COOPERATIVE EFFECTS IN PLASMAS

B. B. Kadomtsev

1. Preliminaries

1.1. Plasma states

The term "plasma" was proposed by Langmuir for a mixture of electrons and positively charged ions to indicate its unusual physical properties. Plasma exists in such circumstances when a gas can be ionized. A very small quantity of ions exist even in a candle flame. But to be considered a plasma the density of the charged gas component must be not too small.

The level of gas ionization can be characterized by the ratio of the ion density to the neutral atom density. When this ratio is not too low and the ion interactions with the electromagnetic fields become important, the charged components of the ionized gas can be considered a plasma. There are many examples of plasma states. The most familiar is ordinary lightning: plasma here is produced for a very short time interval during the giant electric discharge between a cloud and the ground. The mysterious ball lightning is probably another member of the natural plasma states family.

Our Sun and all of the stars are giant plasma spheres. The Sun's surface displays a lot of very complicated physical phenomena reflecting the very rich internal life of solar plasma. No less complicated plasma phenomena occur in near space on the boundary and inside the Earth's magnetosphere. The corresponding plasma activity is produced by the interaction of the solar plasma wind with the magnetic field of the Earth.

Plenty of different plasma patterns are produced artificially by man in various specific devices. The most familiar are the

V. D. Shafranov (ed.), *Reviews of Plasma Physics*

glow discharges used for the illumination of large cities. Similar discharges can be used in gas lasers and in plasma processing devices for the microelectronic industry.

A new branch of plasma physics appeared when the idea of controlled nuclear fusion was proposed. A lot of different concepts for hot plasma magnetic confinement were suggested. A very broad research programme over the world was established and continues to develop with large investment of resources. This research has revealed a lot of new cooperative phenomena in magnetically confined plasmas. All these phenomena are essentially nonlinear. That is why plasma physics is in fact the physics of nonlinearities.

The physics of nonlinear phenomena is complicated enough. To understand it is not a simple matter. To facilitate the discussion of nonlinear effects it is reasonable to start with a very simple subject and then to proceed to more and more complicated phenomena. We shall start our consideration with nonlinear ion sound waves.

1.2. Ion-sound waves

1.2.1. Equations of motion

Ion sound waves can be observed experimentally under very simple conditions. There are many devices in which a lowly ionized plasma can be produced either with the help of an external plasma source or by direct gas ionization in a glow or high-frequency discharge. Such a plasma has usually a low density and its electrons have a Boltzmann distribution of velocities with a temperature of a few electron-volts. As for the ions, they have a very low temperature due to collisions with the neutral gas atoms. Their temperature can be considered as approximately zero compared with the electron temperature.

We shall consider one-dimensional motions when all quantities depend on one coordinate x. Let us assume that the ions are shifted along this coordinate. If the electron shift does not coincide with the ion shift, then some electric field E appears

which is produced by the plasma itself. If ion collisions with neutral gas atoms are not frequent and can be neglected, the ion motion equation can be written in the following form:

$$m_i \frac{dv}{dt} = eE. \tag{1.1}$$

Here m_i is the ion mass and e is its electrical charge, v is the x-component of ion velocity.

The ion density n_i evolves according to the continuity equation:

$$\frac{\partial n_i}{\partial t} + \frac{\partial}{\partial x}(n_i v) = 0. \tag{1.2}$$

Consider now the electron behavior. The motion of ions is slow so that during the compression and decompression the electron component has enough time to equilibrate its temperature along the x-axis. Thus, we can assume the electron temperature $T_e = \text{const}$. The inertial force is negligible for the same reasons so that electrons are in the state of gas dynamic equilibrium:

$$T_e \frac{\partial n_e}{\partial x} = -eEn_e. \tag{1.3}$$

Here n_e is the electron density. The electric field E is produced by the ion and electron charge separation and is subject to the Poisson equation

$$\frac{\partial E}{\partial x} = 4\pi e(n_i - n_e). \tag{1.4}$$

Equations (1.1)–(1.4) describe the one-dimensional motion of plasma. As we see equations (1.2)–(1.3) are nonlinear, so that nonlinear effects are definitely present if the wave amplitude is not small. But before proceeding to nonlinear phenomena, it is reasonable to start from less complicated linear waves.

1.2.2. Linear ion-sound waves

If the amplitude of perturbation is very small, all the set of equations can be linearized by dropping higher order terms. We

assume that the plasma is initially at rest with $v = 0$ and its density is constant: $n_e = n_i = n_0 = \text{const}$. When a small perturbation is present, we have

$$n_i = n_0 + n_i', \quad n_e = n_0 + n_e', \tag{1.5}$$

where the prime indicates small perturbation values.

Dropping the second order terms we obtain linear equations. It is known that any solution of the set of linear equations can be considered as a superposition of the harmonics $\exp(-i\omega t + ikx)$, where ω is the angular frequency and k is the wave number. For such a wave equation (1.3) leads to the relation $ikT_e n_e' = -eEn_0$ and with the help of (1.4) we obtain the following relation between n_e' and n_i':

$$(1 + k^2 d^2) n_e' = n_i'. \tag{1.6}$$

Here $d^2 = T_e/4\pi e^2 n_0$, the d-value is called the Debye length. As we see, the difference between n_e' and n_i' diminishes when $k^2 d^2 \to 0$. Thus, we can claim that plasmas become quasineutral when the wave length is much larger than the Debye length.

With the help of (1.6) and linearized equations (1.2) we can write equation (1.1) in the following form:

$$\frac{\partial v}{\partial t} + \frac{c_s^2}{v_p(1 + k^2 d^2)} \frac{\partial v}{\partial x} = 0. \tag{1.7}$$

We have introduced new notation here: $c_s^2 = T_e/m_i$ is the square of the sound velocity, and $v_p = \omega/k$ is the phase velocity.

In the linear approximation $dv/dt = \partial v/\partial t = -i\omega v$, so that for harmonics of type $\exp(-i\omega t + ikx)$ we obtain from (1.7) the relation between the frequency and the wave number:

$$\omega^2 = \frac{c_s^2 k^2}{(1 + k^2 d^2)}. \tag{1.8}$$

This relation is called the dispersion relation. For very small wave numbers when $k^2 d^2 \to 0$ this relation has the approximate form $\omega^2 = c_s^2 k^2$ similar to the corresponding relation for normal sound propagation in air. The phase velocity

$v_p = \omega/k = c_s$ is the same for all harmonics in this case. In other words there is no dispersion: all harmonics propagate keeping the same phase relations which were imposed on them during the sound generation. We are not surprised that speech can be heard at any distance from the speaker without distortion when the sound is sufficiently loud. But this property off sound would be lost if the phase velocity were dependent upon the wave number.

When k^2d^2 increases, dispersion appears. For $k^2d^2 \ll 1$ it is weak, but if $k^2d^2 \geq 1$ the dispersion is strong. The dependence of the frequency on the wave number for ion sound waves is shown in Fig. 1.1. As we see as $k^2d^2 \to 0$ the frequency is proportional to the wave number, but for large k values saturation takes place. At very large k values the frequency is equal to $\omega = \sqrt{4\pi e^2 n_0/m_i} \equiv \omega_{pi}$. This value is called the ion Langmuir frequency.

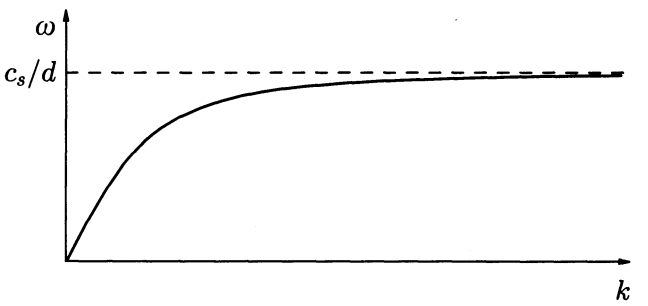

Fig. 1.1. Dependence of the ion wave frequency on the wave number.

The characteristic wave number value k_0 at which the transition to saturation becomes appreciable is equal to $k_0 = d^{-1}$. As we see from Fig. 1.1, the phase velocity $v_p = \omega/k$ is close to c_s as $k^2/k_0^2 \to 0$, but for larger k values the phase velocity decreases when $k \to \infty$.

1.3. Linear waves in dispersive media

1.3.1. Dispersion of waves on a water surface

When we consider harmonic plane wave propagation, wave dispersion looks like a very weak and harmless effect: it appears simply as a phase velocity dependence upon the wave number. But in real conditions monochromatic harmonic waves are excited not so frequently. Usually wave packets are excited or emitted, representing a superposition of plane waves. They look like some localized patterns and in these cases wave dispersion leads to very specific and very interesting effects.

These effects could be considered directly on the basis of plasma waves. However for each of us waves on water surfaces are much more familiar. That is why we start the discussion of wave dispersion effects with gravitational waves on the water.

The most simple case corresponds to gravitational waves on deep water. The compressionless fluid dynamics are described by two equations, namely by the Euler equation with the gravitational force taken into account,

$$\rho \frac{d\boldsymbol{v}}{dt} + \nabla p = \rho \mathbf{g}, \tag{1.9}$$

and the continuity equation

$$\nabla \cdot \boldsymbol{v} = 0, \tag{1.10}$$

where $\mathbf{g}$ is the gravitational acceleration directed downward and ρ is the fluid mass density. At first let us find the dispersion relation for the gravitational waves. We consider a linear wave with a very small amplitude propagating along the horizontal x-axis and having only one harmonic of the type $\exp(-i\omega t + ikx)$. Let the vertical coordinate be z.

In the steady state the water surface is horizontal and is situated at $z = 0$. The equilibrium hydrostatic pressure is equal to $p_0(z) = p_0 - \rho g z$, where p_0 is the atmospheric pressure on the water surface. But if the water surface is perturbed, the gravitational force will try to restore the equilibrium and in combination

with the inertial motion, gravitational waves will be generated on the surface. In addition to the velocity $\boldsymbol{v}$ of fluid we can use the displacement $\boldsymbol{\xi}$ for the small amplitude perturbations. These values are coupled by the evident relation: $\boldsymbol{v} = \partial\boldsymbol{\xi}/\partial t$. Now we can linearize equation (1.9). In the zero approximation the gravitational force is compensated by the hydrostatic pressure gradient. Thus, we obtain the following relation for linear perturbations:

$$\frac{\partial^2 \boldsymbol{\xi}}{\partial t^2} = -\frac{1}{\rho}\nabla p', \tag{1.11}$$

where p' is the pressure perturbation. We can now apply the $\nabla\cdot$ operator to (1.11) taking into account that $\nabla \cdot \boldsymbol{\xi} = 0$. Thus, we obtain

$$\Delta p' = 0. \tag{1.12}$$

For a plane wave of type $\exp(-i\omega t + ikx)$ this equation looks like

$$\frac{d^2 p'}{d^2 z} - k^2 p' = 0. \tag{1.13}$$

If the water is deep we have to take only one of the two solutions of this equation, namely $p' = p'(0)\exp(kz)$, because this function tends to zero when $z \to -\infty$.

Now we have to take into account the fact that the pressure perturbation near the water surface is produced by the thin displaced water layer: when the surface is shifted vertically by ξ_z, the fluid pressure under this surface (at fixed z value) is increased by the value $\rho g \xi_z$. Thus, we obtain a linear boundary condition $p'(0) = \rho g \xi_z(0)$, where the zero argument value corresponds to $z = 0$. Now the z-component of equation (1.11) at $z = 0$ gives

$$\omega^2 \xi_z(0) = \frac{k}{\rho} p'(0) = k g \xi_z(0). \tag{1.14}$$

Therefore we obtain the following dispersion relation:

$$\omega^2 = gk, \tag{1.15}$$

which corresponds to gravitational waves on a deep water surface.

Relation (1.15) is valid for sufficiently long waves when the surface tension is not important. To take the surface tension into account, we have to add an additional force under the curved surface: $\delta p' = \sigma/R = -\sigma\,\partial^2\xi_z/\partial z^2$. Here σ is the specific surface tension and R is the radius of curvature. Thus, we obtain for the harmonic plane wave: $p'(0) = \rho g\xi_z(0) + \sigma k^2\xi_z(0)$. With this new relation we obtain from (1.14):

$$\omega^2 = kg + \frac{\sigma}{\rho}k^3. \tag{1.16}$$

We see that when the wave number increases the gravitational wave is smoothly transformed into a so called capillary wave with frequency $\omega = (\sigma/\rho)^{1/2}k^{3/2}$. This transition takes place for water at $k \sim 10\ \mathrm{cm}^{-1}$, i.e., at a wavelength of about 1 cm.

Now we can obtain the relation for the phase velocity $v_p = \omega/k$. Namely with the help of (1.16) we obtain: $v_p^2 = g/k + \sigma k/\rho$. As we see, water surface waves have a strong dispersion: the phase velocity decreases when k increases for small k-values and increases with k at k values larger than $k_0 = (g\rho/\sigma)^{1/2} \sim 10\ \mathrm{cm}^{-1}$. Only in the narrow region near $k = k_0$ is the dispersion weak and the phase velocity weakly dependent upon k. Note that the minimum value of the phase velocity corresponds to $k = k_0$ and is equal to $(v_p)_{\min} = (g\sigma/\rho)^{1/2}$.

We now consider the case when the water basin is not very deep. Let h be the depth of the water layer. The normal component of displacement, $\xi_z(-h)$, and consequently, $\partial p'/\partial z$, at the bottom must be zero. Hence we have from (1.13) $p' = p'(0)\cosh\big[k(h-z)\big]/\cosh(kh)$, where $\cosh(x) = (e^x + e^{-x})/2$. With the help of this relation and (1.11) we obtain the following dispersion relation for a water layer of finite depth:

$$\omega^2 = \left(kg + \frac{\sigma}{\rho}k^3\right)\tanh(kh). \tag{1.17}$$

An interesting case arises when the water layer is shallow, i.e., $kh \ll 1$. Equation (1.17) can be expanded in this case in a series in powers of kh and so approximately

$$\omega^2 = k^2 h \left(g + \frac{\sigma}{\rho} k^2 \right) \left(1 - \frac{k^2 h^2}{3} \right). \tag{1.18}$$

When k is very small $\omega^2 = k^2 hg$, i.e., there is no dispersion in this case. If we take the next order terms we obtain approximately:

$$\omega^2 \simeq k^2 hg \left(1 - \frac{k^2 h^2}{3} + \frac{\sigma}{\rho g} k^2 \right). \tag{1.19}$$

As we see, weak dispersion appears. Thus, in the case of a shallow water layer the waves have only weak dispersion.

1.3.2. Waves from a local pulsed source

All of us know how beautiful is the picture of rings on pond produced by a dropped stone. After a short period of nonlinear surface rippling the perturbed area expands, the amplitude of perturbation diminishes, and soon a beautiful pattern of many expanding circles develops. This is the most clear and familiar example of wave dispersion.

From the purely mathematical point of view the problem of description of a linear wave is very simple. Any wave pattern can be represented as a superposition of plane waves with a known dependence of the frequency upon the wave number. In particular, the localized source can be considered as a two-dimensional δ-function whose Fourier components are well-known. Thus, the temporal evolution of the Fourier components as well as their superposition are at hand.

We are interested here not in the exact solution, however, but in the qualitative picture of the wave propagation. That is why we shall use a simpler approach which can be applied to many other cases.

Namely we shall argue in the following way. Let the picture of the circular waves look like the one shown in Fig. 1.2

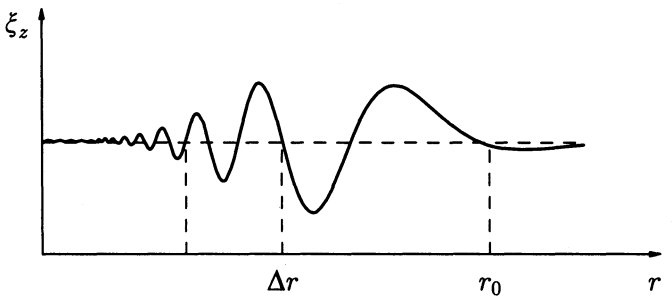

Fig. 1.2. Radial dependence of the surface vertical displacement. The wave pattern is seen inside a circle with radius r_0. The ring with thickness Δr looks like a circular wave packet.

where the radial dependence of the vertical surface displacement ξ_z is drawn as a function of distance r from the point source.

Now let us divide all the circle into rings in such a way that each ring contains at least several humps. Each of these rings can be considered as a wave packet. It is known that the wave packet propagates with the group velocity $v_g = \partial\omega/\partial k$. For gravitational waves with dispersion $\omega = \sqrt{gk}$ the group velocity is equal to $v_g = \frac{1}{2}\sqrt{g/k}$. As we see the group velocity is half the phase velocity $v_p = \omega/k$. Propagating with the group velocity the wave packet with wave number k covers a distance $r = v_g t$ for a time interval t counted from the moment of perturbation generation. Thus, at the moment t this packet reaches the point $r = (t/2)\sqrt{g/k}$ or, alternatively, the wave number at this point is equal to

$$k = gt^2/4r^2. \tag{1.20}$$

Being considered at fixed r value this relation shows that the wave number at a given point increases with time. This means that the wave packets are permanently renewed being replaced by packets with higher k values. These packets run with group velocities and at each r point $v_g = r/t$. If we fix the value t then relation (1.20) describes the dependence of k upon r. As we see the wave number is very large near the center decreas-

ing rapidly with increasing r. When kr becomes of the order of unity, the wave pattern disappears, and at larger values of r we see no waves. In other words, the relation $rk \sim 1$ defines the maximum radius r_0 which represents the boundary of the wave pattern. Using (1.20) we estimate $r_0 \sim gt^2/4$. This is the visible boundary of the wave structure, i.e., the front of perturbations. As we see this front expands with constant acceleration of the order of g — the gravitational acceleration.

Note that the phase velocity is twice the group velocity. Hence, the crests run twice as fast as the groups of waves and after reaching the boundary of the circle at $r = r_0$ they disappear.

In addition to the wave number behavior it is reasonable to consider how the phase of the wave pattern behaves. The phase φ is defined by the relations: $\omega = \partial\varphi/\partial t$, $k = -\partial\varphi/\partial x$. In other words the temporal rate of the phase variation is equal to the angular frequency, and its gradient corresponds to the wave number. Using (1.20) and the relation $\omega = \sqrt{gk}$ it is easy to get the relation

$$\varphi = \frac{gt^2}{4r}. \tag{1.21}$$

In our case when ω and k are functions of the r/t ratio only the phase φ is equal simply to $\varphi = \omega t - kr$. We see again that at a large distance from the center there are no waves: the change of phase there is less than π and no crest can be observed.

1.3.3. Cherenkov emission

Perhaps many readers have had the opportunity to observe the beautiful waves produced by a moving boat or ship. These are shown in Fig. 1.3a. This picture looks very complicated. But in fact this is no more than the simple Cherenkov radiation produced by a body moving with a velocity larger than the phase velocity. All the complication is produced by the dispersion of the gravitational waves.

It is well known that Cherenkov emission arises as a superposition of elementary waves produced by a moving body.

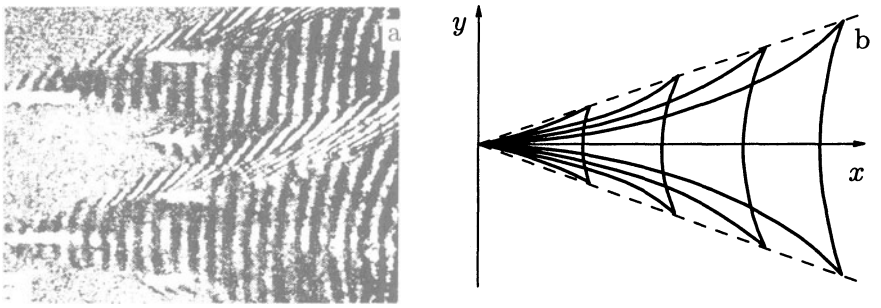

Fig. 1.3. (a) Gravitational waves on deep water produced by a moving ship. (b) Family of curves with constant phase.

Figure 1.4a shows the radiation produced by a body moving with supersonic velocity in air. Elementary waves look like circles here. These elementary perturbations have the same phase on the Mach line representing the envelope of all the circles. As is seen in Fig. 1.4a, the angle of emission θ is defined by the relation

$$\cos\theta = v_p/v_0 \tag{1.22}$$

where v_p is the phase velocity and v_0 is the velocity of the moving body. Note that $\cos\theta = k_z/k$ where k_z is the z-component of the wave vector $\mathbf{k}$. Hence, relation (1.22) can be written in the form $k_z v_0 = k v_p = \omega$. This is the well known Cherenkov emission relation $\omega = k_z v_0$ which shows that emission takes place when the velocity of the moving body coincides with the velocity of the phase along the z-axis: $v_0 = \omega/k_z$. In other words in the moving body frame of reference the wave is seen as a steady state pattern.

Let us now consider how Cherenkov radiation is produced in a dispersive medium. As a simple example of such a medium we consider the emission of gravitational waves produced by a body moving on a deep water surface.

We can say again that Cherenkov emission is produced as a series of small local perturbations. But now each of these per-

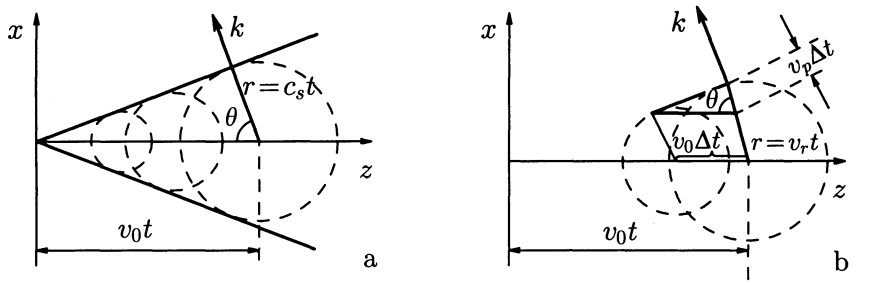

Fig. 1.4. Wave superposition produced by a moving body in the cases of nondispersive (a) and dispersive (b) media.

turbations generates a lot of circular waves as shown in Fig. 1.4. These circular waves are superimposed and they give rise to new lines of constant phase. Namely these lines of constant phase are observed by an external observer as a family of crests. Hence, we have to find the lines of constant phase in order to describe this wave pattern.

How to do this is shown in Fig. 1.4b. We choose two circular lines which correspond to elementary waves emitted by the moving body at times t and $t + \Delta t$. We assume that these lines correspond to the same phase, so that the difference in their radii is equal to $v_p \Delta t$. As for their centers, they are separated by the distance $\Delta z = v_0 \Delta t$ as shown in Fig. 1.4b. Thus, the angle θ is again subject to the relation $\cos\theta = v_p/v_0$ as in the dispersionless case, and again we have $v_0 k_z = v_p k = \omega$. Therefore the local relations are the same.

But now we have to take into account that the wave packets propagate with the group velocity. Hence, we have the relation $r = v_g t$ between the distance from the source and the point of observation, as shown in Fig. 1.4b. This distance is not equal to the $v_p t$ product as it was in the case without dispersion. In particular for the case of gravitational waves $v_g = v_p/2$, so that we have

$$r = v_g t = \frac{1}{2} v_p t = \frac{1}{2} v_0 t \cos\theta \tag{1.23}$$

according to the Cherenkov emission condition (1.22). As is seen from Fig. 1.4b the coordinates of the observation point in the frame of reference, where the moving body is at the center, can be expressed as $x = r\sin\theta$, $z = v_0 t - r\cos\theta$. We can now try to find the coordinates of the points with the same phase φ. With the help of the relations (1.21) and (1.23) we can express both r and t in terms of φ and get the following expression for the coordinates of the observation point:

$$x = r\sin\theta = \frac{v_0^2\varphi}{g}\sin\theta\cos^2\theta, \tag{1.24}$$

$$z = v_0 t - r\sin\theta = \frac{v_0^2\varphi}{g}(2-\cos^2\theta)\cos\theta. \tag{1.25}$$

Here φ is the phase and θ is the angle between the wave vector and the z-axis. If we fix the phase φ, then the formulae (1.24)–(1.25) will define the line of constant phase expressed in the parametric form. Changing φ by the interval 2π we obtain the family of lines which describe the net of crests. As we see all the lines are similar to each other and correspond quite well to the picture observed in Fig. 1.3a.

Qualitatively this net looks like a superposition of two types of waves. The waves with short wavelengths propagate sideways from the wave central line towards the periphery. The phase velocity of these waves is small compared with v_0 so that the angle θ for them is not very much different from $\pi/2$. The second type of wave looks like distinct fragments of the same progressive wave which propagates with a phase velocity v_p equal to the moving body velocity v_0: $v_p = v_0$. Hence, this wave has $k = g/v_0^2$ and propagates together with the body being in resonance with it. As is seen in Fig. 1.4 all the wave pattern is situated inside the angle 2α which can be found from the condition $\partial x/\partial\theta = 0$ and the relation $\tan\alpha = x/z$ for corresponding values of x and z. It turns out that $\alpha \simeq 18°$.

Thus, independent of body size its motion produces a universal net of waves on the surface of the deep water. The structure of this wave family is defined by the dispersion relation only.

No less interesting is the case of capillary waves. Their dispersion relation is $\omega^2 = \sigma/\rho k^3$, so that the group velocity $v_g = \frac{3}{2}\sqrt{\sigma/\rho k}$ is larger than the phase velocity $v_p = \sqrt{\sigma k/\rho}$. It is not difficult to describe the wave propagation from the local pulsed perturbation. Namely, we have again $r = v_g t$ so that

$$k = \frac{4\rho r^2}{9\sigma t^2}. \tag{1.26}$$

As we see at a given t the wave number increases with r and the wavelength decreases with r. For $kr \leq 1$ the wave structure disappears so that we see no waves inside a circle with radius $r_0 \sim (\sigma/\rho)^{1/3} t^{2/3}$. With the help of (1.26) we can find ω and the phase φ. It is more convenient to define φ as $\varphi = -\omega t + kr$ to deal with positive phase values. Thus, we obtain:

$$\varphi = \frac{4}{27}\frac{\rho}{\sigma}\frac{r^3}{t^2}. \tag{1.27}$$

We see again that the real wave structure appears at $r > r_0$ where the phase variation can exceed unity.

Let us consider the Cherenkov emission of capillary waves. The condition that the wave structure propagates with the moving body velocity v_0 is given again by the relation $\cos\theta = v_p/v_0$. The point where this piece of structure is situated is seen in Fig. 1.4b with $r = v_g t = \frac{3}{2} v_p t$. Taking the resonance condition into account we transform this relation into a relation between r and t: $r = \frac{3}{2} v_0 t \cos\theta$. The phase φ at this point can be found with the help of relation (1.27): $\varphi = \frac{1}{3}(\rho/\sigma)\, r \cos\theta$. We can now again express the coordinates $x = r\sin\theta$, $z = v_0 t - r\cos\theta$ through φ and θ:

$$x = 3\frac{\sigma\varphi}{\rho v_0^2}\frac{\sin\theta}{\cos^2\theta}, \quad z = \frac{\sigma\varphi}{\rho v_0^2}\frac{2\sin^2\theta - \cos^2\theta}{\cos^3\theta}. \tag{1.28}$$

The lines of constant phase are shown in Fig. 1.5. As we see the wave pattern exists in front of the moving body. Here the wave number corresponds to the resonance condition $v_p = v_0$. Then

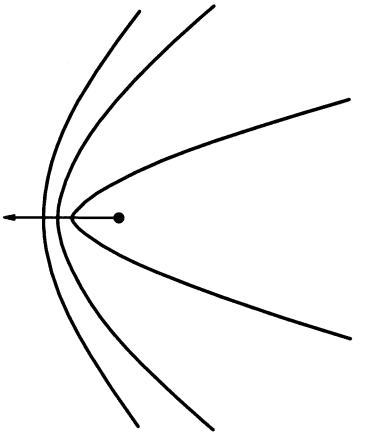

Fig. 1.5. Capillary wave structure produced by a moving body.

the lines of constant phase bend and at large values of z the asymptotic behavior $x \sim \varphi^{2/3} z^{2/3}$ is reached which corresponds to $\cos\theta \to 0$.

1.3.4. Ion-sound waves

The dispersion of sound waves looks like $\omega = c_s k(1 + k^2 d^2)^{-1/2}$ so that the phase velocity $v_p = c_s(1 + k^2 d^2)^{-1/2}$. The group velocity is equal to $v_g = \partial\omega/\partial k = c_s(1+k^2 d^2)^{-3/2}$. We see that the short wave perturbations propagate more slowly than the long wave ones. For the case of a local pulsed source we have $v_g = r/t$ and from the expression for v_g we find

$$k = k_0\sqrt{(c_s t/r)^{2/3} - 1}, \tag{1.29}$$

where $k_0 = 1/d$. As we see this expression is defined in the subsonic region $r < c_s t$ and this is quite natural. With the help of (1.29) the expression for the phase φ can easily be found:

$$\varphi = k_0 r\left[\left(c_s t/r\right)^{2/3} - 1\right]^{3/2}. \tag{1.30}$$

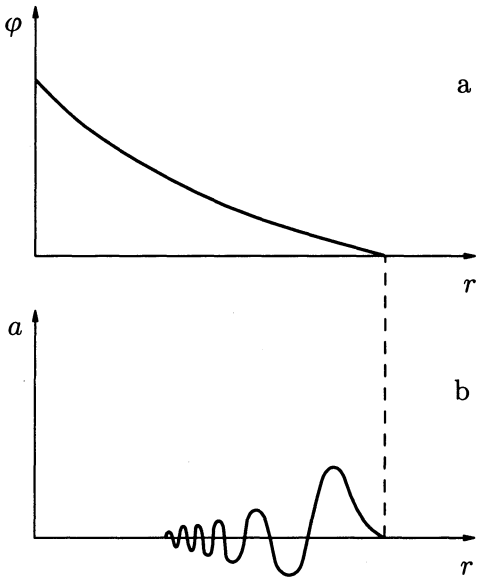

Fig. 1.6. Phase (a) and amplitude (b) dependence upon the distance from a localized pulsed ion sound source.

The dependence upon r is shown qualitatively in Fig. 1.6. The same figure shows the amplitude dependence upon radius. If the source is not a strongly local one, then the short wave harmonics are absent and the wave packet of Fig. 1.6b will be localized near its edge $r = c_s t$. In this case the dispersion of waves is dominated by small kd values so that we can use the approximation

$$\omega = c_s k \left(1 - \frac{k^2}{2k_0^2}\right). \tag{1.31}$$

Thus, in this approximation the ion sound wave dispersion can be considered weak.

In the case of weak dispersion the problem of Cherenkov emission is very much simplified. Namely the dispersion leads to slight broadening of the wave packet near the Mach line. The generation of the slower wave component looks exactly the same as its temporal evolution shown in Fig. 1.6b.

In the general case, however, when all sound wave harmonics are present, the picture of Cherenkov emission is not so simple. Of course it can be treated with the help of the approach discussed earlier.

Namely we can put $v_g = r/t$ and taking into account the relation $v_g = c_s(1+k^2d^2)^{-3/2}$, we obtain the following relations: $1 + k^2d^2 = (c_st/r)^{2/3}$, $\omega = c_sk(r/c_st)^{1/3}$. It is easy to check that these expressions for the wave number and frequency can be found by taking the corresponding derivatives of the phase $\varphi = k_0r\left[(c_st/r)^{2/3} - 1\right]^{3/2}$ where $k_0 = 1/d$. With the help of the Cherenkov radiation condition $\omega/k = v_0\cos\theta$ the variables t, r can be expressed through φ and we obtain the following expressions for coordinates of the point of observation:

$$x = r\sin\theta = \varphi d\frac{v_0^3}{c_s^3}\sin\theta\cos^3\theta\left(1 - \frac{v_0^2}{c_s^2}\cos^2\theta\right)^{-3/2}, \quad (1.32)$$

$$z = v_0t - r\cos\theta = \varphi d\frac{v_0}{c_s}\left(1 - \frac{v_0^2}{c_s^2}\cos^4\theta\right)\left(1 - \frac{v_0^2}{c_s^2}\cos^2\theta\right)^{-3/2}. \quad (1.33)$$

These general expressions are valid for both subsonic and supersonic cases. It is interesting to consider the subsonic case when only the short waves can be emitted. Figure 1.7 shows a picture of Cherenkov radiation for the case $v_0^2/c_s^2 = 1/2$. As is seen, the number of crests is not so large, so that this treatment being based on the wave packet approach is not quite accurate and has to be considered as qualitative.

1.3.5. Weak dispersion

When the wave numbers are small, the ion sound dispersion can be represented in the very simple form $\omega = c_sk(1 - k^2/2k_0^2)$. It turns out that this type of dispersion is more or less universal when the waves have no dispersion in the limit $k \to 0$. Indeed, dispersion in this case can appear when the wave number becomes not very small in comparison with some internal

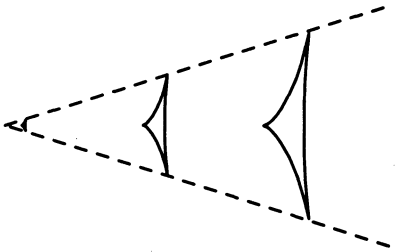

Fig. 1.7. Ion sound wave pattern produced by a subsonic moving body. Crests are shown by solid lines and the boundary of the perturbed area by dashed lines.

reciprocal length. If we denote this value as k_0 and expand the phase velocity value as a function of k in a series of k powers, we obtain in the first nonzero approximation

$$v_p = c_0\left(1 \pm k^2/2k_0^2\right). \tag{1.34}$$

Only the second order term appears in this expansion because the phase velocity is an even function of k. The value c_0 in (1.34) corresponds to the phase velocity of the long wave perturbations in the limit $k \to 0$.

The dependence of the phase velocity upon the wave number (1.34) can have either a plus or minus sign in the brackets. Correspondingly, the group velocity $v_g = c_0(1 \pm 3k^2/2k_0^2)$ either increases or decreases when the wave number increases. In the first case short wave perturbations overtake the long wave ones and in the second case they lag. That is why we shall call the first case a positive weak dispersion and the case with the minus sign as negative weak dispersion. Ion sound has a negative weak dispersion at small values of the wave number. As for waves on the water surface, their dispersion according to (1.19) can have both signs of dispersion: a negative one for a deep water layer and positive one for shallow layers, $h \leq 1$ cm.

When a magnetic field is present in plasma, new branches of plasma waves appear as will be shown in Section 5. They can

have both positive and negative dispersion including the weak dispersion in the limit of very small wave number values.

Thus, weak dispersion is more or less a universal property of waves which have no dispersion in the limit $k \to 0$. The corresponding picture of wave evolution looks like the universal one. The waves propagate mainly with velocity c_0 and only the short wave component either overrun in the case of positive dispersion or lags the main wave packet in the case of negative dispersion.

The picture of Cherenkov emission is again a universal one. It looks like the temporal wave packet evolution near the Mach line. In other words there is no difference between the temporal and spatial evolutions of the wave packets. We will see later that even weak dispersion is a very important property of waves with significant amplitudes. This means that nonlinear wave evolution is drastically changed when dispersion is involved. We shall proceed to understanding this phenomenon step by step starting from the simple case of the waves without dispersion.

2. Nonlinear waves

2.1. Ordinary progressive waves

2.1.1. Beam of non-interacting particles

Nonlinear waves are much more complicated than linear ones. There is no superposition principle here. Roughly speaking it means that there are many more nonlinear phenomena than linear ones. Correspondingly, the physics of nonlinear phenomena is much richer and more complicated.

To step into the nonlinear phenomena world it is reasonable to start with simple nonlinear effects which look like a modest extension of linear ones. We will see that nondispersive and dispersive waves are drastically different. That is why we start our consideration with the nondispersive case and then we shall consider the influence of dispersion.

Before analyzing nonlinear wave evolution we consider a very simple model, namely, a one-dimensional beam of noninteracting particles. The velocity v of each particle is constant so that from the point of view of a single particle this is an extremely simple situation.

But now we consider this beam as continuous matter so that v is the function of time and x-coordinate. The same statement $v = \text{const}$ can be written now as

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} = 0. \tag{2.1}$$

The reason for equation (2.1) is as follows. The continuous function v is equal to $v(t, x)$. When we intend to find the acceleration of a single particle we have to take the derivative

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + \dot{x}\,\frac{\partial v}{\partial x}$$

as for a composed function. But $\dot{x} = v$ and we obtain (2.1).

A system of noninteracting particles is by no means a nonlinear system. But equation (2.1) appears to be nonlinear and has many features of nonlinear equations. In particular, it has a solution which is of a nonlinear nature.

Let us start with very small linear perturbations. Namely we put $v = v_0 + v'$ where v' is extremely small. By linearization of (2.1) we obtain

$$\frac{\partial v'}{\partial t} + v_0\,\frac{\partial v'}{\partial x} = 0,$$

and for harmonic perturbations of the type $\exp(-i\omega t + kx)$ we find the dispersion relation

$$\omega = kv_0. \tag{2.2}$$

As we see the phase velocity $v_p = \omega/k = v_0$ is constant. This means that the beam of particles looks like a nondispersive medium in the linear approximation. Let us assume that this beam has an initial periodic perturbation $a\sin(kx)$ where the amplitude a is not small. It is better to consider the evolution of this

perturbation in a frame of reference moving with velocity v_0. We put $x = x' + v_0 t$, $v = u + v_0$ and drop the prime converting (2.1) into the following relation:

$$\frac{du}{dt} = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0. \tag{2.3}$$

For a description of the beam evolution it is convenient to use an artificial phase space with coordinates u, x. Each point of this phase plane moves horizontally with velocity u. Correspondingly, the upper semiplane moves to the right and the lower one to the left. The velocity of each point is proportional to its distance from the x-axis so that the plane is permanently sheared.

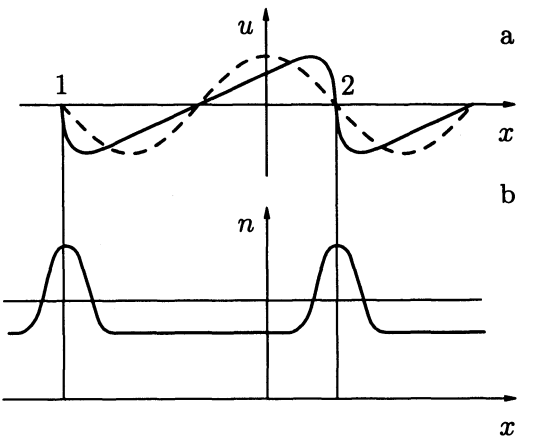

Fig. 2.1. Wave turnover in the case of a beam of noninteracting particles. Evolution of the beam velocity (a) and density (b).

Figure 2.1a shows the evolution of the beam velocity in time. The initial perturbation looks like a dashed line. Later on this line deforms together with the phase plane deformation. Each horizontal layer of this plane moves with the velocity u and this shear flow produces the beam deformation displayed by the continuous curve in Fig. 2.1a.

As we see the wave profile becomes nonharmonical. Points 1 and 2 with zero velocity are at rest but the layers with $u \neq 0$ are moving and distort the initial sinusoidal curve. As we see the wave becomes steeper at the points 1, 2. At some moment steepening reaches such a level that the derivative $\partial u/\partial x$ becomes infinite at the points 1, 2. This is the moment when the wave overturns. Later on this overturning leads to multi beam pattern formation as shown in Fig. 2.2a where we see regions with three beams simultaneously. Here the function $u(x)$ is not single-valued.

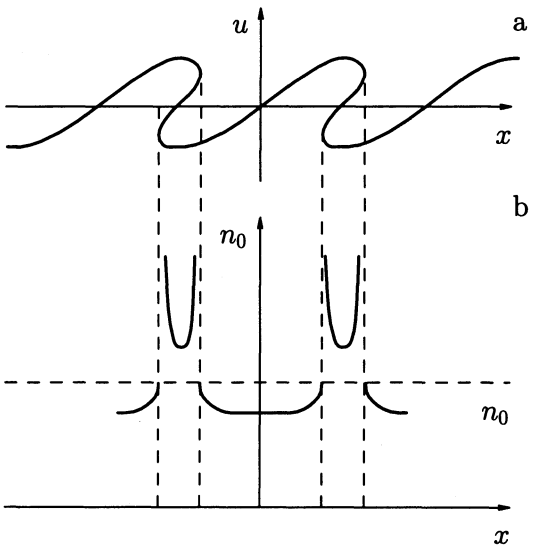

Fig. 2.2. Multibeam pattern as a result of modulated beam evolution. The beam velocity (a) and density (b) are shown.

We see that the curve $u(x)$ has many turn points. These turn points continue moving — the upper points in the positive direction and the lower points in the negative direction. As a result the number of beams increases: new beams are generated, appearing in pairs. After long time of evolution the number of beams becomes so large that we can use the distribution function $f(u)$ for an approximate description.

It is interesting to analyze what happens to the beam density $n(t,x)$. It is easy to see that the harmonic wave distortion is accompanied by the density perturbation. As Fig. 2.1b shows, dense clouds of particles appear at the points 1, 2 where overturning builds up. Bunches of particles arise here. Such bunches of particles are used in klystrons for the generation of electromagnetic waves.

The increase of density at the points 1, 2 continues up to the moment when $\partial u/\partial x$ becomes infinite. Simultaneously the density $n(x)$ at these points becomes infinite. After overturning the number of singularities doubles, as Fig. 2.2b shows. When the number of beams increases the density "spikes" become narrower and finally the density reaches an almost homogeneous distribution along the x-axis. In this limit the homogeneous distribution $f(u)$ can be approximately imagined.

Thus, the consideration of this very simple system reveals typical nonlinear effects. Namely higher order harmonics are generated up to the point of wave overturning accompanied by the corresponding density perturbations. As we shall see these effects are relevant to waves without dispersion.

The nonlinearity is clearly seen in (2.3) equation. In principle it could be treated by the method of subsequent perturbations. For this purpose equation (2.3) should be rewritten in the more convenient form

$$\frac{\partial u}{\partial t} = -u\frac{\partial u}{\partial x}, \tag{2.4}$$

where the second order term is placed on the right-hand side. In the linear approximation this term is put to zero so that a linear wave has no temporal evolution. If we take the linear wave in the form $u^{(1)} = a\sin kx$ then by the second order of perturbations equation (2.4) takes the form

$$\frac{\partial u^{(2)}}{\partial t} = -k^2a^2\sin(kx)\cos(kx) = -\frac{ka^2}{2}\sin(2kx), \tag{2.5}$$

where $u^{(2)}$ is the second order correction. According to (2.5) the second order term increases linearly with time and at $t \sim$

$1/ka$ the amplitude of the second harmonic becomes of the same order of magnitude as $u^{(1)}$. Later on the method of subsequent approximations ceases to be valid.

Using relations (2.4)–(2.5) we can easily understand why the absence of dispersion is so important for the overturning effect. As we see in (2.5) the secondary harmonic can grow up to a large amplitude for the only reason that it is always in resonance with the first harmonic. If dispersion were present, the secondary harmonic would have a different phase velocity so that $\delta v_p = v_p(2k) - v_p(k)$ would not be zero. In this case equation (2.5) would be modified as follows:

$$\frac{\partial u^{(2)}}{\partial t} + \delta v_p \frac{\partial u^{(2)}}{\partial x} = -\frac{ka^2}{2} \sin(2kx). \tag{2.6}$$

Its solution with the initial condition $u^{(2)}(t = 0) = 0$ looks like this:

$$u^{(2)} = \frac{a^2}{4\delta v_p} \left[\cos 2kx - \cos 2x(x - \delta v_p t) \right]. \tag{2.7}$$

This expression is always finite and for $a \ll \delta v_p$ it continues to be small compared with the first harmonic.

From here we see quite clearly why nondispersive and dispersive media differ so much. In nondispersive media all harmonics are always in resonance with each other all the time. Thus, the transfer of energy from one harmonic to another is not limited and continues up to the overturn of the wave. In dispersive media such an energy transfer is absent and higher order harmonics can be maintained in the equilibrium together with the first harmonics even in the case when nonlinearities are present. We will take this effect into account in the analysis of nonlinear wave packet evolution.

With the help of the beam model we can discuss one additional effect, namely, reversibility. This effect in a more complicated form persists in some nonlinear kinetic phenomena, but here it can be seen in a very simple realization.

As we saw, modulated beam evolution looks like a unidirectional process: the number of secondary beams increased

continuously. It seems that eventually we have a practically smooth distribution of particles by velocity so that the distribution function $f(u)$ approach looks like the most appropriate one. However, this approach is not exact. Moreover, it is not the right one.

Indeed, the beam of particles does not contain any dissipative mechanisms and continues to be a completely reversible system. If we were able to reverse a particle velocity at some moment of particle evolution we could observe the backward evolution of a beam in exactly the reversed sense. Namely on the initial background of the smooth particle distribution $f(u)$ we would observe the generation of tiny bunches. These bunches would grow and finally we would come to the initial homogeneous distribution of density with the velocity modulated as $a \sin kx$. In other words we would come to the initial state but with reversed velocities. Later on this modulated beam would experience the same evolution stages as discussed earlier.

We can say that an almost smooth multibeam distribution of particles has, in fact, not a smooth but a very fine discontinuous function $f(u)$. This state keeps all the information which was contained in its initial state. Beam velocity reversal can reveal the information and makes it explicit. Of course to reverse the beam particle velocities directly is not possible. But we will see later that there is another possibility for extracting the hidden information. This possibility relates to a very interesting effect — the echo, which will be considered in this book somewhat later.

2.1.2. Ordinary progressive waves

Now we consider a more realistic case, namely nonlinear ion sound waves with very long wavelengths. We neglect the dispersion assuming $kd \to 0$ and use the quasinuetrality condition. Correspondingly, we have the following set of equations:

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} + c_s^2 \frac{1}{n}\frac{\partial n}{\partial x} = 0, \tag{2.8}$$

$$\frac{\partial n}{\partial t} + \frac{\partial}{\partial x} nv = 0. \tag{2.9}$$

Here $c_s^2 = T_e/m_i = \text{const}$.

The set of equations (2.8)–(2.9) is strongly nonlinear and to solve these equations is not a simple matter. However, it turns out that there is a simplified solution which is very interesting for nonlinearity manifestation. This is the progressive wave. This wave looks like a modest generalization of a linear wave.

So let us start again with linear waves. We put $n = n_0 + n'$, with n' and v small, and then drop all the second order terms. We obtain:

$$\frac{\partial v}{\partial t} + \frac{c_s^2}{n_0} \frac{\partial n'}{\partial x} = 0, \tag{2.10}$$

$$\frac{\partial n'}{\partial t} + n_0 \frac{\partial v}{\partial x} = 0. \tag{2.11}$$

We take the time derivative of equation (2.10) and then exclude $\partial n'/\partial t$ with the help of (2.11). We obtain:

$$\frac{\partial^2 v}{\partial t^2} - c_s^2 \frac{\partial^2 v}{\partial x^2} = 0. \tag{2.12}$$

This equation has a solution $v = v_+(x - c_s t) + v_-(x + c_s t)$, where v_+ and v_- are arbitrary functions of their arguments. The solution v_+ corresponds to a wave propagating along the x-axis in the positive direction, while v_- corresponds to a wave propagating in the negative direction. If the amplitude of both these waves is increased the nonlinear picture will be quite complicated. But let us take only one of these waves, for instance $v_+(x - c_s t)$. Then for this wave $\partial/\partial t = -c_s \partial/\partial x$ so that equation (2.11) can be once integrated and we obtain $n'/n_0 = c_s^{-1} v$. In other words the density is a linear function of v in this case: $n = n_0(1 + v/c_s)$.

Now we are ready to generalize slightly the set of equations (2.8)–(2.9) to the nonlinear case assuming that n can be considered a function of velocity: $n = n(v)$.

With this assumption we select only one progressive wave and obtain instead of (2.8)–(2.9):

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} + c_s^2 \frac{1}{n}\frac{dn}{dv}\frac{\partial v}{\partial x} = 0, \tag{2.13}$$

$$\frac{dn}{dv}\left(\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x}\right) + n\frac{\partial v}{\partial x} = 0. \tag{2.14}$$

Multiplying (2.13) by dn/dv and subtracting the result from (2.14) we obtain

$$c_s^2\left(\frac{dn}{dv}\right)^2 = n^2. \tag{2.15}$$

We find from here $c_s dn/dv = \pm n$. This value can be substituted into (2.13) giving rise to

$$\frac{\partial v}{\partial t} + (v \pm c_s)\frac{\partial v}{\partial x} = 0. \tag{2.16}$$

The plus sign here relates to the wave propagating along the x-axis in positive direction, and the minus sign corresponds to the wave propagating in the opposite direction.

Equation (2.16) is similar to (2.1). With the help of the frame of reference moving with velocity $\pm c_s$ it can be transformed to the form of (2.3) for noninteracting particles. We could say that in this case we deal with a "beam" of noninteracting phonons — sound excitations. According to (2.12) such linear excitations are transfered with the sound velocity and this statement is valid for any function $v(x)$ — they propagate either with velocity c_s or $-c_s$.

In accordance with equation (2.16) the long wave ion sound perturbation evolves in exactly the same way as a beam of noninteracting particles. Namely a periodic perturbation as in Fig. 2.1a is distorted in such a way as to proceed to an "overturn" of the front exactly in the same way as is shown in Fig. 2.1. But the "turnover" itself is not relevant here: in the case of an ion sound wave equation (2.16) is valid in the long wavelength approximation only. When the front of a wave becomes steep enough we are faced with dispersion effects.

2.2. Waves in weakly dispersive media

2.2.1. Nonlinearity and dispersion

In the absence of dispersion nonlinear terms lead to wave overturn. The higher harmonic perturbations accumulate and propagating with the same velocity they produce a coherent effect of increase of $\partial v/\partial x$ on the front of the wave up to the overturn.

But this picture drastically changes in media with dispersion. Indeed, harmonics with different k-values propagate with different velocities. Hence, the harmonic phase coherency is destroyed and overturn can be avoided. In other words the dispersion can compete with the nonlinearity: the nonlinear effects tend to increase the derivative on the wave front whereas the dispersion tends to destroy the coherence and to suppress the derivative.

To discuss the physics of these phenomena it is reasonable to consider at first a simple case when both nonlinearity and dispersion effects are small. They do not interfere very much in this case and can be considered as separate competing effects.

In this respect, small amplitude ion sound is the most appropriate for consideration. But to start with an even simpler example we consider first the waves on a shallow water surface.

2.2.2. Waves on a shallow water surface

The most convenient natural object for weak dispersive wave observation is shallow water. Let us consider perturbations with a wavelength much larger than the layer thickness h_0. We assume that these perturbations depend on the coordinate x only, so that the pressure p and horizontal velocity v depend upon t and x. For shallow water the velocity v can be considered as not dependent upon the vertical coordinate z. Thus, we have the following approximate equation for velocity:

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} + \frac{1}{\rho}\frac{\partial p}{\partial x} = 0. \tag{2.17}$$

Here, by the pressure p we understand its averaged value along z. The pressure p is greater where the layer has greater thickness. Namely the corresponding pressure increase is equal to $(h - h_0)\rho g$ where h is the local thickness of the layer, ρ is the water mass density and g is the gravitational acceleration. Thus, equation (2.17) takes the form

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} + g\frac{\partial h}{\partial x} = 0. \tag{2.18}$$

The thickness h is subjected to the continuity equation:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(vh) = 0. \tag{2.19}$$

This shows that the temporal rate of change $\partial h/\partial t$ is produced by the difference of fluxes hv through the sections at x and $x + dx$.

Equations (2.18)–(2.19) look like the gas dynamics equations at $\gamma = 2$, where γ is the adiabatic constant. This means that the waves on the shallow water surface have no dispersion in the linear approximation and the nonlinear wave front tends to overturn.

Again we can try to find the nonlinear solution for the simple progressive wave. To do this we put $h = h(v)$ and repeat the same algebraic actions as in the previous Section. We obtain in such a manner

$$\frac{\partial h}{\partial v} = \pm\sqrt{\frac{h}{g}}, \quad \frac{\partial v}{\partial t} + \left(v \pm \sqrt{gh}\right)\frac{\partial v}{\partial x} = 0. \tag{2.20}$$

We obtain from the first equation $h^{1/2} = h_0^{1/2} \pm v/2\sqrt{g}$. After substitution of this value into (2.20) we get

$$\frac{\partial v}{\partial t} + \left(\frac{3}{2}v \pm c_0\right)\frac{\partial v}{\partial x} = 0, \tag{2.21}$$

where $c_0 = \sqrt{gh_0}$ is the phase velocity of linear waves. Using the moving frame of reference and notation $u = 3v/2$ we obtain

the equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \tag{2.22}$$

which is quite familiar.

Thus, nonlinear effects in nonlinear nondispersive media can be observed directly in each shallow basin. We will see later that weak dispersion effects can also be observable.

2.2.3. Korteweg–de Vries equation

We now shall take into account weak dispersion effects. For ion sound waves with $k^2d^2 \ll 1$ the weak dispersion can be taken into account by relation (1.31) for the frequency of the harmonic wave or by the following relation for the phase velocity:

$$v_p = c_0\left(1 - k^2/2k_0^2\right). \tag{2.23}$$

As for waves on a shallow water surface the corresponding relation for the phase velocity can be found with the help of expression (1.19) for the frequency. If the surface tension is not important the corresponding relation looks like (2.23) with $c_0 = \sqrt{gh_0}$ and $k_0 = \sqrt{3}/h_0$. But for very shallow basins the surface tension may be important and the sign of dispersion can even change if $h_0 \leq 1$ cm.

To take into account the weak dispersion in the case of weak nonlinearity we can use equation (1.7) in which the driving force was found in the linear approximation. For ion sound waves propagating in the positive direction along the x-axis the phase velocity is equal to $v_p = c_s(1 + k^2d^2)^{-1/2}$. This value can be substituted into (1.7) but since $k^2d^2 \ll 1$ we can use relation (2.23) and obtain approximately

$$\frac{\partial v}{\partial t} + c_s\left(1 - \frac{k^2}{2k_0^2}\right)\frac{\partial v}{\partial x} = 0. \tag{2.24}$$

The second term in this equation is written for a single harmonic of the type $\exp(-i\omega t + ikx)$. It can be easily generalized to

any superposition of harmonics by a very simple replacement: $k^2 v = -\partial^2 v/\partial x^2$.

It is convenient to introduce a new frame of reference moving with velocity c_s. Then the term $c_s \partial v/\partial x$ in (2.24) disappears.

To be freer with the application of (2.24) to other media we shall use the notation c_0 instead of c_s and the notation u instead of v. We have to emphasize that v or u are the velocities of media in the steady state frame of reference whereas the x coordinate is measured in the frame of reference moving with velocity c_0.

