REVIEWS OF TOPICAL PROBLEMS

Millisecond solar radio spikes

G D Fleĭshman, V F Mel'nikov

Contents

Abstract. Currently available models of one of the most intriguing types of unsteady rf solar emission, millisecond solar radio spikes, are discussed. A comparative analysis of the models' implications and of the body of existing data yields an outline of the most realistic radio spike model possible. The spikes are produced by the cyclotron maser mechanism. The cyclotron cone instability is caused by fast electrons distributed over energies according to a (piecewise) power law. The angular part of the distribution function (whose exact form is, as yet, undetermined) suffers fluctuations due to the magnetic field inhomogeneities that arise in the burst loop as a consequence of the original energy release. In some portions of the loop the distribution is not anisotropic enough to secure the development of a cyclotron instability; it is in these 'microtraps' where individual spikes form. Key areas of future theoretical and experimental research are suggested with a view to verifying the adequacy and realizing the diagnostic potential of the model.

1. Introduction

Astrophysical studies play an essential role in modern physics. This is primarily due to the extreme conditions of matter in astrophysical objects, which are unattainable in current laboratory experiments. These conditions include superstrong electromagnetic and gravitational fields, super-high and superlow densities, and powerful processes of energy release.

A powerful energy release is a very general phenomenon. It occurs both in extragalactic (quasars, radiogalaxies, galactic nuclei) and in galactic objects (novae and supernovae, binary systems, neutron stars, and flares in atmospheres of normal stars). Leading to a rich variety of observational manifestations, a powerful energy release practically always demonstrates some general features. First of all is the appearance of accelerated nonthermal particles (electrons and nuclei). The role of these particles in the dynamics and evolution of cosmic objects can hardly be overestimated. In addition to the directly observed nonthermal electromagnetic radiation, such particles can appreciably contribute to the ionization and heating of the surrounding medium, the modification of hydrodynamic flows, the chromospheric evaporation during solar and stellar flares, etc.

Recent progress in observational techniques has brought the studies of the closest star to the Earth, the Sun, especially of solar activity, to a qualitatively new level. The diversity and quality of the currently available physical information about the processes under way in solar flares has approached the level attainable in laboratory experiments.

However, the role of the solar atmosphere (as well as of the terrestrial ionosphere and interplanetary space) as a cosmic laboratory, which allows many physical processes typical for more distant cosmic objects to be studied in detail, has so far been underestimated by astrophysicists. The evolution and dynamics of magnetic fields, current sheet generation, shock wave formation and charged particle acceleration by shock waves are very general phenomena occurring both in atmosphere of the Sun and other stars, binary systems, accretion disc magnetospheres, supernova explosions, etc.

It is worth specially emphasizing that solar flares are accompanied by a genuine firework of observational consequences of the powerful energy release occurring in diverse

G D Fleĭshman A F Ioffe Physical-Technical Institute,
Russian Academy of Sciences
ul. Politekhnicheskaya 26, 194021 St. Petersburg, Russia
E-mail: gregory@sun.ioffe.rssi.ru
V F Mel'nikov Scientific Research Institute of Radio Physics
ul. B. Pecherskaya 25, 603600 GSP-51 Nizhniĭ Novgorod, Russia

physical conditions. Direct observational evidence of magnetic field line 'reconnection', which was theoretically predicted long ago, provides an example. These arguments permit us to hope that some particular problems of solar flares may be of interest to many physicists.

The cyclicity of solar activity not only has an effect on the surrounding natural phenomena, but is of major significance for solar activity studies. For example, during the years of maximum activity experimental studies of solar flares become most important. In recent decades they have covered the total range of the electromagnetic spectrum and also included measurements of neutral and charged particle fluxes. These observations have been carried out both from the ground and from the upper atmospheric layers or from space using balloons and satellites. In the last few years coordinated multiwavelength observations of solar flares with different instruments have played a growing role, providing us with the most complete information on the physical processes taking place.

Since the periods of maximum solar activity are rather short, the development of thoroughly thought out and elaborated observational programs aiming at obtaining desirable physical information is especially important, otherwise we shall have wait for another 10 years.

Years of the minimum solar activity are equally important for solar flare studies. Although much fewer observational data are available during these years, the lack of new data forces us to analyse more deeply the existing data and construct more sophisticated theories and models. Thus, the years of minimum are the time for analysis and discussion of the existing observations and for planning new observational programs for the next solar activity maximum. Such preparation requires knowing what exactly has been obtained in previous studies, what models, hypotheses, mechanisms, etc. are most reliable and most agree with the existing data, what is to be studied and by which means.

The present paper is devoted to one particular, but sufficiently important phenomenon related to solar flares – millisecond radio spikes. Radio spikes represent one of the types of superfine (the characteristic time $t < 1$ s) temporal structure of solar flares radio emission, including also pulsations (regular and irregular), 'antispikes' — millisecond dips of the emission intensity, etc.

Radio spikes with a duration less than 100 ms were detected more than 30 years ago [1–3] at metre wavelengths; the first reviews appearing only in 1967 [4, 5]. In subsequent years the interest in studying spikes (as the emission demonstrating the shortest characteristic time scales and the narrowest spectral widths among the all types of solar electromagnetic radiation) has grown rapidly. In particular, short spikes (at separate frequencies) were detected in the decametre wavelength range (18–26 MHz) [6]. Spikes were also observed at microwave wavelengths [7–11] up to a frequency of 8 GHz. High time resolution observations at shorter wavelengths (10–100 GHz) have not revealed bursts with the characteristic spike durations [10]. As shown in the recent reviews devoted to spikes [12, 13], they typically appear in the decimetre wavelength range 300–3000 MHz. Events with spikes are visualized on the dynamic spectra shown in Fig. 1.

The corresponding particle concentration and magnetic field strength are characteristic of a plasma in which, according to modern concepts, the energy release and particle acceleration occur during solar flares. Usually spikes
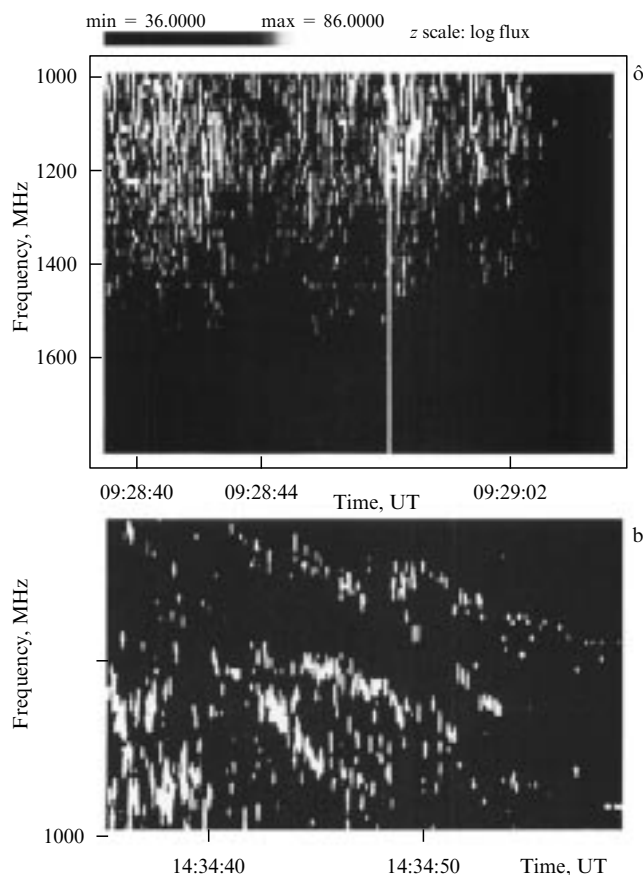
Figure 1. Dynamic spectra of a decimetre millisecond radio spike cluster in the events of 25.12.90 registered with the PHOENIX spectrometer in Zurich [101] (a) and of millisecond spikes associated with a type IV burst in the event of 6.04.80. Several chains of very narrow-band (about 10 MHz) spikes are seen [14] (b).

appear in groups and occupy some finite frequency range (of the order of a few hundred MHz) with a total duration of the order of 1 min. The number of individual spikes during the flare can be as high as a few thousand and even exceed 10000 [12]. These properties (see Ref. [12] for more detail) led to the interpretation of a single spike as a microflare, and the presence of a large number of spikes during the flare led to the concept of energy release fragmentation (each elementary act of energy release gives rise to one spike). In such a model the radio spikes prove to be most closely related (among all types of flare emission) to the process of energy release in solar flares. It is this fact that explains the very high interest of both observers and theoreticians in the problem of radio spikes.

However, along with the fragmentation of the primary energy release, one should point to other possible reasons for the strong radio emission fragmentation observed in events with spikes. Firstly, this fragmentation may be produced by intrinsic properties of the radio emission generation mechanism itself, and secondly, it may result from secondary fragmentation of the radio emission source structure. As we show below, the data collected up to now favour the secondary fragmentation concept.

Considering an extremely short duration of the spikes (down to a few milliseconds [10]), a small proper frequency width ($< 1\%$) [12], a high brightness temperature and (often) the high degree of observed polarization, it becomes clear that

the radio spikes have a high diagnostic potential. However, until recently this potential remained unrealized, mainly because even the very mechanism of radio emission generation in spikes was not established, not to speak of a model of the source.

In this review, we have tried to analyse all the currently available observational data on radio spikes and to compare them with the predictions of existing theoretical models. Such a comparison allows us to separate the most adequate model, according to which radio spikes are generated by nonthermal electrons with an extended energy spectrum by a cyclotron maser mechanism in a coronal loop with strong magnetic field inhomogeneities. It also permits us to design 'crucial studies' (both theoretical and experimental) capable of finally establishing the emission mechanism and developing a realistic model for the source.

In Section 2 we describe different mechanisms and models generating the superfine temporal structure of the radio emission. Section 3 contains the results of a detailed analysis of observational data and a comparison with models. In Section 4 the problems to be solved in the course of the coming 23rd solar activity cycle and ways of their solution are discussed. The last section is devoted to final remarks and the conclusion.

2. Generation mechanisms of the superfine temporal structure of solar radio emission

The source in which an individual spike is generated is very compact. Indeed [13], with the relative emission frequency band $\Delta f/f \sim 1\%$ and the characteristic scale of plasma parameter changes inside the corona $L \sim 2 \times 10^9$ cm we can obtain the linear size of the source

$$l \leqslant L(\Delta f/f) \approx 200 \text{ km} . \tag{1}$$

Correspondingly, the brightness temperatures of the spikes T_b can reach (and even exceed) 10^{15} K [12, 13]. Such a high brightness temperature, along with the extremely short duration of the events and (often) the strong circular polarization of radio emission, can only be provided by some coherent mechanism. Different modifications of plasma and cyclotron maser effects are currently involved as possible spike generation mechanisms. We also mention a still poorly studied possibility of the realization of an instability involving transition emission (maser transition radiation) in solar flares. The difficulty in establishing a reliable spike emission mechanism is due to the similarity of many models' predictions (accessible to observational tests). So in this section we shall focus on the differences between these mechanisms.

2.1 Plasma mechanism of radio emission

The plasma radio emission mechanism in the broad sense is a two-stage process in which at the first stage the eigen modes are generated in a source which cannot leave the plasma. These can be Langmuir, upper- and lowerhybrid modes, Bernstein modes, slow extraordinary modes (Z-mode), etc. At the second stage these waves are transformed into the fast extraordinary (X-mode) or ordinary (O-mode) transverse electromagnetic waves, which leave the source and are observed on the Earth as the radio emission.

2.1.1 The process of transformation of Langmuir waves generated by electron beams into transverse waves near plasma frequencies ω_{pe} and $2\omega_{pe}$ was first proposed almost 40 years ago [15] to interpret type II and III bursts at meter wavelengths. It is presently recognized that type III bursts with a drifting frequency (decreasing or increasing) indicate propagating (upwards or downwards) electron beams, and the plasma mechanism is undoubtedly realized in such bursts. Events similar to meter type III bursts are also observed in decimetre [14, 16] and microwave [17] wavelengths.

Plasma waves excited when an electron beam propagates inside plasma stabilize the instability behind the beam front, however on the front the instability is not stabilized (faster electrons always outrun slower ones so a positive derivative of the distribution function over momentum holds on the front). In such an (unsaturated) regime the speed of the burst frequency drift reflects the speed of the beam front motion in the corona and the sign of the drift reflects the motion direction.

The spectral properties of the radio spikes differ radically from type III bursts so one may neglect the possibility of spike generation by electron beams through the plasma mechanism.

2.1.2 Plasma turbulence is also thought to generate noise storms which consist of a continuum component and type I bursts, narrow-band short bursts. According to one model [18], type I bursts appear as a result of plasma wave conversions on quasi-stationary structures, which are formed by the plasma waves themselves. Such a process is effective at a sufficiently low plasma turbulence level

$$w_{pl} = \frac{W_{pl}}{nT} < 10^{-7} . \tag{2}$$

At higher w_{pl} these bursts turn out to be weaker than the continuum and then type IV bursts appear [18]. The model [18] predicts type I burst emission to arise at the local plasma frequency ω_{pe} (i.e. a harmonic structure for the burst is absent) with a relative band $\Delta f/f \sim 3\%$ and characteristic duration $t \sim 0.1-2$ s. A similar model is proposed in Ref. [19] for the radio spikes. However, significant differences in spatial distribution over the disk, frequency band, and polarization for type I bursts and radio spikes [13], as well as the harmonic structure of radio spikes (see below for more detail) lead to the conclusion that the model [18], although satisfactorily explaining type I burst properties, cannot explain the radio spikes.

2.2 Nonlinear plasma mechanisms

A fraction of the charged particles accelerated during the primary energy release in a solar flare is caught by magnetic traps, the coronal arches (loops). Even if initially these particles are distributed isotropically, the motion inside the trap makes the particle distribution anisotropic. Indeed, during the motion of a particle in an inhomogeneous magnetic field the transverse adiabatic invariant $p_\perp^2/B = \text{const}$ is conserved. Since the background plasma density in the corona (as well as the magnetic field strength) increases toward the solar surface, particles with smaller pitch-angles penetrate into deeper (and hence denser) atmospheric layers, lose more energy and disappear, in contrast to particles with high pitch-angles which are reflected from magnetic plugs at high altitudes. Thus it is natural to expect the formation of electron distributions with a 'loss cone', i.e. with a deficit of electrons with small pitch-angles. The critical angle depends on the plug relation in a given magnetic trap.

