation is characterized by two eigenvectors (directions), one of which corresponds to contraction and the other to extension. As a result of this, in the space of the directions of the transverse flux there are two cones [for a one-dimensional system of type (2.8.1) -- two sectors on the phase plane]: the extension cone and the contraction cone, depending on the variation of the length of the transverse vector $(\Delta\varphi, \Delta\psi)$. In the process of motion these cones may overlap, i.e. cross over into each other partly or

---

[*] When h >> 1, expression (2.8.2) also gives the number of periodic trajectories with a period of T, as can easily be verified immediately. It should be noted that we changed estimate (2.8.2) somewhat as compared to Ref. 40 (T → T - 1) in order to obtain the right asymptotic form for T = 1; h → ∞.

completely, which may lead to limited oscillations of the transverse vector instead of continuous extension. This may in turn lead to the appearance of stable regions.

Let us first of all show that for model (2.8.1) the contraction and extension cones do not overlap in the special case when there are no stable regions. This condition can be written in the form:

$$|f'(\psi)| > c > 0 \qquad (2.8.3)$$

Now, even the minimum value of the stochasticity parameter $K_1 = |k \cdot f'_{min}| > kc$ increases infinitely with k.

Using expression (2.4.14) we find for the direction of the eigenvectors (see Fig. 2.8.1):

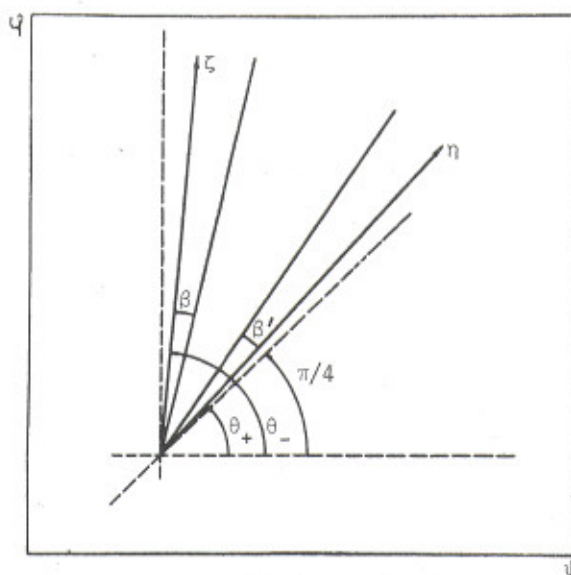$$tg\,\theta_\pm = \frac{K}{\lambda_\pm - 1} \qquad (2.8.4)$$

Fig. 2.8.1: Structure of transversal transformation for elementary model (2.8.1) in the absence of stable regions (2.8.3): $\theta_\pm$ indicate the direction of extension and contraction eigenvectors; $\bar{\beta}$, β' represent the border between contraction and extension cones (sectors) before and after transformation, respectively.

where the eigenvalues $\lambda_\pm$ are determined by formula (2.4.6). When $K_1 \gg 1$ the direction of the vector $\vec{e}_\eta$ is confined to the sector $\pi/4 \pm 2/K_1$, and that of the vector $\vec{e}_\xi$ to the sector: $\pi/2 \pm 1/K_1$. For what follows the minimum value of the angle between $\vec{e}_\xi$, $\vec{e}_\eta$, which is obviously: $\alpha_{\min} = \pi/4 - 3/K_1$, is important.

On the other hand, it is not difficult to show that under the assumed condition that $K_1 \gg 1$ the border between the contraction and extension cones, i.e. the direction for which the length of the transverse vector does not change, is at an angle $\beta \leq 1/\sqrt{2}K_1$ to the contraction axis. Under transformation, the border transverse vector swings round (without any change of length) towards the axis of extension and makes an angle of $\beta' \leq 1/\sqrt{2}K_1$ with it (see Fig. 2.8.1). The minimum angular distance between the latter direction and the new border is:

$$\alpha_{\min} - \beta - \beta' = \frac{\pi}{4} - \frac{3 + \sqrt{2}}{K_1} > 0 \qquad (2.8.5)$$

This is, of course, also the condition that the cones shall not overlap in the process of motion. This means that no transverse trajectory in the extension cone can ever enter the contraction cone. It follows that the transverse flux is asymptotic, and the motion of the system stochastic (Section 2.4).

The stochasticity criterion using the condition of overlapping of the cones, was formulated and applied to model (2.1.14) by Oseledets and Sinai (see also Refs. 42, 150).

Let us now investigate the influence of the regions of stability, not imposing any further limitation (2.8.3) on the function $f(\psi)$. In this case part of the periodic solutions of (2.8.2) may be stable, which leads to the formation of regions ("islets") of stability in the phase plane, i.e. to the appearance of a non-ergodic component.

Let us first consider the special values of the parameter k for the elementary model (2.8.1). We shall limit ourselves to the case of T = 1 (fixed point) which leads to the largest islets of stability.

For the fixed point of transformation (2.8.1) we have:

$$\varphi_0 = 0 \; ; \qquad k \cdot f(\varphi_0) = r \qquad (2.8.6)$$

where r is any integer. The fixed point is stable when (Section 2.4):

$$-4 < k \cdot f'(\varphi) < 0 \qquad (2.8.7)$$

The special values of k are determined from the compatibility [(2.8.6) and (2.8.7)]. This condition is fulfilled within the interval

$$\Delta k \approx \frac{8}{k \cdot f(\varphi_1) \cdot f''(\varphi_1)} \; ; \quad f'(\varphi_1) = 0 \qquad (2.8.8)$$

around the value k, which corresponds to the centre of the stable region (2.8.7) and is:

$$k_r = \frac{r}{f(\varphi_1)} - \frac{2}{r \cdot f''(\varphi_1)} \qquad (2.8.9)$$

The phase area of the stable region is:

$$S(k) \sim \Delta\varphi \cdot \Delta\psi \sim \left(\frac{4}{kf''}\right)^2; \quad \Delta\varphi \sim \Delta\psi \sim \frac{4}{kf''} \qquad (2.8.10)$$

It is essential for any dimension of this region to be arbitrarily small when k >> 1.

Let us note that on turning to the basic model (2.1.11) an additional factor $\Delta I/\Delta\psi = 1/T\omega' = \varepsilon/k$ appears, so that the area of the non-ergodic component becomes even smaller (see Section 2.4):

$$\Gamma(k) \sim \varepsilon/k^3 \qquad (2.8.11)$$

Let us now go over to arbitrary values of k. Supposing that:

$$w_o = \frac{8}{k \cdot f''(\psi_1)} \qquad (2.8.12)$$

is the probability of entering the stable phase region (2.8.7); here it is assumed that there are two stable regions with identical values of f''. Let us further assume that for T, k >> 1 the periodic trajectories are "randomly" situated in the phase plane. This assumption is very important for us, since in the opposite case it is very difficult to obtain any quantitative estimates. It is confirmed intuitively because we are considering an almost stochastic system.

Further reasons in favour of the above assumption can be found by considering the mechanism of the formation of a large number (2.8.2) of periodic trajectories. When k << 1 the set of first order resonances forms a set of periodic solutions $v(T) \sim T^2$ (according to a number of resonances $\varphi = r/q$; $q \leq T$). With regard to the periodic solutions connected with higher order resonances (see Section 2.6), they lie inside the first order resonances. It is therefore possible not to take them into account up to the border of stochasticity. In particular, the stochasticity criterion is determined by the first order resonances only. If k >> 1, the resonances of different orders intermix and spread more or less uniformly over the phase plane.

Several mechanisms of formation of stable periodic trajectories in a stochastic region are possible. The most simple (we shall call it the first) corresponds to the case when all the T of the points of the trajectory are in the stable phase region (2.8.7). The probability of this is w_0^T and the number of such trajectories is (2.8.2):

$$\nu_s^{(1)}(T) \sim w_o^T \cdot \nu(T) \sim w_o \left(w_o e^h\right)^{T-1} \qquad (2.8.13)$$

This estimate is very sensitive to the value of the parameter $\gamma = \omega_0 e^h$. For a "force" $f(\psi) = 1/2\pi \sin 2\pi\psi$ (2.1.15), for instance, $\gamma = 2/\pi < 1$. However, it is easy to construct $f(\psi)$ so that $\gamma > 1$ (see Sections 3.2 and 3.5). At first glance it may appear that in the latter case the fraction of stable regions will be rather considerable, since expression

(2.8.13) diverges when $T \to \infty$. However, experimentation shows that this is not the case; moreover, it turns out that for large γ the fraction of stable regions is even reduced (Section 3.5).

Two effects may produce this result which appears strange at first glance. One of them, apparently secondary, is the fact that estimate (2.8.13) is in reality the upper boundary. This is due to the fact that the periodic trajectory of the transformation, all the points of which are elliptical, is not necessarily stable. Possible instability is explained by the, generally speaking, variable frequency of the phase oscillations around the periodic trajectory, which may lead to parametric resonance.

The main effect is probably that when $\gamma > 1$, considerable overlapping of the stable regions takes place precisely because of the divergence of expression (2.8.13). But in this case we can apply the general criterion of stochasticity according to the overlapping of the resonances (Section 2.1). Indeed, the resonant region of a non-linear system signifies, essentially, a stable region of quasi-periodic motion in the vicinity of the periodic trajectory. The meaning of the overlapping of resonances as a stochasticity criterion in this connection is as follows. First of all, when the resonances overlap the trajectory of motion can cross over from one resonant region to another, i.e. it is no longer localized in the vicinity of the original periodic solution. This feature is also conserved in the case of the overlapping of quasi-resonances, as we shall call the stable regions when $k \gg 1$.

On the other hand, the interaction of neighbouring resonances leads to the formation of a stochastic layer in the vicinity of the resonant separatrix, the width of which increases as the resonances converge and covers the whole of the resonant region at the moment of overlapping (Section 2.6). Something similar also probably takes place for quasi-resonances, although at present it is not clear what is the exact form of the second criterion of stochasticity (2.6.12) and, in particular, what the quantity ω_1 corresponds to in the case of quasi-resonances. A peculiarity of quasi-resonances is that the stochastic component is located among them, and not the invariant Kolmogorov tori as in the case of ordinary resonances. This, of course, facilitates the destruction of quasi-resonances.

Another stability mechanism (the formation of quasi-resonances) is connected with the alternating entry of the transverse vector $(\Delta \varphi, \Delta \psi)$ into the extension and contraction cones. Let us recall that the stable case corresponds to the swinging round of the transverse vector and the unstable (stochastic) one to its extension or contraction.

As shown above, the extension and contraction cones of the elementary model (2.8.1) do not overlap as long as the trajectory of motion remains in the unstable region. Hence it follows that it is impossible for the transverse vector to enter the extension and contraction zones alternately. However, it becomes possible when there is even one point of the periodic trajectory lying in the stable phase region (2.8.7). In this region the transverse vector swings round and may therefore change over into the contraction region.

The transverse motion splits into three phases: extension, contraction and rotation. If the period of the basic motion is T, the duration of the contraction and extension is $(T - 1)/2$. During the extension, the angle of the transverse vector to the asymptote decreases to:

$$\beta \sim \beta_0\, e^{-2h\,(T-1)/2} \sim e^{-h\,(T-1)} \qquad (2.8.14)$$

where $\beta_0 \sim 1$. It is obvious that the rotation must be carried out with the same accuracy by order of magnitude, since the motion along the contracting asymptote is symmetrical. The probability of such rotation $\sim \omega_0 \beta T$, whence the number of quasi-resonances of the second type is[*]:

$$\nu_s^{(2)}(T) \sim \omega_0 \beta T \cdot \nu(T) \sim \omega_0 T \qquad (2.8.15)$$

The total number of quasi-resonances now proves to be infinite independently of the form of the force, and from Ref. 40 it follows that they are located everywhere densely.

The size of a quasi-resonance can be estimated as $\Delta S \sim \Delta \psi (\Delta r)^2$, where ψ, r are the polar transverse coordinates, $\Delta \psi \sim \beta$, and Δr is determined from the condition of the required accuracy of rotation: $\Delta r \sim \beta$. Consequently, even the maximum size of the stable region:

$$\Delta r \sim e^{-h\,(T-1)} \qquad (2.8.16)$$

decreases exponentially. The total area of all the T of the stable regions of the quasi-resonances is:

$$S_1^{(2)} \sim T \cdot e^{-3h\,(T-1)} \qquad (2.8.17)$$

and for all quasi-resonances of the type under consideration:

$$S^{(2)}(k) \sim \sum_{T=T_1}^{\infty} \omega_0 T^2 e^{-3h\,(T-1)} \qquad (2.8.18)$$

The lower limit of summation for an arbitrary k is determined from the condition that $\nu_s^{(2)}(T_1) \sim 1$, i.e. that the quasi-resonances really exist:

$$T_1 \sim (2/\omega_0)^{1/2} \qquad (2.8.19)$$

Whence:

$$S^{(2)}(k) \sim \frac{1}{h} e^{-3h\left(\sqrt{2/\omega_0} - 1\right)} \qquad (2.8.20)$$

---

[*] This estimate is not very reliable in view of some uncertainty in estimate (2.8.2) for the number of periodic trajectories: in particular, it is possible that the exponents ν and β do not fully counterbalance each other. Nevertheless, numerical experiments confirm the order of magnitude of estimate (2.8.20) following from (2.8.15) (see Section 3.5).

For the quasi-resonances just considered, the phase oscillation period coincides with the period of the basic motion: $T_\phi = T$. This condition is not compulsory, the trajectory may enter in the stable phase region several times per period T, say $N = T/T_\phi$ times. In the general case the length of each period of the phase oscillations may be different: $T_\phi = T_i$; $\Sigma_{i=1}^{N} T_i = T$. If the T_i do not differ too greatly from one another, it can be considered approximately that the accuracy of each rotation will be determined by the length of its period (2.8.14): $\beta_i \sim e^{-h(T_i-1)}$. The probability of a specific sequence of stable regions $\sim \Pi_{i=1}^{N} (\omega_0 e^h) \cdot e^{-hT_i}$, and the number of stable trajectories of this (third) type is:

$$\nu_s^{(3)}(T) \sim \gamma^N e^{-h} \cdot C(T, N)$$

(2.8.21)

where $C(T,N)$ is the number of different combinations of T_i. Since our estimates are valid for $T_i \approx T_\phi$, one can put: $C(T,N) \sim n_0^N$, where $n_0 = \alpha T_\phi \leq T/N$ (T >> 1; $\alpha \leq 1$). Consequently (N >> 1) [*]:

$$\nu_s^{(3)}(T) \sim (\alpha T_\phi \gamma)^N$$

(2.8.22)

The total number of such trajectories and also the area of the stable regions diverges for any γ. However, the minimum period from which the divergence begins depends on γ:

$$T_{min} \sim (\alpha \gamma)^{-1}$$

(2.8.23)

By analogy with the ordinary resonances, one can assume that total mutual destruction of the quasi-resonances takes place only for close frequencies of the phase oscillations (see Section 2.6). Hence it follows that there remain undamaged quasi-resonances of all types with $T_\phi < T_{min}$; their number decreases with the growth of γ.

The phase area of the stable region around the periodic trajectory on the assumption that $T_i \approx T_\phi$ is given by the estimate, similar to (2.8.17):

$$S_1^{(3)} \sim T \cdot e^{-3h(T_\phi-1)}$$

(2.8.24)

and the total area of all the quasi-resonances of the third type is:

$$S^{(3)}(k) \sim \sum_{T=T_1}^{\infty} T \cdot (\alpha \gamma T_\phi)^N e^{-3h(T_\phi-1)}$$

(2.8.25)

where T_1 is again determined from the condition that quasi-resonances of this type exist for an arbitrary k: $\Sigma_{T=1}^{T_1} \nu_s^{(3)}(T) \sim 1$. When $T_\phi = $ const, $N = T/T_\phi \to \infty$, the sum diverges if $T_\phi > T_{min}$ (2.8.23).

---

[*] This estimate is an upper one, as (2.8.13) for quasi-resonances of the first type, see explanation on p. 67.

Let us recall that the conclusion concerning the divergence of the sum (2.8.25), from which the destruction of the overwhelming majority of quasi-resonances ensues is not rigorous, since we can obtain only the upper estimate for $S^{(3)}(k)$ (see remark on p. 67). However, an interesting feature of the problem under consideration is the fact that the basic conclusion concerning the stochasticity of motion of model (2.8.1) when $k \gg 1$ does not even depend on the assumption that the sum of (2.8.25) diverges. Indeed, if the sum of the areas of the quasi-resonances converges, it goes to zero when $k \to \infty$, since the size of each stable region decreases with the growth of k (2.8.24); if the above-mentioned sum diverges, mutual destruction of the quasi-resonances takes place, excluding the finite number determined by condition (2.8.23), and their over-all stable area again goes to zero when $k \to \infty$. Of course, one cannot exclude the very special case when all the time $S(k) \sim 1$ when $k \to \infty$, in spite of the fact that $S_1 \to 0$, but such a situation seems to us highly improbable. This result is confirmed by numerical experiments (Section 3.5).

Thus we can now add a third effect to the two previous effects of the overlapping of first order resonances (unification of stable regions and destruction of the separatrix) -- the formation of a large number of quasi-resonances which completely eliminate the last centres of stability.

2.9 Periodic crossing of the resonance

Before going over to the final sections of this chapter, devoted to the general case of the interaction of resonances, let us consider yet another relatively simple system which can be reduced to the basic model (2.1.11) with discrete time. This is the periodic crossing through the resonance of a non-linear oscillator.

If the amplitude of the frequency oscillations considerably exceeds the width of the resonance

$$\Delta\Omega \gg (\Delta\omega)_H \sim \Omega_\phi \tag{2.9.1}$$

the action of the resonance can be considered as a short kick; accordingly, we have a system of the type of the basic model, whose border of stochasticity is determined by condition (2.5.1). On the other hand, as follows from the results of Section 2.6, the general criterion of stochasticity (2.1.4) must be valid.

The system considered in this paragraph is of special interest also because in one of the writer's early papers on stochasticity[10] an erroneous conclusion was drawn about the existence of two independent conditions of stochasticity which had to be fulfilled simultaneously. This conclusion was drawn precisely on the basis of the process in question.

Let the frequency, say, of the perturbation vary according to the law:

$$\Omega = \overline{\Omega} + \frac{\Delta\Omega}{2} \cdot \cos \Omega_o t \tag{2.9.2}$$

Under the condition $\Omega_0 \ll \Delta\Omega$ the perturbation has a locally equidistant spectrum with a distance between resonances of Ω_0 and a number of basic resonances $\sim \Delta\Omega/\Omega_0$. On the basis of Parseval's equation (normalization condition) we obtain the estimate:

$$\varepsilon f_{\kappa} \sim \varepsilon f_{o} \left(\frac{\Omega_{o}}{\Delta \Omega} \right)^{1/2} \tag{2.9.3}$$

where εf_0, εf_k are the amplitudes of the frequency-modulated force and its harmonics, respectively. The general criterion of stochasticity (2.1.4) gives:

$$s^2 = \left(\frac{\Omega_{\kappa}}{\Omega_{o}} \right)^2 \sim \frac{\varepsilon f_o v_o \omega'_w}{\Omega_o^2} \cdot \left(\frac{\Omega_o}{\Delta \Omega} \right)^{1/2} \sim 1 \tag{2.9.4}$$

independently of the rate of crossing the resonance. Here v_0 is the amplitude of the velocity, and $\Omega_k \sim \sqrt{\varepsilon f_k v_0 \omega'_w}$ is the frequency of the phase oscillations near one of the harmonics of the force, unlike $\Omega_\phi \sim \sqrt{\varepsilon f_0 v_0 \omega'_w}$, the frequency of the phase oscillations at the moment of crossing the resonance (see Section 1.5).