The linear equation (2.24) describes the local ion acceleration effect under the action of a driving force. If the amplitude of oscillations is not very small we can replace $\partial v/\partial t$ by dv/dt and then the local ion acceleration will be in balance with the same expression for driving force.

Thus, taking into account all these arguments and the fact that $du/dt = \partial u/\partial t + u\partial u/\partial x$ we obtain an equation which is relevant to the description of weakly nonlinear waves with weak negative dispersion:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{c_0}{2k_0^2}\frac{\partial^3 u}{\partial x^3} = 0. \tag{2.25}$$

This equation was derived by Korteweg and de Vries in 1895 and was named after them.

Equation (2.25) is written for a medium with the negative dispersion. For positive dispersion waves the only difference is that the last term in (2.25) has the opposite sign. It is interesting that both these cases are practically equivalent. Namely if we make the transformation $x \to -x$, $u \to -u$ then in equation (2.25) the first and the second terms keep their signs whereas the third term changes sign. Thus, this simple transformation changes the sign of dispersion. This means that in our frame of reference where all values are referred to the point $x_0 = c_0 t$ the perturbations of the media of both signs are exactly mirror symmetric. That is why it is sufficient to consider the negative

dispersion case only. The positive dispersion pattern can be obtained by a simple mirror symmetry transformation.

2.2.4. Periodic waves, solitons

The Korteweg – de Vries equation (2.25) is essentially nonlinear and seems to be not simple to solve. But in spite of this fact the theory of this nonlinear equation has been developed in detail so that modern mathematics is able to find all its solutions exactly. However, we do not intend to display the rigorous theory here. We would prefer instead to understand the qualitative physics behind equation (2.25). That is why we will try to understand its nature step by step.

At first let us look for some simple particular solutions. Namely we try to find solutions of the type $u = u(x - ct)$ which correspond to simple progressive waves. Here $c = \text{const}$ is the phase velocity. More accurately, c is a small addition to the main phase velocity c_0 in the laboratory frame of reference. For progressive waves we can replace the time derivative by the derivative respect to the x-coordinate: $\partial u/\partial t = -c\partial u/\partial x$. After this replacement the equation becomes an ordinary one and can be once integrated:

$$\frac{c_0}{2k_0^2}\frac{d^2u}{dx^2} = cu - \frac{u^2}{2}. \tag{2.26}$$

This equation could contain a constant after integration of (2.25) but we put it to zero for simplicity. Note that this constant can be always nullified by the appropriate choice of the moving frame of reference.

Equation (2.26) is nonlinear again but it can easily be analyzed with the help of a simple analogy. Namely let us consider the x-variable as time and the u-variable as the coordinate of some material particle. Then equation (2.26) can be considered as the Newton equation with acceleration d^2u/dx^2, mass $c_0/2k_0^2$ and with the force on the right-hand side depending upon the

coordinate u. Now this equation can be written in the form

$$\frac{c_0}{2k^2}\frac{d^2u}{dx^2} = -\frac{\partial W}{\partial u}, \tag{2.27}$$

where

$$W = -c\frac{u^2}{2} + \frac{u^3}{3}. \tag{2.28}$$

We see that the W-function plays the role of the potential energy for our artificial particle. This potential energy is shown in Fig. 2.3.

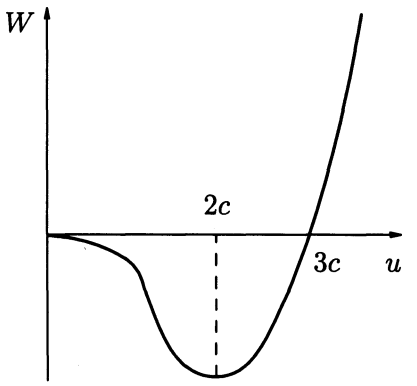

Fig. 2.3. "Potential energy" for a periodic wave.

The W-function has maximum at $u = 0$, reaches a minimum at $u = 2c$ and is twice equal to zero: at $u = 0$ and $u = 3c$. The periodic waves look like periodic oscillations of the particle in this potential well. The amplitude of oscillations is related to the "energy" of the particle in this well, including the kinetic energy.

A particle with "low energy" experiences oscillations near the bottom $u = 2c$. If we put $u = 2c + u'$ then for small u' equation (2.26) can be linearized:

$$\frac{c_0}{2k_0^2}\frac{d^2u'}{dx^2} = -cu'.$$

For a periodic harmonic wave we have $u' = u'_0 \exp\left[ik(x - ct)\right]$ where k is linked to the phase velocity c by the relation $c = c_0 k^2/2k_0^2$. Because $u = 2c + u'$ and the averaged value of u is not zero, the fluid moves with the averaged velocity $2c$. Therefore the wave propagates with respect to fluid with the velocity $-c$. Thus, in the laboratory frame of reference we have the wave propagating with phase velocity $v_p = c_0 - c = c_0 - c_0 k^2/2k_0^2$ which is in exact accordance with the weak dispersion law.

If we increase the "energy" of the oscillating particle, the amplitude increases too and the wave itself becomes asymmetrical. Namely its crests become steeper and the distance between neighboring crests increases because the right slope of the potential well in Fig. 2.3 is steeper than the left slope.

Finally we can consider a perturbation with the maximum possible amplitude when the effective energy is equal to zero. In this case our "particle" starts its motion near the point $u = 0$, $W = 0$, then it "rolls down" to the bottom of the well and continuing to roll "climbs" up to the point $u = 3c$, $W = 0$. Then it experiences the same motion in the backward direction and can reach the stationary state at $u = 0$. Thus, we have only one crest in this case. This special pattern is called the solitary wave or simply soliton. It is easy to check that the soliton-like solution of (2.26) looks like

$$u = \frac{u_0}{\cosh^2\left[(x - ct)/\Delta\right]}, \tag{2.29}$$

where u_0 is the amplitude and Δ is the soliton thickness. Indeed, if we insert function (2.29) into equation (2.26) we obtain:

$$\frac{c_0}{2\Delta^2 k_0^2}\left(\frac{4u_0}{\cosh^2 y} - \frac{6u_0}{\cosh^4 y}\right) = \frac{cu_0}{\cosh^2 y} - \frac{u_0^2}{2\cosh^4 y}, \tag{2.30}$$

where $y = (x - ct)/\Delta$. We see that this equation is satisfied if we choose

$$u_0 = 3c, \quad \Delta^2 = 2c_0/ck_0^2. \tag{2.31}$$

We obtain from here the following statement:

$$\Delta^2 u_0 = 6c_0/k_0^2 = \text{const.} \tag{2.32}$$

This means that the product of the square of the thickness and the soliton amplitude is constant. Thus, there is one parametric family of solitons. For instance, the amplitude may be chosen as a parameter and then the thickness is adjusted: the larger the amplitude, the narrower the soliton.

We underline once more that in accordance with relation (2.31) the amplitude of the soliton is proportional to its phase velocity: $u_0 = 3c$. Thus, in negative dispersion media the phase velocity is positive in the frame of reference moving with c_0. This means that such solitons are supersonic. Both in solitons and in strongly nonlinear waves the nonlinearity and dispersion compensate each other: the nonlinearity tends to overturn the crest while the negative dispersion slows down the higher harmonics and prevents the overturn. As a result a coherent pattern is produced which propagates without distortion or degradation.

A soliton is a supersonic pattern as well as periodic waves with large amplitudes. Thus, the increase of amplitude leads to a change of the velocity of the pattern in the laboratory frame of reference where the fluid is at rest on average: small amplitude perturbations are subsonic and large amplitude patterns are supersonic.

We have considered the negative dispersion case. When the weak dispersion is positive we have to change the sign on the left-hand side of equation (2.26). Then the potential well in (2.28) appears at $c < 0$, $u < 0$. This means that the soliton looks like a negative spike in this case. The same result can be obtained by the simple transformation $x \to -x$, $u \to -u$ which leads to the dispersion sign change. We see again that the positive dispersion soliton looks like a negative spike propagating with subsonic velocity. Both types of solitons are drawn in Fig. 2.4. The positive dispersion soliton of Fig. 2.4b is shown as propagating to the right because the averaged fluid velocity is not zero in this figure.

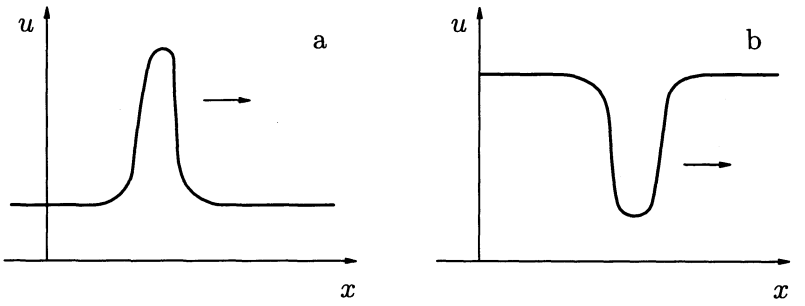

Fig. 2.4. Solitons in a medium with a weak negative (a) and positive (b) dispersion.

2.2.5. Local pulse evolution

We now consider the problem of temporal evolution of the one-dimensional nonlinear perturbation produced by the initial pulse. We assume that the initial perturbation looks like a bell-shape pulse with amplitude u_0 and thickness Δ. For $u_0 \ll c_0$ this perturbation is weakly nonlinear and splits very quickly into pairs of pulses propagating in opposite directions with a velocity close to c_0. We consider only one of these pulses, namely that propagating in the positive x direction. We assume that the medium has a weak negative dispersion. Thus, we can use the Korteweg–de Vries equation for this pattern description.

Nonstationary solutions of Korteweg–de Vries equation were first investigated numerically and later a complete analytic theory was developed. Based on these results we can use the following qualitative argumentation.

For simplicity let us assume that the initial perturbation has a shape exactly the same as that of the soliton. Namely we put $u(t = 0) = u_0 \cosh^{-2}(x/\Delta)$. This perturbation could propagate as soliton if its amplitude were related to its thickness by relation (2.32). In other words the value of σ defined by the relation

$$\sigma = \frac{u_0 k_0^2 \Delta^2}{6c_0} \tag{2.33}$$

is exactly equal to unity for a soliton. Thus, we know one of the solutions at $\sigma = 1$. Now we can consider the cases which are different from the $\sigma = 1$ case. For simplicity we assume that σ is either much less than unity or much larger than unity.

The parameter σ can really be considered as a measure of nonlinearity. It is easy to see that the parameter σ defined by relation (2.33) looks like the ratio of the nonlinear term in the Korteweg–de Vries equation to the dispersion term. Namely in order of magnitude this ratio looks like $\sim u_0^2/\Delta$: $u_0 c_0/k_0^2\Delta^3 = u_0\Delta^2 k_0^2/c_0 = 6\sigma$.

In the first case, $\sigma \ll 1$, we have a small amplitude perturbation. This evolves according to Fig. 1.6. Namely the wave packet has the main spike close to the point $x = c_0 t$ and is accompanied by the retarding oscillation structure which propagates with a subsonic velocity. The packet becomes broader with time and its amplitude decays. The first spike broadens too.

A delicate moment appears here. If the energy of the first spike is not changing very much at initial stage of evolution then the product $u_0^2\Delta$ is almost constant. This means that $u_0 \sim \Delta^{-1/2}$ and hence the parameter $\sigma \sim u_0\Delta^3$ increases with time as $\sigma \sim \Delta^{3/2}$. Thus, we have no guarantee that the σ parameter will continue to be much less than unity. Increasing the nonlinearity parameter σ means that even in the case $\sigma \ll 1$ a broad soliton can be generated. This statement will be clarified somewhat later.

The most interesting perturbation is that with $\sigma \gg 1$. Such a perturbation evolves initially as an ordinary progressive wave without dispersion. Its front slope steepens up to the stage when it becomes similar to the soliton slope with the same amplitude u_0. Then the dispersion starts to play an important role. Computer simulation shows that at this moment a soliton with amplitude u_0 is formed. This soliton, being supersonic, splits off the main pattern and runs away. Then the profile $u(t, x)$ left after the birth of the first soliton continues to steepen and a second soliton splits away. This process continues until complete splitting of the initial pulse into solitons. At this stage

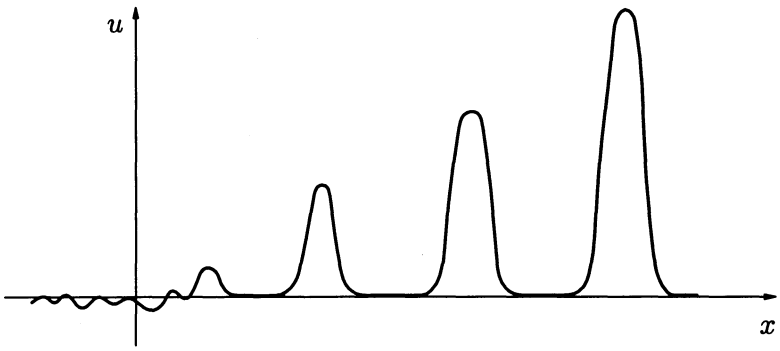

Fig. 2.5. Soliton from an initial pulsed perturbation with $\sigma = 8$. According to calculations by Yu. Berezin and V. Karpman [1].

several solitons are generated and an additional "dispersive tail" appears as is shown in Fig. 2.5.

The perturbation with $\sigma = 8$ in Fig. 2.5 has split into four solitons and a short wave dispersive packet of small amplitude.

Another example of soliton generation is illustrated by Fig. 2.6 where the results of numerical simulation of a harmonic perturbation are shown. These results were obtained by N. Zabusky and M. Kruskal [2]. The initial perturbation looks like $\cos(x/\pi)$. Figure 2.6 displays one period only: the real picture has to be extended periodically in both directions.

As we see, the initial stage of perturbation evolution looks like a simple progressive wave steepening with the tendency to overturn. But when the front becomes steep enough the dispersion steps into play and the first soliton starts to detach. After the split the rest of perturbation repeats the evolution towards the overturn and a second soliton is generated. This process continues until the perturbation pulse has been completely transformed into series of solitons (see Fig. 2.6 stage C).

It is interesting to note that the soliton tops are situated on a straight line. The same is seen in Fig. 2.5, corresponding to the single pulse evolution. This fact has a very simple explanation. All the solitons were born approximately at the same place — at

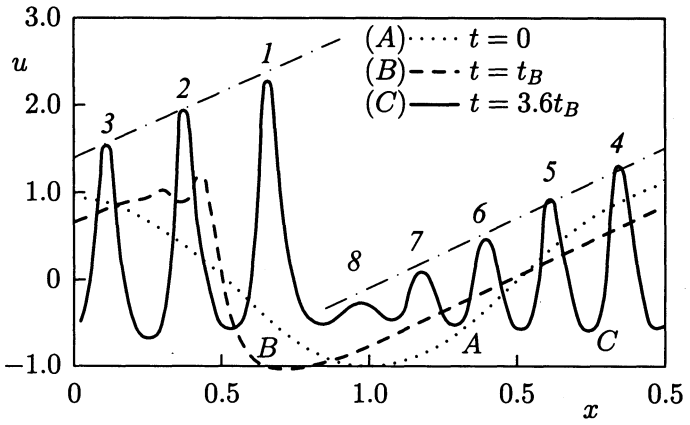

Fig. 2.6. Periodic wave evolution in a weakly dispersive medium: $t = t_B$ corresponds to the start of overturn.

the front of the perturbation. The velocity of soliton propagation is equal $u_0/3$ where u_0 is its amplitude. Therefore the distance of the soliton from the point of birth is equal to $\Delta x = ct = u_0 t/3$. This distance increases with time. For the envelope line of the soliton tops we have a very simple equation

$$\frac{du_0}{dt} = \frac{\partial u_0}{\partial t} + \frac{u_0}{3}\frac{\partial u_0}{\partial x} = 0, \tag{2.34}$$

which is quite similar to the equation for a beam of non-interacting particles. The corresponding picture of the envelope looks like a saw with sharp discontinuity at the soliton with the largest amplitude and is akin to a shock wave without dissipation. Figure 2.6 displays the envelope line at time $t = 3.6t_B$. Later on we have so to say a multi-beam pattern of solitons: solitons with larger amplitudes being quicker overtake the slower solitons as shown in Fig. 2.7.

Figure 2.7 displays a very interesting result of numerical simulation concerning the soliton interaction. As is shown, solitons having different amplitudes cross each other without distortion after interaction. We note that the interaction simply leads

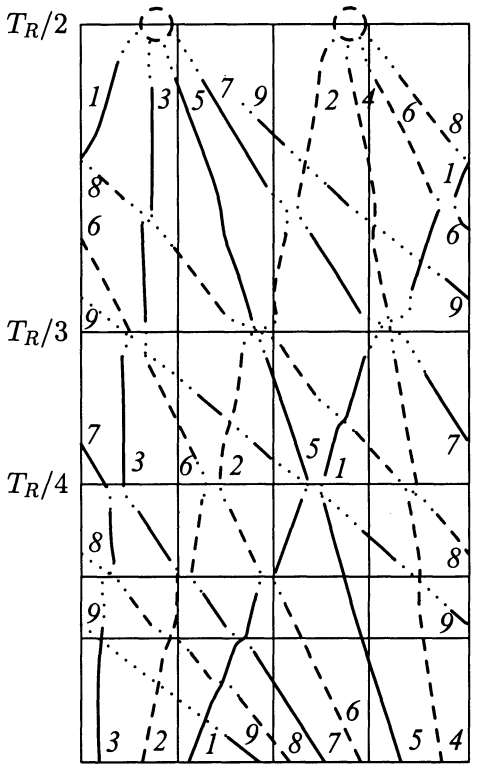

Fig. 2.7. Soliton trajectories. T_R is the time of return to the initial state.

to a phase shift. During the soliton overlap there is no superposition: the net amplitude turns out to be less than the sum of the crossing solitons' amplitudes. It is better to say that a new two-soliton state appears. Later on this two-soliton pulse decays into two separate solitons which are exactly the same as before the collision. The final state looks like a simple penetration of one soliton through the other.

When the difference between the soliton amplitudes is not very high, the character of their interaction changes. The solitons approach each other and the larger soliton "pours out" a part of its perturbation into the slower soliton. After such an en-

ergy exchange the quicker soliton is slowed down and the slower one is accelerated. An example of such an "elastic collision" is shown in Fig. 2.7 for solitons with numbers 2 and 6 at $t = T_R/4$.

Figure 2.7 also displays the collisions of three and even more solitons simultaneously. For instance, at $t = T_R/2$ the fusion of all the even and all the odd solitons takes place. Numerical simulation shows that at $t = T_R$ corresponding to the time of return all the solitons again gather into a single pulse. The perturbation comes back to almost the initial cosine shape. A complete restoration does not occur, however, because the solitons' velocities are not completely conquerable. On subsequent returns at $t = 2T_R$, $3T_R$ and so on the deviation from the initial state increases permanently.

Numerical calculation shows that in media with negative dispersion solitons are born only in the case when the initial perturbation is positive. A negative pulse can not generate solitons: it evolves similarly to a low amplitude perturbation spreading out into a broad wave packet.

The results of pulse evolution in media with negative dispersion can easily be extrapolated to media with positive dispersion. Solitons in the case of weakly positive dispersion look not like spikes but like wells. They propagate with a subsonic velocity which is lower, the larger the soliton's amplitude. As for the linear "dispersive tail," it propagates with supersonic velocity. Correspondingly, the picture of the pulse evolution looks like Fig. 2.6 turned through 180°. Namely, the negative pulsed perturbation generates negative solitons receding from the point $x = 0$ and propagating with velocity $c_0 - |u_0|/3$ in the laboratory frame of reference. As for the small dispersive tail, it is situated on the right-hand side of the point $x = 0$.

2.2.6. *Analytical solution of the Korteweg–de Vries equation*

The Korteweg–de Vries equation looks very elegant and an elegant theory of its analytical treatment was developed by

C. Gardner, L. Green, M. Kruskal and R. Miura [3]. Before explaining their approach we shall try to present some qualitative preliminary remarks.

As we have seen before, the evolution of a non-interacting particle beam is described by a nonlinear equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0.$$

The form of this equation is really nonlinear, but we have simply the free motion of particles with constant velocity: $du/dt = 0$. Thus, in the Lagrange representation it looks like the simplest linear equation: $u = \text{const}$. The coordinate x of the particle is changing linearly in time, $x = x_0 + ut$. Hence, for any initial profile $u(x, t = 0) = u_0(x)$, which can be represented in the reversed dependence of x upon x_0, $x(u, 0) = x_0(u)$, we can find the $x(u, t)$ dependence for any time value: $x = x_0(u) + ut$. In other words, the problem of finding the $x(u, t)$ dependence relates to a linear equation. But to find the reversed profile u upon x at any t, the algebraic equation $x_0(u) + ut = x$ has to be solved. This equation is generally nonlinear. We faced approximately the same situation when considering the progressive simple wave. The corresponding process of evolution looks like independent motion of phonons — excitations of ion sound. The difference is in the presence of a driving force. This force in the linear approximation can be found with the help of equation (1.7). Replacing the derivative $\partial v/\partial t$ with $dv/dt = \partial v/\partial t + v\partial v/\partial x$ we immediately obtain the Korteweg–de Vries equation. Thus, because we have a very simple extension of free particle motion we can not exclude the possibility that the Korteweg–de Vries equation has a hidden symmetry which corresponds to independent motion of some entities of excitations. This is really so and can be argued in the following way.

It turns out that the Korteweg–de Vries equation can be obtained by the classical variational principle of minimum action. For this purpose the Lagrangian representation of this equation is more convenient.

Let φ be the velocity potential defined by the relation $u = \partial\varphi/\partial x$. Then the Lagrangian $\mathcal{L}$ can be defined as

$$\mathcal{L} = \frac{1}{2}\int\left\{\frac{\partial\varphi}{\partial t}\frac{\partial\varphi}{\partial x} + \frac{1}{3}\left(\frac{\partial\varphi}{\partial x}\right)^3 - \frac{c_0}{2x_0^2}\left(\frac{\partial^2\varphi}{\partial x^2}\right)^3\right\}dx. \qquad (2.35)$$

It is not difficult to check that the Lagrange equation

$$\frac{\partial}{\partial t}\frac{\delta\mathcal{L}}{\delta\varphi_t} - \frac{\delta\mathcal{L}}{\delta\varphi} = 0,$$

where $\varphi_t = \partial\varphi/\partial t$ coincides with the Korteweg–de Vries equation (2.25).

Using the well known Nöther theorems, the laws of energy and momentum conservation can be obtained with the help of the Lagrangian (2.35). Namely, the invariance with respect to a shift along the x-axis leads to the following expression for momentum:

$$P = \frac{1}{2}\int u^2 dx, \qquad (2.36)$$

while the invariance of the Lagrangian with respect to a shift in time gives the expression for the energy $\mathcal{E}_0$:

$$\mathcal{E}_0 = \frac{1}{2}\int\left\{-\frac{u^3}{3} + \frac{c_0}{2k_0^2}\left(\frac{\partial u}{\partial x}\right)^2\right\}dx. \qquad (2.37)$$

This energy is measured in the frame of reference moving with sound velocity c_0. In the laboratory frame of reference the wave energy is equal to $\mathcal{E} = \mathcal{E}_0 + c_0 P$. As the value u is assumed to be much smaller than c_0, we have approximately

$$\mathcal{E} \simeq c_0 P = \frac{c_0}{2}\int u^2 dx.$$

In addition to P and $\mathcal{E}$ the Korteweg–de Vries equation has an infinite series of the integrals of motion. The equation itself can be written in the Hamiltonian representation with a completely separable space of variables.

The phase space separability means that different coordinates of a phase point evolve independently of each other. Thus, we indeed have the situation when all the entities are moving independently. Solitons, for instance, look like such entities.

An exact analytical theory of the Korteweg – de Vries equation was developed in the famous paper by C. Gardner, L. Green, M. Kruskal and R. Miura [3]. They found a very original and fruitful method of solution of the nonlinear Korteweg – de Vries equation. Later this method was further developed and refined by L. Faddeev and V. Zakharov [4] with an application to a special class of nonlinear equations. Later on a lot of papers and books were published applying this method to many nonlinear equations.

The analytical theory of the Korteweg – de Vries equation uses an analogy with quantum mechanics. That is why we have to recall some quantum mechanical relations.

Let the function $\psi(x,t)$ satisfy an equation similar to the Schrödinger equation:

$$i\frac{\partial\psi}{\partial t} = H\psi. \tag{2.38}$$

Here H is the Hamiltonian operator playing the role of the Hamiltonian of our problem. It is well known that any Hermitian operator L describing some physical quantity satisfies the following equation:

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + i\big[H, L\big]. \tag{2.39}$$

Here $[H,L]$ denotes the commutator $HL - LH$. If $dL/dt = 0$ then this operator is considered independent of time. This means that its eigenvalues λ defined by the relation

$$L\psi = \lambda\psi \tag{2.40}$$

are not dependent upon time. It turns out that the Korteweg – de Vries equation can be considered the steady state condition

$dL/dt = 0$ for a specially chosen operator L with the special Hamiltonian H. Namely, it is easy to check that by the choice

$$L = \frac{3c_0}{k_0^2}\frac{\partial^2}{\partial x^2} + u, \tag{2.41}$$

$$H = -\frac{2c_0 i}{k_0^2}\frac{\partial^3}{\partial x^3} - \frac{i}{2}\left(u\frac{\partial}{\partial x} + \frac{\partial}{\partial x}u\right), \tag{2.42}$$

the Korteweg–de Vries equation can be expressed in the form of L operator conservation:

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + i\big[H, L\big] = \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{c_0}{2k_0^2}\frac{\partial^3 u}{\partial x^3} = 0. \tag{2.43}$$

This fact allows the development of an analytical scheme for solving the Korteweg–de Vries equation. But even more interesting conclusions can be made directly from the relations above. Let us return to equation (2.40) with operator (2.41). We can represent it in a form akin to the Schrödinger equation:

$$-\frac{3c_0}{k_0^2}\frac{\partial^2\psi}{\partial x^2} - u\psi = -\lambda\psi. \tag{2.44}$$

The first term in this equation can be considered as the kinetic energy and the second term as the potential energy. Hence, the eigenvalue $-\lambda$ corresponds to the total energy. If u decays as $x \to \pm\infty$ quickly enough then positive λ values correspond to a discrete spectrum of bound states and negative λ values relate to the positive energy continuum.

As was claimed above, the eigenvalues of the L operator are conserved with time. Hence, when the profile $u(x,t)$ changes in time according to the Korteweg–de Vries equation then the L operator has constant eigenvalues in accordance with relation (2.43). In other words all the eigenvalues $\lambda_n = \text{const}$ in spite of the fact that $u(x,t)$ varies in time.

In particular, we can try to find eigenvalues for solitons. It turns out that for the soliton profile $u = u_0 \cosh^{-2}(x/\Delta)$ equation (2.44) can be solved exactly. There is only one eigenvalue

$\lambda = u_0/2$ for a soliton with amplitude u_0. Next, the eigenvalue $\lambda = 0$ corresponds to the eigenfunction $\psi = \tanh(x/\Delta)$, which does not decay as $x \to \pm\infty$ and therefore corresponds to continuous spectrum.

When solitons evolve in time all the eigenvalues λ_n will be time-independent. In particular, if the initial pulse evolves and splits into solitons, all the λ_i values are conserved. But when solitons are separated and are sufficiently far from each other, then all the discrete λ_n values are equal to $u_n/2$ where u_n is the amplitude of soliton number n.

It follows from here that for any initial profile $u(x,0)$ it is possible to predict all the types of solitons which will born during the profile's evolution. For this purpose it is sufficient to solve equation (2.44) and to find all the discrete eigenvalues λ_n for a given $u(x,0)$ perturbation. Then the amplitudes of solitons can be found using the relation $u_n = 2\lambda_n$. We can now estimate the number of solitons generated by a given initial pulse $u(x,0)$. As is known, the n-th energy level in a potential well of thickness Δ increases with n as $\lambda_n \simeq n^2/\Delta^2$. There is only one level for a soliton. Therefore we can estimate the number of solitons N generated by initial perturbation as $N \simeq \sqrt{\sigma}$.

We can make one more conclusion. Equation (2.44) shows that solitons can be produced only by positive perturbations, $u > 0$. If $u < 0$ there are no bound states, and the perturbation will evolve similarly to a linear one at its final stages. Vice versa any negative perturbation generates at least one soliton even if it is very small $\sigma \ll 1$. This conclusion follows from the well known quantum mechanical statement that any one-dimensional potential well has at least one discrete level.

Thus, the total number of solitons together with their amplitudes can be found for any initial perturbation without solving the nonlinear Korteweg–de Vries equation but simply with the help of the linear equation (2.44). It turns out that this equation can serve as a tool for exact solution of the Korteweg–de Vries equation [5, 4]. The asymptotic behavior of the ψ-function is used for this purpose. Namely, if the u profile decreases suffi-

ciently quickly as $x \to \pm\infty$ the asymptotic solution of (2.44) looks like a superposition of bound and continuous eigenfunctions:

$$\begin{aligned} \psi_n &= a_n(t)\exp(-\kappa_n x), \\ \psi_k &= a_k(t)\exp(-ikx) + b_k(t)\exp(ikx). \end{aligned} \tag{2.45}$$

Here $\kappa_n = k_0\sqrt{\lambda_n/3c_0}$, $k = k_0\sqrt{-\lambda/3c_0}$, while $\lambda_n > 0$ is the discrete eigenvalue and $\lambda < 0$ corresponds to the continuous spectrum of L operator. The temporal dependence of amplitudes in expression (2.45) can easily be derived using the Korteweg–de Vries equation which claims in fact the conservation of the L operator as is seen from (2.43). As $x \to \pm\infty$ the $u\,\partial u/\partial x$ term in this equation can be dropped and then the linear equation obtained gives the following temporal dependence for amplitudes:

$$\begin{aligned} a_n(t) &= a_n(0)\exp\left(2c_0k_0^{-2}\kappa_n^3 t\right), \\ a_k(t) &= a_k(0)\exp\left(2ic_0k_0^{-2}k^3 t\right), \\ b_n(t) &= b_n(0)\exp\left(-2ic_0k_0^{-2}k^3 t\right). \end{aligned} \tag{2.46}$$

The ratio $S_k(t) = b_k(t)/a_k(t)$ can be considered as the reflection coefficient for a quantum mechanical plane wave which comes from infinity and is scattered on the $u(x,t)$ potential in accordance with equation (2.44). From the quantum mechanical scattering theory we know that the potential $u(x,t)$ can be found if S_k is known as a function of k for the continuous spectrum and the characteristics κ_n and a_n of the discrete spectrum are known as well. This problem is called the inverse scattering problem.

Hence, the following procedure exists for the analytical solution of the Korteweg–de Vries equation. At first the discrete values λ_n and initial values of the collision matrix $S_k(0)$ have to be found with the help of the linear equation (2.44) for a given initial profile $u(x,0)$. Then the temporal dependence $S_k(t)$ is found in accordance with relation (2.46). After that the $u(x,t)$ profile can be found from the inverse scattering theory [6] using

the known values $a_n(t)$, $S_k(t)$ as functions of time. As for the inverse scattering problem, its solution relates to solution of a linear integral equation. Thus, at any stage of such procedure only linear equations have to be solved in spite of the fact that eventually the solution of the nonlinear Korteweg – de Vries equation is found. It is evident that this solution depends nonlinearly upon the initial function $u(x, 0)$.

The possibility to solve the Korteweg – de Vries equation analytically is deeply related to the nature of this equation as being a Hamiltonian system with separable variables. At present a lot of equations are known which are fully integrable by the inverse scattering method. Details of this theory can be found in references [7].

2.2.7. Bow waves and bow solitons

We can now consider nonlinear Cherenkov emission in the weakly dispersive media. The corresponding perturbations propagate as waves in the vicinity of the Mach line defined by the well known Cherenkov condition $\cos\theta = v_p/v_0 \simeq c_0/v_0$. Across the Mach line the evolving perturbation looks one-dimensional and can be described with the help of the Korteweg – de Vries equation where t stands for the evolution time along the Mach line. If we assume a weak dispersion then shallow water can be accepted as an example. Nonlinearity means in this case that the amplitude of the water layer perturbation is not very small compared with the initial depth of the basin layer.

When observed along the Mach line the pulse evolution looks like the temporal evolution of a perturbation described by the Korteweg – de Vries equation. Namely it can produce supersonic solitons for which $\cos\theta = v_p/v_0$ is slightly larger than c_s/v_0. Such solitons or a single soliton run in front of the Mach line. Then a low amplitude dispersive tail can be produced retarding the Mach line. But due to dispersion the amplitude of this tail decays sufficiently rapidly with distance so that at long distances only solitons can be observed.

Soliton generation can be observed on any shallow stream. An especially striking picture arises on streams on smooth road surfaces, for instance, asphalt. Solitons are striking as a network of solitary waves crossing such a broad stream at some permanent angle to the stream flow. Sometimes the roughness of the bottom is seen which generates these solitons. When the perturbation is strong enough, produced by a stone for instance, not only one but a pair of solitons are seen.

2.2.8. Shock waves in a weakly dispersive medium

It was assumed above that the all dissipative mechanisms are absent or are negligible compared with the dispersion. In reality dissipation is present. If it is a wave on the water surface then viscosity can play a certain role. Dissipation produces monotonic decay of waves. But this is not such an interesting effect. Much more interesting is the possibility of appearance of a new wave pattern. This possibility really exists. We mean shock wave generation by an external energy supply. Such waves in weakly dispersive media have new quite interesting features.

To take viscosity into account we have to add a new term in the Korteweg–de Vries equation:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{c_0}{2k_0^2}\frac{\partial^3 u}{\partial x^3} = \nu\frac{\partial^2 u}{\partial x^2}. \tag{2.47}$$

Here ν is an effective kinematic viscosity. Let us again try to find a solution of the type $u(x-ct)$, i.e., looking like a stationary wave. For such functions the derivative with respect to time is again proportional to the space derivative and (2.47) can be once integrated. We obtain:

$$\frac{c_0}{2k_0^2}\frac{d^2u}{dx^2} - \nu\frac{du}{dx} = cu - \frac{u^2}{2} = \frac{dW}{du}. \tag{2.48}$$

Here W is again the potential for an artificial "particle" with the coordinate u. Thus, (2.48) again describes a nonlinear oscillator but now with dissipation, if $-x$ is considered as the time

variable. "Time" $-x$ flows from $x = +\infty$ to zero and then towards $x = -\infty$. For this interpretation the dissipative term in (2.48) indeed provides the decay of u in time. The evolution of the particle in time is shown in Fig. 2.8.

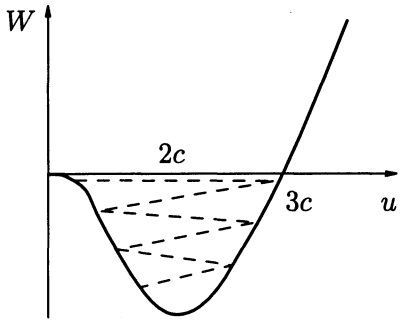

Fig. 2.8. Particle motion in a potential well $W(u)$ in the presence of dissipation.

As we see the motion starts at u close to zero where the particle is in unstable equilibrium. Due to instability the particle deviates from the point $u = 0$ and slides down to the bottom of the well. Then due to inertia it climbs up the right slope and slides down again. These oscillations repeat several times until complete relaxation of the particle at the well bottom. Thus, this solution is definitely asymmetrical: $u(x) \to u_0$ at $x \to -\infty$ and $u(x) \to 0$ at $x \to +\infty$. It is called a shock wave. The phase velocity of such wave propagation is equal to c and we see in Fig. 2.8 that this value is related to the value u_0 corresponding to the potential energy minimum as $c = u_0/2$. The shock wave structure in a medium with weak negative dispersion is shown in Fig. 2.8a.

The shock wave is shown in the laboratory frame of reference and that is why the phase velocity is equal to $(c_0 + v_0/2)$ and the variable v is used instead of u to avoid misunderstanding.

The shock wave in Fig. 2.9a propagates in the positive direction of the x-axis with phase velocity $c = v_0/2$ (the notation v instead of u is used here to cover both negative and positive

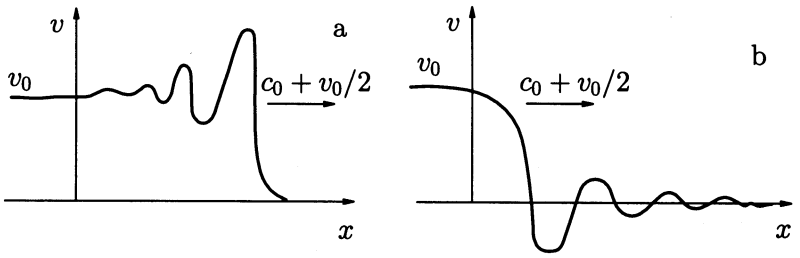

Fig. 2.9. Shock wave structures in media with negative (a) and positive (b) dispersion.

dispersions). In front of the shock the media is motionless, $v = 0$ at $x = \infty$. On the front itself the first spike has an amplitude, as is seen in Fig. 2.8, close to $3c = 3v_0/2$. Then the oscillation structure follows and as $x \to -\infty$ the velocity tends to v_0. To maintain the nonzero value of velocity v_0 at $x = -\infty$ it is necessary to have something like a piston pushing the fluid in spite of energy dissipation. Such a jump with $v = 0$ at $x = +\infty$ and $v = v_0 \neq 0$ at $x = -\infty$ is called a shock wave. In dispersive media this jump is accompanied by an oscillatory train just after the first spike.

If we increase the dissipation ν the number of oscillations diminishes and for sufficiently large ν a smooth slide of particle from the $u = 0$ point to $u = u_0 = 2c$ takes place. In this case it is a simple shock wave with low amplitude because of the $u \ll c_0$ condition.

Figure 2.9a relates to a medium with the weak negative dispersion. In this case the first spike-like soliton has the maximum amplitude and is accompanied by the oscillation structure. In the case of positive dispersion the oscillation structure runs in front of the main shock discontinuity as is shown in Fig. 2.9b. As we know the positive dispersion pattern can be obtained from the negative dispersion picture by a simple rotation through 180°. If we rotate Fig. 2.9a through 180° in the x, u variables then the shock structure will have $u < 0$ everywhere: $u = -u_0$ at $x \to \infty$ and $u \to 0$ at $x \to -\infty$. But in the laboratory frame of

reference the wave structure continues to propagate to the right with the subsonic phase velocity $c_0 - c$. Thus, in the picture obtained by such a simple rotation the fluid in front of the shock is moving with velocity $-u_0$. To obtain a more common situation with an initially unmoving fluid we can add an additional velocity $v_0 = u_0$ by changing the frame of reference. Now the fluid is at rest in front of the shock and the shock itself propagates with velocity $c_0 - c + u_0 \equiv c_0 - c + v_0 = c_0 + v_0/2$ as is shown in Fig. 2.9b. The velocity of the fluid behind the shock as $x \to -\infty$ is again equal to v_0 and the first minimum has a velocity value not far from $-v_0/2$. Thus, the main discontinuity is followed by an oscillating structure and a dip with $v \simeq -v_0/2$. Figure 2.10 demonstrates the experimental realizations of both these cases.

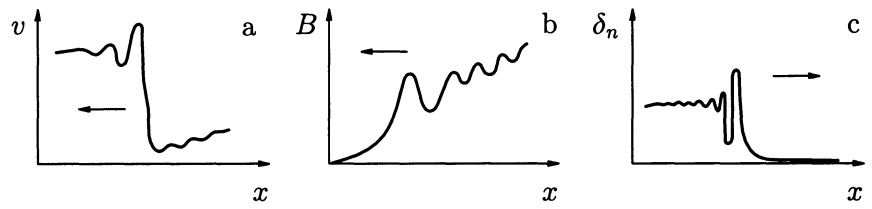

Fig. 2.10. Shock waves in a weakly dispersive medium. (a) Electromagnetic wave in a coaxial-spiral line with ferrite; (b) collisionless shock wave; (c) electrostatic ion-sound shock structure.

Figure 2.10a demonstrates the structure of the electromagnetic shock wave in a coaxial-spiral nonlinear transmission line [8]. The nonlinearity is produced by ferrite and the weak dispersion of long wave perturbations arises due to the finite cross-section. Waves in such a transmission line are described by a Korteweg–de Vries equation with a dissipation similar to (2.47).

Another example is represented by Fig. 2.10b, where the profile of a collisionless shock wave in plasma is shown [9]. As we see the dissipation in such a shock is not very large and does not prevent the existence of an oscillation structure. Figure 2.10c gives the profile of an electrostatic ion-sound shock [10]. As we

see this profile is very close to the theoretical one displayed by Fig. 2.9b. It corresponds to a weakly dispersive shock wave with a small amplitude.

2.2.9. Ion-sound wave overturn

We started the discussion of the nonlinear waves with a very simple example of a noninteracting particle beam. We saw that any perturbation in such a beam tends to overturn and to generate a multibeam pattern. Exactly the same situation arises in nonisothermal plasma ($T_i \ll T_e$) when the initial ion velocity perturbation has an amplitude much larger than the ion sound velocity c_s. In other words if the Mach number $M = v_0/c_s \gg 1$. Evidently, the interaction of ions with the electrostatic field is low in this case and the ions can be considered as neutral particles. On the other hand if the amplitude is small and the Mach number is much less than unity any perturbation evolves according to the Korteweg–de Vries equation without a hint of an overturn. The question arises of how these two cases relate each other and why the dispersion can not prevent the overturn of a wave with large amplitude. This matter can be illustrated with the help of an ion sound wave of large amplitude.

For this purpose we have to use the complete set of equations (1.1)–(1.4) for electrostatic perturbations in plasma with cold ions, $T_i \ll T_e$. We can represent them in the form:

$$\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x} = -c_s^2\frac{\partial \psi}{\partial x}, \tag{2.49}$$

$$\frac{\partial n'}{\partial t} + \frac{\partial}{\partial x}(n'v) = 0, \tag{2.50}$$

$$d^2\frac{\partial^2 \psi}{\partial x^2} = e^{\psi} - n'. \tag{2.51}$$

Here v is the ion velocity, $n' = n_i/n_0$ is the ion density related to its unperturbed value, $c_s^2 = T_e/m_i$, e^{ψ} is the electron density related to n_0 as defined by the Boltzmann distribution, $\psi = e\varphi/T_e$, and φ is the electric potential defined by the relation

$E = -\partial\varphi/\partial x$. We can look again for a periodic wave solution travelling with constant phase velocity, v_p. It is convenient to use the frame of reference where the wave is at rest i.e., the time derivative is zero. Then equations (2.49)–(2.50) can be integrated with the following result:

$$v^2 + 2c_s^2\psi = v_p^2 = \text{const}, \quad n'v = v_p = \text{const}. \tag{2.52}$$

In the first relation we have chosen the arbitrary constant of integration as being equal to v_p^2 prescribing a definite choice for the arbitrary constant in the electric potential. The second relation shows that v_p plays the role of the phase velocity: the fluid moves with this velocity where $n' = n_i/n_0 = 1$. With the help of relation (2.52) we can represent equation (2.51) in the following form:

$$d^2\frac{d^2\psi}{dx^2} = e^{\psi} - \frac{v_p}{(v_p^2 - 2c_s^2\psi)^{1/2}} = -\frac{dW}{d\psi}. \tag{2.53}$$

We can again consider this equation as an equation of motion for some artificial particle with coordinate ψ if the x variable is considered as time. The force on the left-hand side can again be considered as $-dW/d\psi$ where the potential energy W is equal to

$$W = -e^{\psi} - \frac{v_p}{c_s^2}\left(v_p^2 - 2c_s^2\psi\right)^{1/2}. \tag{2.54}$$

The derivative $dW/d\psi$ is equal to the force in (2.53) with the opposite sign. As we see it turns out to be zero at $\psi = 0$. Hence, W has minimum at $\psi = 0$. For small ψ the force in (2.53) can be taken in a linear approximation and we obtain:

$$d^2\frac{d^2\psi}{dx^2} = -\frac{c_s^2 - v_p^2}{v_p^2}\psi. \tag{2.55}$$

The solution of this linear equation can be taken in the simplest form $\psi \sim \exp(ikx)$ and we obtain the dispersion relation $v_p^2 = c_s^2/(1 + k^2d^2)$ of a linear sound wave. It is easy to check that the value of v_p in (2.53)–(2.54) can be either positive or negative

corresponding to the sign of the velocity in the moving frame of reference where the wave is in a steady state.

With the help of relations (2.53)–(2.54) the nonlinear waves can be analyzed using again the interpretation of a nonlinear oscillator. For instance, if we take into account the second order terms equation (2.53) will be similar to the Korteweg–de Vries equation but written in another moving frame of reference.

If we increase the amplitude of ψ further a difficulty will be faced: we see that ψ can not exceed $v_p^2/2c_s^2$. At $\psi = v_p^2/2c_s^2$ the ion velocity drops to zero and at the larger ψ values the ions would be reflected from the wave crest. A single ion sound stream ceases to exist for $\psi > v_p^2/2c_s^2$. Detailed analysis [11] shows that this critical case corresponds to the critical Mach number $M_* = 1.6$. In other words a single stream stationary nonlinear wave can not exist for $M_* > 1.6$. Thus, we see that the approximation of a weakly nonlinear wave with weak dispersion wave described by the Korteweg–de Vries equation is really relevant to the small amplitude case. For large values of the Mach number ion sound waves overturn.

2.3. Wave self-focusing and self-compressing

2.3.1. Soliton oscillations and self-focusing

Up to now we have considered one-dimensional waves only. All the physical values have been considered as functions of two variables, x and t, only. Now we proceed with the consideration of the two-dimensional waves.

We start our discussion with soliton behavior in two dimensions. We assume that a soliton propagates along the x-axis but is extended along the y-direction as is shown in Fig. 2.11a.

If the amplitude u_0 and the phase x_0 of this soliton do not depend upon y we have a normal one-dimensional soliton as discussed previously. It may be, for instance, a soliton on the water surface in a shallow basin with a thickness larger than 1 cm when the surface tension can be neglected. Thus, it is a negative dispersion case.

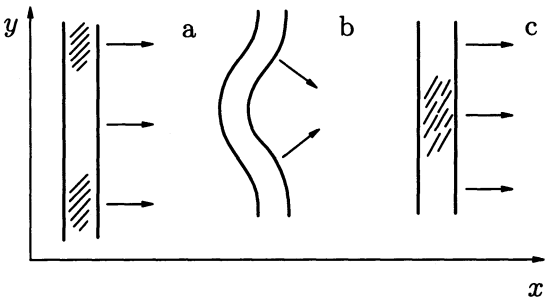

Fig. 2.11. Soliton oscillations in the case of weak negative dispersion.

Let us now assume that this soliton is slightly disturbed along the y-axis. Its amplitude u_0 depends upon y periodically as shown in Fig. 2.11a: the hatched regions there correspond to a slightly larger amplitude value u_0. As is known the velocity of soliton propagation $c_0 + c$ depends upon its amplitude and is equal to $c_0 + u_0/3$. Hence, its motion along the x-axis looks like

$$\frac{\partial x_0}{\partial t} = c_0 + \frac{u_0}{3}.$$

If u_0 depends upon y the different soliton phases x_0 move with different velocities so that the soliton is progressively curved. Its incline to the y-axis $\partial x/\partial y$ varies as

$$\frac{\partial}{\partial t}\frac{\partial x_0}{\partial y} = \frac{\partial}{\partial y}\left(\frac{\delta u_0}{3}\right), \tag{2.56}$$

where δu_0 is dependent upon the y part of u_0. The curved soliton has both concave and convex regions as shown in Fig. 2.11b. Due to the concave curvature of the central part of the soliton drawn in Fig. 2.11b its amplitude increases. Therefore this part starts to move more quickly and at a certain moment the soliton straightens as shown in Fig. 2.11c. Now the hatched area is displaced along the y-axis by half a wavelength. Thus, the soliton in a medium with negative dispersion oscillates near the equilibrium homogeneous position.

If the modulation is small, $\delta u_0 \ll u_0$, these oscillations can be considered in a linear approximation. When the soliton curvature is small, $\partial x/\partial y \ll 1$, its y-component of velocity can be written as

$$\delta u_y = -\frac{\partial x_0}{\partial y} c_0.$$

The transfer of energy $\mathcal{E}$ with this velocity leads to a change with time:

$$\frac{\partial \mathcal{E}}{\partial t} = -\frac{\partial}{\partial y}\left(\delta u_y \mathcal{E}\right). \tag{2.57}$$

The energy is proportional to $u_0^2 \Delta \sim u_0^{3/2}$ so that we have

$$\frac{\partial}{\partial t} u_0^{3/2} = \frac{3}{2} u_0^{1/2} \frac{\partial}{\partial t} \delta u_0 = -\frac{\partial}{\partial y}\left(\delta u_y u_0^{3/2}\right) \simeq c_0 u_0^{3/2} \frac{\partial^2 x_0}{\partial y^2}. \tag{2.58}$$

Here $u_0^{3/2}$ in the last equality can be considered as $u_0^{3/2} \simeq \text{const}$. By differentiating (2.58) with respect to time and using relation (2.56) we obtain:

$$\frac{\partial^2}{\partial t^2} \delta u_0 = \frac{2 c_0 u_0}{9} \frac{\partial^2}{\partial y^2} \delta u_0. \tag{2.59}$$

We see that in the case of negative dispersion, when $u_0 > 0$, we obtain a hyperbolic equation describing string oscillations. Thus, such a soliton is stable.

On the contrary when the soliton amplitude is negative equation (2.59) leads to instability. Hence, a one-dimensional soliton in a positively dispersive medium is unstable in two dimensions. The cause of this instability is very simple. The regions with larger amplitude lag behind the rest of the soliton and create a concavity. The new portions of perturbation rush towards this region leading to amplitude amplification of it. This instability looks like soliton self-focusing in a two-dimensional medium.

It is not difficult to derive the nonlinear equation which describes a two-dimensional soliton [12]. This equation looks

like a generalization of the Korteweg–de Vries equation for two-dimensional media. It is convenient to start with the linear dispersion relation for the negative dispersion medium:

$$\omega = c_0 k - c_0 k^3 / 2k_0^2. \tag{2.60}$$

The second term in this relation is small. When the soliton amplitude and phase are functions of the variable y the k number is not equal to k_x but has a dependence upon k_y: $k = \sqrt{k_x^2 + k_y^2}$. When the dependence upon y is a slow one, we can put approximately $k \simeq k_x + k_y^2/2k_x$. In the second small term of (2.60) k^3 can be replaced simply by k_x^3 so that the linear dispersion relation can be written in the form

$$k_x \omega - c_0 k_x^2 - \frac{1}{2} c_0 k_y^2 - \frac{c_0 k_x^4}{2k_0^2} = 0. \tag{2.61}$$

Now we can use the frame of reference moving with velocity c_0 along the x-axis and replacing ω by $i\partial/\partial t$ and $k_x = -i\partial/\partial x$, $k_y = -i\partial/\partial y$ we obtain the linear equation for a two-dimensional weakly dispersive medium:

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial t} + \frac{c_0}{2x_0^2}\frac{\partial^3 u}{\partial x^3}\right) = \frac{c_0}{2}\frac{\partial^2 u}{\partial y^2}. \tag{2.62}$$

To obtain the nonlinear equation we can recall that according to (1.7) the $\partial u/\partial t$ derivative is simply equal to the local ion acceleration and in the case of finite amplitude it has to be changed into $du/dt = \partial u/\partial t + u\partial u/\partial x$. Thus, we obtain the following nonlinear equation for the nonlinear wave in a weakly dispersive medium:

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{c_0}{2k_0^2}\frac{\partial^3 u}{\partial x^3}\right) = \frac{c_0}{2}\frac{\partial^2 u}{\partial y^2}. \tag{2.63}$$

This equation is called the Kadomtsev–Petviashvily equation.

For the case of positive dispersion the sign of the dispersive term has to be changed. But we can again use the transformation

$u \to -u$, $x \to -x$ and obtain

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{c_0}{2k_0^2}\frac{\partial^3 u}{\partial x^3}\right) = -\frac{c_0}{2}\frac{\partial^2 u}{\partial y^2}. \tag{2.64}$$

This equation differs from (2.63) by the negative sign in the right-hand side. It turns out that equation (2.64) has solutions localized both in the x and y directions. Such two-dimensional solitons will be considered in Section 2.5. Equations (2.63)–(2.64) are fully integrable. They conserve the number of solitons during their nonlinear interaction.

2.3.2. Self-focusing

Self-focusing can develop in the simpler cases of nonlinear wave packets when they are not very much different from harmonic plane waves. The amplitude and phase of such waves can slowly change in space and time. If the wave amplitude is not very small then effects of back influence on the packet propagation can arise. There are two most interesting effects of this nature: self-focusing and self-compression of waves packets.

The self-focusing effect was predicted theoretically by Askarian [13]. His starting idea was very simple. Let a very intensive laser beam propagate through an optically transparent medium. Several nonlinear effects like nonlinear electric susceptibility, electrostriction, thermal expansion, etc., can change the refractive index at the place of beam propagation. If this change is positive, i.e., the medium becomes optically denser, the beam itself can produce a local lens. This lens can focus this beam as is shown in Fig. 2.12.

Let us consider this effect in more detail. Let k_0 be the wave number for an almost plane wave propagating along the x-axis. If the wave front inclines then a small y-component of the wavenumber, k_y, appears. We assume that this wave front distortion is produced by the change of the refractive index N by a small variation δN which is proportional to the square of

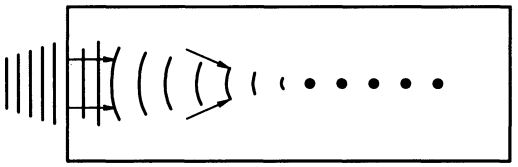

Fig. 2.12. Optical self-focusing.

the amplitude: $\delta N = qa^2N$. Here $q = \text{const}$ and a is the wave amplitude.