Anisotropic electron distributions (as well as electron beams) can be unstable with respect to plasma wave generation. In such a case a cone instability is said to be realized. If an unstable electron distribution with a loss-cone forms in a plasma, the plasma wave energy density grows exponentially so that non-linear effects become significant (in contrast to the beam front that sweeps fresh plasma regions where plasma waves have no time to accumulate). These nonlinear effects include the back reaction of amplified waves on fast electrons and proper non-linear wave effects. In the weak turbulence approximation the evolution of the plasma wave energy density $W_{\pm}(\mathbf{k})$ and fast electron distribution function $f(p, \mu, t)$ are described by equations [20–27]

$$\frac{\mathrm{d}}{\mathrm{d}t} W_{\pm}(\mathbf{k}) = 2\gamma_{\pm}^{\mathrm{QL}} W_{\pm}(\mathbf{k}) + J(l + i \Rightarrow l') , \qquad (3)$$

$$\frac{\partial f}{\partial t} = \frac{16\pi^3 e^2 \gamma}{m_{\mathrm{e}}^2 p} \sum_{\omega_{\pm}} \sum_{n=-\infty}^{\infty} \int \omega_{\pm}^2 \, \mathrm{d}k \, \mathrm{d}\eta$$

$$\times \delta(kp\eta\mu + n\omega_{\mathrm{Be}} - \gamma\omega_{\pm}) \left[\frac{\partial}{\partial p} + \frac{\partial}{\partial \mu} \frac{\gamma}{p} \left(\frac{k\eta p}{\gamma\omega_{\pm}} - \mu \right) \right]$$

$$\times \left\{ \frac{J_n^2(\lambda) R_{\pm}(\mathbf{k}) W_{\pm}(\mathbf{k})}{p} \left[\frac{\partial f}{\partial p} + \frac{\gamma}{p} \left(\frac{k\eta p}{\gamma\omega_{\pm}} - \mu \right) \frac{\partial f}{\partial \mu} \right] \right\}$$

$$+ F_{\mathrm{i}}(\mathbf{r}, \mathbf{p}, t) + \hat{L}f(\mathbf{r}, \mathbf{p}, t) , \qquad (4)$$

where $\gamma_{\pm}^{\mathrm{QL}}$ is the quasi-linear growth rate of the plasma wave

$$\gamma_{\pm}^{\mathrm{QL}} = \frac{2\pi^2 \omega_{\mathrm{pe}}^2 \omega_{\pm}^2(\mathbf{k})}{k^2} \frac{n_{\mathrm{b}}}{n_0} R_{\pm}(\mathbf{k})$$

$$\times \sum_{n=-\infty}^{\infty} \int J_n^2(\lambda) \left[p \frac{\partial f}{\partial p} + \left(\frac{k\eta p}{\gamma\omega_{\pm}} - \mu \right) \frac{\partial f}{\partial \mu} \right]$$

$$\times \delta(kp\eta\mu + n\omega_{\mathrm{Be}} - \gamma\omega_{\pm}) \, \mathrm{d}p \, \mathrm{d}\mu , \qquad (5)$$

$J(l + i \to l')$ is the term describing stimulated wave scattering on plasma ions

$$J(l + i \Rightarrow l') = -\frac{e^2 \omega_{\pm}(\mathbf{k})}{\pi^2 m_{\mathrm{e}}^2 \omega_{\mathrm{pe}}^4} W_{\pm}(\mathbf{k}) R_{\pm}(\mathbf{k})$$

$$\times \int k''^2 \, \mathrm{d}\mathbf{k}' W_{\pm}(\mathbf{k}') R_{\pm}(\mathbf{k}') \frac{(\mathbf{k}\mathbf{k}')^2}{k^2 k'^2}$$

$$\times \frac{\delta\epsilon_{\mathrm{e}}'^2(\omega'', \mathbf{k}'') \delta\epsilon_{\mathrm{i}}''(\omega'', \mathbf{k}'')}{\left| \epsilon^l(\omega'', \mathbf{k}'') \right|^2} , \qquad (6)$$

$\hat{L}$ is the operator describing the magnetic field focusing, Coulomb losses, and diffusion due to Coulomb encounters

$$\hat{L} = -\beta c\mu \frac{\partial}{\partial s} + \beta c \frac{d\ln B}{ds} \frac{\partial}{\partial \mu} \left[\frac{(1-\mu^2)}{2} \right] + \frac{c}{\lambda_0} \frac{\partial}{\partial E} \left(\frac{1}{\beta} \right)$$

$$+ \frac{c}{\lambda_0 \beta^3 \gamma^2} \frac{\partial}{\partial \mu} \left[(1-\mu^2) \frac{\partial}{\partial \mu} \right] , \qquad (7)$$

and $F_{\mathrm{i}}(\mathbf{r}, \mathbf{p}, t)$ is the source of fast electrons. The nonthermal electron distribution function $f(p, \mu, t)$ is normalized to 1. The signs $+$ and $-$ relate to upper- and lowerhybrid waves, respectively, which in a magnetoactive plasma have the spectra

$$\omega_{\pm}^2 = \frac{1}{2}(\omega_{\mathrm{pe}}^2 + \omega_{\mathrm{Be}}^2) \pm \frac{1}{2} \left[(\omega_{\mathrm{pe}}^2 + \omega_{\mathrm{Be}}^2)^2 - 4\omega_{\mathrm{pe}}^2 \omega_{\mathrm{Be}}^2 \cos^2\theta \right]^{1/2} . \qquad (8)$$

Here n_{b}, n_0 denote the concentrations of the fast and background plasma electrons, respectively; ω_{Be} is the gyrofrequency; $\eta = \cos\theta$; θ is the angle between vectors $\mathbf{k}$ and $\mathbf{B}$; $J_n(\lambda)$ is the Bessel function; $\lambda = k_{\perp} p_{\perp}/\omega_{\mathrm{Be}}$; $\mu = \cos\vartheta$; ϑ is the electron pitch-angle; $\beta = v/c = p/\gamma c$ is the dimensionless electron velocity; $p = P/m_{\mathrm{e}}$ is the specific momentum per unit mass; $\gamma = \sqrt{1 + p^2/c^2}$ is the Lorenz-factor; E is the kinetic energy of electrons; $R_{\pm}(\mathbf{k}) = \{\partial(\omega\epsilon^l)/\partial\omega\}_{\omega=\omega_{\pm}}^{-1}$; $\omega'' = \omega - \omega'$; $\mathbf{k}'' = \mathbf{k} - \mathbf{k}'$; $\epsilon^l(\omega'', \mathbf{k}'')$ is the longitudinal dielectric permittivity; $\delta\epsilon_{\mathrm{e}}'(\omega'', \mathbf{k}'')$, $\delta\epsilon_{\mathrm{i}}''(\omega'', \mathbf{k}'')$ is the partial real electron and imaginary ion dielectric permittivity, respectively, with account of thermal motion of particles; s is the coordinate along the magnetic field; $\lambda_0 = (n(s)\ln\Lambda)^{-1} \times 10^{24}$ cm, and $\ln\Lambda$ is the Coulomb logarithm.

2.2.1 Let us suppose that (with the plasma wave energy density increase) the back reaction of the waves on fast particles starts earlier than the proper non-linear wave interactions become significant. In this case the latter can be neglected and one may consider the 'waves – particles' system in the so-called quasilinear approximation.

Then a quasilinear relaxation is the solution of Eqns (3) and (4), so that the particles are isotropized and the plasma wave density first increases to a maximum W_{sat} (the saturated energy density) on a time scale

$$t_{\mathrm{gr}} \sim \frac{1}{\gamma} \ln \left(\frac{W_{\mathrm{sat}}}{W_0} \right) , \qquad (9)$$

where W_0 is the initial plasma wave energy density, and $\gamma \sim (10^{-2} - 10^{-3})(n_{\mathrm{b}}/n_0)\omega_{\mathrm{pe}}$ is the characteristic plasma wave growth rate determined by Eqn (5), and then decreases due to, for example, collisional damping on a time scale [28]

$$\tau_{\mathrm{ei}} = 0.18 \frac{T^{3/2}}{n_0 \ln\Lambda} , \qquad (10)$$

where $\ln\Lambda \approx 11.2$. Therefore, in the quasilinear regime a pulse of plasma waves appears which are then transformed into electromagnetic waves at frequencies ω_{pe} or $2\omega_{\mathrm{pe}}$ generating a radiopulse either without any harmonic structure or with a 1:2 harmonic ratio. Applied to the radio spikes, such a model with radiowave emission at a double plasma frequency has been proposed in Ref. [29].

When a permanent source of fast particles is present, this process can multiply repeat, otherwise the isotropization will lead to an enhanced precipitation of fast electrons from the trap and to violation of the instability condition. We shall discuss quasilinear effects in more detail in Sections 2.3.7 and 2.3.8.

2.2.2 In the magnetoactive plasma with $\omega_{\mathrm{pe}} > \omega_{\mathrm{Be}}$ the upperhybrid waves with $\omega_{\mathrm{UH}} = \omega_+$ (8) are unstable. Correspondingly, at the first harmonic $\omega \approx \omega_{\mathrm{UH}} = \omega_+$ they can transform only into ordinary waves, i.e. the radio emission will be 100% polarized. At the second harmonic $2\omega_{\mathrm{UH}}$ the sign and degree of polarization can be arbitrary. The generation of the upperhybrid waves in plasma with a strong magnetic field $\omega_{\mathrm{Be}} \geqslant \omega_{\mathrm{pe}}$ is considered in paper [30]. In such a situation the upperhybrid frequency ω_{UH} is close to the electron cyclotron frequency ω_{Be} and the radio emission arises due to the coupling of two upperhybrid waves at frequency $2\omega_{\mathrm{UH}} \approx 2\omega_{\mathrm{Be}}$. In that paper a system of model

equations for amplitudes and phases of longitudinal and transverse waves was analysed. The solutions of this system demonstrate irregular spike-like behaviour in the radio emission pulse sequence.

The radio emission generated by the above plasma mechanisms either has no harmonic structure or shows two harmonics with frequency ratio 1:2. It is very difficult to provide third or higher harmonics. For this to be realized either a simultaneous coupling of three or more plasma waves is required, which is improbable, or the coupling of a plasma wave at ω_{pe} with a transverse wave at $2\omega_{\mathrm{pe}}$ should occur:

$$l(\omega_{\mathrm{pe}}) + t(2\omega_{\mathrm{pe}}) \rightarrow t(3\omega_{\mathrm{pe}}) \, . \tag{11}$$

It is very difficult to satisfy the momentum conservation law for this process due to the different dispersion laws for transverse and longitudinal waves.

2.2.3 However, in a magnetoactive plasma with

$$0.1 \leqslant Y \leqslant 1 \, , \tag{12}$$

where $Y = \omega_{\mathrm{pe}}/\omega_{\mathrm{Be}}$, the growth rates of quasi-transverse slow extraordinary waves (Z-mode) can be quite large. Since such waves are unable to leave the plasma (due to total internal reflection), their transformation into fast extraordinary or ordinary waves is required to generate the observed radio emission. So we consider such a two-step process as a variant of the plasma mechanism. Melrose [31] studied a model in which Z-waves are generated by electrons due to the cone instability and then are transformed into transverse t-waves (t denotes X- or O-wave) in the processes

$$Z + Z' \rightarrow t, \qquad Z + t \rightarrow t' \, . \tag{13}$$

Since the refraction coefficients of Z-waves are close to unity (although slightly higher), it is rather easy to meet the kinematic conditions

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 \, , \qquad \omega_1 + \omega_2 = \omega_3 \, , \tag{14}$$

which should be fulfilled for any of the three-wave processes (13). Thus the first act of the wave coupling $Z + Z' \rightarrow t_2$ results in a wave near the second gyrofrequency harmonic, the next act $Z + t_2 \rightarrow t_3$ forms a wave at the third harmonic, etc. The number of harmonics formed in this process, their intensity and polarization can be derived from the corresponding transfer equations [31], whose solutions are ultimately determined by the free energy stored in the system and the characteristic time of the wave transformations. The brightness temperature of emission at the n-th harmonic is maximal if its generation occurs in the saturated regime, i.e. when the number of couplings generating t_n equals the number of decays. Clearly, for the saturated regime of t_n generation to set in it is necessary that t_{n-1} be generated in the saturated regime as well. Since the probability of X-wave formation is higher than O-wave formation, the former will saturate faster. This means that in the fully saturated regime the emission at the n-th harmonics is unpolarized. In the unsaturated regime mostly X-waves are generated and the degree of polarization can reach up to 100%.

The predictions of the model [31] which can be tested observationally are as follows:

a) radio emission is generated at the second, third, and (possibly) fourth harmonics of the gyrofrequency; higher harmonics are very difficult to generate; special conditions are required for emission at the fundamental frequency;

b) the degree of polarization varies from 0 to 100%;

c) the degree of polarization increases from lower to higher harmonics;

d) the direction of rotation of the polarization vector corresponds to X-waves.

Thus the properties of this mechanism differ significantly from the plasma mechanisms considered above, which are based on generation of Langmuir or upperhybrid waves.

2.2.4 Now consider the case where the nonlinear wave interactions become effective significantly earlier than the quasilinear effects of back reaction of waves on fast particles are important. Then in the first approximation we cannot consider the equation for the particle distribution function (4) and instead focus on the plasma wave energy density equation (3). In typical conditions the main nonlinear effect is the so-called induced wave scattering on plasma ions [20–22, 27]. The corresponding nonlinear term $J(l + i \rightarrow l')$ of Eqn (3) has the form (6).

Thus the evolution of the energy density of quasi-longitudinal (plasma) waves is described by the equation

$$\frac{\partial W_{\pm}}{\partial t} + \mathbf{v}_{\mathrm{g}} \frac{\partial W_{\pm}}{\partial \mathbf{r}} = (2\gamma_{\pm}^{\mathrm{QL}} - v_{\mathrm{eff}}) W_{\pm} + J(l + i \Rightarrow l') \, , \tag{15}$$

where $\mathbf{v}_{\mathrm{g}}$ is the wave group velocity. The types of solutions of Eqn (3) mostly depend on the relation of the kernel 'width' in integral (6) and the excited wave function $W_{\pm}$. If the kernel is significantly narrower than the function $W_{\pm}$, one can try to transit from the integral form to a differential approximation by expanding $W_{\pm}$ in series. However solutions with steep fronts appear in this case, for which differential approximation conditions are violated due to the high values of derivatives on the front. This means that it is essential to account for the integral character of this nonlinear interaction.

2.2.5 In another limiting case when the instability region [i.e. the region where $(2\gamma_{\pm}^{\mathrm{QL}} - v_{\mathrm{eff}}) > 0$] is appreciably narrower than the kernel in Eqn (6), Eqn (3) can be significantly simplified as follows [32]. We isolate a resonance region in the wave vector space (where the waves are generated by particles) and a non-resonant region (where the waves are absorbed either due to Coulomb collisions or by fast electrons themselves). After integrating Eqn (3) first over the resonance and then nonresonance regions and by denoting the plasma wave energy densities inside these regions by W and W^*, respectively, we arrive at

$$\frac{\partial W}{\partial t} + \mathbf{v}_{\mathrm{g}} \frac{\partial W}{\partial \mathbf{r}} = \gamma W - \eta W W^* \, , \tag{16}$$

$$\frac{\partial W^*}{\partial t} + \mathbf{v}_{\mathrm{g}} \frac{\partial W^*}{\partial \mathbf{r}} = -\tilde{\gamma} W + \eta W W^* \, . \tag{17}$$

Here

$$\gamma = \langle (2\gamma^{\mathrm{QL}} - v_{\mathrm{eff}}) \rangle \tag{18}$$

is some effective ('mean') growth rate in the instability region, $\tilde{\gamma}$ is the effective decrement of the wave damping inside the non-resonant region, and the nonlinear interaction constant η is estimated as [32]

$$\eta \approx 10^{-1} \frac{\omega_{\mathrm{pe}}}{n_0 T_{\mathrm{i}} (1 + T_{\mathrm{e}}/T_{\mathrm{i}})^2} \, , \tag{19}$$

where T_i, T_e are the temperature of ions and electrons in the plasma, respectively. Note that the accuracy of system (16), (17) is poorly controlled, since it is derived by applying the mean value theorem [32] and its coefficients γ, $\tilde{\gamma}$, η are determined only to an order of magnitude.

If the source is sufficiently extended,

$$L \gg \frac{v_g}{\sqrt{\gamma\tilde{\gamma}}}\,, \tag{20}$$

we can neglect the spatial derivatives in Eqns (16), (17) and reduce these equations to a Lottka–Volterra type system, which describes nonlinear periodical oscillations corresponding to looped curves on the phase plane $W^* - W$ encircling a singular 'centre'-like point:

$$W_0 = \frac{\tilde{\gamma}}{\eta}\,, \qquad W_0^* = \frac{\gamma}{\eta}\,. \tag{21}$$

On the phase trajectories we have

$$W - \frac{\tilde{\gamma}}{\eta}\ln\left(\frac{\eta W}{\gamma}\right) + W^* - \frac{\gamma}{\eta}\ln\left(\frac{\eta W^*}{\gamma}\right) = \widetilde{W}\,, \tag{22}$$

where the integration constant $\widetilde{W}$ relates to the initial conditions and determines the modulation of oscillations. The pulsation period is found by integration along the trajectory (22) and is [33]

$$T = \frac{\eta}{\gamma}\int \frac{dW^*}{\eta W(W^*) - \tilde{\gamma}}\,. \tag{23}$$

In the weak modulation regime

$$T = \frac{2\pi}{\sqrt{\gamma\tilde{\gamma}}}\,, \tag{24}$$

and for strong modulation (when $\gamma \gg \tilde{\gamma}$)

$$T = \frac{1}{\tilde{\gamma}}\ln\left(\frac{\gamma}{\eta W^*(t=0)}\right)\,. \tag{25}$$

Note that if the wave absorption is determined by collisions of thermal electrons with ions so that $\tilde{\gamma} \sim \nu_{ei}$, the oscillation period (25) is practically independent of the fast electron parameters and hence can be very stable. For typical conditions in coronal arches the characteristic oscillation periods can vary from 10 milliseconds to several seconds. Thus the considered nonlinear plasma mechanism for radio emission (in the given case the emission at the second harmonic $\omega \approx 2\omega_{pe}$ dominates [34, 35]) can produce a superfine millisecond time structure for the solar radio emission like millisecond pulsations but not spikes. Possible reasons for solar radio emission pulsations are discussed in Refs [36–42] in more detail (see also the review [43] and references therein).