Now let us consider another approach to the problem. When the resonance is crossed rapidly (V >> 1, see below) the change of the frequency (and energy) of the oscillator is given by the expression (1.5.7), which leads to the first of the difference equations of the type of (2.1.11). The phase equation can also be obtained from (1.5.7) in the following way. Removing the brackets we find (k = 1):

$$\psi = \psi_o + \frac{\dot{\Omega} t^2}{2} \mp \frac{\Delta \omega}{2} \cdot t + \frac{(\Delta \omega)^2}{8 \dot{\Omega}} \tag{2.9.5}$$

Here the third term gives the ordinary phase change due to the change in the frequency of the oscillator after crossing the resonance. The factor $\frac{1}{2}$ is explained by the fact that this term takes into account only half the frequency change after the moment of exact resonance (see Section 1.5). The other half is included in the second term, which when there is arbitrary frequency variation $\Omega(t)$ is replaced by $\int (\Omega(t) - \omega_0) dt = \int \Omega dt - \omega_0 t$, where ω_0 is the value of the frequency of the oscillator at the moment of exact resonance. The last term of (2.9.5) is small under condition (2.9.1). In order to obtain the phase of the next resonance, it is necessary to sum up expression (2.9.5) after the first resonance (upper sign) and until the second (lower sign). Taking into account also the rules of the changing of signs when the direction in which the resonance is crossed changes (Section 1.5), we obtain:

$$\psi_1 = \psi_o + \int_0^{t_{o1}} \Omega dt' - \omega_1 t_{o1} + \frac{(\Delta \omega)_o^2}{\dot{\Omega}_o} + \frac{(\Delta \omega)_1^2}{\dot{\Omega}_1} \tag{2.9.6}$$

where $t_{01} \approx T/2 = \pi/\Omega_0$ is the interval of time between two successive crossings of the resonance and it is assumed that $\omega \approx \bar{\Omega}$ (2.9.2). Equality $t_{01} \approx T/2$ is violated, firstly as a result of the frequency variation: $\Delta t_{01} \sim (\Delta \omega)^2 / \dot{\Omega}$; this effect can be ignored, like the two last terms in (2.9.6) under condition (2.9.1). Secondly, it is necessary to take into account the finite amplitude of the phase oscillations, so that the moment of resonance is determined by the intersection of the straight line and the sinusoid in Fig. 1.5.1. This leads to additional change of the phase of the resonance by $\Delta \psi \sim V^{-1}$. This effect becomes considerable when there is slow crossing of the resonance (see below).

Taking into account the above-mentioned approximations, the crossing of the resonance can be described by the following transformation:

$$\omega_{u+1} \approx \omega_u - \sqrt{\frac{2\pi}{V}} \cdot \Omega_\varphi \cdot \cos\left(\psi_u \pm \frac{\pi}{4}\right)$$

$$\psi_{u+1} \approx \psi_u + \theta - \frac{\pi \omega_{u+1}}{\Omega_0}$$

(2.9.7)

where the constant phase $\theta = \int_0^{T/2} \Omega \, dt$ and the sign is determined by the direction in which the resonance is crossed (see Section 1.5). The stochasticity parameter for (2.9.7) is found in a similar way to (2.4.9) and (2.4.3):

$$K_0 \approx \pi \sqrt{2\pi} \, \Omega_\varphi^2 \big/ \Omega_0 \sqrt{\dot{\Omega}} \sim 1$$

(2.9.8)

The last estimate gives the criterion of stochasticity (2.5.1). Since $\dot{\Omega} \sim \Delta\Omega \cdot \Omega_0$ and $\Omega_\Phi^2 \sim \Omega_k^2 \, (\Delta\Omega/\Omega_0)^{\frac{1}{2}}$ (2.9.3), then $K_0 \sim (\Omega_k/\Omega_0)^2 \sim s^2$ and both forms of the stochasticity criterion (2.9.4) and (2.9.8) agree.

With slow crossing of the resonance, when

$$V = \frac{\dot{\Omega}}{\Omega_\varphi^2} \sim \frac{\Delta\Omega \cdot \Omega_0}{\Omega_\varphi^2} \sim \sqrt{\frac{\Delta\Omega}{\Omega_0}} \cdot s^{-2} \ll 1$$

(2.9.9)

the change of the oscillator frequency is given by expression (1.5.9):

$$\Delta\omega(\omega,\xi) = 2V\Omega_\varphi \cdot \ln(V+\xi)(4\pi V - \xi^2) + $$
$$+ \frac{8}{\pi}\Omega_\varphi \sqrt{1 + 2\pi V - \xi^2/4}$$

(2.9.10)

The phase change is determined, as usual, by relation (2.9.6), but the additional finite amplitude of the phase oscillations must now be taken into account, as noted above. This is equivalent to changing over from the continuous phase ψ_1 determined by relation (2.9.6) to the phase $\xi = \psi - \pi/2$ limited by the interval: $-V < \xi < \sqrt{4\pi V}$. Since in this interval $\cos\xi \approx 1 - \xi^2/2$, from (1.5.3) we find (see also Fig. 1.5.1):

$$V(\psi - \psi_1) \approx 1 - \xi^2/2$$

(2.9.11)

In what follows we shall need the derivative:

$$\frac{d\xi}{d\psi_1} = \frac{V}{V+\xi} \sim \sqrt{V}$$

(2.9.12)

where we used the condition $d\psi/d\xi = 1$ (Fig. 1.5.1). The latter estimate is also easy to obtain directly if it is taken into account that the range of variation of ψ is equal to 2π and of ξ to $\sqrt{2\pi V}$.

Now we can approximately describe the slow crossing of the resonance by the transformation:

$$\omega_{n+1} \approx \omega_n \pm \Delta\omega\,(\omega_n,\,\xi_n)$$

$$\psi_{n+1} \approx \psi_n + \theta - \frac{\pi\,\omega_{n+1}}{\Omega_0} \qquad (2.9.13)$$

Here the function $\Delta\omega(\omega_n,\xi_n)$ is given by expression (2.9.10), the link between ψ_n and ξ_n by equation (2.9.11) and the constant phase θ is defined above. The stochasticity parameter is:

$$K_0 \approx \left| \frac{d\xi_{n+1}}{d\psi_{n+1}} \cdot \frac{d\psi_{n+1}}{d\xi_n} \right| \sim \frac{\dot\Omega}{\Omega_0\Omega_\phi} \sim \frac{\Delta\Omega}{\Omega_\phi} \qquad (2.9.14)$$

where the first of the derivatives is given by expression (2.9.12) and the second $\sim \sqrt{V\Omega_\phi/\Omega_0}$.

Relation (2.9.14) shows that in the approximation of short kicks (2.9.1) $K_0 \gg 1$, i.e. slow crossing always leads to stochasticity.

At first glance, criterion (2.9.14) is in no way connected with the parameter of the overlapping of the resonances s. However, it should not be forgotten that the condition of slow crossing (2.9.9) must be fulfilled, from which it follows that s > 1 always when $\Delta\Omega > \Omega_0$. If $\Delta\Omega \ll \Omega_0$, the parameter s loses its sense, since in this case there is in fact a single resonance $\Omega = \bar\Omega$, while the width of the remaining ones is considerably smaller and they can simply be ignored (see Section 2.7).

We still have to consider the case $\Delta\Omega \lesssim \Omega_\phi$, when the short kick approximation is not valid. Instead of this, let us turn to the phase equation originating from the equations in slow variables of the form of (1.3.15):

$$\dot I \approx -\varepsilon\,U_0\,\cos\psi$$
$$\dot\psi \approx \omega - \Omega(t) \qquad (2.9.15)$$

where $\Omega(t)$ is the periodic function of (2.9.2). We obtain:

$$\ddot\psi + \Omega_\phi^2\,\cos\psi \approx -\dot\Omega(t) \qquad (2.9.16)$$

This system has a separatrix in the vicinity of which a stochastic layer is formed under the action of the perturbation in the right-hand side of (2.9.16). According to the results of Section 2.6 the relative width of the stochastic layer in energy is given by the estimate:

$$\delta \sim \mu \sim \frac{\dot\Omega}{\Omega_\phi^2} = V \qquad (2.9.17)$$

under the condition that the perturbation frequency is sufficiently small: $\Omega_0 \lesssim \Omega_\phi$. In the opposite case, the width of the layer is exponentially small (2.6.17). The results of Section 2.6 are applied when there is small perturbation ($\mu \ll 1$), i.e. only for slow crossing of the resonance. When the crossing is rapid, system (2.9.16) simply does not

have a separatrix (Section 1.5). As regards stochasticity, which according to (2.9.4) is
also possible for rapid crossing, it is connected with a completely different mechanism,
namely mixing of the phase from one crossing of the resonance to the next.

For slow crossing of the resonance there are thus two mechanisms of stochasticity:
one which is the same as for rapid crossing and another connected with the stochastic
layer. The influence of the latter mechanism depends on the ratio between the range of
frequency variation ($\Delta\Omega$) and the width of the layer: $\Delta\Omega/\Omega_\phi \delta \sim \Omega_\phi/\Omega_0$. The maximum influence
corresponds to the condition $\Omega_0 \sim \Omega_\phi$ (when $\Omega_0 > \Omega_\phi$ the width of the layer decreases exponen-
tially). In this case stochasticity occurs, which is not at variance with the general
criterion according to the parameter s, since $\Delta\Omega \ll \Omega_0 \sim \Omega_\phi$ and this parameter loses its
sense, as noted above.

If $\Omega_0 \ll \Omega_\phi$ (and $\Delta\Omega \ll \Omega_\phi$), the case when $\Delta\Omega \gg \Omega_0$ is possible, so that the stochas-
ticity parameter $s \geq 1$ (2.9.9) has its ordinary meaning. From the point of view of equa-
tion (2.9.16) in this case stochasticity may also be expected, since the system passes
through a stochastic layer during approximately one phase oscillation period $[(1.5.3)$ and
$(2.9.17)]$. Moreover, with slow crossing capture is possible (Section 1.6), which increases
the time the system spends in the stochastic layer and consequently also the over-all
stochasticity of the motion.

Thus the general criterion of stochasticity according to the overlapping of resonances
(2.1.4) also applies to periodic crossing of the resonance as well as to rapid and slow
crossing. In the latter case stochasticity always occurs, in contradiction to Ref. 10, in
which it was assumed on the contrary that stochasticity is always absent for slow crossing,
on account of the approximate reversibility of the process (see Section 1.5). As we shall
see later this last effect leads only to a reduction of the diffusion rate and the K-entropy
(Section 2.11).

2.10 Kinetic equation

If the motion of a dynamical system becomes stochastic, it no longer makes any physi-
cal sense to describe it in terms of a trajectory, because of local instability. The
changeover to a statistical description, the meaning of which also, in our opinion, lies
precisely in re-establishing the stability of the description, is usually carried out in
two stages. First of all, as the basic physical quantity, one introduces the distribution
function or phase density $f(x,t)$ (x is the complete set of phase space coordinates) of an
ensemble of identical systems, differing only by the initial conditions. The variation of
f is determined by Liouville's equation:

$$i \frac{\partial f}{\partial t} = \hat{L} f \tag{2.10.1}$$

where $\hat{L}$ is a linear differential operator[49].

It is customary to emphasize the equivalence of Liouville's equation to the dynamical
equations. However, it should not be forgotten that their solutions are physically iden-
tical only for singular initial conditions: $f(x,0) = \delta(x - x_0)$, since, at least within
the limits of classical mechanics, we are concerned with only one single system of the
statistical ensemble. This fact, which is often underestimated (see for example Ref. 50)

is of paramount importance when discussing the nature of statistical laws[10] (see also Section 2.13). The use of the continuous density of f(x,t) already means introducing into the mechanics some random element and, in particular, excluding a set of special zero-measure trajectories. These special trajectories should not be regarded as absolutely exceptional. For instance, the periodic trajectories of a stochastic system which form an everywhere dense set[40] (Section 2.8) are related to them. In addition, the introduction of continuous density automatically excludes any fluctuation in the limit t → ∞.

With the above reservations, Liouville's equation is equivalent to the dynamical equations and its solution for a stochastic system also proves unstable in the following sense. Let us introduce so-called coarse-grained phase density $\bar{f}(x,t,\lambda)$, which is obtained by averaging f(x,t) over the phase space cells, of a size $\lambda \to 0$ [32]. Only such a density also has a direct physical sense. Indeed, we always have to do with a finite, although possibly also very large, number of systems N. Hence it is clear that the density is determined only for finite cells of the phase space containing many systems: $f(x,t) \cdot \lambda^k \gg 1$. If, on the other hand, there is only one system and the density is found according to the relative time the system stays in the phase space cell, the system must enter the cell again several times, i.e. the cell must have a finite size for any finite time of motion.

It is evident that the properties and behaviour of the coarse-grained density $\bar{f}$ depend to a certain extent on the choice of one or another set of phase space cells. This is why the notion of subdivision (of the phase space into cells) is one of the basic notions of the ergodic theory. In particular, as Sinai[52] recently showed, special (Markovian) subdivisions can be chosen, which enable one to change over rigorously from a dynamical description to a statistical (random) one in the form of a Markovian process.

Let us assume that

$$f(x,t) = \bar{f}(x,t,\lambda) + \tilde{f}(x,t,\lambda) \qquad (2.10.2)$$

where $\tilde{f}(x,t,\lambda)$ is the fine-grained density with a wave length λ. It turns out that however small $\lambda \to 0$ is, the fine-grained $\tilde{f}$, generally speaking, has a considerable influence over the development of the coarse-grained density $\bar{f}$ according to Liouville's equation. This follows directly from the qualitative picture of the mixing process given in Section 2.4. · It is obvious that we have to do with the instability of the trajectories of the stochastic system expressed in other terms. Let us recall that the time of development of such instability very weakly depends on the scale of λ: $\tau_\lambda \sim |\ln \lambda|$ (Section 2.4).

In order finally to get rid of this instability, it is necessary to change over from Liouville's equation to another one, which automatically excludes the fine-grained density. This equation will be called kinetic[*]. In order to exclude the fine-grained density it is natural to add to Liouville's equation the operation of periodic (for some characteristic

---

[*] This definition is not generally accepted. Sometimes, for example, Liouville's equation[32], or one related to it[69], is called kinetic. On the other hand, the term "kinetic equation" is also used, since the work by Bogolyubov[54], in a narrow sense to designate only the equation for a so-called single-particle distribution function.

interval of time Δt) averaging of the equation (for $\bar{f}$) over all possible functions of $\tilde{f}$, or more precisely over a complete set of such functions. If as the latter one chooses δ-functions in each of the phase space cells $\left[f(x,t) = \Sigma_i \, f_i \cdot \delta(x - x_i) \right]$ this will give us a clear picture of averaging over the position of the trajectory of the system in a phase space cell. Moreover, the trajectory distribution inside the cell is considered to be uniform. This latter hypothesis, necessary for carrying out the averaging operation, is the basis, explicit or implicit, of all the methods of obtaining a kinetic equation[30]. The "complexity" of the usual systems of statistical mechanics and the "complication" of their trajectories afford intuitive justification of this hypothesis. These intuitive considerations are formulated mathematically[49] by going to the limit $N \to \infty$, $V \to \infty$, $N/V =$ const, where N is the number of particles in the system, and V its volume (for further details of this method see Section 2.13). Other justification can be obtained by means of modern ergodic theory (see below).

Let us explain the physical meaning of the averaging in terms of trajectories. As already observed above, in reality we are always concerned with a single trajectory of a single system (a finite number of systems, interacting or not, equivalent to one system in the unified phase space -- so-called $\underline{\Gamma\text{-space}}$). The motion along this trajectory can be split into two processes: mixing in a small section of the phase space ($\lambda \to 0$) in the immediate vicinity of a given point of the trajectory, and transition from one such section to another throughout the whole of the accessible region of the phase space. The latter process is exactly described by the kinetic equation, while the first is equivalent to averaging in time or, because of the ergodicity, per phase cell. From the physical, or rather, mechanical point of view, the initial process (in Γ-space, see Section 2.12) is averaging in time (first process).

The instability of the solutions of Liouville's equation for a stochastic system generalizes the standard notion of an improper problem for the equation in partial derivatives, a notion introduced by Adamar (see for instance Ref. 58). This means that there is no continuous dependence of the solution on the initial conditions.

The Cauchy problem for Liouville's equation -- the wave type equation -- is always proper in the usual sense, namely for a finite interval of time and with the "distance" between the functions determined through their difference [for instance, $\rho(f,\phi) =$ max $|f(x) - \phi(x)|$, see Ref. 58]. However, the parameter λ of fine-grained density (2.10.2) can be taken as the "distance" between the distribution functions. When $\lambda \to 0$ the functions $f(x,t)$, $\phi(x,t)$ are considered to be close independently of the values $|\tilde{f}|$, $|\tilde{\phi}|$, provided $\bar{f} \to \bar{\phi}$ in the usual sense. It is evident that the distribution functions that are close in the sense indicated are characterized by an infinitely small trajectory shift; this is precisely the physical meaning of the new definition of "distance" between the functions. If, further, the asymptotic solution of Liouville's equation with $t \to \infty$ is considered, the problem becomes improper and therefore requires special methods of solution. One of them is precisely the kinetic equation method.

Recently, completely independently of statistical mechanics, various methods of solving improper problems of a completely different kind have been developed. The most complete survey of this work is to be found in a report by Lavrent'ev[59]. One of the

methods, using a regularizing operator, proposed by Tikhonov[60], recalls the kinetic equation method in its approach. It would be interesting to make a more systematic comparison of both classes of improper problems.

In connection with the averaging operation, the notion of the probability of transition (between phase space cells) naturally arises, and this can be calculated on the basis of the dynamical equations and the above-mentioned hypothesis concerning the uniform "spreading" of the trajectory over the cells.

The transition probability enables us not only to obtain a most general kinetic equation of an integro-differential type[51,43], but also to describe the fluctuations neglected by the standard kinetic equation. Description of the motion of the system in terms of transition probability is called the Markovian process. Its characteristic feature is independence from the previous history of the motion. For arbitrary subdivision of the phase space the probability of transition between the cells of the subdivision, determined by the measure of the corresponding regions, generally speaking depends on the previous states. This did not allow of rigorous transition from the dynamical equations to the Markovian process[30] in spite of numerous attempts. Only recently Sinai succeeded in constructing special subdivisions for which such transition proved possible[52]. These Markovian subdivisions have, generally speaking, a very complicated structure. Therefore, in the present paper, we shall restrict ourselves to obtaining the kinetic equation only, as the simplest method of describing a stochastic process.

The solutions of the kinetic equation are generally speaking stable and thus again have the usual physical sense. However, this stability is bought at the cost of part of the solutions of the original Liouville equation, which are unstable. They describe the growth of the fluctuations. A priori it is not at all obvious that the stable solutions of the kinetic equation (of the type with relaxation to an equilibrium state) generally exist and, moreover, describe in some sense the overwhelming majority of processes observed. This is due mainly to the fortunate fact that our world is in a strongly non-equilibrium state. If we had to describe the miserable phenomena which could still occur in a state of statistical equilibrium it would perhaps be just these processes of formation of large fluctuations that would be the most important. We should be faced with the very difficult dilemma of nevertheless devising some way of making a stable description of the growth of the fluctuations, or of generally rejecting the requirement for stability when describing physical processes. In any case, the statistical physics of such processes would appear rather unusual from the modern point of view. As an instance, one can cite the present method of describing the growth of large fluctuations a posteriori, namely under the condition (afterwards) of the formation of fluctuations with given parameters[61,62]. Such a description is made by means of an equation similar to the kinetic one and is stable. However, it is clear that the most important feature of the law of physics -- the possibility of prediction -- will be lost. These questions will be discussed further in Section 2.13.

Since the processes of relaxation and growth of the fluctuations are reciprocally reversible in time, the kinetic equation which describes only relaxation is of necessity irreversible. It is clear that this in no way means the physical disparity of both directions in time, or the existence of "the time arrow", the current expression[70], but is the

consequence of deliberate exclusion of reverse processes which are in principle possible, but are unstable (for more detail see Section 2.13). It is interesting to note that this point of view, which is perfectly natural for present-day ergodic theory (see Section 2.4) is not shared by many of the physicists (see for example Ref. 70).