Now we have to find an equation for the wave vector k and the amplitude. The first equation can be obtained with the help of very simple arguments. If the wave is not very much different from a plane wave it can be described by the function $a\exp(-i\varphi)$, where φ is its phase. We have $\omega = \partial\varphi/\partial t$, $\mathbf{k} = -\nabla\varphi$. Thus, from these relations we find:

$$\frac{\partial \mathbf{k}}{\partial t} = -\frac{\partial}{\partial t}\nabla\varphi = -\nabla\frac{\partial\varphi}{\partial t} = -\nabla\omega. \tag{2.65}$$

It is sufficient for us to use only the transverse y-component of this equation:

$$\frac{\partial k_y}{\partial t} = -\frac{\partial\omega}{\partial y}.$$

With the nonlinearity taken into account the expression for the frequency looks like

$$\omega = kv_p\left(1 - \frac{\delta N}{N}\right) = kv_p - qkv_pa^2$$

where v_p is the phase velocity along the x-axis. This phase velocity is a function of y and, namely, this dependence produces the wave front curvature (see Fig. 2.12). Remembering that $\partial\omega/\partial k = v_g$ we can write the equation for k_y in the form:

$$\frac{\partial k_y}{\partial t} + v_g\frac{\partial k_x}{\partial y} = k_0v_pq\frac{\partial a^2}{\partial y}. \tag{2.66}$$

Here k_0 is the unperturbed value of the wave number. Now using the relation $\partial k_x/\partial y = \partial k_y/\partial x$ which follows from the fact that the vector field $\mathbf{k} = -\nabla\varphi$ is vortexless we represent (2.66) in the more convenient form

$$\frac{\partial k_y}{\partial t} + v_g \frac{\partial k_y}{\partial x} = k_0 v_p q \frac{\partial a^2}{\partial y}. \tag{2.67}$$

This equation has to be supplemented by an equation for a^2. If we remember that a^2 is proportional to the density of energy which is transported with the group velocity we can write:

$$\frac{\partial a^2}{\partial t} + v_g \frac{\partial a^2}{\partial x} + \frac{\partial}{\partial y}\left(\frac{k_y}{k_0} v_g a^2\right) = 0. \tag{2.68}$$

We have taken into account here the presence of a small transverse component $k_y v_g/k_0$ of the group velocity, where v_g is its unperturbed value.

In the stationary case we can omit the time derivatives in equations (2.67)–(2.68). Excluding k_y with the help of (2.68) and differentiating (2.68) with respect to x we obtain:

$$\frac{\partial^2 a^2}{\partial x^2} = -q \frac{v_p}{v_g} \frac{\partial^2 a^2}{\partial y^2}. \tag{2.69}$$

Thus, for $q > 0$ the light beam is unstable with respect to small initial perturbations dependent upon y. Namely, as a result of this instability, self-focusing develops.

The question may be raised of why this nonlinear effect is not accommodated by the higher harmonics generation as took place for the wave overturn. The answer is that the overturn can be prevented by a very small dispersion which does not much change the sinusoidal shape of the wave pattern.

2.3.3. Wave packet self-compression — envelope instability

Compression of the wave energy can occur not only in the transverse direction but in the longitudinal one as well. To

describe this effect we assume that the amplitude and phase of this packet are slowly varying functions of the variables x, t. We can put again

$$\omega = \omega_k + \alpha a^2, \tag{2.70}$$

where ω_k is the linear frequency of the wave with wavenumber k and αa^2 is the nonlinear correction proportional to the square of the amplitude. According to (2.65) we have

$$\frac{\partial k}{\partial t} = -\frac{\partial \omega}{\partial x} = -v_g \frac{\partial k}{\partial x} - \alpha \frac{\partial a^2}{\partial x}. \tag{2.71}$$

Here it is taken into account that $v_g = \partial \omega_k / \partial k$. The wave energy transfer is described by the relation

$$\frac{\partial a^2}{\partial t} + \frac{\partial}{\partial x}\left(v_g a^2\right) = 0. \tag{2.72}$$

These equations describe the envelope instability when a packet can be broken by itself into separate clots.

To reveal this instability we assume that the envelope of the initially pure harmonic wave is perturbed in such a way:

$$\begin{aligned} k &= k_0 + k' \exp(-i\nu t + i\kappa x), \\ a &= a_0 + a' \exp(-i\nu t + i\kappa x). \end{aligned} \tag{2.73}$$

Here $\nu \ll \omega$ and $\kappa \ll k_0$ are the frequency and the wavenumber of the small envelope perturbation. Assuming that $k' \ll k_0, a' \ll a_0$ we can use a linear approximation for equations (2.71)–(2.72). Then we obtain:

$$-\nu k' + v_g \kappa k' + \alpha \kappa 2 a' a_0 = 0, \tag{2.74}$$

$$\kappa v_g' k' - \nu 2 a_0 a' + \kappa v_g 2 a_0 a' = 0. \tag{2.75}$$

Here $v_g' = \partial v_g / \partial k = \partial^2 \omega_k / \partial k^2$. The dispersion relation for equations (2.79)–(2.80) gives

$$\nu = v_g \kappa \pm \sqrt{\alpha v_g' a_0^2}\, \kappa. \tag{2.76}$$

As we see there are two solutions for ν. When $\alpha = 0$ both solutions merge into one, $\nu = v_g\kappa$, and correspond to wave packet propagation with the group velocity. At $\alpha v_g' > 0$ we have two types of envelope waves with different relations between the amplitude and phase modulations. But if $\alpha v_g' < 0$ the wave packet is unstable with respect to the envelope instability. This instability was discovered by M. Lighthill [14] and the condition $\alpha v_g' < 0$ is called the Lighthill criterion.

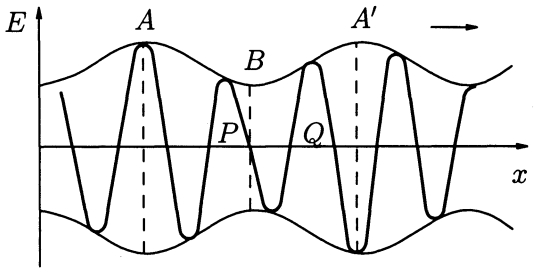

Fig. 2.13. Wave packet selfcompression — envelope instability.

The physics of the envelope instability is clarified with the help of Fig. 2.13. Suppose that $\alpha > 0$. Then the phase velocity at point B is less than at points A and A', so the wavenumber at the spot P will increase since it is proportional to the number of wave nodes per unit space interval, while at the spot Q it will decrease. Hence, provided $v_g' < 0$, the wave packet at P will be decelerated and its power will be added to the wave at spot A while at spot Q it will be accelerated and will amplify the wave at spot A'.

Gravitational waves on deep water do meet the condition $\alpha v_g' < 0$: the nonlinear contribution to the phase velocity (and, consequently, to the frequency) appears to be positive whereas the derivative $v_g' = -(1/4)\sqrt{g/k^3}$ is negative. At first the conclusion that periodic gravitational surface waves must be unstable caused a sensation and seemed wrong. But in time it was rederived by different methods and was confirmed experimentally [15]. Now nobody doubts that gravitational waves on

deep water are unstable and so there are physical grounds for the belief in "the tenth wave."

2.4. Nonlinear wave interaction

2.4.1. Three-wave processes

Let us consider three-dimensional nonlinear oscillations. Certainly we do not expect to succeed in the most general case for the mathematical theory of nonlinear partial differential equations is not complete yet. Nevertheless we may assume the amplitude of oscillations to be not large and use the perturbation theory. In the leading approximation we have a linear theory and the superposition principle, so any perturbation of the steady state may be represented as a set of eigenmodes.

In the framework of the linear approximation it does not matter which physical quantity (the electric field, particle velocities, density etc.) is chosen for the description of the oscillations since all of them are proportional to each other. So we shall use a scalar function $\psi(\mathbf{r}, t)$ that is linear in all other characteristics of the wave and for a while shall not specify its physical meaning. In the linear approximation the function ψ behaves as a collection of plane waves:

$$\psi(\mathbf{r}, t) = \sum \left[a_{\mathbf{k}} e^{-i\omega_{\mathbf{k}} t + i\mathbf{k}\cdot\mathbf{r}} + (\text{c.c.}) \right]. \tag{2.77}$$

Here (c.c.) denotes the complex conjugated quantity and the summation is taken over all the wavenumbers such that $a_{\mathbf{k}} \neq 0$. The frequency $\omega_{\mathbf{k}}$ is the eigenfrequency of the wave with wavenumber $\mathbf{k}$ and is determined by the linear dispersion law. Namely, for any elementary plane wave $\psi_{\mathbf{k}} = a_{\mathbf{k}} \exp(-i\omega t + i\mathbf{k}\cdot\mathbf{r})$ we have

$$(\omega - \omega_{\mathbf{k}}) a_{\mathbf{k}} = 0. \tag{2.78}$$

We assumed the root $\omega = \omega_{\mathbf{k}}$ of the dispersion equation to be simple and that is why the dispersion relation can be represented in the form (2.78) for all values of the frequency ω within a

small vicinity of $\omega_{\mathbf{k}}$. In accordance with (2.78) $\omega = \omega_{\mathbf{k}}$ provided $a_{\mathbf{k}} \neq 0$.

Note that the terms on the right-hand side of (2.77) appear in complex conjugated pairs so that ψ should be real. In other words every elementary wave is represented by the sum

$$a_{\mathbf{k}} \exp(-i\omega_{\mathbf{k}} + i\mathbf{k} \cdot \mathbf{r}) + a_{\mathbf{k}}^* \exp(i\omega_{\mathbf{k}} - i\mathbf{k} \cdot \mathbf{r}),$$

where the star denotes complex conjugation. As we see, the second term corresponds to the wavenumber $-\mathbf{k}$, so

$$a_{-\mathbf{k}} = a_{\mathbf{k}}^*, \quad \omega_{-\mathbf{k}} = -\omega_{\mathbf{k}}. \tag{2.79}$$

Thus, every plane wave can be characterized either by the pair $(\mathbf{k}, \omega_{\mathbf{k}})$ or by $(-\mathbf{k}, \omega_{-\mathbf{k}} = -\omega_{\mathbf{k}})$, while only the phase velocity $\boldsymbol{v}_p = \mathbf{k}\omega_{\mathbf{k}}/k^2$ appears to be invariant. It is convenient to choose the pair with the positive frequency (if the wave energy is positive). Then the directions of the wavevector $\mathbf{k}$ and the phase velocity will coincide and one may simply speak of a plane wave $(\mathbf{k}, \omega_{\mathbf{k}})$ of amplitude $a_{\mathbf{k}}$. So far the quantity $a_{\mathbf{k}}$ is unspecified since we assumed ψ to be an arbitrary linear function of the linear wave amplitude. It turns out that all the needed formulae will simplify significantly if we choose $a_{\mathbf{k}}$ such that the wave energy density $\mathcal{E}_{\mathbf{k}}$ is expressed in the form $\mathcal{E}_{\mathbf{k}} = \omega_{\mathbf{k}}|a_{\mathbf{k}}|^2 = \omega_{\mathbf{k}} a_{\mathbf{k}} a_{\mathbf{k}}^*$. We shall see in Section 5 that the momentum density $P_{\mathbf{k}}$ of a plane wave is proportional to its energy density $\mathcal{E}_{\mathbf{k}}$ and is connected to it by the simple formula $P_{\mathbf{k}} = \mathcal{E}/v_p$, with v_p being the wave's phase velocity. The momentum density $P_{\mathbf{k}}$ is directed along the phase velocity, so the formula $\mathbf{P}_{\mathbf{k}} = \boldsymbol{v}_p/v_p^2 \mathcal{E}_{\mathbf{k}}$ is more precise.

At this moment it is useful to draw a quantum analogy. A classical linear wave may be regarded as a collection of a large amount of a free field quanta that appear to be Bose-particles. Let $N_{\mathbf{k}}$ be the number of identical particles placed in a unit volume. Then their energy and momentum can be represented respectively in the form $\hbar\omega_{\mathbf{k}} N_{\mathbf{k}}$ and $\hbar\mathbf{k} N_{\mathbf{k}}$. Certainly we do not need the Plank constant while considering classical fields, so

henceforth we shall omit it assuming that $\hbar = 1$ in the appropriate system of units. Correspondingly, we shall use the classical quantity $N_{\mathbf{k}} = a_{\mathbf{k}} a_{\mathbf{k}}^*$ that differs from the number of quanta by the factor $\hbar^{-1}$.

We can summarize that with the forgoing choice of the amplitude $a_{\mathbf{k}}$ the following expressions are valid:

$$\mathcal{E}_{\mathbf{k}} = \omega_{\mathbf{k}} N_{\mathbf{k}}, \quad \mathbf{P}_{\mathbf{k}} = \mathbf{k} N_{\mathbf{k}}, \quad N_{\mathbf{k}} = a_{\mathbf{k}}^* a_{\mathbf{k}}. \tag{2.80}$$

It was assumed here that the wave energy is positive and, consequently, $\omega_{\mathbf{k}} > 0$. In a stable stationary medium waves really do have positive energy. But a plasma is usually far from equilibrium. It makes sense to choose $\omega_{\mathbf{k}}$ negative if the wave energy is negative since $N_{\mathbf{k}} > 0$. Therefore one may ask whether the expression for the momentum in (2.80) is correct in the case of a negative frequency. As will be shown in Section 5, the relation $\mathcal{E}_{\mathbf{k}} = \boldsymbol{v}_p \cdot \mathbf{P}_{\mathbf{k}}$ is universal, and by definition the phase velocity $\boldsymbol{v}_p$ is equal to $\mathbf{k}\omega_{\mathbf{k}}/k^2$. Consequently, the formula $\mathbf{P}_{\mathbf{k}} = \mathbf{k} N_{\mathbf{k}}$ is also valid for waves of negative energy since it implies $\boldsymbol{v}_p \cdot \mathbf{P}_{\mathbf{k}} = \omega_{\mathbf{k}} N_{\mathbf{k}} = \mathcal{E}_{\mathbf{k}}$. We see that $\boldsymbol{v}_p \cdot \mathbf{P}_{\mathbf{k}} = \mathcal{E}_{\mathbf{k}} < 0$ for a negative-energy wave, i.e., the momentum $\mathbf{P}_{\mathbf{k}}$ and the phase velocity $\boldsymbol{v}_p$ are oppositely directed. Thus, we should use the pair $(\omega_{\mathbf{k}}, \mathbf{k})$ with negative $\omega_{\mathbf{k}}$ for the description of waves with negative energy. For these waves the elementary portion $\mathbf{k}$ of the momentum (with $\hbar = 1$) and the total momentum $\mathbf{P}_{\mathbf{k}} = \mathbf{k} N_{\mathbf{k}}$ appear to point in the direction opposite to that of the phase velocity.

The foregoing immediately brings up the question about which option should be taken: $(\mathbf{k}, \omega_{\mathbf{k}})$ or $(-\mathbf{k}, -\omega_{\mathbf{k}})$? Recall that these two pairs according to (2.79) are exactly equivalent! The answer is provided by relation (2.80). Since $N_{\mathbf{k}}$ was defined as the number of quanta, the direction of the vector $\mathbf{k}$ is to be chosen in such a way that $\mathbf{k} N_{\mathbf{k}}$ should coincide with the wave momentum density. Then the expression $\omega_{\mathbf{k}} N_{\mathbf{k}}$ will give the energy density with the proper sign corresponding to the energy sign of the wave under consideration.

Let us return to expressions (2.77)–(2.78). The superposition (2.77) of eigenmodes appears, in fact, to be a solution to the dispersion equation (2.78), since we forced the frequency ω of an elementary wave $\psi_{\mathbf{k}}$ to be equal to the value $\omega_{\mathbf{k}}$ determined by the dispersion relation. Equation (2.78) may be regarded as an extremely simplified version of the primary system of linear equations: we converted the latter to a simple relation in the neighborhood of the frequency eigenvalue. Relation (2.78) can be also regarded as the simplified wave equation in the $(\mathbf{k}, \omega)$-representation. To return to the $(\mathbf{r}, t)$-representation one is to treat the function $\psi(\mathbf{r}, t)$ as a superposition of plane waves of the type $\psi_{\mathbf{k}} = a_{\mathbf{k}} \exp(-i\omega t + \mathbf{k} \cdot \mathbf{r})$. Replacing ω and $\mathbf{k}$ on the right-hand side of expression (2.78) by the operators $i\partial/\partial t$ and $-i\nabla$, we transform it to the form

$$i\frac{\partial \psi}{\partial t} = \hat{\omega}_{\mathbf{k}} \psi, \tag{2.81}$$

where the operator $\hat{\omega}_{\mathbf{k}}$ is obtained by the substitution $\mathbf{k} \to -i\nabla$ in the function $\omega_{\mathbf{k}}(\mathbf{k})$.

For instance, the frequency $\omega_{\mathbf{k}}$ of a free quantum particle with mass m is equal to $\mathcal{E}_{\mathbf{k}}/\hbar = \hbar|\mathbf{k}|^2/2m$ and equation (2.81) assumes the form

$$i\hbar\frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi,$$

i.e., coincides with the Schrödinger equation for a free particle. If ψ corresponds to one of the equivalent Bose-particles, the squared amplitude corresponds to the number of particles per unit volume and the classical formulae (2.80) for the energy and momentum densities seem natural from the quantum mechanics viewpoint.

Let us turn to slightly nonlinear waves. If the wave amplitudes are not infinitesimally small, we are to take the nonlinear terms of the equations of motion into account. It makes sense to arrange them in ascending order with respect to the powers of amplitude, i.e., to collect the quadratic, cubic, etc. terms in

distinct blocks. If the amplitude is extremely small, it is sufficient to leave only the quadratic terms and to neglect the others. Expressing all the physical variables in terms of the previously chosen one, we can reduce all the equations to the form (2.78), where the sum of quadratic terms (and of cubic, fourth-power and all other terms if they have not been neglected) will appear on the right-hand side. The quadratic terms are proportional to $a_{\mathbf{k}'}a_{\mathbf{k}''}$ and the summation is apparently to be taken over the indices $\mathbf{k}', \mathbf{k}''$ satisfying the restriction $\mathbf{k}' + \mathbf{k}'' = \mathbf{k}$ in order to provide the same dependence $\exp(i\mathbf{k} \cdot \mathbf{r})$ on the spatial coordinate as that of the left-hand side.

One may treat the frequency on the left-hand side of (2.78) as the operator $i\partial/\partial t$. Therefore in the case of the elementary plane wave $a_{\mathbf{k}} \exp(-i\omega t + i\mathbf{k}\cdot\mathbf{r})$ the bracketed expression $(\omega - \omega_{\mathbf{k}})$ transforms as follows:

$$(\omega - \omega_{\mathbf{k}})a_{\mathbf{k}} \exp\Big(-i\omega_{\mathbf{k}}t + i\mathbf{k} \cdot \mathbf{r}\Big) = i\frac{\partial a_{\mathbf{k}}}{\partial t} \exp\Big(-i\omega_{\mathbf{k}}t + i\mathbf{k} \cdot \mathbf{r}\Big).$$

Thus, if quadratic terms are taken into account, equation (2.78) assumes the form

$$\frac{\partial a_{\mathbf{k}}}{\partial t} = \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'\mathbf{k}''}\, a_{\mathbf{k}'}a_{\mathbf{k}''} \exp\Big[-i(\omega_{\mathbf{k}'} + \omega_{\mathbf{k}''} - \omega_{\mathbf{k}})t\Big] \qquad (2.82)$$

with $\mathbf{k}'' = \mathbf{k} - \mathbf{k}'$. The time-dependent factor arose from the insertion of the superposition of eigenmodes into the nonlinear terms. $V_{\mathbf{k}\mathbf{k}'\mathbf{k}''}$ is a certain function of the wavevectors and is uniquely determined by the nonlinear terms of the equations of motion. It is natural to refer to $V_{\mathbf{k}\mathbf{k}'\mathbf{k}''}$ as the matrix element of the wave interaction.

We see that the quadratic terms

$$a_{\mathbf{k}'}a_{\mathbf{k}''} \exp\Big[-i(\omega_{\mathbf{k}'} + \omega_{\mathbf{k}''})t + i(\mathbf{k}' + \mathbf{k}'') \cdot \mathbf{r}\Big]$$

that contribute to the nonlinear external current play the role of a power source with frequency $\omega_{\mathbf{k}'} + \omega_{\mathbf{k}''}$ and wavevector $\mathbf{k}' + \mathbf{k}''$. Note that this power source is inevitably accompanied by a similar source with frequency $\omega_{\mathbf{k}'} - \omega_{\mathbf{k}''}$ and wavevector

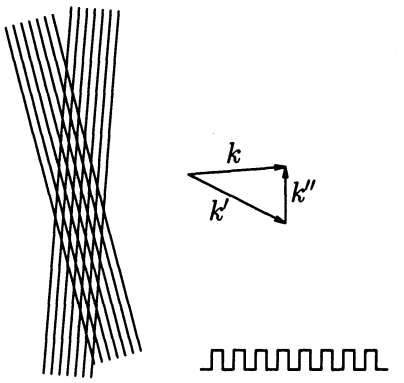

Fig. 2.14. A moire as the result of a nonlinear superposition of two plane waves.

$\mathbf{k}' - \mathbf{k}''$ since different terms in (2.77) are to be paired up with the complex-conjugated ones because ψ is to be real.

One can visualize the "pulsation" at the combinative frequencies $\omega_{\mathbf{k}'} \pm \omega_{\mathbf{k}''}$ and wavevectors $\mathbf{k}' \pm \mathbf{k}''$ with the help of Fig. 2.14, which presents the two plane wave superposition pattern. Each hatching in Fig. 2.14 resembles a meander-wave whose amplitude hops from zero to unity: it equals to zero on the black lines and unity between them (see the bottom of Fig. 2.14). The superposition takes place in the intersection area — the black lines run through the points where at least one of the amplitudes is nonzero with white spots between them. One can clearly see through the figure that the wave superposition forms a moire — a wave with wavevector $\mathbf{k}' - \mathbf{k}''$. (In addition a small-scale ripple appears at the wavevector $\mathbf{k}' + \mathbf{k}''$).

It is even more easy to get a view of a moire with the help of two combs. Put one of them on the other and look through. If the combs are identical, a moire will appear only under the condition that the combs are placed at a nonzero angle as is shown in Fig. 2.14, but if they are different, a moire will appear even if one of them is put along the other. Shift the combs a little and watch the moire moving. A similar picture can be also seen in nylon folds.

The moire effect is provided by the wave superposition and moving waves produce a moving moire. If this moving moire coincides with an eigenmode, i.e., if the combinative frequency $\omega_{\mathbf{k}'} \pm \omega_{\mathbf{k}''}$ is equal to the eigenfrequency $\omega_{\mathbf{k}}$ corresponding to the pulse wavevector $\mathbf{k} = \mathbf{k}' \pm \mathbf{k}''$, the nonlinear power source will amplify the $\mathbf{k}$-wave and its amplitude $a_{\mathbf{k}}$ will change with time according to (2.82).

In the resonant case the difference $\Delta = \omega_{\mathbf{k}'} + \omega_{\mathbf{k}''} - \omega_{\mathbf{k}}$ is zero and in time the oscillation amplitudes can change drastically even when the nonlinearity is small. The interaction processes involving three waves are known as three-wave interactions and the dispersion laws (the dependence of the frequency $\omega_{\mathbf{k}}$ on $\mathbf{k}$) satisfying the decay conditions

$$\mathbf{k} = \mathbf{k}' + \mathbf{k}'', \quad \omega_{\mathbf{k}} = \omega_{\mathbf{k}'} + \omega_{\mathbf{k}''}, \tag{2.83}$$

are called decay spectrums. In a plasma the decay constraints are easily settled because of the great number of different oscillation modes.

Three-wave processes provide an opportunity for eigenwave transformation and, therefore, lead to a complicated picture of energy transfer in the wavenumber phase space. In particular, these triad processes take place in certain hydrodynamic systems [16].

Let us consider the simplest case when there are only three waves satisfying the resonance condition (2.83). For the following we shall denote their amplitudes by $a_1 = a_{\mathbf{k}}$, $a_2 = a_{\mathbf{k}'}$, $a_3 = a_{\mathbf{k}''}$ and the corresponding frequencies by ω_1, ω_2, ω_3, and shall assume that $\omega_1 = \omega_2 + \omega_3 > \omega_2 > \omega_3$.

With the aid of (2.82) we obtain the equation

$$\frac{\partial a_1}{\partial t} = V_1 a_2 a_3 \tag{2.84}$$

for a_1 with $V_1 = V_{\mathbf{k},\mathbf{k}',\mathbf{k}''}$. The complementary equations for a_2 and a_3 may be obtained in the same way. Interchanging $\mathbf{k}$ and $\mathbf{k}'$ in (2.82), we come to

$$\frac{\partial a_2}{\partial t} = V_2 a_1 a_3^*, \tag{2.85}$$

and substituting $\mathbf{k}''$ for $\mathbf{k}$ and $\mathbf{k}$ for $\mathbf{k}'$ we obtain

$$\frac{\partial a_3}{\partial t} = V_3\, a_1 a_3^* \tag{2.86}$$

with $V_2 = V_{\mathbf{k}',\mathbf{k},-\mathbf{k}''}$ and $V_3 = V_{\mathbf{k}'',\mathbf{k},-\mathbf{k}'}$.

In the framework of hydrodynamics when thermal motions are neglected, equations (2.82)–(2.86) appear to be the Fourier transform of the partial differential equations of motion and the coefficients $V_{\mathbf{k},\mathbf{k}',\mathbf{k}''}$ all together are either real or purely imaginary. For simplicity we shall assume them to be real (if they were imaginary one could multiply the amplitudes a_i by i and return to the case under consideration). Equations (2.82)–(2.86) are given in the dynamic variables $a_{\mathbf{k}}$ and we can rescale each by an appropriate factor to simplify them. Choose again this normalization so that the energy $\mathcal{E}_{\mathbf{k}}$ of the $\mathbf{k}$-th wave is $\omega_{\mathbf{k}}|a_{\mathbf{k}}|^2$. Then its momentum $\mathbf{P}_{\mathbf{k}}$ will be equal to $\mathbf{k}|a_{\mathbf{k}}|^2$ and the quantity $N_{\mathbf{k}} = |a_{\mathbf{k}}|^2$ can be interpreted as the number of waves.

Because of the energy and momentum conservation laws the matrix elements involved in equations (2.84)–(2.86) must meet certain restrictions. Really, multiplying them correspondingly by $\omega_i a_i^*$ and $\mathbf{k}_i a_i^*$ ($\mathbf{k}_1 = \mathbf{k}$, $\mathbf{k}_2 = \mathbf{k}'$, $\mathbf{k}_3 = \mathbf{k}''$) and adding to the complex conjugated equations, we obtain

$$\begin{aligned} \omega_2(V_1 + V_2) + \omega_3(V_1 + V_3) &= 0, \\ \mathbf{k}_2(V_1 + V_2) + \mathbf{k}_3(V_1 + V_3) &= 0 \end{aligned} \tag{2.87}$$

in view of the energy and momentum conservation laws

$$\omega_1 N_1 + \omega_2 N_2 + \omega_3 N_3 = \text{const}, \tag{2.88}$$

$$\mathbf{k}_1 N_1 + \mathbf{k}_2 N_2 + \mathbf{k}_3 N_3 = \text{const}, \tag{2.89}$$

and of the relations $\omega_1 = \omega_2 + \omega_3$, $\mathbf{k}_1 = \mathbf{k}_2 + \mathbf{k}_3$. Formulae (2.87) represent a system of four scalar equations in two variables $V_1 + V_2$, $V_1 + V_3$. There is only the zero solution to an overdetermined system, so $V_1 = -V_2 = -V_3$. Moreover, since the sign reversal before all the wavevectors is equivalent to complex conjugation and we assumed all the V_i to be real, the rela-

tion $V_{\mathbf{k},\mathbf{k}',\mathbf{k}''} = V_{-\mathbf{k},-\mathbf{k}',-\mathbf{k}''}$ is valid. Equation (2.82) also implies that $V_{\mathbf{k},\mathbf{k}',\mathbf{k}''} = V_{\mathbf{k},\mathbf{k}'',\mathbf{k}'}$. Thus, we have obtained the following relations representing the symmetries of the matrix elements:

$$V_{\mathbf{k},\mathbf{k}',\mathbf{k}''} = V_{\mathbf{k},\mathbf{k}'',\mathbf{k}'} = V_{-\mathbf{k},-\mathbf{k}',-\mathbf{k}''} = -V_{\mathbf{k}',\mathbf{k},-\mathbf{k}''} = -V_{\mathbf{k}'',\mathbf{k},-\mathbf{k}'}. \tag{2.90}$$

In view of these properties the equations for the amplitudes of the interacting waves take the simple form

$$\frac{\partial a_1}{\partial t} = V a_2 a_3, \quad \frac{\partial a_2}{\partial t} = -V a_1 a_3^*, \quad \frac{\partial a_3}{\partial t} = -V a_1 a_3^*. \tag{2.91}$$

Now it is easy to find a new integral of motion. Multiply the first equation by a_1^*, the second by a_2^*, sum up and add to the complex conjugated relations. Obviously the right-hand side will be equal to zero and hence

$$N_1 + N_2 = \text{const}. \tag{2.92}$$

In the same manner one can deduce that

$$N_1 + N_3 = \text{const}. \tag{2.93}$$

But the last relation appears to be the consequence of (2.88) and (2.92) because $\omega_1 = \omega_2 + \omega_3$. The integrals (2.92)–(2.93) are known as the Manley–Rowe conservation laws. Their physical sense can be easily displayed in the framework of quantum mechanics. Suppose, for instance, that there is a one-quantum decrease of N_1. The Manley–Rowe relations imply that an additional quantum is to appear in each of the states 2 and 3, i.e., the quanta number deviations δN_i satisfy the constrains $\delta N_2 = -\delta N_1$, $\delta N_3 = -\delta N_1$. Moreover, the energy and momentum conservation laws appear to be fulfilled:

$$\hbar\mathbf{k}_1 = \hbar\mathbf{k}_2 + \hbar\mathbf{k}_3, \quad \hbar\omega_1 = \hbar\omega_2 + \hbar\omega_3. \tag{2.94}$$

This ensures that we are really dealing with the decay of the quantum $(\mathbf{k}_1, \omega_1)$ into the sum of quanta $(\mathbf{k}_2, \omega_2)$ and $(\mathbf{k}_3, \omega_3)$.

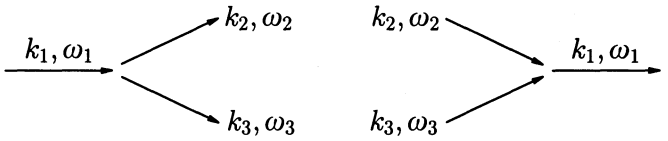

Fig. 2.15. Three-wave processes.

Relation (2.94) can be read in the opposite direction. Then it states that the quanta $(\mathbf{k}_2, \omega_2)$ and $(\mathbf{k}_3, \omega_3)$ can annihilate giving birth to the quantum $(\mathbf{k}_1, \omega_1)$. In other words the three-wave processes are formed by the decays of a single quantum into a pair and by the backward processes of pair merging (see Fig. 2.15).

Thus, we have obtained two integrals of motion, (2.88) and (2.92), that impose significant restrictions on the wave interaction. To clarify the essence of these constrains suppose that all the amplitudes a_i are real. Then relations (2.88) and (2.92) can be rewritten in the form

$$\omega_1 a_1^2 + \omega_2 a_2^2 + \omega_3 a_3^2 = \text{const}, \quad a_1^2 + a_2^2 = \text{const}. \tag{2.95}$$

The first represents an ellipsoid of half-diameters $\omega_1^{-1/2}$, $\omega_2^{-1/2}$, $\omega_3^{-1/2}$ in the Cartesian coordinates (a_1, a_2, a_3) and the other corresponds to a cylinder with the axis directed along a_3. The amplitudes a_i can evolve only in such a way that the point (a_1, a_2, a_3) should move along the intersection line of these surfaces. In our case of $\omega_1 > \omega_2 > \omega_3$ these lines are of the form shown in Fig. 2.16.

As we see, the trajectories in the neighborhood of the axes a_2 and a_3 are like small ellipses, i.e., each amplitude a_2, a_3 oscillates near its initial value provided the imposed perturbation is small (Fig. 2.17a). On the contrary the wave of the maximum frequency ω_1 can "decay," i.e., its energy can be completely transferred to the waves a_2, a_3 even if their initial amplitudes were small: $a_2, a_3 \ll a_1$ (Fig. 2.17b). Figures 2.16 and 2.17b show that the decay process is periodic: at first the wave a_1 is

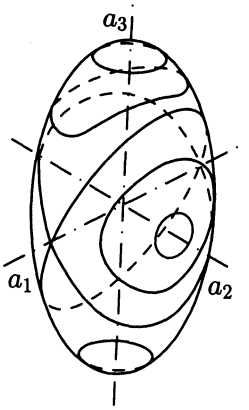

Fig. 2.16. Trajectories of three interacting waves.

converted into the waves a_2, a_3 and then the system returns to the initial state.

2.4.2. Analogy to the Euler equations

The above process is similar to free solid body motion. In particular, the decay of wave a_1 into a_2, a_3 is similar to the rotation of a free asymmetric body around the axis of intermediate angular moment of inertia I_2, with $I_1 > I_2 > I_3$ being its angular moments of inertia (see Fig. 51 in [17]). The governing equations are also alike.

Recall that the angular momentum $\mathbf{M}$ of a free body remains constant and satisfies the equations

$$\frac{d\mathbf{M}}{dt} = \frac{d'\mathbf{M}}{dt} + \boldsymbol{\Omega} \times \mathbf{M} = 0. \tag{2.96}$$

Here the primed differentiation in time corresponds to the differentiation of the $\mathbf{M}$-components related to the Cartesian coordinate system frozen to the body, while the second term arose due to the differentiation of the corresponding moving orthogonal unit bench-vectors. Since $\mathbf{M} = \{I_1\Omega_1, I_2\Omega_2, I_3\Omega_3\}$ in the moving coordinates, the angular velocity components in accordance

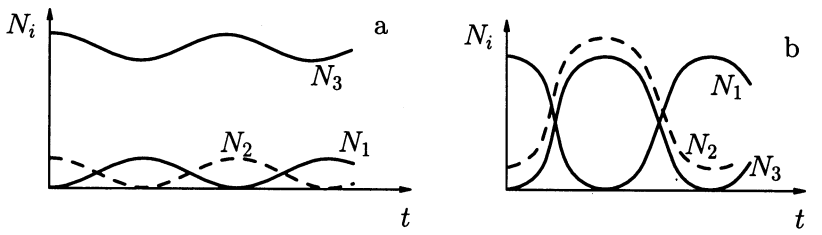

Fig. 2.17. Dependence of amplitudes of three interacting waves upon time in the case of stable (a) and unstable (b) initial states.

with (2.96) satisfy the equations

$$\frac{d\Omega_1}{dt} = \frac{I_2 - I_3}{I_1}\Omega_2\Omega_3, \tag{2.97}$$

$$\frac{d\Omega_2}{dt} = \frac{I_3 - I_1}{I_2}\Omega_3\Omega_1, \tag{2.98}$$

$$\frac{d\Omega_3}{dt} = \frac{I_1 - I_2}{I_3}\Omega_1\Omega_2. \tag{2.99}$$

These equations are similar to equations (2.91) for the amplitudes a_i. The integrals $2E = M_1^2/I_1 + M_2^2/I_2 + M_3^2/I_3$ and $M^2 = M_1^2 + M_2^2 + M_3^2$ are conserved and the intersections of their level surfaces in the Cartesian coordinate system $(\Omega_1, \Omega_2, \Omega_3)$ provide the same picture as the wave decay: the rotations around the axes of extreme angular moments of inertia I_1, I_3 are stable, while the rotation about the axis of intermediate angular moment of inertia I_2 is unstable. The dependence of the angular velocity components Ω_i and of the amplitudes a_i upon time also appears to be described by similar formulae.

2.4.3. Explosive instability

So far, while considering the three-wave interaction we assumed the energy of all the waves to be positive. The above symmetries of the interaction matrix elements were derived on this assumption. However, in a non-steady plasma the wave

energy can be negative (see Section 5.4.3) and the interaction of such waves with waves of positive energy leads to a nonlinear instability of "explosive" type [18]. The cause is that the energy transfer from a wave of negative energy to waves of positive energy leads to an increase the negative-energy wave amplitude, but not to a decrease.

Suppose, for instance, that wave a_1 of the highest frequency is also of negative energy. Since we have agreed to preserve the formulae $\mathcal{E}_{\mathbf{k}} = \omega_{\mathbf{k}} N_{\mathbf{k}}$, $\mathbf{P}_{\mathbf{k}} = \mathbf{k} N_{\mathbf{k}}$ for the energy and momentum of waves of negative energy but to choose $(\omega_{\mathbf{k}}, \mathbf{k})$ as the pair with negative frequency and wavevector directed against the phase velocity $\boldsymbol{v}_p = \mathbf{k}\omega_{\mathbf{k}}/k^2$, the resonance conditions are now to be read as

$$\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0, \quad \omega_1 + \omega_2 + \omega_3 = 0. \tag{2.100}$$

Energy and momentum conservation imply that $V_1 = V_2 = V_3$ and hence the equations for the amplitudes of the three interacting waves take the form

$$\frac{\partial a_1}{\partial t} = V\, a_2 a_3, \quad \frac{\partial a_2}{\partial t} = V\, a_1 a_3^*, \quad \frac{\partial a_3}{\partial t} = V\, a_1 a_2^*. \tag{2.101}$$

One may question the difference between formulae (2.83) and (2.100). As we know, any plane wave can be characterized by the pairs $(\mathbf{k}, \omega_{\mathbf{k}})$ and $(-\mathbf{k}, -\omega_{-\mathbf{k}})$ on equal grounds. Relations (2.83) and (2.100) do formally coincide up to this interchange. The answer appears to be not formally mathematical, but rather physical. Namely, it is the structure of equations (2.101) that gives evidence of the opposite energy signs of the first and other waves. Indeed, the Manley–Rowe relations assume the form $N_1 - N_2 = \text{const}$, $N_1 - N_3 = \text{const}$, $N_2 - N_3 = \text{const}$. In other words the addition of a single quantum to each state does not violate these constraints and is in compliance with the energy conservation law only under the condition that the sum of the quanta energies is zero.

To simplify the analysis of equations (2.101), let the amplitudes be real. Then it is clearly seen that all of them grow

simultaneously. In view of the similarity of these equations to the equation $dx/dt = x^2$ one can expect their solutions to be analogous to the solution $x = (t_0 - t)^{-1}$, i.e., that $a_i \sim 1/(t_0 - t)$. Thus, we have encountered an instability of "explosive" nature: the waves amplitudes grow to infinite values in a finite time.

2.4.4. Random-phase approximation

It is clear from the equations governing the time evolution of three interacting waves that the intrinsic time scale of their change is $t \sim 1/Va$. This scale is large if the amplitudes are small and the synchronization conditions are to be supported during all this time in order for the resonance interaction to occur. Apparently the lower the amplitude, the lower the acceptable divergence Δ, i.e., the synchronization constraint $\Delta = \omega_1 - \omega_2 - \omega_3 \ll Va$ becomes more severe.

But in real experiments on a plasma it is very difficult to reach exact synchronization since there are always small disturbances that slightly vary the eigenfrequencies in space (or in time, if these disturbances are non-steady). Under these circumstances the wave interaction becomes much more complicated, but on the other hand its irregularity permits one to use another method involving the so-called random-phase approximation. We shall apply it to the problem of three wave interaction, assuming the energies to be positive. Multiplying each equation for the amplitudes a_j by the corresponding a_j^* and adding to the complex-conjugated relations, we obtain the equations in squared amplitudes:

$$\frac{\partial N_1}{\partial t} = \frac{\partial}{\partial t}|a_1|^2 = V\, a_1^* a_2 a_3 + (\text{c.c.}), \tag{2.102}$$

$$\frac{\partial N_2}{\partial t} = -V\, a_1 a_2^* a_3^* + (\text{c.c.}), \tag{2.103}$$

$$\frac{\partial N_3}{\partial t} = -V a_1 a_2^* a_3^* + (\text{c.c.}), \tag{2.104}$$

where (c.c.) stand for the corresponding complex conjugated quantities. Each of the amplitudes a_j can be represented in the

form $|a_j|^2 \exp(i\varphi_j)$ with φ_j being its phase. Notice that there is the same phase factor $\exp\left[\pm i(\varphi_1 - \varphi_2 - \varphi_3)\right]$ on the right-hand side of all equations (2.102)–(2.104). The phase difference $\varphi_1 - \varphi_2 - \varphi_3$ can vary with time because of the wave interactions (these variations are of the same time scale as the amplitude variations) and due to inhomogeneities of the medium. The latter can cause even more rapid phase variations. Undoubtedly the plasma inhomogeneities do cause amplitude fluctuations as well, but one can expect them to be much less than the phase fluctuations since the squared amplitude is related to the conserved quantity, namely, to the energy density, and hence only a slight spatial redistribution of the wave packet energy can take place. As for the phase, there are no constraints imposed on its variations and it can be easily altered at any spot by a mere shift of a wave packet portion.

Thus, it seems quite reasonable to use the random-phase approximation, assuming all the amplitudes to be constant and the phases to be arbitrary random functions of time. In other words one is to insert $a_j \approx \sqrt{N_j}\exp(i\varphi_j)$ in equations (2.102)–(2.104) and to average over the phases φ_j. But this averaging, if formally applied, will turn the right-hand sides to zero, so while dealing with them we have to take care of small deviations of the amplitudes from their zero-approximation value, i.e., to put

$$a_j = \sqrt{N_j} e^{i\varphi_j} + \delta\tilde{a}_j. \tag{2.105}$$

Here $\delta\tilde{a}_j$ stands for the above deviation and can be obtained with the aid of the first-order approximation

$$\frac{\partial}{\partial t}\delta\tilde{a}_1 = V\sqrt{N_1 N_2}\, e^{i(\varphi_1+\varphi_2)}$$

to the amplitude evolution governing equation. Apparently

$$\delta\tilde{a}_1 = V\sqrt{N_1 N_2}\int_{t_0}^{t} e^{i(\varphi_1+\varphi_2)}dt \tag{2.106}$$

with t_0 being a large negative instant, such that $\delta\tilde{a}_1(t_0)$ can be considered zero. We put $\sqrt{N_1 N_2}$ outside the integral, assuming

N_j to be constant over the time scale of the phase variation. Substituting successively the deviations $\delta\tilde{a}_j$ for the corresponding amplitudes on the right-hand side of equations (2.102)–(2.104), we get the equations for N_j in the leading-order approximation:

$$\frac{\partial N_1}{\partial t} = W\big(N_2N_3 - N_1N_2 - N_1N_3\big), \tag{2.107}$$

$$\frac{\partial N_2}{\partial t} = \frac{\partial N_3}{\partial t} = -\frac{\partial N_1}{\partial t}. \tag{2.108}$$

Here

$$W = 2V^2\tau, \quad \tau = \langle \exp[i\theta(t)] \int_{-\infty}^{t} \exp[-i\theta(t')]\, dt' \rangle,$$

with θ being the sum of any two random phases, say of φ_1 and φ_2: $\theta = \varphi_1 + \varphi_2$. The quantity τ determines the phase correlation scale.

The equations for N in the random-phase approximation imply the Manley–Rowe identities $N_1 + N_2 = \text{const}$, $N_1 + N_3 = \text{const}$ and, consequently, energy and momentum conservation. There is a stationary solution $N_j = \text{const}$ corresponding to the equipartition of the energy, i.e., to the Rayleigh-Jeans law:

$$\mathcal{E}_j = \omega_j N_j = T_0 = \text{const}.$$

Indeed, substituting $N_j = T_0/\omega_j$ in the right-hand side of equation (2.107) and moving the common factor $N_1N_2N_3/T_0$ out of the brackets, we obtain the factor $(\omega_1 - \omega_2 - \omega_3)$ that is apparently equal to zero due to the synchronization condition for the frequencies.

The equations for N_j can be naturally regarded from the quantum mechanical standpoint. The processes of combinative decay of a quantum $\hbar\omega_1$ into the pair $\hbar\omega_2 + \hbar\omega_3$ and the backward merging are described by the equation

$$\frac{\partial N_1}{\partial t} = W\left[N_2N_3(N_1+1) - N_1(N_2+1)(N_3+1)\right]. \tag{2.109}$$

Here W is the transition probability and the units are added to the final states in order to take into account spontaneous transitions of Bose-particles from one state to another. If the numbers N_j of quanta are essentially larger than unity, it is sufficient to keep only quadratic terms on the right-hand side of this equation and it will obviously coincide with (2.107).

2.4.5. Low-frequency wave interaction with high-frequency waves

So far we have considered only three wave interaction. In reality the number of interacting waves can be essentially larger. In the three-wave approximation we treated the superposition of three engaged waves, considering the interaction to be either under the condition of fixed phases provided the resonance conditions are perfectly supported, or under the condition of random phases in the case of a slightly inhomogeneous medium or irregular waves.

Let us now consider the interaction of high-frequency waves with a low-frequency one to offer an illustration of the case of a large number of interacting waves of fixed phases [19]. For instance, a high-frequency electromagnetic wave interacting with ion-sound. This case corresponds to Fig. 2.14 when the moire produced by a high-frequency wave appears to be in resonance with a low-frequency wave.

Denote the low-frequency wave amplitude by b, and its wavevector and frequency by κ and ν. Let the wavenumber of the second wave be k_0, with $k_0 \gg \kappa$. Then this primary wave can interact with a wave of $k_1 = k_0 + \kappa$, while the latter wave can in turn interact with a wave of $k_2 = k_0 + 2\kappa$ and so on. Thus, besides the two primary waves k_0 and κ, the entire set of waves with the wavenumbers $k_n = k_0 + \kappa$, n being any integer, can be engaged in the interaction. Let the amplitude of the wave $k_0 + n\kappa$ be a_n, so that a_0 corresponds to the primary high-frequency wave. We shall assume that the waves are in exact resonance, i.e., that $\omega_n = \omega_{n-1} + \nu$ or $\omega_k = \omega_{k-\kappa} + \nu$.

Now we can put $\omega_{k-\kappa} \approx \omega_k - (\partial\omega/\partial k)\kappa$ since κ was supposed to be small and rewrite the resonance condition in the form

$$\frac{\partial\omega}{\partial k}\kappa = v_g\kappa = \nu. \tag{2.110}$$

It is obvious that in our case of small κ the resonance condition corresponds to the equality of the group velocity of the high-frequency wave packet with the phase velocity of the low-frequency waves. Since κ is small, the matrix element in (2.82) can be treated as if it were independent of n, so the equation for a_n assumes the form

$$\frac{\partial a_n}{\partial t} = iV\left(a_{n+1}b^* + a_{n-1}b\right), \tag{2.111}$$

where V is a certain constant that we shall assume to be real. (For the following it will be convenient to use such a normalization of the amplitudes that the interaction matrix element iV is purely imaginary). By analogy we obtain the equation

$$\frac{\partial b}{\partial t} = iV\sum_n a_n a_{n-1}^* \tag{2.112}$$

for b. The interaction matrix elements in (2.111) and (2.112) were assumed to coincide. This surely can be achieved by an appropriate normalization of the amplitudes a_n and b.

Suppose that initially there was the only a high-frequency wave of amplitude $a_0(0)$ and a weak low-frequency wave of amplitude b. Let us assume for simplicity that b is real. Then one may easily verify with the help of the well-known identity

$$2\frac{dJ_n}{dx} = J_{n-1} - J_{n+1}$$

for the Bessel functions that

$$a_n = i^n a_0(0)\, J_n\left(2V\int b\, dt\right) \tag{2.113}$$

gives the solution to (2.111). This ensures that due to the interaction with the low-frequency wave higher and higher modes a_n are exited and that on average their amplitudes become equal as $t \to \infty$ since $J_n(x) \sim \text{const}/\sqrt{x}$ at large x (provided $n < \sqrt{x}$). In view of the relation

$$\begin{aligned} \psi &= \sum_n a_n \exp(-i\omega_n t + ik_n r) \\ &= \exp(-i\omega_0 t + ik_0 r) \sum_n a_n \exp(-in\nu t + in\kappa r) \\ &= \exp\left[-i\omega_0 t + ik_0 r - i\lambda \cos(\nu t - \kappa r)\right] \end{aligned} \tag{2.114}$$

(with $\lambda = 2V \int b\, dt$) it is obvious that the excitation of the modes a_n is nothing more than the initial wave modulation. Thus, in the considered case of $\kappa \ll k_0$ and of matrix elements independent of n, the wave modulation appears to be a purely phase one.

Substituting the obtained formulae (2.113) for a_n into equation (2.112) for b we arrive at

$$\frac{\partial b}{\partial t} = -V \sum_n J_n(\lambda) J_{n-1}(\lambda). \tag{2.115}$$

But the right-hand side is equal to zero due to the well-known addition theorem for the Bessel functions. Thus, the high-frequency wave modulation takes place on the background of the low-frequency wave of constant amplitude. Incorporation of a slight asymmetry of the interaction matrix elements, i.e., of their dependence on n, will lead to a low-frequency amplitude variation in time — it will either grow or oscillate at low frequency. Everything depends on which direction (to the low or to the high-frequency range) the high-frequency wave transformation predominates.

The nonlinear interaction of a high-frequency wave with plasma oscillations was observed experimentally by Stern and Tzoar [20]. The corresponding data are plotted in Fig. 2.18. As we see, this interaction excites ion-sound and Langmuir oscillations in the plasma (the pumping wave frequency was near

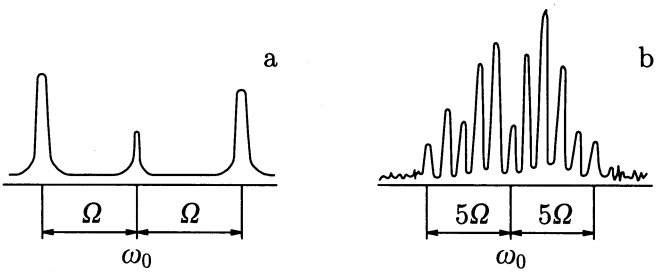

Fig. 2.18. Parametric generation of ion-sound and Langmuir oscillations by weak (a) and strong (b) pumping (ω_0 is the pumping frequency, Ω the ion-sound frequency).

the Langmuir one), and the number of satellites grows simultaneously with the pumping power in accordance with the wave interaction theory.

2.4.6. Wave front reversal

There is one more family of nonlinear phenomena that is connected to the so-called wave front reversal [21]. At first consider a simple "gedanken" experiment. Let a monochrome acoustic wave run through a dry substance, say through quicksand. We may regard the wave as a set of plane waves $a_{\mathbf{k}} \exp(-i\omega t + i\mathbf{k} \cdot \mathbf{r})$ of the same frequency. Now imagine that a uniform electric field is superimposed on the medium. Then there will be a quadratic effect of the form $b_{\mathbf{k}} \exp(i\mathbf{k} \cdot \mathbf{r})$ with $b_{\mathbf{k}}$ being independent of the frequency and proportional to $a_{\mathbf{k}}$. This motionless disturbance of a dry medium can be memorized in the form of a weak immovable deformation of the medium.

If we remove the signal sources, the frozen wave will be conserved by the medium and represent saved information. If somewhat later we impose the uniform electric field again, the quadratic effect will result in waves of the form $c_{\mathbf{k}} \exp(\pm\omega t + i\mathbf{k} \cdot \mathbf{r})$. One of these waves will run exactly in the opposite direction to the previously transmitted acoustic wave: this is wave

front reversal. Detectors placed on the sand surface will feel a signal right at the spots where the transmitting sources had been situated initially.

Apparently the above example of frozen wave memory is analogous to hologram formation known in optics: the alternating electric field plays the role of the reference wave while the acoustic wave interferes with it and results in a hologram type disturbance.

Front reversal can also occur during induced wave scattering, i.e., in a process similar to wave decay. Suppose that the electromagnetic wave (ω, k) decays into two waves: scattered electromagnetic (ω', k') and low-frequency (Ω, κ). If the frequency Ω is small, the scattered wave frequency will be near the incident wave frequency, i.e., $\omega' = \omega - \Omega \approx \omega$. The momentum conservation law $\mathbf{k} = \mathbf{k}' + \kappa$ will be satisfied at $|k'| \simeq |k|$ if $\mathbf{k}' \simeq -\mathbf{k}$ and $\kappa \simeq 2\mathbf{k}$. The scattered electromagnetic wave will run in the direction opposite to that of the incident wave, i.e., there is front reversal.

2.5. Solitons

2.5.1. Electric domains

In the above Sections we encountered the simplest solitons in a weakly dispersive medium. We observed them to be of a universal form. In different cases, for instance in the case of ion-sound, solitary pulses can be of an essentially more complicated shape. Nevertheless they are also solitons, i.e., solitary waves.

In a plasma and in similar non-steady media soliton type disturbances appear to occur frequently. So now we shall try to get acquainted with solitons by considering different cases by way of example. Let us start off with solitons that arise under certain conditions with an electric current in semiconductors and are called electric domains. Although this phenomenon is naturally studied under topics of semiconductor physics, its physical nature is closely related to plasma.

In 1963 Gunn discovered an extremely interesting effect: if an electric field of about 3 kV cm^{-1} is imposed on an AsGa n-type crystal sample, a coherent alternating current of large amplitude is generated [22]. The oscillation frequency, which was about 10^9 Hz, appeared to be inversely proportional to the sample length. Later on Gunn experimentally determined [23] that these oscillations were connected with the motion from the cathode to the anode of domains with a strong electric field: each hump of the electric current in the circuit corresponded to the birth of such a domain on the cathode followed by its passage through the sample. These domains were called electric domains. Fig. 2.19 presents their forms, with two curves corresponding to different voltages applied to the sample. It can be seen that the domains are asymmetric — their rare slope is steeper than the front one. If the domain amplitude (i.e., the voltage drop on the domain) becomes less, the asymmetry vanishes and the shape becomes symmetric.

Fig. 2.19. Electric field within a stationary moving domain in AsGa. The domains are moving to the left with the speed 1.2×10^7 cm s^{-1}.

The origin of electric domain formation in semiconductors was clarified on the grounds of the notion of negative differential conductance. Suppose that there is a gradual slope between E_{c1} and E_{c2} in the current-voltage diagram, such that the differential conductance $\sigma_d = dj/dE$ is negative (Fig. 2.20).