2.2.6 However, as already noted, the reduction of Eqn (3) to a Lottka-Volterra type system is a rather crude approximation and cannot describe the real situation even qualitatively. In paper [44] a more precise numerical analysis of the nonlinear Eqn (3) is made for two different distributions of the upperhybrid wave growth rate in the wave vector space for

$$Y = \frac{\omega_{pe}}{\omega_{Be}} \gg 1\,. \tag{26}$$

In the first case a narrow enough distribution of growth rates on the plane $k_\perp$, $k_\parallel$ was considered (Fig. 2a), which is realized for the following fast electron distribution function

$$f(\mathbf{p}) = F(p) \cdot \Phi(\mu)\,, \tag{27}$$

where μ is the cosine of the electron pitch-angle and

$$\Phi(\mu) \propto \exp\left(-\frac{\mu^2}{\mu_0^2}\right)\,, \tag{28}$$

μ_0 is the cosine of the loss-cone angle,

$$F(p) \propto p^{-\alpha}\,, \qquad 0.2 \leqslant \frac{p}{mc} \leqslant 1\,, \qquad \alpha = 4\,. \tag{29}$$

The result of numerical solution of Eqn (3) for the plasma parameters

$$T_e = T_i = 10^6 \text{ K}\,, \quad Y = 5\,, \quad \omega_{pe} = 10^9 \text{ c}^{-1} \tag{30}$$

is shown (for the total energy of upperhybrid waves) in Fig. 2c. It is clear that after some transition process a regime of almost periodic oscillations of the plasma wave energy density sets in. Similar oscillations will be observed in radio emission, which in this case (like in Refs [34, 35]) is mostly generated near the second plasma frequency harmonic $2\omega_{pe}$. Thus these calculations confirm the qualitative predictions deduced from the analysis of the approximate Lottka-Volterra system, although the numerical value of the oscillation period can be several times higher than the approximate value derived under the same initial parameters from Eqn (25) (which, strictly speaking, is not applicable because $\tilde{\gamma} > \gamma$). The correct period may be obtained by substituting γ into Eqn (25) instead of $\tilde{\gamma}$. The reason for this becomes clear noticing that the characteristic times of increase and decrease of a single pulse are of the same order of magnitude (Fig. 2c).

Alternatively, a wider instability region (Fig. 2d) was considered, which is realized when the momentum dependence of the distribution function (27) is taken, instead of Eqn (29), in the form of a function with positive derivative with respect to p:

$$F(p) \propto \frac{p^4}{(p_0^2 + p^2)^4}\,, \quad 0.02 < \frac{p}{mc} < 3\,, \tag{31}$$

with the same angular distribution $\Phi(\mu)$. In this case the dependence of the total plasma wave energy density on time (Fig. 2e) is radically different from that predicted by the Lottka–Volterra system. Periodical solutions are absent, and irregular sequences of flares appear. Note significant variations in individual burst time profiles in the two regimes of the nonlinear plasma mechanism operation discussed above. The logarithmic derivative of the wave energy density as a function of time is shown in Fig. 2c, f. The horizontal parts of these curves correspond to an exponential change of energy density with time (increase or decrease depending on the sign of the logarithmic derivative). It is seen that for the periodic regime almost exponential laws are typical both on the increasing and decreasing parts of an individual pulse. Conversely, in the nonperiodic regime these curves demonstrate a much more complex behaviour including clearly distinguished over-exponential growth at some parts.

Assuming the solar radio spikes to result from such an irregular nonlinear plasma mechanism, the following model predictions should be tested observationally:
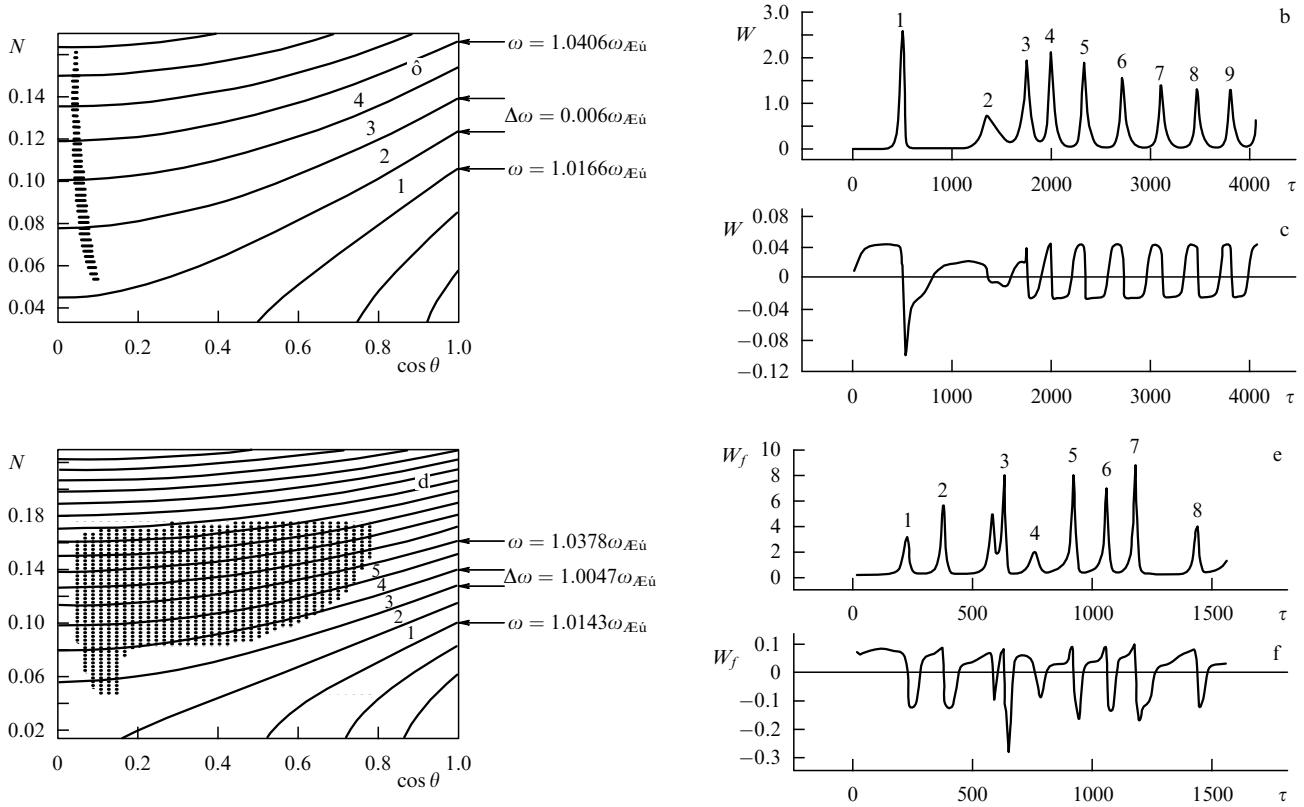
Figure 2. (a) Instability region for distribution (29) (thick hatching) on the plane $N, \cos(\theta)$. The solid lines mark the constant frequency levels. A non-linear wave transition mainly occurs in regions marked with figures $1-4$ [44]. (b) The plasma wave energy density W as a function of the dimensionless time W for distribution (29). (c) The logarithmic derivative W as a function of time [44]. (d) The instability region for distribution (31) (thick points) on the plane $N, \cos(\theta)$. The solid lines correspond to constant frequencies. A non-linear wave transition mainly occurs in regions $1-5$ [44]. (e) The dependence of W on the dimensionless time for distribution (31). (f) The logarithmic derivative W as a function of time [44].

a) the radio emission shows no harmonic structure (radiowaves are mostly generated at the second plasma frequency harmonic, at the first harmonic wave generation is less probable);

b) the bursts at a given frequency ω are emitted from the same place in the coronal arch with $\omega_{pe} \sim \omega/2$;

c) more or less smooth transitions from oscillations to aperiodic bursts and vice versa are possible during slow changes of the fast particle distribution function parameters (due to injection, coming out of the trap, isotropization, etc.);

d) time profiles of individual spikes have a complex structure with occasional 'over-exponential' intensity growth.

2.2.7 To conclude this section, consider one more model based on a plasma mechanism and leading to the formation of a superfine time structure of the solar flare radio emission. Let a quasi-stationary loss-cone fast electron distribution form in the coronal arch providing the continuum emission (type IV bursts [45]). Let now an additional beam of fast electrons be injected into the trap (as a result of some particle acceleration process). Then the situation is possible where the total anisotropy of the 'trapped-and-injected-electron' system can decrease (due to loss-cone filling by the injected electrons). Here either the instability growth rate decreases or is fully stabilized and develops to the same level only after the passing of the beam [46, 47]. As a result, the radio emission intensity shows a strong dip with a drifting frequency as the injected electron beam propagates. If such an injection takes a sufficiently long time and does not suppress but only

decreases the instability growth rate, oscillations relative to some reduced level can appear in the system [48, 49]. Thus the plasma mechanism with episodic injection of fast electrons into the coronal arch explains the appearance of drifting millisecond 'antispikes' and pulsations below the main burst level. A fragment of an event with such features was registered on November 17, 1991, at the Crimean Astrophysical Observatory and is shown in Fig. 3.

2.3 Cyclotron maser emission
If the electron distribution in the magnetoactive plasma is unstable relative to the generation of transverse (X or O) waves near the lower cyclotron frequency harmonic, the corresponding emission is called cyclotron maser emission. Such an instability was shown to be possible quite long ago [50, 51]. Much progress in studies of cyclotron maser emission (CME) was made after the discovery [52] and subsequent observations [53] of auroral kilometre radiation (AKR) in the terrestrial magnetosphere. In paper [54] a semi-relativistic CME theory was proposed (in application to AKR) and the principal necessity of taking into account the relativism of overthermal radiating electrons was shown, even if they can be considered non-relativistic in other respects ($v/c \ll 1$). Later on the CME theory was intensively developed [55 – 62] and experimentally confirmed [63 – 68] by space measurements of particle distribution functions and electromagnetic fields in the terrestrial magnetosphere. CME is also thought to operate in some other planets' radio emission — Jupiter [69,70], Saturn [71], Uranus and Neptune [72, 73]. These
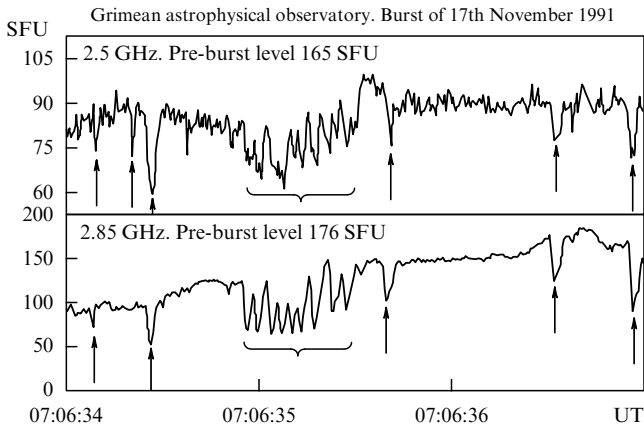
Figure 3. Time profiles of the event of 17.11.91 at frequencies of 2.5 and 2.85 GHz with a time resolution of about 10 ms. A decrease in the 'zero' level during the flux density pulsations at 09:06:35 UT is clearly seen (the interval is marked with the curved bracket). Antispikes are marked with arrows [48].

studies are very important for the CME theory since they directly confirm the existence of such radiation in nature.

Theoretical studies [54–62] have shown that CME properties and the very possibility of its generation strongly depend on both the background plasma parameters and the fast (overthermal) electron distribution details. The plasma temperature, magnetic field strength, ratio $Y = \omega_{pe}/\omega_{Be}$, characteristic overthermal electron energies, and apparently their distribution are strongly different in the terrestrial magnetosphere and solar corona. So the CME studies in solar flares essentially represent an independent problem.

A cyclotron cone instability was proposed in Refs [74, 75] to explain 100%-polarized rapidly varying microwave bursts [8], and the first semi-quantitative model for the radio spike generation based on CME was analysed in Ref. [76].

The analysis of the cyclotron instability in Ref. [76] is based on estimates of the corresponding linear growth rate at the n-th gyrofrequency harmonic

$$\gamma_n^\sigma(\mathbf{k}) = \int \mathrm{d}^3 p A_n^\sigma(\mathbf{p}, \mathbf{k}) \sigma \left(\omega - \frac{n\omega_{Be}}{\gamma} - k_\parallel v_\parallel \right)$$
$$\times \left(\frac{n\omega_{Be}}{\gamma v_\perp} \frac{\partial}{\partial p_\perp} + k_\parallel \frac{\partial}{\partial p_\parallel} \right) f(\mathbf{p}), \quad (32)$$

where $A_n^\sigma(\mathbf{p}, \mathbf{k})$ is the probability of spontaneous emission of a σ-type wave at the n-th harmonic

$$A_n^\sigma(\mathbf{p}, \mathbf{k}) = \frac{4\pi^2 e^2 c^2 \beta_\perp^2}{\omega n_\sigma \partial(\omega n_\sigma)/\partial\omega(1 + K_\sigma^2)}$$
$$\times \left| \frac{\Gamma_\sigma \sin\theta + (\cos\theta - n_\sigma\beta_\parallel)K_\sigma}{n_\sigma\beta_\perp \sin\theta} J_n(z) + J_n'(z) \right|^2. \quad (33)$$

The indices $\parallel$ and $\perp$ stand for longitudinal and transverse with respect to the magnetic field vector components, respectively, n_σ is the refraction coefficient for the σ-type wave ($\sigma = Z, X, O$),

$$n_\sigma^2$$
$$= 1 - \frac{2v(1-v)}{2(1-v) - u\sin^2\theta + \sigma[u^2\sin^4\theta + 4u(1-v)^2\cos^2\theta]^{1/2}}, \quad (34)$$

where $v = \omega_{pe}^2/\omega^2$, $u = \omega_{Be}^2/\omega^2$, $\sigma = 1$ for O-waves, $\sigma = -1$ for Z- and X-waves, $\beta = v/c$; Γ_σ, K_σ are the wave polarization vector components, θ is the angle between the wave vector $\mathbf{k}$ and the magnetic field $\mathbf{B}$; and $J_n(z), J_n'(z)$ are the Bessel functions and their first derivatives with respect to the argument, respectively, $z = k_\perp p_\perp/\omega_{Be}$. The integration in Eqn (32) is performed along a 'resonance ellipse' [66], which corresponds to the resonance condition

$$\omega - \frac{n\omega_{Be}}{\gamma} - k_\parallel v_\parallel = 0. \quad (35)$$

In paper [76], as well as in the pioneering paper [54], a semi-relativistic approximation is used

$$\gamma^{-1} \approx 1 - \frac{v^2}{2c^2}, \quad (36)$$

and the Bessel functions (as well as their derivatives) are substituted by the leading terms of their expansion in series for small arguments, which also corresponds to the semi-relativistic approximation. The estimates obtained in such an approach allowed the authors of Ref. [76] to conclude that:

a) CME is generated inside a narrow quasi-transverse region with $\theta = 70 - 85°$;

b) the brightness temperature of CME can be as high as $T_b \sim 10^{16} - 10^{17}$ K;

c) the wave absorption at frequency ω in the gyroresonance layers ($\omega_{Be} = \omega/s$, $s = 2, 3$) can be very high (and their optical depth $\tau_s \gg 1$) and impede the escaping radiation generated by the corona; the radiation at the second gyrofrequency harmonic is proposed since it can more easily escape the corona being absorbed only inside the third gyrolayer ($s = 3$);

d) the instability saturation time, which is several inverse growth rates, $t_{sat} \sim 10(\gamma_n^\sigma)^{-1} \sim 10$ ms, is smaller than the formation time of the loss-cone electron distribution by multiple reflections of particles from the magnetic plugs.

On these grounds, an anisotropic distribution with a 'one-side loss cone' (Fig. 4) is assumed to appear as a result of a single reflection from one of the magnetic plugs.

The first CME studies in solar flare conditions, which demonstrated the principal possibility of this mechanism operating in solar corona, have posed many questions. How is the anisotropic electron distribution generating the CME formed? Which modes and gyrofrequency harmonics are mostly enhanced? Is it possible to form a harmonic structure of radio emission? How does the radio emission pass through the gyroresonance layers? These and many other questions (some of which still await a final answer) caused a large number of papers on the CME theory to appear.
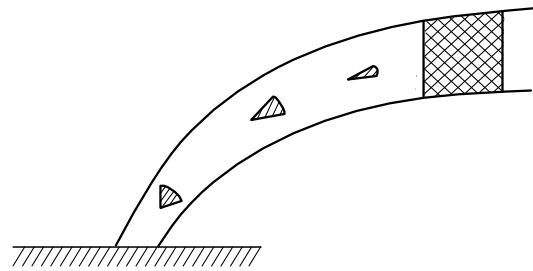
Figure 4. One-side loss-cone distribution resulting from a single reflection from a magnetic mirror [76].

The analysis of linear growth rates of the intrinsic modes generated by the CME has used rather simple model fast electron distributions, which had already been used earlier to model laboratory experiments and the auroral kilometre radiation, but taking into account the specific conditions of the solar corona.