As far as we know, so far no-one has attempted to obtain a kinetic equation by means of ergodic theory. Generally the statistical element when obtaining a kinetic equation is just to postulate in one form or another, for instance the assumption of the absence of pair correlations in Boltzmann's first kinetic theory, the Bogolyubov condition of correlation relaxation[54,55] or the random phase hypothesis in the quasi-linear plasma wave theory[56].

In this paragraph the kinetic equation for the basic model (2.1.11) will be obtained without any a priori statistical hypotheses, on the basis of the results of the ergodic theory (Sections 2.3 and 2.4). A comparison of our approach to the problem and that of present-day statistical mechanics will be made in Sections 2.11 and 2.13. Now let us note only that the kinetic equation which we are just about to obtain is similar to the so-called master equation of statistical mechanics, since it relates to Γ-space. This is the only possibility for a one-dimensional system of the type of the basic model. In the many-dimensional case a kinetic equation of another type in so-called μ-space is possible (see end of Section 2.12).

Since for our model the perturbation is small ($\varepsilon \to 0$), the kinetic equation must have the form of an FPK diffusion equation (Fokker-Planck-Kolmogorov)[43,44] (2.10.10). It is determined, as is known, by the two first moments $\langle (\Delta I)_1 \rangle$, and $\langle (\Delta I)_1^2 \rangle$, where the index 1 signifies that the mean value is taken per unit of time (one step in our case) and averaging is carried out by means of transition probability.

Let us recall that the first moment $\langle (\Delta I)_1 \rangle$ describes the systematic variation of the momentum and is equal to the mean rate of its variation. The diffusion itself is described by the second moment $\langle (\Delta I)_1^2 \rangle$, which is equal to the mean rate of I "dispersion". All the statistical properties of this process are linked just with the latter quantity and it is only owing to this quantity that the FPK equation becomes kinetic in the sense indicated above. Hence it can be concluded that for transition to the kinetic equation it is necessary for the quantity $(\Delta I)^2$ to increase (on the average) $\propto t$.

As noted above, averaging over the transition probability or the phase space cells is not different from an approximate description of the development in time of one of the systems of the statistical ensemble determined by a single dynamical trajectory. Therefore when calculating $\langle (\Delta I)_1^2 \rangle$ the original averaging should be carried out in time. Let us write the variation I_n for one of the systems of the ensemble in the form:

$$I_{n+1} = I_0 - \varepsilon \cdot \sum_{k=0}^{n} h_\theta(\theta_k) \qquad (2.10.3)$$

where for the sake of simplicity we neglect the dependence of h_θ on I_n, which will be taken into account when calculating $\langle (\Delta I)_1 \rangle$ (see below). We find

$$(\Delta I)^2 = \varepsilon^2 \cdot \left[\sum_{k=0}^{n} h_\theta^2(\theta_k) + \sum_{k \neq \ell} h_\theta(\theta_k) \cdot h_\theta(\theta_\ell) \right] \qquad (2.10.4)$$

In order to obtain the kinetic equation it is necessary for this value to increase (on the average) in proportion to time ($\propto$ n), namely it is sufficient to ignore the second sum. It is precisely that sum which includes all the instability of the stochastic motion and also its reversibility (and hence the fluctuations), since this sum depends on the correlation of phases that are successive in time. Accordingly, it is sufficient for us to require the absence of pair autocorrelations (Section 2.3). If, moreover, the motion is ergodic in phase, then:

$$\overline{(\Delta I)_1^2} = \langle (\Delta I)_1^2 \rangle = \varepsilon^2 \cdot \langle h_\theta^2 \rangle \qquad (2.10.5)$$

Let us note that the rejection of the second sum (Σ_2) in (2.10.4) is not a trivial mathematical operation, since Σ_2 is of the order of the first sum ($\Sigma_k^n h_\Theta(\theta_k) \sim h_\Theta \cdot \sqrt{n}$). The rejection is possible only on the average, since Σ_2 oscillates. In principle, this procedure is similar to the averaging method in non-linear mechanics[3]. Moreover, we simply reject such special initial conditions when the second sum is much larger than the first for a long time (fluctuations).

It would seem that it is the ergodic theory that provides justification for the rejection of Σ_2 in (2.10.4), even weak mixing being sufficient (Section 2.3). However, in reality the situation is more complicated. The point is that according to the ergodic theory mixing, and in particular the disappearance of pair correlations, takes place only asymptotically when $t \to \pm\infty$, i.e. it takes place parallel to relaxation to statistical equilibrium. But we wish to describe the relaxation process itself. Another aspect of this difficulty is that mixing with respect to the phase Θ, which we should like to regard as a "simply random" ("microscopic") parameter, is necessarily accompanied, by virtue of the equations of motion (2.1.11) by mixing with respect to the momentum I, which should play the part of a diffusion ("macroscopic") variable. Therefore, at first glance it seems generally impossible to apply the kinetic equation to inhomogeneous distributions and this means that it completely loses its sense. These difficulties were thoroughly analysed by Krylov[30] who came to a pessimistic conclusion.

It seems to us, however, that a solution can be found by using an idea of Bogolyubov[54,55] of introducing different time scales into the problem. In the case of the basic model it is a question of two kinds of time -- the dynamical mixing time $\tau_n \sim 1$ (one step) and the I diffusion time: $\tau_D \sim \varepsilon^{-2}$ (steps, see below). Asymptotically when $\varepsilon \to 0$ we can thus separate both processes with an arbitrary degree of accuracy[*].

Let us make a more accurate formulation of the conditions of such separation of mixing from diffusion. For this we will calculate (2.10.4), having split it into sub-sums with a

---

*) A comparison with Bogolyubov's theory, where other time scales are introduced, will be given in Section 2.13.

given difference $m = k - \ell$ ($\ell = 1,2, \ldots$) and introducing the pair correlation coefficient $\rho(m) = \langle h_\Theta (\Theta_k) \cdot h_\Theta (\Theta_{k+m}) \rangle / \langle h_\Theta^2 (\Theta_k) \rangle$. The sum which is of interest to us turns out to be asymptotically ($n \to \infty$) equal to: $\Sigma_2 \approx n \cdot \langle h_\Theta^2 \rangle \cdot \Sigma_{m=1}^{\infty} \rho(m)$, namely proportional to n like the first sum in (2.10.4). Therefore the (FPK) kinetic equation remains valid also when the correlation exists (with another diffusion coefficient), provided the sum $\Sigma \rho(m)$ converges. In the opposite case Σ_2 increases faster than t and the kinetic equation is inapplicable.

The sum convergence condition can be written in the form:

$$\rho(\bar{m}) < \bar{m}^{-\alpha} \; ; \; \alpha > 1 \; ; \; \bar{m} \to \infty \qquad (2.10.5a)$$

This is the more accurate condition of application of the kinetic equation for systems of the type under consideration. From this condition it can be seen, in particular, that weak mixing is insufficient since generally speaking there is no limitation of the rate of decrease of the correlations. Stochasticity is a sufficient condition because it includes a requirement for positive K-entropy (Section 2.3). In the latter case $\rho(m) = e^{-h \cdot m} = \beta^m$ ($\beta < 1$; h = K-entropy) and for the second moment the following expression is obtained, defining (2.10.5) more exactly:

$$\langle (\Delta I)_1^2 \rangle = \varepsilon^2 \langle h_\theta^2 \rangle (1 - \beta)^{-1} \qquad (2.10.5b)$$

When the K-entropy is sufficiently great (h >> 1) correction is small. If, however, h << 1, it is necessary to take into account not only the change of the second moment and consequently also of the diffusion coefficient (2.10.12), but also of the dynamical scale of time $\tau_n \sim h^{-1}$.

It should be noted, finally, that the exponential decrease in the correlations when h > 0, although not a necessary condition, considerably increases the accuracy of the statistical description by means of the kinetic equation.

As noted above, the mixing process covers not only the phase Θ but also the momentum I, which on the other hand is an independent variable of the distribution function. In order to overcome this difficulty let us split I into two parts:

$$I = \bar{I} + \tilde{I} \qquad (2.10.6)$$

a diffusion ($\bar{I}$) and a dynamical ($\tilde{I}$) part, so that the mixing process affects only $\tilde{I}$ and not $\bar{I}$. On the other hand the diffusion is now determined only with an accuracy of the order of the value $\tilde{I}$, which must of course be sufficiently small for this whole procedure to have a physical sense. The value of $\tilde{I}$ can best be estimated from the slope of the extension eigenvector, i.e. from the slope of the extending asymptote (2.4.14):

$$(\Delta I)_o = \tilde{I} \sim \frac{\varepsilon}{K_o} << 1 \qquad (2.10.7)$$

Thus there is a minimum size (ΔI_0) of subdivision cell beyond which a kinetic description, i.e. a description by means of a kinetic equation, becomes inapplicable. The distribution function $\bar{f}$ inside the minimal cell should be considered constant.

The size of the cell (2.10.7) may be decreased if one considers the transformation to be in N steps. Then $\varepsilon \to \varepsilon\sqrt{N}$; $K_0 \to K_0^N$ (see Section 2.11), but the ratio $\tau_D/\tau_n \sim (N\varepsilon^2)^{-1}$ decreases. The condition $\tau_D \sim \tau_n$ determines the maximum permissible length of the dynamical mixing process: $N_{max} \sim \varepsilon^{-2}$. Whence the absolute minimum size of cell is:

$$(\Delta I)_{\text{мин}} \sim e^{-ch/\varepsilon^2} \tag{2.10.8}$$

where $h \sim \ln K_0$ is the K-entropy and $c \sim 1$ is a constant depending on the required accuracy of description.

It should be stressed that the limitation of the size of the phase space cell, and consequently also the permissible subdivision, relates only to the description of the relaxation process by means of a kinetic equation and does not extend to the asymptotic theorems of the ergodic theory. In particular, this limitation no longer applies for an equilibrium state. Let us note, however, that in any case statistical mechanics has to do with finite, although arbitrarily small, phase space cells, which is equivalent to using a coarse-grained distribution function[*]. Often it is not specially stipulated [see for instance Ref. 65], but simply implied that the kinetic equation gives an incomplete description according to part of the variables (say, according to the momenta), while the remaining variables (the phases) determine the transition probability. In certain problems these "random" variables are outwardly camouflaged, as for instance in the Boltzmann type of kinetic equation. Sometimes imperfections of this kind lead to direct ambiguities, and in particular to the erroneous assertion that the "exact" entropy of the closed system does not increase[67].

Returning to the FPK equation, let us note that the first moment $\langle(\Delta I)_1\rangle$ must be calculated with the same accuracy as the second (2.10.5), namely with an accuracy $\sim \varepsilon^2$. Let us use expression (2.1.12) for this and take into account that $\tau \to 0$ ($\varepsilon\tau = $ const); $\langle h_\Theta \rangle = 0$; $\langle h_{\Theta I} \cdot h_\Theta - h_{\Theta\Theta} \cdot h_I \rangle = 2\langle h_{\Theta I} \cdot h_\Theta \rangle - \langle(h_\Theta \cdot h_I)_\Theta\rangle = (\langle h_\Theta^2 \rangle)_I$, where we assume that all the functions are continuous. We obtain:

$$\langle(\Delta I)_1\rangle = \frac{1}{2} \cdot \frac{\partial}{\partial I} \langle(\Delta I)_1^2\rangle \tag{2.10.9}$$

Taking into account this relation, the FPK equation

$$\frac{\partial \bar{f}}{\partial t} = -\frac{\partial}{\partial I}\left(\bar{f} \cdot \langle(\Delta I)_1\rangle\right) + \frac{1}{2}\frac{\partial^2}{\partial I^2}\left(\bar{f} \cdot \langle(\Delta I)_1^2\rangle\right) \tag{2.10.10}$$

comes as is known[**], to the standard diffusion equation

$$\frac{\partial \bar{f}}{\partial t} = \frac{\partial}{\partial I}\left(\mathcal{D}(I) \cdot \frac{\partial \bar{f}}{\partial I}\right) \tag{2.10.11}$$

---

[*] The recent revival of what is called symbolic dynamics[68] already described by Birkhoff (Ref. 16, Chapter 8, Section 11) is connected with this.

[**] See for instance Ref. 65.

with a diffusion coefficient

$$\mathcal{D}(I) = \frac{\langle (\delta I)_1^2 \rangle}{2} = \frac{\varepsilon^2 \langle h_5^2 \rangle}{2} \qquad (2.10.12)$$

The last expression shows, in particular, the validity of the estimate made above of the diffusion time scale $\tau_D \sim \varepsilon^{-2}$.

In what follows we shall omit the bar above the function f and shall consider, if this has not been done by special reservation, that all the distribution functions are coarse-grained ones ($\bar{f} \to f$).

Relation (2.10.9) considerably simplifies obtaining the FPK equation, since for calculating the second moment a first approximation is sufficient. Landau showed[66] that relation (2.10.9) follows directly from the principle of detailed balancing, i.e. from the symmetry of transition probability in relation to the initial and final states. Unfortunately, the principle of detailed balancing is far from always being valid, even when there is a symmetry condition in relation to time reversal. In the latter case the reversal of all velocities is implied, which generally speaking is not assumed when formulating the principle of detailed balancing.

In the general case it is necessary to add to the diffusion equation (2.10.11) the term $-(\partial/\partial I) [f \langle (\Delta I)_a \rangle]$, where $\langle (\Delta I)_a \rangle$ is the additional (anomalous) rate of variation of the momentum I. The general relation between the moments in the absence of detailed balancing was obtained by Belyaev[65]. A simple example is the crossing of the resonance by a non-linear oscillator (Sections 1.5 and 2.9). However, in this case it can also be said that there is no time reversal, since the process is considered for crossing the resonance in a given direction. A similar situation arises when charged particles move in a given external magnetic field[65]. Even for rapid crossing, generally speaking, a systematic shift appears, which can be calculated from expression (1.5.7). However, (1.5.7) cannot be averaged simply over the phase ψ_0, which is no longer, generally speaking, canonically conjugate to the momentum I. This is due to the fact that the transformation (2.9.7) relates to a variable interval of time, which itself depends on the dynamical variables. The scale of this effect, leading to non-uniform mixing in ψ_0, is characterized by the value $V^{-1} \ll 1$ (Section 2.9) and proves to be of the same order as the constant shift in (1.5.7). Let us note that this effect is unimportant when calculating the diffusion coefficient.

Dissipation can serve as an even simpler example. Let us go over from the momentum I to the energy W and put: $\langle (\Delta W)_a \rangle = \dot{W} < 0$. Then there is a steady solution of the FPK equation, which can be written in the form:

$$f(w) = f(0) \cdot \exp \left\{ \int_0^w \frac{\dot{w}}{\mathcal{D}} \, dw \right\} \qquad (2.10.13)$$

This means that there are steady state stochastic oscillations of the energy of the system under the action of the perturbation, which can be characterized by the effective "temperature":

$$T_{\partial \varphi}(W) = - \frac{\mathcal{D}}{\dot{w}} \qquad (2.10.14)$$

The energy spectrum of the system depends on the form of the functions D(W), $\dot{W}$(W): in particular, the Maxwellian spectrum is obtained when there is "balanced" (for the given dissipation) perturbation: $-D(W)/\dot{W}(W) = T = $ const.

This simple example shows that even when there is damping in a non-linear oscillatory system processes are possible which are described neither by the classical perturbation theory nor by the KBM theory (see Section 2.2).

Let us note, in conclusion, that when there is an additional condition for the symmetry of the momenta I_i of the system with respect to the sign of the velocities, the principle of detailed balancing and with it also relation (2.10.9), which for several degrees of freedom is written in the form[65]:

$$\langle(\Delta I_i)\rangle = \sum_{k}^{'} \frac{1}{2}\frac{\partial}{\partial I_k} \langle(\Delta I_i)(\Delta J_k)\rangle \qquad (2.10.15)$$

directly follows from the time reversibility. As a simple example one can take a system of weakly coupled oscillators for which the unperturbed canonical momenta depend on the squares of the velocities. In particular, for the one-dimensional case this follows directly from the results of Ref. 38.

2.11 Transition to continuous time, or the general case of the interaction of resonances

In Section 1.1 we began by studying the motion of a one-dimensional non-linear oscillator under the action of external perturbations. In the special but perhaps more interesting case of stochastic conditions we had to simplify the problem and go over to the basic model (2.1.11). The most important feature of the latter is discrete time. Although in itself the transition to the transformation in place of the differential equations does not limit the generality of the problem, since such a transition can always be carried out by means of an ordinary S-operator[*], the specific form of the basic model (2.1.11) is undoubtedly a certain special case of the original problem.

In this paragraph we shall endeavour to extend the results concerning the stochasticity of the basic model to the general case of the interaction of resonances (Section 2.1), namely to a one-dimensional non-autonomous non-linear oscillator of type (1.1.1). What concerns us mainly are the three basic parameters of a stochastic system -- the criteria of stochasticity, K-entropy and the diffusion coefficient.

For the starting point for our argument we will take the elementary model (2.8.1) which is almost stochastic when $k \gg 1$. The term "almost" implies the existence, generally speaking, of small "islets" of stability for any $k \to \infty$ (Section 2.8). This fact has so far prevented a rigorous study of the stochastic properties of the elementary model (see Section 2.4). Let us point out that our idea of its stochasticity is based not only on physical intuition but also on the results of various numerical experiments, which will be described in the following chapter.

With regard to extending the stochasticity criterion to the case of a arbitrary non-linear oscillator (1.1.1) or to the case of continuous time, as we shall say henceforth

---

[*] See for example Ref. 55.

for the sake of brevity, this has in fact already been done in Section 2.6. Let us merely recall briefly that this extension was possible because the mechanism of stochasticity is connected with the expansion and overlapping of the stochastic layers of resonances, which always exist in one form or another, the motion inside which, in the final analysis, amounts to the elementary model.

From the point of view of the set of resonances (continuous time) the basic model represents a special case in the sense that there are completely defined phase relations between the different resonances. The possibility of directly transferring the stochasticity criterion for the basic model (2.5.1) and (2.5.2) to the general case of the interaction of resonances shows that the stochasticity criterion does not depend on phase relations. This conclusion is also confirmed, in particular, for the very special case of periodic crossing of the resonance (Section 2.9). Unfortunately the same cannot be said of the two other characteristics of stochasticity -- K-entropy and the diffusion coefficient.

Let us begin with the entropy, considering a transformation of the basic model type:

$$\varphi_1 = \varphi + k \cdot f(\psi)$$
$$\psi_1 = \psi + \varphi + k \cdot f(\psi)$$

$$(2.11.1)$$

where $f(\psi)$ is a function that is periodic according to ψ with a period of one. In what follows we shall call the transformation a cascade, in order to stress the discreteness of the time[31]. We shall give the description of the motion in terms of continuous time the standard name of flux. What now interests us is the transition from the flux to the cascade, the stochastic parameters of which we know how to calculate.

It is not difficult to verify that in the general case the cascade will not have the form of the basic model. For this let us consider transformation (2.11.1) in two steps:

$$\varphi_2 = \varphi + k f(\psi) + k f(\psi + \varphi + k f(\psi))$$
$$\psi_2 = \psi + 2(\varphi + k f(\psi)) + k f(\psi + \varphi + k f(\psi))$$

$$(2.11.2)$$

In the general case it does not amount to (2.11.1), especially as this concerns a transformation with an arbitrary number of steps -- N. In the special case of k >> 1 the last term plays the main role in (2.11.2) so that the effective value of the stochasticity parameter is $K_2 \sim k^2$. In the same way, for a transformation with N steps $K_N \sim k^N$. This result could also have been obtained directly from the expression for the K-entropy of the basic model (2.4.19), which, of course, should not depend on the interval of the transformation: $h \simeq (1/N) \ln K_N \simeq \ln k$.

The quantity $K_N \sim s_N^2$ is the parameter of stochasticity of the cascade, which thus depends exponentially on the interval of the transformation. However, if we have a flux this stochasticity parameter is given: $s \sim \Omega_\phi/\Delta$, and consequently the expression for the K-entropy of the cascade (2.4.21) cannot in the general case be used for the flux[*]. However,

---

[*] In other words, it is not clear what is the characteristic interval of time of the cascade to which the parameter $s^2 \sim k_0$ of the flux corresponds.

in the special case of short kicks or periodic crossing of the resonance the K-entropy is actually determined by expression (2.4.19) with $k_0 \sim s^2$. This shows that the K-entropy essentially depends on the phase relations between the resonances and therefore a general estimate of it is impossible.