At a fixed value of the current through the circuit, say at j_0, there are three available values of the applied voltage — E_1, E_2, E_3. One of these values, namely E_2, appears to be un-

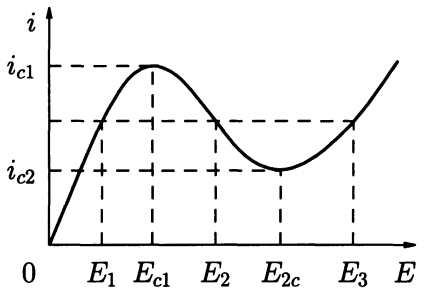

Fig. 2.20. Dependence of the current on the applied voltage for a semiconductor with a negative differential conductance.

stable. Indeed, suppose that $E = E_2$ and that there is a small fluctuation n' of the electron density. Due to the presence of electrons the fluctuation area will be of negative potential. But in view of $\sigma_d < 0$ the electrons will be attracted by this area and the fluctuation will grow. So because of instability charged domains must appear. Since the negative differential conductance is created only along the external field lines, these domains will be of the form of thin layers oriented perpendicularly to the current lines. Assuming that the current is directed along the x axis and that all the quantities depend only on x, we can describe this instability with the equations

$$j = en\mu(E)E + eD\frac{\partial n}{\partial x}, \tag{2.116}$$

$$\varepsilon\frac{\partial E}{\partial x} = -4\pi e(n - n_0), \tag{2.117}$$

$$j + \frac{\varepsilon}{4\pi}\frac{\partial E}{\partial t} = j_0 = \text{const}, \tag{2.118}$$

where $\mu(E)$ is the electron mobility, D — their diffusion coefficient, ε — the medium dielectric permeability, n_0 — the donor density, j_0 — the external current. We preserved the displacement current in equation (2.118), thus it resembles the current conservation. Equations (2.116)–(2.118) in the linear approximation yield the following expression for the frequency of a small

perturbation proportional to $\exp(-i\omega t + ikx)$:

$$\omega = ku - i\frac{4\pi\sigma_d}{\varepsilon} - iDk^2. \tag{2.119}$$

Here $u = -\mu(E_0)E_0$ is the electron drift velocity and $\sigma_d = en_0 d(\mu E)/dE$ stands for the differential conductance. One can see from (2.119) that local linear perturbations with sufficiently small k grow with time provided $\sigma_d < 0$.

While the amplitude grows, the nonlinear effects come to the fore. We pass to the area $E < E_{c1}$ or to $E > E_{c2}$ (see Fig. 2.20), where $\sigma_d > 0$, and the oscillations will stabilize. The steady oscillations can be treated with the help of the same equations (2.116)–(2.119). Assuming the perturbation to be of the form $E(x - v_p t)$ of a progressive wave and eliminating n and j, we obtain the only nonlinear differential equation of second order in E:

$$D\frac{\partial^2 E}{\partial x^2} + \left[\mu(E)E + v_p\right]\frac{\partial E}{\partial x} = \frac{4\pi e n_0}{\varepsilon}\mu(E) - E - j_0 = -\frac{\partial W}{\partial E}. \tag{2.120}$$

As in the above investigation of the nonlinear Korteweg – de Vries equation we may regard (2.120) as the evolution equation of a nonlinear oscillator, such that the friction force is described by the term with the first-order derivative on the left-hand side. In order that the friction force power average out to zero, the magnitude of v_p should apparently be of the order of the electron drift velocity.

The equivalent potential $W(E)$ on the right-hand side of (2.120) slightly depends on j_0 as is shown in Fig. 2.21. There is a certain value j_b, such that the potential curve touches the zero level twice. This corresponds to the case when there exists a solution of the form of a wide soliton: at first the field E is kept at the value E_1, then it passes to the point E_3 and stays there for a long time, and finally it quickly returns to the initial value. The cases $j_0 > j_b$ and $j_0 < j_b$ correspond respectively to the cases of the strong-field domain (soliton) and to the weak-field

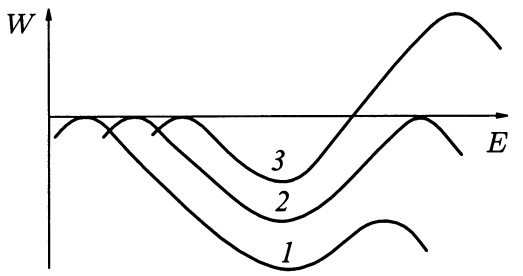

Fig. 2.21. Potential $W(E)$ for an electric domain. *1* — $j_0 < j_b$; *2* — $j_0 = j_b$; *3* — $j_0 > j_b$.

soliton. There are also plenty of periodic and dissipative shock-wave types solutions, but the real Gunn domain corresponds to a simple strong-field soliton.

Since then electric domains have been found in different semiconductors. It also turned out that there were plenty gearing to negative differential conductance. For instance, the condition $\sigma_d < 0$ for AsGa, where the domains were found for the first time, is related to electron heating and their transfer to a valley of the Fermi surface with lower effective mass. For details and further examples see review [24].

2.5.2. The nonlinear Schrödinger equation

It turns out that the previously considered nonlinear phenomena of self-focusing and self-compression also lead to solitons, this time to envelope solitons. In order to describe the corresponding solutions, we have to treat the slightly non-linear waves, whose frequency ω depends on the amplitude a as $\omega = \omega_k + \alpha|a|^2$, more precisely than before. Let us rewrite the dispersion law in a form analogous to (2.78):

$$\left(\omega - \omega_{\mathbf{k}} - \alpha|a|^2\right)a_{\mathbf{k}} = 0. \qquad (2.121)$$

This equation states that only the modes corresponding to the zero bracketed expression do not vanish.

Multiplying each equation of system (2.121) for the elementary wave amplitude $a_{\mathbf{k}}$ by the factor $\exp(-i\omega t + i\mathbf{k}\cdot\mathbf{r})$ and summing over $\mathbf{k}$, we obtain the equation for the function $\psi(\mathbf{r},t) = \sum a_{\mathbf{k}} \exp(-i\omega t + i\mathbf{k}\cdot\mathbf{r})$. Certainly we are to substitute the operators $i\partial/\partial t$ and $-i\nabla$ for ω and $\mathbf{k}$.

Let us now apply this method to a weakly modulated monochromatic wave. Suppose that the function $\psi(\mathbf{r},t)$ takes the form $\psi(\mathbf{r},t) = \Psi_a(\mathbf{r},t)\exp(-i\omega_0 t + ik_0 x)$, with ω_0 corresponding to the eigenfrequency of the plane harmonic wave with wavenumber k_0, and that the wave itself propagates along the x axis. If the amplitude $\Psi_a(\mathbf{r},t)$ is not constant, the wave is not monochromatic in the lateral sense of the word. In order to take this fact into account we have to use a superposition of equations (2.121). Formally, the amplitude $a_{\mathbf{k}}$ in equation (2.121) should be replaced by the function $\psi(\mathbf{r},t)$, while the frequency ω and the wavevector $\mathbf{k}$ — by the operators $i\partial/\partial t$ and $-i\nabla$. One can easily see that

$$i\frac{\partial\psi}{\partial t} = \left(\omega_0\Psi_a + i\frac{\partial\Psi_a}{\partial t}\right)\exp\left[-i\omega_0 t + ik_0 x\right], \quad (2.122)$$

$$-i\nabla\psi = \left(\mathbf{k}_0\Psi_a - i\nabla\Psi_a\right)\exp\left[-i\omega_0 t + ik_0 x\right]. \quad (2.123)$$

Let us assume the medium to be isotropic so that $\omega_{\mathbf{k}} = \omega_k$ shall depend only on the absolute value of the wavevector. Near $\mathbf{k} = \mathbf{k}_0$ the frequency $\omega_{\mathbf{k}}$ may be expanded in a power series in $|\mathbf{k} - \mathbf{k}_0|$. Keeping the terms up to the second order, we have

$$k = \sqrt{k_x^2 + |\mathbf{k}_\perp|^2} \simeq k_0 + (k_x - k_0) + \frac{1}{2k_0}|\mathbf{k}_\perp|^2.$$

The corresponding expansion of the frequency is

$$\omega_k = \omega_0 + v_g(k_x - k_0) + \frac{1}{2}v_g'(k_x - k_0)^2 + \frac{v_g}{2k_0}|\mathbf{k}_\perp|^2 \quad (2.124)$$

with $v_g = \partial\omega_k/\partial x$, $v_g' = \partial v_g/\partial k = \partial^2\omega_k/\partial k^2$. These relations should be substituted into the superposition of the equations

(2.121), i.e., into the equation for ψ. Omitting the index a of the amplitude Ψ_a, we obtain

$$i\frac{\partial\Psi}{\partial t} + iv_g\frac{\partial\Psi}{\partial x} + \frac{v_g}{2k_0}\Delta_\perp\Psi + \frac{1}{2}v_g'\frac{\partial^2\Psi}{\partial x^2} - \alpha|\Psi|^2\Psi = 0. \quad (2.125)$$

This equation for the envelope amplitude was called the nonlinear Schrödinger equation. Its linear version, i.e., an equation of the form (2.125) with $\alpha = 0$, was proposed earlier by M. A. Leontovich [25].

Let us return to the problem of a plane wave instability. Put $\Psi_0 = a_0\exp(i\nu_0 t)$ with a_0 and ν_0 being the initial amplitude and frequency. According to (2.125) we have $\nu_0 = \alpha a_0^2$. Let us now suppose that the wave Ψ_0 is slightly perturbed, so that its amplitude and phase alter in space and with time:

$$\Psi_0 = (a_0 + a')\exp\left[-i\nu_0 t + i\varphi'\right].$$

The small quantities a' and φ' are to be treated as real. Substituting this expression into (2.125) and omitting the terms quadratic in a' and φ', we obtain the system

$$\frac{\partial a'}{\partial t} + v_g\frac{\partial a'}{\partial x} - a_0 L\varphi' = 0, \quad (2.126)$$

$$a_0\frac{\partial\varphi'}{\partial t} + a_0 v_g\frac{\partial\varphi'}{\partial x} + La' + 2\alpha a_0^2 a' = 0 \quad (2.127)$$

of equations for a' and φ', with the operator L given by

$$L = -\frac{v_g}{2k_0}\Delta_\perp - \frac{1}{2}v_g'\frac{\partial^2}{\partial x^2}.$$

Equations (2.126)–(2.127) are linear and one can look for a solution proportional to $\exp(-i\nu t + i\boldsymbol{\kappa}\cdot\mathbf{r})$. Under this condition L may be treated as a number equal to

$$L = \frac{v_g}{2k_0}\kappa_\perp^2 + \frac{1}{2}v_g'\kappa_x^2.$$

The condition that equations (2.126)–(2.127) be compatible results in the following expression for the frequency:

$$\nu = v_g \kappa_x \pm \sqrt{L(2\alpha a_0^2 + L)}. \tag{2.128}$$

At $\kappa_\perp = 0$ and for small κ_x it gives the previous result (2.76), i.e., self-compression instability in the case of $\alpha v_g' < 0$. But (2.128) indicates that at $\alpha v_g' < 0$ self-compression takes place only under the condition that κ_x is sufficiently small. If $\kappa_x^2 > -4\alpha a_0^2 / v_g'$, the self-compression process is canceled by the diffraction that tends to increase the wave packet dimensions. In the limit $\kappa_x = 0$ relation (2.128) implies self-focusing for $\alpha < 0$. Nevertheless the instability develops only if $\kappa_\perp$ is sufficiently small, namely if $\kappa_\perp^2 < 4|\alpha| a_0^2 k_0 / v_g$. Since the minimum value of $\kappa_\perp$ is of the order $1/R$ of the inverse radius of the cylindrical beam, one may state that in the framework of the given approximation, when only the terms quadratic in amplitude are kept in expression (2.70) for the frequency deviation, the self-focusing process begins only if the laser power exceeds a certain critical value proportional to $a_0^2 R^2$.

Stationary focused beams and self-compressed wave packets are feasible due to the stabilizing effect of diffraction [26, 27]. Suppose, for instance, that $\alpha v_g' < 0$ and that there is no dependence on y and z. On the grounds of (2.125) we conclude that

$$\Psi = \exp(-i\nu_0 t)\, u(x - v_g t)$$

is a solution of the form of a running wave packet provided u satisfies the equation

$$\frac{\partial^2 u}{\partial x^2} = \frac{2\alpha}{v_g'} u^3 - \frac{2\nu_0}{v_g'} u. \tag{2.129}$$

But we treated equations of this type while considering periodic solutions to the Korteweg–de Vries equation. So we shall again regard (2.129) as the equation for a nonlinear oscillator with the potential energy

$$W = -\frac{\alpha}{2v_g'} u^4 + \frac{\nu_0}{v_g'} u^2.$$

For $\alpha v_g' < 0$ and $\nu_0 v_g' < 0$ the potential W is of the form of a well and so both periodic envelope waves and localized packets like solitons are feasible. In a similar manner one can obtain solutions local in the lateral direction in the case $\alpha < 0$. These solutions correspond to beams that provide themselves with a waveguide and look like thin threads running through the medium. However, they appear to be unstable [28].

2.5.3. Vortex filament solitons

We shall now consider a very interesting example of soliton formation in the case of special vortex motion in an incompressible nonviscous fluid. Vortex filaments in an ideal fluid are known to preserve their individuality and their evolution in three-dimensional space can be approximately described by the so-called localized induction equation. Under this approximation the motion of a thin vortex filament is supposed to be locally identical to the motion of a thin cylindrical vortex of the same curvature. Let $\mathbf{r} = \mathbf{r}(s,t)$ be the radius-vector of the vortex line, where s stands for the arclength measured along the filament and t for the time. The first asymptotic equation governing the filament motion is

$$\frac{\partial \mathbf{r}}{\partial t} = G\kappa\mathbf{b}. \tag{2.130}$$

Here $G \simeq (\Gamma/4\pi)\log(1/\kappa\delta)$ is the coefficient of the local induction that is proportional to the circulation Γ around the filament, κ is the curvature, and $\mathbf{b}$ is the unit vector of the binormal direction. The magnitude of G may be considered to be constant as long as the logarithm variations can be neglected, i.e., when $\kappa\delta \ll 1$ with δ being the vortex diameter. By rescaling the time t we transform equation (2.130) to the nondimensional form

$$\dot{\mathbf{r}} = \kappa\mathbf{b} \tag{2.131}$$

with the dot denoting $\partial/\partial t$. This equation should be supplemented with the Frenet–Seret equations

$$\begin{aligned} \mathbf{r}' &= \mathbf{t}, & \mathbf{t}' &= \kappa\mathbf{n}', \\ \mathbf{n}' &= -\kappa\mathbf{t} + \tau\mathbf{b}, & \mathbf{b}' &= -\tau\mathbf{n} \end{aligned} \tag{2.132}$$

of differential geometry. Here the prime denotes $\partial/\partial s$, κ is the curvature, τ is the torsion, whereas $\mathbf{t}, \mathbf{n}, \mathbf{b}$ are orthogonal unit vectors pointing along the tangent, the principal normal and the binormal directions, respectively.

Following Hasimoto [29], we can essentially simplify these equations. At first we are to evaluate $\dot{\kappa}$ and $\dot{\tau}$. Differentiating (2.131) with respect to s and taking account of (2.132), we obtain

$$\dot{\mathbf{t}} = \kappa'\mathbf{b} - \kappa\tau\mathbf{n}. \tag{2.133}$$

By applying the above procedure to this relation once more, we find that

$$\dot{\kappa}\mathbf{n} + \kappa\dot{\mathbf{n}} = \kappa^2\tau\mathbf{t} - (2\kappa'\tau + \kappa\tau')\mathbf{n} + (\kappa'' - \kappa\tau^2)\mathbf{b}. \tag{2.134}$$

The $\mathbf{n}$-component of this equation reads as

$$-\dot{\kappa} = 2\kappa'\tau + \kappa\tau'. \tag{2.135}$$

Therefore,

$$\dot{\mathbf{n}} = \kappa\tau\mathbf{t} + \left(\frac{\kappa''}{\kappa} - \tau^2\right)\mathbf{b}. \tag{2.136}$$

Differentiating this relation with respect to s and projecting the result on $\mathbf{b}$, we obtain in view of (2.133) and of the Frenet–Seret equations the following identity:

$$\dot{\tau} = \left(\frac{\kappa''}{\kappa} - \tau^2 + \frac{\kappa^2}{2}\right)'. \tag{2.137}$$

To simplify the dynamic equations (2.135) and (2.137), Hasimoto proposed using the complex function

$$\Psi = \kappa \exp\left(i\int \tau ds\right), \tag{2.138}$$

where the integral is to be taken along the filament. It is easy to verify that the equation

$$i\frac{\partial\Psi}{\partial t}+\frac{\partial^2\Psi}{\partial s^2}+\frac{1}{2}|\Psi|^2\Psi=0 \tag{2.139}$$

is equivalent to equations (2.135)–(2.137).

Thus, the equations governing vortex filament curvature and torsion evolution in an incompressible nonviscous fluid are reduced to a special form of the nonlinear Schrödinger equation, corresponding to the envelope instability. It therefore describes soliton formation.

It is easy to reveal the shape of these solitons. Assuming κ and τ to be unknown functions of the only variable $\xi = s-ct$, we find from (2.135) that at $\kappa^2 \neq 0$

$$\tau=\tau_0=\frac{c}{2}=\text{const}. \tag{2.140}$$

Thus, the torsion is constant along the filament. As long as $\tau = \text{const}$ any solution to equation (2.137) looks like

$$\kappa=\frac{2\nu}{\sinh(\nu s)}, \tag{2.141}$$

where ν stands for an arbitrary constant. Thus, we obtained a localized soliton type solution. With $\tau = \text{const}$ and the explicit dependence of the curvature on s at hand the form of the vortex filament can be found.

In particular, when the torsion is zero, the vortex filament appears to be like a cord in a vertical two-dimensional plane that is permanently spinning at a constant angular velocity around the z-axis. With the notation θ for the angle between the tangent vector $\mathbf{t}$ and the axis z of revolution and ω for the angular velocity we have

$$\dot{\mathbf{t}}=-\omega\sin\theta\,\mathbf{b}. \tag{2.142}$$

Recalling that $\kappa = d\theta/ds$ and utilizing (2.133), we obtain

$$\frac{d^2\theta}{ds^2}=-\omega\sin\theta. \tag{2.143}$$

This is the nonlinear pendulum equation, which, in particular, has a kink-like solution. It was shown by Hasimoto [30] that the same equation describes the shape of a plane elastic filament. Indeed, suppose for instance that the tension of such a filament at rest is T and is directed along the z axis as $z \to \pm\infty$. Then due to internal stresses there is a force of constant magnitude applied to any filament cross-section. This force produces in turn a torque $T \sin\theta$ per filament portion of unit length. In view of the balance between this torque and the turning momentum $-\alpha d^2\theta/ds^2$ produced by the stresses, with α being proportional to the flexural modulus, we have

$$\frac{d^2\theta}{ds^2} = -\frac{T}{\alpha}\sin\theta. \tag{2.144}$$

A soliton corresponds to a single loop on the elastic filament with $\sin\theta \to 0$ as $s \to \pm\infty$. This pattern is familiar to all of us.

2.5.4. Two-dimensional solitons

It has been shown earlier (Section 2.3.1) that nonlinear waves in a two-dimensional dispersive medium may be described by the Kadomtsev–Petviashvily equation. To proceed with its investigation let us convert it to the dimensionless form

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - \alpha\frac{\partial^3 u}{\partial x^3}\right) = \frac{\partial^2 u}{\partial y^2}. \tag{2.145}$$

The only remaining parameter α takes the value $+1$ in the case of a medium of positive dispersion (the corresponding equation being called KP1) and -1 in the opposite case, corresponding to the KP2 equation. The intrinsic properties of these equations are radically different. We have just observed this while considering line solitons in a two-dimensional medium: in a medium of negative dispersion such solitons are stable with respect to a wave front bend, whereas in a positive-dispersion medium a line soliton is unstable. In the linear approximation the KP1-soliton instability turns out to be of a self-focusing nature. Certainly in

the framework of the linear approximation the question of the nonlinear development of self-focusing remains open. But, fortunately, the Kadomtsev – Petviashvily equations appear to be in the class of completely integrable equations, and can be solved exactly by means of the inverse scattering method. Moreover, the KP1 equation for a medium of positive dispersion possesses solutions of the form of two-dimensional solitons local in all directions and decreasing at infinity as x^{-2}, y^{-2}. Let us examine them more thoroughly.

As we have seen before, a one-dimensional soliton in a medium of positive dispersion proved to be subsonic and of negative amplitude. It is natural to expect a two-dimensional soliton to exhibit the same characteristics. Therefore it is convenient in the case of positive dispersion, i.e., at $\alpha = 1$, to rewrite equation (2.145) in a slightly different form, using the transformation $x \to -x$, $u \to -u$. On performing the substitution, we find out that both cases are described by the unique equation

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3}\right) = \alpha\frac{\partial^2 u}{\partial y^2}. \tag{2.146}$$

As before, $\alpha = 1$ corresponds to the case of positive dispersion and $\alpha = -1$ to the case of negative dispersion. In both cases one-dimensional solitons of the equation (2.146) are of positive amplitude and of positive phase speed. Putting $\alpha = 1$, $u = u(x - t)$, we obtain the following equation for a two-dimensional soliton:

$$\frac{1}{2}\frac{\partial^2 u^2}{\partial x^2} + \frac{\partial^4 u}{\partial x^4} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}. \tag{2.147}$$

The analytic expression for its solution is rather simple [31, 32]. Namely, one may verify by direct calculations that the solution is presented by the formula

$$u = 24\frac{(3 - x^2 + y^2)}{(3 + x^2 + y^2)^2}. \tag{2.148}$$

We see that at $x = 0$ the soliton section along the y axis is of the form of a hump, monotonically decreasing as $u \sim y^{-2}$

as $y \to \infty$. The dependence on x is more complicated: the function u is positive in the central domain of the line $y = 0$, but for $x^2 > 3$ its sign is reversed and the areas of negative u are formed.

Let $u(x, y)$ be the soliton (2.148) and, hence, a solution to equation (2.147). Apparently the function $c^2 u(cx - c^2 t, cy)$ is also a solution to (2.146) in the case of a positive dispersion, i.e., at $\alpha = +1$. In fact, this can be easily proved for a uniformly moving soliton with the aid of the symmetries of equation (2.146).

Thus, if the soliton amplitude undergoes rescaling by the factor c^2, its phase velocity will be rescaled by the same factor, whereas the dimensions will be rescaled by the inverted c-factor only. For instance, for $c^2 > 1$ the amplitude and phase velocity will increase while the dimensions decrease [in comparison to the standard soliton (2.148)]. In other words, we have a one-parameter family of self-similar solitons with the "tall" solitons being thin and mobile and the "short" ones being "fat" and "lazy." Thus, solitons appear to follow the classic Don Quixote and Sancho Pansa example.

The analytic theory [31, 32] of the KP1 equation enables one to construct not only single-soliton, but multi-soliton solutions as well. Investigation of the formulae describing multi-soliton evolution revealed the surprising property that the KP1-solitons not only preserve their form and initial parameters (amplitude, velocity, dimensions), but also appear to be subjected to a zero phase shift. However, even the two-soliton interaction is not reduced to a mere superposition of their fields. Moreover, two identical solitons appear to exhibit one more unusual aspect of behavior [33]. Namely, two solitons starting to move, say in the direction of x, one after the other, but not exactly in a line, experience attraction even with a large distance between them (compared to their dimensions). Thus, they are drawn together with the front soliton amplitude being decreased and the rear soliton amplitude being increased. Simultaneously they repel each other in the y direction and so, while being drawn together

in x, they disperse in y. With increasing distance the solitons become equal in amplitude and as $t \to \infty$ they become two identical solitons an infinite distance apart in y moving in the x direction. The speed of the initial approach in x and of the final dispersion in y changes with time as $|t|^{-1/2}$.

There is also a solution describing the reverse process. In this case two identical solitons infinitely separated in y, start moving in the direction of x. For some time they appear to attract each other in the direction of y, but later on they line up and as $t \to \infty$ separate from each other in the x direction, acquiring the initial velocities and amplitudes.

Besides the above solutions there are also steady multi-soliton structures resembling bound states. Thus, the KP1 equation analytically describes the complicated soliton "life" in a two-dimensional medium of weak dispersion and quadratic nonlinearity.

Now let us pass to the nonlinear Schrödinger equation — the simplest representative of wave equations with cubic nonlinearity. Assuming that there is no dependence on x in equation (2.125), we put $\psi = \psi(\mathbf{r}) \exp(-i\Omega t)$ and obtain the standard two-dimensional equation depending on $\rho = \sqrt{y^2 + z^2}$. For $\alpha < 0$ it looks like the Schrödinger equation with the potential well $\alpha|\psi|^2$ depending upon ψ. It is known that there is always a bound state in a deep well. Suppose, in the first instance, that we revealed the corresponding eigenfunction $\psi_0(\rho)$ and eigenvalue Ω_0 and, therefore, are aware of the "number of particles" N, i.e., of the integral of $|\psi|^2$ over a cross-section of our two-dimensional axially-symmetric soliton. With the notation ρ_0 for its lateral dimension we can easily construct a single-parameter family of solitons at the fixed value of N. Indeed, consider the function $\psi_0(\rho c)$ of the lateral dimensions rescaled by c. Then the values of $\Delta_\perp \psi_0$ and of $|\psi|^2$ at fixed N will be multiplied by c^2 and, hence, we obtain a soliton type solution with the eigenvalue $c^2\Omega$. This single-parameter family corresponds to a certain critical value N_c of the "number of particles."

Thus, we have become acquainted with two-dimensional solitons in two-dimensional media described by the Kadomtsev–Petviashvily and the nonlinear Schrödinger equations.

2.5.5. Wave collapse

Let $\psi = \psi_0(\rho)\exp(-i\Omega_0 t)$ be a soliton-like solution to the nonlinear Schrödinger equation. This solution corresponds to a certain value $N = N_c$ of the "number of particles" defined as the integral of $|\psi|^2$ over the soliton cross-section. We have established before that the function $\psi = c\psi_0(\rho c)\exp(-i\Omega_0 c^2 t)$ also appears to be a solution to (2.125), with the constant c being proportional to the lateral dimensions of the soliton. It turns out that the function $\psi_0(\rho)$ enables one to construct a steady solution describing soliton self-focusing [34]. Suppose for simplicity that $v_g' = 0$. This is not a restriction since this quantity is related to the self-compression and has no role in self-focusing. One can easily verify that

$$\psi = \frac{1}{x}\psi_0\left(\frac{\rho}{x}\right)\exp\left(\frac{i\Omega_0}{v_g x}\right) \tag{2.149}$$

is a solution to the stationary Schrödinger equation (2.125). We see that this solution describes a soliton whose lateral dimension tends to zero as $x \to 0$. In other words a wave propagating along x is focused at the point $x = 0$.

If it were an envelope wave, the governing equation would not be valid on scales of the order of the wavelength. But if the nonlinear Schrödinger equation describes a particular physical process down to infinitesimal scale, one may speak about wave field compression into a "point." This process has been called wave collapse [35].

Wave collapse can also occur in a three-dimensional medium of positive dispersion. In this case the KP1 equation (2.146) is apparently transformed to the form

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3}\right) = \alpha\Delta_\perp u \tag{2.150}$$

with $\Delta_\perp = (\partial^2/\partial y^2 + \partial^2/\partial z^2)$ and $\alpha = +1$ for a medium of positive dispersion. Suppose that there is a two-dimensional soliton (2.148) with no dependence on z. Repeating the arguments that led to the instability of a one-dimensional soliton in a two-dimensional medium, one can verify that the soliton under consideration is unstable with respect to self-focusing. Moreover, we shall show that all attempts to construct a steady three-dimensional soliton at $\alpha = +1$ will fail. The cause is that in the present case the energy flows to the areas of maximum amplitude in two dimensions: along y and along z. The nonlinearity and dispersion in x fail to stop the self-focusing process and the soliton compresses into a "point." On offering these physical arguments, let us proceed to a more quantitative analysis.

It is easy to verify that the three-dimensional equation (2.150) conserves the following integrals:

$$P = \int u^2 d\mathbf{r}, \tag{2.151}$$

$$H = \int \left\{ \left(\frac{\partial u}{\partial x} \right)^2 + (\alpha \nabla_\perp \mu)^2 - \frac{2}{3} u^3 \right\}. \tag{2.152}$$

Here the quantity μ is defined by the identity $\partial\mu/\partial x = u$, whereas P and H give the momentum and the energy of the wave field. One can easily verify that equation (2.150) can be represented in the Hamiltonian form

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(\frac{\delta H}{\delta u} \right). \tag{2.153}$$

It turns out [36] that if a soliton solution existed, the sign of its energy could be reversed by an appropriate variation of the shape with P kept fixed. Therefore, according to Lyapunov's theorem, the soliton under consideration is unstable. Computer modeling shows [37] that a localized solution to the KP1 equation collapses in a finite time.

We shall encounter the collapse once more in considering Langmuir turbulence and solitons.

3. Waves and particles

3.1. Landau damping

3.1.1. Langmuir waves

In a plasma, ion-sound waves are in the class of oscillations for which the ions play a key role. It is not difficult to consider electron modes, which are of higher frequencies and appear to be unaffected by the ions. Let us consider the most simple type of electron oscillation — Langmuir waves.

The Langmuir waves are excited when the quasi-neutrality of a plasma is violated, i.e., when the electrons are displaced with respect to the ions. The arising electric field results in a quasi-elastic force that tends to return the electrons to the state of equilibrium. Since the electrons are much lighter than the ions, the electron oscillations caused by this quasi-elastic force take place on a background of practically motionless ions.

If the electron thermal motion is neglected, hydrodynamical language can be used. Let $\boldsymbol{\xi}$ be the electrons' displacement from their equilibrium position, which we assume to be small. Then the equation of motion for an electron can be written in the form

$$\frac{\partial^2 \boldsymbol{\xi}}{\partial t^2} = \frac{e}{m_e} \nabla \varphi, \tag{3.1}$$

where φ is the electric field potential defined by

$$\Delta \varphi = 4\pi e n'. \tag{3.2}$$

In the linear approximation the electron density perturbation n' is equal to

$$n' = -n \nabla \cdot \boldsymbol{\xi} \tag{3.3}$$

Evaluating the divergence of (3.1) and excluding $\boldsymbol{\xi}$ and φ with the aid of (3.2)–(3.3), we obtain

$$\frac{\partial^2 n'}{\partial t^2} = -\omega_{pe}^2 n' \tag{3.4}$$

with

$$\omega_{pe} = \sqrt{4\pi e^2 n/m_e}. \tag{3.5}$$

This value is usually called the Langmuir frequency. From equation (3.5) it follows that any violation of quasi-neutrality results in electron oscillations with the frequency ω_{pe}. If the initial disturbance has the shape of a plane wave with wavenumber k, the Langmuir wave will also be plane with a phase velocity equal to $v_p = \omega_{pe}/k$.

It is seen that the phase speed diminishes with an increase of k and becomes of the same order in magnitude as the electron thermal velocity at $k \sim 1/d$. Since we neglected the thermal motion, equation (3.4) is valid for $kd \ll 1$. When the phase speed approaches the thermal velocity, a kinetic description of plasma should be used.

3.1.2. Vlasov Equation

To describe oscillations of a rarefied plasma under the condition that binary collisions are not important, Vlasov [38] proposed a kinetic equation with a self-consistent field and eliminated the collision terms. For particles of mass m and charge e in the absence of a magnetic field this equation takes the form

$$\frac{\partial f}{\partial t} + \boldsymbol{v}\nabla f + \frac{e}{m}\mathbf{E} \cdot \frac{\partial f}{\partial \boldsymbol{v}} = 0. \tag{3.6}$$

The function f is known as the velocity distribution and represents the particle density in the phase space, so that $f d\mathbf{r} d\boldsymbol{v}$ is the number of particles located inside an infinitesimal volume $d\mathbf{r}$ and having velocities within $d\boldsymbol{v}$.

According to Vlasov, the self-consistent field $\mathbf{E}$ in (3.6) is to be determined from the equation

$$\operatorname{div} \mathbf{E} = 4\pi e(n_i - n_e), \tag{3.7}$$

where

$$n_i = \int f_i d\boldsymbol{v}, \quad n_e = \int f_e d\boldsymbol{v}, \tag{3.8}$$

and f_e and f_i are the distribution functions of the electrons and ions.

If the magnetic field $\mathbf{B}$ is taken into account, the kinetic equation for particles of type α with a self-consistent field will take the form

$$\frac{\partial f_\alpha}{\partial t} + \boldsymbol{v} \cdot \nabla f_\alpha + \frac{e_\alpha}{m_\alpha} \left\{ \mathbf{E} + \frac{1}{c} \boldsymbol{v} \times \mathbf{B} \right\} \cdot \frac{\partial f_\alpha}{\partial \boldsymbol{v}} = 0. \tag{3.9}$$

The self-consistent electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ are to be determined from the Maxwell equations that should involve the electrical charge ρ and current $\mathbf{j}$ densities in the form

$$\rho = \sum_\alpha e_\alpha \int f_\alpha \, d\boldsymbol{v}, \quad \mathbf{j} = \sum_\alpha e_\alpha \int \boldsymbol{v} f_\alpha \, d\boldsymbol{v}. \tag{3.10}$$

When binary collisions have no role in a rarefied plasma, it is refered to as a collisionless plasma. Correspondingly, the collective processes are called collisionless when Coulomb collisions are not important. The Vlasov equation with a self-consistent field serves as a foundation for the theoretical description of such processes.

3.1.3. Landau damping

We shall now proceed to the consideration of small-amplitude Langmuir waves in the case when the thermal motion of particles begins to play an active role. As was originally shown by Landau [39], a completely new specific damping mechanism appears under these conditions and it operates even in the absence of collisions. In view of the fundamental role of Landau damping in cooperative effects in a collisionless plasma, we shall use a more precise mathematical approach in contrast to the previous Sections where we simplified matters to some extent.

Let us assume that the ions constitute a motionless homogeneous background with density $n = \text{const}$ and consider a Langmuir wave propagating in the direction of x, so that the problem is one-dimensional.

The equations for small-amplitude oscillations can be linearized. To do this, put $f = f_0 + f'$ with f_0 standing for the equilibrium function and f' for the deviation from equilibrium, and neglect the quadratic term $(e/m)\mathbf{E} \cdot \partial f'/\partial \boldsymbol{v}$ in (3.6). A plasma is neutral and homogeneous in the state of equilibrium, so that $\int f_0 d\boldsymbol{v} = n$. Due to the homogeneity, these linear equations can be written in terms of amplitude components of the expansions of f' and φ in Fourier integrals. Therefore, it is sufficient to consider the evolution of a separate Fourier-component. Assuming the functions f' and φ to be of the form $f'(v,t)\exp(ikx)$, $\varphi(t)\exp(ikx)$, we write down the linearized equations for electrons with the negative charge $-e$:

$$\frac{\partial f'}{\partial t} + ikvf' + ik\varphi \frac{e}{m_e}\frac{\partial f_0}{\partial v} = 0, \tag{3.11}$$

$$k^2\varphi = -4\pi e \int f' dv. \tag{3.12}$$

It is natural to start off with a search for an eigenmode corresponding to some frequency ω. To do this, we put

$$f'(v,t) = f_1(v)\exp(-i\omega t), \quad \varphi(t) = \varphi_1 \exp(-i\omega t). \tag{3.13}$$

From equation (3.11) we might have

$$f_1 = \frac{k}{\omega - kv}\frac{e}{m_e}\frac{\partial f_0}{\partial v}\varphi_1. \tag{3.14}$$

Substituting this expression into (3.12), we obtain the dispersion equation, relating the eigenfrequency ω to the wavenumber k:

$$\varepsilon(k,\omega) \equiv 1 + \frac{4\pi e^2}{km_e}\int \frac{(\partial f_0/\partial v)dv}{\omega - kv} = 0. \tag{3.15}$$

We have used ε to denote the plasma dielectric permeability.

As we see, expression (3.15) has a singularity in the integrand, hence it can not be used until we specify how to treat it. Vlasov [40] proposed taking the principal value of the integral, but there are not sufficient grounds to do this.

Landau [39] pointed out the correct approach to the solution of the problem of small amplitude plasma oscillations. This approach has also shown the way to overcome the ambiguity caused by the divergence in (3.15). Landau noticed that a naturally posed problem about small oscillations usually deals with the given initial or boundary conditions, and showed the way to solve both problems.

Consider, for instance, the initial condition problem. It should be assumed in this case that the disturbance vanishes for $t < 0$. At $t = 0$ the external source is switched on and the initial perturbation $g(v)$ of the distribution function is created. The problem is to determine the temporal evolution of that perturbation. To do this, we use equations (3.11)–(3.12) supplied with the external source $g(v)\delta(t)$ on the right-hand side of (3.11). To solve this system, we may use the Laplace transform technique. For this purpose let us introduce

$$f_p(v) = \int_0^\infty \exp(-pt) f'(v,t) dt \tag{3.16}$$

and a similar representation for $\varphi(t)$.

Multiplying equation (3.11) with the source $g(v)\delta(t)$ on the right-hand side by $\exp(-pt)$ and integrating, we obtain the relation

$$f_p = \frac{ig(v)}{ip - kv} + \frac{k}{ip - kv} \frac{e}{m_e} \frac{\partial f_0}{\partial v} \varphi_p. \tag{3.17}$$

As we see, this expression differs from (3.14) by the additional term with g and in ω being replaced by ip. Substitution of (3.17) into (3.12) results in

$$\varphi_p = -i \frac{4\pi e}{k^2 \varepsilon(k, ip)} \int \frac{g(v) dv}{ip - kv}, \tag{3.18}$$

where $\varepsilon(k, \omega)$ is given by expression (3.15). As φ_p is known, it is not difficult to find $\varphi(t)$:

$$\varphi(t) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \exp(pt)\, \varphi_p dp. \tag{3.19}$$

It appears to be convenient to replace the variable p by $-i\omega$, so that integration over the complex variable ω should be performed in the upper half-plane:

$$\varphi(t) = -i\frac{2e}{k^2}\int_{-\infty+i\sigma}^{+\infty+i\sigma}\frac{\exp(-i\omega t)}{\varepsilon(k,\omega)}\left[\int_{-\infty}^{+\infty}\frac{g(v)dv}{\omega - kv}\right]d\omega. \qquad (3.20)$$

This expression gives a complete solution to the problem of small plasma oscillations generated by an initial disturbance $g(v)$. We see that in general there is no definite dependence of ω upon k: for a given k the integration is performed over all the frequencies. But if $g(v)$ is not singular, the asymptote of integral (3.20) at large t will be determined only by the zeros of $\varepsilon(k,\omega)$. In fact, $\varphi(t) \sim \exp(-i\omega_k t)$ with ω_k standing for a root of the equation $\varepsilon(k,\omega_k) = 0$. Thus, when t is large, the oscillatory mode defined by

$$\varepsilon(k,\omega_k) = 0 \qquad (3.21)$$

is deduced from the solution (3.20). Since the integration in (3.20) is to be performed along a horizontal line in the upper half-plane, one should assume ω to be in the upper half-plane when evaluating ε by means of (3.15). More definitely, one should assume

$$\frac{1}{\omega - kv} \to \frac{1}{\omega - kv + i\nu} \to \frac{P}{\omega - kv} - \pi i\delta(\omega - kv)$$

when ω is close to the real axis. In the above expression the notation P is used to denote the principal value. This rule of pole circumvention is usually called the Landau rule. With regard to this rule the dielectric permeability in (3.15) proves to be complex:

$$\varepsilon = 1 + \frac{4\pi e^2}{km_e}\int\frac{P}{\omega - kv}\frac{\partial f_0}{\partial v}dv - \frac{4\pi e^2}{km_e}\frac{\pi i}{|k|}\frac{\partial f_0}{\partial v}\Big|_{v=\omega/k}. \qquad (3.22)$$

The emergence of the imaginary part of ε corresponds to the Landau damping, which is proportional to the derivative $\partial f_0/\partial v$

at $v = \omega/k$. At this point the particle velocity coincides with the phase speed of the wave. It may be stated that the Landau damping is connected with the wave energy absorption by the resonant particles. If the electrons are subjected to the Maxwellian temperature distribution law, the dielectric permeability given by (3.22) can be represented in the form

$$\varepsilon = 1 + \frac{m_e \omega_{pe}^2}{k^2 T}\left[1 + i\sqrt{\pi} z W(z)\right], \tag{3.23}$$

where T is temperature, ω_{pe} the Langmuir frequency, $v_{T_e} = \sqrt{2T/m_e}$ the mean thermal velocity, $z = \omega/kv_{T_e}$, and $W(z)$ is defined by

$$W(z) = \exp(-z^2) + \frac{2i}{\sqrt{\pi}} \int_0^z \exp(-z^2 + \xi^2)\, d\xi. \tag{3.24}$$

Equating (3.23) to zero, one can find the real and imaginary parts of the complex frequency. For small $k \ll 1/d$, where d is the Debye radius, the effective value of z proves very large in comparison with unity, $z \gg 1$, and integral (3.24) can be evaluated via an expansion in inverse powers of z. After some simple algebra we find that the expression for the real part of the frequency has the form

$$\omega^2 = \omega_{pe}^2 + \frac{3T}{m_e} k^2, \tag{3.25}$$

while its imaginary part, i.e., the damping decrement γ, is equal to

$$\gamma = -\omega_{pe} \sqrt{\frac{\pi}{8}} \frac{1}{(kd)^2} \exp\left\{-\frac{1}{2(kd)^2}\right\}. \tag{3.26}$$

As we see, the damping decrement is exponentially small at low kd.

At very large t the pole $\omega - kv = 0$ also makes a certain contribution to the integral (3.20). This pole corresponds to the free propagation of particles from the area of initial disturbance, so that the dependence of the potential φ upon the time may be rather complex.

3.1.4. Van Kampen waves

Thus, the Langmuir waves generated by an initial disturbance must be damped in time. But this does not mean that undamped eigenmodes do not exist in a plasma. Indeed, let us return to expression (3.14) that relates the disturbance of the distribution function f_1 to the potential φ. According to the Landau rule, a small imaginary part should be added to ω in the denominator, $\omega \to \omega + i\nu$, so that (3.14) may be rewritten in the form

$$f_1 = \frac{P}{\omega - kv} \frac{ke}{m_e} \frac{\partial f_0}{\partial v} \varphi_1 - i\pi\delta(\omega - kv) \frac{ke}{m_e} \frac{\partial f_0}{\partial v} \varphi_1. \qquad (3.27)$$

The character P means that the corresponding integrals involving f_1 should be treated at the singular points in the sense of the principal value. Expression (3.27) represents the disturbances of the distribution function caused by the potential φ_1 of the electric field of the wave. As is seen, the disturbance becomes larger as the velocity approaches the wave phase speed ω/k. In a small neighborhood of this point the major contribution is given by the second term, which is connected with the imaginary part as $\nu \to 0$ by the following relation:

$$\mathrm{Im} \frac{1}{\omega - kv - i\nu} = -\frac{\nu}{(\omega - kv)^2 + \nu^2} = \pi\delta(\omega - kv). \qquad (3.28)$$

Note that the simplest way to take account of the damping due to collisions is to introduce a small term $-\nu f'$ into the left-hand side of equation (3.11). This will also result in the Landau rule of pole circumvention and, as a consequence, in an expression identical to (3.27). This seems completely natural because both causes — the collisions and the finite time of evolution — result in qualitatively similar bounds on the disturbance amplitude at the resonant point.

If (3.27) is substituted into the Poisson equation (3.12), the Landau dispersion equation will be obtained. In other words, the only disturbance of the distribution function (3.27) surviving

in Langmuir oscillations at large t is that created by the wave's potential.

Assume now that not only a smooth disturbance $g(v)$ is imposed on the plasma, but a modulated beam of particles with a velocity equal to the phase speed of the wave is simultaneously injected. If the intensity and phase of that beam are chosen properly, one may completely balance the resonant electron disturbance that is described by the second term of (3.27). But in this case we obtain expression (3.14) without the δ-function, i.e., the Vlasov dispersion equation with the principal value of the integral, which describes the undamped Langmuir waves. Thus, the Vlasov solution also has a certain physical sense: it describes a wave coupled to a group of resonant particles.

Nevertheless this solution seems rather particular since it corresponds to a special choice of density of the additional particles. As shown by Van Kampen [41], equations (3.11)–(3.12) describe a much broader class of eigenmodes. In order to find these oscillations we have to eliminate the incorrectness which was admitted in the derivation of expression (3.14) for a perturbation of the distribution function. Naturally, if ω is a real number, the homogeneous equation for f_1, $(\omega - kv)f_1 = 0$, has a non-trivial solution $\lambda\delta(\omega - kv)\varphi_1$, where λ is an arbitrary constant depending upon ω and k. This solution of the homogeneous equation must be added on to (3.14). Assume, for definiteness (it is obviously not a restriction), that the integral of $1/(\omega - kv)$ is treated in the sense of the principal value. Then instead of (3.14) we must write the following relation for the eigenmodes:

$$f_1 = \frac{P}{\omega - kv}\frac{ke}{m_e}\frac{\partial f_0}{\partial v}\varphi_1 + \lambda\delta(\omega - kv)\,\varphi_1. \tag{3.29}$$

Substituting this expression into the Poisson equation, we find the dispersion law:

$$1 + \frac{4\pi e^2}{m_e k}\int \frac{P}{\omega - kv}\frac{\partial f_0}{\partial v}\,dv + \frac{4\pi e}{k^2|k|}\lambda = 0, \tag{3.30}$$

Since this equation contains two unknown values, ω and λ, it does not give any definite relation between ω and the wavenumber k. It should be treated as an equation that defines λ at a given ω. This means that the frequency may be of an arbitrary value at any given k, i.e., that the spectrum of eigenvalues ω is continuous. In other words, for any given frequency ω one can find such a density λ of the resonant particles that the solution will have the form of an undamped wave with frequency ω. This is a Van Kampen wave. A Van Kampen wave looks like a modulated beam of particles propagating with a velocity equal to the phase velocity $v_p = \omega/k$ of the wave. This beam moves together with the polarized cloud arising as the result of its interaction with the plasma electrons. The perturbation f_1 within this cloud is described by the first term of (3.29).

If the perturbation wavelength is large ($kd \ll 1$) and its frequency ω is close to the plasma frequency from (3.23), the value of λ defined by (3.30) is very small because the sum of the first two terms in (3.30) is close to zero. In that case we are dealing with a plasma wave supplied with a small quantity of resonant particles. So it is reasonable to refer to the solutions appreciably different from Langmuir waves as Van Kampen waves. For these waves the second term of (3.29) is greater or comparable with the first. Therefore, we may say that the eigenmodes of an electron plasma may be divided into two classes: Van Kampen waves (modulated beams) and Langmuir waves.

Van Kampen showed that the system of functions (3.29) is complete, i.e., that any initial disturbance $g(v)$ can be represented as a sum of these functions. Hence, all the waves excited in a plasma may be regarded as a superposition of eigenmodes. The corresponding superposition coincides completely with Landau's solution.

Van Kampen waves account for the Landau damping. Indeed, let us consider a disturbance with a characteristic wavelength much less than the Debye radius. Such a disturbance is a superposition of Van Kampen waves, and it is not surprising that it will be damped in time due to the phase divergence of

the oscillations involved. One may say that this damping is a consequence of the continuity of the eigenvalue spectrum.

The picture is a bit more complicated for Langmuir waves, because most of the particles move coherently within the area filled with oscillations. But the resonant particles are permanently diverging, and their contribution to the collective motion is so large that they do completely "drag" the Langmuir wave apart into phases, i.e., the wave is damped again. It is easy to notice that this is again a consequence of the continuity of the eigenvalue spectrum.

3.1.5. Nonlinear Landau damping

The previous treatment of nonlinear waves was rather general and could be applied to any continuous medium. We took no account of the marked characteristics of a plasma, whereas it appears to be a system of many identical particles. In fact, a plasma is a system of particles subjected to long-range forces and, therefore, nonlinear collective phenomena include not only processes of wave-wave interactions, but of wave-particle interaction as well. The so-called nonlinear Landau damping [42]–[45] serves as the simplest, but important example of interaction of particles with a wave of finite amplitude.

Let us begin with a Langmuir wave of a small, but not infinitesimal, amplitude. Suppose that the wavelength is sufficiently large in comparison with the Debye length d, so that the phase velocity $v_p = \omega/k$ of the wave essentially exceeds the thermal motion velocity. On these assumptions, one the one hand, we may consider the correction of the frequency to be small and put $\omega = \omega_{pe}$; on the other hand we may assume that there are only few resonant particles.

At first we shall show that a plane simple-harmonic Langmuir wave of finite amplitude can easily travel in a cold plasma. For this case the electron motion can be described by the La-

grange equations

$$\frac{d^2}{dt^2}\xi(x,t) = -\frac{e}{m_e}E(x+\xi,t), \tag{3.31}$$

where ξ stands for the displacement of an electron from its equilibrium position at x. But a displacement of an electron layer, say to the right, will produce an electric field E. This field will be determined by the excess of ions on the left of the given layer, i.e., $E(x+\xi,t) = 4\pi e n_0 \xi$, and equation (3.31) takes the form

$$\frac{d^2\xi}{dt^2} = -\omega_{pe}^2\,\xi \tag{3.32}$$

with ω_{pe} being the Langmuir electron frequency. Thus, we see that in the one-dimensional case the electrons oscillate harmonically with a finite amplitude at the frequency ω_{pe}, as long as the amplitude remains so small that there are no trajectory intersections and multi-flow movements.

Proceeding now to the problem of nonlinear damping, we may assume that the Langmuir wave in a cold plasma has a small, but finite amplitude, and its potential is simple-harmonic: $\varphi = \varphi_0 \cos(\omega_{pe}t - kx)$. In view of thermal motion this wave can interact with resonant particles whose velocities are nearly equal to the phase velocity ω_{pe}/k of the wave. If the phase velocity is essentially greater than the average thermal velocity, the number of resonant particles is exponentially small and, therefore, one may expect that they will produce a negligible effect on a wave of finite amplitude. Thus, we can apply the perturbation theory. At first we shall assume the wave to be stationary and shall consider the behavior of resonant particles in its field, and only then shall we try to give an account of their backward influence on the wave itself.

Thus, let us consider resonant particles in a stationary wave. It appears to be convenient to utilize the frame of reference moving together with the wave. When refered to these coordinates, the wave itself looks like a stationary disturbance of the electric field: $\varphi = \varphi_0 \cos kx$. Electrons immersed in such

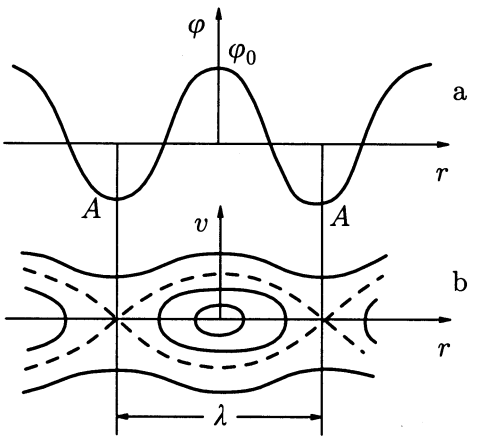

Fig. 3.1. Phase trajectories of electrons in a wave.

a field may be divided into two groups: the trapped ones that oscillate in the vicinity of the potential maxima, and the transitory ones, whose energy is high enough to fly over the potential humps.

It is useful to get a view of the particle motion with the help of the phase plane (x, v). Figure 3.1a presents the dependence of φ on x, while Fig. 3.1b shows the phase trajectories of the electron motion in the wave field. The electrons with small velocities are trapped by the wave. For instance, an electron trapped in the neighborhood of $x = 0$, is subjected to the force

$$-eE = e\frac{\partial \varphi}{\partial x} = -e\varphi_0 k \sin kx \approx -e\varphi_0 k^2 x,$$

and hence it oscillates with the frequency

$$\Omega = k\sqrt{\frac{e\varphi_0}{m_e}}. \tag{3.33}$$

As the amplitude of electron oscillations grows, the frequency decreases, becoming zero on the border (the dashed line in Fig. 3.1b) separating the trapped and the transit particles, since

the corresponding electrons can stay at the point A of the minimum potential energy $-e\varphi$ infinitely long. The average velocity of the transitory particles is also very small near the separatrix, increasing with the distance from it.

Since the total energy of an electron is

$$\frac{m_e(v - v_p)^2}{2} - e\varphi = \text{const}$$

and the difference $v - v_p$ is zero at point A, the dimension of the finite motion area at $x = 0$ is obviously equal to $\Delta v = 4\sqrt{e\varphi_0/m_e}$, i.e., it decreases with the amplitude essentially more slowly than it would in the case of a linear dependence. This means that even a wave of a very small amplitude can trap a lot of particles.

Let us now consider the evolution of the resonant particle distribution function f. We may apparently neglect the small linear perturbations of f in the area of interaction, since the dimension Δv of this area is not large. Thus, we assume that the initial distribution of the resonant particles coincides with the unperturbed one. Its dependence on v is depicted in Fig. 3.2a by the solid line.

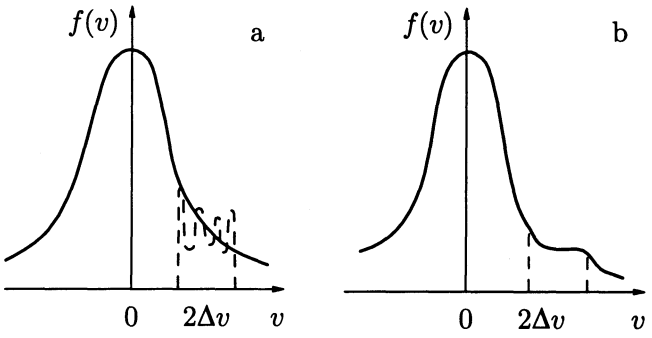

Fig. 3.2. Formation of small-scale oscillations (a) and of a "plateau" (b) on the distribution function profile.

Now let us regard the Vlasov equation

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} + \frac{e}{m_e}\frac{\partial \varphi}{\partial x}\frac{\partial f}{\partial v} = 0 \tag{3.34}$$

as a conservation law for some substance with density f flowing in the phase plane. Since the velocity v does not depend on x nor the acceleration on v, we apparently have

$$\frac{\partial v}{\partial x} + \frac{\partial}{\partial v}\left(\frac{e}{m_e}\frac{\partial \varphi}{\partial x}\right) = 0, \tag{3.35}$$

i.e., this substance turns out to be incompressible. Therefore, the quantity f is constant on the stream-lines. Since we are aware of the function f and of the stream-lines' initial shape, we can easily reveal the evolution of f.

Let us start off with the trapped particles. We hatch the area of small velocities where the function f is relatively large (Fig. 3.3a). Since the trapped particles oscillate with a frequency of the order of Ω, the internal part of the corresponding area in the phase plane will rotate, so that in a half period the picture will be transformed into Fig. 3.3b. The particles on the separating lines do not revolve and these lines remain unchanged, whereas most of the particles appear to be permutated — the density of the fast electrons becomes higher than the density of the slow ones, i.e., the sign of the derivative $\partial f/\partial v$ is reversed for the trapped particles. Following the process, we shall see that this sign reversal will occur periodically, so that the picture will more and more resemble small-scale oscillations with increasing time (see Fig. 3.3c and the dashed line in Fig. 3.2a). The transitory resonant particles near the separatrix are mixed in a similar manner (see the dotted shading in Fig. 3.3).