2.3.1 The fast electron distribution far from the acceleration region is assumed in Ref. [77] to have the form

$$f_b(v_\perp, v_\parallel) = \frac{1}{(2\pi a^2)^{3/2}} \frac{n_b}{n_0} \exp\left\{-\frac{(v_\parallel - v_b)^2 + v_\perp^2}{2a^2}\right\}, \quad (37)$$

where v_b is the beam velocity, a is the thermal velocity dispersion inside the beam, and n_b, n_0 are the concentration of particles in the beam and the background plasma, respectively. As a result of reflection from the magnetic plug, a group of reflected particles is formed with the distribution function

$$f_r(v_\perp, v_\parallel) = f_b(v_\perp, v_\parallel)\Theta\left(v_\perp - \frac{v_\parallel}{\sigma}\right), \quad (38)$$

where $\Theta(x)$ is the Heaviside theta-function, and σ is the plug ratio determined as

$$\sigma = \left(\frac{B_{max}}{B_{min}} - 1\right)^{1/2}. \quad (39)$$

The main attention in Ref. [77] is given to the background plasma effects on X- and O-wave growth rates at the lowest gyrofrequency harmonic. In the vacuum limit $Y \to 0$, which can approximately be used in the terrestrial magnetosphere, where $Y \ll 1$ and possibly $Y \sim 0.005$ [78], the X-wave growth rate at the first harmonic γ_{X1} is notably higher than at the second harmonic and than the O-wave growth rates at all harmonics γ_{On}. In the solar corona $Y > 0.1$ and accounting for the finiteness of Y in calculations [77] significantly changes the ratio between the wave growth rates at different harmonics. It is found that for distributions like (38) (in Ref. [77] the following parameters were used: $v_b/c = 1/3$, $v_b/a = 3$, $\sigma = 2$) the value of γ_{X1} is appreciably reduced at $Y \sim 0.1$, and for $Y \sim 0.4$ the X1-mode stops its rapid growth. At $Y > 0.4$ the ordinary wave at the first harmonic has the largest growth rate (O1-wave), although X2-wave growth rates are of the same order of magnitude as γ_{O1}. The presence of a hot background plasma in the corona with $T > 10^6$ K additionally modifies the wave growth rates

$$\gamma^\sigma \to (\gamma^\sigma - \nu_{eff}), \quad (40)$$

where ν_{eff} includes both collisional and gyroresonance damping in the source of emission. The frequency range near gyroharmonics, where cyclotron damping is important, increases with plasma temperature. According to Ref. [77], this damping can strongly reduce, and for $T > 10^7$ K totally suppress, the cyclotron instability at the second harmonic. If CME does appear in the source, it must pass through the gyroresonance absorbing layers ($s = 2, 3$) before reaching an observer on Earth. It is assumed in Ref. [77] that nonthermal electrons with a nonstationary distribution are also present in these layers, which can also reduce the absorption by thermal particles or even additionally enhance the waves inside the gyroresonance layers.

2.3.2 In paper [79] a more detailed study is performed of the proper plasma modes, both high-frequency (X and O) and low-frequency (Z), upperhybrid (UH), whistlers (W)), for various Y and T. The nonthermal electron distribution in [79] was assumed to have the form of the production of long-itudinal and transverse functions

$$f = n_b f_\perp(v_\perp) f_\parallel(v_\parallel), \quad (41)$$

where the DGH-distribution is chosen for the transverse one [80]:

$$f_\perp = \frac{1}{2\pi} \frac{1}{a_\perp^2 l!} \left(\frac{1}{2}\right)^l \left(\frac{v_\perp}{a_\perp}\right)^{2l} \exp\left(-\frac{v_\perp^2}{2a_\perp^2}\right), \quad (42)$$

and the longitudinal distribution is taken in the form of a beam distribution with the thermal dispersion

$$f_\parallel = \frac{1}{(2\pi)^{1/2}} \frac{1}{a_\parallel} \exp\left[-\frac{(v_\parallel - v_b)^2}{2a_\parallel^2}\right], \quad (43)$$

where l is the DGH-distribution index; and $a_\parallel$, $a_\perp$ are the thermal dispersion of the transverse and longitudinal velocity, respectively. The following parameters were taken: $n_b/n_0 = 10^{-2}$, $l = 1$, $\epsilon_\parallel = 3$ keV, $\epsilon_\perp = 30$ keV; the value of Y varied over the range 0.1 to 2.5, and the temperature from 2×10^6 to 10^7 K. It was found that an X1-mode is the most rapidly growing for $Y < 0.35$, for $0.35 < Y < 1$ a O1-mode dominates, and then upperhybrid waves and whistlers prevail. Growth rates γ_{X1} and γ_{O1} are insensitive to the temperature changes within the above limits. X1- and O2-waves never become the most rapidly growing. It was noted that in the interval $0.4 < Y < 0.6$ the instability growth rate for Z-waves γ_Z is comparable with γ_{O1}, but the value of $\gamma_{eff} = (\gamma_Z - \nu_{eff})$ is strongly temperature-dependent.

In paper [81] a similar distribution is analysed with the DGH index $l = 1$ but with $v_b = 0$, $a_\parallel = a_\perp = a$, $a/c = 0.075$, ($\epsilon \sim 3$ keV). The value of Y varied within the limits $0.4 < Y < 2.4$. In this case an O1-mode dominates for $Y < 0.5$, then an O2-mode for $0.5 < Y < 1.5$ and a Z-mode for $Y > 1.5$, i.e. extraordinary wave growth rates at the second harmonic can be maximal, in contrast to the conclusions of paper [79]. We see that the dominating mode hierarchy strongly depends on the energy of emitted electrons even for the same distribution.

2.3.3 The conclusions we arrived at above are confirmed in paper [82] which studies the 'synchrotron' maser emission of particles with an energy of $\epsilon \sim 1$ MeV and a DGH-distribution with $a_\parallel = a_\perp$, $l = 2$ (larger l correspond to a larger anisotropy of the distribution). X-wave growth rates are maximal up to $Y \approx 2.3$, and the Z-mode dominates afterwards. In addition, the growth rates are sufficiently high over a wide range of angles $\theta \sim 70-90°$, the instability bands being appreciably wider than before $\Delta\omega/\omega \sim 10-20$%. The following properties of the model are considered in Ref. [82] as advantageous:

1) at high Y no problem of the radiation passing through the gyroresonance layers arises, although an X-mode remains dominating;

2) the instability bands lie far from the gyroharmonics (due to Doppler effect), so that the enhancement does not depend on the background plasma temperature.

2.3.4 CME generation by the DGH-distribution was also studied using numerical modelling [83] (particle-in-cell). No

background plasma was assumed so that both the wave dispersion and wave instability are determined by hot particles with thermal velocity $v_T = c/14$ and a DGH-distribution index $l = 2$.

At $Y = 0.1$ all the radiated energy comes in an X1-mode, and a small additional extraordinary wave emission at the second harmonic appears for $Y = 0.2$ (at the 1%-level by energy). However, by $Y = 0.3$ longitudinal Berstein modes appear, with their growth rates being close to those of the extraordinary waves at the first and second harmonics:

$$\gamma_{\mathrm{B}} : \gamma_{\mathrm{X1}} : \gamma_{\mathrm{X2}} = 10 : 9 : 8. \qquad (44)$$

At larger Y ($Y = 0.5$) mostly Berstein waves are generated, although a small addition of transverse Z- and X2-waves is present. These results are in general agreement with earlier calculations of the linear growth rates for the DGH-distribution [54, 79, 81, 84], although they have some important restrictions. First of all, they do not include the background (cold) plasma and take into account only transverse wave propagation with respect to the magnetic field direction, although in real conditions a maximum growth rate can be reached with an oblique propagation as well.

2.3.5 The DGH-distribution used in most papers (with the exception of paper [79], in which the longitudinal function corresponds to the electron beam) is an even function. For such a distribution function to form in a coronal magnetic trap, fast electrons must suffer several reflections from magnetic plugs. According to estimates [76], the duration of such a process is significantly larger than the quasilinear relaxation time, so it is rather difficult to form an even fast electron distribution in a coronal arch. For this reason, in Ref. [76] a one-side loss-cone electron distribution formed after single reflection of accelerated electrons from one magnetic plug was proposed as a CME source. Such a distribution can frequently be approximated by a $\sin - N$ — Maxwellian like distribution with an anisotropic angular dependence $f(\mathbf{p}) = F(p) \times \Phi(\theta)$ [85]

$$F(p) = \frac{n_{\mathrm{b}}}{(\pi^{1/2} ma)^3} \exp\left(-\frac{p^2}{m^2 a^2}\right), \qquad (45)$$

$$\Phi(\theta) = a_N \times \begin{cases} \sin^N\left(\dfrac{\theta}{\theta_c}\dfrac{\pi}{2}\right), & \theta < \theta_c, \\ 1, & \theta > \theta_c, \end{cases} \qquad (46)$$

where $a = 2^{1/2} v_{T_{\mathrm{h}}}$, $v_{T_{\mathrm{h}}} = (T_{\mathrm{h}}/m)^{1/2}$ is the thermal velocity of hot electrons, T_{h} is their temperature, N is the distribution exponent, and a_N is the normalization constant such that distribution (46) is unitary.

According to calculations [59, 86–88], this distribution is preferable to DGH for ordinary wave generation. Extra-ordinary waves dominate for $Y < 0.24$ (at the first harmonic, X1) and for $1 < Y < 1.4$ (at the second harmonic, X2). In the intermediate case $0.24 < Y < 1$ an ordinary wave at the first harmonic (O1) dominates. In the cited papers wave damping on thermal particles has also been analysed, and proved unimportant if the hot and cold electron temperatures are strongly different. It was noted in Ref. [87] that a maximum X1-wave amplification can be achieved far from the transverse (and even almost longitudinal) wave propagation directions at $Y \sim 0.25$.

2.3.6 Another popular distribution used for the CME analysis is a hollow-beam (HB) distribution:

$$f(\mathbf{p}) = A \exp\left\{ -\frac{(p_\parallel - p_{\parallel 0})^2}{m^2 a_\parallel^2} - \frac{(p_\perp - p_{\perp 0})^2}{m^2 a_\perp^2} \right\}, \qquad (47)$$

where the normalization constant A is expressed as

$$A = n_{\mathrm{b}} \left\{ \pi^{3/2} a_\perp^2 a_\parallel \left[\exp\left(-\frac{p_{\perp 0}^2}{m^2 a_\perp^2} \right) + \pi^{1/2} \left(\frac{p_{\perp 0}}{a_\perp} \right) \right. \right.$$
$$\left. \left. \times \left(1 + \mathrm{erf}\left(\frac{p_{\perp 0}}{a_\perp} \right) \right) \right] \right\}^{-1}. \qquad (48)$$

This distribution was used in Refs [66, 89, 90] and has not revealed any new CME properties in addition to those discussed above, with the exception of paper [90], in which the case of 'synchrotron' maser emission at high gyrofrequency harmonics near $2\omega_{\mathrm{p}}$ for $Y \gg 1$ was considered. Under such conditions longitudinal waves can be generated, which are barely discussed in [90].

2.3.7 In CME analysis an important role is played by the instability suppression mechanisms which determine both the temporal characteristics of the radiation emitted and the efficiency of transformation of the free energy of fast electrons into transverse wave energy. Paper [91] analyses analytically the quasilinear relaxation due to the electromagnetic waves amplified by the following fast electron model distribution

$$f(\mathbf{v}) = \frac{n_{\mathrm{b}}}{\pi^{3/2} a(a_1^2 - a_2^2)} \exp\left(-\frac{v_\parallel^2}{a^2} \right)$$
$$\times \left[\exp\left(-\frac{v_\perp^2}{a_1^2} \right) - \exp\left(-\frac{v_\perp^2}{a_2^2} \right) \right], \quad a_1^2 > a_2^2. \qquad (49)$$

The free energy of electrons is determined as

$$W_{\mathrm{F}} = \int \mathrm{d}^3 v \left[\frac{mv^2}{2} - \frac{3}{2} T_{\mathrm{eff}} \right] f(\mathbf{v}), \qquad (50)$$

where T_{eff} is the effective temperature.

The free energy value for distribution (49) with account of Eqn (50) is determined as

$$\frac{W_{\mathrm{F}}}{n_{\mathrm{b}} m a_1^2} = \frac{1}{4} \left\{ \left[\frac{a^2}{a_1^2} + 2(1 + R) \right] - 3 \left[\frac{a}{a_1}(1 - R) \right]^{2/3} \right.$$
$$\left. \times \exp\left\{ \frac{2R}{3(1 - R)} \left[2 - R - \sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+1-nR)} \right] \right\} \right\}, \qquad (51)$$

where

$$R = \frac{a_2^2}{a_1^2} > 1. \qquad (52)$$

For typical terrestrial magnetosphere conditions it is found [91] that

$$\frac{W_{\mathrm{F}}}{n_{\mathrm{b}} m a_1^2} \approx 0.1. \qquad (53)$$

Some simplifying assumptions are made in Ref. [91] to analyse the quasilinear equation system [23]. Considering that

the maximum wave amplification in a rarefied plasma occurs in quasitransverse (but still different from transverse) directions, it is thought that

$$k_\| \frac{v_\|}{\omega} \ll 1 \,, \tag{54}$$

which allows one to neglect derivatives with respect to $v_\|$ in the equation for $f(v)$ and to cast it into the form

$$\frac{\partial f}{\partial t} = \frac{1}{v_\perp} \frac{\partial}{\partial v_\perp} D v_\perp \frac{\partial f}{\partial v_\perp} \,, \tag{55}$$

where

$$D = \frac{4\pi^2 e^2}{m^2} \int \mathrm{d}^3 k \, W_\sigma(t) \delta \left[\omega - \omega_{\mathrm{Be}} \left(1 - \frac{v^2}{2c^2} \right) - k_\| v_\| \right] \,, \tag{56}$$

and $W_\sigma(t)$ is the energy density of the amplified waves (σ modes) determined by the equation

$$\frac{\partial W_\sigma}{\partial t} = 2\gamma(\mathbf{k}, t) W_\sigma \,, \tag{57}$$

in which the growth rate $\gamma(\mathbf{k}, t)$ is related through Eqn (32) with the electron distribution function $f(\mathbf{v})$.

In paper [91] it is also assumed that

$$\frac{D}{v_\perp} \gg \frac{\partial D}{\partial v_\perp} \,, \tag{58}$$

since the dependence on $v_\perp$ enters D only through a small correction to the argument in the δ-function. Then Eqn (55) takes the form

$$\frac{\partial f}{\partial t} = D \frac{\partial^2 f}{\partial v_\perp^2} + \frac{D}{v_\perp} \frac{\partial f}{\partial v_\perp} \,, \tag{59}$$

with the solution

$$f(\mathbf{v}, t) = n_{\mathrm{b}} \frac{\exp(-v_\|^2/a^2)}{\pi^{3/2} a (a_1^2 - a_2^2)} \left[\frac{1}{1 + \tau/a_1^2} \exp\left(-\frac{v_\perp^2}{a_1^2 + \tau} \right) \right. \\ \left. - \frac{1}{1 + \tau/a_2^2} \exp\left(-\frac{v_\perp^2}{a_2^2 + \tau} \right) \right] \,, \tag{60}$$

where

$$\tau = 4 \int_0^t \mathrm{d}t' D(t', \mathbf{v}) \tag{61}$$

and

$$W_\sigma = W_\sigma(0) \exp\left[2 \int_0^t \mathrm{d}t' \gamma(\mathbf{k}, t')\right] \,. \tag{62}$$

According to Ref. [91], the wave amplification stops when the derivative $\partial f/\partial v_\perp$ ceases to be positive for all $v_\perp$. Last of all this occurs at $v_\perp = 0$, and since

$$\left. \frac{\partial f}{\partial v_\perp^2} \right|_{v_\perp = 0} \propto \left\{ \frac{1}{a_1^2 (1 + \tau/a_1^2)^2} - \frac{1}{a_2^2 (1 + \tau/a_2^2)^2} \right\} \,, \tag{63}$$

the maser saturation occurs at $\tau = a_1 a_2$. With account of Eqns (61), (63) this condition can be approximately rewritten in the form

$$\frac{\pi \omega_{\mathrm{pe}}}{n m a_1 a_2} \int_0^{t_s} \mathrm{d}t \int \mathrm{d}^3 k \, W_\sigma(t) I = 1 \,, \tag{64}$$

where t_{s} is the saturation time,

$$I \equiv \int \mathrm{d}^3 v f_0(\mathbf{v}) \delta \left[\omega - \omega_{\mathrm{Be}} \left(1 - \frac{v^2}{2c^2} \right) - k_\| v_\| \right] \,. \tag{65}$$

In the subsequent estimates of CME efficiency in Ref. [91], the authors neglect the quadratic correction $v^2/2c^2$ to the δ-function argument (note that this can be done if the most amplification occurs not strictly in the transverse direction, $k_\| \neq 0$). Assuming that W_s forms a narrow peak with the characteristic $k_\| = k_{\|0}$ so that

$$\frac{(\omega - \omega_{\mathrm{Be}})^2}{k_{\|0} a^2} < 1 \,, \tag{66}$$

one can obtain

$$\frac{W_s}{n_{\mathrm{b}} m a_1^2} = \frac{2}{\pi} \frac{a_2}{a_1} \frac{\omega_{\mathrm{Be}}^2}{\omega_{\mathrm{pe}}^2} \left(\frac{k_{\|0} a}{\omega_{\mathrm{Be}}} \right) \left(\frac{\gamma_{\sigma \max}}{\omega_{\mathrm{Be}}} \right) \left(\frac{n_0}{n_{\mathrm{b}}} \right) \,, \tag{67}$$

where $\gamma_{\sigma \max}$ is the maximum s-wave growth rate, and $W_s = \int \mathrm{d}^3 k W(t)$ is the σ-wave energy density in saturation.