Let us note that this problem does not arise for determining the border of stochasticity, since at the border of stochasticity $K_N \sim k \sim 1$. This illustrates the remark made above concerning the non-dependence of the stochasticity criterion on the resonance phase relations.

In order to estimate the K-entropy of the flux one can, however, consider a case that is in a sense "typical", when the phases of various resonances are "random", i.e. when there are no special relations between them. Then the only interval of time characterizing the non-linear interaction of the resonances will be the inverse frequency of the renormalized phase oscillations (Ω_Σ^{-1}) and the renormalization, by virtue of the assumed randomness of the phases, should be carried out with a power $n = 4$ (Section 2.7). It is easy to obtain the law of renormalization of a system of resonances of the same order of width, by analogy with (2.7.13):

$$\Omega_\Sigma^{n-1} \sim \Omega_\phi^n \cdot \frac{(\Delta\omega)_\Sigma}{\Omega_\Sigma \cdot \Delta} \sim \frac{\Omega_\phi^n}{p \cdot \Delta}$$

$$(\Delta\omega)_\Sigma^{n-1} \sim \frac{(\Delta\omega)_\Sigma^n}{\Delta} \tag{2.11.3a}$$

where $(\Delta\omega)_\Sigma$, $(\Delta\omega)_H$ are the renormalized and non-renormalized width of the resonance $p\omega = q$, respectively; Δ is the distance between the resonance values of the frequency ω. Here we used the relation $\Omega_\phi \sim p(\Delta\omega)_H$ (Section 1.4). In what follows we shall assume for the sake of simplicity: $p \sim 1$. It is easy to see that as a rough estimate of the entropy one can put:

$$h \sim \Omega_\Sigma \sim \Omega_\phi^{4/3} \cdot \Delta^{-1/3} \tag{2.11.3}$$

There is hardly any sense in making estimate (2.11.3), which we shall call "typical", more accurate, because of its dependence on the phase relations. In particular, in the case of the basic model:

$$h = \frac{\Delta}{\pi} \ln \frac{\Omega_\phi}{\Delta} \tag{2.11.4}$$

Local instability, characterized by K-entropy, determines the process of phase mixing [for a system of the type of the basic model (2.11.1)]. The latter can also be described by means of the phase autocorrelations in a similar way to that mentioned in Section 2.3 for the elementary example of stochasticity. For reasons to be explained below we will slightly generalize the correlation coefficient determined above, putting:

$$\rho_n^{(2)}(p, q) = \langle e^{2\pi i(p\psi_n + q\psi)} \rangle \tag{2.11.5}$$

Let us first consider the special case of the cascade (2.11.1) with $f(\psi) = (1/2\pi) \times \sin 2\pi\psi$ and assume that n = 1. Expressing ψ_1 through φ,ψ and integrating by virtue of the ergodicity (k >> 1) over φ,ψ, we find:

$$\rho_1^{(2)}(p,q) = \frac{e^{2\pi ip} - 1}{2\pi i p} \cdot J_{p+q}(kp) \sim \frac{1}{p^{3/2}\sqrt{k}} \qquad (2.11.6)$$

From this it can already be seen that the laws of correlation relaxation and of the development of local instability do not agree, as was the case in the elementary example (Section 2.3). Moreover, the dependence of the correlation coefficient on q, which occurs when $k \lesssim q$ (p $\sim$ 1), is of a completely different nature from that described in Section 2.3. These peculiarities are explained mainly by the fact that the correlation coefficient is now determined primarily by the region near the stable phases (2.8.7): $\Delta\psi \sim k^{-\frac{1}{2}}$.

Let us now calculate $\rho_n^{(2)}$ (p,q). For this let us express ψ_n through the previous phases $\psi_{n-1}, \ldots, \psi$ by applying (2.11.1) successively:

$$\psi_n = \psi + n\varphi + k\{nf(\psi) + (n-1)f(\psi_1) + \ldots + f(\psi_{n-1})\} = \psi + n\varphi + k \cdot S_n'(\psi) \qquad (2.11.7)$$

When k >> 1 the successive phases can be considered to be random[*]. Therefore the random quantity $S_n(\psi)$ when n >> 1 is distributed normally with the parameters:

$$\langle S_n \rangle = 0; \quad \langle S_n^2 \rangle = \sigma^2 \approx \langle f^2 \rangle \cdot \frac{n^3}{3} \qquad (2.11.8)$$

When calculating $\rho_n^{(2)}$, integration over ψ can now replace integration over S, having assumed that $d\psi/dS \sim e^{-S^2/2\sigma^2}$, since $d\psi$ is proportional to the measure in phase. Assuming that q = -p to eliminate the term with ψ in the exponent (2.11.5), we obtain the estimate:

$$\rho_n^{(2)}(p,-p) \sim \exp\left\{-\frac{2\pi^2 p^3}{3} \cdot \langle f^2 \rangle \cdot k^2 n^3\right\} \qquad (2.11.9)$$

·The characteristic relaxation time of the correlations proves to be of the order of:

$$\tau_{cor} = n_{cor} \sim k^{-2/3} \qquad (2.11.10)$$

which agrees with the "typical" estimate for the inverse K-entropy (2.11.3) $[\Omega_\phi^2 \sim k; \quad \Delta = 2\pi$ for discrete time (2.11.1)]. Meanwhile, in the case under consideration ("atypical") the K-entropy is defined by expression (2.11.4) or, in discrete time h $\approx$ ln k.

This difference is probably explained by the fact that the K-entropy is determined by the behaviour of the system only on the asymptote, while the correlation coefficient is some integral quantity. From the point of view of the mixing process the correlation

[*] The small residual correlations (2.11.6) can be taken into account in the following approximation, as was done in the previous paragraph (2.10.5b).

coefficient is a more direct characteristic, so that it enhances the role of the "typical" estimate (2.11.3) and (2.11.10). It should, however, be remembered that for a cascade of type (2.11.1) with k >> 1 the difference between the K-entropy and the correlations can have real value only for some very fine details of the mixing structure (q >> 1) which will not be discussed here. For q ≤ k the residual correlations are small, even after one step (2.11.6).

Returning to the flux, it can be concluded that the characteristic damping time of the correlations, and hence also the mixing, will be determined by the "typical" estimate: $\tau_n \sim h^{-1}$ (2.11.3).

Let us take as an example yet another model, which can serve as a link between a cascade of type (2.11.1) and a flux. The model is given by the transformation:

$$\varphi_{n+1} = \varphi_n + \frac{k}{2\pi} \cdot \sin 2\pi (\psi_n + \theta_i)$$

$$\psi_{n+1} = \psi_n + \varphi_{n+1}$$

(2.11.11)

where θ_i is the sequence of T random phases, which then recur periodically. We shall from now on call this model <u>quasi-random</u>. On the one hand it recalls the basic model, since it is given by a transformation and the perturbation has a period T. On the other hand, when $\tau_0 \to 0$, where the interval τ_0 corresponds to one step, the quasi-random model changes over to the "typical" flux with a random discrete perturbation spectrum and the distance between the lines of the spectrum is $\Delta = 2\pi/T$. Bearing in mind the transition mentioned, it is assumed that k << 1.
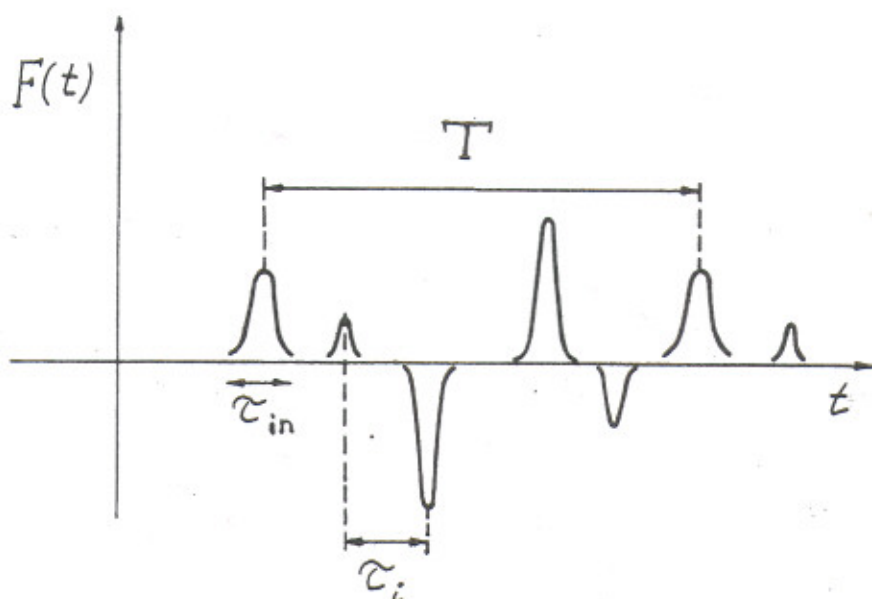

Fig. 2.11.1: Time dependence of the perturbation F(t) for the quasi-random model: τ_{in} is the duration of the interaction; T is the perturbation period.

Figure 2.11.1 is a schematic diagram of the time dependence of the perturbation for the quasi-random model. The model has been further extended by introducing variable (in particular random) distances between the kicks τ_i ($\bar{\tau}_i = \tau_0$) and a finite width of the kick τ_{in}, keeping a constant period T. Transformation (2.11.11) corresponds to the case of $\tau_i = 1$; $\tau_{in} = 0$. A flux results when $\tau_i \sim \tau_{in}$.

Using the results of Section 2.4 one can find the eigenvalues of transformation (2.11.11):

$$\lambda \approx 1 \pm \sqrt{K_i} \; ; \quad K_i = k \cdot \cos 2\pi (\psi_i + \theta_i) \qquad (2.11.12)$$

and the directions of the eigenvectors (Fig. 2.8.1):

$$tg\, \theta_{\pm} \approx \pm \sqrt{K} \qquad (2.11.13)$$

If transformation (2.11.11) were unstable ($K_i > 0$) in each step, then the K-entropy would be: $h \approx k^{\frac{1}{2}} \cdot \langle \sqrt{\cos 2\pi\psi} \rangle$. However, in reality in roughly half the cases $K_i < 0$, i.e. the transverse vector turns, and hence it can change over from the extension cone to the contraction cone (Section 2.8). When $k \ll 1$, the difference between the two cones is insignificant and itself depends on k. We can therefore write: $h \sim k^m$, where $m > \frac{1}{2}$. For what follows it is significant that h cannot depend on T, since under conditions of stochasticity (in a period T) all the phases ($\psi_i + \theta_i$) are random.

Let us now turn to estimates in terms of continuous time, having formally represented the cascade (2.11.1) as a flux with a δ-function in time. By virtue of the periodicity, the spectrum of the flux is equidistant: $\omega_n = n \cdot \Delta = 2\pi n/T$. Taking into account further that the sum of $\Sigma_1^T \sin 2\pi \, \theta_i \sim T^{\frac{1}{2}}$, we obtain the following estimate for the amplitude of the perturbation harmonic and the non-renormalized frequency of the phase oscillations:

$$F_o \sim \frac{k}{\sqrt{\pi}} \cdot T^{-1/2} \sim \Omega_\phi^2 \qquad (2.11.14)$$

Since the value Ω_ϕ depends on T, it cannot determine the characteristic time for the development of stochasticity, and that implies also K-entropy (see above). Renormalization with a power of n gives: $\Omega_{\phi\Sigma}^n \sim \Omega_\phi^n \cdot (\Omega_{\phi\Sigma}/\Delta)$, or:

$$\Omega_{\phi\Sigma} \sim \left[\left(\frac{k}{\pi} \right)^{\frac{n}{2}} \cdot \frac{T^{1-\frac{n}{4}}}{2\pi} \right]^{\frac{1}{n-1}} \qquad (2.11.15)$$

The dependence on T disappears when $n = 4$ and we thus arrive at a "typical" estimate (2.11.3) for a special case of the quasi-random model (2.11.11).

Let us find the phase correlations for this model. It is not difficult to find out that they are given by the same estimate of (2.11.10) as for the basic model (2.11.1). The sole difference lies in the fact that for the quasi-random model this estimate is valid for any k, and also when $k \to 0$ (changeover to flux). In reality the only requirement in order to obtain (2.11.9) and (2.11.10) is that the quantity $S_n(\psi)$ be the sum of the random functions (2.11.7). However, this is automatically provided for in the quasi-random model for any k (2.11.11), provided the stochasticity criterion is satisfied in a

period T $\left[k T^{3/2} \geq 1;\quad \text{see } (2.11.14)\right]$. Using (2.11.10) and (2.11.14) we again arrive at a "typical" estimate (2.11.3).

Let us now turn to the calculation of the diffusion coefficient of the flux. As shown in the previous paragraph, for this it is sufficient to find the second moment $\langle (\Delta I)_1^2 \rangle$. Let the variation of I be given by the equation:

$$\frac{dI}{dt} = \sum_{n}' F_n \cdot \cos\left[\int \omega_n \, dt + \theta_n(t)\right] \tag{2.11.16}$$

where the quantities $\omega_n(t)$, $\theta_n(t)$ slowly vary under the action of the perturbation during a time $\sim \tau_{cor}$. Since ω_n in fact signifies the differences of resonance frequencies, it will be convenient for us to consider that they may be both positive and negative: $-\Omega_1 < \omega_n < \Omega_2$. In what follows we shall assume for the sake of simplicity that the basic part of the spectrum is situated symmetrically to the resonance: $\Omega_1 \sim \Omega_2 \sim \omega_{max}$. The stochasticity of the flux (2.11.16) corresponds to the condition $\tau_{cor} \cdot \Delta_0 \leq 1$, where Δ_0 is the mean distance between the lines of the spectrum ω_n.

When $t \ll \tau_{cor}$ the spectrum can be considered to be discrete. We find:

$$\Delta I \approx \sum_{n}^{N} \frac{F_n}{\omega_n} \left[\sin(\omega_n t + \theta_n) - \sin \theta_n \right] \tag{2.11.17}$$

whence

$$(\Delta I)^2 \approx \sum_{n}^{N} \left(\frac{2 F_n}{\omega_n}\right)^2 \cdot \sin^2 \frac{\omega_n t}{2} \cdot \cos^2\left(\frac{\omega_n t}{2} + \theta_n\right) +$$
$$+ \sum_{\bar{n} \neq n} \frac{4 F_{\bar{n}} F_n}{\omega_{\bar{n}} \cdot \omega_n} \cdot \sin \frac{\omega_{\bar{n}} t}{2} \cdot \sin \frac{\omega_n t}{2} \cdot \cos\left(\frac{\omega_{\bar{n}} t}{2} + \theta_{\bar{n}}\right) \cdot \cos\left(\frac{\omega_n t}{2} + \theta_n\right) \tag{2.11.18}$$

Let us first of all consider the most simple case of a discrete spectrum -- an equidistant spectrum ($\omega_n = \pm n \cdot \Delta_0$), corresponding to periodic motion with a period of $T = 2\pi/\Delta_0$. Let us further assume that in the limit $N \to \infty$ ($\Delta_0 \to 0$) the phases θ_n are distributed in a circle "randomly", uniformly and independently of F_n. The last condition is fulfilled, in particular, for sufficiently continuous dependence of F_n on n. Then the second sum (Σ_2) in (2.11.18) is equal to zero and $\langle \cos^2\left[(\omega_n t/2) + \theta_n\right] \rangle = \frac{1}{2}$.

The first sum is transformed in the usual way into an integral (see for example Ref. 49):

$$(\Delta I)^2 \approx 2 \int_{-\infty}^{\infty} F_n^2 \cdot \frac{\sin^2 \omega_n t}{\omega_n^2} \, dn = \pi \cdot \frac{F_\omega^2}{\Delta_0} \cdot t \tag{2.11.19}$$

whence the diffusion coefficient

$$\mathcal{D} = \pi \, J(0) \tag{2.11.20}$$

where

$$J(\omega) = \frac{F_\omega^2}{\Delta_0} \tag{2.11.21}$$

is the spectral density of the perturbation. Let us note that the value $J(\omega)$ remains approximately constant when there is local transformation of the form of the spectrum, for instance upon transition from a discrete to a continuous spectrum, provided the frequency variation is not very big. Hence it follows, in particular, that expression (2.11.20) for the diffusion coefficient is universal. The other relations sometimes mentioned in the literature in the final analysis can be transformed into the form of (2.11.20). For instance, in Ref. 74 the diffusion coefficient is determined by the sum (in our symbols): $D \sim \Sigma_n F_n^2 V(\omega_n)$, where $V(\omega)$ is the spectrum of the correlation function. In this case, however, F_n defines a spectral line of finite width $\sim \tau_{cor}^{-1}$, so that the sum $\Sigma_n F_n^2 \cdot V(\omega_n) \sim$ $\sim J(0)$ is simply the renormalization of the spectrum.

Another example is connected with the extension of expressions (2.11.20) and (2.11.21) to the case of a complete set of resonances (Section 2.7), when the amplitudes F_{mn} are of a different order for different values of m, although the corresponding frequencies ω_{mn} may be very near. In this case one can divide the components of the sum (2.11.19) (this time over m,n) into groups, in which the F_{mn} functions are fairly smooth. Relation (2.11.19) is valid for each of such groups, but the general result is given by the sum[46]:

$$ J(\omega) = \sum_{\bar{n}}{}' \left(\frac{F_{\bar{m}n}^2}{\Delta \bar{n}} \right)_{\omega_{\bar{m}n} \approx \omega} \tag{2.11.21a} $$

which as usual signifies the spectral density of the perturbation.

The result (2.11.20) was first obtained, apparently, by Bogolyubov[71]. Strictly speaking it is valid only in the limit $N \to \infty$, but it can also be used approximately for finite, sufficiently large N. However, in the latter case there is a limitation on the maximum permissible time:

$$ \tau_{min} \lesssim t \lesssim \omega_{min}^{-1} \tag{2.11.22} $$

where $\omega_{min} \sim \Delta_0$. For greater times ($t \cdot \Delta_0 \gg 1$) expression (2.11.19) becomes invalid, but from (2.11.17) it can be seen that the motion is in that case periodic ($\omega_n = \pm n\Delta_0$), so that the kinetic equation is of course inapplicable.

In order to complete the picture a lower limit was added to (2.11.22); in the present case (2.11.19) $\tau_{min} \sim \omega_{max}^{-1}$, where ω_{max} is the width of the perturbation spectrum, connected by the uncertainty relation ($\omega_{max} \cdot \tau_{in} \sim 1$) to the duration of the interaction τ_{in} -- one of the characteristic time scales introduced by Bogolyubov[54] *). For our basic model $\tau_{in} = 0$. When $t \lesssim \tau_{in}$, expression (2.11.19) is determined by the whole perturbation spectrum and not only by the value F_0. This leads to the diffusion coefficient depending on time, i.e. dynamic correlations make their appearance and the FPK equation is no longer valid.

---

*) In the general case this scale is determined by the correlation time τ_Λ (see below).

The requirements for the phases Θ_n formulated above are essential, at least the uniformity. The elementary model (2.11.1) with $k \ll 1$ (Kolmogorov stability) will serve as a most simple example. Whatever the length of the interval (perturbation period) of $T \to \infty$, the kinetic equation will not be applicable to this system, since it is not stochastic. Even in the stochastic case the value of the diffusion coefficient, generally speaking, depends considerably on the phase relations between the resonances, as will be directly seen from the examples given below.

The requirement for "randomness" of the phases θ_n can be replaced, changing over to the general case of a discrete spectrum, by a requirement for "randomness" of the frequencies ω_n. Let $\Delta\omega_r$ characterize the order of the "random" displacement of the line of the equidistant spectrum. Then the additional phase displacements $\Delta\Theta_n \sim t \cdot \Delta_r$, and for $t > \Delta\omega_r^{-1}$, the phases become "random". It is evident that in this case the lower limit of the kinetic interval (2.11.22) is:

$$\tau_{min} \sim \Delta\omega_r^{-1} \tag{2.11.22a}$$

In particular, in order to obtain the maximum interval it is necessary that $\Delta\omega_r \sim \omega_{max}$.