Due to the "mixing effect" the distribution f soon (in comparison with the average collision frequency) becomes so broken, that Coulomb collisions come to the fore. It is known that Coulomb collisions lead to diffusion in the velocity space, therefore the effective relaxation time of small-scale oscillations in the

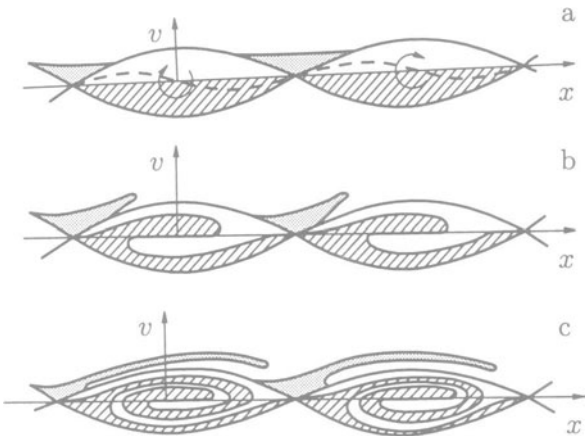

Fig. 3.3. Motion of resonant particles in a wave field.

phase space is much less than the average collision time. Because of the mixing and Coulomb collisions the distribution function will average out in the resonance area to a smooth plateau (see Fig. 3.2b). But the wave damping due to collisions will be weak for a long time, and all this time the wave will be almost periodical.

Note that this wave corresponds to a stationary solution of the Vlasov equation, and that its singularity at the point $\omega - kv = 0$ is canceled by the emergence of a plateau on the distribution function profile with $\partial f_0/\partial v = 0$ at the resonance point. This cancellation corresponds to the choice of the principle value of the singular integral in the dispersion equation (3.15).

Thus, the evolution of a wave of a finite amplitude can result in a stationary wave. This steady state regime occurs when the distribution function becomes constant (with an accepted accuracy) on the stream lines in the phase space.

This steady wave can be generated from the outset. To do so it would have been sufficient to implement an appropriate correction of the distribution of the resonant particles, such that

it should remain unchanged with respect to the frame of reference moving with the wave. For instance, if the distribution function is constant in the domain filled with trapped particles (i.e., inside of area encircled by the separatrix in Fig. 3.3), their movement does not vary it (a similar picture also holds for the transitory particles near the separating line). We see that the stationary Langmuir waves correspond to the partitions of the resonant particles, which rapidly evolve to a uniform distribution in the resonance area. These waves are called the BGK waves after Bernstein, Green and Kruskal [46]. If the number of resonant particles is increased, we shall get a wave coupled to a spatially modulated beam. This is the transition to the Van Kampen wave. In fact, the Van Kampen waves correspond to spatially modulated beams running through the plasma.

On considering the motion of resonant particles, we can proceed with their influence on the wave. Apparently the wave amplitude will decrease (Fig. 3.4) during the first half of the resonant particle oscillation period (Fig. 3.3b), for they borrow energy from the wave. Then the resonant particles will lose energy and the wave amplitude will grow, and so on. In time the distribution function will be averaged out to a constant due to the resonant particle oscillations at frequencies lying in the band from zero to Ω, the oscillations of the wave amplitude will cease and its amplitude will tend to a certain value φ_∞ as $t \to \infty$.

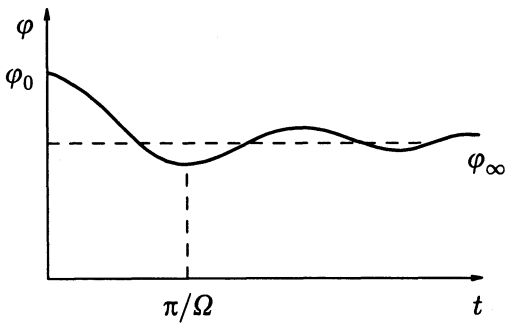

Fig. 3.4. Temporal evolution of a wave of a finite amplitude.

The fewer the number of the resonant particles, the nearer φ_∞ is to φ_0. If the initial amplitude φ_0 of the wave tends to zero, the period of the amplitude oscillations will grow infinitely (put $\Omega \to 0$ in Fig. 3.4), i.e., there are no oscillations in the linear approximation at all and the amplitude decreases monotonically. Certainly we can not regard the amplitude as a constant, as we did it before while considering the motion of trapped particles in the field of a wave of finite amplitude. Moreover, the trapped particles themselves do not oscillate — their initial displacement, shown in Fig. 3.3a by the dashed line, leads to an essential perturbation of the resonant particle density, such that the wave monotonically decays. One may presume that a weak wave appears to be overloaded with resonant particles.

It is not difficult to evaluate the critical value of the amplitude corresponding to the transition from a linear wave to a nonlinear one. This transition obviously occurs when the wave energy density

$$\mathcal{E} = \frac{E_0^2}{8\pi} = \frac{k^2\varphi_0^2}{8\pi}$$

becomes of the order of the density of the energy absorbed by the resonant particles due to the averaging of the distribution function. The extent of the resulting plateau is about $\Delta v \sim \sqrt{e\varphi_0/m_e}$, while the shift of the energy of a resonant electron caused by the increase Δv of its velocity is apparently equal to $m_e\omega\,\Delta v/k$, for these electrons move at a speed nearly equal to the phase velocity ω/k of the wave.

On plateau formation the distribution function undergoes the change

$$\Delta f \sim \Delta v \left.\frac{df_0}{dv}\right|_{v=\omega/k}$$

and hence the change of the density of the trapped particle energy is of the order of

$$\Delta\mathcal{E} \sim m_e\frac{\omega}{k}\Delta v(\Delta f\Delta v) \sim m_e\left(\frac{e\varphi_0}{m_e}\right)^{3/2} v\left.\frac{df_0}{dv}\right|_{v=\omega/k}. \qquad (3.36)$$

Comparing (3.36) with the wave energy density in a plasma with a Maxwellian distribution of electrons, we obtain the condition that the Langmuir wave be linear:

$$\sqrt{\frac{e\varphi_0}{T_e}} \ll \frac{1}{(kd)^4} \exp\left(-\frac{1}{2k^2d^2}\right). \tag{3.37}$$

Now it is evident that if kd is extremely small, a wave of even a very small amplitude can not be considered linear. This is related to the fact that the number of resonant electrons for these waves is exponentially small and, therefore, the wave damping is negligible. On the contrary, for $kd > 1$ any wave may be regarded as linear even at moderate values of $e\varphi_0/T_e$.

Our consideration of finite amplitude wave damping clarifies the problem of temporal reversibility. This problem may be posed as follows. On the one hand, the kinetic equation with a self-consistent field is completely reversible with respect to time: it is not changed by the substitution $-t$ for t, provided the velocities of all particles have been reversed. Therefore, it seems to describe only reversible processes. But on the other hand the exponential decay of linear waves in a system of charged particles apparently reveals the irreversibility: over a sufficiently large time the wave may be lost! This problem can be solved with the help of Fig. 3.3.

If there were no collisions, the distribution function evolution would be completely reversible. Namely, if the velocities of all the particles were reversed, the spiral in Fig. 3.3c would untwist. Then it would pass through the initial state and begin to twist again in the opposite direction. Correspondingly, the amplitude $\varphi_0(t)$ would oscillate at first, then, at $t = 0$, φ_0 would reach its maximum value, and later oscillate and tend to the saturation value φ_∞.

A similar process would take place in the case of a wave of small amplitude, i.e., in the case of a large quantity of resonant particles. Certainly there can be no trapped particles now, for the wave decays in such a short time that the particles fail to swing from their initial positions to the turning points: by that

time the potential well would disappear. However, the resonant particles pull apart the initial components of the perturbation, permanently borrowing the wave energy. When the oscillations finally cease, their energy is transferred to the resonant particles. But this process also appears to be reversible. The medium stores information about the initial disturbance for a long time (more precisely, as long as collisions have no role), and if the particle velocities were reversed, the process would run backward. Certainly one can not make the particles reverse their velocities, but the stored information about the initial oscillations provides echo effects which will be considered somewhat later.

3.1.6. Experiments on Landau damping

In all appearance the first detailed and straightforward investigation of the Landau damping for longitudinal oscillations was the experiment of Malmberg, Wharton and Drummond [47]. In this experiment the spatial damping of the longitudinal electron wave of the form $\exp(-i\omega t + ik_r x - k_i x)$, excited by high frequency oscillations applied to the Langmuir probe, was measured. Plasma with a density $10^8 - 10^9\,\mathrm{cm}^{-3}$ and temperature from 5 to 20 eV was created by means of a plasmotrone. To register the oscillations a second Langmuir probe was used. The results of the measurement of the damping (the ratio of the imaginary part k_i of the wavenumber to the real part k_r) as a function of the squared ratio of the phase velocity $v_p = \omega/k$ to the average thermal velocity $v_{Te} = \sqrt{2T_e/m_e}$ are presented in Fig. 3.5. Since k_i is proportional to the damping decrement, the ratio k_i/k_r should decrease exponentially with $(\omega/kv_{Te})^2$. Apparently this dependence really exists, the experimental points being in good agreement with the theoretical curve.

To check whether this damping is really connected with the resonant electrons, the authors of the work [47] changed the potential of the butt electrodes, thus "cutting the tail" of the Maxwell distribution, i.e., getting rid of the fastest electrons. Once the boundary of the cutting had reached some value v_*,

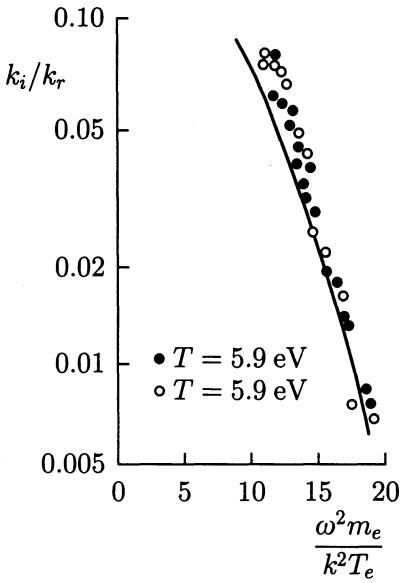

Fig. 3.5. Damping decrement dependence on the squared ratio of the phase velocity to the mean thermal velocity.

an abrupt change of the damping was observed. The value v_* appeared to be close to the phase velocity $v_p = \omega/k$ of the wave, giving convincing evidence for wave damping by the resonant electrons.

In subsequent experiments [48]–[51] a more detailed investigation of the dispersion relation for the plasma waves was conducted. In Fig. 3.6 the results of Derfler and Simonen on the dependence of the wave frequency ω and the damping (the imaginary part of the wavenumber k_i) upon k_r are presented, compared with the exact dispersion relation $\varepsilon = 0$ [see (3.15)] for arbitrary kd. (The plot in Fig. 3.6 should be treated rather as a dependence of k_i and k_r on ω). We see that the experimental points are in good accord with the theoretical curves. Thus, at present there is reliable experimental evidence not only for the damping, but also for the entire dispersion relation $\varepsilon(\omega, k) = 0$.

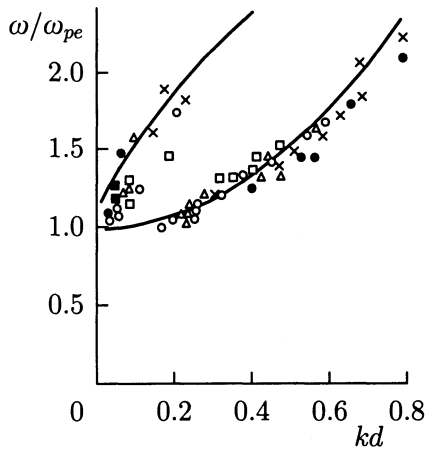

Fig. 3.6. Dispersion curves for Langmuir waves.

Later Malmberg and Wharton [52] investigated finite amplitude wave damping. Their results are shown in Fig. 3.7, where the experimentally measured dependence of the amplitude of oscillations upon the distance between the probe-radiator and the probe-receiver is shown.

The amplitude is presented in relative units. Curve 1 corresponds to the small-amplitude wave. If one does not take into account the range of distances less than 5 cm, where the plane wave is not yet formed, it will be clear that the wave decays monotonically by an exponential law.

As the amplitude of the wave increases, oscillations of the dependence of the amplitude upon the distance appear, with the period decreasing as the amplitude increases. The experiment showed that the wavenumber K of the oscillations of the amplitude increases as $\sqrt{V}$ with a change of the potential V of the probe-radiator. This dependence is in close agreement with the theoretical prediction, according to which the frequency Ω of the oscillations and, consequently, K should increase as the square root of the amplitude. Thus, the Landau damping effect may be considered reliably proved experimentally.

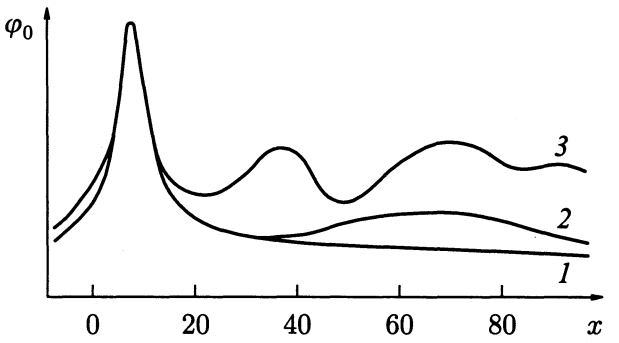

Fig. 3.7. Dependence of oscillation amplitude upon the distance between the probe-radiator and the probe-receiver for the following values of the potential (in Volts) of the probe-radiator: *1* — 0.9; *2* — 2.85; *3* — 9.

There is a question of whether it is possible to check experimentally that the Landau damping does not lead directly to the irreversibility and that in a collisionless plasma a "memory" of the damped oscillations is kept. This check was really accomplished in experiments with echo.

3.2. Echo in a plasma

3.2.1. Spin echo

At first let us recall what an echo is. The echo effect was discovered by Hahn [53] in experiments on nuclear magnetic resonance. It looks as follows. If one applies two short pulses separated by a time interval τ with a frequency close to the resonant one, a new pulse corresponding to a spontaneous emission by the nuclear spins at the resonant frequency will appear a time τ after the second pulse. This effect, called spin echo, was explained by Hahn himself.

Let us consider a system of nuclear spins in a strong magnetic field and suppose that at the initial instant all the magnetic moments are directed along the field (the axis of z in Fig. 3.9).

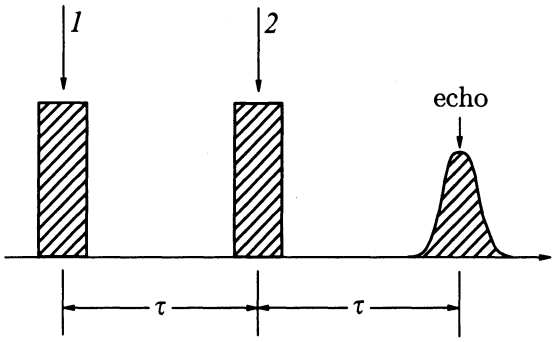

Fig. 3.8. Spin echo.

As is known, the application of a resonant high-frequency electromagnetic field leads to a deviation of the spins from the z axis. Suppose, for simplicity, that the first pulse is a 90°-pulse, i.e., that the amplitude of the pulse is chosen so that the spins alter their direction by 90° and transfer to the x axis. After the cessation of the pulse, radiation at the resonant frequency should take place due to the precession of the magnetic moment. But since the inhomogeneity of the external magnetic field is small, this radiation stops very soon. As the frequency of the precession varies slightly from point to point, the phases of the different spins will soon drift apart, and their mean distribution in the sample will become almost uniform (Fig. 3.9).

Let us now assume that a second pulse is applied at the instant $t = \tau$. The picture is simplest if the second pulse has a twofold amplitude, i.e., if it is a 180°-pulse. In this case "the fan" of spins in Fig. 3.9 simply turns by 180° about the axis of y, which is equivalent to the reversal of the time: then all the spins will revolve in the plane of "the fan" in the opposite direction and after a time τ will meet together at the x axis. Just at this instant the echo will be observed as strong radiation from the spins with the same phase.

An echo effect with smaller amplitude is also observed in the case of pulses not exactly of 90 and 180 degrees.

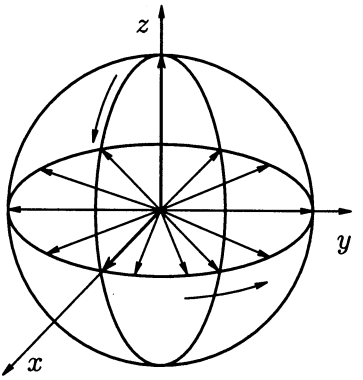

Fig. 3.9. Rotation and spread of spins under the action of a 90°-pulse.

It is easy to see that the echo effect is closely related to the existence of a continuous spectrum of the modes of oscillation. Indeed, the spreading of spins into the fan takes place just because the frequency of spin precession at every point of space has its own value $\omega(\mathbf{r})$. The frequency $\omega(\mathbf{r})$ is a continuous function of coordinates, and if it changes in space, the oscillations of spins at different points will eventually diverge in phase, so that a more and more jagged wave field with larger and larger wavenumbers will be created. Indeed, (2.65) yields

$$\mathbf{k} = \frac{\partial \omega}{\partial \mathbf{r}} t \to \infty.$$

Because of the phase divergence of the neighboring perturbations the macroscopic effect of radiation of waves through spin precession disappears.

The application of the second pulse again leads to a second order response, strongly jagged in space, but developing so that now the second-order wavenumber decreases with time and goes to zero at $\delta t = \tau$ after the second pulse.

3.2.2. A simple demonstration of the echo effect

An extremely simple demonstration of the echo effect was proposed by Vedenov and Dychne. Suppose that we have two rasters with different periods (for example, two combs with a different number of teeth per unit length). Let the wavenumbers of these rasters, corresponding to the main harmonics, be k_1 and k_2. Then, if one casts a diffused light on them (e.g., from a window), the angular distribution of the modulation of the light intensity at a distance x from each raster will be evidently equal to $\exp(iky + ik\theta x)f(\theta)$, where $k = k_1$ or k_2, and $f(\theta)$ is the angular distribution of intensity of light before the raster (Fig. 3.10).

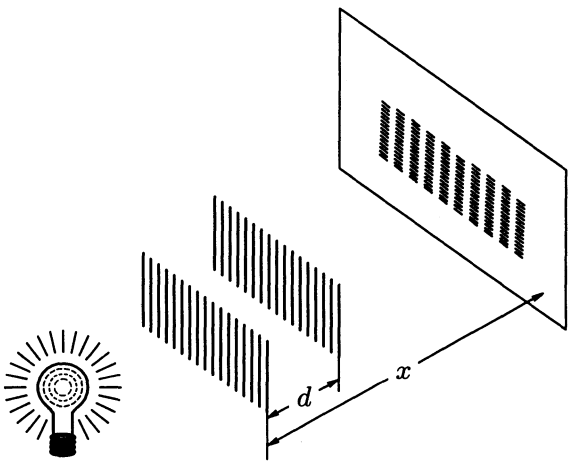

Fig. 3.10. Echo for diffused light.

As one moves away from the raster, its shadow will be more and more dispersed, because the distribution of the illuminance, proportional to $\int \exp(iky + ik\theta x)f(\theta)d\theta$, approaches a uniform distribution due to the spreading with respect to θ. But if the rasters are placed one after the other with a distance d between them so that the shadow of one raster falls on the other, then the distribution of the illuminance will be

$$\int \exp\left\{ik_1 y + ik_1\theta x \pm \left[ik_2 y + ik_2\theta(x - d)\right]\right\} f(\theta)d\theta. \quad (3.38)$$

At the distance $x = k_2 d/(k_2 - k_1)$ from the first raster the exponential function under the integral becomes independent of θ, and a moire corresponding to the difference $k_2 - k_1$ of the wavenumbers appears on a screen. From the expression $x = k_2 d/(k_2 - k_1)$ for the position of the screen it follows that the effect takes place only if $k_2 > k_1$.

From the examples considered above it is clear that the echo effect appears when damping of linear perturbations is produced by the phase dispersal. In the case of a spin echo the macroscopic average effect disappears due to the accumulation of phase differences at different points, whereas in the case of rasters it is due to the spreading along the y-axis of different light beams. But a completely analogous effect takes place in the case of Landau damping: eventually the electric field of the wave decays due to spreading of the resonant particle beams in the velocity space. Thus it is natural to expect an echo type effect for plasma waves. But first we shall get acquainted with the cyclotron echo in plasma, which is more similar to the spin echo phenomenon.

3.2.3. Cyclotron echo in a plasma

Cyclotron echo in a plasma was discovered by Hill and Caplan [54]. The experiment was set in the following way [54]–[56]. Two pulses of high-frequency electromagnetic waves with a time interval τ between them were transmitted across the core of a decaying plasma created by a high-frequency discharge. The pulses were applied when there was no current in the plasma and it was sufficiently quiet. The plasma was placed into a magnetic field of about 3 kGs, and the generator frequency was close to the cyclotron frequency of the electrons. The amplitude of the pulses was sufficiently high for the electrons to acquire an energy considerably exceeding the thermal energy. After the transmission of the two pumping pulses a train of pulses was observed at the cyclotron frequency from the plasma with the time interval τ between them (Fig. 3.11).

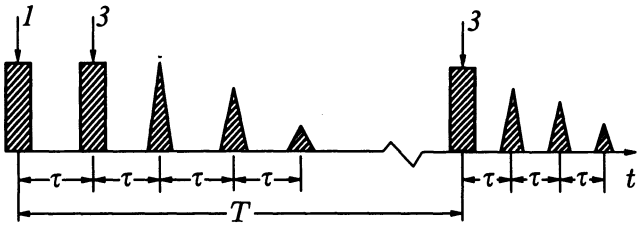

Fig. 3.11. Cyclotron echo.

In addition to this two-pulse echo, a three-pulse echo was observed by the authors, when the third pulse was applied after a long delay T, considerably exceeding the mean collision time of electrons with the atoms of the neutral gas, but smaller then the energy relaxation time for the electrons. As is shown in Fig. 3.11, after that a series of pulses from the plasma was again observed.

An explanation of the cyclotron echo mechanism and its main peculiarities was given by Gould [57] (see also [58]). In principle the cyclotron echo is similar to spin echo. While under nuclear magnetic resonance the action of a high-frequency field leads to deviation of the spins from the z axis, under cyclotron resonance an accumulation of transverse energy by the electrons takes place. Provided that the duration of the pulse is not too large, it has a sufficiently wide spectrum and we may assume that all the electrons (initially being cold) acquire the same energy even if some inhomogeneity of the magnetic field is present. Let the position of the electrons just after the first pulse correspond to the point $(v_p, 0)$ in the space of velocities (v_x, v_y), i.e., a plane perpendicular to the magnetic field (Fig. 3.12a). In the following instants the electrons will orbit with the cyclotron frequency, and in the presence of a small inhomogeneity of the magnetic field they will very soon reach, on average, a homogeneous phase distribution throughout the sample.

Let us consider a group A of electrons having the same phase ξ just before the second pulse (Fig. 3.12b). To be more precise, let this group consist of all the electrons with the phase

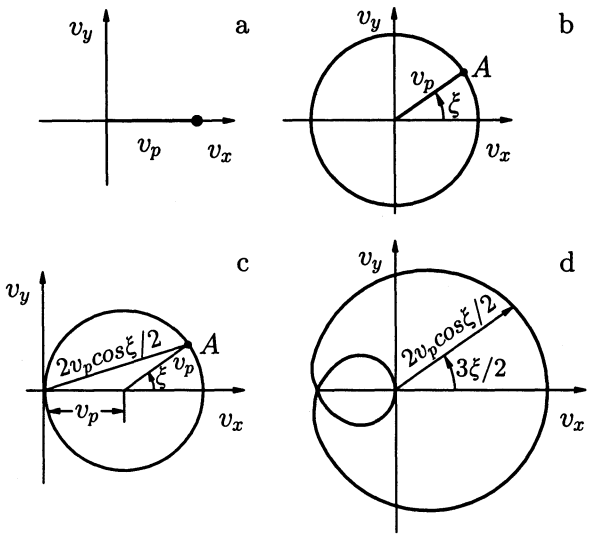

Fig. 3.12. Distribution of velocities of the electrons. (a) — for a two-pulse echo just after the first pulse; (b) — before the second pulse; (c) — just after the second pulse; (d) — at the moment of the first echo.

$\xi+2\pi n$, where n is an integer. During the second pulse the electrons acquire or lose energy in accordance with the values of their phases, and the action of the pulse can be taken into account by a mere shift of all the distribution function by the same value v_p, as was done for the first pulse. In this case the group A of electrons that were picked out will have the velocity $2v_p\cos(\xi/2)$, as is shown in Fig. 3.12c. Since this group includes electrons from different points of space, the particles of this group will eventually diverge in the space of velocities. However, since the electrons are rotating with the same angular velocity as before, in a time τ they will group together again. For example, at the instant $\delta t = \tau$ after the second pulse, i.e., just at the first echo signal, the particles of this group will meet together at point A with the phase $3\xi/2$, and all the distributions of velocities of the electrons will turn into the curve shown in Fig. 3.12d.

This means that the electrons in the process of cyclotron oscillations are potentially able to create an echo effect. But in fact the distribution in Fig. 3.12d does not give any echo: it is easy to check that under this distribution the total current expressed through the mean values $\langle v_x \rangle$, $\langle v_y \rangle$ is equal to zero. It is no wonder, because the system considered is linear — we simply summed up the effects of the two pulses. But the echo is substantially a nonlinear effect: the total signal is not a superposition of the responses to each pulse, it is determined by both pulses in their nonlinear interplay.

In fact, for a cyclotron echo to appear, some nonlinear mechanism has to come into play to violate the picture of Fig. 3.12d and to lead to the absence of exact compensation of all the currents. For example, a nonlinear mechanism of this kind can be provided by the dependence of mass on velocity (the relativistic effect), by the nonlinearity of the wave and so on. But the most simple and natural mechanism is just a dependence of the collision frequency on the velocity. The collisions of the electrons with the atoms of a neutral gas or between themselves lead to the departure of some electrons from coherent motion. If the collision rate depends upon the velocity, the number of electrons departing from the curve of Fig. 3.12d will be different at distinct points and, consequently, a macroscopic current with the cyclotron frequency will appear, i.e., an echo will occur. The experimental data are in good agreement with this mechanism [59, 60]. The three-pulse echo is also explained by it.

As was stated before, in the case of the three-pulse echo the electrons completely lose the mean momentum by the instant of the third pulse, but do not have enough time to lose their energy (in a process of elastic collisions with the atoms of the neutral gas the relaxation time for the energy is m_i/m_e times more than the relaxation time for the momentum). This means that before the third pulse the electrons of group A in Fig. 3.12 are distributed uniformly on a sphere with radius $2v_p \cos(\xi/2)$ (Fig. 3.13a). Just after the third pulse this sphere will be found displaced by the value v_p (Fig. 3.13b), and then the phases of

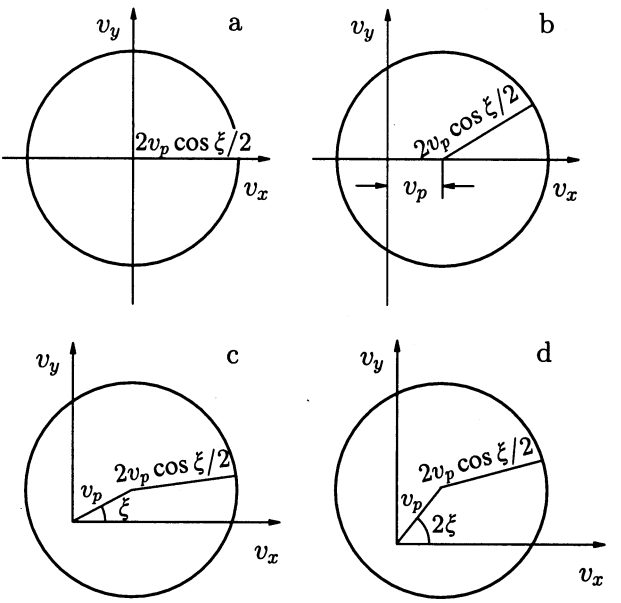

Fig. 3.13. Distribution of velocities of the electrons. (a) — for a three-pulse echo before the third pulse; (b) — just after the third pulse; (c) — at the instant of the first echo; (d) — at the instant of the second echo.

the electrons of this group will diverge. However, in a time τ a grouping will take place in the space of velocities with a phase shift ξ from pulse to pulse (Fig. 3.13, c and d). But again some mechanism of dependence of the collision rate on velocity must be activated for the total current to be nonzero. And the effect itself, as it is easily seen, reaches the maximum when the collision rate reaches the value of about the inverse time τ^{-1} between the pulses. These conclusions are also in satisfactory agreement with the experimental data.

The cyclotron echo is of great interest by itself as a new nonlinear effect in a plasma, but it may be also used for the purposes of diagnostics, for instance for investigation of the relaxation processes in a plasma.

3.2.4. Echo for plasma waves

Let us return now to plasma waves. As was shown by Gould, O'Neil and Malmberg [59], the echo effect occur with these, too. At first let us consider the more simpler case of modulated Van Kampen beams, which propagate in the form of plane waves along the x axis. These waves can be excited in a plasma with the help of a grid, to which a periodic signal is applied. Let the frequency of this signal be considerably greater than the plasma frequency ω_{p_e}. Then the dielectric permittivity ε of the plasma may be considered to be equal to unity, i.e., the polarization of the medium by the waves may be neglected. In this case the perturbation of the partition function of the electrons that cross the grid is merely equal to

$$f_1(x, v, t) = f_1(v) \exp\left(-i\omega_1 t + i\omega_1 \frac{x}{v}\right), \tag{3.39}$$

where $f_1(v)$ is the perturbation of the distribution function near the grid. The form of the perturbation (3.39) follows from the fact that at a large distance from the grid, f_1 should satisfy the equation of free motion of particles:

$$\frac{\partial f_1}{\partial t} + v\frac{\partial f_1}{\partial x} = 0. \tag{3.40}$$

The distribution function (3.39) may be treated as a set of modulated beams. Near the grid all these beams oscillate coherently, but as one moves away the phases of the beams with different velocities will diverge considerably, so the charge density

$$\rho = -e \int f_1 dv \tag{3.41}$$

has to decrease rapidly as x increases (f_1 becomes a rapidly oscillating function of v). This means that the oscillations of the electric potential should decrease rapidly as one moves away from the grid, and this effect is quite similar to the dephasing of magnetic moments before the second pulse when the magnetic echo is observed.

Assume that the second grid is placed at a distance d from the first grid, and that an alternating potential with frequency $\omega_2 > \omega_{p_e}$ is applied to it. Then Van Kampen waves, appearing because of the perturbation of f, will also propagate from the second grid. But the second grid will also modulate function (3.39), so a nonlinear response at the combinative frequency will appear:

$$f_2(x,v,t) = f_2(v)\exp\left[-i\omega_1\left(t-\frac{x}{v}\right)\pm i\omega_2\left(t-\frac{x-d}{v}\right)\right]. \tag{3.42}$$

For $x = d\omega_2/\omega_2 - \omega_1$ the exponent corresponding to the frequency $\omega_1 - \omega_2$ becomes independent of v.

That means that at the point

$$x = d + \frac{\omega_1}{\omega_2-\omega_1}d = \frac{\omega_2}{\omega_2-\omega_1}d \tag{3.43}$$

appreciable oscillations of the charge density with frequency $\omega' = \omega_2 - \omega_1$ should be observed. In other words, an electric probe placed at this point would reveal an echo of the combinative frequency $\omega_2 - \omega_1$. Just this kind of response is observed in experiment (Fig. 3.14).

This conclusion also holds in the case of $\omega \sim \omega_{p_e}$, because it is connected only with the behavior of the resonant particles. Of course the quantitative formulation of the echo effect changes in this case because one should take into account the disturbances for other electrons and, correspondingly, the expression for the second-order response will include ε — the dielectric permittivity of the plasma. One may treat the ion-sound wave echo, the echo for the transverse waves and so on in a similar way. In addition to the second-order echo, plasma echo effects of higher orders also exist.

The theoretically predicted echo effect for plasma waves was first observed by Malmberg, Wharton, Gould and O'Neil [60] for electron oscillations, and then investigated theoretically and in experiments for various types of waves in a plasma

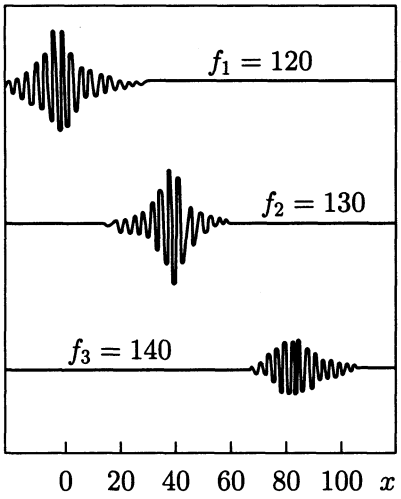

Fig. 3.14. Third order echo in a plasma (i.e., $2f_2 - f_1 = f_3$) [60].

[61]–[66]. The echo effects show that the collisionless damping of waves in a plasma is in essence a reversible process connected with the dispersal of the phases of resonant particles and with the corresponding decrease of their coherent contribution to the perturbation of the electric field.

3.3. Scattering of waves by particles

Considering nonlinear Landau damping, we restricted ourselves to one harmonic wave of finite amplitude. A more complicated picture appears in the case when there may be several waves simultaneously in the plasma. For instance, two waves traveling in one direction with close phase velocities may exchange with the trapped particles, and so a rather complicated indirect interaction between them will be brought about.

But if the waves are not parallel, then this interaction will be absent or considerably attenuated. Of course it will also be very weak if both waves have large phase velocities and there are few trapped particles. However, in this case the interaction

of finite-amplitude waves through resonant particles may also take place, though not for the direct resonance $\omega - \mathbf{k} \cdot \boldsymbol{v} = 0$, but for the combinative one. Let us explain why. As we have seen before, beats with combinative wavenumbers $\mathbf{k} \pm \mathbf{k}'$ and frequencies $\omega_{\mathbf{k}} \pm \omega'_{\mathbf{k}}$ appear in the process of superposition of two weakly nonlinear waves due to the nonlinear terms. For instance, in the case of nearly coinciding wavevectors $\mathbf{k}$ and $\mathbf{k}'$, the wave $\mathbf{k} - \mathbf{k}'$ corresponds to the moire in Fig. 2.14. Even if a free wave with frequency $\omega_{\mathbf{k}} \pm \omega_{\mathbf{k}'}$ and wavevector $\mathbf{k} - \mathbf{k}'$ cannot propagate in the plasma, i.e., if the decay is prohibited by the laws of conservation of momentum and energy, the forced oscillations will, nevertheless, be present. These oscillations may also interact with the resonant particles, if the velocity of these particles coincides with the phase velocity of the combinative wave, i.e., if

$$\omega_{\mathbf{k}} - \omega'_{\mathbf{k}} - (\mathbf{k} - \mathbf{k}') \cdot \boldsymbol{v} = 0 \tag{3.44}$$

[to be definite we have chosen the sign to be minus, a wave with $(-\mathbf{k}', -\omega'_{\mathbf{k}})$ would correspond to a plus sign].

As the amplitude of forced oscillations for weakly nonlinear waves is, generally speaking, small, the corresponding interaction of the particles with these oscillations will be analogous to the linear Landau damping. In other words, if the derivative $\partial f/\partial v$ at the point of resonance is negative, the particles will have to take energy away from the combinative wave. But the amplitude of this wave is simply proportional to the product of the amplitudes of the initial waves $a_{\mathbf{k}}$ and $a'_{\mathbf{k}}$. So one can change the amplitude of the combinational wave only by changing the amplitudes of the initial waves, i.e., the numbers $N_{\mathbf{k}}$, $N_{\mathbf{k}'}$ of the initial waves.

These waves, as we know, have a total energy $\mathcal{E} = \omega_{\mathbf{k}} N_{\mathbf{k}} + \omega_{\mathbf{k}'} N_{\mathbf{k}'}$ (the frequencies are assumed to be positive) and momentum $\mathbf{P} = \mathbf{k} N_{\mathbf{k}} + \mathbf{k}' N_{\mathbf{k}'}$. In the process of interaction with the resonant particles some energy and momentum should be transferred to the particles, with the momentum being transferred only in the direction of propagation of the combinative wave, since the resonant field of the wave acts upon the particles in

this direction. Thus, the component of the wave momentum transverse to the vector $\mathbf{k} - \mathbf{k}'$ should be conserved, i.e.,

$$(\mathbf{k} - \mathbf{k}') \times \delta\mathbf{P} = (\mathbf{k} \times \mathbf{k}')\,\delta N_{\mathbf{k}'} + (\mathbf{k} \times \mathbf{k}')\,\delta N_{\mathbf{k}} = 0. \qquad (3.45)$$

So it is evident that $\delta(N_{\mathbf{k}} + N_{\mathbf{k}'}) = 0$, i.e.,

$$N_{\mathbf{k}} + N_{\mathbf{k}'} = \text{const}. \qquad (3.46)$$

That means that in fact we are dealing with the scattering of the wave by particles: as the number of waves in one direction increases, the number of waves in the other direction decreases by the same number. Since the change of energy in this case is equal to

$$\delta\mathcal{E} = \omega_{\mathbf{k}}\,\delta N_{\mathbf{k}} + \omega_{\mathbf{k}'}\,\delta N_{\mathbf{k}'} = (\omega_{\mathbf{k}} - \omega_{\mathbf{k}'})\,\delta N_{\mathbf{k}}, \qquad (3.47)$$

it is clear that if the energy is transferred from the waves to the particles, i.e., if the derivative $\partial f/\partial v$ is negative for the resonant particles, the scattering goes in the direction of the frequency decrease: the number of waves with higher frequency decreases, and the number of waves with lower frequency increases.

Wave-waves resonant decay (three-wave interaction) and their induced scattering by resonant particles constitute the two main mechanisms of interaction of weakly nonlinear waves in a plasma.

4. Plasma in a magnetic field

4.1. Magneto-hydrodynamics

4.1.1. MHD equations

The Swedish physicist H. Alfvén was the first to notice the extremely interesting properties that a conducting gas or liquid exhibits in a magnetic field. These peculiarities are due to fact that the motion of a conducting volume induces an electric current that interacts with the magnetic field and affects the

fluid motion. The particular characteristics of the fluid are of no importance, it is only necessary that it should be conducting. A plasma surely satisfies this condition and hence in many aspects it behaves like a conducting liquid. In particular, H. Alfvén himself applied his investigation to cosmic plasma.

Thus, we shall proceed with the phenomena that can be described on the grounds of magneto-hydrodynamics. This approximation is based on the assumption that plasma behavior is similar to the behavior of an ideal gas with the equation of state $p = 2nT$, where p stands for the pressure, T for the temperature (measured in energy units) and n for the electron or ion density (we assume the ions to be of unit charge). Since the mass m_i of an ion essentially exceeds the electron mass m_e, the mass density ρ of the plasma is approximately equal to $m_i n$. Let $\mathbf{B}$ be the magnetic field and $\mathbf{j}$ — the electric current density.

On account of the Ampère force the equation of the fluid motion takes the form

$$m_i n \frac{d\boldsymbol{v}}{dt} + \nabla p = \frac{1}{c}\mathbf{j} \times \mathbf{B}. \tag{4.1}$$

To this dynamical equation we must add, in the first place, the continuity and Maxwell equations. Since the displacement current is negligible as long as we consider slow plasma motion, i.e., as long as the fluid velocity is small in comparison with the velocity of light, we shall add the following:

$$\frac{\partial n}{\partial t} + \nabla \cdot n\boldsymbol{v} = 0, \tag{4.2}$$

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{j}, \tag{4.3}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{4.4}$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}. \tag{4.5}$$

In accordance with Ohm's law the current density $\mathbf{j}$ in a conducting medium is equal to $\sigma \mathbf{E}_*$, where $\mathbf{E}_*$ is the electric field

refered to the frame that is moving with the liquid volume. On the condition that $v \ll c$ this field is given by

$$\mathbf{E}_* = \mathbf{E} + \frac{1}{c}(\boldsymbol{v} \times \mathbf{B}).$$

Expressing $\mathbf{E}$ in (4.5) in terms of $\mathbf{E}_*$ and putting

$$\mathbf{E}_* = \frac{\mathbf{j}}{\sigma} = \frac{c}{4\pi\sigma} \nabla \times \mathbf{B},$$

we transform (4.5) into

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \mathbf{B}) + \frac{c^2}{4\pi\sigma} \Delta \mathbf{B}. \tag{4.6}$$

Equations (4.1)–(4.4) and (4.6) are important in magnetohydrodynamics. Note that we have neglected the viscosity in this approximation. As for the temperature T, its evolution can be revealed with the help of the adiabatic condition $T \sim n^{\gamma-1}$ with γ being the adiabatic coefficient. But if the space scale of plasma disturbances is small, the isothermal approximation $T = \text{const}$ will be more accurate, since the electron thermal conductivity is high.

The hydromagnetic equations describe large-scale plasma motion rather well, when the difference in motion of different particles is not important. In particular, these equations may be successfully applied to the investigation of plasma equilibrium in a magnetic field when the anisotropy is inessential, i.e., when the distribution of particle velocities is nearly Maxwellian.

Let us examine equation (4.6) more thoroughly and try to reveal its physical meaning. Note that the last term on the right-hand side describes the diffusion of the magnetic field inhomogeneities in a conductive medium caused by imperfect conductivity. The quantity $D_m = c^2/4\pi\sigma$ may be called the diffusion coefficient of the magnetic field. If the conductivity is sufficiently high, the field diffusion may be neglected and we have ideal hydrodynamics without dissipation. Equation (4.6) then takes

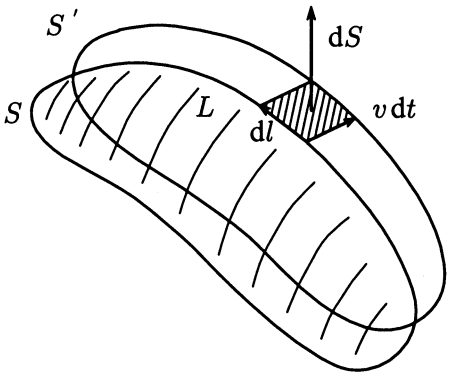

Fig. 4.1. Motion of a contour in a plasma.

the form

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \mathbf{B}). \tag{4.7}$$

In this limit equation (4.7) ensures that the magnetic lines are loaded with the mass of the plasma, or, so to speak, are frozen to the conductive medium. To verify the statement, let us consider an arbitrary closed curve L moving with the plasma. Since $\nabla \cdot \mathbf{B} = 0$, the magnetic lines cannot begin or end at any point in the interior of the plasma, and we can unambiguously define the magnetic flux $\Phi = \int_S \mathbf{B} \cdot d\mathbf{S}$ through a surface embraced by the contour L: Φ apparently will not depend on the shape of this surface and will be determined only by the contour itself. Now let us evaluate the change $\Delta\Phi$ of the flux in the time Δt:

$$\Delta\Phi = \int_{S'} \mathbf{B}' \cdot d\mathbf{S} - \int_S \mathbf{B} \cdot d\mathbf{S}. \tag{4.8}$$

Here $\mathbf{B}' = \mathbf{B}(t + \Delta t, \mathbf{r})$ is the magnetic field at the instant $t+\Delta t$, S is the initial and S' is the shifted surface (see Fig. 4.1), while $d\mathbf{S}$ is an infinitesimal area, regarded as a vector directed perpendicularly to the surface. Figure 4.1 shows that we may treat the surface S' as if it consisted of two parts: of the initial surface S and of a narrow band ΔS of width $v\Delta t$, connecting the contours L and L' and being glued to the boundary of S.

Therefore, we can represent the first term on the right-hand side of (4.8) as the sum of integrals over S and ΔS. The first of them can be added to the second integral in (4.8), and the sum appears to be equal to

$$\Delta t \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S}$$

up to terms of the second power in Δt. As for the integral over ΔS, it can be evaluated as a certain contour integral taken along L, since an infinitesimal element of area of the band ΔS is apparently equal to $d\mathbf{S} = -d\mathbf{l} \times \boldsymbol{v}$. With the foregoing notation we have:

$$\Delta\Phi = \Delta t \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} - \Delta t \int_L (\boldsymbol{v} \times \mathbf{B}) \cdot d\mathbf{l}.$$

But the last integral is equal to $\int_S \nabla \times (\boldsymbol{v} \times \mathbf{B}) \cdot d\mathbf{S}$ due to the well-known formula of vector calculus. In view of equation (4.6) for $\mathbf{B}$ we obtain $d\Phi/dt = 0$ in the limit $\Delta t \to 0$, i.e., that $\Phi = \text{const}$. Since this statement is valid for any contour L moving with the fluid, we have proved that the magnetic lines are frozen to a perfectly conducting liquid. This phenomenon essentially simplifies the analysis of the magnetic field behavior in an ideal plasma.

4.1.2. Plasma equilibrium

A plasma can easily be set into motion and in fact rarely appears to be in a steady state. Nevertheless it is worthwhile to start off with the investigation of a steady state since things are simpler. Certainly we can imagine that a uniform plasma fills all space and is in thermal equilibrium. This extremely idealized picture is frequently used for a theoretical treatment of plasma oscillations. But we are aware that in a laboratory a plasma is always inhomogeneous and that the confinement of such plasma requires the use of some forces. These forces may be either due to a magnetic field in the case of a completely ionized plasma, or to the friction caused by a neutral gas, if the ionization is weak.

We shall start off with a completely ionized plasma. To simplify things, we shall regard the plasma as an ideal gas with the equation of state $p = 2nT$, and shall consider it to be a perfect electric conductor. Under the additional assumption of a vanishing viscosity such a plasma is called ideal.

It follows from (4.1) that an ideal plasma can be in a steady state only if the pressure gradient is balanced by the Ampère force:

$$\nabla p = \frac{1}{c}\mathbf{j} \times \mathbf{B}. \tag{4.9}$$

Here $\mathbf{j}$ is the electric current density and $\mathbf{B}$ is the magnetic field.

Under the condition of perfect conductivity there are no restrictions imposed on the current density $\mathbf{j}$ with the only exception being the continuity relation:

$$\nabla \cdot \mathbf{j} = 0. \tag{4.10}$$

The magnetic field $\mathbf{B}$ is related to $\mathbf{j}$ by the Maxwell equation (4.3) and is also divergence-free due to (4.4). Equation (4.9) obviously implies

$$\mathbf{B} \cdot \nabla p = 0, \quad \mathbf{j} \cdot \nabla p = 0. \tag{4.11}$$

These relations may be regarded as the condition that the pressure is to be constant both on the magnetic field lines and on the current stream lines: there are no forces in these directions and the plasma can move freely along $\mathbf{B}$ and $\mathbf{j}$. But since ∇p is directed perpendicularly to the surfaces $p = \text{const}$, relations (4.5) imply that the magnetic and the current lines must lie on surfaces of constant pressure.

In particular, if we wish to separate the plasma from the walls with the aid of a magnetic field (this problem has emerged in connection with attempts to provide a controlled nuclear fusion process), we should form closed surfaces of constant pressure and place them one inside another in such a way that the pressure be zero on the plasma boundaries and reach its maximum value within the plasma volume.

One can fill these surfaces with field and current lines in such a way that there are no singularities (corresponding, for instance, to a conductor carrying an electric current, piercing through the surface of constant pressure) only when these surfaces are torus-like. Thus, we are faced with the problem of plasma equilibrium in a torus. This problem will be considered somewhat later.

With the aid of (4.3) we may express $\mathbf{j}$ in terms of $\mathbf{B}$ and rewrite the equilibrium equation (4.9) as follows:

$$\nabla p + \nabla \frac{B^2}{8\pi} = \frac{1}{4\pi} (\mathbf{B} \cdot \nabla) \mathbf{B}. \tag{4.12}$$

The second term on the left-hand side of this equation can be interpreted as the field pressure gradient.

The equilibrium equation can be written in a more useful form with the help of the normalized vector $\mathbf{h} = \mathbf{B}/B$, pointing in the direction of the magnetic field:

$$\nabla p + \nabla_\perp \frac{B^2}{8\pi} = \frac{B^2}{4\pi} (\mathbf{h} \cdot \nabla) \mathbf{h} = \frac{B^2}{4\pi R} \mathbf{n}. \tag{4.13}$$

Here $\nabla_\perp = \nabla - \mathbf{h} (\mathbf{h} \cdot \nabla)$ is the transverse gradient component, $\mathbf{n}$ is the unit vector of the normal direction to the field line and R is its inverted curvature. Equations (4.13) shows that because of the field lines tension the magnetic field exerts a pressure $B^2/8\pi$ on the plasma in the transverse direction and applies an additional force in the direction, normal to the magnetic lines.

Now let us consider some particular steady state configurations.

1. A cylindrically-symmetric column. Let us begin with the simplest case of plasma equilibrium when there is a cylindric symmetry. In fact, such configurations can be of large but finite dimensions and may result from plasma compression that can be caused either by a longitudinal magnetic field (the so-called θ-pinch) or by the field of an electric current, running through the plasma (z-pinch). To simplify computations, we shall consider

the plasma to be like a uniform cord of infinite length, i.e., shall neglect the edge effects.

In the cylindrical coordinates r, θ, z with z being directed along the cord's axis, the equilibrium equation takes the form

$$\frac{dp}{dr} + \frac{d}{dr}\frac{B_z^2}{8\pi} + \frac{B_\theta}{r}\frac{d}{dr}\frac{rB_\theta}{4\pi} = 0. \tag{4.14}$$

Since there is only one equation and three unknown functions p, B_z, B_θ, there must be plenty of solutions corresponding to stationary configurations.

2. θ-pinch. At first let us examine the case when there is no longitudinal current j_z and, consequently, $B_\theta = 0$. In these circumstances the plasma is confined by the azimuthal part j_θ of the current and that is why the configuration is called a θ-pinch. In view of (4.14) we have

$$p + \frac{B_z^2}{8\pi} = \text{const}, \tag{4.15}$$

i.e., the plasma pressure is balanced by the magnetic pressure. This equation shows that the plasma appears to be diamagnetic: it widens the magnetic lines out and decreases the field. If B_0 is the field outside the plasma and $p = B_0^2/8\pi$, there will be no field within the plasma and it will be confined only by the current running over its surface. In the opposite low-pressure case, $p \ll B^2/8\pi$ and the perturbation δB of the magnetic field inside the plasma is small and is apparently given by $\delta B = -\beta B/2$, with $\beta = 8\pi p/B^2$. On the assumption that $B \approx$ const and that the plasma pressure is constant within the cord and vanishes rapidly on its surface, we can integrate the equilibrium equation

$$dp/dr = -(1/c)\, j_\theta B_z$$

and formulate the equilibrium condition in the form

$$p = -(1/c) j_s B \tag{4.16}$$

with j_s being the surface current density. Equation (4.16) states that the plasma pressure is balanced by the Ampère force applied to the cord boundary.

3. z-pinch. The other particular case $B_z = 0$ corresponds to the ordinary pinch effect, i.e., to the compression of a plasma by the magnetic field of the electric current flowing along the plasma cord. The current density may be expressed in terms of B_θ:

$$j_z = \frac{c}{4\pi r}\frac{d}{dr}(rB_\theta).$$

Multiplying the equilibrium equation by $r^2 dr$ and integrating from $r = 0$ till $r = \infty$, we obtain

$$\int_0^\infty \frac{dp}{dr} r^2 dr = -\frac{1}{8\pi}(r^2 B_\theta^2)\Big|_0^\infty. \tag{4.17}$$

With the notation I for the total current flowing along the plasma cord, we have $rB_\theta = 2I/c$ as $r \to \infty$; therefore the right-hand side of (4.17) is equal to $I^2/2\pi c^2$. On the other hand, partial integration of the left-hand side results in $\int 2pr\, dr$. Thus, we obtain the integral equilibrium relation

$$4c^2 N\bar{T} = I^2 \tag{4.18}$$

with $N = \int 2\pi r n\, dr$ standing for the number of electrons per segment of the cord of unit length, and $\bar{T} = (\bar{T}_i + \bar{T}_e)/2$ for the mean temperature. Relation (4.18) is usually called the Bennet relation.

If there is a longitudinal magnetic field, relation (4.18) must be supplied with an additional term to account for the difference between the pressure inside and outside the cord. In particular, if the plasma pressure is small, the compression of the plasma by the magnetic field of its own current can be balanced by an increase of the field inside the cord. In these circumstances the plasma turns out to be paramagnetic.

4. Force-free configurations. In a conducting plasma when the plasma pressure is small in comparison with the magnetic field pressure, certain equilibrium configurations with no pressure may exist. It follows from the force balance equation (4.9) that in these circumstances the electric current must run along

the magnetic lines. The magnetic configurations with $\nabla p = 0$ are called pressureless or force-free equilibria. In a laboratory they can occur when the energy losses are high and the plasma is not heated. Force-free configurations often arise on the Sun when the magnetic fields that have been generated inside of the Sun come to the surface.

Force-free configurations are simplest in the case of cylindrical symmetry. Since we still have one equation in two unknown functions $B_\theta(r)$, $B_z(r)$, there are still plenty configurations available.

We shall consider only one example that is interesting because of its relation to experiments on high-current discharges being stabilized by a moderate longitudinal magnetic field. The plasma pressure in these circumstances is not large due to different anomalous losses (high thermal conductivity and diffusion), so the field configuration is almost pressureless. It turns out that in toroidal experiments it is not only strainless, but satisfies the additional condition that the magnetic field is proportional to the current, i.e.,

$$\nabla \times \mathbf{B} = \mu \mathbf{B}, \quad \mu = \text{const.} \tag{4.19}$$

With the notation J_0, J_1 for the Bessel functions a general solution to these equation may be presented in the form

$$B_z = B_0 J_0(\mu r), \quad B_\theta = B_0 J_1(\mu r). \tag{4.20}$$

Let a be the plasma column radius. We see from (4.20) that the longitudinal magnetic field flux shrinks towards the axis as μa increases, i.e., we are faced with a "paramagnetic effect." When μa reaches the first zero $\alpha_0 \approx 2.4$ of the Bessel function J_0, the field component B_z becomes zero on the column boundary, and appears to be negative for $\mu a > \alpha_0$.

For the following we shall use the integral parameter $\Theta = 2\pi I a / c\Phi$ with I standing for the total current and Φ for the magnetic flux. It can be shown that $\Theta \cong \mu a/2$. In experiments a discharge is usually performed inside a casing with perfectly

conducting walls, so that $\Phi = \pi a^2 B_0$ where B_0 is the initial longitudinal magnetic field.