Estimates [91] demonstrate that W_s can be as high as a few percent of the free energy W_{F} (60). The wave amplification is then very significant

$$\frac{W(t = t_s)}{W(t = 0)} \sim 10^{13} \,, \tag{68}$$

with the saturation time being of order

$$t_s \approx \frac{15}{\gamma_{\sigma \max}} \,. \tag{69}$$

In paper [92] it is noted that in a sufficiently rarefied plasma, where $Y < v_{\mathrm{Te}}/c$, relativistic corrections are important not only for fast electrons, but for the background plasma as well. It turned out that the maximum enhancement of transverse waves occurs in the direction strictly perpendicular to the magnetic field and the growth rates can be several times larger than those obtained without taking into account the relativistic corrections. In Ref. [92], a DGH-distribution (41), (42) with different indices was taken as the initial one. Since condition (54) is obligatorily fulfilled during the maximum wave amplification across the magnetic field, the fast electron distribution function satisfies the same equation (55) as in Ref. [91]. Interestingly, its solution for a DGH-distribution with index l can be derived from the solution for a double-Maxwellian distribution (60) with $a_2 = 0$ [91] by differentiating l times with respect to the parameter (a_1^{-2}):

$$f(\mathbf{v}, t, l) = (-1)^l (l! a_1^{2l+2})^{-1} \frac{\mathrm{d}^l}{\mathrm{d}(a_1^{-2})^l} \left[a_1^2 f(\mathbf{v}, t, l = 0) \right] \,. \tag{70}$$

At the beginning the DGH-distribution has a positive slope for $v_\perp^2 < l a_1^2$, and a sufficient condition for stability is, as before, the vanishing of the derivative with respect to $v_\perp$ at $v_\perp = 0$. The CME efficiency estimates are made in Ref. [92] in a way similar to Ref. [91] (see above) with one important exception. Since in this case $k_\| = 0$, when evaluating I (65) one needs to take the quadratic relativistic correction to the δ-function argument (we assume below that $a_1 = a$):

$$\omega_{\mathrm{Be}} I = \frac{2^l}{\pi^{1/2} (2l+1)!!} \frac{4c^2}{a^2} x_0^{2l+1} e^{-x_0^2} \,, \tag{71}$$

where

$$x_0 = 2\frac{c^2}{a^2}\left(1 - \frac{\omega}{\omega_{Be}}\right). \tag{72}$$

As a result, the transverse wave energy density at saturation related to the electron energy density due to electron transverse motion is

$$\varepsilon = \frac{W_s}{(n_b m l a^2/2)} = \frac{2}{\pi}(\omega_{Be} I)^{-1}\frac{\omega_{Be}^2}{\omega_{pe}^2}\frac{\gamma_{\sigma\,max}}{\omega_{Be}}. \tag{73}$$

Thus the efficiency ε is proportional to the linear growth rate in Y^{-2}. With decreasing Y the growth rate γ_{max} decreases [92], but the efficiency ϵ increases due to the factor Y^{-2}.

This conclusion agrees with the numerical modelling results obtained in Ref. [92] for a DGH-distribution with $l = 1, 2, 4$. As seen from Fig. 5a, at smaller values of the linear growth rate a higher saturation energy density is reached. The result of the evolution of the transverse part of the distribution function can be seen in Fig. 5b. The efficiency ε varies in Ref. [92] within the limits

$$1\% < \varepsilon < 7\%. \tag{74}$$

As already noted [83], the numerical modelling was performed for values $Y > 0.1$ higher than in Ref. [92]. The background plasma was ignored, so that the plasma frequency ω_{pe} and the wave dispersion were determined by the fastest electrons with a DGH-distribution. Although the results of the modelling depend both on Y and the presence or absence of a continuous particle injection into the emission source, the efficiency ε is about a few percent, as before.

An analytical consideration of the quasilinear relaxation, as described above, can be done making a number of approximations and assumptions, not all of which are obviously valid in real conditions. The CME modelling also has shortcomings, in particular, it requires significant parameter scaling when confronting the corresponding results with observations.

In Ref. [88] a detailed study of CME saturation is performed based on the fully relativistic numerical calculations of the quasilinear equation system for the wave energy density and particle distribution function over a wide range of parameters. The initial distribution function in Ref. [88] was taken in the form of a Maxwellian distribution multiplied by the $\sin - N$ function (46). The background plasma was described by an isotropic Maxwellian function with a lower temperature.

In Ref. [88] the following set of 'standard' parameters was used: $Y = 0.1$, the dominating mode is X1, the background plasma density $n_0 = 1.25 \times \times 10^8$ cm^{-3} ($\omega_{pe}/2\pi = 100$ MHz), its temperature $T_0 = 10^6$ K, the hot particle density $n_b = 1.25 \times 10^6$ cm^{-3}, their temperature $T_h = 10^8$ K, the plug ratio $\sigma = 4$ (the loss cone angle $\theta_c = 30°$), the angular distribution parameter $N = 6$, and the initial radiowave level $N(k) = 2.1 \times 10^{15}$, which corresponds to the brightness temperature $T_b = 10^{14}$ K.

The diffusion rate of the fast electrons in the momentum space is mainly determined by the product of the wave energy density in the resonance region by the second derivative of the distribution function with respect to the pitch-angle. So the diffusion is small both at the beginning (due to the small wave energy density) and at the end of the process (due to the small
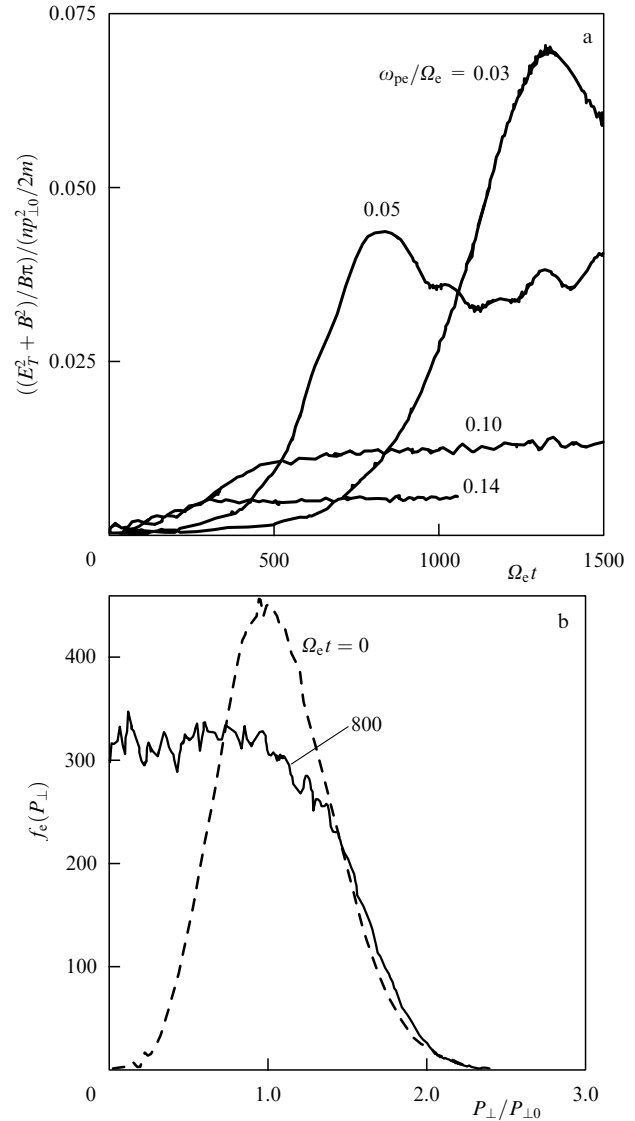
Figure 5. (a) Total electromagnetic energy density (normalized to the initial density of the perpendicular component of the electron kinetic energy) in one-dimensional calculations at $\theta = 90°$ for different ratios ω_{pe}/ω_{Be}. The DGH parameters: $c^2/(l\alpha^2) = 50$ and $l = 2$ [92]. (b) The numerical calculations of the unstable distribution relaxation due to transverse waves [92]. The dashed line shows the initial distribution, the solid line the saturated distribution.

value of the derivative). In these calculations, the diffusion rate reached a maximum by the time $t \sim 10/\gamma_{max}$. The loss cone then starts filling with particles, the wave amplification growth rates decrease and the exponential radiation growth slows down (Fig. 6a). At the intermediate stages of the quasilinear relaxation process, flux-type singularities appear in the particle distribution, which could lead to beam-like instabilities (not accounted for in Ref. [88]).

The transformation of the electron kinetic energy into wave energy mostly occurs on a time scale $t = (10-20)/\gamma_{max}$, the growth rate decreasing by three times by $t \sim 20/\gamma_{max}$. Particle relaxation is practically reached by the moment $t \sim 50/\gamma_{max}$ when the loss cone is filled with particles. The spectral half-width of the generated wave packet changes with time from 2.3% at the beginning of the diffusion to 3.5% at saturation. The efficiency of the kinetic into electromagnetic
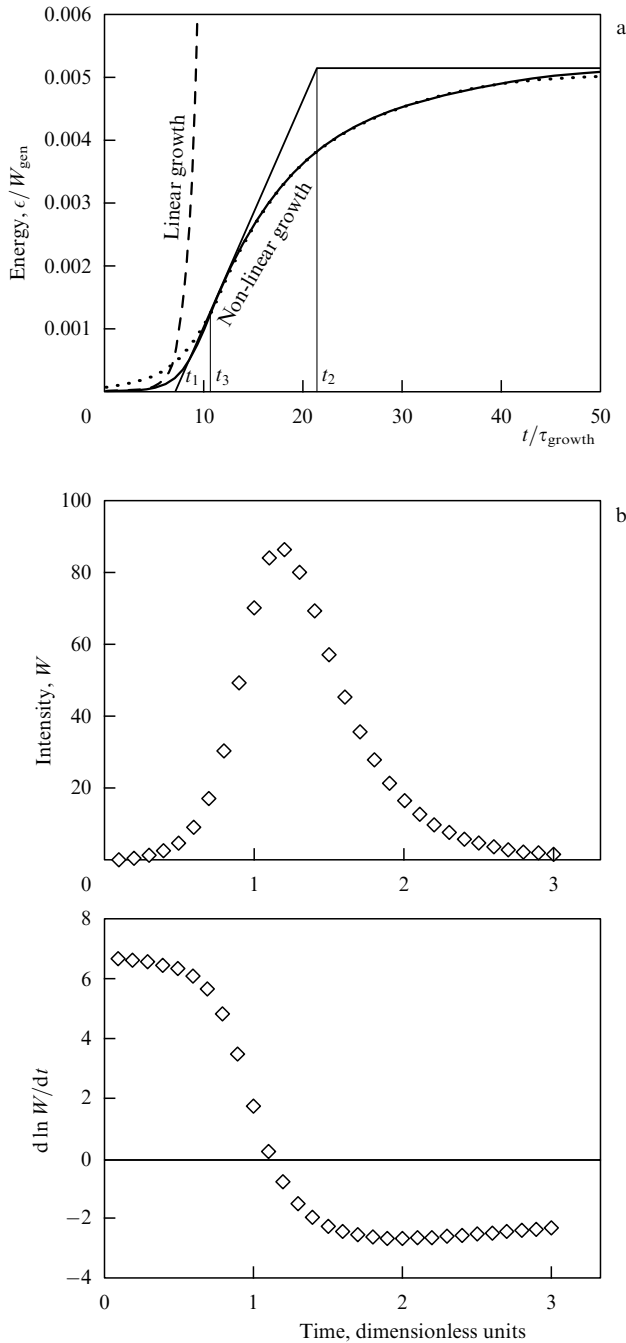
Figure 6. (a) X1-wave energy density as a function of time taking account of the quasi-linear relaxation [88]. The dashed line indicates the exponential energy growth in the linear approximation. The actual growth proceeds more slowly. (b) The time profile and the time logarithmic derivative of the energy density of the dominating (lowerhybrid) wave taking account of the quasi-linear relaxation of an initially unstable anisotropic fast electron distribution.

energy conversion (determined as $\varepsilon_{kin} = [E_{kin}(t \to \infty) - E_{kin}(t = 0)]/E_{kin}(t = 0)$) is about 0.5% in calculations with standard parameters.

In Ref. [88] the dependence of the quasilinear relaxation on the main parameters of the problem by successive varying of each of the 'standard' parameters was studied.

The ratio n_b/n_0 varied from 0.001 to 0.1 and the dependence of the saturation time t_s, brightness temperature T_T (at the beginning of diffusion) and T_s (in the saturation),

and the efficiency ϵ_s on n_b/n_0 was studied. The quantity $t_s\gamma_{max}$ lying in the range $12-16$ weakly depends on n_b/n_0, the efficiency ϵ is about $0.47-0.54\%$ and almost constant. The brightness temperatures T_T and T_s increase with n_b/n_0 by slightly more rapid than linear power laws (in the zero approximation one can expect linear growth).

When the fast electron temperature drops to $T_h/T_0 = 10-20$ the growth rate strongly decreases due to thermal cyclotron damping, and the efficiency ϵ_s sharply falls as well. The optimal for CME is $T_h/T_0 \sim 200$; for large values the growth rate and efficiency decrease slowly, the growth of brightness temperatures T_T and T_s with (T_h/T_0) slows down.

The derivative of distribution (46) with respect to angle is mainly determined by the distribution exponent N

$$\frac{\partial f}{\partial \theta} \propto N^{1/2}, \qquad N \gg 1. \tag{75}$$

Correspondingly, $\gamma_{max}t_s$, ϵ_s, T_T, T_s depend on N by power-laws with a degree less than 0.5.

Different parameters of the problem depend differently on the critical angle θ_c (determined by the plug relation). For example, the growth rate γ_{max} reaches a maximum at $\theta_c \sim 30°$ and the efficiency ϵ_s increases with θ_c and reaches 1.5% at $\theta_c = 50°$. The saturation time $t_s\gamma_{max}$ slowly decreases with θ_c while the brightness temperatures are almost constant for $20° < \theta_c < 50°$.

It is important to realize how the results of calculations depend on Y, as the dominating emission mode changes with increasing Y. Despite the fact that the maximum growth rate γ_{max} decreases by three orders of magnitude on changing Y from 0.1 to 1.4, the efficiency ϵ_s and the saturation time $t_s\gamma_{max}$ remain practically constant. The brightness temperatures are also practically constant for each of the dominating modes, but they are higher in the regime when the O1-mode grows more rapidly than others. Simple approximate formulas are derived in Ref. [88] for all the dependences discussed qualitatively above.

In the all cases described above, only the most rapidly growing mode of electromagnetic radiation, which determined the total relaxation process, was taken into account. It is frequently thought for this reason that in order to evaluate the level of other modes one can use the relation

$$W_s^{(\sigma)} = W_0 \exp\left\{\beta \frac{\gamma_s^{(\sigma)}}{\gamma_{max}}\right\}, \quad \beta = \ln\frac{W_s^{(\sigma)}}{W_0}, \tag{76}$$

i.e. if the dominating mode growth rate σ is several times smaller than the maximum growth rate, such waves will not be significantly enhanced.