On the other hand, for any $\Delta\omega_r$ the phases become random asymptotically when $t \to \infty$. The interesting theorem of Kac[72] relates to this case; it states that the sum

$$\zeta = \sqrt{\frac{2}{N}} \sum_{n=1}^{N} \cos \omega_n t \tag{2.11.23}$$

is an asymptotically ($N \to \infty$, $t \to \infty$) random quantity, distributed normally with the parameters $(0,1)$ (mean and dispersion), provided the frequencies ω_n are linearly independent, i.e. if $\sum_n \lambda_n \omega_n \neq 0$; $\lambda_n \neq 0$ are integers. In this connection it should be noted that the measure of the linearly dependent frequencies is equal to zero[20].

For the kinetic equation to be valid it is necessary, however, for the following condition to be fulfilled (also when going to the limit $N \to \infty$):

$$\Delta\omega_r \gg \Delta_0 \tag{2.11.24}$$

It would be natural to call just such a spectrum (in the limit $\Delta_0 \to 0$) continuous. It is a sligthly stronger property than weak mixing (Section 2.3). The latter is equivalent to a continuous spectrum in the sense that there is no eigenfunction of Liouville's equation[26] *). From the results of this paragraph it follows that a continuous spectrum with condition (2.11.24) leads not only to a decrease of pair auto-correlations but also to integrability of the correlation coefficient, i.e. it ensures the validity of the kinetic equation (see below). In what follows, the term <u>continuous spectrum</u> should be understood to mean with condition (2.11.24).

---

*) Apparently this is equivalent also to the general notion of a continuous spectrum ($\Delta_0 \to 0$) without additional conditions imposed in the phases or frequencies.

By virtue of the foregoing, inequality (2.11.24) can be called the continuity condition of the sequence ω_n. The latter is also equivalent to the notion of a completely uniformly distributed sequence (Section 2.3).

Thus, for the kinetic equation to be applicable in a discrete spectrum the set of frequencies ω_n (or phases Θ_n) must be "continuous". Let us see what this condition means in terms of the time dependence of the perturbation $F(t) = dI/dt$ (2.11.16). We require the quantity $(\Delta I)^2 = (\int F \, dt)^2 = \int F(t)F(t')dt \, dt'$ to vary approximately in proportion to time in some interval (2.11.22). Let us introduce the autocorrelation coefficient[*]

$$\rho_\Lambda(u, t, t_1) = \left[\frac{1}{t} \int_{t_1}^{t_1+t} dt \, F(t+u) \, F(t) \right] \cdot \left(\overline{F^2} \right)^{-1} \tag{2.11.25}$$

by means of which the second moment can be presented, as usual, in the form:

$$\overline{(\Delta I)^2} = \overline{F^2} \int_0^\infty \rho_\Lambda \, du \tag{2.11.26}$$

As in the case of discrete time (Section 2.10) the latter integral should converge and should not depend on t_1, t [in the interval (2.11.22)]. In the special case:

$$\rho_\Lambda(u) = e^{-u/\tau_\Lambda} \tag{2.11.27}$$

we obtain

$$\mathscr{D} = \overline{(\Delta I)^2_1} = \overline{F^2} \cdot \tau_\Lambda \tag{2.11.28}$$

where the bar, as usual, means averaging in time. The latter expression is valid when $t \gtrsim \tau_\Lambda$ [**] and thus the lower boundary of interval (2.11.22) is now determined by the correlation time:

$$\tau_{min} \sim \tau_\Lambda \ll \omega_{max}^{-1} \tag{2.11.29}$$

This inequality is a necessary and sufficient condition for the validity of the kinetic equation for a discrete spectrum; it simply means that the interval (2.11.22) is non-null. For "typical" perturbation with "random" phases Θ_n the correlation time becomes minimal: $\tau_\Lambda \sim \omega_{max}^{-1} \sim \tau_{in}$.

---

[*] The index "Λ" signifies the linear model. The correlation coefficient for a flux is usually called the correlation function; we retain the term "correlation coefficient", however, because it is convenient to have a single designation.

[**] For continuously acting perturbation. For example, for a quasi-random model with a kick duration of τ_{in} and an interval between kicks of τ_0 we obviously have $\tau_{min} \sim \tau_0$, although $\tau_\Lambda \sim \tau_{in} \ll \tau_0$. However, in this case also the correlation time can be considered to be $\sim \tau_0$, since during this time the perturbation is equal to zero; the correlation coefficient (2.11.23), however, vanishes for $\tau_0 > u > \tau_{in}$ by virtue of the peculiarity of its definition.

By comparing expressions (2.11.29) and (2.11.22a) we arrive at the interesting estimate:

$$\tau_\lambda \cdot \Delta\omega_r \sim 1 \tag{2.11.30}$$

which shows that the "random" displacement of the line $\Delta\omega_r$ fills the role of its effective width (in the discrete spectrum!).

So far we have considered the discrete spectrum, i.e. perturbation (2.11.16) with constant frequencies ω_n and phases Θ_n. Such a case occurs, for example, in a pure linear system, say in a system of linear oscillators with linear coupling. Such a system is certainly non-ergodic, since it can be transformed into normal coordinates, i.e. into a system of independent oscillators. Nevertheless, it is possible for such a system to have the statistical behaviour described by the kinetic equation, in the interval (2.11.22). For the reasons given, the case of a discrete spectrum will be called the linear (statistical) model. As already noted, such a model was first introduced by Bogolyubov[71] and at present is the most widespread in statistical mechanics (see for instance Ref. 49). This model will be more thoroughly discussed in Section 2.13 and we will do no more than note in passing that it bears no relation to the ergodic theory and its main drawback is an upper time limit (2.11.22).

The upper limit of the interval (2.11.22) is often called the Poincaré cycle and is believed to be connected with his recurrence theorem. This conclusion is valid, however, as we shall see, only in the case of a discrete spectrum, i.e. for a linear model. Taking into account the non-linearity and the resulting stochasticity, the spectrum becomes continuous and the upper time limit (2.11.22) no longer exists.

Let us make a more detailed study of the foregoing case. Let us represent the perturbation F(t) in the form

$$F(t) = f(t) \cdot v(t) \tag{2.11.31}$$

where $f(t)$ is the given external force (divided by the frequency) having a discrete (in particular equidistant) spectrum with a mean distance between lines of Δ; $v(t)$ is the velocity of the oscillator. It is obvious that the formation of a continuous spectrum is connected precisely with the last quantity $v(t) = v_0(t) \cdot \cos\phi(t)$ and is the result of the mixing process in the phase $\phi(t)$.

Taking into account the fact that for small perturbation $v_0(t)$ varies insignificantly as compared to the phase, the velocity correlation can be expressed through the phase correlation (2.11.5):

$$\rho_v(u) = \frac{\overline{v(t+u) \cdot v(t)}}{\overline{v^2}} \sim \rho_u^{(2)}(p,q) \tag{2.11.32}$$

An explicit estimate of the latter correlation coefficient cannot be given in the general form, but from the consideration mentioned at the beginning of this section it follows that $\rho_u^{(2)}$ decreases with u exponentially with a characteristic time, determined by the "typical" estimate of the K-entropy (2.11.3).

The perturbation correlations F(t) depend, generally speaking, in a complicated way on the linear correlations (2.11.25) of the force f(t) and the non-linear correlation of the velocity v(t) (2.11.32). Let us consider the two limiting cases.

First let

$$\tau_\Lambda \cdot h \ll 1 \qquad (2.11.33)$$

We will call this case quasi-linear by analogy with the corresponding approximation in plasma wave theory[56]. In this approximation the correlations are determined by the external force: $\rho(F,F') \approx \rho(f,f')$ when $u \leq T [f = f(t); \quad f' = f(t + \mu)]$ and decrease with a characteristic time τ_Λ (see above). When $u \geq T$ the linear correlations $\rho(f,f')$ increase again on account of the quasi-periodicity of f(t). In the simple case of a periodic force f(t), the correlation coefficient $\rho(f,f')$ is also periodic. This leads to a strong increase in linear correlations in the intervals: $kT < u < kT + \tau_\Lambda; \quad k = 1, 2, \ldots$. It is here that the velocity correlation: $\rho(F,F') \approx \rho(v,\sigma')$ becomes significant with a characteristic time $\tau_{cor} \sim h^{-1}$ (2.11.3). The schematic variation of the total correlation coefficient is shown in Fig. 2.11.2 as a continuous line; the dotted line represents the non-linear correlations $\rho(v,v')$.

It can be said that in the quasi-linear case there is a region of applicability of the linear model, confirmed for t > T by the non-linear model.

In the opposite limiting case

$$\tau_\Lambda \cdot h \gg 1 \qquad (2.11.34)$$

the linear model is not applicable at all, but the decrease of the correlations is characterized by the dotted line in Fig. 2.11.2, provided the following additional condition is fulfilled

$$\omega_{max} \gtrsim h \sim \Omega_{\phi\Sigma} \qquad (2.11.35)$$

The physical meaning of this condition is that there must be several renormalized resonancés which destroy each other. In the opposite case only one renormalized resonance is formed and the maximum perturbation frequency $\sim \omega_{max} \ll \Omega_{\phi\Sigma}$, so that only a narrow stochastic layer forms near the separatrix of this resonance (2.6.16).

Let us note that conditions (2.11.35), (2.11.33) and (2.11.34), are generally speaking independent, since $\tau_\Lambda^{-1} \sim \Delta\omega_r \lesssim \omega_{max}$. But for "typical" ("random") initial phases Θ_n we have: $\Delta\omega_r \sim \omega_{max}$, so that the developed stochasticity corresponds only to the quasi-linear case (2.11.33).

In conclusion, let us consider a few examples of calculating the diffusion coefficient.

Let us begin with the basic model (2.1.11) which was thoroughly studied in the previous Section. Let us express the diffusion coefficient for it through the continuous time (flux) parameters, introducing the δ-function into the transformation. Let us assume $h_\Theta(\Theta) = \cos \Theta = \cos \omega t$ to be definite, then $F_n = \varepsilon/T$, and from (2.11.20) we obtain:
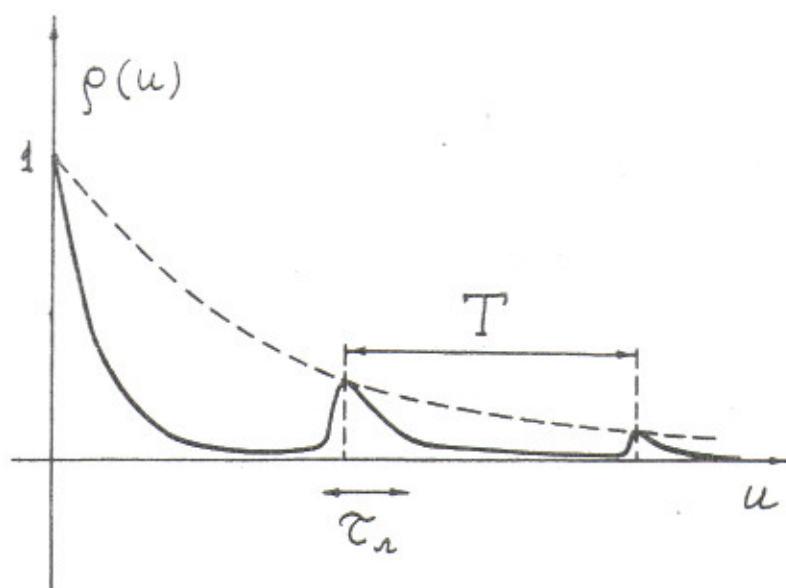
Fig. 2.11.2: Schematic diagram of the autocorrelations of the perturbation F(t) in the quasi-linear case. The dotted curve represents the nonlinear correlations $\rho(v,v')$ with a characteristic time $\sim h^{-1} \sim \Omega_\phi^{-4/3} \cdot \Delta_0^{1/3}$; the continuous curve represents the total correlations $\rho(F,F')$, determined in the interval of applicability of the linear model (t < T) by the linear correlations $\rho(f,f')$ with a characteristic time $\tau_\Lambda \sim \Delta\omega_r^{-1}$.

$$\mathcal{D} = \frac{\varepsilon^2}{8T} \tag{2.11.36}$$

which agrees exactly with (2.10.12), if it is taken into account that T characterizes the duration of one step.

The diffusion coefficient can be estimated in another way, which can be useful in some cases. Keeping in mind the picture of "touching" renormalized resonances, we can write:

$$\mathcal{D}_\omega \sim (\Delta\omega)^2_\Sigma \cdot \Omega_{\phi\Sigma} \sim \Omega_{\phi\Sigma}^3 \sim \frac{\Omega_\phi^4}{\Delta} \sim \frac{\varepsilon^2}{T}(\omega')^2 \tag{2.11.37}$$

The index ω shows that the diffusion coefficient is obtained according to the frequency. The latter expression shows that it agrees in order of magnitude with (2.11.36).

The first of the estimates (2.11.37) is valid for any "typical" system, and the second only for resonances in the first harmonic: $p\omega = q$; $p = 1$. Comparing it with the "typical" estimate of K-entropy (2.11.3), we arrive in this case at an interesting relation:

$$\mathcal{D}_\omega \sim h^3 \tag{2.11.37a}$$

Let us note that if ω is not a canonical variable it is necessary to change over to the I diffusion coefficient: $D_I \equiv D = D_\omega \cdot (dI/d\omega)^2$, before inserting it in the kinetic equation.

Let us consider the quasi-random model

$$\varphi' = \varphi + \frac{k}{2\pi} \cdot \sin 2\pi (p\psi + q)$$
$$\psi' = \psi + \varphi'$$

modified so as to obtain the resonances of the pth harmonic: $p\omega = q\Omega$, where $\omega = 2\pi\dot{\psi}/T$; $\Omega = 2\pi/T$. The distance between resonances, i.e. between the values of the perturbation frequency: $\Delta_0 = 2\pi/T$. In terms of continuous time [compare with (2.11.14)] we have: $\Omega_\phi^2 \sim pF_0 \sim pk/\sqrt{T}$; $(\Delta\omega)_H^2 \sim k/p\sqrt{T}$. Renormalization gives (2.11.3a): $\Omega_\Sigma \sim (pk)^{2/3}$; $(\Delta\omega)_\Sigma \sim k^{2/3} \cdot p^{-1/3}$. Using the general estimate (2.11.37) we find: $D_\omega \sim (\Delta\omega)_\Sigma^2 \Omega_\Sigma \sim k^2$, which agrees in order of magnitude with the general formulae [(2.11.20) and (2.11.21)]: $D_\omega \sim F_0^2/\Delta_0 \sim k^2$.

Let us now study the diffusion by periodic crossing of the resonance.

The simplest case is fast crossing of the resonance, which is described by transformation (2.9.7) of the basic model type. From expression (2.10.12) we find directly (for symbols see Section 2.9):

$$\mathcal{D}_\omega = \frac{\langle (\Delta\omega)_1^2 \rangle}{8\pi} \cdot \Omega_0 = \frac{\Omega_0}{8} \cdot \frac{\Omega_\phi^2}{V} \tag{2.11.38}$$

or in the flux parameters:

$$\mathcal{D}_\omega \sim \frac{\Omega_k^4}{\Omega_0} \sim \frac{\varepsilon^2 F^2}{\Omega_0} \cdot (\omega')^2 \tag{2.11.39}$$

The latter estimate agrees with the general expression (2.11.20). However, the region of applicability of this expression, as for the previous example, is determined by the condition $t > T = 2\pi/\Omega_0$, so that the quasi-linear region in the present case is completely absent because of the special phase relations between the resonances.

. For slow crossing of the resonance the frequency change is almost constant and almost reversible (Section 1.5). Let us therefore consider the total frequency change $\Delta\omega_\pm$ per modulation period, i.e. after two crossings (there and back). From expression (2.9.10) we find

$$\Delta\omega_\pm \approx 2V\Omega_\phi \ell n \left(\frac{V + \xi_+}{V + \xi_-} \cdot \frac{\sqrt{4\pi V} - \xi_+}{\sqrt{4\pi V} - \xi_-} \cdot \frac{\sqrt{4\pi V} + \xi_+}{\sqrt{4\pi V} + \xi_-} \right) - \frac{\Omega_\phi}{\pi} \left(\xi_+^2 - \xi_-^2 \right) \tag{2.11.40}$$

Since the ξ distribution is now already completely non-uniform (see Section 2.9), without further calculation one can obtain only the estimate:

$$\mathcal{D}_\omega = \frac{(\Delta\omega_\pm)^2}{8\pi}\Omega_o \sim 16\,\Omega_o\,V^2\Omega_\phi^2\,\big|\ell n\,V\big|^2 \sim$$

$$\sim 16\,\frac{(\Delta\Omega)^{3/2}}{\Omega_k^2}\,\Omega_o^{7/2}\,\big|\ell n\,V\big|^2 \sim \frac{16\,|\ell n\,V|^2(\Delta\Omega)^{3/2}\,\Omega_o^{7/2}}{|\varepsilon F\omega'|}$$

$$(2.11.41)$$

Here the numerical factor takes into account the number of combinations from the logarithmic terms in (2.11.40).

Let us note that for slow crossing the dependence of the diffusion coefficient on the parameters (2.11.41) is roughly the inverse of what it is for fast crossing (2.11.39). In particular there is the seemingly paradoxical result that the diffusion coefficient decreases with the increase of the perturbation. The explanation is that the parameter V decreases simultaneously (with a given $\dot{\Omega}$), and with it also the phase interval ξ, in which a difference effect (2.11.40) ensues.

Thus for periodic crossing of the resonance the diffusion rate as a function of $\Omega_k^2 \sim F$ (the perturbation) or V (rate of crossing) has its maximum in the region of $V \sim 1$.

Above, we ignored the possibility of capture upon slow crossing of the resonance. In fact, when $V \to 0$ capture plays a considerable part (Section 1.6) and the diffusion process becomes extremely complicated. On the one hand, as a result of the decrease in the phase oscillations the process is found to be partly reversible and this leads to a reduction of the diffusion. On the other hand, stable capture is possible only for one direction of resonance crossing, and this causes a systematic frequency displacement. The biggest possible value of the diffusion coefficient can, apparently, be estimated as:

$$\mathcal{D}_{max} \sim (\Delta\Omega)^2 \cdot \Omega_o \qquad (2.11.42)$$

It does not depend on the perturbation at all.

2.12 Many-dimensional non-linear oscillator. Arnold diffusion

In this section we shall try to extend the results obtained above to a many-dimensional autonomous system consisting of a number (N) of weakly coupled non-linear oscillators. As noted in Section 1.1, a many-dimensional oscillator can be reduced in a first approximation to a one-dimensional non-autonomous oscillator of type (1.1.1). Since, as Anosov[31] showed, a stochastic system is coarse[*], the higher approximations cannot jeopardize the stochasticity or even substantially change its parameters. However, such a conclusion does not apply to the region of the Kolmogorov stability, where the many-dimensional system is considerably different from a one-dimensional one[20].

Let us first consider a more accurate transition to a one-dimensional oscillator in a first approximation. The essential difference from the given external perturbation is that the perturbation frequencies now are not constant, since they represent a combination of the frequencies of other non-linear oscillators.

---

[*] i.e. structurally stable; a notion introduced by Andronov and Pontryagin[84]. See also Ref. 85.

The main resonance of a many-dimensional oscillator is the so-called simple resonance, or resonance with a multiplicity of one, i.e. simply one resonant relation between all the frequencies:

$$\sum_{i}^{N} n_i \omega_i (I_j) \equiv (n, \omega) = 0 \qquad (2.12.1)$$

This relation defines the series of intersecting surfaces (for arbitrary integers n_i) in the momentum space or a beam of planes intersecting at the origin in the frequency space. In what follows, frequency space should always be understood, unless a special reservation is made.