Thus, the force-free Bessel-function model predicts that the longitudinal field has to reverse its sign on the plasma boundary when Θ passes over $\alpha_0/2$. This phenomenon has been observed experimentally (see Fig. 4.2). The critical value of Θ corresponding to the sign reversal is also in good agreement with experiment: the experimental value is $\Theta \approx 1.4$ while the theoretical is $\Theta \approx 1.2$.

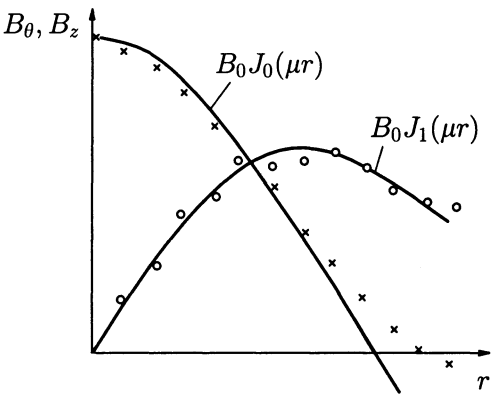

Fig. 4.2. Magnetic field components inside a diffusive pinch.

The reason why only special force-free configurations of the form (4.19) arise in experiments was clarified by J. Taylor [67]. Following his arguments based on experiments, we shall assume that a diffusion pinch is always accompanied by strong hydromagnetic turbulence. The turbulent processes can transfer the plasma and the magnetic energy into a thermal form, which will be lost later due to different leakages.

Note that a plasma in such a column is usually a well-conducting one, therefore there are some constraints imposed on its motion. If the conductivity is perfect, the variations of the magnetic field must meet the restrictions (4.7). Correspondingly, we have for the vector potential $\mathbf{A}$, defined by the relation $\mathbf{B} =$

$\nabla \times \mathbf{A}$, the identity

$$\frac{\partial \mathbf{A}}{\partial t} = \boldsymbol{v} \times \mathbf{B} + \nabla \chi, \tag{4.21}$$

where χ stands for an arbitrary scalar function.

With the aid of (4.7) and (4.21) it may easily be verified that the integral $\int(\mathbf{A} \cdot \mathbf{B})d\mathbf{r}$ taken over any volume embraced with magnetic lines does not depend on the time, i.e., represents an integral of motion.

If the conductivity is imperfect, this statement will not be totally valid, since the magnetic field lines will be able to interconnect. But if the conductivity is sufficiently high, the magnetic field will be disturbed slightly and the integral $K_0 = \int(\mathbf{B} \cdot \mathbf{A})\, d\mathbf{r}$ over the whole plasma volume will be nearly conserved, since reconnection of the field lines leads to redistribution of different magnetic tubes, but does not change their net contribution to the integral K_0.

Now we are in a position to look for a magnetic field of minimal energy that satisfies the constraint $K_0 = \text{const}$. The solution to the corresponding variational problem is provided by the formulae (4.19). In summary, we may say that the observed force-free equilibrium in a diffusive pinch is related to the small-scale hydromagnetic turbulence that makes the magnetic field energy dissipate with K_0 being kept fixed.

4.2. Two-fluid magneto-hydrodynamics

4.2.1. Basic equations

The previous treatment was based on the assumption that the plasma was at rest, i.e., that its average velocity v was zero. This velocity was implicitly identified with the velocity of all plasma components, i.e., with the velocities v_i and v_e of the ions and of the electrons. But if there is an electric current in a plasma, these two velocities will be different since $\mathbf{j} = en(\boldsymbol{v}_i - \boldsymbol{v}_e) \neq 0$.

To estimate the electron velocity v_e let us assume, in the first instance, that there is no longitudinal current. Then the equilibrium equation (4.9) enables us to evaluate the current density:

$$\mathbf{j}_\perp = \frac{c}{enB^2}\mathbf{B} \times \nabla p. \tag{4.22}$$

So if the ion velocity is zero, the electron velocity due to the pressure gradient (this velocity will henceforth be refered to as the drift velocity) is given by

$$\boldsymbol{v}_{e\perp} = -\frac{c}{enB^2}\mathbf{B} \times \nabla p. \tag{4.23}$$

Thus, we have the following estimate of the drift velocity: $v_\perp \sim cT/eBa \sim v_{Ti}\,\rho_i/a$. Here $v_{Ti} = \sqrt{T/m_i}$ stands for the average thermal velocity of the ions, $\rho_i = v_{Ti}/\Omega_i \equiv v_{Ti}\,m_i c/eB$ for their gyroradius, $\Omega_i = eB/m_i c$ for the ion cyclotron frequency and a for the plasma spatial scale. The gyroradius ρ_i and, correspondingly, the drift velocity $v_\perp$ are very small in a strong magnetic field.

But this estimate is not very helpful in our analysis of the implicit assumptions that were made in the previous Sections. In fact, the drift velocity has not been involved in the consideration and, therefore, we are not aware of the physical meaning of the approximation used and of the effects that may have been lost. In order to clarify this problem we have to change the viewpoint and to regard the steady state of a plasma as a stationary flow. Thus, we find ourselves in a position to incorporate the macroscopic (hydrodynamical) velocities of the electrons and ions into the basic equations. If we retrieve the assumption that the electrons and the ions behave from the standpoint of kinematics as two independent fluids, we shall accept the so-called two-fluid hydrodynamical approximation. The grounds for this approximation were long ago pointed out by Landau [68] and are provided by the essential difference of time scales relating to the rate at which the Maxwellian distribution is achieved inside each plasma component (electron and ion) and the rate at which heat

is transferred from one component to the other [69]. In view of the friction between the ions and electrons we shall accept the equations of the two-fluid hydrodynamics in the following form:

$$m_i n \frac{d\boldsymbol{v}_i}{dt} + \nabla p_i = +en\mathbf{E} + \frac{en}{c}\boldsymbol{v}_i \times \mathbf{B} + \frac{m_e n}{\tau_e}(\boldsymbol{v}_e - \boldsymbol{v}_i), \quad (4.24)$$

$$m_e n \frac{d\boldsymbol{v}_e}{dt} + \nabla p_e = -en\mathbf{E} - \frac{en}{c}\boldsymbol{v}_e \times \mathbf{B} - \frac{m_e n}{\tau_e}(\boldsymbol{v}_e - \boldsymbol{v}_i). \quad (4.25)$$

Here m_i, m_e are the masses of the ions and electrons, $\boldsymbol{v}_i, \boldsymbol{v}_e$ are their velocities, $p_i = nT_i$, $p_e = nT_e$ and T_i, T_e are their pressures and temperatures, while $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields and τ_e is the mean-free time between collisions of the electrons with the ions. We also assume as before that the electron and ion densities are equal (the quasi-neutrality condition).

Note that the derivatives $d\boldsymbol{v}/dt = \partial \boldsymbol{v}/\partial t + \boldsymbol{v}\nabla \boldsymbol{v}$, following the motion of the fluid, are involved in both equations (4.24) and (4.25). But since the electron mass is very small and can be neglected in many circumstances, we shall omit the electron inertial term in the second equation to simplify things.

Summing up equations (4.24)–(4.25) and neglecting the electron inertia, we obtain the equation

$$m_i n \frac{d\boldsymbol{v}}{dt} + \nabla p = \frac{1}{c}\mathbf{j} \times \mathbf{B} \quad (4.26)$$

of the one-fluid hydrodynamics with $p = p_e + p_i$, $\mathbf{j} = en(\boldsymbol{v}_i - \boldsymbol{v}_e)$ and $\boldsymbol{v}$ being equal to the ion velocity.

Under this approximation equation (4.25) may be rewritten in the form

$$\nabla p_e = -en\mathbf{E} - \frac{en}{c}\boldsymbol{v} \times \mathbf{B} + \frac{1}{c}\mathbf{j} \times \mathbf{B} - \frac{en\mathbf{j}}{\sigma} \quad (4.27)$$

with $\sigma = e^2 n \tau_e / m_e$ standing for the conductivity. This equation relates the current to the electric field and can be regarded as a generalization of Ohm's law. One may say that the two-fluid plasma origin results in additional terms in Ohm's law (4.27).

The equations of motion (4.26)–(4.27) are to be joined with Maxwell's equations

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{j}, \tag{4.28}$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \tag{4.29}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{4.30}$$

(we neglected the displacement current) and with the continuity equation

$$\frac{\partial n}{\partial t} + \nabla \cdot n\boldsymbol{v} = 0. \tag{4.31}$$

Moreover, one should add the heat balance equations, governing the evolution of T_i and T_e. But on short time scales when the heat transfer is not of vital importance, one may assume that T_i and T_e are subject to an adiabatic law.

Let us suppose that a plasma moves at a velocity that exceeds essentially the drift velocity $v_\perp \sim cT/eBa$. Then the terms ∇p_e and $(1/c)\,(\mathbf{j}\times\mathbf{B})$ in equation (4.27) will be negligible in comparison with the second term on the right-hand side and hence there must be a strong electric field $\mathbf{E}$. Thus, we can omit the second and the third terms on the right-hand side and treat such processes with the ordinary Ohm's law of one-fluid magneto-hydrodynamics. Therefore, rapid motion of a plasma is governed by one-fluid hydrodynamics. But what about the equilibrium? Note that according to the one-fluid model the ions will be at rest.

In order to answer this question let us try to balance the contributions of the electrons against the ions in the one-fluid model. It follows from (4.26) that the inertial term $m_i n d\boldsymbol{v}/dt$ becomes dominant only if the ion velocity is of the same order of magnitude as the sound velocity. If the plasma velocity v is small, say of the order of the drift velocity, this term can be neglected and we return to the problem of the plasma equilibrium. One may also imagine particular stationary flows, such

that the plasma slides along the magnetic surfaces and the magnetic field itself is frozen to the plasma and does not change. Varying the longitudinal part of the flow velocity (i.e., the component directed along the magnetic field), one can make the flow is divergence-free. For instance, we may put $\boldsymbol{v} = \mathbf{j}/en$ since this vector lies on the magnetic surface and $\nabla \cdot n\boldsymbol{v} = 0$. But this example relates to the case when the current is transported only by the ions while all electrons are motionless.

In a more general case there is an essential freedom to rearrange the ion and electron contributions to the electric current. Therefore, the one-fluid stationary hydromagnetic configurations form a wide class of steady plasma flows corresponding to slow sliding of the ion and electron components along the magnetic surfaces.

4.2.2. The Hall effect

The previous consideration seems convincing, but nevertheless the following question may be posed. As we have found, one-fluid magneto-hydrodynamics and the condition that the magnetic field is frozen to the plasma are valid only as long as the velocities are large in comparison with the drift velocity. But we applied this approximation to the cases when the flow velocity was small. Was it correct and under what conditions is it correct? In order to answer this question we have to take the two-fluid plasma nature into account. Assuming the conductivity to be large, let us rewrite Ohm's law in the form

$$\nabla p_e = -en\mathbf{E} - \frac{en}{c}(\boldsymbol{v} \times \mathbf{B}) + \frac{1}{c}(\mathbf{j} \times \mathbf{B}). \qquad (4.32)$$

The last term on the right-hand side may be called the Hall term since it results in the well known Hall effect in metals and semiconductors. In summary, we may say that the generalized Ohm's law takes the electron pressure gradient and Hall effect into account.

Now let us consider the particular case when the electron temperature T_e is constant throughout the whole space. Re-

trieving the electron velocity $\boldsymbol{v}_e$, we may write the equation (4.32) in the form

$$\mathbf{E} = -\frac{1}{c}(\boldsymbol{v}_e \times \mathbf{B}) - \nabla\left(\frac{T_e}{e}\log n\right). \tag{4.33}$$

Inserting this expression for the electric field into the Maxwell equation (4.29) and recalling that the curl of a potential vector field vanishes, we obtain

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\boldsymbol{v}_e \times \mathbf{B}). \tag{4.34}$$

This relation also proves to be valid in the case when T_e is not a constant but depends on the density, for instance, when T_e and n are constant on magnetic surfaces but change across them.

As we have already stated, equation (4.34) means that the magnetic lines are frozen to the fluid moving at the speed $\boldsymbol{v}_e$, i.e., to the electrons in this particular case. Thus, one may expect that one-fluid hydrodynamics will considerably diverge from two-fluid only when the difference of $\boldsymbol{v}_e$ and $\boldsymbol{v}_i$ is sufficiently high.

This statement can be applied to a particular case. Suppose that we treat a typical steady state in the framework of magneto-hydrodynamics, when the pressure p and the temperatures T_i and T_e (because of high thermal conductivity along the magnetic field) are constant on magnetic surfaces. Since the electric current is flowing along the magnetic surfaces, we need not bother which plasma component — the electron or the ion — is frozen to the magnetic field: the magnetic surfaces will not be changed in any way. Therefore, the one-fluid approximation is still valid.

If there is no exact equilibrium, we shall find ourselves in a different position. Even if there are magnetic surfaces and the electrons appear to be confined to them (because of the magnetic fluxes that are to be frozen to the electrons), the condition that the equilibrium is not complete results in transverse ions flows with velocities of the order of the drift velocity. We shall consider a few cases of such flows by way of example.

To proceed with steady states in two-fluid hydrodynamics, let us begin with the investigation of electron flows loaded with a magnetic field under the condition that the ions are motionless. We shall take the particular case of an electric current in a rigid corrugated metal conductor [70]. The last problem may also be related to the example of an inhomogeneous pinch in a rarefied plasma. Indeed, if the density is low, the electron velocity $v = -j/en$ due to the current may essentially exceed the sound velocity $c_s = \sqrt{T/m_i}$. Namely, the Bennet relation (4.18) implies that $v/c_s \sim D_i^{-1/2}$, with $D_i = Nm_ic^2/e^2$. Therefore, if $D_i \ll 1$, we have $v \gg c_s$ and may regard the electron flow on short time scales (more precisely, as long as the plasma column remains undisturbed by the ions moving with the sound velocity) as if the ion density (and hence the electron density) were constant. One may also note that since v is of the order of the drift velocity $c_s\rho_i/a$, the condition $v \gg c_s$ implies $\rho_i \gg a$, i.e., that the magnetic field appears to be too weak to affect the ions.

Thus, take a corrugated conductor and suppose that the charge density $n(r, z)$ is a periodic function of z. The problem is to describe stationary electron flows under the condition that the density $n(r, z)$ is kept fixed.

For the sake of simplicity we assume the conductivity to be infinitely large and the temperature T to be constant throughout the conductor. Thus, we need only to reveal the influence of the Hall effect on the electron velocity.

On the condition that there is axial symmetry, i.e., that all the quantities depend only upon the two variables r, z in cylindrical coordinates, the law

$$\nabla \cdot \mathbf{j} = \frac{1}{r}\frac{\partial}{\partial z} r j_r + \frac{\partial}{\partial z} j_z = 0 \tag{4.35}$$

of current conservation implies that the radial and longitudinal parts j_r, j_z of the current density may be expressed in terms of an arbitrary scalar function $I(r, z)$:

$$j_r = -\frac{1}{2\pi r}\frac{\partial I}{\partial z}, \quad j_z = \frac{1}{2\pi r}\frac{\partial I}{\partial r}. \tag{4.36}$$

In view of the expression for j_z we have

$$I = \int_0^r 2\pi J_z r' dr', \tag{4.37}$$

i.e., that $I(r, z)$ coincides with the total current through the circular area of the radius r of the conductor cross-section at the level z.

Now recall that the magnetic field is frozen to the electrons. Since $B_z = B_r = 0$, we need to consider only the θ-component of equation (4.34), which in the stationary case takes the form

$$\frac{\partial}{\partial r} v_r B_\theta + \frac{\partial}{\partial z} v_z B_\theta = 0. \tag{4.38}$$

Substituting v for $-j/en$ and expressing j in terms of I with the help of (4.36), we obtain

$$\frac{\partial}{\partial r}\left(\frac{I}{nr^2}\frac{\partial I}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{I}{nr^2}\frac{\partial I}{\partial r}\right) = 0 \tag{4.39}$$

given that $B_\theta = 2I/cr$ due to the axial symmetry. Carrying out the differentiation we find that

$$\frac{\partial I}{\partial z}\frac{\partial}{\partial r} nr^2 - \frac{\partial I}{\partial r}\frac{\partial}{\partial z} nr^2 = 0. \tag{4.40}$$

Thus, the function I must depend only on nr^2. In other words, the condition that the magnetic field is to be frozen to the electrons imposes significant constraints upon the electron flow. For instance, in the case of a uniform conductor with $n = \text{const}$ the current stream-lines appear to be straight. In particular, there must be no current in convexities (in the "shadow zones") of a corrugated cord; the current is nonzero only in the central part of the conductor which electrons can "look through" [70].

4.2.3. Plasma equilibrium in a tilted magnetic field

Let us now take the more important case of a strong magnetic field when the average gyroradius is small in comparison

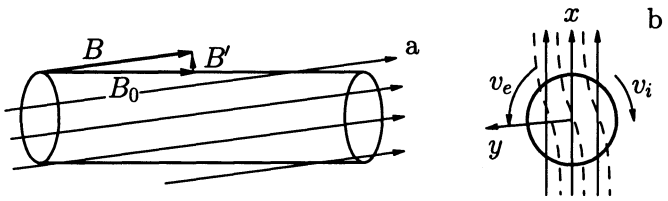

Fig. 4.3. Plasma rotation in a tilted magnetic field.

with the plasma column cross dimensions. Although the drift velocity also appears to be small, under certain circumstances we have to take it into account. Suppose, for instance, that a hydromagnetic system is very near equilibrium, so that the Ampère force is nearly balanced by the pressure gradient. Thus, the plasma will expand slowly due to the this imperfect balance, and if the transverse part of the velocity of expansion flow is comparable to the drift velocity, the one-fluid magneto-hydrodynamics and the simplified Ohm's law become invalid and we must take account of the Hall effect.

A plasma column of very low pressure ($\beta = 8\pi p/B^2 \ll 1$) in a tilted magnetic field (see Fig. 4.3a) may serve as a good example. Assuming the column to be infinitely long, we represent the magnetic field as $\mathbf{B} = \mathbf{B}_0 + \mathbf{B}'$, with $\mathbf{B}_0$ being directed along the axis and $\mathbf{B}'$ standing for a small transverse component. When $B' \ll B_0$ the plasma is nearly at rest, but, nevertheless, not totally in a steady state. Therefore, in the one-fluid approximation the plasma will expand along the magnetic lines at the speed c_s of sound and an observer will see that the column widens in the direction of $\mathbf{B}'$ at the speed $v_\perp \sim (B'/B_0)c_s$. If this velocity becomes less or of the order of the drift velocity, i.e., if

$$\frac{B'}{B_0} \leq \frac{\rho_i}{a}, \tag{4.41}$$

we should use two-fluid hydrodynamics. Suppose B' to be not

too small, so that

$$\frac{B'}{B_0} > \frac{\rho_e}{a} = \sqrt{\frac{m_e}{m_i}}\frac{\rho_i}{a}. \tag{4.42}$$

Then the electron leakage velocity $(B'/B_0)\sqrt{T_e/m_e}$ essentially exceeds the drift velocity. Therefore, we may regard the electrons as being subjected to the equilibrium distribution law along the magnetic lines. Correspondingly, we may neglect the inertial term in the equations of motion and rewrite them in the form (4.33). Multiplying (4.33) by $\mathbf{B}$, we obtain

$$\mathbf{B} \cdot \mathbf{E} = -\mathbf{B} \cdot \nabla \left(\frac{T_e}{e} \log n \right). \tag{4.43}$$

But a stationary electric field is always a potential one: $\mathbf{E} = -\nabla\varphi$. Therefore, we may integrate equation (4.43) along each line of force. This results in

$$\varphi = \frac{T_e}{e} \log n + \varphi_0 \tag{4.44}$$

with φ_0 standing for a certain constant that may depend upon the magnetic line. In our case the column is uniform with respect to z, thus φ_0 can depend only on y (see Fig. 4.3b, presenting a view from above; the z axis is directed along the column, the x axis — along the transverse component $\mathbf{B}'$ of the magnetic field, while the y axis is normal to the magnetic surfaces $y =$ const that contain the field lines). For a stationary cylindrically-symmetric configuration φ_0 is apparently equal to zero.

Thus, the electric field is given by

$$\mathbf{E} = \nabla\varphi = -\frac{T_e}{en}\nabla n \tag{4.45}$$

Insert this expression into the equation

$$m_i n \frac{d\boldsymbol{v}}{dt} + \nabla n T_i = en\mathbf{E} + \frac{en}{c}(\boldsymbol{v} \times \mathbf{B}) \tag{4.46}$$

of the ions motion and neglect the inertia term. Taking the vector product of (4.46) with $\mathbf{h} = \mathbf{B}/B$ and introducing the notation $p = n(T_e + T_i)$ for the total pressure, we obtain

$$\boldsymbol{v}_\perp = \frac{c}{enB}\,\mathbf{h} \times \nabla p. \tag{4.47}$$

Thus, in a stationary state the plasma column revolves at a velocity determined by formula (4.47), whereas the azimuthal part of the electron velocity appears to be zero since the electrons are "glued" to the magnetic lines.

The plasma rotation prevents its expansion along the field lines. Indeed, the longitudinal component of the pressure gradient tries to shift the ions at the top of Fig. 4.3a to the right, but the rotation brings them to the bottom of the picture, where they are subjected to a force that pushes them to the left. If the rotation is fast, i.e., if $c_s\rho_i/a \gg c_sB'/B_0$, the shift at the top is balanced against the shift at the bottom and there is no net displacement. In summary, we may say that due to the Hall effect a stationary flow develops instead of an equilibrium state. But this flow appears to be unstable.

The above investigation is valid only under the condition that the pressure is very low, i.e., if $\beta \ll 1$, or, more precisely, if the plasma kinetic energy $m_iv^2n/2$ due to the rotation is far less than the energy $B^2/8\pi$ of the transverse magnetic field: $8\pi p/B'^2 \ll a^2/\rho_i^2$. In the opposite limit $B'^2 \ll 8\pi p\rho_i^2/a^2$ the transverse field fails to provide a velocity of the order of the drift velocity and we may regard the ions as motionless. Let us consider the process from the very beginning. Since the magnetic field is frozen to the electrons, the field lines will revolve together with the electron component of the plasma (see Fig. 4.4a, b, where we have neglected the dependence of the angular velocity of the electrons upon the distance from the axis to simplify things). On performing the first half of the revolution, the rotating field will be torn from the external one. Thus, the transverse magnetic field may be represented as if produced by the superposition of a dipole-like plasma field on the constant

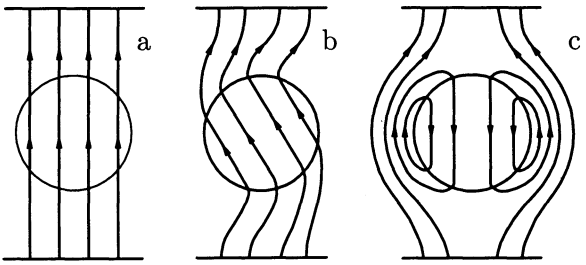

Fig. 4.4. Plasma rotation in a θ-pinch due to the Hall effect in the presence of a weak transverse magnetic field.

external field that does not penetrate into the column (see Fig. 4.4c). Since the magnetic lines of the plasma field were glued to the external lines during the initial revolution, the plasma gets a portion of the angular momentum and starts rotating at a certain circular velocity v_θ. On balancing the kinetic energy of rotation against the magnetic energy, we obtain the following estimate: $v_\theta \approx B'(4\pi m_i n)^{-1/2}$.

This model of plasma rotation due to an external magnetic field was first considered by Hains [71, 72] and then by Thonemann and Kolb [73] in connection with experiments on θ-pinches. The centrifugal instability of the plasma column revealed itself in such experiments, and indicated clearly that the column had been rotating.

4.3. Plasma diffusion

4.3.1. Plasma in a strong magnetic field

Our investigation, so far, was conducted under the assumption that the conductivity σ was infinitely high. All the steady states that were previously described were obtained in the framework of this approximation. But if the conductivity is imperfect, a plasma will slowly diffuse transversely to the magnetic field. To verify the statement, let us consider a particular case of a low-pressure ($\beta \ll 1$) plasma column in a magnetic

field. Neglecting weak diamagnetic effects, we shall assume the field to be uniform. Then by virtue of the azimuthal component of the generalized Ohm's law (4.27) we obtain

$$v_r = \frac{c}{B}E_\theta - \frac{c}{B\sigma}j_\theta = \frac{c}{B}E_\theta - \frac{c^2}{\sigma B^2}\frac{dp}{dr}. \tag{4.48}$$

Here we expressed j_θ in terms of dp/dr with the aid of the equilibrium equation. If the magnetic field does not change with time, E_θ will vanish and according to (4.48) the radial part of the velocity will correspond to the plasma expansion. If $T = \text{const}$, the velocity v_r is proportional to the pressure gradient and the plasma expansion is of a diffusion type with the transverse diffusion coefficient being equal to

$$D_\perp = \frac{2T}{m_e\Omega_e^2\tau_e} = \beta\frac{c^2}{8\pi\sigma} = \frac{\rho_e^2}{\tau_e}, \tag{4.49}$$

where $\beta = 8\pi p/B^2$, $\Omega_e = eB/m_ec$ is the electron cyclotron frequency, τ_e is the mean-free time between the electron-ion collisions, $\sigma = e^2n\tau_e/m_e$ is the conductivity and $\rho_e^2 = 2T/m_e\Omega_e^2$ is the squared electron gyroradius. Expression (4.49) for the diffusion coefficient has a simple physical meaning: it states that on the average an electron covers a distance of the order of the gyroradius between collisions.

If β is not small, the magnetic field will change during the process of diffusion and hence $E_\theta \neq 0$. Thus, it will be more precise to say that the magnetic field and the plasma diffuse through each other.

In the opposite limit of a motionless plasma we are faced with a skin-effect problem, i.e., the problem of magnetic field penetration into a conducting medium with diffusion coefficient $D_m = c^2/4\pi\sigma$.

4.3.2. Toroidal plasma diffusion

If the plasma is of the shape of a torus, the diffusion process is more complicated and is like convection. This effect relates to the fact that even in a stationary magnetic field the

component E_θ of the electric field is not zero because of the current flowing along the helical magnetic lines:

$$E_\theta = \frac{B}{B_\theta}\frac{\tilde{j}_{||}}{\sigma_{||}}. \tag{4.50}$$

Here $\tilde{j}_{||}$ stands for the alternating (with θ) part of the longitudinal current and $\sigma_{||}$ for the longitudinal plasma conductivity, which may be different from the transverse conductivity involved in (4.48) (according to the kinetic theory $\sigma_{||} \approx 2\sigma_\perp$ in a hydrogen or deuterium plasma due to the Coulomb collisions). Relating $j_\perp$ to the pressure gradient with the aid of the equilibrium equation and using the identity $\nabla \cdot \mathbf{j} = 0$, we may evaluate the current $\tilde{j}_{||}$:

$$\frac{B_\theta}{rB}\frac{\partial \tilde{j}_{||}}{\partial \theta} = -\nabla \cdot \mathbf{j}_\perp = -\nabla \cdot \frac{c}{B^2}\left(\mathbf{B} \times \nabla p\right) = \frac{2cr}{BR_0}\frac{dp}{dr}\sin\theta. \tag{4.51}$$

Note that we also made use of $B \approx B_T$ (under the assumption that $B_\theta^2 \ll B_T^2$) and that B_T, although slightly, must depend on θ because it is inversely proportional to the distance from the axis of symmetry. It follows from (4.50)–(4.51) that

$$E_\theta = \frac{2cB}{B_\theta^2 \sigma_{||}}\frac{dp}{dr}\frac{r}{R_0}\cos\theta. \tag{4.52}$$

Let us introduce the quantity $q = rB_T/R_0B_\theta$ coinciding with the ratio of the number of turns of a magnetic line around the major circumference of the torus to the number of turns around the minor circumference. Inserting expression (4.52) into (4.48), we see that the electric field E_θ results in a flow, whose radial velocity v_r is Rq^2/r times greater than the diffusion velocity in the case of a straight cylinder. However, the average value of this velocity with respect to the azimuth is equal to zero, i.e., there is no diffusion across magnetic surfaces due to the field (4.52) in the leading approximation. Thus, we are faced with a stationary laminar convection: the plasma flows away from the inside of the torus at the outer part of its surface (i.e., where $\cos\theta < 0$) and returns at the inner side.

Nevertheless, a more accurate approximation reveals that this convection causes a diffusion flow. In fact, in order to find the mean flow through a magnetic surface, we have to average out nv_r with the weight factor $\left[1 - (r/R_0)\cos\theta\right]d\theta$ since an element of area of the torus surface is

$$dS = 2\pi R r d\theta \cong 2\pi R_0 \left(1 - \frac{r}{R_0}\cos\theta\right) r d\theta.$$

Evaluating the corresponding integral and taking special care of small terms of the order of r^2/R_0^2, we obtain

$$n\bar{v}_r = -\frac{c^2}{\sigma_\perp B^2}\frac{dp}{dr}\left(1 + \frac{2\sigma_\perp}{\sigma_\parallel}q^2\right). \tag{4.53}$$

Putting $\sigma_\parallel = 2\sigma_\perp$, we find that in the case of a hydrogen or a deuterium plasma the toroidal diffusion coefficient is given by the formula

$$D_\perp = D_{\perp 0}(1 + q^2) \tag{4.54}$$

with $D_{\perp 0}$ being the diffusion coefficient for a straight magnetic field. This relation was obtained by Pfirsh and Schlüter [74].

Note that all these results are valid only in the case of a dense plasma, when the ion and electron mean free-paths are far less than the characteristic length qR_0 of a magnetic line and the plasma may be treated by means of hydrodynamics.

4.3.3. Neoclassical transport

In order to perform a controlled nuclear fusion process we need a high-temperature plasma of moderate density. The mean free-path of particles in such a plasma is very long, so we may speak of a highly rarefied plasma. If we restrict ourselves to the particular case of a torus, i.e., to the tokamak geometry, we may state that only the case when the ion and electron free-paths exceed the characteristic length qR_0 of a magnetic line is of practical interest. Under these circumstances hydrodynamics are invalid. The transport theory for a rarefied plasma in a torus

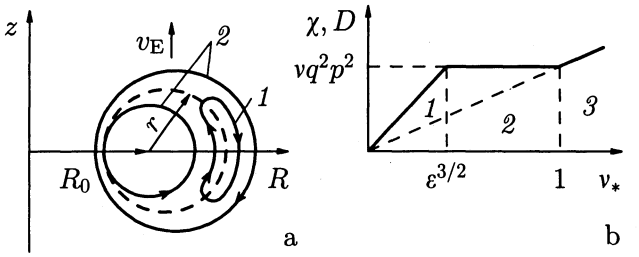

Fig. 4.5. (a) Trajectories of trapped *1* and transitory *2* particles. (b) Dependence of the transport coefficients on the dimensionless collision frequency $\nu_* = \nu q R/v_T$.

subjected to a strong magnetic field was developed by Galeev and Sagdeev [75]. A simplified version is the following.

Let us consider, in the first instance, the motion of a separate charged particle in the tokamak magnetic field. It is well-known that a charged particle moves in a strong uniform magnetic field along a helical trajectory. In fact, the particle tends to follow a magnetic line at the speed $v_{||}$ and orbits around it along a circle of the radius $\rho = v_\perp/\Omega$, where $\Omega = eB/mc$ is the gyroradius (the mass and the charge of the particle are denoted by m and e). If the magnetic field is non-uniform, we shall observe two effects. First, the velocities $v_{||}, v_\perp$ will vary along magnetic lines, and, second, there will be a slow drift across the magnetic field. In a tokamak the toroidal component of the magnetic field is the governing one because of its strength, so only its inhomogeneity is important.

Consider a transverse cross-section of a magnetic surface and denote its major radius with R_0 and the minor one with r, as in Fig. 4.5a. Assuming $\varepsilon = r/R_0$ to be small, we ensure that the approximate relation $B = B_0(1 - r\cos\theta/R_0)$ is valid on the magnetic surface. Thus, the magnetic field when compared to its value at the point R_0 appears to be weaker on the outer part of the torus surface and stronger on the inner one.

A charged particle in a magnetic field conserves its total energy $\mathcal{E} = mv_{||}^2/2 + mv_\perp^2/2$. Moreover, it also conserves

the "transverse adiabatic invariant" $\mu_\perp = mv_\perp^2/2B$ that can be interpreted as the magnetic momentum. We see that the transverse kinetic energy grows and the longitudinal decreases when the particle penetrates into an area of stronger magnetic field. If the longitudinal part of the energy is sufficiently small in comparison with the transverse part, the particle may be reflected. To relate this effect to our case, assume the longitudinal and transverse velocities to be equal to $v_\parallel^{(0)}$ and $v_\perp^{(0)}$ at the point $\theta = 0$, i.e., at the outer part of the torus. Motion to an area of a stronger magnetic field will result in an increase of the transverse energy. But since the total energy is conserved, the longitudinal energy will change proportionally to $v_\parallel^2 = |v_\parallel^{(0)}|^2 - \varepsilon(1 - \cos\theta)|v_\perp^{(0)}|^2$. This relation ensures that the particles are reflected by the strong magnetic field and do not reach the inner parts of the torus provided $|v_\parallel^{(0)}|^2 < 2\varepsilon|v_\perp^{(0)}|^2$. Particles with $|v_\parallel^{(0)}|^2 > 2\varepsilon|v_\perp^{(0)}|^2$ are called transitory particles for they move freely along the field lines.

Now we are prepared to take the transverse drift into account. Representing a particle with its gyroorbit regarded as a small circle moving with a velocity $v_\parallel$ along a magnetic line, we may say that such a particle is subjected to two forces directed along the major torus radius, namely to the centrifugal one, $mv_\parallel^2/R$, and to the diamagnetic one, equal to $-\mu_\perp \nabla B = mv_\perp^2/2R$ and trying to push the gyrocircle out. They result in the force $F_\perp$ that causes the particle to drift at the velocity v_d, such that $F_\perp$ should be balanced by the Lorentz force: $F_\perp = ev_dB/c$. In other words, $v_d = (v_\parallel^2 + v_\perp^2)/2\Omega R$. Note that the drift velocity is directed upwards and hence at the upper part of the torus the particle crosses the magnetic surfaces from the inside, whereas at the lower part from the outside. It may also be easily seen that the projection of a trapped particle trajectory on the torus cross-section looks like a "banana" (see Fig. 4.5a). The "banana" width is $\Delta = v_{dr}t$, where v_{dr} is the component of the drift velocity along the minor radius of the torus and t is half a period of the bounce along the banana-like trajectory. For a typical trapped particle of average thermal energy and longi-

tudinal velocity $v_{||} \approx \varepsilon v_{\perp}$ the "banana" width is of the order of $\Delta \approx q\rho/\varepsilon$, where ρ is the average gyroradius, q — the safety factor (recall that the characteristic scale of a field line on the outer part of the magnetic surface is proportional to qR_0).

The transitory particles also deviate from the magnetic surfaces. The deviations of the particles that are near to becoming trapped (see Fig. 4.5a) appear to be of the order of those of the trapped ones. But on average the longitudinal and transverse velocities of the transitory particles are of the same order in magnitude and this deviation is small, namely about $q\rho$.

With the particle trajectories in hand we are able to estimate the transport coefficients. It is convenient to begin with the case of an extremely rarefied plasma, when the mean free-path is very large compared with qR_0. It turns out that in this case the transport process is predominantly due to the trapped particles, so we may speak of diffusion and thermal conductivity on the "banana." The area of very rare collisions is marked with *1* in Fig. 4.5b. The character *3* marks the hydrodynamical domain of a sufficiently dense plasma, when the free-path is less than qR_0, while area *2* is related to the intermediate collision frequencies. The ordinate axis in Fig. 4.5b corresponds to the values of the transport coefficients, i.e., to the diffusion D and to the heat conductivity χ, while the points of the horizontal axis correspond to the collision frequency ν divided by the average bounce frequency, $\nu_T = v_T/qR_0$, with v_T being the average velocity of the thermal motion. Thus, the condition that $\nu = \nu_T$ means the equality of the mean free-path and the quantity qR_0. Correspondingly, this point sets a boundary between the areas of frequent and rare collisions.

In order to estimate the transport coefficients in the area of very rare collisions, note that collisions can make a trapped particle free and vice versa, and that the characteristic displacement of such a particle is about Δ. Moreover, this transition may evidently be caused by a mere deviation of the longitudinal velocity by a quantity of the order of $v_{||} \sim \sqrt{\varepsilon} v_{\perp}$, which is small in comparison with $v_{\perp}$. On average the magnitude of such a de-

viation represents a small, $\sim \sqrt{\varepsilon}$, part of the thermal velocity. Recall now that the Coulomb collisions due to their long-range character cause diffusion in the space of velocities. Therefore, we are able to introduce the effective collision frequency $\nu_{\text{eff}} \approx \nu/\varepsilon$, which controls the conversion rate of the trapped particles into the transitory ones due to a diffusive shift over a distance of the order of $\sqrt{\varepsilon} v_\perp$ in the space of longitudinal velocities. Since the number of trapped particles is of the order of $\sqrt{\varepsilon}$ of the total number, we obtain the estimate $\sqrt{\varepsilon}\Delta^2\nu_{\text{eff}} = \varepsilon^{-3/2}\nu q^2\rho^2$ for the thermal conductivity χ.

This expression is related to both the ion and the electron temperature conductivities: one needs only to insert the appropriate values of ν and ρ. The diffusion coefficient is of the same order of magnitude as the electron thermal conductivity, i.e., it is essentially less than the total thermal conductivity.

The expressions for the transport coefficients in the area of rare collisions are valid only up to the values $\sqrt{\varepsilon} v_T/qR_0$ of the effective frequency ν_{eff}, equal to the oscillation frequencies of the trapped particles. Therefore, the boundary between areas *1* and *2* lies somewhere near $\nu/\nu_{eff} \approx \varepsilon^{3/2}$, and the corresponding value of χ is approximately equal to $\nu_T q^2\rho^2$. Almost the same value of χ is apparently achieved at the point $\nu \approx \nu_T$, i.e., at the boundary between areas *2* and *3*. Therefore, the conductivity χ must be independent of the collision frequency in area *2* of intermediate collision frequencies and this is why this frequency band is often called the plateau area. Area *3* of frequent collisions is also known as the Pfirsh – Schlüter area due to relation (4.54) for the diffusion coefficient. The dependence of transport coefficients on the collision frequency is schematically plotted in Fig. 4.5b.

Not only did the neoclassical transport theory reveal the complicated dependence of the transport coefficients on the collision frequency, but predicted some new phenomena as well. One of them, perhaps the most interesting, was called a bootstrap-current. It is related to the possibility to drive a tokamak current by means of diffusive plasma expansion [76]. There are no

difficulties in showing that provided the plasma pressure is sufficiently large, the neoclassical diffusion coefficient exceeds the coefficient $c^2/4\pi\sigma$ of magnetic field diffusion. In these circumstances the plasma expands more rapidly than the magnetic field penetrates into it. Correspondingly, the diffusion flux can push away the magnetic lines and support the current. One may argue for this in a different manner, saying, for instance, that the confined "banana" gives rise to the drift current

$$j_d \sim \varepsilon^{3/2} \frac{c}{B_0} \frac{dp}{dr}$$

because of the toroidal expulsion. Due to collisions with the effective frequency $\nu_{\text{eff}} \sim \nu/\varepsilon$ their angular momentum is transferred to the transitory particles, which, in turn, pass it to the ions at frequencies of the order of ν. Correspondingly, the current due to the transitory particles is something like

$$j \sim -\sqrt{\varepsilon} \frac{c}{B_\theta} \frac{dp}{dr}$$

and grows with pressure p. If a continuous diffusive plasma flux towards the torus periphery is provided (for instance, by an injection of tiny pellets of a cold substance into the central part of the plasma core), one may balance the current decay caused by the imperfect plasma conductivity. We may say that the plasma supports the current that confines it by the magnetic field. In analogy with L. Carroll's story about Alice who succeeded in keeping herself in the air by pulling on her shoe-laces, the English physicists called this effect a bootstrap-current.

4.3.4. Diffusion of a weakly ionized plasma

Whereas in the case of a fully ionized plasma the problem of equilibrium may be reduced to the question of whether the plasma can be confined by a magnetic field, it appears to be essentially more complicated as soon as a third component, namely a neutral gas, is involved. But there is a particular interesting

case of a weakly ionized plasma that can be treated relatively easily. Assuming the ionization to be so small that the plasma pressure is negligible compared with the neutral gas pressure, we see that the electron collisions with one another have no role in comparison with the collisions of the charged particles with the molecules of the neutral gas. In these circumstances the gas may be regarded as a motionless uniform medium in which the ions and electrons diffuse. Therefore, the question about a steady or slowly changing state may be reduced to an investigation of the plasma diffusion.

At first let us consider the case of a vanishing magnetic field. Assuming the ions and the electrons to be subjected to the Maxwellian law of distribution at the temperatures T_i and T_e, we may write the following equations of motion for each component:

$$\nabla q_j = e_j n\mathbf{E} - \frac{m_j}{\tau_j} n\boldsymbol{v}_j. \tag{4.55}$$

Here $p_j = nT_j$ is the corresponding pressure, e_j — the charge, m_j — the mass, v_j — the mean velocity, and τ_j — the average free-time between the collisions with the molecules of the neutral gas. Under the condition that $T_j = \text{const}$ this equation implies

$$\boldsymbol{q}_j \equiv n\boldsymbol{v}_j = -D_j \nabla n + \frac{e_j}{e} n\mu_j \mathbf{E} \tag{4.56}$$

with $\boldsymbol{q}_j$ standing for the flux of the particles, $D_j = T_j\tau_j/m_j$ for the diffusion coefficient and $\mu_j = e\tau_j/m_j$ for the mobility (we shall put all μ_j down as positive both for the electrons and the ions, taking account of the sign of the charges by the factors e_j/e). Inserting (4.56) into the continuity equation, we obtain the relation

$$\frac{\partial n}{\partial t} = D_j \Delta n - \nabla \cdot \left(\frac{e_j}{e} n\mu_j \mathbf{E} \right) \tag{4.57}$$

of diffusive type. Let us consider, for instance, a cylindrically-symmetric plasma column. Then only the radial part of $\mathbf{E}$ will be nonzero, and we may exclude it by multiplying equation (4.57)

by μ_j, (4.56) by μ_e and summing the results. Thus, we obtain

$$\frac{\partial n}{\partial t} = D_a \Delta n, \tag{4.58}$$

where the quantity

$$D_a = \frac{D_e \mu_i + D_i \mu_e}{\mu_i + \mu_e} \tag{4.59}$$

is called the ambipolar diffusion coefficient. Usually $\mu_e \gg \mu_i$, so that $D_a \approx D_i + (\mu_i/\mu_e) D_e$. From the physical viewpoint the origin of the coefficient D_a is the result of the radial electric field E_r that keeps the electrons near the ions but, nevertheless, "pushes on" the ions.

If the plasma is confined to the interior of a cylinder of the radius a, all solutions to equation (4.58) decay with time. Under the condition that $n(a) = 0$ (Shottky's boundary condition) the solution with the maximum modulus τ of decay appears to be of the form $n = n_0 J_0(\alpha_0 r/a)$, with n_0 standing for the density on the axis, J_0 for the Bessel function of zero order, $\alpha_0 \approx 2.4$ for the first zero of J_0. The corresponding decay modulus is given by the formula $1/\tau = D_a \alpha_0^2/a^2$.

Relation (4.59) may easily be extended to the case of a nonzero longitudinal magnetic field; one needs only to replace D_j, μ_j with their "magnetized" values $D_{\perp j} = D_j/(1 + \Omega_j^2 \tau_j^2)$, $\mu_{\perp j} = \mu_j/(1 + \Omega_j^2 \tau_j^2)$. For instance, in the case $\Omega_e \tau_e \gg 1$, $\Omega_i \tau_i < 1$ we have

$$D_\perp = \frac{D_a}{1 + (\mu_i/\mu_e)(\Omega_e \tau_e)^2}. \tag{4.60}$$

In a strong magnetic field $D_\perp$ decreases with the field as $1/B^2$.

Let us now take an example of a weakly ionized plasma in a toroidal tube with the magnetic field directed along the torus. The totally ionized plasma can not be confined by such a field: it will be pushed to the area of weaker field. But we do not need a total equilibrium, for the weakly ionized plasma will in any case

diffuse across the field due to collisions. We intend only to take account of the torus geometry.

Let the direction of the magnetic field decrease be taken as the x axis. Consider a plasma tube formed by the magnetic lines and denote the volume of its interior with V. This tube must exert the action of a certain force F, trying to push it in the direction of x. In order to evaluate F recall that if there were no friction caused by the neutral gas, the plasma tube would be able to move only in such a way that the magnetic flux $\Phi = BS$ through any cross-section S of the tube should remain constant. Denoting the field line radius by R and an infinitesimal displacement of the tube in the direction of x by dx, we may represent the above statement in differential form:

$$0 = d\Phi = S\frac{dB}{dx}dx + BdS = -\frac{BS}{R}dx + BdS.$$

This relation ensures that $dS = (S/R)dx$. But the displacement dx is followed by the increase $2\pi dx$ of the tube length $2\pi R$, therefore the increase of the tube volume is $dV = (2V/R)dx$. Equating the work $p\,dV$ with Fdx, we obtain the value of the force F, corresponding to the coordinate x: $F = 2pV/R$.

A displacement of the tube in the direction of x does not alter the flux Φ through the tube cross-section. Therefore, the force F will cause the plasma motion, such as if there were no magnetic field at all, i.e., at the speed $v_x = 2D_a/R$, with D_a being the ambipolar diffusion coefficient. This estimate may easily be established by comparing v_x to the ambipolar flow $q = nv = -D_a\nabla n$ caused by the pressure gradient $T\nabla n$. Also taking the transverse diffusion into account (note that the relative coefficient is $D_\perp$), we obtain the equation

$$\frac{\partial n}{\partial t} + \frac{2D_a}{R}\frac{\partial n}{\partial x} = D_\perp \Delta n. \tag{4.61}$$

The second term on the left-hand side describes the displacement of the plasma as a whole to the outer boundary.

If we consider the plasma decay, i.e., the decrease of its density with the time caused by diffusion, it will be sufficient in

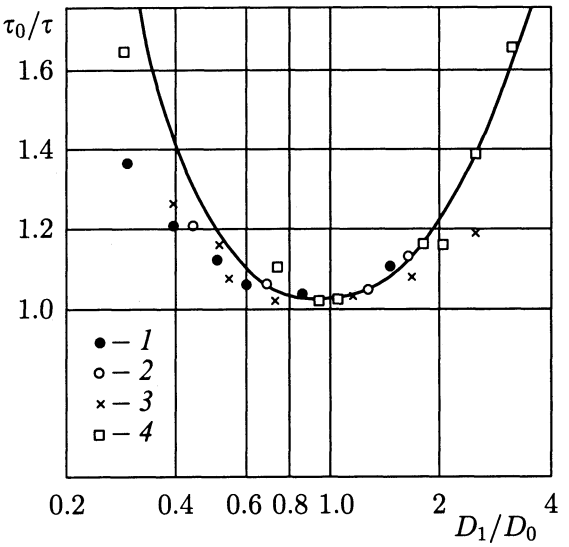

Fig. 4.6. Dependence of the lifetime of a decaying plasma in a toroidal tube upon the diffusion coefficient [Helium, $R = 28\,\text{cm}$, $a = 0.3\,\text{cm}$, pressure (in mm of mercury): *1* — 0.025, 2 — 0.040, *3* — 0.055, *4* — 0.120].

the limit $t \to \infty$ to take account of only a single solution to equation (4.61), namely of that with the maximum lifetime τ. This solution is given by the formula

$$n = n_0 \exp\left(-\frac{t}{\tau} + \frac{D_a}{RD_\perp}x\right) J_0\left(\frac{r\alpha_0}{a}\right) \tag{4.62}$$

with J_0 being the Bessel function of zero order and $\alpha_0 \approx 2.4$ — its first zero. The decay decrement is determined by the relation

$$\frac{1}{\tau} = \frac{D_\perp \alpha_0^2}{a^2} + \frac{D_a^2}{R^2 D_\perp}. \tag{4.63}$$

Thus, with increasing magnetic field (i.e., with decreasing $D_\perp$) the decay decrement $1/\tau$ at first goes down, but later begins to grow. Formula (4.63) is in full agreement (see Fig. 4.6) with the experiments on a plasma decay in a toroidal magnetic field [77].

5. Linear waves

5.1. Alfvén and magneto-acoustic waves

5.1.1. Equation for weak oscillations

Long-range forces between charged particles cause a kind of resiliency in a plasma, which provides the propagation of various types of waves. Linear waves are primarily distinguished among them. The amplitude of these waves is so small that they propagate independently of each other. In other words, arbitrary small-amplitude oscillations can be represented as a superposition of certain waves usually referred to as eigenmodes of the medium.

The major principle of the theoretical treatment of linear oscillations is the following. A stationary state of the object under consideration is found in the framework of a certain system of equations. A disturbance is then imposed on this state. The equations are linearized under the assumption of small magnitude of the perturbation. The resulting system of linear equations is that describing weak oscillations.

It is natural to start studying weak oscillations in a plasma in the the framework of the single-fluid magneto-hydrodynamic approximation. We once more write down the equations of magneto-hydrodynamics under the assumption of ideal conductivity:

$$m_i n \frac{d\boldsymbol{v}}{dt} + \nabla p = \frac{1}{4\pi} (\nabla \times \mathbf{B}) \times \mathbf{B}, \tag{5.1}$$

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\boldsymbol{v}) = 0, \tag{5.2}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \mathbf{B}). \tag{5.3}$$

This system should be completed by the heat transport equation to define the plasma temperature and, consequently, the pressure $p = n(T_e + T_i)$. However, we assume for simplicity

that the pressure is defined by the plasma density, namely

$$\frac{p}{n^\gamma} = \text{const.} \tag{5.4}$$

Equation (5.4) describes an isothermal plasma at $\gamma = 1$ and corresponds to its adiabatic oscillations when $\gamma = 5/3$. In general (5.4) can be treated as a polytropic type equation of state.

Let us now consider the simplest equilibrium state assuming that the plasma is uniformly distributed in space with density n_0 and pressure p_0 while the homogeneous magnetic field $\mathbf{B}_0$ is directed along the z axis of the Cartesian coordinate system (x, y, z). We assume the plasma velocity $\boldsymbol{v}_0$ to be zero.

Impose a small perturbation $\boldsymbol{v}', p', n', \mathbf{B}'$ on this equilibrium state and linearize equations (5.1)–(5.4), i.e., leave only the terms containing small quantities in the first power of the primed variables. We get

$$m_i n_0 \frac{\partial \boldsymbol{v}'}{\partial t} + \nabla p' = \frac{1}{4\pi} (\nabla \times \mathbf{B}') \times \mathbf{B}_0, \tag{5.5}$$

$$\frac{\partial n'}{\partial t} + \nabla \cdot n_0 \boldsymbol{v}' = 0, \tag{5.6}$$

$$\frac{\partial \mathbf{B}'}{\partial t} = \nabla \times (\boldsymbol{v}' \times \mathbf{B}_0), \tag{5.7}$$

$$p' = \frac{\gamma p_0 n'}{n_0} \tag{5.8}$$

instead of (5.1)–(5.4). This system of equations can be essentially simplified if we turn to the plasma displacement $\boldsymbol{\xi}$ from the stationary position, which is connected to $\boldsymbol{v}'$ by the obvious relation

$$\boldsymbol{v}' = \frac{\partial \boldsymbol{\xi}}{\partial t}. \tag{5.9}$$

If we insert this expression for $\boldsymbol{v}'$ into equations (5.6)–(5.7), then after integration in time, i.e., after omitting temporal derivatives, we obtain

$$n' = -\nabla \cdot (n_0 \boldsymbol{\xi}), \quad \mathbf{B}' = \nabla \times (\boldsymbol{\xi} \times \mathbf{B}_0). \tag{5.10}$$

The meaning of the first of these relations is completely obvious. It simply states that the density at a given point decreases proportionally to the fluid volume leaving the domain.

Consider now the second relation of (5.10). Note first that (5.10) contains only the normal component $\boldsymbol{\xi}_\perp$ of the displacement $\boldsymbol{\xi}$ because the vector product of the parallel vectors $\boldsymbol{\xi}$ and $\mathbf{B}_0$ is zero. Using the well known relationship

$$\nabla\times(\mathbf{a}\times\mathbf{b}) = (\mathbf{b}\cdot\nabla)\,\mathbf{a} - (\mathbf{a}\cdot\nabla)\,\mathbf{b} + \mathbf{a}\,(\nabla\cdot\mathbf{b}) - \mathbf{b}\,(\nabla\cdot\mathbf{a}) \quad (5.11)$$

of vector analysis and taking into account that $\mathbf{B}_0 = \text{const}$, we write down the expression for $\mathbf{B}'$ in the form

$$\mathbf{B}' = -\mathbf{B}_0\,\nabla\cdot\boldsymbol{\xi}_\perp + B_0\frac{\partial\boldsymbol{\xi}_\perp}{\partial z}. \quad (5.12)$$

The first term on the right-hand side is analogous to the above expression for n'. It shows that the magnetic field lines are drawn apart and its strength decreases as the substance expands in the (x, y) plane (Fig. 5.1a). The second term on the right-hand side of (5.12) is also of a simple sense.

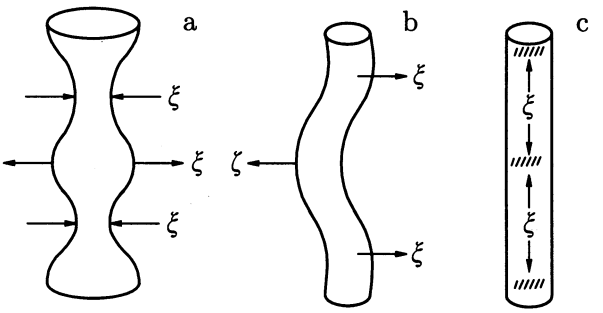

Fig. 5.1. Plasma displacements corresponding to magneto-sound (a), Alfvén (b) and ion-sound (c) oscillations.

If $\boldsymbol{\xi}_\perp$ changes in z, the magnetic field lines, which are frozen to the substance, are distorted to some extent (see Fig. 5.1b). As a result, a magnetic field cross component appears, which is proportional to $\partial\boldsymbol{\xi}_\perp/\partial z$.

Inserting the relations for n', $\mathbf{B}'$ and p' into the linearized equation (5.5), one can obtain the single vector equation for the displacement $\boldsymbol{\xi}$. But at first we shall simplify the right-hand side of (5.5). This can be done most clearly in the following way.