2.3.8 Nevertheless, different groups of fast particles are responsible for different modes of amplification. So far we have considered 'quasimaxwellian' fast particle distributions in which the number of particles with velocities exceeding thermal is exponentially small. Usually in astrophysical situations the accelerated particles have not thermal but much more energy extended power-law spectra. For this reason we turn to a model of CME generated by a power-law momentum distribution with the loss cone [93−95]

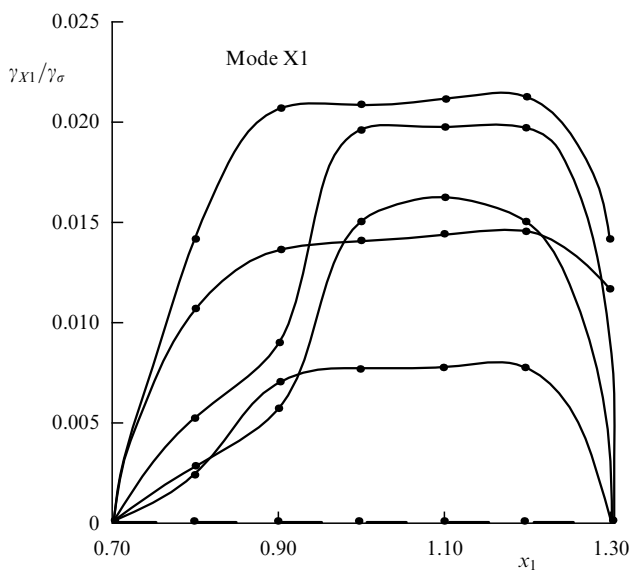$$f(\mathbf{p}) = \frac{\xi - 3}{2\pi p_0^3}\left(\frac{p_0}{p}\right)^{\xi} F(\mu), \quad p_0 < p < p_1, \tag{77}$$

$$F(\mu) = \frac{1}{N}\exp\left(-\frac{\mu^2}{\mu_0^2}\right), \quad \int_{-1}^{1} F(\mu)\,d\mu = 1. \tag{78}$$
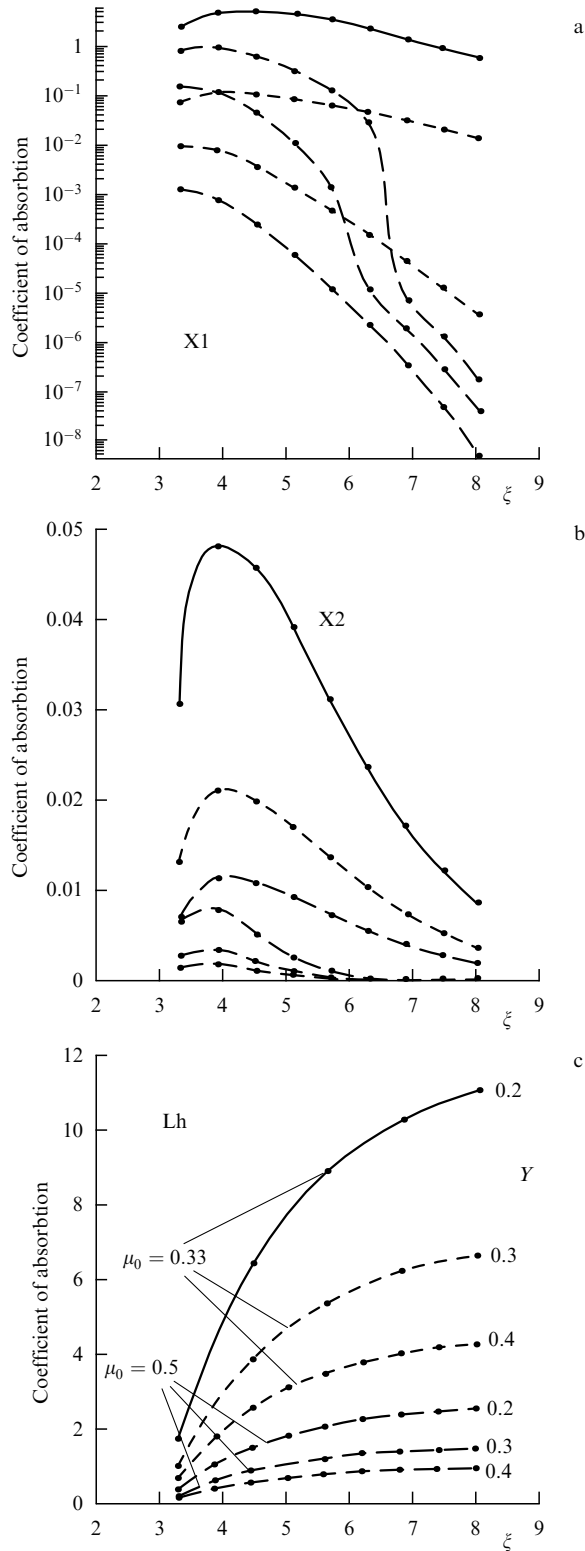
The dependence of time growth rates of different proper plasma modes on the parameter Y and fast electron distribution parameters was studied in paper [93] and the corresponding dependences for spatial growth rates (negative reabsorption coefficients) were analysed in paper [94]. It was shown that neglecting thermal wave damping an X1-wave would dominate for $Y < 0.2-0.3$ [depending on the parameters of distribution (77), (78)], a lowerhybrid wave would dominate for $0.2-0.3 < Y < 0.5-0.6$, and a Z-wave would dominate for $Y > 0.5-0.6$. However thermal damping can fully suppress the lowerhybrid wave instability, and in this case the growth rates of O1- and X2-waves turn out to be maximum and comparable to each other. Note that for $Y = 0.2-0.3$ maximum X1-wave amplification occurs in the longitudinal (and not in the quasi-transverse) direction relative to the magnetic field. This property was discovered also in Ref. [88] for a $\sin-N$-like distribution.

The dependence of the transverse wave growth rates on the electron energy spectrum expansion (the quantity $x_1 = p_{max}/mc$) is very important. It turns out that for sufficiently large values of x_1 corresponding to $E_{kin} > 0.5$ MeV, the growth rates decrease, and the cyclotron instability is suppressed (Fig. 7). This occurs because at each frequency several terms of the series in n start to contribute to the growth rate

$$\gamma_\sigma^{QL} = \frac{\pi^2 \omega_{pe}^2 (n_b/n_0)}{\omega n_\sigma^2 [\partial(\omega n_\sigma)/\partial\omega](1 + K_\sigma^2)|\eta|} \sum_{n_{min}}^{n_{max}} \int_{p_0}^{p_1} \frac{p^2 \, dp}{\gamma\beta} (1 - \mu_*^2)$$

$$\times \, \Theta(1 - \mu_*^2) \left[J_n' + \frac{\Gamma_\sigma\sqrt{1 - \eta^2} + K_\sigma(\eta - \beta\mu n_\sigma)}{n_\sigma\beta\sqrt{(1 - \mu^2)(1 - \eta^2)}} J_n(z) \right]^2$$

$$\times \left[p\frac{\partial}{\partial p} + (n_\sigma\beta\eta - \mu_*)\frac{\partial}{\partial\mu} \right] f(p,\mu) \bigg|_{\mu = \mu_*}, \qquad (79)$$
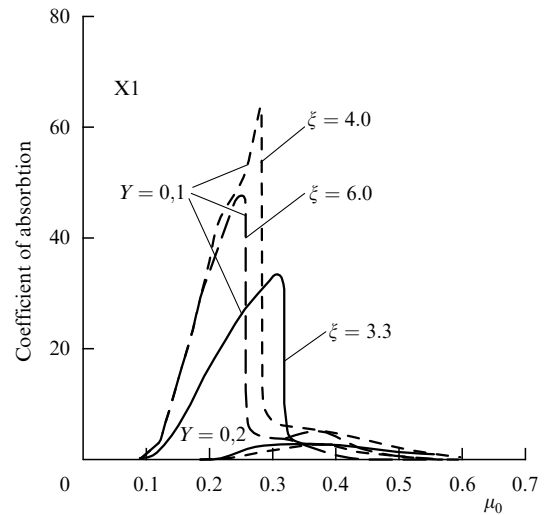
where $\mu_* = (\gamma - n/s)/\gamma\beta n_\sigma\eta$, $s = \omega/\omega_{Be}$, $z = (\omega/\omega_{Be}) \times \gamma n_\sigma\beta\sqrt{(1 - \mu^2)(1 - \eta^2)}$, K_σ, are Γ_σ the polarization vector components, due to a significant Doppler frequency shift for relativistic particles, with the negative contribution

Figure 7. Growth rate of an extraordinary wave (the fundamental harmonic) as a function of the parameter $x_1 = p_{max}/mc$ [93]. Different curves correspond to different parameters of the initial distribution (77), (78).

of some terms being compensated by the positive contribution of others.

From here an experimentally tested conclusion follows immediately: the CME cannot be generated simultaneously with soft gamma-ray emission, which is produced by electrons with $E_{kin} > 0.5$ MeV. However, if the number of electrons with $E_{kin} > 0.5$ MeV is sufficiently small, the harder the energy spectrum (the smaller the power-law index ξ), the larger the transverse wave growth rates (Fig. 8). This is explained by the negative contribution to the growth rate (which is related to the derivative from the distribution function with respect to the absolute value of momentum) being proportional to $(-\xi)$. Thus the CME conditions are more easily met for harder electron spectra.

Distribution (78) depends on one parameter of the anisotropy μ_0. The dependence of the X1-wave growth rates on μ_0 shows a maximum (Fig. 9), with higher μ_0 corresponding to a higher electron isotropy. This means that two forms of the electromagnetic emission pulse are possible in principle. If at the beginning of the CME process the system is at the ascending branch of the plot (see Fig. 9), an over-exponential emission growth is possible. Indeed, approximating the ascending part of the curve $\gamma_{X1}(\mu_0)$ by a linear function

$$\gamma_{X1} \simeq b_0 + b_1\mu_0, \qquad (80)$$

and assuming the isotropization process, which is related to quasilinear relaxation, in the form of a (model) increase of μ_0 with time

$$\mu_0(t) = a_0 + a_1 t + \dots, \qquad (81)$$

we get

$$I \propto \exp\left\{ \int_0^t \gamma_{X1}(\mu_0(t')) \, dt' \right\} = c_0 \exp\left\{ c_1(t - t_0)^2 \right\}. \quad (82)$$

In contrast, on the descending branch of the plot the similar considerations lead to a Gaussian form of the CME growth phase

$$I \propto \exp\left\{ -c_2(t - t_0)^2 \right\}. \qquad (83)$$

Note that numerical analysis of the quasilinear relaxation [88] on the transverse waves also leads to a non-exponential phase of the CME growth (although the decreasing phase is exponential). It is natural to expect that the non-exponentiality of the energy increase of amplifying waves will always appear during the relaxation phase of the originally unstable distribution of the emitting particles due to the corresponding wave growth rate decrease.

Preliminary results of a numerical analysis of the quasilinear relaxation of distribution (77), (78) on the dominating lowerhybrid wave [95a] confirm this statement. Figure 6b shows the time dependence of the energy density of generated lowerhybrid waves (at the top) and the corresponding logarithmic derivative (at the bottom). It is seen that a significant part of the pulse (starting from 20% of the maximum value in the growing phase up to 80% at the decaying phase) is described by a Gaussian dependence of the signal on time.

Note some other important results obtained in [95a]. The quasilinear relaxation of a power-law (over momentum) distribution does not lead to the total isotropization of the

Figure 8. Maximal amplification coefficient versus the electron energy spectral slope ξ for (a) an X1-mode (extraordinary wave, first harmonic), (b) an X2-mode (extraordinary wave, second harmonic), (c) an Lh-mode (lowerhybrid wave) [94].

Figure 9. Maximal amplification coefficient versus the anisotropy degree μ_0 of the electron spectrum for an X1-mode (extraordinary wave, first harmonic) [94].

Also important is that such a modified particle distribution remains unstable relative to the transverse wave generation. The characteristic saturation time of the pulse is of order $t \sim 50/\gamma_{\mathrm{max}}$, in contrast to estimate (69).

Paper [95] is devoted to the analysis of harmonic instability bands of CME. For mildly relativistic particles with distribution (77), (78) the enhancement can occur in some frequency bands shifted relative to the gyrofrequency harmonics due to the 'interplay' between two factors, the Doppler shift and the resonance conditions. As a result, the ratio of frequencies at which amplification occurs is not, in general, equal to the ratio of small integer numbers (Fig. 10).

It is more difficult to provide the wave amplification for the extended electron energy spectrum (under other equal conditions) than for quasimaxwellian distributions. For this reason the instability regions (see Fig. 10) prove to be narrower than, for example, in paper [82].

The model predictions for a CME generated by a power-law electron distribution are as follows.

1. The radiation is generated in separate narrow regions within the frequency range $\omega = (1-2.5)\omega_{\mathrm{Be}}$.

2. The direction of the preferential wave amplification varies from longitudinal to quasi-transverse depending on the parameter Y (the longitudinal direction appears for $Y = 0.2-0.3$).

3. The CME is necessarily accompanied by hard X-ray emission, while soft gamma-ray emission is absent or has a low intensity.

4. The probability of meeting the cyclotron instability conditions is higher for harder electron spectra for energies $E_{\mathrm{kin}} \sim 10-200$ keV.

5. The typical emission pulse growing phase is described by a Gaussian law (83).

6. The ratio of harmonic emission bands is generally not an integer.

Unfortunately, detailed calculations of the saturation of CME generated by electrons with power-law spectra have not yet been done. One may assume that if the quasilinear particle relaxation is determined by a dominating transverse wave, the corresponding results must be close to those

initial distribution. This distribution is modified (isotropized) only such as to provide stabilization (and the subsequent decay) of the initially unstable lowerhybrid wave. More than one pulse of these waves can be generated during this process.
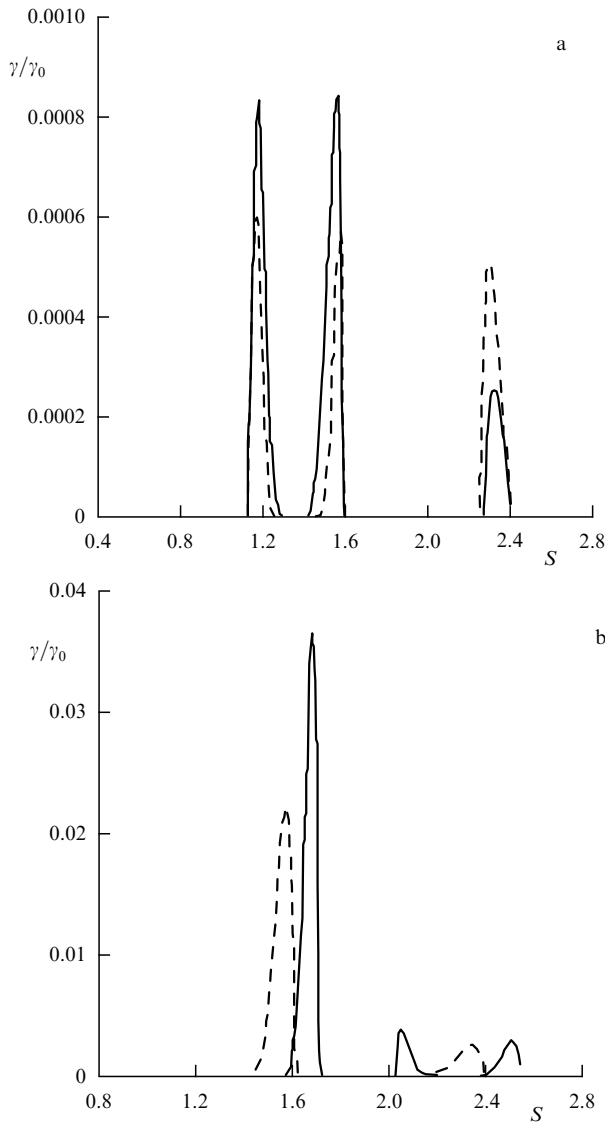
Figure 10. (a) Growth rate of extraordinary waves versus frequency [95]. Parameters: $Y = 0.3$, $\eta = 0.8$, $\mu_0(|\mu_0|)$, $\mu_0(-|\mu|) = 1.42\mu_0(|\mu|)$, $\xi = 4.0$ the frequency ratio 3:4:6. (b) The growth rate of extraordinary waves versus frequency [95]. Parameters: $Y = 0.3$, $\eta = 0.45$, $\mu_0 = 0.3$, $\xi = 4.0$, $x_0 = 0.4$, $x_1 = 0.5$ (the solid line), the frequency ratio 4:5:6.15; $Y = 0.3$, $\eta = 0.5$, $\mu_0 = 0.3$, $\xi = 3.3$, $x_0 = 0.4$, $x_1 = 0.65$ (the dashed line), the frequency ratio 3:4:4.5.

obtained in Ref. [88] (as described above). Nevertheless, a detailed study of the quasilongitudinal wave (when they are formed) effects on the spectral particle evolution, as well as of nonlinear wave processes which can be very important [27], is still to be made.

When interpreting the solar corona radio emission using the CME theory, many questions arise about the geometrical model of the source, the formation of unstable distributions, the propagation of the radio emission in solar corona, etc. We return to these points below, after discussion of the observational properties of millisecond radio spikes.

2.4. Transition maser emission

The transition radiation is generated in any medium with refraction coefficients changing in space or in time. As shown in papers [96, 97], the spectrum of this radiation in a plasma

with density inhomogeneities has a large narrow peak near the plasma frequency (resonance transition radiation, RTR), the width and intensity of which depends on the magnetic field strength. The source optical depth at these frequencies can be notably larger than unity. If the electron distribution strongly depends on the pitch-angle for small pitch-angles, the RTR reabsorption coefficient can be negative [98]. In this case maser transition radiation will occur. Its properties are different from CME as follows.