Let us write the equations of motion in the form

$$\dot{I}_k = - \varepsilon \frac{\partial H(I_j, \theta_i)}{\partial \theta_k} = - i\varepsilon \sum_{(n)} n_k \tilde{H}_{(n)} e^{i(n,\theta)}$$

$$\dot{\theta}_k = \omega_k (I_j) + \varepsilon \frac{\partial H(I_j, \theta_i)}{\partial I_k} \qquad (2.12.2)$$

where $H = \Sigma_{(n)} \tilde{H}_{(n)} \cdot e^{i(n,\theta)}$ is the Hamiltonian; all the quantities without indices and in brackets represent N-dimensional vectors, and $k, j = 1, 2, \ldots, N$. Ignoring in the first approximation the term with $\partial H/\partial I_k$ (Section 2.3) and introducing the resonance phase:

$$\psi_{(n)} = (n, \theta) \qquad (2.12.3)$$

we can write near one of the resonances in the absence of overlapping:

$$\dot{I}_k \approx - i\varepsilon n_k \tilde{H}_{(n)} \cdot e^{i\psi_{(n)}}$$

$$\dot{\psi}_{(n)} = (n, \omega) \qquad (2.12.4)$$

and the phase equation is:

$$\ddot{\psi}_{(n)} \approx - \left(i\varepsilon \sum_{k,j} n_k \cdot \frac{\partial \omega_k}{\partial I_j} \cdot n_j \tilde{H}_{(n)} \right) e^{i\psi_{(n)}} = - i \Omega_{(n)}^2 e^{i\psi_{(n)}} \qquad (2.12.5)$$

where $\Omega_{(n)}$ is the phase oscillation frequency of the resonance concerned, which is determined by a certain mean non-linearity for all the oscillators. The phase oscillations change the plane (2.12.1) into a resonant layer with a thickness of

$$(\Delta \omega)_{(n)} \sim \frac{\Omega_{(n)}}{|(n)|} \qquad (2.12.6)$$

where $|(n)|$ is the modulus of the vector (n). The latter expression generalizes the notion of the width of a one-dimensional resonance (Section 1.4).

A resonance with a multiplicity k occurs when there is simultaneous fulfilment of k resonant conditions (2.12.1), which occurs in the region of intersection of k resonant layers. The motion near the multiple resonance is described by a system of k phase equations of type (2.12.5):

$$\ddot{\psi}^{\,\ell} = -i \sum_j \Omega_{\ell j}^2 \, e^{i\psi^j} \tag{2.12.7}$$

where the indices $1, j = 1, 2, \ldots, k$ signify the number of the vector (n), and the phase frequencies are given by the expressions:

$$\begin{aligned}
\Omega_{\ell j}^2 &= \sum_{i,k} n_i^{\ell} \cdot \frac{\partial \omega_i}{\partial I_k} \cdot n_k^j \cdot \widetilde{H}^j \\
\dot{I}_k &= -i\varepsilon \sum_{\ell} n_k^{\ell} \cdot \widetilde{H}^{\ell} \cdot e^{i\psi^{\ell}} \\
\dot{\psi}^{\ell} &= (n^{\ell}, \omega)
\end{aligned} \tag{2.12.8}$$

In particular, for small oscillations $e^{i\psi j} \rightarrow \mp \psi^j$ (the sign depends on the choice of a stable or unstable fixed point respectively, see Section 1.4) and a system of linear equations is obtained.

As already noted, in the stochastic case there is no reason to expect any new effects, but it is of interest to obtain many-dimensional estimates in a more explicit form.

First let us consider the total set of resonances and determine the critical value of the smoothness parameter of the ℓ_c (Section 2.7). As in the one-dimensional case it depends on the convergence of the sum $\Sigma_s \sim \Sigma_{(n)} (\Delta\omega)_{(n)} \sim \Sigma_{(n)} \Omega_{(n)}/|(n)|$. As usual we shall consider only first order resonances and moreover restrict ourselves to non-renormalized resonances. As explained in Section 2.7, in this way we obtain the lower estimate for ℓ_c, which supplements Moser's upper estimate[28]. What follows depends on the type of perturbation chosen. Let us study two cases. In the first we put:

$$\widetilde{H}_{(n)} \sim \prod_i n_i^{-(\ell_i + 3)} \tag{2.12.9}$$

This expression generalizes the one-dimensional (2.7.6) and means that the smoothness of the perturbation is characterized independently for each degree of freedom by its own parameter ℓ_i. Inserting (2.12.9) in (2.12.6) and using (2.12.5) we find:

$$\Sigma_s' \infty \sum_{(n)}{}' \left(\sum_{i,k}{}' n_i n_k \cdot \prod_j n_j^{-(\ell_j + 3)} \right)^{1/2} \Big/ |(n)| \tag{2.12.10}$$

where only the quantities on which the convergence of Σ_s depends are left. The convergence is determined by the limit $n_i \rightarrow \infty$. For a given (ω) all $n_i \propto n_k \propto |(n)|$. Therefore the sum $\Sigma_s \propto \Sigma_{(n)} n_j^{-(\ell_j + 3)/2}$ converges, if:

$$\ell = \ell_{min} > -1 \tag{2.12.11}$$

which agrees with the result of Section 2.7 for continuous time (2.7.31).

However, the perturbation can also be determined otherwise:

$$\tilde{H}_{(n)} \sim |n|^{-(\ell+3)} \qquad (2.12.12)$$

where $|n| = \Sigma_i |n_i|$. Moser accepted just this determination[28] (see also Ref. 20). It signifies that the combined derivatives of the force are continuous only up to number ℓ inclusive. The sum in which we are interested now takes the form:

$$\Sigma_s \infty \sum'_{(n)} \left(\sum_{i,k} n_i n_k |n|^{-\ell-3} \right)^{1/2} \Big/ |(n)| \infty \sum'_{(n)} |n|^{-\frac{\ell+3}{2}}$$

and converges under the condition:

$$\ell > 2N - 3 \qquad (2.12.13)$$

Comparing this value with Moser's result[28], we find that the critical value of the smoothness of the perturbation lies in the interval:

$$2N - 3 < \ell_c \leq 2N + 2 \qquad (2.12.14)$$

Let us note that the width of the interval for ℓ_c, which is five if ℓ is a real number, and four if ℓ is an integer, agrees with the width for the one-dimensional transformation: $1 < \ell_c \leq 6$ (Section 2.7). This question will be further discussed in the section devoted to numerical experiments (Section 3.3).

Let us now make a more accurate estimate of the border of stochasticity according to the overlapping of the first order resonances. We shall limit ourselves to the case of almost harmonic oscillations, i.e. we shall assume that in a zero approximation there is only the basic harmonic ω_i for each degree of freedom, and the amplitude of the higher harmonics $n_i \omega_i$ are of the order of ε^{n_i-1}.

Let us further assume that an m-fold interaction takes place[*]. When calculating the number of frequencies of the perturbation (and of the resonances) only those combinations of oscillators which include the one whose motion interests us should be taken into account. Then in a first approximation over ε the quantity of resonances is:

$$N_1 = C_{N-1}^{\bar{m}-1} \cdot 2^{\bar{m}-1} \approx \frac{(2N)^{\bar{m}-1}}{(\bar{m}-1)!} \qquad (2.12.15)$$

This expression is easily extended to the case when there are n_0 first harmonics of comparable amplitude:

$$N_1 = C_{N-1}^{\bar{m}-1} \cdot \frac{(2n_0)^{\bar{m}}}{2} \approx n_0 \cdot \frac{(2n_0 N)^{\bar{m}-1}}{(\bar{m}-1)!} \qquad (2.12.15a)$$

In the second approximation (ε^2) the number of resonances increases considerably for two reasons: firstly on account of the higher harmonics, and secondly on account of the appearance in each of the oscillators of the basic harmonic of all the rest.

---

[*] i.e. direct interaction only between m degrees of freedom.

The first effect can be ignored when $2m \sqrt{\epsilon} < 1$. Indeed, it increases the number of resonances $2m$-fold, but in return the width of the resonance is reduced by $\sqrt{\epsilon}$ times. In the opposite case $(2m \sqrt{\epsilon} \gg 1)$ the total width of the resonances increases the maximum by $2m! \; \epsilon^{m/2} \sim (m \sqrt{\epsilon})^m \cdot e^{-m}$ times independently of N. This effect is neglected in what follows.

The second effect is the main one and it leads to an increase of the number of resonances by $m \cdot 2^{m-1} N^{m-1}/(m-1)! \sim N_1$ times, so that $N_2 \sim N_1^2$. Similarly it can be shown that $N_k \sim N_1^k$. It is not difficult to verify that the position of the border of stochasticity by order of magnitude is the same in any approximation. Indeed, $\Sigma_s^{(k)} \propto \epsilon^{k/2} N_k \propto (\sqrt{\epsilon} N_1)^k \sim 1$. The latter estimate follows from the fact that $\sqrt{\epsilon} N_1$ is a dimensionless quantity (for further details, see below). It is true that this cannot be said of the stochasticity parameter $s^{(k)} \propto \Sigma_s^{(k)}$, which diverges when $k \to \infty$, if $\sqrt{\epsilon} N_1 > 1$. However, this divergence is fictitious since it is necessary to renormalize the resonances (see below).

Thus it is sufficient for us to determine the border of stochasticity in the first approximation. In the subsequent estimates we shall encounter sums of the form: $\Sigma_0 = \Sigma_i^m n_i f_i$, where $n_i = \pm 1$ and f_i are some quantities. For such sums we shall take a "typical" estimate, corresponding to a "random" set n_i: $\Sigma_0 \sim f \sqrt{m}$, where f is a certain mean value of f_i. By putting: $\partial \omega / \partial I \sim \alpha \omega / I$, $\tilde{H}_{(n)} \sim \omega I$, we obtain for an m-fold interaction: $\Omega_{(n)} \sim \omega (\epsilon \alpha m)^{\frac{1}{2}}$ (2.12.5); $|(n)| \sim \sqrt{m}$ and $(\Delta \omega)_{(n)} \sim \omega (\epsilon \alpha)^{\frac{1}{2}}$; $\Omega \sim \omega \sqrt{m-1} \sim \omega \sqrt{m}$ is the frequency interval of the quantity $\Sigma_i^m n_i \omega_i$ occupied by N_1 resonances. The border of stochasticity is determined by the relation: $s \sim N_1 \cdot (\Delta \omega)_{(n)} / \Omega \sim 1$ or[*]

$$(\varepsilon \alpha)_s \sim \frac{\bar{m}}{N_1^2} \approx \frac{\bar{m} \cdot [(\bar{m}-1)!]^2}{(2N)^{2\bar{m}-2}} \qquad (2.12.16)$$

Thus the critical value of the perturbation decreases, at least $\propto N^{-2}$ ($m = 2$). Hence it is clear that macroscopic molecular systems -- typical objects of statistical mechanics -- are always far inside the region of stochasticity. In particular the size of the non-ergodic component decreases, at least $\propto N^{-6}$ (Section 2.8).

Let us note that for the normalization used above $(\tilde{H}_{(n)} \sim \omega I)$ the small parameter ϵ characterizes only one resonance. Since the total number of resonances with $N_1 \gg 1$ (2.12.15) is very great, the following additional condition must be fulfilled:

$$\zeta^2 \equiv \varepsilon^2 N_1 \ll 1 \qquad (2.12.17)$$

---

[*] It is interesting to note that the value $(\epsilon \alpha)_s$ is here m times greater than in Ref. 76 (see also Ref. 13). This difference can be explained by the fact that in Ref. 13 and 76 the border of stochasticity was determined from the overlapping of the resonant layers in N-dimensional space. However, this condition is not sufficient, since the phase oscillations inside the layer take place in a fully determined direction $(\Delta \omega_{(n)})$ (2.12.4), while the motion along the layer is, generally speaking, slow (see below). Therefore for stochasticity it is necessary for the chain of vectors $(\Delta \omega_{(n)})$ corresponding to the various resonances, to be closed or, in other words, for the one-dimensional widths of the resonances to overlap (2.12.6).

where the new parameter ξ characterizes the total perturbation, written by means of Parseval's equality.

Let us now find the renormalized width of the resonance $(\Delta\omega)_{(n)\Sigma} \equiv (\Delta\omega)_{\Sigma}$ which, according to the results of Section 2.11 determines[**] the dynamical time scale τ_n (local instability, correlations, mixing). Renormalizing in the "typical" case with a power $n = 4$, we obtain: $(\Delta\omega)_{\Sigma} \sim (\Delta\omega)_{(n)}^{4/3} \cdot \Delta^{-1/3}$ (2.11.3). But according to our previous estimates $(\Delta\omega)_{(n)} \sim \omega\sqrt{\epsilon\alpha}$ and:

$$\Delta \sim \frac{\omega\sqrt{\bar{m}}}{N_1} \sim \omega\sqrt{\bar{m}} \cdot \frac{(\bar{m}-1)!}{(2N)^{\bar{m}-1}} \qquad (2.12.18)$$

Whence we find:

$$(\Delta\omega)_{\Sigma} \sim \omega \cdot \bar{m}^{-1/6} \cdot (\alpha\xi)^{2/3} \qquad (2.12.19)$$

Let us verify that the renormalized width of the resonance does not diverge in the higher approximations: $(\Delta\omega)_{\Sigma}^{(k)} \sim (\epsilon^{2k}N_k)^{1/3} \sim \xi^{2k/3}$, but according to condition (2.12.17) $\xi \ll 1$ and this series rapidly converges.

According to the results of Section 2.11 the entropy and diffusion coefficient[*] are expressed in the "typical" case through the renormalized width of the resonance (2.11.37)[**]:

$$\mathcal{D}_\omega \sim h^3 \sim (\Delta\omega)_{\Sigma}^3 \sim \frac{\omega^3}{\sqrt{\bar{m}}}(\alpha\xi)^2 \qquad (2.12.20)$$

The diffusion coefficient depends only on the square of the total perturbation (ξ^2), which acts as if it were completely random. In Section 2.5 it was said precisely in this sense that stochastic instability leads to the most rapid diffusion possible for the given perturbation.

Let us now go over to the region of Kolmogorov stability. As already noted (Section 2.2) the situation in this region is essentially different from the one-dimensional case. The most important difference is that the invariant tori, whose dimensionality is obviously equal to N, do not divide the (2N - 1)-dimensional energy surface in the phase space of the system[20] [***]. In the momentum (frequency) space, the invariant tori are represented simply by points distributed among the everywhere dense web of interwoven and intersecting resonant surfaces (planes) (2.12.1).

Each of the resonant surfaces represents, as we know, a layer of thickness (2.12.6) inside which are invariant tori similar to the ones outside (this already follows from the results of Arnold's paper[77] and has been thoroughly investigated by Moser[78]), but outside is the stochastic layer (Section 2.6). It is precisely these numerous intersecting

---

*) In the general case we have the diffusion tensor (2.10.15); the following estimate relates to the diffusion along one of the axes: $\langle(\Delta\omega)_i^2\rangle = 2D_\omega \cdot t$; for the total N-dimensional vector: $\langle|(\Delta\omega)|^2\rangle = 2ND_\omega \cdot t$.

**) When $n_i \sim 1$ (2.11.3a).

***) Let us recall that the case we call one-dimensional is that where N = 2 for an autonomous system (Section 2.2).

stochastic layers which form an unstable (ergodic) component of the motion in the region of Kolmogorov stability. The first example of such instability was studied by Arnold[21] and subsequently it was learnt that a similar instability mechanism is very general for many-dimensional motion[79] (see also Section 2.6). It leads to a peculiar diffusion along the system of intersecting resonances, which we shall henceforth call Arnold diffusion.

In order to understand the mechanism of Arnold diffusion let us return to the basic equations of the many-dimensional resonance (2.12.1) to (2.12.5). Let us note first of all that for each given resonant term the phase factor $e^{i(n,\theta)}$ is identical for all the components of the vector (I) and therefore the variation (ΔI) is directed along the vector (n). For the main resonance, which we will call guiding resonance henceforth this gives the direction of the phase oscillations, and the phase factor takes the form: $e^{i\psi(n)}$, where $\psi_{(n)}$ is the resonance phase (2.12.3). Each of the remaining terms in (2.12.2) characterizes the perturbation of the resonant torus directed parallel to its vector (n). The phase factor for each of these perturbing resonances can be written in the form: $e^{i(\omega_1 t+\psi(n))}$ where $\omega_1 = (\Delta n, \omega)$ is the detuning of the frequency in relation to the guiding resonance. If the system is inside the guiding resonance, the perturbation of the neighbouring resonances leads only to the deformation of the resonant torus. However, at the edge of the resonance, inside the stochastic layer, the increments of the integral: $\int dt\, e^{i(\omega_1 t+\psi(n))} \sim \mu_n\, \Omega_{(n)}^{-1}$; $\mu_n \sim e^{-c/s_n}$; $s_n \sim \Omega_{(n)}/\omega_1$ for each half-period of the phase oscillations with a frequency of $\Omega_{(n)}$, form a random sequence on account of the random phase shift of the phase oscillation with respect to the perturbation. A thorough analysis of this stochasticity mechanism was made in Section 2.6. The momentum perturbation is of the order of

$$|\Delta I|_{nn'} \sim \varepsilon_{n'}/\mu_n \cdot \Omega_{(n)}^{-1} \tag{2.12.21}$$

where the vectors (n), (n′) relate to the guiding and perturbing resonance respectively; $\varepsilon_{n'} \sim \varepsilon\, \tilde{H}_{(n')}$.

As already mentioned above, the direction of the vector $(\Delta I)_{nn'}$ is along (n′) and generally speaking this is not identical to the direction of the stochastic layer. Diffusion (2.12.21) is therefore possible only over a small distance of the order of the thickness of the stochastic layer. For long-distance diffusion at least two perturbing resonances with non-parallel $(n')_1$, $(n')_2$ are necessary. Then one of them will certainly have a component along the stochastic layer, which is the one that gives Arnold diffusion proper, and the other will have a component across the layer, which ensures reflection from the border of the layer. The diffusion will thus go along the line of intersection of the stochastic layer with the plane of the vectors $(n')_1$, $(n')_2$. For the diffusion to go in any direction along the (N - 1)-dimensional resonant surface, there must obviously be N linearly independent perturbing resonances.

This is possible only for a non-autonomous system, i.e. under external perturbation. If the system in question is closed there are (N - 1) perturbing resonances only since they form, together with the guiding resonance, the full set of N linearly independent resonances. Thus, for a closed system Arnold diffusion can go only along some (N - 2)-dimensional surface which must be, obviously, the intersection of the stochastic layer with the energy surface of the system. No other limitations of the Arnold diffusion seem to exist. At least we can assume it as a hypothesis which is in accordance with Poincaré's theorem[36] on the absence,

in the general case, of analytical integrals of motion except the energy (see also Ref. 152 and Section 2.6). In the light of the KAM theory it is natural to assume that the destruction of all the other integrals occurs precisely in the stochastic layers of the resonances as a result of Arnold diffusion.