Return again to the equation (5.1) and recall that according to (4.13) we can rewrite the right-hand side in the form

$$\frac{1}{4\pi}(\nabla\times\mathbf{B})\times\mathbf{B} = -\nabla_\perp\frac{B^2}{8\pi} + \frac{B^2}{4\pi}(\mathbf{h}\cdot\nabla)\,\mathbf{h}, \tag{5.13}$$

where $\mathbf{h} = \mathbf{B}/B$ is the unit vector directed along $\mathbf{B}$. If we linearize this expression, then $B_0^2 + 2\mathbf{B}_0\cdot\mathbf{B}'$ should be substituted for B^2 in the first term on the right-hand side and the linear term will contain only the z-component of the disturbance $\mathbf{B}'$ determined by the first term of (5.12). The second term in (5.13) is reduced to $(B^2/4\pi)\,\partial\mathbf{h}'/\partial z$ in the linear approximation, where $\mathbf{h}'$ is the magnetic field disturbance direction. Since

$$\mathbf{h} = \frac{\mathbf{B}_0 + \mathbf{B}'_\parallel + \mathbf{B}'_\perp}{\sqrt{(B_0 + B'_\parallel)^2 + (B'_\perp)^2}}$$

and since $(B'_\perp)^2$ under the square root can be neglected in the linear approximation, we obtain $\mathbf{h} = \mathbf{h}_0 + \mathbf{B}'_\perp/B_0$, with $\mathbf{h}_0$ being the unit vector along the z axis. Consequently, $\mathbf{h}' = \mathbf{B}'_\perp/B_0$ is determined only by the second term of expression (5.12) for the magnetic field disturbances. Inserting the expressions obtained for the Ampère force (5.13) and the disturbances n', p' into the equation of motion (5.5), we obtain

$$\frac{\partial^2\boldsymbol{\xi}}{\partial t^2} = c_s^2\,\nabla\nabla\cdot\boldsymbol{\xi} + c_A^2\,\nabla_\perp\nabla\cdot\boldsymbol{\xi}_\perp + c_A^2\frac{\partial^2\boldsymbol{\xi}_\perp}{\partial z^2}, \tag{5.14}$$

where $c_s = \sqrt{\gamma p_0/n_0}$ denotes the sound and $c_A = B_0/\sqrt{4\pi m_i n_0}$ the Alfvén velocity. Equation (5.14) describes small magnetohydrodynamic oscillations of an ideally conductive medium.

5.1.2. Alfvén waves

The vector equation (5.14) corresponds to three scalar equations, hence it describes three types of waves. Alfvén waves, named after H. Alfvén, are the most interesting and essential among them for understanding of the properties of a conductive fluid in a magnetic field.

In order to obtain the equation for Alfvén wave propagation let us consider a narrower class of displacements $\boldsymbol{\xi}$ which is characterized by the absence of a displacement ξ_z along the field and by the incompressibility $\nabla\cdot\boldsymbol{\xi}_\perp = 0$ of the cross-field motion. It is obvious that also $\nabla\cdot\boldsymbol{\xi} = 0$ and the first two terms on the right-hand side of equation (5.14) vanish, yielding the equation for $\boldsymbol{\xi}_\perp$:

$$\frac{\partial^2\boldsymbol{\xi}_\perp}{\partial t^2} = c_A^2\,\frac{\partial^2\boldsymbol{\xi}_\perp}{\partial z^2}. \tag{5.15}$$

This equation coincides precisely with the equation of small-amplitude vibrations of a string. Its arbitrary solution can be represented as a superposition of waves propagating along the z axis at the speed $\pm c_A$. Analogy with a string arises not only in connection with the formal similarity of these equations, but bears a more keen physical sense. Actually, the last term of (5.14) is connected with the field line tension that is described by the second term of expression (5.13) for the Ampère force. This tension creates the quasi-resilient force that returns the plasma to the state of equilibrium.

Alfvén waves propagate only along the field lines as is seen from equation (5.15). If we generate a torsion-like plasma displacement (Fig. 5.2a) or a more complex incompressible motion with $\nabla\cdot\boldsymbol{\xi}_\perp = 0$ in some plane, then the corresponding disturbance will run along the field lines at the Alfvén velocity and will be concentrated within the cylinder where the initial displacement was not zero.

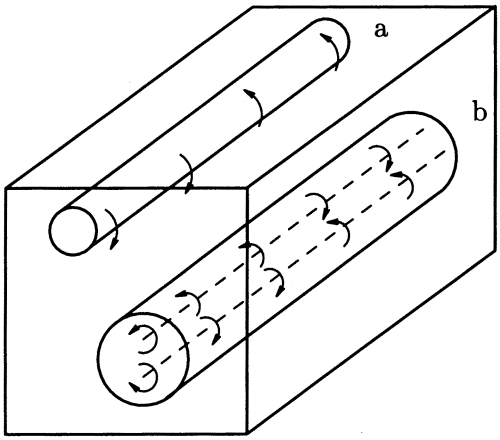

Fig. 5.2. Plasma displacements in Alfvén waves.

5.1.3. Magneto-acoustic waves

Consider now two other types of oscillations with nonzero ξ_z and $\nabla\cdot\boldsymbol{\xi}_\perp$. It is convenient to derive the equations exactly for these quantities. We have from (5.14) for the longitudinal displacement component

$$\frac{\partial^2\xi_z}{\partial t} = c_s^2\frac{\partial^2\xi_z}{\partial z^2} + c_s^2\frac{\partial}{\partial z}(\nabla\cdot\boldsymbol{\xi}_\perp)\,. \tag{5.16}$$

The equation for $\nabla\cdot\boldsymbol{\xi}_\perp$ can be obtained by evaluating the divergence of the transverse component of equation (5.14) and reads as

$$\frac{\partial^2}{\partial t^2}\nabla\cdot\boldsymbol{\xi}_\perp = c_A^2\,\Delta\,(\nabla\cdot\boldsymbol{\xi}_\perp) + c_s^2\,\Delta_\perp\,(\nabla\cdot\boldsymbol{\xi}_\perp) + c_s^2\,\Delta\frac{\partial\xi_z}{\partial z}, \tag{5.17}$$

where $\Delta_\perp = \nabla\cdot\nabla_\perp$, $\Delta = \Delta_\perp + \partial^2/\partial z^2$.

Consider first the case $\beta \sim c_s^2/c_A^2 \ll 1$ of a low pressure plasma. Then the terms proportional to c_s^2 can be neglected in equation (5.17) and we come to the wave equation

$$\frac{\partial^2 \psi}{\partial t^2} = c_A^2 \, \Delta \psi, \tag{5.18}$$

where $\psi = \nabla \cdot \boldsymbol{\xi}_\perp$. This equation describes the so-called magneto-sound. The elasticity of the medium in magneto-acoustic oscillations is created by the magnetic field pressure while the Alfvén velocity $c_A = \sqrt{B^2/4\pi m_i n}$ can be treated as the sound velocity $c_s = \sqrt{\gamma p/\rho}$ in a medium with pressure $B^2/8\pi$, density $\rho = m_i n$ and adiabatic coefficient $\gamma = 2$. The condition $\gamma = 2$ is connected with the frozenness of the magnetic field to the plasma. As a consequence, at $\xi_z = 0$ the field should change proportionally to the density while its pressure should be proportional to the squared plasma density.

Equation (5.16) enables us to find a small (the magneto-sound velocity being great) longitudinal displacement ξ_z in magneto-acoustic waves when $\nabla \cdot \boldsymbol{\xi}_\perp \neq 0$. Moreover, equation (5.16) describes a distinct oscillatory mode when $\beta \ll 1$. For these oscillations $\nabla \cdot \boldsymbol{\xi}_\perp$ can be assumed to be zero and we obtain

$$\frac{\partial^2 \xi_z}{\partial t^2} = c_s^2 \frac{\partial^2 \xi_z}{\partial z^2}.$$

These oscillations are known as ion-sound waves. They represent sound propagation as $\beta \to 0$ along the magnetic field with only an along-field plasma displacement.

Thus, three types of displacements (polarizations) correspond to the three wave modes. The displacement in Alfvén waves causes distortion of the field lines (Fig. 5.1b), in magneto-acoustic waves it compresses and decompresses them (Fig. 5.1a), while in ion-sound waves at $\beta \ll 1$ the magnetic field is almost undisturbed (Fig. 5.1c). In the latter case the plasma is, so to speak, one-dimensional, i.e., it is broken into tubes with "rigid magnetic walls."

If β is not small, the above separation of magneto- and ion-sound can not be made. Equations (5.16) and (5.17) should be treated jointly in this case. It is useful to utilize plane waves

for their solution, since in a homogeneous medium any disturbance can be represented as a superposition of plane waves. Consequently, it is sufficient to consider a plane wave of type $\exp(-i\omega t + i\mathbf{k}\cdot\mathbf{r})$. Equations (5.16) and (5.17) become the simple algebraic relations

$$(\omega^2 - c_s^2 k_z^2)\xi_z - k_z c_s^2(\mathbf{k}_\perp \cdot \boldsymbol{\xi}_\perp) = 0, \quad (5.19)$$

$$-k_z c_s^2 k_\perp^2 \xi_z + (\omega^2 - c_A^2 k^2 - c_s^2 k_\perp^2)(\mathbf{k}_\perp \cdot \boldsymbol{\xi}_\perp) = 0 \quad (5.20)$$

for such a wave. For this system to have nontrivial solutions, its determinant should be equal to zero. This condition leads to a relationship between the frequency ω and the wavevector $\mathbf{k}$, which is usually referred to as the dispersion equation. It is easy to find that in our case the dispersion equation has the form

$$\omega^4 - (c_A^2 + c_s^2)k^2\omega^2 + c_A^2 c_s^2 k_z^2 k^2 = 0. \quad (5.21)$$

The greater root ω_+ of this equation corresponds to the so-called fast magneto-acoustic wave while the smaller one ω_- describes the slow wave. When $c_s^2 \ll c_A^2$, the fast wave turns to magneto-sound while the slow wave is transformed into ion-sound.

Equation (5.21) simplifies essentially in two particular cases of purely longitudinal and transverse propagation. In the case of longitudinal propagation $k^2 = k_z^2$ and it follows from (5.21) that $\omega_+^2 = c_A^2 k_z^2$ for the fast wave and $\omega_-^2 = c_s^2 k_z^2$ for the slow wave for $c_A > c_s$. In the case of transverse propagation $k_z = 0$ and in accordance with (5.21) $\omega_+^2 = (c_A^2 + c_s^2)k^2$, $\omega_-^2 = 0$.

In the general case the expressions for ω_+^2 and ω_-^2 can be easily obtained from (5.21), but they are hard to picture. Polar diagram of phase velocities gives a much better visual aid.

5.1.4. Polar diagram and Frederick's diagram

The polar phase diagram represents the dependence of phase velocity $v_p = \omega/k$ on the angle θ between the vector $\mathbf{k}$ and the magnetic field $\mathbf{B}$ in the polar coordinate system v_p, θ. It is of the form shown in Fig. 5.3a.

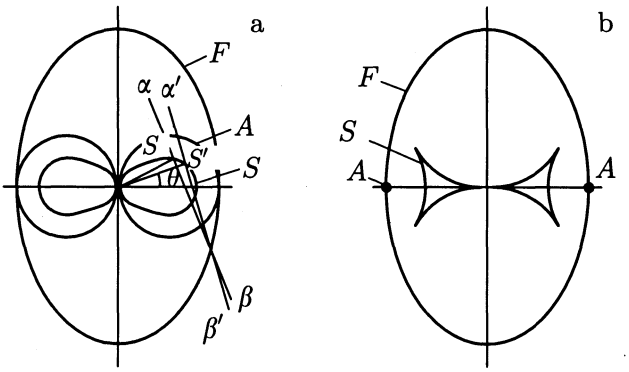

Fig. 5.3. Phase (a) and group (b) diagrams: A — the Alfvén wave, S and F — slow and fast magneto-acoustic waves.

For a fast magneto-sound wave the phase diagram has the shape of an oval flattened along the the magnetic field. For a slow wave it looks like a pair of ovals touching each other and flattened in the transverse direction (Fig. 5.3a). When $\beta \to 0$, the big oval transforms into a circle of radius c_A, and the ion-sound wave diagram turns into two contiguous circles of diameter c_s. In the limit $c_s/c_A \to \infty$ we come to an incompressible fluid, and the ion-sound polar diagram transforms into two contiguous circles of diameter c_A. Exactly the same diagram corresponds to the Alfvén wave at arbitrary β.

In addition to the phase diagram one can construct the phase velocity surface for arbitrary values of $\mathbf{k}$ in a three-dimensional space. For simplicity we shall restrict ourselves, however, to the two-dimensional case.

As we shall see further, the group velocity $\boldsymbol{v}_g = \partial\omega/\partial\mathbf{k}$ phase diagram, i.e., the group velocity dependence on the direction, is practically more interesting. The group polar curve is usually called Frederick's diagram (sometimes both the phase and group velocity curves are thus called).

As is known, the group velocity determines the speed of a wave packet. In the case of a non-dispersive medium, when the

frequency ω is proportional to k, the group velocity is independent of the absolute value of $\mathbf{k}$. This means that all disturbances propagating in a given direction have the same speed. In particular, a pulse-shaped perturbation will conserve its form in time. Therefore, the group polar diagram represents the pattern of a pulse propagation source: a small wave packet propagating at the group speed can be attributed to each piece of the diagram.

Frederick's diagram can be obtained analytically by evaluating $\partial\omega/\partial\mathbf{k}$. However, a graphical construction appears to be more visual and, besides, enables us to take another view on the process of disturbance propagation from a point source. Taking into account that $\omega = kv_p$, where v_p depends only upon the direction of $\mathbf{k}$, we have

$$\boldsymbol{v}_g = \frac{\partial\omega}{\partial\mathbf{k}} = \frac{\mathbf{k}}{k}v_p + \boldsymbol{v}_{\perp g},$$

where the component $\boldsymbol{v}_{\perp g} = k\partial v_p/\partial\mathbf{k}$ is normal to the wavevector $\mathbf{k}$. As we see, the group velocity component normal to the wave surface coincides with the phase velocity. To proceed with interpretation of the diagram, let us break the disturbance into small coin-like pieces oriented perpendicularly to $\mathbf{k}$. Every "coin" may be treated as a wave packet, i.e., as an area occupied with plane waves with approximately coinciding wavevectors. Each packet moves at the speed v_p along $\mathbf{k}$ and in addition "slides" in the transverse direction at the speed $\boldsymbol{v}_{\perp g}$ (the phase speed component transverse to $\mathbf{k}$ of course has no sense).

Consequently, the polar group curve can be approximately regarded as a "mosaic" consisting of separate small "coins" — the wave packets. Using these considerations, it is easy to find out how to construct the polar group curve via the phase one. Take an arbitrary point M of the phase curve and draw the line $\alpha\beta$ normal to the radius vector of M, i.e., to $\mathbf{k}$ (Fig. 5.3a). The group packet $\mathbf{k}$ "coin" should be somewhere on that line. Take now some point M' of the polar phase curve near M and draw a line $\alpha'\beta'$ normal to the radius vector of M'. A wave packet containing plane waves with wavenumbers from the neighbor-

hood of M' should lie on the line $\alpha'\beta'$. But the wave packets of the corresponding planes will overlap for closely located M and M' and form a single packet. Therefore this packet will lie on the intersection of the lines $\alpha\beta$ and $\alpha'\beta'$ that correspond to the near-by points M and M'.

Thus, for constructing the polar group curve via the phase one it is sufficient to draw a family of strait lines orthogonal to the radius vectors of the points of the polar phase curve and to find the envelope of these lines. A Frederick diagram built in such a way is shown in Fig. 5.3b. It has the form of an oval for the fast magneto-acoustic wave, of two "sharpened" triangles for the slow wave and degenerates into two points of the outer oval in the case of the Alfvén waves. The polar group curve shows how a pulse propagates from a momentary point source: the corresponding disturbance has nonzero values in the area between the slow and the fast polars with an increase in amplitude near the boundaries, i.e., near the polars.

An interesting peculiarity of the Frederick diagram arises in the case $\beta = 8\pi p/B^2 = 1$ when the Alfvén velocity is equal to the sound velocity [78]. In this case according to (5.21) $\omega_{\pm}^2 = c_A^2 k^2(1 \pm \sin\theta)$ where θ is the angle between $\mathbf{k}$ and $\mathbf{B}$, i.e. $\cos\theta = k_z/k$. A corner point at $\theta = 0$ is formed on the polar phase curve and the group velocity vector does not tend to horizontal position when $\theta \to 0$, but constitutes the finite angle of approximately $\theta_0 \simeq 26°$ with the horizon. Under these conditions a phenomenon analogous to internal conic refraction in optics should take place: a "beam" of magneto-hydrodynamic waves propagating along $\mathbf{B}_0$ normally to the boundary should open in the form of a cone with a spreading angle of $2\theta_0$. We do not know if anybody has observed this phenomenon in an experiment.

5.1.5. Small body streamlining

The Frederick diagram provides an easy analysis of the streamlining pattern of a thin needle-like body moving in the

cross-field direction. Assume that the body's length is so great that it can be considered infinite and the magnetic field is directed normally to the body's axis. We also assume that the body conductivity is not large and it does not "reel" the field lines. Under these conditions the streamlining pattern is purely two-dimensional and any disturbance far from the body may be assumed small. By analogy with ordinary gas dynamics we can suppose that a moving body can be represented as a series of momentary point sources that produce pulses in the form of expanding Frederick diagrams. At the spots where these pulses are properly phased Mach type waves arise, while at other points much smaller perturbations occur because of interference (Fig. 5.4a, b).

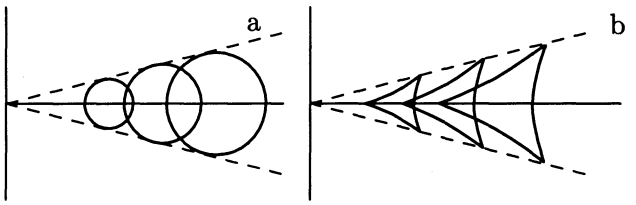

Fig. 5.4. Cherenkov emission in an isotropic medium (a) and for gas hydromagnetics (b).

In order to determine the Mach waves, it is sufficient to put the body's velocity vector on the Frederick diagram and draw the lines tangent to the diagram as shown in Fig. 5.5 for various cases. The flows in Fig. 5.5a, c are double-hyperbolic, while the flows in Fig. 5.5b, d are of mixed hyperbolic-elliptic type with the only type of radiated waves.

As is seen from Fig. 5.5c, d, perturbations in magnetohydrodynamics may propagate upstream, i.e., ahead of the moving body. It should be stressed that these waves correspond to ordinary Cherenkov emission of a body moving at a speed greater than the phase speed of the transmitted waves. The unusual possibility of wave radiation ahead of the moving body results from the group, but not the phase diagram's, pulse interference.

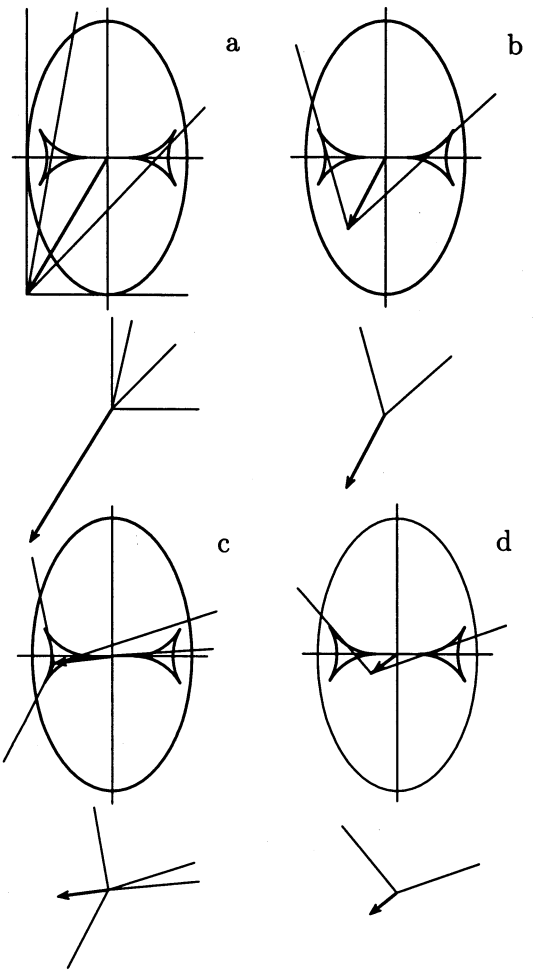

Fig. 5.5. Four cases of Mach line formation. Lines leave the generating body whose velocity is denoted by a bold vector.

5.2. Wave dispersion in two-fluid hydrodynamics

5.2.1. Ion-sound

Magneto-hydrodynamics describes large-scale low-frequency motion in a plasma sufficiently well, but it is still a kind of

approximation. It is natural that certain effects may occur in plasma wave motion that are not covered by one-fluid hydrodynamics. One of these effects is dispersion, i.e., the dependence of the phase speed on the frequency.

As we have found before, an ideally conductive gas is an anisotropic but non-dispersive medium: the phase speed of its low-amplitude oscillations depends only upon the wavevector direction, but not upon its magnitude. This fact has a major role in the evolution of finite-amplitude perturbations. Therefore one should have a clear understanding of the dispersion behavior up to which a plasma can be treated as a non-dispersive medium.

We consider ion-sound assuming that $\beta \ll 1$ and, consequently, neglecting the magnetic field perturbations caused by oscillations. This means that the oscillatory electric field can be treated as a potential one:

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \simeq 0.$$

To determine the sound dispersion, we should utilize two-fluid instead of the rougher one-fluid hydrodynamics. In the considered case of potential oscillations the equations are essentially simplified because the magnetic field oscillations are not to be taken into account. We shall further simplify our analysis by assuming that the ions are cold, i.e., $T_i = 0$ (in practice this means that the ion temperature is much lower than the electron one). The more general case of a rarefied plasma with $T_i \neq 0$ demands, generally speaking, a kinetic consideration because the ion-sound phase speed is not large and is of the order of the thermal ion velocity at $T_i \sim T_e$.

Thus, neglecting the electron inertia and their collisions with the ions, we have at $T_i = 0$:

$$\frac{d\boldsymbol{v}}{dt} = -\frac{e}{m_i}\nabla\varphi + \frac{e}{m_i c}(\boldsymbol{v} \times \mathbf{B}), \tag{5.22}$$

$$\frac{\partial n_i}{\partial t} + \nabla \cdot n_i \boldsymbol{v} = 0, \tag{5.23}$$

$$\nabla n_e T_e = e n_e \nabla\varphi - \frac{e n_e}{c}\,(\boldsymbol{v}_e \times \mathbf{B}), \tag{5.24}$$

$$\Delta\varphi = -4\pi e(n_i - n_e). \tag{5.25}$$

The Poisson equation (5.25) relates the potential φ to the ion and electron density perturbations.

We shall assume for simplicity that the electrons have constant temperature $T_e = \text{const}$. This is a completely admissible assumption for waves propagating at a phase speed much lower than that of the electron thermal motion. From the longitudinal component

$$T_e \frac{\partial n_e}{\partial z} = e n_e \frac{\partial\varphi}{\partial z} \tag{5.26}$$

of the electron equilibrium equation (5.25) it follows that the electrons are subjected to Boltzmann's distribution along the magnetic field lines. For the considered low-amplitude oscillations in a homogeneous plasma we have from (5.26)

$$n_e' = n_0 \frac{e\varphi}{T_e}. \tag{5.27}$$

Let us now find the ion density perturbation n_i'. To do this we can utilize the equation of motion (5.22) and the continuity equation (5.23). From the longitudinal component of the equation of motion we obtain the relation

$$v_z = \frac{e}{m_i}\frac{k_z}{\omega}\varphi \tag{5.28}$$

for a perturbation in the form of a plane wave.

In order to find the transverse ion velocity it is convenient to introduce the coordinate system (x, y, z) with the z axis directed along the magnetic field and the y axis normal to $\mathbf{k}$. Then only the components $k_x = k_\perp$ and k_z of $\mathbf{k}$ will be nonzero. As a result, the ion equation of motion projected onto the axes x, y assumes the form

$$-i\omega v_x - \Omega_i v_y = \frac{e}{m_i}E_x = -\frac{e}{m_i}ik_\perp\varphi, \tag{5.29}$$

$$\Omega_i v_x - i\omega v_y = 0 \tag{5.30}$$

with $\Omega_i = eB/m_i c$. Using these equations, we easily find the velocity components

$$v_x = -\frac{i\omega}{\Omega_i^2 - \omega^2}\frac{e}{m_i}E_x = \frac{\omega}{\Omega_i^2 - \omega^2}\frac{e}{m_i}k_\perp \varphi, \tag{5.31}$$

$$v_y = \frac{\Omega_i}{\Omega_i^2 - \omega^2}\frac{e}{m_i}E_x = \frac{\Omega_i}{\Omega_i^2 - \omega^2}\frac{e}{m_i}k_\perp \varphi \tag{5.32}$$

together with

$$\nabla \cdot \boldsymbol{v}_\perp = -\frac{i\omega}{\Omega_i^2 - \omega^2}\frac{e}{m_i}k_\perp^2 \varphi.$$

Inserting the expression obtained into the continuity equation

$$-\omega n_i' + n_0\,(\nabla \cdot \boldsymbol{v}_\perp) + n_0 i k_z v_z = 0 \tag{5.33}$$

and taking formula (5.28) for the longitudinal velocity component into account, we obtain the ion density perturbation

$$n_i' = \left(\frac{k_z^2 c_s^2}{\omega^2} - \frac{k_\perp^2 c_s^2}{\Omega_i^2 - \omega^2}\right) n_0 \frac{e\varphi}{T_e} \tag{5.34}$$

with $c_s^2 = T_e/m_i$.

Insert now the above expressions (5.27) and (5.34) for the electron and ion densities into the Poisson equation (5.25), rewritten for a plane wave in the form

$$k^2 d^2 n_0 \frac{e\varphi}{t_e} = n_i' - n_e', \tag{5.35}$$

where $d^2 = T_e/4\pi e^2 n_0$ is the square of the Debye radius. After substitution we obtain the dispersion equation that determines the frequency ω of oscillations:

$$1 + k^2 d^2 = \frac{k_z^2 c_s^2}{\omega^2} - \frac{k_\perp^2 c_s^2}{\Omega_i^2 - \omega^2}. \tag{5.36}$$

At first let us consider the case of longitudinal propagation, when $k_\perp = 0$, $k = k_z$. Then

$$\omega^2 = \frac{k_z^2 c_s^2}{1 + k_z^2 d^2}. \tag{5.37}$$

For $k_z d \ll 1$ we get the familiar expression $\omega = c_s k_z$ for the ion-sound frequency, which follows from one-fluid hydrodynamics. But at high k_z the frequency grows more slowly than k_z and the phase speed ω/k_z decreases. It tends to zero when $k_z \to 0$. The ion oscillations are going on with practically immobile electrons in this case. The electron pressure balances the electric field and prevents the electrons from being displaced.

An analogous scenario takes place in the absence of the magnetic field as well. If we put $\Omega_i = 0$ in (5.34), we get the same expression with only k_z replaced by k:

$$\omega^2 \equiv \omega_1^2 = \frac{k^2 c_s^2}{1 + k^2 d^2}. \tag{5.38}$$

Consider now the more general case of magneto-sound oscillations in a magnetic field at arbitrary values of k. Using the notation ω_1^2 for the squared frequency (5.38) in the absence of the magnetic field, we can rewrite the dispersion equation (5.36) in the form

$$\omega^4 - (\Omega_i^2 + \omega_1^2)\omega^2 + \Omega_i^2 \omega_1^2 \frac{k_z^2}{k^2} = 0. \tag{5.39}$$

It follows from here that

$$\omega^2 = \frac{1}{2}(\Omega_i^2 + \omega_1^2) \pm \sqrt{\frac{1}{4}(\Omega_i^2 + \omega_1^2)^2 - \frac{k_z^2}{k^2}\Omega_i^2 \omega_1^2}. \tag{5.40}$$

When $k \to 0$, there are two roots: $\omega^2 = \Omega_i^2$ and $\omega^2 = c_s^2 k_z^2$. As the wavenumber k increases, the slow-wave phase speed appears to depend on k. The frequency dependence upon k shown in Fig. 5.6 is plotted for the case of a dense plasma when $c_A < c$ (c is the light velocity), i.e., when $\Omega_i < \omega_{pi} = \sqrt{4\pi e^2 n_0/m_i}$.

At large k the slow-wave frequency tends to $\Omega_i \cos\theta = \Omega_i k_z/k$ and the fast-wave frequency approaches ω_1. The slow-wave of intermediate propagation starts to disperse at lower wavenumbers $k \sim \Omega_i/c_s$ than that propagating along the magnetic field. If a wave propagates along the magnetic field, the

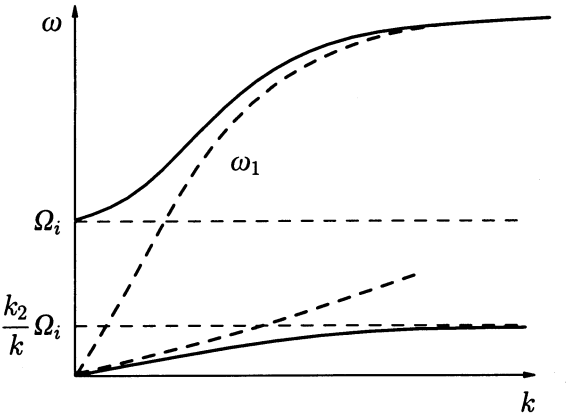

Fig. 5.6. Ion-sound dispersion.

dispersion law $\omega = \omega(k)$ is represented by a pair of intersecting curves $\omega = \Omega_i$ and $\omega = \omega_1$ shown by the dashed lines in Fig. 5.6.

5.2.2. Magneto-sound waves (cross-field propagation)

Consider now magneto-sound dispersion. Assume that a wave propagates precisely across the magnetic field, say, along the x axis. As we shall see, the dispersion in this case is connected with the electron inertia, which should be taken into account within the generalized Ohm's law written in the form of the electron equation of motion. Neglecting the electron pressure, we may write this equation for a plane wave [proportional to $\exp(-\omega t + i\mathbf{k}\cdot\mathbf{r})$] in the following way:

$$-i\omega \boldsymbol{v}_e = -\frac{e}{m_e}\mathbf{E} - \frac{eB_0}{c}\,(\boldsymbol{v}_e \times \mathbf{h})\,. \tag{5.41}$$

Here $\mathbf{h} = \mathbf{B}_0/B_0$ is the unit vector along the z axis. Using (5.41) and recalling that $\omega \ll \Omega_e = eB_0/m_e c$, we obtain approximately

$$\boldsymbol{v}_e = \frac{c}{B_0}\,(\mathbf{E} \times \mathbf{h}) + \frac{i\omega}{\Omega_e}\frac{c}{B_0}\,\mathbf{E}. \tag{5.42}$$

If we used only the first term of the right-hand side of (5.42), we would have neglected the electron inertia in comparison to their drift. In accordance with (5.41) this case corresponds to ideal electron conductivity and, as a consequence, to the magnetic field frozenness to the electron component. But the electrons are displaced together with the ions within a purely cross-field wave in a quasi-neutral plasma, hence the magnetic field frozenness to the electrons implies its frozenness to the ions as well. The latter property, as we have already stated in Section 1, leads to a non-dispersive magneto-sound wave.

Summing up, for dispersion to appear, the electron velocity x-component should differ from that of the electron drift. In other words, a small term that corresponds to the so-called inertial drift should not be neglected in the expression for the electron velocity x-component:

$$v_{ex} = \frac{cE_y}{B_0} + \frac{i\omega}{\Omega_e}\frac{cE_x}{B_0}. \tag{5.43}$$

As for the y-component of the velocity, its former value can be left unchanged:

$$v_{ey} = -\frac{c}{B_0}E_x. \tag{5.44}$$

Using this relation, E_x can be expressed in terms of $v_{ey} = -j_y/en_0$, where the y-component of the current density is connected with the magnetic field perturbation B' via the formula

$$j_y = -\frac{c}{4\pi}\frac{\partial B'}{\partial x} = -i\frac{ck}{4\pi}B'. \tag{5.45}$$

It is easy to express B' in terms of E_y using the Maxwell equation $\partial\mathbf{B}/\partial t = -c\,\nabla\times\mathbf{E}$:

$$B' = \frac{ck}{\omega}E_y. \tag{5.46}$$

With the help of relations (5.44)–(5.46) one can express E_x in terms of E_y and substitute the result into expression (5.43) for

the electron velocity. Then we get

$$v_{ex} = \left(1 + \frac{c^2 k^2}{\omega_{pe}^2}\right) \frac{cE_y}{B}, \tag{5.47}$$

where

$$\omega_{pe} = \sqrt{\frac{4\pi e^2 n_0}{m_e}} \tag{5.48}$$

is the Langmuir frequency.

In order to obtain the dispersion relation, let us use the linearized equation of motion

$$m_i n_0 \frac{\partial v_x}{\partial t} = \frac{\partial}{\partial x} \frac{B_0 B'}{4\pi}, \tag{5.49}$$

which is simply the sum of the equations of motion for the ions and the electrons with the small inertial electron term omitted. Using the expression for B', we obtain for a plane wave

$$v_x = \frac{c_A^2 k^2}{\omega^2} \frac{cE_y}{B_0}. \tag{5.50}$$

Since the velocities v_x and v_{ex} appear to be equal due to the quasi-neutrality of a plasma, we get from (5.47) and (5.50)

$$\omega^2 = c_A^2 k^2 \frac{\omega_{pe}^2}{\omega_{pe}^2 + c^2 k^2}. \tag{5.51}$$

The corresponding dependence of ω upon k is shown in Fig. 5.7. The frequency grows linearly with k when k is not large, while as $k \sim \omega_{pe}/c$ the growth rate decreases and at $k \to \infty$ the frequency tends to its limiting value $c_A \omega_{pe}/c = \sqrt{\Omega_e \Omega_i}$, i.e., to the geometrical mean of the ion and electron frequencies. This frequency is called the lower hybrid frequency.

The picture considered relates only to the case of a low pressure plasma with $\beta \ll 1$. At $\beta \simeq 1$ the dispersion appears much sooner, namely at $\omega \sim \Omega_i$. The cause is that the adiabatic coefficient $\gamma = 2$ for $\omega \ll \Omega_i$ in a collisionless plasma, while for

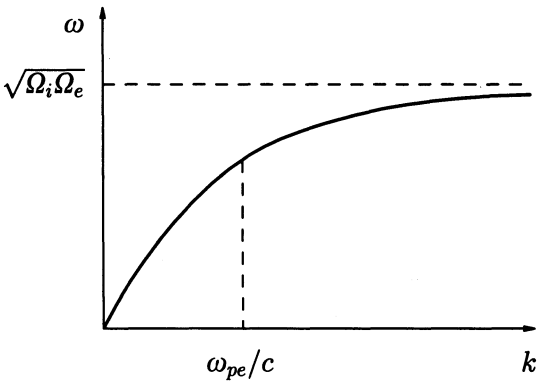

Fig. 5.7. Magneto-sound dispersion in the case of transverse propagation.

$\omega \gg \Omega_i$ it becomes equal to 3. This is because only one degree of freedom participates in oscillations in the first case but a pair participate in the second. Consequently the phase speed starts growing when $\omega \sim \Omega_i$. Wavelengths of the order of gyroradius correspond to that frequency.

5.2.3. Oblique magneto-sound wave

The dispersion reveals itself much earlier in an oblique magneto-acoustic wave as well as in an oblique ion-sound wave. This is a result of the fact that quasi-neutrality can be provided by an electron along-field flow and the dispersion should appear as soon as the difference between the ion and electron transverse velocities becomes significant.

The electron inertia has no role in this case. As we found while considering plasma equilibrium, this is also the case when the magnetic field is frozen to the electron gas at $T_e = \text{const}$:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\boldsymbol{v}_e \times \mathbf{B}) . \tag{5.52}$$

Under the assumption of zero plasma pressure the equation of

motion takes the form

$$m_i n \frac{d\boldsymbol{v}}{dt} = \frac{1}{c}(\mathbf{j} \times \mathbf{B}) = -\nabla \frac{B^2}{8\pi} + \frac{1}{4\pi}(\mathbf{B} \cdot \nabla)\mathbf{B}. \tag{5.53}$$

Recalling that $\boldsymbol{v}_e = \boldsymbol{v} - \mathbf{j}/en$ and taking equation of motion (5.53) into account, we can rewrite (5.52) in the form

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\boldsymbol{v} \times \mathbf{B}) - \frac{m_i c}{e} \nabla \times \frac{d\boldsymbol{v}}{dt}. \tag{5.54}$$

The last term of this equation appears due to the Hall effect and causes the dispersion of magneto-acoustic waves.

If we linearize the equation of motion (5.53) and the magnetic field equation (5.54), introduce the plasma displacement $\boldsymbol{\xi}$ and use the transformations from Section 1, a system of equations for a plane wave will be obtained:

$$-m_i n_0 \omega^2 \boldsymbol{\xi} = -i\mathbf{k}_\perp \frac{B_0 B'_z}{4\pi} + i\frac{b_0}{4\pi} k_z \mathbf{B}'_\perp, \tag{5.55}$$

$$\mathbf{B}' = -i\mathbf{B}_0(\mathbf{k} \cdot \boldsymbol{\xi}_\perp) + iB_0 k_z \boldsymbol{\xi}_\perp + \frac{\omega m_i c}{e}(\mathbf{k} \times \boldsymbol{\xi}). \tag{5.56}$$

From the first equation it is seen that $\xi_z = 0$, while the equations in ξ_x and ξ_y can be obtained on substituting the right-hand side of (5.56) for $\mathbf{B}'$ in (5.55). For a plane wave of the form

$$\exp(-i\omega t + ik_\perp x + izk_z)$$

with $k_y = 0$ and $k_x = k_\perp$ the outlined equations take the form

$$(\omega^2 - c_A^2 k^2)\xi_x - i\omega \frac{k^2 c B_0}{4\pi e n_0} \xi_y = 0, \tag{5.57}$$

$$i\omega \frac{k_z^2 c B_0}{4\pi e n_0} \xi_x + (\omega^2 - c_A^2 k_z^2)\xi_y = 0. \tag{5.58}$$

Putting the determinant of the system equal to zero, we obtain the dispersion relation:

$$(\omega^2 - c_A^2 k^2)(\omega^2 - c_A^2 k_z^2) - \omega^2 k^2 k_z^2 c_A^4 / \Omega_i^2 = 0. \tag{5.59}$$

Consider, for simplicity, the case of nearly transverse propagation at $k_z^2 \ll k^2$. Then for a magneto-acoustic wave with $\omega \gg c_A k_z$ we have

$$\omega^2 = c_A^2 k^2 \left(1 + \frac{c_A^2 k_z^2}{\omega_i^2}\right). \tag{5.60}$$

The corresponding dispersion curve is shown in (Fig. 5.8). As is seen, the frequency grows linearly with $\theta = k_z/k$ at small k and begins to grow much faster when $k \sim \Omega_i/c_A\theta$.

In the case of $\theta = k_z/k \gg \sqrt{m_e/m_i}$ an oblique wave begins dispersing much earlier than a purely transverse one with the same k. It is the only case when neglecting the electron inertia is correct. The only waves we shall call oblique are those characterized by $k_z/k \gg \sqrt{m_e/m_i}$. It should be noted that at large k the unity within the brackets in (5.60) can be neglected and the frequency (5.60) becomes equal to

$$\omega = \frac{c_A c k k_z}{\omega_{pi}} = \frac{c B_0 k k_z}{4\pi e n_0}. \tag{5.61}$$

This value does not depend upon the ion mass m_i. In this case the electrons oscillate together with the magnetic field lines being frozen to them, and the ions remain practically motionless. These oscillations are called helicons or whistlers.

5.2.4. Alfvén waves

The second root of equation (5.59) corresponds to the Alfvén wave. We shall again restrict ourselves to the case of almost transverse propagation with $k_z^2 \ll k^2$. At $\omega^2 \ll c_A^2 k^2$ we have the following estimate for that root:

$$\omega^2 = c_A^2 k_z^2 \frac{\Omega_i^2}{\Omega_i^2 + c_A^2 k_z^2}. \tag{5.62}$$

This relation gives us the dispersion law of the Alfvén waves. As we see, the dependence of ω upon k_z departs from linear with

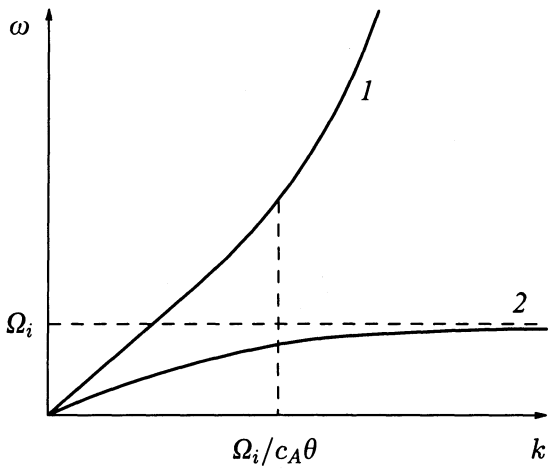

Fig. 5.8. Magneto-sound *1* and Alfvén wave *2* dispersion in the case of oblique propagation.

an increase of k_z, and as $k_z \to \infty$ the frequency tends to the ion cyclotron frequency Ω_i. The corresponding dependence of frequency upon k is also shown in (Fig. 5.8). The Alfvén waves at large k_z with frequencies close to the cyclotron frequency are sometimes referred to as cyclotron waves.

5.3. Electromagnetic waves in a plasma

5.3.1. Dielectric permeability, dispersion equation

So far we have been dominantly using the approach of gas dynamics in the analysis of magneto-hydrodynamic waves in a plasma: we specified displacement of the particles, then found forces caused by that displacement, and, finally, determined the types of oscillations. But these oscillations can be treated from another — purely electrodynamical — viewpoint. Naturally, any motion in a plasma generates electric and magnetic fields, and from the viewpoint of electrodynamics wave propagation in a plasma can be treated as the propagation of electromagnetic fields in a continuous medium. Mathematically

this corresponds to another succession of equation reduction: instead of determining the fields and forces via the displacements of the plasma components, we must first find the charges and the currents through the laws of motion of the particles in a given external fields and then consider how such fields can propagate in the plasma. Such an approach is of course equivalent to the first but, being more general, it enables us to take a broader view of the waves in a plasma and gives us the possibility to obtain a more complete description of all types of linear oscillations in a plasma.

Thus, we shall treat a plasma as a continuous medium which is capable of transfering electromagnetic waves. This time taking the displacement current into account, write down once more the Maxwell equations:

$$\nabla \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\mathbf{j}, \tag{5.63}$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \tag{5.64}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{5.65}$$

$$\nabla \cdot \mathbf{E} = 4\pi\rho. \tag{5.66}$$

Here ρ is the charge density, satisfying the equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \tag{5.67}$$

that is a consequence of (5.63) and (5.66).

If linear waves (i.e., of a small amplitude) are considered, ρ and $\mathbf{j}$ can be linearly expressed in terms of the fields $\mathbf{E}$ and $\mathbf{B}$ with the aid of the equations of motion for the charged plasma particles. Maxwell's equations will therefore be linear and may be rewritten in a more compact form by introducing the electric displacement $\mathbf{D}$, which is defined by the relation

$$\frac{\partial \mathbf{D}}{\partial t} = \frac{\partial \mathbf{E}}{\partial t} + 4\pi\mathbf{j}. \tag{5.68}$$

Correspondingly, equations (5.63) and (5.66) take the form

$$\nabla \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t}, \tag{5.69}$$

$$\nabla \cdot \mathbf{B} = 0. \tag{5.70}$$

Let us consider again, for simplicity, a homogeneous medium. In these circumstances an arbitrary solution to the linear equations is a superposition of plane waves $\exp(i\omega t + i\mathbf{k} \cdot \mathbf{r})$. For a plane wave the Maxwell equations become algebraic:

$$\mathbf{k} \times \mathbf{B} = -\frac{\omega}{c}\mathbf{D}, \tag{5.71}$$

$$\mathbf{k} \times \mathbf{E} = \frac{\omega}{c}\mathbf{B}. \tag{5.72}$$

By virtue of relations (5.71) and (5.72), equations (5.65) and (5.70) are satisfied:

$$\mathbf{k} \cdot \mathbf{D} = 0, \quad \mathbf{k} \cdot \mathbf{B} = 0. \tag{5.73}$$

Expressing $\mathbf{B}$ in terms of $\mathbf{E}$ with the help of (5.72) and substituting the result into (5.71), it is easy to obtain

$$k^2\mathbf{E} - \mathbf{k}(\mathbf{k} \cdot \mathbf{E}) - \frac{\omega^2}{c^2}\mathbf{D} = 0. \tag{5.74}$$

As we have already stated, the vector $\mathbf{D}$ can be expressed via $\mathbf{E}$ through the equations of motion for charged particles. For a plane wave the linearized equations of motion also take an algebraic form. Therefore $\mathbf{D}$ and $\mathbf{E}$ must be related by a linear transformation. It is known that in the general case a linear relationship between vectors can be written in the form of $\mathbf{D} = \hat{\varepsilon}\,\mathbf{E}$, or, using components, as

$$D_\alpha = \sum_\beta \varepsilon_{\alpha\beta} E_\beta, \tag{5.75}$$

where $\hat{\varepsilon}$ stands for the matrix $\varepsilon_{\alpha\beta}$. The tensor $\hat{\varepsilon}$ is referred to as the dielectric permeability of the medium. Substituting

(5.75) into (5.74), we obtain a system of linear homogeneous equations, which has a non-trivial solution only in the case when the determinant of the system vanishes:

$$\Delta \equiv \mathrm{Det}\left\{k^2\delta_{\alpha\beta} - k_\alpha k_\beta - \frac{\omega^2}{c^2}\varepsilon_{\alpha\beta}\right\} = 0. \tag{5.76}$$

This dispersion equation defines the eigenfrequencies of the medium, i.e., the waves that can propagate in the medium.

The case of an isotropic plasma without a magnetic field is especially simple. The vector $\mathbf{D}$ must be directed along $\mathbf{E}$ for any wave, that is $\mathbf{D} = \varepsilon\mathbf{E}$. But the proportionality coefficient ε may differ for longitudinal ($\mathbf{E}$ along $\mathbf{k}$) and transverse waves. For a longitudinal wave in accordance with (5.73) we have

$$\varepsilon_{\parallel} = 0, \tag{5.77}$$

while for a transverse wave we obtain from (5.74)

$$\frac{k^2c^2}{\omega^2} = \varepsilon_{\perp}. \tag{5.78}$$

According to (5.78) the transverse wave propagates at the phase speed $\omega/k = c/N$, where $N = \sqrt{\varepsilon_{\perp}}$ is the refraction coefficient.

5.3.2. Ion-sound, spatial dispersion

Let us consider the ion-sound in the absence of a magnetic field as the simplest example of a longitudinal wave in a plasma. The dispersion law for longitudinal waves is given by the relation $\varepsilon_{\parallel} = 0$, hence we must find $\varepsilon_{\parallel}$. Let a plane monochromatic wave propagate along the x-axis. To determine ε, we utilize expression (5.68):

$$\varepsilon E_x = E_x + \frac{4\pi i}{\omega} j_x. \tag{5.79}$$

Here the current density is equal to $j_x = en(v_{ix} - v_{ex})$, i.e., it is determined by the ion and the electron x-components of the

velocity. For a cold ion gas ($T_i = 0$) the velocity v_{ix} may be obtained from the equation of motion $m_i \partial \boldsymbol{v}_i / \partial t = e\mathbf{E}$:

$$v_{ix} = \frac{ie}{m_i \omega} E_x. \tag{5.80}$$

As for the electrons, in ion-sound they do have enough time to reach the Boltzmann distribution. Therefore the electron density perturbation is related to the potential by the equation

$$n'_e = \frac{e\varphi}{T_e} n = \frac{ien}{kT_e} E_x \tag{5.81}$$

with k being the wavenumber. This perturbation is connected with the velocity v_{ex} through the continuity condition

$$-i\omega n'_e + iknv_{ex} = 0. \tag{5.82}$$

Expressing the velocity v_{ex} in terms of E_x with the aid of (5.81) and (5.82) and substituting the result into (5.79), we obtain

$$\varepsilon = 1 + \frac{4\pi e^2 n}{k^2 T_e} - \frac{4\pi e^2 n}{m_i \omega^2} = 1 + \frac{1}{k^2 d^2} - \frac{\omega_{pi}^2}{\omega^2}, \tag{5.83}$$

where ω_{pi} is the plasma ion frequency and d is the Debye radius.

The second term of (5.83) corresponds to the electron contribution to ε while the third term gives the ion one.

Putting ε equal to zero, we obtain the already known expression for the ion-sound frequency:

$$\omega^2 = \frac{k^2 d^2}{1 + k^2 d^2} \omega_{pi}^2. \tag{5.84}$$

This electrodynamical approach has the advantage that the expressions once obtained for a contribution to ε of a particular type of particles can be used in the consideration of other vibrational modes.

It is seen from expression (5.83) that the dielectric permeability depends upon the wavenumber as well as upon the frequency. A dependence upon k arises in cases when the thermal motion of particles begins to play an active role. A dependence upon k is usually called a spatial dispersion by analogy with the frequency dispersion which is a dependence of ε upon ω.

5.3.3. Waves in a cold plasma

So far we have been considering only longitudinal waves in a plasma. Transverse waves can be treated in an analogous way. In particular, the dielectric permeability for a transverse wave in a cold plasma without a magnetic field and with the ions at rest has the same form as for a longitudinal wave, i.e., $\varepsilon = 1 - \omega_{pe}^2/\omega^2$. In accordance with (5.78) we obtain

$$\frac{c^2k^2}{\omega^2} = \varepsilon = 1 - \frac{\omega_{pe}^2}{\omega^2}. \tag{5.85}$$

Hence, the expression for the frequency may be found easily:

$$\omega^2 = \omega_{pe}^2 + c^2k^2. \tag{5.86}$$

As we see, the only waves which can propagate in a plasma are those whose frequency is greater than the plasma frequency. In the opacity domain $\omega < \omega_{pe}$ the wavenumber k is purely imaginary in accordance with (5.86), i.e., $k = i\kappa$. The value of κ determines the thickness of the skin-layer that screens the external fields. According to (5.86)

$$\kappa^2 = (\omega_{pe}^2 - \omega^2)/c^2, \tag{5.87}$$

so that for $\omega \ll \omega_{pe}$ the depth of penetration of a transverse field into the plasma is equal to c/ω_{pe}.

In the presence of an external magnetic field the pattern of electromagnetic wave propagation in a plasma becomes more complicated, especially in the case when the phase velocity of the waves is small and the thermal motion of the particles begins to play a key role. The wave propagation pattern is simpler in a cold plasma. In this case

$$\mathbf{D} = \varepsilon\mathbf{E} = \mathbf{E} + \frac{4\pi in}{\omega}\sum_{\alpha} e_\alpha \boldsymbol{v}_\alpha \tag{5.88}$$

and the charged particle velocities can be found from the equations of motion

$$-i\omega\boldsymbol{v}_\alpha = \frac{e_\alpha}{m_\alpha}\mathbf{E} + (\boldsymbol{v}_\alpha \times \boldsymbol{\Omega}_\alpha), \quad \boldsymbol{\Omega}_\alpha = \frac{e_\alpha \mathbf{B}}{m_\alpha c}. \tag{5.89}$$

Using (5.88) and (5.89), it is not difficult to find the expressions for the components of the tensor $\hat{\varepsilon}$. When referred to coordinates with the z-axis being directed along the magnetic field, the tensor $\hat{\varepsilon}$ assumes the form

$$\hat{\varepsilon} = \begin{pmatrix} \varepsilon & ig & 0 \\ -ig & \varepsilon & 0 \\ 0 & 0 & \eta \end{pmatrix}, \tag{5.90}$$

where

$$\varepsilon = 1 - \sum_{\alpha} \frac{\omega_{p\alpha}^2}{\omega^2 - \Omega_{\alpha}^2}, \quad g = -\sum_{\alpha} \frac{\omega_{p\alpha}^2 \Omega_{\alpha}}{\omega(\omega^2 - \Omega_{\alpha}^2)}, \quad \eta = 1 - \sum_{\alpha} \frac{\omega_{p\alpha}^2}{\omega^2},$$

and Ω_{α} is the cyclotron frequency of an α-type particle.

We do not intend to make a penetrating analysis of all possible wave motions in a plasma, but shall consider the simplest case of an Alfvén wave propagating along the magnetic field lines, when $k_x = k_y = 0$, $k_z = k$. Since $\omega \ll \Omega_i$, $c_A \ll c$ for Alfvén waves, the expressions for ε and g are simplified: $\varepsilon \approx \omega_{pi}^2/\Omega_i^2 = c^2/c_A^2$, $g \approx 0$. The dispersion equation (5.76) takes the form

$$\left(k^2 - \frac{\omega^2}{c^2}\varepsilon\right)^2 \left(k^2 - \frac{\omega^2}{c^2}\eta\right) = 0 \tag{5.91}$$

and the Alfvén wave frequency appears equal to

$$\omega = \frac{ck}{\sqrt{\varepsilon}} = c_A k. \tag{5.92}$$

As we see, the Alfvén wave in a dense plasma corresponds to a strongly slowed electromagnetic wave with a huge refraction coefficient $N = c/c_A$.

It might not be difficult to consider other modes of electromagnetic oscillations, in particular, the electron ones. But we shall not dwell on this topic in the book because voluminous literature is devoted to this matter.

5.4. Wave energy and momentum

Physics belongs to the family of exact sciences, but its approximate relationships or even qualitative considerations play a role of no less importance than the precise quantitative formulae. Sometimes qualitative considerations are even more necessary because they develop physical intuition and enable us to understand phenomena, which may demand much more efforts for a quantitative description.

That is the situation we encounter when considering wave processes. Treating linear problems with the help of the Fourier technique provides us with a complete solution to the problem. Decomposition into plane waves in the case of a homogeneous medium may serve as a typical example. The formal solution, however, does not enable us to obtain directly any qualitative image of the process. The wave packet language seems to be more adequate for that purpose, and we have already got acquainted with it in the preceding Sections. Now we shall make the next step and proceed from a purely qualitative description of wave propagation made in terms of spatial and temporal phase variations to the quantitative, but still approximate, description of wave phenomena.