1. The waves are mostly amplified along the magnetic field direction.

2. Only radiation near the local plasma frequency is emitted, with no harmonic structure.

3. The maser RTR properties are insensitive to the electron energy spectral features for energies $E_{\mathrm{kin}} > 0.5$ MeV, so it may or may not be accompanied by soft gamma-ray emission.

3. Observational data on radio spikes

The most general properties of radio spikes have already been listed in the Introduction. As we mentioned, the short duration, high polarization, high brightness temperature, and narrow frequency band of pulse radio emission can be explained by several models and mechanisms of radio emission. However, various models and mechanisms make different detailed predictions. In this section we focus on discussing the observational data obtained so far and analyzing their possible interpretations by the different theoretical models described in Section 2.

3.1 Temporal properties of radio spikes

As mentioned above, radio spikes have the shortest time-scales among all known types of solar electromagnetic radiation. It was noted quite long ago [13] that the duration of spikes t depends on their frequency. In particular, $t \sim 50-100$ ms near $f \sim 250$ MHz, $t \sim 10-50$ ms at $f \sim 500$ MHz, $t < 10$ ms at $f \sim 3$ GHz.

A detailed study of the temporal properties of the radio spikes was done in Ref. [99]. In this paper, both the dependences of the time constants on frequency and the time profile were analysed.

3.1.1 To describe the temporal properties of the radio emission pulses, different quantities can be used: the total duration d (determined, for example, by the level $1/e$, d_e), the characteristic e-fold rising time of the signal τ_{re}, the characteristic decay time τ_{de}. All these quantities decrease with increasing frequency but with different power laws.

The decay time τ_{de} depends on f^{-1} almost linearly:

$$\tau_{de}(f) = 0.0165 \left(\frac{f}{661 \text{ MHz}} \right)^{-1.06 \pm 0.06} \text{s} , \quad (84)$$

while the total duration decreases faster

$$d_e(f) = 0.0348 \left(\frac{f}{661 \text{ MHz}} \right)^{-1.34 \pm 0.13} \text{s} . \quad (85)$$

Interestingly, in Ref. [99] no correlation was revealed between τ_{de} and d_e at a given frequency, so the conclusion was made that the decay mechanism is independent of the excitation one. The observed signal $S(t)$ is assumed in Ref. [99] to result from the superposition of a source function $f(t)$ with a permanent wave decay mechanism with the time

constant τ (the decay function $g(t) = \exp(-t/\tau)$) so that

$$S(t) = \frac{1}{\tau} \int_{-\infty}^{t} \exp\left(-\frac{t-t'}{\tau}\right) f(t') \, dt' . \qquad (86)$$

In this case the source function $f(t)$ has its own proper duration D obeying a power-law

$$D(f) = 0.0325 \left(\frac{f}{661 \text{ MHz}}\right)^{-1.49 \pm 0.17} \text{s} . \qquad (87)$$

If the spikes are generated by a plasma mechanism, the most probable reason for the wave decay is electron-ion collisions. In Ref. [99] from the constant τ_{de} and frequency f for different radio spikes it is found that the background plasma temperature varies in the range $T = (2-4) \times 10^6$ K and a correlation $T \propto n_0^{1/3}$ exists, i.e. the higher the frequency of observations, the higher the plasma temperature.

In the CME model, the spike duration is inversely proportional to the growth rate of the amplified mode, $d_e \propto \gamma_{on}^{-1}$, and the emission frequency is proportional to the magnetic field strength, $f \propto B$ (see Section 2.3). The dependences $\gamma(Y)$ calculated in Refs [93, 94] are different for different emission modes. This means that if Y correlates with the magnetic field strength B, $Y(B)$, the spike duration as a function of frequency is different for spikes generated at different gyrofrequency harmonics (and also for X- and O-modes). If Y increases with B, the probability of wave emission near the second harmonic must increase with frequency, and vice versa. So it is very important to try to establish the dependence $d_e(f)$ for different harmonic bands (in events showing the harmonic structure) both to determine the number of harmonics at which the spikes are generated and to find the dependence $Y(B)$ in the spike sources. It is equally important to investigate $d_e(f)$ over a wider frequency range (up to 8 GHz) and with the use of fuller observational data than in Fig. 11a. A break in such a dependence would be indicative of changing the dominating spike emission harmonics with frequency.

3.1.2 The time profile of an individual spike usually has a simple form easily separated into the rising and decaying phases. Observations with millisecond time resolution [99, 100] did not reveal statistically significant evidence of variability on a time scale shorter than the characteristic pulse duration (tens of milliseconds). The dynamics of the intensity increases and decreases of the studied spikes does not follow a simple exponential law, which should be expected from the linear theory of plasma or cyclotron instabilities. The phase of the radio emission decay in an individual spike is described by an exponential law with the time constant (84).

The law of the flux density change at the rising phase is slower than exponential growth and is close to Gaussian $F \propto \exp[-(t - t_{max})^2 / \sigma^2]$. The physical sense of the Gaussian time profile of the spikes becomes clearer considering its direct relation to the linearity of the logarithmic derivative of the flux. Its time dependence (Fig. 11 b,c) is linear for almost all spikes studied. The linearity holds for a significant part of the spike (from the moment when $F = 0.2 F_{max}$ at the flux growth phase up to the moment when $F = 0.8 F_{max}$ at the decay phase). At the same time, the slope of the dependence can be different for different spikes apparently reflecting the difference in physical conditions in the corresponding radio
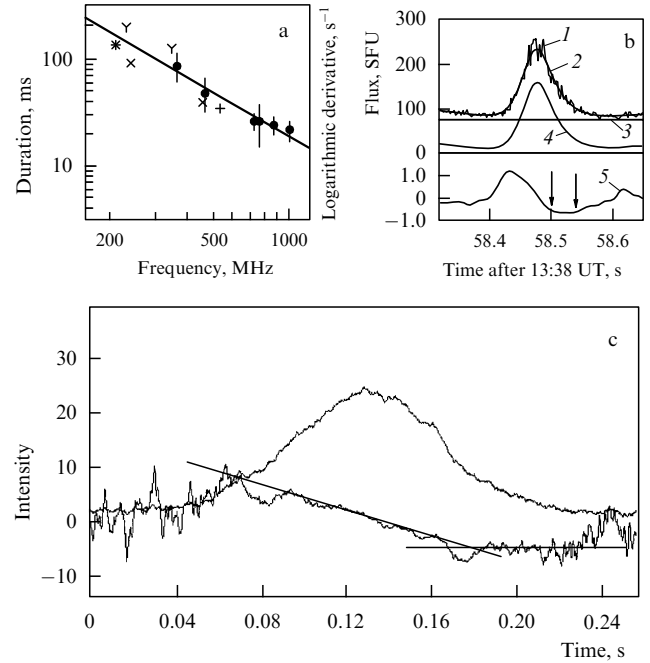
Figure 11. Spike duration as a function of frequency according to different observations [99] (a). The temporal evolution of the flux density from a single spike (b, c). The event of 4.06.82 [99] (b): 1 — the spike time profile record; 2 — the spline-interpolated time profile; 3 — the smoothed background decreased by 10 sfu; 4 — the spike time profile after subtraction of the background; 5 — the logarithmic derivative of 4. The event of 26.10.94 [100] (c). The logarithmic derivative of the flux density linearly decreases with time over most of the spike.

sources, although a detailed dependence of σ on frequency and other parameters has not been studied as yet. The linear dependence of the logarithmic derivative is in a good agreement with the CME mechanism considering the effect of quasilinear relaxation of an anisotropic energetic electron distribution (83). A numerical analysis of CME [88] also suggests a non-exponential (slower) growth phase and exponential decay. In the framework of the CME mechanism, the observed temporal characteristics of the logarithmic derivative can be used to test physical parameters in the spike generation region [100].

Apparently, a similar influence of the quasilinear relaxation on the radiation time profile should be expected for plasma generation mechanisms as well. Note, however, that the Gaussian pulse form cannot be reproduced in some models. For example, the time profile of a separate spike generated by the nonlinear plasma mechanism (see Section 2.2) strongly differs from the observed spike forms. Calculations show that in the quasiperiodic regime (pulsation regime) both the pulse growth and decay phase have an exponential time dependence, and in the non-periodical regime no universal law has been found.

3.1.3 Clusters of radio spikes were analysed in Ref. [101a] using the methods of nonlinear dynamics. The fractal dimension of the emission source was found to be very high (almost ∞). Such a behaviour of time series corresponding to the radio spike clusters can be realized when different radio emission pulses are produced by different (independent) local sources. If radio spikes were generated by a nonlinear plasma mechanism (operating in an irregular regime, see Section 2.2),

the fractal dimension of the time series should be a finite value.

3.2 Spectral properties of radio spikes

Radio spikes are most frequently observed in decimetre wavelengths at frequencies $f = 300 - 3000$ MHz [12, 13], and on some occasions in the microwave region [10] up to frequencies $f \sim 8$ GHz [11]. Individual spike clusters can cover a frequency range of several hundred MHz, and the number of separate spikes can exceed 10000 [12]. This is apparently due to the spikes being generated at different altitudes. In this section, we discuss the spectral line profile and the emissions harmonic structure.

3.2.1 The spectral bandwidth of the spikes was discussed in Ref. [12]. In decimetre spectral range the line bandwidth at the half power level is $10 - 15$ MHz, with the relative width being $\Delta f / f \sim 1.5 - 2\%$ and somewhat decreasing at $f > 900$ MHz. The highest resolution in solar observations (20 kHz) was reached in studying spikes registered in VLBI experiments [113]. No fine spectral structures with line widths smaller than 2 MHz were revealed.

A detailed study of the spike spectral line profiles was done in Ref. [101] using a lot of observational data. Eight events (from 208 available) were selected for the analysis, which covered the frequency range from 300 MHz to 8 GHz where the radio spikes are observed. The total number of individual spikes was 196.

The best (although not the most accurate) approximation of the spectral line profile was a Gaussian curve

$$S(f) = S_0 \exp\left\{ -\frac{1}{2} \frac{(f - f_0)^2}{\Gamma^2} \right\}, \tag{88}$$

where the 'decay' constant Γ is related to the line bandwidth at half power as

$$\Delta f_{0.5} \approx 2.35\,\Gamma. \tag{89}$$

Note that the spectra analysed are obtained with spectrometers with an exposure time at the given frequency of 0.5 ms and a spectral scanning time of 100 ms. This means that these spectra are not fully simultaneous. The error due to this non-simultaneity was mostly less than 1% and was ignored in Ref. [101].

The spectral line width for individual events was found to be practically independent of frequency and characterised by a large dispersion. Moreover, the characteristic widths at a fixed frequency vary strongly from one event to another, i.e. Γ is a characteristic of the event and not a function of frequency. The smallest relative width found was 0.23% (7 MHz, which is close to the instrumental resolution limit) in the event of 10.07.1991, the mean value being 0.38%, although paper [10] reported on even smaller widths in the microwave range. Paper [102] also points to a significant dispersion of this value. Maximum values of Γ in an individual event exceed the minimum ones by $2 - 3$ times (in one of the events, by 8 times). The width distribution is asymmetric, with a slow decrease towards larger Γ. No correlation between Γ and the central spike frequency f_0 was revealed. In four of eight events, a correlation between the spectral line width and the spike intensity was found. In three cases, the width increased with intensity, in one decreased (with a large dispersion around regression curves).

The correlation between the spectral width and spike duration has not been studied. However the tight correlation between the spike duration and frequency allows one to suppose that the spectral width is independent of the spike duration (because it is independent of frequency).

When analyzing data on all eight events (covering a wider frequency range than in any individual case) a tendency of increasing Γ with frequency is found. The positive slope of the regression curve is small and the constant coefficient is relatively large:

$$\Delta f_{0.5} = (0.044 f + 10.7) \text{ MHz}, \tag{90}$$

and individual values are strongly dispersed around curve (90) (Fig. 12). Figure 13 demonstrates the dependence of the radio flux at the spike maximum on the central frequency. The microwave spike intensity is smaller on average than in the decimetre wavelengths.
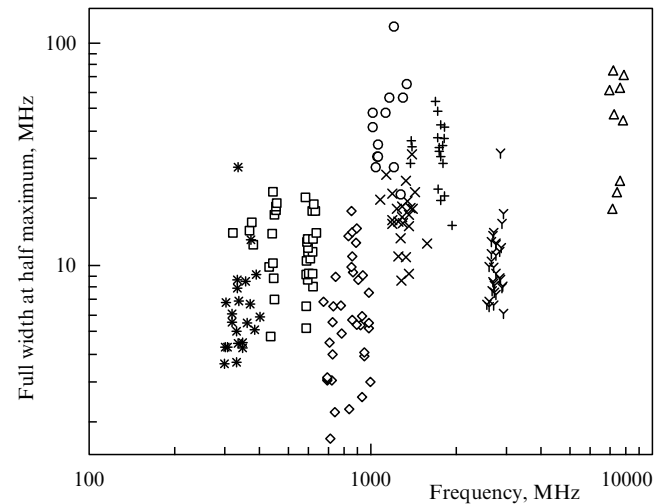
Figure 12. Spectral bandwidth of separate spikes as a function of the central frequency [101]. The regression line is $\Delta v_{0.5} = 0.66 v^{0.42}$. Different spike events are marked with different symbols: 03.04.82 — by squares, 0.4.06.82 — by asterisks, 06.06.83 — by rhombuses, 22.04.89 — by triangles, 08.0190 — by crosses, 25.12.90 — by '0's, 10.07.91 — by 'Y's.
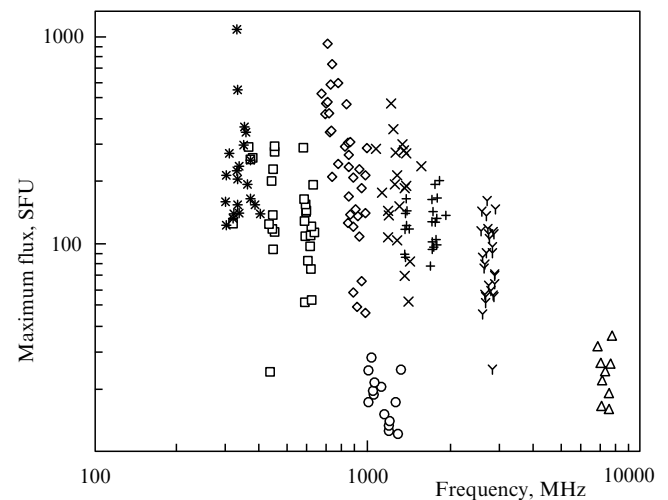

Figure 13. Maximal spike flux density versus the central frequency of individual spikes [101]. The symbols are as in Fig. 12.

The large dispersion of the spike spectral widths and the absence of a clear correlation between the width and other spike parameters (frequency, intensity) lead to the conclusion that observations provide not a natural spectral line width, which is an intrinsic property of the emission mechanism itself, but reflect the spatial inhomogeneity of the radio emission source [101]. In this case the natural width must be sufficiently small (say, less than 0.23% of the minimum width registered). The simple estimates made in paper [101] show that such narrow spectral lines can be provided both by the plasma and cyclotron maser mechanisms (however, in Ref. [101] a correction to the plasma wave dispersion due to the finiteness of the magnetic field strength was not taken into account). Nevertheless some models can be refuted on the basis of the data obtained: the model [19] based on the plasma mechanism ($\Delta f/f > 3\%$, see Section 2.1) and the model based on the 'synchrotron' maser emission ([82], Section 2.3) yielding $\Delta f/f \sim 10-20\%$. Difficulties also appear in the CME model [88] with a Maxwellian hot electron distribution ($\Delta f/f \sim 2\%$). More narrow lines ($\leqslant 1-2\%$) appear for a power-law fast electron distribution [95] (see Section 2.3).

There is one more possible reason for a strong dispersion of Γ which was not discussed in Ref. [101]. As follows from Section 2.3, CME spectral properties strongly depend on details of the fast electron distribution, while in Ref. [101] the dependence of Γ on macroscopic parameters (for example, the emission frequency determined by the background plasma density or the magnetic field strength) and not on the parameters related to the fast electron distribution was examined. Here new information could have been obtained by studying the correlation between Γ and hard X-ray emission characteristics. In addition, the emission line can be broadened by wave scattering (for example, Raman scattering) on inhomogeneities during the emission propagation in the corona. In this case the spike spectral width can depend on the source position on the solar disk, namely, increase from the centre to the limb. The degree of polarization is also sensitive to propagation effects (see Section 3.3), so it is interesting to study the degree of (anti)correlation between the spectral width and the polarization of the spikes. These possibilities have so far not been discussed in the literature. These correlations would open new possibilities for testing short-scale turbulence in the solar corona from the analysis of fine structure of spectral lines in the radio spikes.