The diffusion coefficient can be roughly estimated as: $D_A \sim |\Delta I|^2_{nn'} \cdot \Omega_{(n)}$, or taking into account (2.12.5) and (2.12.21):

$$D_A \sim I^2 \omega \cdot \frac{\varepsilon_{n'}^2}{n\sqrt{\alpha \varepsilon_n}} \cdot e^{-2c/s_n} \tag{2.12.22}$$

In order to obtain more explicit estimates of the diffusion rate let us put:

$$\varepsilon_n \sim \varepsilon \cdot e^{-\bar{m}n/2n_0} \tag{2.12.23}$$

where n is now the maximum harmonic number. We consider that for each degree of freedom the amplitude of the perturbation harmonic decreases as e^{-n/n_0}, the interaction being m-fold (Section 2.12). It would be more accurate to write: $\exp(-\Sigma_{i=1}^{m} |n_i|/n_0)$, but for our rough estimates we shall put: $\Sigma_i^m n_i = \bar{m}n \approx mn/2$, whence (2.12.23) also follows [*]. In fact the parameter m now characterizes the number of frequencies for which harmonics are taken. The diffusion coefficient is determined, mainly by the exponent, which has the form[**]:

$$\exp\left(-\frac{n'\bar{m}}{n_0} + \frac{n}{4n_0} - \frac{2c\, e^{\bar{m}n/4n_0}}{n\sqrt{\alpha\varepsilon}\cdot(n')^{N-1}}\right)$$

The argument of the exponent reaches a maximum when:

$$n' = \left(\frac{2c(N-1)n_0 e^{\bar{m}n/4n_0}}{\bar{m}n\sqrt{\alpha\varepsilon}}\right)^{\frac{1}{N}} \gtrsim n \tag{2.12.24}$$

The latter inequality is necessary for the validity of the approximation used (Section 2.6). It is violated in a certain interval n, which can be determined from the condition:

$$\left(\frac{\varepsilon_s}{\varepsilon}\right)^{\frac{1}{2N}} < \frac{\bar{m}n}{2n_0} \cdot e^{-\frac{\bar{m}n}{4n_0 N}} \tag{2.12.25}$$

Here we used estimate (2.12.29a) for ε_s. If $\varepsilon/\varepsilon_s$ is not too small, the unknown interval is:

$$n'_c = \frac{2n_0}{\bar{m}} \lesssim n \lesssim \frac{4n_0 N}{\bar{m}} = n_c \tag{2.12.25a}$$

When $\varepsilon/\varepsilon_s < (2N/e)^{2N}$ always n' > n. In the interval (2.12.25a) optimal n' = n (Section 2.6) and in the exponent only the first term may remain. Indeed the second term in the exponent is always small in comparison with the first, and the relation of the third to the first is: (n'/n)/(N - 1). On the border, when n' = n, the third term gives the correction factor [N/(N - 1)]. With only the first term remaining in the exponent, we obtain:

---

[*] This estimate makes sense if different frequencies and n_i are of the same order of magnitude.

[**] Here we use estimate (2.12.28) for ω_1, see below.

$$\mathscr{D}_A' \sim I^2 \omega \cdot \frac{\varepsilon^{3/2}}{n\sqrt{\alpha}} \cdot e^{-\frac{\bar{n}n}{n_0}} \tag{2.12.26}$$

From the validity condition of this expression (2.12.25) it is seen that always $D_A' < D_A$ (2.12.29). This estimate on both borders of interval (2.12.25a) changes approximately into estimate (2.12.29), which is easy to verify, using (2.12.25).

Let us now turn to the region $n > n_c$ (or $n < n_c'$). It should be explained that in the majority of problems one is required to estimate the rate of Arnold diffusion due to the resonances with a given n, or more precisely an n of less than a certain value. The last condition determines the mean distance between resonances Δ_n. The specific form of the function $\Delta_n(n)$ depends on the form of the interaction, see for example formulae (2.12.15a) and (2.12.18) which are valid for $m \ll N$.

Now let us consider the opposite limiting case $m \approx N$. A rough estimate can be made as follows. The total number of different combinations of the components of the vector (n) is $(2n)^N$, since each component can assume values from $-n$ to n. Assuming that for large n the distribution of the vectors (n) and (ω) is on the average isotropic, one can estimate the mean distance between resonances as[*]:

$$\Delta_n \sim \frac{\omega}{n^N} \tag{2.12.27}$$

The main error of this estimate is due to the non-uniform density of the resonances (see Fig. 4.3.1), which can be taken into account in (2.12.27) by introducing a special factor: $\omega \rightarrow k\omega$.

Expression (2.12.27) gives, in particular, the estimate of the perturbation frequency ω_1, which we used above:

$$\omega_1 \sim n \cdot \Delta_n \sim \frac{\omega}{n^{N-1}} \tag{2.12.28}$$

The mean "gap" between the resonances Δ_n determines the density of the network of stochastic layers along which the Arnold diffusion spreads. From estimate (2.12.22), taking into account (2.12.23) and (2.12.24) we obtain

$$\mathscr{D}_A \sim I^2 \omega \cdot \frac{\varepsilon^{3/2}}{\sqrt{\alpha}} \cdot \mathscr{M}_n ; \tag{2.12.29}$$

$$\mathscr{M}_n = n^{-1} exp \left\{ -\frac{4N}{\bar{n}(N-1)} \left(\frac{2cn_0^2(N-1)}{n\bar{u}^2} \right)^{\frac{1}{N}} \left(\frac{\varepsilon_s}{\varepsilon} \right)^{\frac{1}{2N}} e^{\frac{n\bar{u}}{4n_0 N}} - \frac{n}{4n_0} \right\}$$

where the critical value ε_s is determined from the expression:

$$\mathscr{Q}_{(n)} \sim \frac{2n_0}{\bar{u}} \cdot \sqrt{\alpha \varepsilon_s} \cdot \omega \sim \omega \left(\frac{\bar{u}}{2n_0} \right)^N \tag{2.12.29a}$$

In the latter estimate we used the effective value of the parameter n_0, which follows from the appearance of the exponential factor (2.12.23) $e^{-mn/2n_0}$, whence $n_{0eff} \approx 2n_0/m$.

_____

[*] We omit here the numerical factor since it depends greatly on particular specification of the set of resonances in question.

Thus the rate of Arnold diffusion decreases with the growth of n according to the double exponent law. This clearly characterizes the degree of Kolmogorov stability in the many-dimensional case. Let us recall that in the one-dimensional case the stability is eternal (Section 2.2). In practice a dependence of type (2.12.29) determines a limit:

$$\Pi_2 = \frac{\bar{m}\eta}{4 n_o N} \sim 1 \tag{2.12.29b}$$

beyond which the rate of Arnold diffusion becomes unobservably small. It should be noted that this limit corresponds just to the condition $n \sim n_c$ (see above). This means that estimate (2.12.29) can in practice be used only near the border (2.12.29b).

It turns out, however, that in some cases more rapid diffusion along the set of resonances is also possible. The appearance of the double exponent in (2.12.29) is due, as we saw, to the fact that the value $s_n \sim \Omega_{(n)}/\Delta_n$ in estimate (2.12.22) itself becomes exponentially small, since Ω_n decreases with the growth of n exponentially, and Δ_n only as n^N (2.12.27). But this does not apply to resonances with a multiplicity of two, i.e. at the intersection of two resonant surfaces. In this case $s_n \sim 1$ always and the exponent disappears from the estimate (2.12.22). Furthermore, since the majority of resonances with a multiplicity of two consists of resonances of the same order of n ($n \sim n_{max}$), total destruction of the resonance takes place, i.e. the width of the stochastic layer becomes of the same order as the width of the resonance itself. Thus there forms a relatively wide channel along which the diffusion spreads at a comparatively high rate. In order to distinguish this special kind of diffusion we shall call it streamer diffusion. This name is connected with the fact that for the minimal dimensionality, when this diffusion is possible, the stochastic layer of a resonance with a multiplicity of two in the frequency space has the shape of a narrow tube (streamer) along which comparatively fast diffusion spreads, a picture which recalls streamer breakdown in gas.

Streamer diffusion is possible only when resonances with a multiplicity of two form an intersecting network in the frequency space. From geometrical considerations it is clear that this is possible under the condition:

$$N_a \geqslant 4; \qquad N_H \geqslant 3 \tag{2.12.30}$$

where N_a, N_H is the number of degrees of freedom (of the dynamical frequencies) for an autonomous and non-autonomous system respectively. Thus for streamer diffusion one more degree of freedom is required than for ordinary Arnold diffusion.

An estimate of the velocity of streamer diffusion is obtained from (2.12.22) and (2.12.23), taking into account that $s_n \sim 1$; $n \sim n'$:

$$D_c \sim I^2 \omega \cdot \frac{\varepsilon^{3/2}}{n \sqrt{\alpha}} \cdot e^{-\frac{3}{4} \cdot \frac{\bar{m}\eta}{n_o}} \tag{2.12.31}$$

The quantity D_c, of course, also decreases rapidly with the increase of the harmonic number of the resonance, but not as catastrophically as D_A. Except for a numerical factor ~ 1 in the exponent, D_c is identical with D_A' (2.12.26). Both mechanisms give roughly the same diffusion rate when $n \sim 1$, provided it comes within the region (2.12.25). In the opposite case streamer diffusion proves to be even slower in this region. It is significant, however, that the law (2.12.31) is valid with any n, whereas in ordinary Arnold diffusion a double exponent appears for large n. Therefore streamer diffusion plays an important part only in the region $n \gtrsim n_c$ (2.12.25a).

Let us note that for streamer diffusion two resonances are sufficient, instead of three as for ordinary Arnold diffusion (see above). This is due to the fact that both resonances now coincide in space and their vectors (n) are always non-parallel. However, in the present case the requirement for a component of the vector (ΔI) [or (n)] along the streamer is non-trivial. In particular, this condition does not apply when (ΔI), ($\Delta\omega$) are parallel: $(\Delta I)||(\Delta\omega)$. Since $(\Delta I)||(n)$ (2.12.2), from the resonance condition (2.12.1) it follows that: $(\Delta\omega, \omega) = 0$, i.e. the vector ($\Delta\omega$) is perpendicular to the resonant plane, and that means also to the streamer, so that diffusion does not occur. Since the non-linearity matrix $\partial\omega_i/\partial I_k = \partial^2 H/\partial I_i \partial I_k$ is symmetrical and can be transformed to the principal axes, the condition for non-parallelism of the vectors (ΔI), ($\Delta\omega$), which is necessary for streamer diffusion, amounts to a requirement for the eigenvalues of the non-linearity matrix to be different.

Resonances with a multiplicity > 2 do not lead to qualitatively new effects.

The diffusion coefficient (2.12.29), like (2.12.31), does not yet determine real diffusion in the momenta space. Indeed, Arnold diffusion spreads along the resonant surfaces, which in the general case form a very complicated system; in places where the surfaces intersect, "random" (on account of the stochasticity of the motion) transition from one surface to another will take place, so that, as a whole, Arnold diffusion represents a combination of two random processes: diffusion along the stochastic layer and transition from one layer to another. If the mean length between two intersections ℓ_I is sufficiently small, the total length of the diffusion trajectory L along the system of intersecting layers can be estimated by the ordinary formulae of the random walk theory[43]:

$$L \sim (\Delta I)^2 / \ell_I \qquad (2.12.32)$$

where ΔI is the total variation of the momentum in the diffusion process. Then the diffusion time can be estimated according to the formula:

$$t_A \sim \frac{L_i^2}{D_A} \sim \frac{\omega}{\Delta_n^2} \cdot \left(\frac{\Delta I}{I}\right)^4 \cdot \frac{\alpha^{5/2}}{\varepsilon^{3/2} \mu_n} \qquad (2.12.33)$$

where we put $\ell_I \sim \Delta_n/\omega' = \Delta_n \cdot I/\alpha\omega$.

Let us note that the law of this "double diffusion" is unusual, since the diffusion time t_A is proportional not to the square but to the fourth power of "spacing" ΔI. Let us introduce the "double diffusion" coefficient:

$$\mathcal{D}_{\mathcal{A}} \equiv \frac{d}{dt} \langle (\Delta I)^4 \rangle \sim I^4 \cdot \frac{\Delta_n^2}{\omega} \cdot \frac{\varepsilon^{3/2} \cdot \mathcal{M}_u}{\alpha^{5/2}} \qquad (2.12.34)$$

The relations obtained remain valid also for streamer diffusion, substituting $D_A \to D_c$.

The rate of Arnold diffusion (2.12.29) decreases exponentially with the decrease of the stochasticity parameter $s^2 \sim \varepsilon/\varepsilon_s$. Moreover, the diffusion takes place only for special initial conditions, the relative measure of which $\sim \delta \ll 1$ (when $s < 1$; Section 2.6). However, the total system of resonances, and that means also the stochastic layers, is everywhere dense. Therefore the problem of the motion in the region of Kolmogorov stability is asymptotically (when $t \to \infty$) improper, since any arbitrarily small variation in the initial conditions displaces the trajectory from the stable component to the stochastic one and vice versa. Let us note that in the present case we cannot simply average over a small volume of phase space, as was done when solving the kinetic equation. This is due to the fact that in a large part of the phase space the system is stable and therefore such averaging does not correspond to any real process in the system, and reference to the "practical" uncertainty of the initial conditions is insufficient in mechanics.

It is possible, however, to regularize the problem as follows. Let us add to the dynamical system some "external" diffusion process with a diffusion coefficient D_0. For example, in the case of the motion of a particle in a magnetic trap (Section 4.4) the scattering always present in residual gas is such a process. This additional diffusion eliminates the singularity of the initial conditions and, moreover, enables us to neglect the resonances of very high harmonics, leaving only a finite number of resonances. However, contrary to the behaviour in a stochastic region, the motion will now substantially depend on the additional diffusion, also in the limit $D_0 \to 0$.

The diffusion process will take place in two stages. In the first there occurs "external" diffusion with a coefficient D_0 up to the nearest resonant surface, i.e. over a distance $\sim \ell_I$. In the second stage the "external" diffusion occurs "parallel" to the Arnold diffusion. In the most simple case, when the number of resonances is not great, so that $\ell_I \sim (\Delta I)$, one can neglect the "double diffusion" and assume that the diffusion along the stochastic layers with a coefficient D_A is roughly the same as the diffusion in I. Then the total diffusion coefficient in the second stage of the process is $D_0 + D_A \cdot \omega$, where $\omega < 1$ is a reduction factor of the diffusion rate, because the system spends only a small part of the time inside the stochastic layer. If $n' \sim n$, estimate (2.6.13) gives $\Delta \sim s_n \cdot e^{-c/s_n}$. Comparing this with estimate (2.12.22) one can conclude that the reduction factor ω is equivalent to some change of numerical coefficient in the first exponent of the expression for D_A^-.

Let us note that the total diffusion coefficient depends essentially in any case on the "auxiliary" parameter D_0. The latter should not be too large, otherwise the stochastic layers stop functioning at all. The critical value of D_0 is determined from the condition of leaving the layer in a time of the order of one phase oscillation, which is just the order of the diffusion rate in the stochastic layer (Section 2.10). Hence the conditions for the existence of Arnold diffusion:

$$\mathcal{D}_o \lesssim \mathcal{D}_A \tag{2.12.35}$$

However, an observable effect of this diffusion takes place for considerably smaller $D_0 \lesssim D_A \, \omega$.

Let us note that the dependence of the diffusion time on the perturbation parameter will have the characteristic shape of a transition curve with two plateaux for small and large perturbation. In fact, in both limits the diffusion coefficient is equal to D_0, but the diffusion distance is different: (ΔI) and ℓ_I, respectively. The ratio of the diffusion times at the plateaux is therefore:

$$k \sim \left(\frac{\Delta I}{\ell_I}\right)^2 \tag{2.12.36}$$

Let us now consider the more interesting case of a large number of resonances: $\ell_I \ll (\Delta I)$, when "double diffusion" takes place $[(2.12.33),\ (2.12.34)]$. This means that the kinetic equation takes on a more complex form than usual (2.10.10). In view of the roughness of the estimates relating to "double diffusion", we shall not solve this equation but will use the simple estimate for ordinary diffusion: $d(\Delta I)^2/dt \sim D_0$ and for "double diffusion": $d(\Delta I)^4/dt \sim D_D$ (2.12.34). Hence the total diffusion rate is:

$$\frac{d(\Delta I)^2}{dt} \sim \mathcal{D}_o \left(1 + \frac{\gamma}{(\Delta I)^2}\right); \quad \gamma = \frac{\mathcal{D}_D \cdot \omega}{2\mathcal{D}_o} \tag{2.12.37}$$

By integrating this equation in the most simple case $\omega = $ const, we obtain the total diffusion time:

$$t_A \sim \mathcal{D}_o^{-1}\left[(\Delta I)^2 - \gamma \ln\left(1 + \frac{(\Delta I)^2 - \ell_I^2}{\gamma}\right)\right] \tag{2.12.38}$$

which again depends essentially on D_0.

Some experimental data on Arnold diffusion will be given in Sections 3.6 and 4.4.

For streamer diffusion the picture remains qualitatively the same, but the effective diffusion coefficient $D_C^0 = D_C \cdot \omega$ decreases not so much:

$$\mathcal{D}_c^o \sim \mathcal{D}_c \cdot s_n^2 \sim I^2 \omega \cdot \frac{\varepsilon^{5/2}}{\alpha^{3/2}} \cdot n^{\mathcal{N}-1} \cdot e^{-\frac{5}{4} \cdot \frac{\pi n}{n_\omega}} \qquad (2.12.39)$$

as D_A (see above). In the latter estimate we used relation (2.12.31) and the expression for s_n:

$$s_n \sim \frac{\Omega_n}{\Delta_n} \sim \sqrt{\frac{\varepsilon}{\alpha}} \cdot n^{\mathcal{N}} \cdot e^{-\frac{\pi n}{2 n_o}} \qquad (2.12.40)$$

The ratio of the diffusion times on the plateaux will be smaller than (2.12.36), since instead of the "absorbent" resonance surfaces which are necessarily intersected in the diffusion process D_0, there are now "absorbent" tubes (streamers), which may be by-passed ("missed"). However, the increase in the lifetime is slight, since the probability of a miss rapidly decreases as the streamer is approached.

As a model, one can examine the diffusion between two concentric cylinders with absorption only on the inner one. Simple calculations show that the diffusion time is proportional only to the logarithm of the ratio of the cylinder radii. In our case this relation $\sim s_n$, since the mean distance between streamers is of the same order as that between the resonance surfaces. Thus, instead of (2.12.36) we find:

$$k \sim \left(\frac{\Delta I}{\ell_i}\right)^2 / |\ell_n s_n| \qquad (2.12.41)$$

In conclusion, let us make a few remarks about the kinetic equation in the many-dimensional case. If the number of degrees of freedom is not great, the distribution function as usual describes the ensemble of identical N-dimensional systems in the 2N-dimensional Γ-space. The equation for such a distribution function is called the master equation[49]. However, as a rule the complex systems of statistical mechanics consist of a very large number ($n \to \infty$) of identical elements ("particles"), interacting with each other. In this case a new possibility appears: besides the master equation one can write the equation for a so-called single particle distribution function describing the density of "particles" in the phase space of one "particle", which is called μ-space. Since the total number of particles n is always finite, then in μ-space (as in Γ-space) only a coarse-grained distribution function has direct physical meaning (Section 2.10).

In a similar way one can introduce the kinetic equation for s-particle phase density ($s \ll n$), describing the distribution of the subsystem of s "particles".

For a many-dimensional oscillator a "particle" is a one-dimensional oscillator, weakly coupled to the others, for instance a phonon in a crystal lattice. If such a one-dimensional oscillator is considered to be non-autonomous with a given external perturbation, we arrive at the master equation in the most simple Γ-space (Section 2.10). But if the same one-dimensional oscillator is assumed to be a "typical" representative of the system of interacting oscillators, we obtain the kinetic equation in μ-space. In both cases one of course obtains the same equation (Section 2.10) and the only difference is the physical meaning of the phase density.

2.13 Remarks on the nature of statistical laws

Since we have to do with statistical mechanics and in particular kinetic equations, it is difficult to resist the temptation to make a few general remarks on the nature of statistical laws, irreversibility and other such problems that are still somewhat mysterious. It is hoped that these remarks will not prove to be a mere repetition of well-known arguments. In this question we have the advantage of the detailed investigation made in this paper into the transition from dynamical to statistical behaviour for a very simple, probably the simplest, mechanical system -- the elementary model, which represents a one-dimensional non-linear oscillator under the action of given periodic perturbation.

To the main question of whether the motion of such a system is a "true" random process we reply in the affirmative, unlike many other authors engaged in investigating this problem. Among them is Krylov[30], whose point of view in other respects is very close to ours and whose ideas are in fact extended and developed in the present paper.

If this assertion is accepted, it opens the way to a general explanation of the statistical laws of nature on the basis of the classical mechanical model. In this case the statistical laws are valid in a dynamical system, in so far as the motion of the system is stochastic in the sense given to this term in the present paper (Section 2.3). This point of view is perfectly natural at present for mathematicians with their ergodic theory (see, for example Ref. 42) but, strangely enough, apparently alien to physicists, in any case in current statistical mechanics[27,49] in which the so-called linear model exercises completely sway (see Section 2.11 and below).