We shall again discuss matters in terms of the wave packets slowly varying in space and time. A separate portion of such a packet can be regarded as a plane wave. The electric field of such a wave takes the following form:

$$\mathbf{E} = \mathbf{e}_{\mathbf{k}} E \exp(-i\omega_{\mathbf{k}} t + i\mathbf{k}\cdot\mathbf{r}) + (\text{c.c.}). \tag{5.93}$$

Here $\mathbf{k}$ is the wavevector, ω is the eigenfrequency, $2E$ is the amplitude of the electric field oscillations, $\mathbf{e}_{\mathbf{k}}$ is the unit vector of the wave polarization and (c.c.) denotes the complex conjugated quantity (for $\mathbf{E}$ is to be real). The eigenfrequency $\omega_{\mathbf{k}}$ is a solution to the dispersion equation (5.76), namely to $\Delta(\omega, \mathbf{k}) = 0$. The polarization vector $\mathbf{e}_{\mathbf{k}}$ can be obtained from equations (5.74) by setting $\omega = \omega_{\mathbf{k}}$. The vector $\mathbf{e}_{\mathbf{k}}$ may contain the multiplier $\exp(i\varphi_0)$, which corresponds to the constant

phase shift φ_0, so that $\mathbf{e_k}$ is complex. The phase shift φ_0 is not so important, however, for a wave packet substantially extended in space. The most important characteristics of the wave packet are $\mathbf{k}$ and $\langle E^2 \rangle$, the latter being the oscillation period average of the squared electric field which is proportional to the wave energy density at the given point.

It is natural that equations describing the temporal evolution of wave packets should be formulated in terms of the major quantities $\mathbf{k}$ and $\langle E^2 \rangle$. The first of these equations is obtained quite easily, especially in the case of a homogeneous medium. If we really have an almost sinusoidal wave with parameters slowly varying in space and time, it will be more convenient to regard the phase φ as a function of $(\mathbf{r}, t)$, but not of $(\omega_{\mathbf{k}} t - \mathbf{kr})$. So expression (5.93) for a wave packet will take the form

$$\mathbf{E} = \mathbf{e_k} E \exp(-i\varphi) + (\text{c.c.}). \tag{5.94}$$

The rate of the phase variation at a given point, i.e., $\partial\varphi/\partial t$, obviously corresponds to the frequency of oscillations, while the phase gradient corresponds to the wavenumber since it gives the magnitude and the direction of the fastest variation of the phase in space. Therefore,

$$\omega = \frac{\partial \varphi}{\partial t}, \quad \mathbf{k} = -\nabla\omega. \tag{5.95}$$

Evaluating the time derivative of the second relation and using the first one, we obtain the identity

$$\frac{\partial \mathbf{k}}{\partial t} = -\nabla\omega. \tag{5.96}$$

In a homogeneous medium the frequency of oscillations depends only upon $\mathbf{k}$: $\omega = \omega_{\mathbf{k}} = \omega(\mathbf{k})$. Therefore equation (5.96) is a differential equation with respect to $\mathbf{k}(\mathbf{r}, t)$. Thus, we can find the wavevector temporal evolution at an arbitrary point of space. For instance, equation (5.96) takes a simple form

$$\frac{\partial k}{\partial t} + v_g \frac{\partial k}{\partial x} = 0 \tag{5.97}$$

in the one-dimensional case when the wavenumber depends only upon x. Here we used the notation $v_g = \partial\omega/\partial k$ for the group velocity. As we see, the packet wavenumber is merely transported with the group velocity.

Let us now proceed to the derivation of the equation for $\langle E^2 \rangle$. Since this quantity is proportional to the wave energy density, it seems more reasonable to utilize this variable as the one that has a definite physical sense.

5.4.1. Wave energy

Let the Maxwell equation (5.63) be supplied with some external current $\mathbf{j}_s$ able to interact with the wave. This current is added on to the self-consistent current $\mathbf{j}$ that is expressed in terms of $\mathbf{D}$ via (5.68). Using the routine procedure of multiplying (5.68) by $\mathbf{E}$, (5.64) by $\mathbf{B}$ and adding the results, it is easy to obtain

$$\frac{1}{4\pi}\left(\mathbf{E}\cdot\frac{\partial \mathbf{D}}{\partial t}+\mathbf{B}\cdot\frac{\partial \mathbf{B}}{\partial t}\right)+\frac{c}{4\pi}\nabla\cdot(\mathbf{E}\times\mathbf{B})=-\mathbf{E}\cdot\mathbf{j}_s. \tag{5.98}$$

The right-hand side of this equation corresponds to the external source power that works against the wave field $\mathbf{E}$. Hence, the first term in (5.98) may be interpreted as the rate of electromagnetic field energy variation:

$$\frac{\partial \mathcal{E}}{\partial t}=\frac{1}{4\pi}\mathbf{E}\cdot\frac{\partial \mathbf{D}}{\partial t}+\frac{1}{4\pi}\mathbf{B}\cdot\frac{\partial \mathbf{B}}{\partial t}. \tag{5.99}$$

The second term in (5.98) may be treated as the energy transport along the Pointing vector

$$\mathbf{\Pi}=\frac{c}{4\pi}(\mathbf{E}\times\mathbf{B}). \tag{5.100}$$

For simplicity we shall restrict ourselves to the case of a transparent medium where the energy is not absorbed.

In the absence of external sources expression (5.98) takes the form of the energy conservation law

$$\frac{\partial \mathcal{E}}{\partial t}+\nabla\cdot\mathbf{\Pi}=0. \tag{5.101}$$

It follows from this law that the integral of the energy density $\mathcal{E}$ over the whole space does not change with time:

$$\int \mathcal{E} d\mathbf{r} = \text{const.} \tag{5.102}$$

Equation (5.101) corresponds to the equation needed for the squared wave amplitude. It should only be reduced to a simpler form that is more convenient for the description of wave packet propagation. The electric and magnetic fields in a wave packet are of the form of a quasi-monochromatic wave (5.93) with slowly varying amplitudes $E(\mathbf{r}, t)$, $B(\mathbf{r}, t)$. Accordingly,

$$\frac{\partial \mathbf{E}}{\partial t} = \left(\frac{\partial E}{\partial t} - i\omega_{\mathbf{k}} E \right) \mathbf{e}_{\mathbf{k}} \exp(-i\omega_{\mathbf{k}} t + i\mathbf{k} \cdot \mathbf{r}) + (\text{c.c.}). \tag{5.103}$$

On the other hand, the electric field $\mathbf{E}$ can be decomposed into the Fourier integral

$$\mathbf{E} = \int \mathbf{e}_{\boldsymbol{q}} E_{\boldsymbol{q}\omega} \exp(-i\omega t + i\mathbf{q} \cdot \mathbf{r}) \, d\mathbf{q} \, d\omega + (\text{c.c.}). \tag{5.104}$$

In this expression the "packet-like" form of the wave, i.e., its proximity to a plane monochromatic wave, is reflected only by the fact that the Fourier components $E_{\boldsymbol{q}\omega}$ substantially differ from zero only in a small neighborhood of the point $(\mathbf{q}, \omega) = (\mathbf{k}, \omega_{\mathbf{k}})$. The closer the packet to a monochromatic plane wave, the smaller the dimensions of this band are. It is clear from (5.104) that in the Fourier representation the differentiation with respect to time looks like multiplication by $-i\omega$. In the zeroth order, i.e., for a monochromatic wave, this multiplier will be simply equal to $-i\omega_{\mathbf{k}}$. But this is not true for a packet, i.e., $-i\omega + i\omega_{\mathbf{k}} \neq 0$. By virtue of (5.103) we see that the factor $-i(\omega - \omega_{\mathbf{k}})$ corresponds to the time derivative of the wave amplitude (because $\partial/\partial t$ on the left-hand side corresponds to $-i\omega$).

More complex operators can be treated in an analogous way. For example, the derivative $\partial \mathbf{D}/\partial t$ from (5.98) corresponds to the expression $-i\omega\hat{\varepsilon}\mathbf{E}$, where $\hat{\varepsilon}$ is the square matrix

$\varepsilon_{\alpha\beta}(k,\omega)$. For a wave packet this expression will obviously be equal to

$$-i\omega_{\mathbf{k}}\hat{\varepsilon}(\mathbf{k},\omega_{\mathbf{k}}) + \frac{\partial}{\partial\omega}(\hat{\varepsilon}\omega)\frac{\partial}{\partial t}. \tag{5.105}$$

If this is substituted into (5.99), we realize that the term containing $-i\omega_{\mathbf{k}}\hat{\varepsilon}(\mathbf{k},\omega_{\mathbf{k}})$ vanishes after being added to its complex conjugate. This is valid for the case of a transparent medium when the matrix $\hat{\varepsilon}$ is Hermitian. Under these conditions the only terms that survive are those that involve temporal derivatives, and after the transform we obtain the expression

$$\mathcal{E}_{\mathbf{k}} = \frac{\langle E^2\rangle}{8\pi}\frac{\partial}{\partial\omega}(\varepsilon\omega) + \frac{\langle B^2\rangle}{8\pi} \tag{5.106}$$

for the wave packet energy density with $\varepsilon = \mathbf{e}_{\mathbf{k}}^{*}\hat{\varepsilon}\,\mathbf{e}_{\mathbf{k}}$. Since $\mathbf{B}$ can be expressed in terms of $\mathbf{E}$ via expression (5.72), we have

$$\mathcal{E}_{\mathbf{k}} = \frac{\langle E^2\rangle}{8\pi}\frac{\partial}{\partial\omega}\left(\varepsilon\omega - \frac{c^2k^2}{\omega} + \frac{c^2}{\omega}|\mathbf{k}\cdot\mathbf{e}_{\mathbf{k}}|^2\right). \tag{5.107}$$

The derivative of the expression within round brackets is to be taken at the point $\omega = \omega_{\mathbf{k}}$.

So we have related the squared wave electric field average to a more physically sensible quantity $\mathcal{E}_{\mathbf{k}}$, which is the wave packet energy density. Hence, we may choose a more convenient quantity to normalize the wave amplitude than the squared electric field. As we saw before, it is more convenient to take $a_{\mathbf{k}}$ as the wave amplitude which is connected to the packet energy density by the relation

$$\mathcal{E}_{\mathbf{k}} = \omega_{\mathbf{k}}a_{\mathbf{k}}^2 \equiv \omega_{\mathbf{k}}N_{\mathbf{k}}, \tag{5.108}$$

with the frequency $\omega_{\mathbf{k}}$ being positive for a positively defined wave energy. By analogy with quantum mechanics it is convenient to interpret the squared amplitude $a_{\mathbf{k}}^2$ as the number of waves $N_{\mathbf{k}}$ with fixed $\mathbf{k}$.

Let us return to equation (5.101). It is possible to proceed with the spatial derivatives in the same manner as we did with

temporal derivatives. So the derivative $\partial \mathbf{E}/\partial x$, for example, is equal to

$$\frac{\partial}{\partial x}\mathbf{E} = \left(\frac{\partial E}{\partial x} + ik_x E\right)\mathbf{e}_\mathbf{k} \exp(-i\omega_\mathbf{k} t + i\mathbf{k}\cdot\mathbf{r}) + (\text{c.c.}).$$

in accordance with (5.104). In the Fourier representation this derivative corresponds to multiplication by iq_x. Consequently, we see again that the factor $i\mathbf{q} - i\mathbf{k}$ corresponds to the gradient of the wave amplitude. In particular, the operator

$$-i\omega\hat{\varepsilon} = -i(\omega\hat{\varepsilon})_\mathbf{k} - \frac{\partial}{\partial \mathbf{k}}(\omega\hat{\varepsilon})\frac{\partial}{\partial \mathbf{r}}$$

when acting on the field of the wave packet, gives a contribution proportional to the gradient of the oscillation amplitude. Correspondingly, in the presence of a spatial dispersion equation (5.101) takes the form

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla\cdot\mathbf{S}_\mathbf{k} = 0. \tag{5.109}$$

The energy flux $\mathbf{S}_\mathbf{k}$ has a contribution coming not only from the Pointing vector, but from the spatial part of the first term of (5.101) as well. Following the same algebra we used while computing $\mathcal{E}_\mathbf{k}$, it is not difficult to find the expression for $\mathbf{S}_\mathbf{k}$:

$$\mathbf{S}_\mathbf{k} = \frac{\langle E^2\rangle}{8\pi}\frac{\partial}{\partial \mathbf{k}}\left(\varepsilon\omega - \frac{c^2k^2}{\omega} + \frac{c^2}{\omega}\,|\,\mathbf{k}\cdot\mathbf{e}_\mathbf{k}|^2\right). \tag{5.110}$$

As we see, $\mathcal{E}_\mathbf{k}$ and $\mathbf{S}_\mathbf{k}$ are expressed via the derivatives of the same function

$$A(\mathbf{k},\omega) = \omega\,\mathbf{e}_\mathbf{k}^*\hat{\varepsilon}\mathbf{e}_\mathbf{k} - \frac{c^2k^2}{\omega} + \frac{c^2}{\omega}\,|\,\mathbf{k}\cdot\mathbf{e}_\mathbf{k}|^2. \tag{5.111}$$

This function becomes zero at $\omega = \omega_\mathbf{k}$, since it is obtained through multiplying equation (5.74) by $\mathbf{e}_\mathbf{k}^*/E$. In other words, $A = 0$ along the whole dispersion curve $\omega = \omega_\mathbf{k}$, i.e.,

$$dA = \frac{\partial A}{\partial \omega}\,d\omega + \frac{\partial A}{\partial \mathbf{k}}\,d\mathbf{k} = 0 \tag{5.112}$$

everywhere on that curve. But the ratio $d\omega/d\mathbf{k}$ is equal to the group velocity, hence $\mathbf{S}_\mathbf{k} = \mathcal{E}_\mathbf{k} \boldsymbol{v}_g$. Thus, the wave packet energy transport equation (5.109) may be written in the simple form

$$\frac{d\mathcal{E}}{dt} + \nabla \cdot (\boldsymbol{v}_g \mathcal{E}_\mathbf{k}) = 0. \tag{5.113}$$

This is the equation needed for the squared wave amplitude. It may be easily generalized for the case of a weakly damped wave, i.e., for the case when the eigenfrequency acquires a small imaginary part, so that $\omega = \omega_\mathbf{k} + i\gamma_\mathbf{k}$ ($\gamma_\mathbf{k} < 0$ for a damped wave). This damping is connected with the anti-Hermitian constituent of the tensor $\hat{\varepsilon}$ and can be found directly either from the dispersion relation $A = 0$ with A given by (5.111), or from the energy transport equation involving the anti-Hermitian part of $\hat{\varepsilon}$. In the case $\gamma \ll \omega$ of weak damping, i.e., when the imaginary part of $\varepsilon = \mathbf{e}_\mathbf{k}^* \hat{\varepsilon}_\mathbf{k} \mathbf{e}_\mathbf{k}$ is small, we obtain by virtue of (5.111) and $A(\omega_\mathbf{k}, \mathbf{k}) = 0$ the relation

$$\gamma = -\frac{\operatorname{Im} A}{\frac{\partial}{\partial\omega} \operatorname{Re} A} = -\frac{\omega^2 \varepsilon''}{\frac{\partial}{\partial\omega}(\omega^2 \varepsilon')} \tag{5.114}$$

with $\varepsilon' = \operatorname{Re}(\mathbf{e}_\mathbf{k}^* \hat{\varepsilon} \mathbf{e}_\mathbf{k})$, $\varepsilon'' = \operatorname{Im}(\mathbf{e}_\mathbf{k}^* \hat{\varepsilon} \mathbf{e}_\mathbf{k})$. Thus, in the case of a slightly absorbing medium equation (5.113) takes the form

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot (\boldsymbol{v}_g \mathcal{E}_\mathbf{k}) = 2\gamma_\mathbf{k} \mathcal{E}_\mathbf{k}. \tag{5.115}$$

We have got 2γ precisely here, because the square of the amplitude is damped twice faster than the amplitude itself.

5.4.2. Wave momentum

In addition to the energy a wave packet has a certain momentum. The latter could be found from the general relations of electrodynamics of a continuous medium, but here we shall use a simpler but no less general approach based upon the Galilean invariance of physical processes at velocities much less than that of light.

Let a wave packet propagate in a medium which moves at a constant speed $\boldsymbol{V}$. Let the packet energy with respect to that medium (i.e., in the moving frame of reference) be $\mathcal{E}_{\mathbf{k}}$. In general, in the motionless reference frame the packet energy $\mathcal{E}^0_{\mathbf{k}}$ will differ from $\mathcal{E}_{\mathbf{k}}$. The relation between both values can be found from the principle of Galilean invariance. To do this, consider a particle of a mass m moving at the velocity $\boldsymbol{v}_0 = \boldsymbol{v} + \boldsymbol{V}$, where $\boldsymbol{v}$ is the relative velocity. The energy of this particle is equal to $\mathcal{E}_p = mv_0^2/2 = mv^2/2 + m\boldsymbol{v}\cdot\boldsymbol{V} + mV^2/2$. Thus, up to a constant $mV^2/2$, the particle energy is equal to $\mathcal{E}^0_p = \mathcal{E}_p + \mathbf{p}\cdot\boldsymbol{V}$, where $\mathcal{E}_p$ stands for the energy in the reference frame moving with the medium, and $\mathbf{p} = m\boldsymbol{v}$ for the momentum of the particle.

Let us suppose that the particle is capable of energy and momentum exchange with the wave. For the energy and momentum conservation laws to be satisfied simultaneously in the moving and stationary reference frames, the relation between the energy and momentum of a packet in a moving medium should coincide with that of a particle, i.e.,

$$\mathcal{E}_{\mathbf{k}} = \mathcal{E}^0_{\mathbf{k}} + \mathbf{P}_{\mathbf{k}}\cdot\boldsymbol{V}. \tag{5.116}$$

Now let us take into account that a free particle is able to interact intensively with a wave only in the case when the Cherenkov resonant condition $\omega_{\mathbf{k}} - \mathbf{k}\cdot\boldsymbol{v} = 0$ is satisfied. Moreover, the particle energy variation in respect to the considered medium is connected with its momentum variation by an obvious relation: $\delta\mathcal{E}_p = m\boldsymbol{v}\cdot\delta\boldsymbol{v} = \boldsymbol{v}\cdot\delta\mathbf{p}$. The wave packet energy and momentum variations should also be coupled by the same relation. But they are simply proportional to the squared amplitude and, therefore, are proportional to each other, i.e., $\mathcal{E}_{\mathbf{k}} - \boldsymbol{v}\cdot\mathbf{P}_{\mathbf{k}} = 0$ at $\omega_{\mathbf{k}} - \mathbf{k}\cdot\boldsymbol{v} = 0$. Since the particle velocity component normal to $\mathbf{k}$ is arbitrary, the wave momentum $\mathbf{P}_{\mathbf{k}}$ is directed along $\mathbf{k}$ and is equal to

$$\mathbf{P}_{\mathbf{k}} = \frac{\mathbf{k}}{\omega_{\mathbf{k}}}\mathcal{E}_{\mathbf{k}} \tag{5.117}$$

because $\boldsymbol{v}\cdot\mathbf{k} = \omega_{\mathbf{k}}$. Thus, we have found the expression for $\mathbf{P}_{\mathbf{k}}$. Now we see the reason for introducing the quantity $N_{\mathbf{k}}$ of the

number of waves in the packet. Using it, we are able to give quite simple expressions for the packet energy and momentum:

$$\mathcal{E}_{\mathbf{k}} = \omega_{\mathbf{k}} N_{\mathbf{k}}, \quad \mathbf{P}_{\mathbf{k}} = \mathbf{k} N_{\mathbf{k}}. \tag{5.118}$$

Accordingly, the transformation of the wave energy due to the change of reference frame is also simplified. Comparing (5.118) with (5.116), we see that the wave energy is transformed in the same manner as the frequency due to the Doppler shift, i.e.,

$$\mathcal{E}^0_{\mathbf{k}} = \omega^0_{\mathbf{k}} N_{\mathbf{k}}, \tag{5.119}$$

where $\omega^0_{\mathbf{k}} = \omega_{\mathbf{k}} + \mathbf{k} \cdot \boldsymbol{V}$. The number $N_{\mathbf{k}}$ of waves and the momentum $\mathbf{P}_{\mathbf{k}}$ are naturally conserved under reference frame transformations.

It is worth noting once more that the ratio of the wave energy to its momentum is equal to the phase speed: $\mathcal{E}_{\mathbf{k}} = \boldsymbol{v}_p \cdot \mathbf{P}_{\mathbf{k}}$. Thus, the lower the phase speed, the "heavier" the wave packet is, i.e., the greater its momentum at a given energy level. Hence, the "fastest" waves appear to be the "lightest" ones.

5.4.3. Waves of negative energy

At first glance it seems to be rather natural to assume that a wave packet energy is always positive since usually we need to spend some energy for wave excitation. For equilibrium media this property can be strictly proved [79]. For a medium far from the equilibrium state, however, this is not true. Waves can often have a negative sign of energy in a plasma that is sometimes far from equilibrium. This negative energy sign of a wave packet means that one should not spend, but extract some amount of energy from the medium in order to excite a wave.

The simplest example of a wave of a negative energy arises when considering waves in a moving medium. Let a plane wave of frequency $\omega = \omega_{\mathbf{k}}$ propagate in the direction of x, so that $k = k_x$. The energy of such a wave is equal to $\mathcal{E} = \omega_{\mathbf{k}} N$. Assume now that the whole medium starts moving in the opposite

direction, i.e., to the left, with the velocity $V_x = -v_0$. Then a Doppler frequency shift $\omega \to \omega_k^0 = \omega_k - kv_0$ will take place and, accordingly, the energy will diminish: $\mathcal{E}_\mathbf{k} \to \mathcal{E}^0 = (\omega_k - kv_0)N$. Now let the medium velocity v_0 increase. When the velocity v_0 reaches the wave phase speed, the wave will stop in the motionless frame of reference and its frequency and energy will become zero. If the velocity v_0 exceeds the phase speed, the wave energy will change its sign and become negative. The wave is swept by such a "supersonic" medium to the left while it still propagates to the right relative to the medium itself. The energy of such a "lagging" wave is negative, i.e., one should not spend, but extract energy from the medium in order to excite it.

Note, in particular, that the well-known relation of Landau for the speed limit of superfluid helium capillary motion, $\mathcal{E}_k - p_k v_0 > 0$, where $\mathcal{E}_k$ is the energy and p_k is the momentum of an elementary helium excitation, is equivalent to the statement of the absence of negative energy excitations in the laboratory reference frame, i.e., relative to the walls of the capillary that carries helium.

Taking the wave energy sign into account, one can take a different view on the phenomenon of Cherenkov emission or body streamlining by flows. Denoting the particle velocity (or the reversed velocity of the streamlining flow) by $\boldsymbol{v}_0$, it is easy to see that the Cherenkov condition $\omega_\mathbf{k} - \mathbf{k} \cdot \boldsymbol{v}_0 = 0$ means that the frequency of the corresponding wave, and, therefore, its energy, vanishes in the reference frame connected with the particle. One may say that a body transmits only zero-energy bow waves in the streamlining flow. This seems rather natural, because a body at rest does not perform any work.

If there are internal degrees of freedom and a body can vibrate, i.e., if it is an oscillator with frequency ω_0, then energy exchange between the wave and the oscillator can take place. The latter occurs under the resonance condition:

$$\omega - \mathbf{k} \cdot \boldsymbol{v}_0 = \omega_0 \quad \text{or} \quad \omega - \mathbf{k} \cdot \boldsymbol{v}_0 = -\omega_0. \tag{5.120}$$

In the first case the resonance condition is satisfied when the

wave frequency $\omega - \mathbf{k} \cdot \boldsymbol{v}$ and, consequently, the energy, are positive. This case is referred to as an ordinary Doppler effect and an ordinary resonance. In the second case the energy of a resonant wave appears to be negative and it is usually referred to as anomalous. The transmitting oscillator does not lose, but gains energy in the case of an anomalous Doppler effect since the wave has a negative energy in the reference frame of the oscillator. The particles orbiting the magnetic lines in a plasma can serve as an example of such oscillators. In this case ω_0 can be equal to an arbitrary integer multiple of the cyclotron frequency, and, accordingly, the resonance condition takes the form

$$\omega - k_z v_z = n\Omega_\alpha, \tag{5.121}$$

where n is an integer that is positive in the case of the normal and is negative in the case of the anomalous Doppler effect, while v_z is the particle velocity along the field line (or the streamlining medium velocity taken with the opposite sign).

We have considered only the simplest case of motion of the medium, when the wave possessed a positive energy relative to the medium. When considering an unstable plasma, one can encounter a more complicated situation when the wave energy defined by the general expression (5.107) appears negative even for a medium at rest. Of course this can only be the case when the plasma is far from equilibrium. For example, waves of negative energy can exist in a plasma confined within a mirror trap with magnetic mirrors when the particle distribution function corresponds to the inverse population caused by the leakage of plasma components with low transverse velocities. Negative energy waves can also exist in a plasma with beams of charged particles.

5.4.4. The parabolic equation

Equations (5.96) and (5.115) obtained for $\mathbf{k}$ and $\mathcal{E}$ correspond to the approximation of geometrical optics, when the wavelength is assumed to be infinitely small in comparison with

characteristic dimensions of the wave packet. But it is possible to take into account the small terms of the order of the ratio of the wavelength to the wave packet dimensions. With this generalization we can obtain an approximate description of wave diffraction. In the case of a nearly monochromatic wave this procedure leads to the so-called parabolic equation, which was put forward by Leontovich [25].

Let us consider again a quasi-monochromatic wave and assume for simplicity that the medium is isotropic, so that the frequency depends only upon the absolute value of the wavenumber. We represent the electric field in the form

$$\mathbf{E} = \mathbf{e}_{\mathbf{k}_0} E \exp(i\omega_0 t + ik_0 x) + (\text{c.c.}), \tag{5.122}$$

which differs from (5.93) by the weak dependence of both the amplitude and phase upon the spatial and temporal coordinates. Here k_0 is the wavenumber, $\omega_0 = \omega(k_0)$ is the corresponding eigenfrequency and $\mathbf{e}_{\mathbf{k}_0}$ is the polarization vector. To simplify things, we shall assume the polarization vector to be constant. This is surely valid for a smooth packet with k_0 essentially exceeding the inverted packet length. We shall also assume that the x axis of the rectangular coordinate system is directed along k_0.

The complex amplitude E, which alternates slowly in space and time, can be represented with the Fourier integral:

$$E = \int E_{\boldsymbol{\kappa}\nu} \exp(-i\nu t + i\boldsymbol{\kappa} \cdot \mathbf{r})\, d\boldsymbol{\kappa}\, d\nu. \tag{5.123}$$

With the notation L for the typical wave packet dimensions and $T \sim L/v_g$ for the corresponding time scale we may state that the extension of the integration domain over κ is $\sim 1/L$ and over ν is $\sim 1/T$.

If (5.123) is substituted into (5.122), the field $\mathbf{E}$ will be represented in the form of a Fourier integral with the component $\mathbf{E}_{\mathbf{k}\omega}$ being equal to $\mathbf{e}_{\mathbf{k}_0} E_{\boldsymbol{\kappa}\nu}$ at $\mathbf{k} = \mathbf{k}_0 + \boldsymbol{\kappa}$, $\omega = \omega_0 + \nu$. But the only components that can be involved in the Fourier expansion are those whose frequency ω coincides with the eigenfrequency,

i.e., $(\omega - \omega_{\mathbf{k}})E_{\mathbf{k}\omega} = 0$. Therefore, $(\omega_0 + \nu - \omega_{\mathbf{k}_0+\boldsymbol{\kappa}})E_{\boldsymbol{\kappa}\nu} = 0$ and hence the relation

$$\int \left(\nu - \omega_{\mathbf{k}_0+\boldsymbol{\kappa}} + \omega_{\mathbf{k}_0}\right) E_{\boldsymbol{\kappa}\nu} \exp(-i\nu t + i\boldsymbol{\kappa}\cdot\mathbf{r})\, d\nu\, d\boldsymbol{\kappa} = 0 \quad (5.124)$$

holds. Since $\boldsymbol{\kappa}$ is small, we shall expand $\omega_{\mathbf{k}_0+\boldsymbol{\kappa}} - \omega_{\mathbf{k}_0}$ in a series in $\boldsymbol{\kappa}$ up to terms of the second power. In accordance with our assumption $\omega_{\mathbf{k}}$ depends only upon the magnitude of $\mathbf{k}$. Hence,

$$\omega_{\mathbf{k}_0+\boldsymbol{\kappa}} \approx \omega_{\mathbf{k}_0} + \left(|\mathbf{k}_0 + \boldsymbol{\kappa}| - k_0\right)v_g + \frac{1}{2}\left(|\mathbf{k}_0 + \boldsymbol{\kappa}| - k_0\right)^2 v_g', \quad (5.125)$$

where

$$v_g = \frac{\partial\omega}{\partial k}, \quad v_g' = \frac{\partial^2\omega}{\partial k^2},$$

$$|\mathbf{k}_0 + \boldsymbol{\kappa}| = \sqrt{(k_0 + \kappa_x)^2 + \kappa_\perp^2} \approx k_0 + \kappa_x + \kappa_\perp^2/2k_0.$$

If we substitute these expressions into (5.124) and take into account that $\nu = i\partial/\partial t$, $\boldsymbol{\kappa} = -i\partial/\partial\mathbf{r}$, we shall obtain the equation for the complex amplitude:

$$i\frac{\partial E}{\partial t} + iv_g\frac{\partial E}{\partial x} + \frac{v_g}{2k_0}\Delta_\perp E + \frac{v'}{2}\frac{\partial^2 E}{\partial x^2} = 0. \quad (5.126)$$

This is the parabolic equation. Omitting the last two small terms, we obtain the transport equation $\partial E/\partial t + v_g\partial E/\partial x = 0$ in the approximation of geometrical optics. Keeping these terms provides us with a more accurate description of the wave packet. In particular, it gives an approximate account of the diffraction effects.

5.5. Waves in inhomogeneous media

5.5.1. WKB-approximation

The problem of a plasma oscillation description becomes much more complicated in the case of an inhomogeneous medium. From the purely mathematical viewpoint this complication

is related to the fact that one has to treat differential (or integral) equations with nonconstant coefficients, which faces severe computational difficulties. But the real aspect of the situation is not the tedious specification of the corrections of the description of homogeneous plasma oscillations, but the emergence of new, and sometimes of rather unexpected, plasma modes. These new modes frequently appear to be unstable and therefore the theory of inhomogeneous plasma oscillations includes the whole theory of plasma instability. But this subject is very extensive and is not the topic of this Section, so we shall only make brief comments.

The oscillatory modes of a homogeneous plasma are surely among the modes of an inhomogeneous one. Therefore the corresponding wave propagation can be described by means of the wave packet formalism, provided the wavelength is small compared with the inhomogeneity scale. The corresponding approximation is known as the WKB-approximation or the approximation of geometrical optics and proved universal and applicable to any waves, especially waves in a medium with a one-dimensional inhomogeneity.

Let, for the first instance, the plasma characteristics depend only on the z coordinate and the y-component k_y of the wavevector $\mathbf{k}$ be zero. Then the frequency of an eigenmode is to satisfy the local dispersion equation determined on the grounds of the theory of a homogeneous medium and may be regarded as a function of the wavevector with the coordinates k_x, k_z: $\omega = \omega(k_x, k_z)$. Consider the field of a point source vibrating at the frequency ω_0. If the wave field has been settled, the wavenumber will be stationary at any point, therefore $\nabla\omega = 0$ according to (5.96), i.e., $\omega(k_x, k_z) = \omega_0 = \text{const}$. For a wide packet the components k_x, k_z should be constant at any fixed z and should not depend on x. But the vector field $\mathbf{k}$ is curlless, i.e., $\mathbf{k} = \nabla\varphi$, and hence $\partial k_x/\partial z = \partial^2\varphi/\partial z\partial x = \partial k_z/\partial x = 0$. Thus, the independence of k_z upon x implies that k_x is constant: $k_x = \text{const}$. Therefore, in summary we may say that the relation $\omega(k_x, k_z, z) = \omega_0$ implies the functional dependence of

k_z on z: the wavenumber k_z always takes such a value that the local frequency should be equal to the external fixed frequency.

Now let us turn to the second equation of wave packets dynamics — the energy transfer equation (5.113). In the steady one-dimensional case it assumes the form $\partial(v_{gz}\mathcal{E})/\partial z = 0$. Since $\mathcal{E}$ is proportional to the squared oscillation amplitude E^2, we may conclude that $E^2 \sim \text{const}/v_{gz}$. The packet "condenses" at the spots of zero group velocity and its amplitude tends to infinity there. These spots correspond to the turning points, i.e., to the points of reflection from the opaque area.

Let us now consider, by way of example, a transverse electro-magnetic wave penetrating into a plasma from the outside. In this case we have $\omega_0^2 = \omega^2 = c^2k^2 + \omega_{pe}^2$. Let the wave be travelling upwards in the positive direction of z and assume that the plasma density increases with z. Then the dispersion equation implies

$$k_z = \frac{1}{c}\sqrt{\omega^2 - \omega_{pe}^2 - k_x^2c^2}. \tag{5.127}$$

We see that the component k_z of the wavevector decreases when the plasma frequency $\omega_{pe} = \sqrt{4\pi e^2 n/m}$ increases and tends to zero at the point $\omega^2 = \omega_0 - c^2k_x^2$. But $\omega^2 = c^2(k_x^2 + k_{z0}^2)$, with k_{z0} being the value of the wavenumber in the vacuum, and hence this point corresponds to the frequency

$$\omega_{pe}^2 = \omega^2\frac{k_{z0}^2}{k_x^2 + k_{z0}^2} = \omega_0^2\sin^2\theta, \tag{5.128}$$

where θ is the angle of incidence of the wave on the plasma. The point $k_z = 0$ (v_{gz} is also equal to zero there) corresponds to the vanishing of the electro-magnetic wave: it cannot penetrate into the deeper layers because the plasma appears to be opaque. If the angle θ is near $\pi/2$, this spot will be somewhere in the vicinity of the resonant point $\omega_0 = \omega_{pe}$, where the external frequency coincides with the Langmuir frequency.

This results in the flexibility of transformation of a transverse wave into a longitudinal one. Assume, in the first instance,

that the electro-magnetic wave is polarized so that its electric part lies in the (x, z) plane. Then this electric vector will be almost parallel to the z axis near the vanishing point, for $k_x \gg k_z$ there and, consequently, the wave travels in the direction of x. Behind the turning point k_z appears to be purely imaginary and the electric field decays exponentially fast as z increases. But nevertheless a weak portion reaches the resonant point $\omega = \omega_{pe}$. The external field amplifies the Langmuir waves at this spot and the transverse wave energy is transferred into a longitudinal wave. This effect is called wave transformation. Other vibration modes can experience similar transformations.

The WKB-approximation can be successfully applied to problems of propagation of waves of different types, in particular, of hydromagnetic waves. Taking the appropriate boundary conditions into account, one can obtain the eigenfrequencies of a plasma in the case when the standing waves appear to exist due to reflection by the boundaries or by the turning points. However, this method is not completely universal; in fact it fails to reveal one of the most interesting peculiarities of plasma eigenfrequency spectra, in particular of hydromagnetic modes — the existence of continuous spectra.

5.5.2. The continuous spectra

The WKB-approximation essentially utilizes the picture of wave packet propagation and is therefore valid only under the condition that these wave packets really do run through a plasma. But in the case when the group velocity is infinitesimal or exactly equal to zero, the oscillation pattern can be somewhat different. To reveal the corresponding picture, we shall proceed with Langmuir oscillations in a cold plasma. The frequency $\omega = \omega_{pe}$ of these oscillations does not depend on k and therefore their group velocity is equal to zero. Let us consider what these oscillations will look like in an inhomogeneous plasma.

We can describe the oscillations with the help of the equation

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} + \omega_{pe}^2 \mathbf{E} = -c^2 \nabla \times \left(\nabla \times \mathbf{E} \right). \tag{5.129}$$

It can easily be derived by differentiation of (5.63) in time and the successive elimination of $\mathbf{B}$ and $\mathbf{j}$ with the aid of (5.64) and the linearized equation

$$\frac{\partial \mathbf{j}}{\partial t} = -en \frac{\partial \boldsymbol{v}_e}{\partial t} = \frac{e^2 n}{m_e} \mathbf{E}$$

of the electron motion. Equation (5.129) describes both transverse and longitudinal waves in a plasma. For longitudinal waves $\mathbf{E} = -\nabla\varphi$ and the $\nabla\times$-operation transforms (5.129) into $\nabla\varphi \times \nabla n = 0$. Thus, the levels of constant density should be equipotential.

Let us start off with these oscillations. In view of $\nabla\varphi = \partial\varphi/\partial n \, \nabla n$, equation (5.129) assumes the form

$$\frac{\partial^2}{\partial t^2} \frac{\partial \varphi}{\partial n} + \omega_{pe}^2 \frac{\partial \varphi}{\partial n} = 0. \tag{5.130}$$

It is evident that at any spot the potential oscillates with the local plasma frequency. For every available value $\omega = \omega_{pe}^{(0)}$ of the frequency there is a discontinuous solution to equation (5.130), corresponding to $\partial\varphi/\partial n = \delta(n - n^{(0)})$. It is given by a piecewise constant function with a single jump on the surface corresponding to the chosen value of the eigenfrequency. This mode describes oscillations of the electrons inside of the thin layer containing the surface $n = \text{const}$, with other electrons being at rest. The specified layer behaves as a harmonically oscillating electric double-layer distributed over the surface of the potential jump. Thus, the problem under consideration is related to the continuous oscillation spectra, and the eigenfrequencies can take any value between the minimum $\omega_{pe\,\min}$ and maximum $\omega_{pe\,\max}$ values of the local plasma frequency, determined respectively by the minimum and maximum values of the electron density.

Suppose that there is an initial disturbance of the electron density in a plasma, such that all the electrons were simultaneously shifted in the same direction. At first we shall see the

coherent electron oscillations near the equilibrium position. But in time the phases of oscillations in different layers will diverge from one another. In other words, the wavevector $\mathbf{k}$ will grow linearly with time due to the equation $\partial\mathbf{k}/\partial t = \nabla\omega$. Namely we have $\mathbf{k} = t\nabla\omega_{pe}$. It is evident that the higher the density gradient, the higher the growth rate. We can say that a "wind" blows in the k-space from small to high wavenumbers. The amplitude of the electro-magnetic oscillations will decrease monotonically since the oscillations will be synchronized only in narrow contracting areas.

The above picture proved typical for any system with a continuous spectrum, not only for an inhomogeneous plasma in respect to the Langmuir oscillations. As for the decay of oscillations due to the phase disbalance, it appears to be analogous to the Landau damping, which takes place in the space of velocities with Van Kampen waves of the continuous spectrum.

Let us now pass to Langmuir oscillations of a more general type. Suppose the length of these oscillations is much less than c/ω_{pe}. Then the right-hand side of equation (5.129) may be considered large in comparison with the left-hand one, and therefore the equation can be treated by means of successive iterations. As a first approximation we have $\mathbf{E} = -\nabla\varphi$. To obtain the next approximation one should impose the restriction that the small left-hand side be orthogonal to the solution of the homogeneous equation of the first approximation. In our case this restriction is just the condition that the expression on the left-hand side is divergence-free, i.e.,

$$\frac{\partial^2}{\partial t^2}\Delta\varphi + \nabla\cdot(\omega_{pe}^2\nabla\varphi) = 0. \tag{5.131}$$

On assuming the density n to be dependent only on the coordinate z, we can look for a solution of the form $\varphi(z)\exp(-i\omega t + ikx)$. In the particular case $n = n_0(1 - z/L)$ of linear increase of the density with z, equation (5.131) takes the form

$$\frac{d}{dz}\left(\nu^2 + \frac{z}{L}\right)\frac{d\varphi}{dz} - k^2\left(\nu^2 + \frac{z}{L}\right)\varphi = 0 \tag{5.132}$$

with $\nu^2 = (\omega^2 - \omega_{pe}^{(0)2})/\omega_{pe}^{(0)2}$, $\omega^{(0)} = \sqrt{4\pi e^2 n_0/m_e}$. This equation also generates a continuous spectrum of eigenfrequencies [80].

In the case of Langmuir oscillations the continuous spectrum appeared to be the consequence of the vanishing of the group velocity. As a result all the oscillations were local in space and did not move. It turns out that other plasma oscillation modes, in particular the hydromagnetic modes, can exhibit similar characteristics.

Let us consider, by way of example, an Alfvén wave. We know that its group velocity is directed along the magnetic field. Therefore in a plasma layer, such that the Alfvén velocity depends, for instance, on the lateral coordinate x (say, due to the dependence of the density upon x), but which is uniform with respect to translations along the magnetic lines, things will be just the same as in the case of Langmuir waves. Indeed, since the transverse part of the group velocity vanishes, the waves propagating inside different layers will not affect each other, each one running at its own local Alfvén velocity.

Consequently, any initial disturbance will be split into fiber-like wave packets and the net amplitude will, so to speak, decay with time.

This effect was observed in experiments on Alfvén wave propagation along a plasma column created by a θ-pinch contraction. One may introduce small initial perturbations into the plasma column by a pulsed disturbance of the transverse magnetic field. It could be expected that this distortion would run away as a pulse of an Alfvén wave. But in fact it vanished rapidly and almost did not move. This effect was related to the inhomogeneities of the density in the direction transverse to the plasma column axis [81], which resulted in a difference of phase velocities at different distances from the axis. Running at different velocities, the disturbance pieces become desynchronized in a short time and the perturbation disappears. Later on special experiments on the evolution of periodic perturbations of a plasma column in a magnetic field displayed full agreement with

the theory based on the grounds of the continuity of the Alfvén wave spectrum.

It turned out that the spectrum of other hydromagnetic modes in an inhomogeneous plasma appeared to be continuous as well [82]–[84]. A simple example is furnished by the case of a cylindrical plasma column carrying an electric current and immersed in a longitudinal magnetic field. With the notation B_z for the longitudinal and B_θ for the azimuthal parts of the magnetic field we find that the frequency of the Alfvén wave of a helical-like form $\exp(ik_z z + im\theta)$ is given by

$$\omega_A^2 = c_A^2 k_{||}^2 = \frac{1}{4\pi n m_i}\left(k_z B_z + \frac{m}{r} B_\theta\right)^2 .$$

As already stated, these waves run independently within thin coaxial layers.

The slow wave spectrum is also continuous [83]. The dispersion equation for such a strongly localized wave may be obtained by putting $k \to \infty$ in (5.21) and reads as

$$(c_A^2 + c_s^2)\omega^2 - c_A^2 c_s^2 k_{||}^2 = 0. \tag{5.133}$$

Thus, in addition to Alfvén waves there are two families of hydromagnetic oscillation modes with continuous spectra. These modes cannot be rigorously described by the WKB-approximation, for they correspond to singular surfaces where the WKB-approximation is not valid. The continuous spectra are a very interesting peculiarity of one-fluid magneto-hydrodynamics.

REFERENCES

1. Yu. A. Berezin and V. I. Karpman, *Zh. Eksp. Teor. Fiz.*, **51**, 1557 (1966)
2. N. J. Zabusky and M. D. Kruskal, *Phys. Rev. Lett.*, **15**, 240 (1965)
3. C. S. Gardner, J. M. Green, M. D. Kruskal, and R. M. Miura, *Phys. Rev. Lett.*, **19**, 1095 (1967)
4. V. E. Zakharov and L. D. Faddeev, *Funktsional'nyy Analiz i ego Prilozheniya*, **5**, 18 (1971)
5. R. M. Miura, C. S. Gardner, and M. D. Kruskal, *J. Math. Phys.*, **6**, 1204 (1968)
6. I. M. Gel'fand and B. M. Levitan, *Izv. Akad. Nauk SSSR, Matematika*, **15**, 309 (1951); V. L. Marchenko, *Dokl. Akad. Nauk SSSR*, bf 194, 695 (1954)
7. A. S. Scott, F. Y. Chu, and D. W. McLaughlin, *Proc. IEEE*, **61**, 1443 (1973); V. E. Zakharov, S. V. Manakov, S. P. Novikov, and L. P. Pitaevskiy, *Teoriya Solitonov: Metod Obratnoy Zadachi* (Soliton Theory: Inverse Problem Method) [in Russian], Nauka, Moscow (1980)
8. A. V. Gaponov, L. A. Ostrovskiy, and M. I. Rabinovich, *Izv. Vuzov, Radiofizika*, **13**, 163 (1970)
9. S. G. Alikhanov, N. I. Alinovskiy, G. G. Dolgov-Savel'ev, et al., in: *Plasma Phys. and Controlled Nuclear Fusion Research*, Vol. 2, IAEA, Vienna (1969), p. 47.
10. R. J. Taylor, D. R. Baker, and H. T. Ikezi, *Phys. Rev. Lett.*, **24**, 206 (1970);
11. R. Z. Sagdeev, in: *Voprosy Teorii Plazmy* (Problems of Plasma Theory) [in Russian], Vol. 4, Atomizdat, Moscow (1964), p. 20.
12. B. B. Kadomtsev and V. I. Petviashvili, *Dokl. Akad. Nauk SSSR*, **192**, 753 (1970)
13. G. A. Askar'yan, *Zh. Eksp. Teor. Fiz.*, **42**, 1567 (1962); *Usp. Fiz. Nauk*, **111**, 249 (1973)
14. M. J. Lighthill, *J. Inst. Math. Appl.*, **1**, 269 (1965); *Proc. Roy. Soc.*, **A229**, 28 (1967)
15. T. B. Benjamin and J. E. Feir, *J. Fluid Mech.*, **27**, 417 (1967); *Proc. Roy. Soc.*, **A229**, 59 (1967)

16. A. M. Obuhov, et al., *Nelineynye Sistemy Gidrodinamicheskogo Tipa* (Nonlinear Systems of Hydrodynamics Type) [in Russian], Nauka, Moscow (1974)
17. L. D. Landau and E. M. Lifshitz, *Mekhanika* (Mechanics) [in Russian], Fizmatgiz, Moscow (1958)
18. B. Coppi, M. N. Rosenbluth, and R. Sudan, *Ann. Phys.*, **55**, 207 (1969)
19. A. S. Bakai, *Nucl. Fusion*, **10**, 53 (1970)
20. R. A. Stern and N. Tzoar, *Phys. Rev. Lett.*, **17**, 903 (1966)
21. B. Ya. Zeldovich, N. F. Pilipetskiy, and V. V. Shkunov, *Obraschenie Volnovogo Fronta* (Wave Front Inversion) [in Russian], Nauka, Moscow (1985)
22. J. B. Gunn, *Solid State Comm.*, **1**, 88 (1963)
23. J. B. Gunn, "Plasma Effects in Solids," in: *Proc. 7th Int. Conf. Phys. of Semiconductors*, Dunod, Paris (1965)
24. A. F. Volkov and Sh. M. Kogan, *Usp. Fiz. Nauk*, **96**, 633 (1968)
25. M. A. Leontovich, *Izv. Akad. Nauk SSSR, Fizika*, **8**, 16 (1944)
26. R. J. Chiao, F. Gardmire, and C. H. Townes, *Phys. Rev. Lett.*, **13**, 479 (1964)
27. V. I. Bespalov, A. G. Litvak, and V. I. Talanov, *Nelineynaya Optika* (Nonlinear Optics) [in Russian], Nauka, Novosibirsk (1968)
28. V. E. Zakharov and A. M. Rubenchik, *Zh. Eksp. Teor. Fiz.*, **65**, 997 (1973)
29. H. Hasimoto, *J. Fluid Mech.*, **51**, 477 (1972)
30. H. Hasimoto, *J. Fluid Mech.*, **31**, 293 (1971)
31. S. V. Manakov, V. E. Zakharov, L. A. Bogard, A. R. Its, and V. B. Matveev, *Phys. Lett.*, **A63**, 205 (1977)
32. M. J. Ablowitz and J. Satsuwa, *J. Math. Phys.*, **19**, 2180 (1978)
33. K. A. Gorshkov, D. E. Pelinovsky, and Yu. A. Stepanyants, *Zh. Eksp. Teor. Fiz.*, **77**, 237–245 (1993)
34. S. N. Vlasov, V. A. Petrischev, and V. I. Talanov, *Izv. Vuzov, Radiofizika*, **14**, 1353 (1976)
35. V. E. Zakharov, *Zh. Eksp. Teor. Fiz.*, **62**, 1745 (1972)
36. V. I. Petviashvili and O. A. Pokhotelov, *Solitary Waves in Plasmas and in the Atmosphere*, Gordon & Breach Publ. Group (1992); G. H. Derric, *J. Math. Phys.*, **5**, 1252–1254 (1964)

37. E. A. Kuznetsov, S. L. Musher, and A. V. Shafarenko, *Pis'ma v Zh. Eksp. Teor. Fiz.*, **37**, 204–207 (1983)

38. A. A. Vlasov, *Zh. Eksp. Teor. Fiz.*, **8**, 291 (1938); *Usp. Fiz. Nauk*, **93**, 444 (1967)

39. L. D. Landau, *Zh. Eksp. Teor. Fiz.*, **16**, 574 (1946); *Usp. Fiz. Nauk*, **93**, 527 (1967)

40. A. A. Vlasov, *J. Phys. (USSR)*, **9**, 25 (1945)

41. N. S. Van Campen, *Physica*, **21**, 949 (1955)

42. A. A. Vedenov, E. P. Velikhov, and R. Z. Sagdeev, *Yaderny Syntez*, **3**, 1049 (1962)

43. R. K. Mazitov, *Prikl. Mekh. Tekh. Fiz.*, **1**, 27 (1965)

44. T. M. O'Neil, *Fhys. Fluids*, **8**, 2255 (1965)

45. L. M. Altshul and V. I. Karpman, *Zh. Eksp. Teor. Fiz.*, **49**, 515 (1965)

46. I. Bernstein, J. Green, and M. Kruskal, *Phys. Rev.*, **108**, 546 (1957)

47. J. H. Malmberg, C. B. Wharton, and W. E. Drummond, *Plasma Phys. and Contr. Nucl. Fusion Res.* **1**, 485 (1966)

48. J. H. Malmberg and C. B. Wharton, *Phys. Rev. Lett.*, **17**, 175 (1966)

49. G. Van-Hoven, *Phys. Rev. Lett.*, **17**, 169 (1966)

50. H. Derfler and T. C. Simonen, *Phys. Rev. Lett.*, **17**, 172 (1966)

51. H. Derfler and T. C. Simonen, in: *Proc. 8th Int. Conf. on Phenomena in Ionized Gases*, Contributed papers, Vienna (1967), p. 335.

52. J. H. Malmberg and C. B. Wharton, *Phys. Rev. Lett.*, **19**, 775 (1967)

53. E. L. Hahn, *Phys. Rev.*, **80**, 580 (1950)

54. R. M. Hill and D. E. Kaplan, *Phys. Rev. Lett.*, **14**, 1061 (1965)

55. R. M. Hill, D. E. Kaplan, and A. Y. Wong, *Phys. Lett.*, **22**, 585 (1966)

56. R. M. Hill, D. E. Kaplan, and G. F. Hermann, in: *Proc. 8th Int. Conf. on Phenomena in Ionized Gases*, Contributed papers, Vienna (1967), p. 424.

57. R. M. Gould, *Phys. Lett.*, **19**, 477 (1965); *Am. J. Phys.*, **37**, 585 (1969)
58. R. F. Crouford and R. S. Houp, *J. Appl. Phys.*, **37**, 4405 (1966)
59. R. M. Gould, T. M. O'Neil, and J. H. Malmberb, *Phys. Rev. Lett.*, **19**, 219 (1967)
60. J. H. Malmberg, C. B. Wharton, R. M. Gould, and T. M. O'Neil, *Phys. Rev. Lett.*, **20**, 95 (1968)
61. H. Ikezi and N. Takahashi, *Phys. Rev. Lett.*, **20**, 140 (1968)
62. D. R. Baker, N. R. Ahern, and A. Y. Wong, *Phys. Rev. Lett.*, **20**, 318 (1968)
63. C. H. Su and C. Oberman, *Phys. Rev. Lett.*, **20**, 427 (1968)
64. J. H. Malmberg, C. B. Warton, R. M. Gould, and T. M. O'Neil, *Phys. Fluids*, **11**, 1147 (1968)
65. M. Porkolab and J. Sinnes, *Phys. Rev. Lett.*, **21**, 1227 91968)
66. H. Ikezi, N. Takahashi, and K. Nishikiwa, *Phys. Fluids*, **12**, 853 (1969)
67. J. B. Taylor, *Phys. Rev. Lett.*, **33**, 1139 (1974)
68. L. D. Landau, *Zh. Eksp. Teor. Fiz.*, **7**, 203 (1937)
69. D. V. Sivukhin, in: *Reviews of Plasma Physics* (edited by M. A. Leontovich), Vol. 1, Consultants Bureau, New York (1965)
70. A. I. Morozov and A. P. Shubin, *Zh. Eksp. Teor. Fiz.*, **46**, 710 (1964)
71. M. G. Hains, *Phys. Lett.*, **6**, 313 (1963)
72. M. G. Hains, *Advances in Phys.*, **14**, 167 (1965)
73. P. C. Thonemann and A. C. Kolb, *Phys. Fluids*, **7**, 1455 (1964)
74. D. Pfirsh and A. Schlüter, Preprint MPI/PA/7/62, Max-Plank Institute für Physik und Astrophysik, Münich (1962)
75. A. A. Galeev and R. Z. Sagdeev, *Voprosy Teorii Plazmy* (Problems of Plasma Theory) [in Russian], Gosatomizdat, Moscow (1973)
76. R. J. Bickerton, J. W. Connor, and J. B. Taylor, *Nature Phys. Science*, **229**, 110 (1971); B. B. Kadomtsev and V. D. Shafranov, *Nucl. Fusion Res.*, **1**, 172 (1971)
77. V. E. Golant and O. B. Danilov, *Zh. Eksp. Teor. Fiz.*, **33**, 1043 (1963)

78. D. V. Sivukhin, *Magnitnaya Gidrodinamika*, No. 1, 35 (1966); V. P. Demutskii and R. V. Polovin, *Fundamentals of Magnetohydrodynamics*, Kluwer Academic/Plenum Publishers, New York (1990)

79. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, Pergamon, New York (1960)

80. A. V. Timofeev, *Usp. Fiz. Nauk*, **102**, 185 (1970)

81. J. Tataronis and W. Grossman, *Z. Phys.*, **261**, 1 (1973)

82. H. Grad, *Proc. Nat. Acad. Sci. USA*, **70**, 3277 (1973)

83. K. Appert and R. Gruber, *Phys. Fluids*, **17**, 1471 (1974)

84. J. P. Goedbloed, *Phys. Fluids*, **18**, 1258 (1975)