The most unexpected result of the above point of view proved to be the possibility of statistical behaviour of extremely simple systems right down to the elementary model, which has only one degree of freedom. However, for the ergodic theory this was not unexpected. Hopf[29] already pointed out this possibility for a system with two degrees of freedom, although it was only recently that Sinai succeeded in demonstrating the stochasticity of the motion of a real mechanical system -- a system of hard balls in a box, which in the simplest case has only two degrees of freedom[106]. This result sharply contradicts the idea which is of very long standing in physics, that the statistical laws are valid only in a very complex system with an enormous number of degrees of freedom $N \to \infty$.

Let us now turn to the basic assertion made above, that the statistical laws correspond to a certain special case of motion of a classical mechanical system, namely stochastic motion.

There are two kinds of possible objection to this assertion. One of them, the less important, is connected with the "islets" of stability which always exist in the stochastic region of the elementary model (Section 2.8). This means that as a rule one always finds special initial conditions of finite, even though very small, measure, for which the motion is not stochastic. With regard to this objection it can be said only that such stable regions, generally speaking, really exist and can be observed for simple systems (Section 3.5). Let us note, however, that such "islets" of stability are characteristic just for oscillatory

systems, for instance, for a crystal lattice, if one turns to the typical macroscopic molecular objects of statistical physics. At the same time there are no stable regions at all in gas, and probably in liquid. At least this was demonstrated by Sinai for the gas model as a system of hard balls, mentioned above[106]. Furthermore, even for oscillatory molecular macroscopic systems the stable regions are extremely small (Sections 2.8, 2.12):

$$\frac{\Delta I}{I} \lesssim N^{-\gamma} \tag{2.13.1}$$

At this point one is tempted to use phase space quantization to prove that very small ΔI are completely impossible[50]. Such a proof is not possible, however, as explained by Krylov[30]. Roughly speaking, it is a question of the type of description of the motion changing simultaneously with the quantization, namely it is necessary to change over from the phase space of classical mechanics to the Hilbert space of the wave functions, in which the motion of the quantum system is described, as usual, by a trajectory.

Finally, and this is our main argument, both types of statistical physics system (with and without stable regions) always interact with each other through molecular collisions, and also through electromagnetic (thermal) radiation. Under these conditions the stable regions can remain only for very simple macroscopic or molecular systems with a small number of degrees of freedom (2.13.1), and during a short interval of time as compared to the time of relaxation with the surrounding medium. Similar effects are actually observed, in particular, in so-called unimolecular reactions, for instance thermal dissociation, if the number of atoms in the molecule is greater than 2. At present there are two contradictory theories on such reactions, one of which [that of Landau[153] and Kassel[154]] is based on the unlimited stochasticity of intramolecular motion, whereas the other [that of Slater[155]], on the contrary, assumes the existence of a full set of integrals of motion. In reality, as confirmed by direct experiments and numerical calculations[156], when collisions are rare a certain intermediate case occurs that is typical for a system with divided phase space (Section 2.5). In order to avoid confusion, let us point out once again that divergences from statistical behaviour are limited in this example by the very short interval of time between two successive collisions.

The second considerably more profound objection to our general conception of the statistical laws is connected with the very nature of mechanical motion as motion along a trajectory reversible in time. This problem has been most thoroughly studied by Krylov[30]. It is also closely connected with the Loschmidt paradox, arising from the contradiction between the dynamical reversibility and the statistical irreversibility of the motion.

Krylov's main objection to the classical mechanical model of the statistical laws amounts to the following. Since dynamical motion is reversible, its irreversible statistical properties (for instance the increase or decrease of the entropy) will wholly depend on the initial conditions. Let us further consider the usual organization of a statistical experiment as a multiple repetition of a process under given macroscopic initial conditions. Then in order to obtain statistical behaviour the distribution function of the microscopic initial conditions for a given macroscopic state must be uniform in sufficiently small

regions of the phase space. Meanwhile the evolution of the distribution with time leads on the other hand, to an increasingly singular state which is easy to verify by investigating the inverse motion. Thus we find it virtually impossible not only to prove, but even to introduce a postulate concerning the initial microscopic conditions.

Sometimes this objection is "developed" still further and it is asserted that from the point of view of the dynamical model the probability of the entropy increasing or decreasing is generally identical, since by virtue of Liouville's theorem the phase volume of the two states with different entropy through which the system passes in the process of motion is identical. It will be easy for us to begin by refuting this explicitly incorrect assertion, which is in fact based on a misunderstanding. The point is that if the term "phase space volume" is interpreted literally, in any real situation it is equal to zero, since we always have a finite number of systems (and of particles in a system), and also a finite number of repetitions of the experiment (see Section 2.10). If some indefinite phase volume is introduced, determined approximately according to a finite number of points in phase space, it is not conserved in the process of motion.

Let us examine this latter case more thoroughly, using as an example a system of $N \gg 1$ oscillators of the type of the elementary model, weakly interacting with each other. We shall describe the state of this system in μ space (Section 2.12), which in the present case is identical with the phase square of the elementary model. The state of the system is represented by N points in this square. Let them be distributed statistically uniformly and independently so that the system initially occupies "all" the phase space. This does not prevent it, however, from congregating in the process of motion in a very small region ξ of the phase square. Let us estimate the probability of this, assuming that the region ξ has a simple form. Let us consider the inverse process of mixing, when the region ξ is transformed into a system of narrow strips with an over-all area of ξ, uniformly distributed over the square (Fig. 2.4.1). It is evident that all the N points must lie on one of the strips of this system: the probability of this is $\omega_\xi = \xi^N$ or, if we are interested in the congregation in any small region ξ (decrease of the entropy): $\omega = \xi^{N-1}$. But this is exactly the probability corresponding to the fluctuation. Thus Liouville's theorem does not contradict the smallness of the fluctuations, which lead, in particular, precisely to a decrease of the entropy.

The example taken is also the answer to Krylov's second objection regarding the evolution of the distribution function. Undoubtedly, the continuous [fine-grained (Section 2.10)] distribution function tends with time to become singular, as can easily be understood from the picture of the mixing in Fig. 2.4.1. However, as already noted, such a function does not correspond to any real experiment, i.e. it is essentially unobservable and should therefore be excluded from the theory and replaced by a coarse-grained distribution function (Section 2.10). The latter tends to become uniform according to the ergodic theory. This gives the possibility not only of eliminating the contradiction pointed out by Krylov, i.e. of introducing a postulate concerning the initial microscopic conditions, but also opens the way to proving this postulate. It is now natural to assume that the initial state

(t = 0) of the process concerned (t > 0) in a system is determined by the finite state of the previous process (t < 0) in the same system.

Here, however, there are two difficulties. The first is due to the fact that uniform distribution occurs, according to the ergodic theory, only in the limit $t \to \infty$. This difficulty is not important, since when there is positive K-entropy the mixing process runs exponentially fast and in practice is completed in a comparatively short time. However, there is still the other difficulty, connected with the organization of the statistical experiment. As noted above, multiple repetition of the process under study is assumed, with the same macroscopic conditions. This requirement is not very definite, in the sense that it does not mention the microscopic state. This gives rise at least to the suspicion, if not the certainty, backed essentially by our somewhat hazy idea of our freedom of will, that we can "create" any initial micro-state and so obtain any course of the process in contradiction to the statistical experiment. Of course, on the other hand there exists an intuitive idea that the microscopic co-ordinates are in fact "inaccessible" to the experimenter, so that in practice it is not in our power to influence the microscopic state of the macroscopic system. But this does not constitute a proof, and in any event there is always the chance that we shall somehow learn how to do this in the future, or, to quote a popular modern catch-phrase: "Nothing is impossible for science!"

It seems to us that these doubts can be banished on the basis of an analysis of the most important property of a stochastic system -- the local instability of motion developing exponentially with time. It is not difficult to verify that owing to this property there is no system in nature that is closed in relation to its dynamical motion, except the whole Universe. As an example let us consider the motion of gas molecules in a model with hard balls of radius r, with a mean free path ℓ. Let us study the perturbation of this motion by the gravitational field of a single proton at the "other end" of the Universe, i.e. at a distance $R \sim 10^{28}$ cm. Taking into account the fact that the perturbation is tidal and that a change in the gravitational field of the proton is essential on account of its displacement, we obtain the additional angular deviation of the gas molecule:

$$(\Delta\theta)_o \sim \frac{\varkappa \bar{u}}{R v^2} \cdot \left(\frac{\ell}{R}\right)^3 \tag{2.13.2}$$

where κ is the gravitational constant and v the velocity of the molecule. This perturbation will grow according to the law:

$$(\Delta\theta)_n \sim (\Delta\theta)_o \cdot \left(\frac{\ell}{r}\right)^n \tag{2.13.3}$$

where n is the number of successive collisions of the molecule and in

$$n_1 \sim - \frac{\ell_n (\Delta\theta)_o}{\ell_n \ell/r} \tag{2.13.4}$$

collisions becomes $(\Delta\theta)_1 \sim 1$, i.e. the trajectory of dynamical motion changes considerably. If one takes a gas under normal conditions: $m \sim 10^{-24}$ g; $v \sim 10^5$ cm/sec; $\ell \sim 10^{-5}$ cm;

$r \sim 10^{-8}$ cm, then estimate (2.13.4) gives: $n_1 \approx 60$, which requires only $\sim 10^{-8}$ sec. This limiting example easily shows that from the point of view of molecular dynamics there is only one closed system -- the Universe as a whole, which naturally also includes the experimenter. The latter thus has no control either over his own or any other microscopic state. This state is determined by the initial conditions of the Universe at $t = -\infty$, and not at all by "creating" the initial state in a specific statistical experiment. The violation of statistical laws in such a model in an infinite interval of time is possible therefore only for initial conditions of zero measure. The fact that the Universe is not in this special state is the minimal hypothesis of our model.

Being minimal, this hypothesis is not trivial, for the same reasons as those behind Krylov's second objection: when $t \to +\infty$ the phase point of the system tends towards a certain exceptional position, whereas its initial value ($t \to -\infty$) should not be exceptional. It seems to us, however, that this difficulty is psychological rather than physical. The point is that the two exceptional regions ($t \to \pm\infty$) are completely different. Formally they differ only (!) by the change of the sign of all the velocities, and this has a kind of hypnotising effect. But we know that even negligible variation of the initial conditions of a stochastic system leads to a complete change in the trajectory of motion. With regard to the above-mentioned exceptional regions when $t \to \pm\infty$, then as can easily be seen from the picture of mixing in Fig. 2.4.1, they represent two systems of intersecting strips. Therefore any exceptional region $t \to +\infty$ is uniformly distributed over all the exceptional regions $t \to -\infty$, in other words it completely loses all its exceptionality when the velocity is reversed.

It seems to us that in this lies the answer to the Loschmidt paradox concerning irreversibility in statistical mechanics.

With regard to the predominant direction of the thermodynamical processes in the Universe, this is determined by its strongly macroscopic non-uniformity of cosmological origin. The most important thing here is the dominating role of gravitational interaction in the Universe. When there is such interaction there is no steady state at all, on account of so-called collapse (unlimited contraction), which terminates the development of both the Universe as a whole and of individual sufficiently massive stars[157]. Let us note in passing that the absence of thermodynamical equilibrium makes thermal death of the Universe impossible. As far as we know, this simple consideration was put forward very recently by Zel'dovich and Novikov[157]. If singularity does not cease to exist upon the collapse of the Universe*), a state will occur which could be called the cosmological death of the Universe. If singularity ceases, as one can, apparently, conclude from the work of Lifshits, Sudakov and Khalatnikov[159] **) then the Universe periodically has the chance of starting life "all over again" (and of course of making a better job of it!).

---

*) For the closed model of the Universe, which in our opinion is more probable[158].

**) See moreover p. 551, of Ref. 157.

Let us now briefly discuss a few possibilities of eliminating the minimal hypothesis formulated above. Is such a hypothesis really necessary? Can one select such special initial conditions, even of zero measure, that statistical behaviour proves impossible regardless of the stochasticity of the system? It may turn out that such conditions simply do not exist independently of their measure. In order to understand this, let us return to the numerical experiment with the elementary model and assume that there are absolutely no stable regions in it, as in the case of molecular collisions. Since this relates to <u>real</u> numbers there are always exceptional initial conditions for which the motion does not obey any statistical law. The measure of such trajectories is of course equal to zero, but they exist. However, a numerical experiment is always limited in principle by <u>rational</u> numbers because of the finite number of digits of the computer mantissa, the measure of which is also equal to zero. But two sets of zero measure and completely different nature certainly do not intersect. In any event here there is a theoretical possibility of rigorous proof that <u>all</u> initial conditions lead to stochastic motion. The question arises as to whether the same effect does not also occur in nature as the result of space-time quantization, if the latter really exists. The answer to it is not at all evident, as was seen in the example of phase-space quantization mentioned above. However, this possibility is not excluded. It is interesting that in this case even a reversal of the velocity does not lead to violation of the statistical laws, since the trajectories of the forward and return motion are not at all identical, on account of "round-off" ("quantization"). It is interesting to note that the motion is nevertheless in a certain sense reversible, since the dynamical equations including the "round-off" procedure do not change when time is reversed.

It is significant that the absolute value of the space-time quantum in practice is of no importance, as a result of the exponential development of the local instability of the stochastic motion. So, for example, even though the quantum has an order of gravitational length ($\sim 10^{-55}$ cm for an electron) its influence on the dynamics of the motion of the gas under normal conditions (see above example) will be effective already after ~ 16 collisions, or $\sim 10^{-9}$ sec.

A new peculiar phase-space quantization has been studied by Krylov[30]. As noted above (Section 2.3) the usual quantization ($\Delta\mu \sim \hbar$) does not lead to the expected effect, as a result of the change in the type of description. In order to avoid this difficulty, Krylov put forward the hypothesis that macroscopic systems do not have a definite Ψ function, because of the special complementarity assumed by Krylov between the microscopic (quantum) state and its macroscopic (thermodynamical) characteristic. This leads to the quantization $\Delta\mu \gg \hbar$ and gives the possibility of explaining the statistical behaviour even in the classical formulation of the problem (see above). This hypothesis cannot be examined more thoroughly here, since the present paper is restricted exclusively to classical mechanics. Let us only say that the development of this hypothesis seems to us extremely interesting and, furthermore, that it means essentially that quantum mechanics is inapplicable to a macroscopic system and consequently that there is no continuous transition through a quasi-classical region.

The other possibility of eliminating the minimal hypothesis is connected with the fact that the exact dynamical laws of nature may prove to be irreversible in time. Now this is one of the possible explanations of the anomalous decay of the K-meson[160]. It is again significant that under conditions of exponential local instability of the stochastic system arbitrarily small irreversibility of the dynamical equations might be sufficient.

Sometimes it is assumed that quantum-mechanical motion is essentially irreversible as a result of the so-called "reduction" of the Ψ function (wave packet) by measurement, i.e. by interaction with a macroscopic object, which is not described by the Schroedinger equation[27]. In fact in present-day quantum mechanics there is no clear understanding of the process of measurement, so that this whole question remains open. However, in our opinion there is a more plausible hypothesis, which is in a sense the opposite: that the "reduction" of the Ψ function is itself due to the statistical properties of the macroscopic measuring apparatus. This hypothesis is based on the following consideration. "Reduction" of the Ψ function must not necessarily be accompanied by the transformation of the original Ψ function into one of the states whose superposition it was before the measurement. It is sufficient for the original pure (or coherent) state to have been transformed into the mixed (or incoherent) one, i.e. for the phase relations between the superposed states to have become indefinite. In the latter case there is no interference between the states and the Ψ function gives the classical probability, when the system "in reality" is in one of the states before, but we do not know exactly which. In contrast to this, before the measurement the system was "in reality" in all the states simultaneously (pure state). But the destruction of the phase relations between the superposed states, necessary for the transformation of the pure state into the mixed one, is also apparently inevitable as a result of the inter-action of the micro-system with the statistical apparatus. This hypothesis of course needs detailed investigation, which does not come within the scope of the present paper. Let us point out that this problem has been discussed for quite a long time in the literature (see, for example, Ref. 162).

In conclusion let us compare our point of view with the linear model of present-day statistical mechanics, introduced by Bogolyubov[71] and more fully developed by Prigogine and his school[49] (see also Ref. 161). As already mentioned in Section 2.11, the latter model does not need ergodicity but it is valid, i.e. it leads to statistical laws, rigorously speaking, only in the limit of a very large number of degrees of freedom $N \to \infty$. For a finite N the application of the model has an upper time limit (Section 2.11). For macro-scopic molecular systems this upper limit is very great[163] and in practice is insignificant. It is essential, however, that even when $N \to \infty$ a statistical description is possible only for the small sub-system $N_1/N \to 0$ [49]. This condition can be formally satisfied by taking only the retarded solutions of Liouville's wave equation (2.10.1) (with the additional requirement N, $V \to \infty$; $N/V = $ const., V is the volume of the system) thus excluding the reaction of the whole system on the sub-system studied. For the reason mentioned, this condition is sometimes incorrectly linked with the principle of causality[49]. It seems to us better to speak of definite initial conditions [the absence of incoming (advanced) waves],

since in fact the total solution of Liouville's equation (retarded + advanced) must also satisfy the principle of causality.

In the light of the above condition it can be said that the linear model makes it possible to obtain the Gibbs canonical distribution for a sub-system in a thermostat without a microcanonical distribution of the whole closed system. The sub-system achieves statistical behaviour because of the additional demands on the parameters of the whole system (thermostat) of the type of a requirement for random phases or frequencies (Section 2.11).

A linear model is possible in statistical physics and is very convenient by virtue of the relative simplicity of its mathematical technique. However, it is not necessary and, in fact, does not correspond to real molecular dynamics, since real macroscopic molecular systems are stochastic (Section 2.12). This is contradictory to Prigogine's assertion that such systems are not ergodic and have a full set of integrals of motion[49]. The inaccuracy of the latter assertion is evident, if only from Sinai's example[106]. The origin of the error is that the series representing these integrals, generally speaking, diverge. According to the KAM theory they converge only when the perturbation is sufficiently small, outside the stochastic region. Let us note in passing, that in the stochastic region, and more precisely, even for weak mixing (Section 2.3), Liouville's equation (2.10.1) does not have eigenfunctions at all (except a constant).

Nevertheless, there is a region in which the linear model is very important. Let us explain by taking a gas as an example. In his approach to statistical mechanics Bogolyubov[54] introduced two characteristic time scales: the duration of the interaction upon collision (τ_{in}) and the time between collisions (τ_{cl}). The latter turns out to be just of the order of the mixing time ($\sim h^{-1}$) [30]. Therefore our non-linear model works only for $t \gg \tau_{cl}$, i.e. only in a diffusion (hydrodynamical) region, where gas relaxation (diffusion) takes place in co-ordinates (compare with the diffusion of the basic model, scale τ_D, Section 2.10). Gas relaxation in momenta takes place just in a time $\sim \tau_{cl}$, so that it cannot be described by a non-linear model. In the best case the latter gives only the order of the relaxation time[30]. For the basic problem of the foundation of statistical physics such a lower time limit is unimportant -- what is more important is the absence of an upper limit. However when it comes to applications it is very important to extend the region of applicability of the kinetic equation in the direction of lower times. This can be done precisely by means of the linear model with an additional special limitation on the conditions of the system. The most general limitation of this type was obtained by Sandri[63] and called by him "the principle of the absence of parallel motion", which means the absence of strong correlations at the initial moment (t = 0). According to our way of thinking, this "principle" can be validated on the basis of the previous motion of the system (t < 0) taking into account the mixing. Let us note that "the absence of parallel motion" according to Sandri does not at all mean, as is sometimes supposed[55], the total absence of collective processes. It is only necessary for there also to be random relative motion of the particles, or more precicely, for the pair correlations to grow no faster than v_{12}^2 when $v_{12} \to 0$, where v_{12} is the relative velocity of two particles[63].