Nevertheless, upper limits on the spike source size can be inferred from observations. They were found to vary from 25 to 250 km and to decrease with frequency. No significant differences in properties of decimetre and microwave spikes (except for a somewhat smaller flux density at $f > 2$ GHz) were discovered, which allows them to be considered as events of the same type.

It is interesting to study the dynamics of spectral line time characteristics, which cannot be addressed using the given observational data set (one spike appeared only in one of several 100 ms scans). However, in paper [99] the time resolution was much higher (due to a smaller number of frequency channels), which enabled typical values of the spike frequency drift to be measured. The average drift value df/dt changes from -100 to -400 MHz/s at $f = 362$ MHz to -1000 to -4000 MHz/s at $f = 770$ MHz. Histograms of the number of spikes in the form of the inverse frequency drift (s/MHz) (Fig. 14) correspond to a more or less isotropic velocity distribution of the emitting matter (assuming the drift to
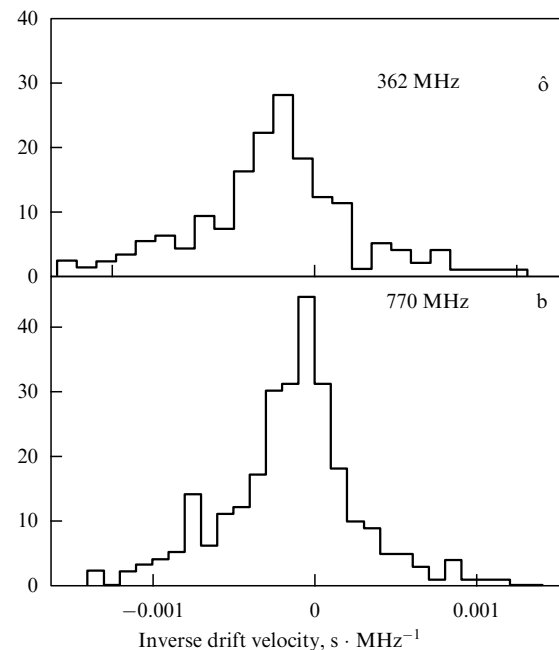
Figure 14. Histogram of the inverse drift velocities of spikes in the events of 04.06.82 (a) and of 17.07.82 (b) [99]. The distributions are almost symmetric.

result from motion of the emitting region) superimposed on a negative (i.e. upward-directed) drift. Nevertheless, it remains unknown whether the observed drift is caused by the motion of the emitting matter or by the internal spectral line dynamics.

3.2.2 The problem of the spike emission harmonic structure was formulated quite long ago since it is one of the key problems in understanding the spike radio emission mechanism. Paper [10] reported on dedicated observations at frequencies of 3.47 and 5.205 GHz which relate as 3:2. No simultaneous generation of emission at the second and third harmonics was found.

Decimetre spikes observed with a spectrometer in the spectral range $200-1100$ MHz were analysed in Ref. [103]. In spite of the spikes frequently forming clusters on the dynamical spectrum (i.e. on the frequency – time plane), clearly frequency-separated regions with the ratio $f_1 : f_2 \sim 1 : 1.4$ were observed only in one (16.05.1981) of 36 events. In this case a correlation between individual spikes from these two regions was present, but the frequency ratio was less stable.

A later study [104] revealed nine events with harmonic structure, with the central spike cluster frequencies being related as small integer ratios, in particular, 2:3, 3:4, 2:3:4, 3:4:5. The frequency ratio in the event of 16.05.1981, which was discussed above [103], is found to be 2.15:3 [104]. It was noted that the most intensive emission changes towards the higher harmonic numbers n with an increase of the main tone. This is probably a consequence of a systematic increase of Y with B (see Section 3.1.1). Emission at the fundamental tone ($n = 1$) has never been detected. The relative spectral width of the spikes was $\Delta f/f \sim 3\%$ and did not depend on the harmonic number.

A correlation analysis of time profiles describing two adjacent harmonic regions demonstrates a notable peak of the correlation coefficient ($r = 0.5$) for zero time lag between

records (an ideal correlation would give $r = 1.0$). This result corresponds to a peak-to-peak correlation between individual spikes at the neighbouring harmonics and not only to the presence of harmonic bands as a whole. This is naturally due to the generation of different harmonics in the same (local) source by the same mechanism.

The integer harmonic ratio in radio spikes was explained in Ref. [31] (see Section 2.2) by generation of Z-waves (near ω_{Be}) and their subsequent coupling with each other and transverse waves. We return to the discussion of this model below in Section 3.3.

The study of the harmonic ratio in radio spikes was appreciably advanced in Ref. [105]. This paper did not address the problem of finding events with the harmonics related as small integer numbers but tried to analyse a maximum statistical set of events with clearly separated clusters on the dynamical spectrum. Frequency correlation analysis was used. From data collected over 15 years of observations, 17 events were selected (of which three had three and one had four harmonic bands).

The harmonic ratios $h = f_1 : f_2$ were generally found not to reduce to a ratio of small integer numbers and h is a well determined value for each event: an increase in the integration time increased the determination accuracy of h and did not increase the dispersion. Only in one case (of 17) was h found to change with time (Fig. 15).

The harmonic ratio does not depend on frequency. The latter strongly varies from one event to another from 1.17 to 1.52 (Fig. 15 d) and is on average 1.4 ($= 7:5$). Six events in the list [105] were the same as in Ref. [104]. However, the values of h derived in these two papers are somewhat different apparently due to the more accurate analysis performed in Ref. [105]. In the events with three harmonic bands h decreased for a higher-frequency pair, but the frequency difference $\Delta f_2 = f_3 - f_2$ was higher than $\Delta f_1 = f_2 - f_1$. If the observed frequencies were an integer multiple of the fundamental frequency, one should obtain $\Delta f_2 = \Delta f_1$.

In spite of the absence of an obvious peak-to-peak correspondence of emission in two neighbouring harmonic bands, a cross-correlation analysis confirms the conclusion [104] of the presence of a correlation (with the same correlation coefficient 0.5). This result strengthens the conclusion that the radio source emits simultaneously at two (or more) frequencies. Only in one event with 4 harmonic bands (02.05.1993) was the correlation coefficient for $2-3-4$ bands about 0.3, and the correlation between the first and the last bands was absent. These deviations of the correlation coefficient from unity in the CME model may be due to the emission time characteristics having different frequency dependences near the first and second harmonics, as discussed in Section 3.1.1. One of the events with three harmonic bands (06.04.1980) was different from all others: it
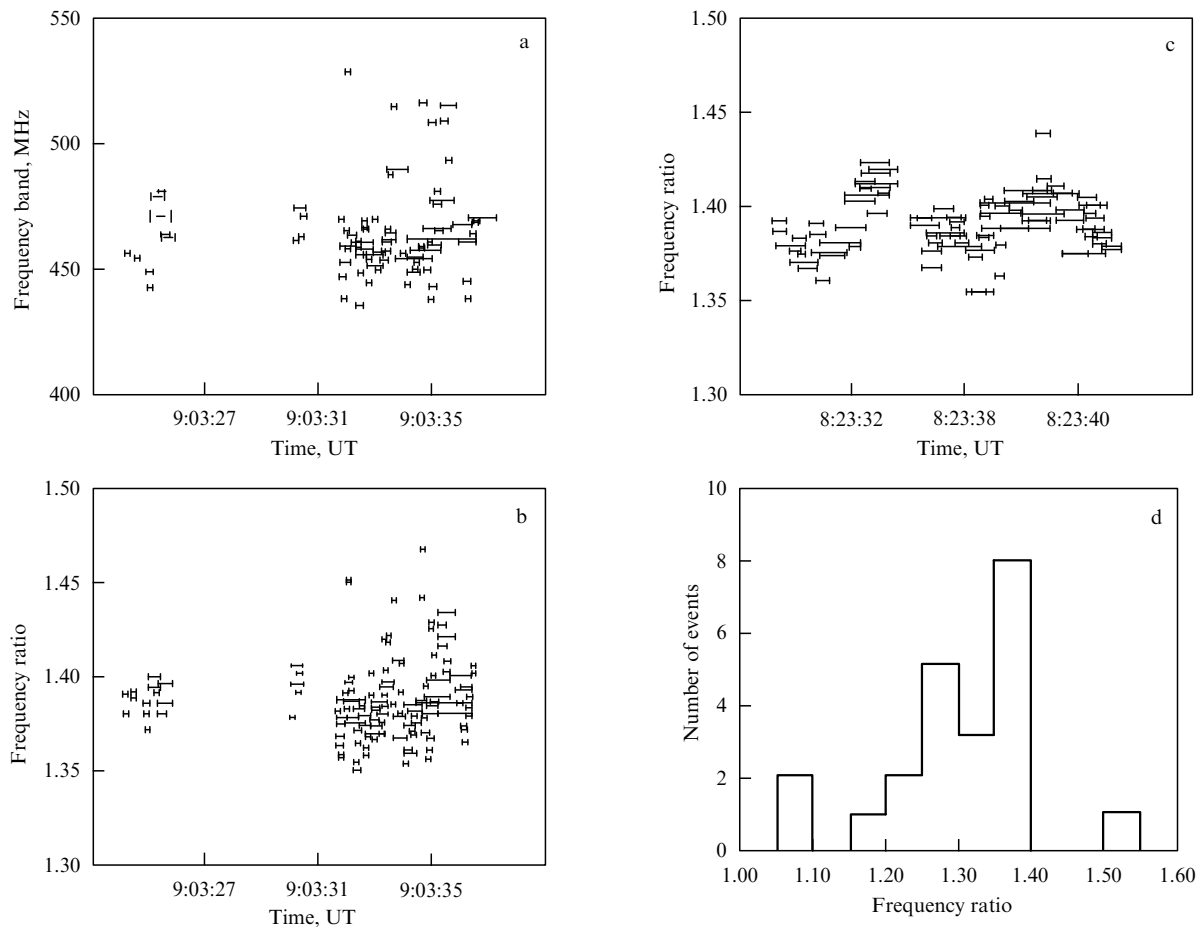
Figure 15. The spike harmonic structure [105]: (a) the frequency separation between two harmonic bands in the event of 14.09.92 is determined by the anticorrelation for different time intervals, the error intervals are determined by the integration time; (b) the frequency ratio of the harmonic bands for the same time intervals; (c) the frequency ratio in different parts of the spike event of 16.05.81, the change with time is clearly seen;(d) the harmonic ratio distribution for all events.

was the only one in which the spikes were generated in the high-frequency tail of a strong type IV burst and the harmonic bands were characterized by a fast drift ~ 4 MHz/s towards the high frequencies (see Fig. 1b).

Among all theoretical models for radio spike generation considered in Section 2, non-integer harmonic ratios appear only in the model of CME formed by a power-law fast electron distribution [95], although accurate calculations of the relative intensity of different harmonics are still to be done.

### 3.3 Polarization properties of the radio spikes

3.3.1 The radio spike polarization is known to vary from 0 to 100% [13]. In the first event with a harmonic structure registered (16.05.1981) [103] both components had left circular polarization, the degree of polarization of the low-frequency and high-frequency harmonics being ~ 16% and ~ 7%, respectively. When the magnetic field direction in the radio source was determined by the leading spot in the photosphere, the left-polarized waves corresponded to an ordinary mode. Of nine events with harmonic structure discussed in Ref. [104] the polarization data is given only for one (31.08.1989). The polarization degree in all harmonic bands was ~ 20% (and the polarization was right). Melrose's model [31] predicts non-polarized radiation at low harmonics and an increase in polarization (possibly, up to 100%) at the highest harmonics. Clearly, the present observations are inconsistent with this model (as well as the harmonic ratios themselves, see Section 3.2).

In paper [105] the polarization of emission in different harmonic bands was not discussed at all. So the question of whether the polarization in the two events [103, 104] is typical for all cases with harmonic structure remains open. It is unclear, in particular, whether different harmonics can have opposite polarization and what is the characteristic degree of polarization at different harmonics.

3.3.2 Significant efforts have been made to determine the dominating emission mode in spikes (O or X). Paper [36] analysed 36 events showing type III activity at low energy. Of them, in 13 cases both the spike polarization and the corresponding H_α flare were known. In the framework of the leading spot hypothesis the dominating mode was determined: in 10 events it was an O-mode. As a rule, the polarization direction in the spikes and in the accompanying type III bursts was the same (type III bursts are thought to be polarized as O-waves).

However, an analysis made in Ref. [106] using a larger data set yielded the opposite result: in 98% of cases the rotation direction of the polarization vector in the spikes is opposite to that in the associated type III bursts; the polarization degree in the spikes is higher than in the type III bursts (Fig. 16a). In this paper, 38 spike events with known polarization and 32 events with known corresponding H_α flare location were selected from the total list of 163 spike clusters. Clusters connected with both type III and type IV activity were present among the selected events. For this reason the type of spike dominating mode was determined not only from the polarization of the corresponding type III bursts but from other criteria as well, viz. with account of the polarity of the 'leading' spot and of the closest spot to the H_α flare.

According to the leading spot hypothesis, 68% of spikes have X-polarization (71% of type III bursts have O-polarization), and according to the closest spot hypothesis the fraction of X- and O-polarized spikes amounts to 79%.
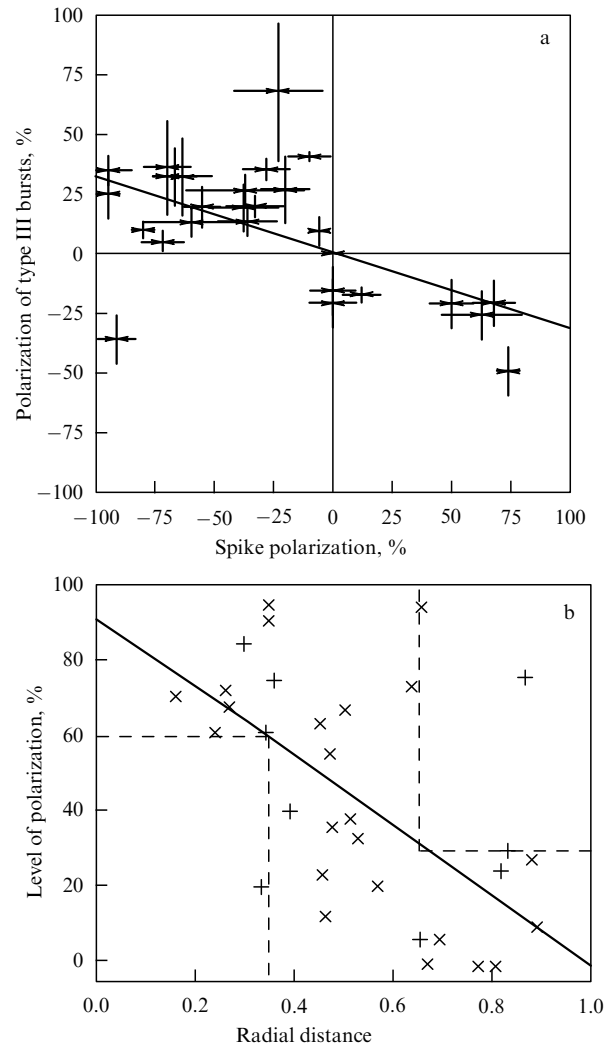
Figure 16. (a) Type III burst polarization in comparison with spike polarization for 29 spike events associated with type III bursts. The diagonal line shows the calculated regressive dependence [106]. (b) The correlation diagram of the polarization degree dependence as a function of the radial distance of the corresponding H_α flare projected on the solar disk. The events associated with type III bursts are labelled with crosses, other events with pluses [106].

Thus in most cases an X-mode dominates in the spike emission, although there are also O-polarized spikes. This is indirectly supported by the sometimes observed change of polarization sense inside one cluster.

3.3.3 In addition, in Ref. [106] a dependence of the spike polarization degree on the location on the disk was found. The spikes formed near the disk centre are strongly polarized and those near the limb are weakly polarized (Fig. 16b), which indicates the important role of the propagation effects in the resulting degree of polarization.

3.3.4 The spike polarization in the microwave region [10] varies from −100% to +100%, whereas the polarization of the associated continuum microwave flare changes from −20% to +20%. These two polarizations do not correlate either in intensity or sign. The polarization sense was sometimes observed to change between consecutive spikes at one frequency.

Note that from the theoretical point of view it is hard to expect the microwave continuum to correlate with microwave radio spikes at the same frequency. The point is that the microwave continuum is formed by a gyrosynchrotron radiation in an optically thin source. In contrast, the radio spikes are generated by a coherent mechanism, which requires a large (and negative) optical thickness of the source. Clearly, these two conditions cannot be met simultaneously in one source at a given frequency. This implies that the continuum and radio spikes are generated in spatially separated regions, which explains the lack of correlation between these two types of radio emission. Conversely, for the same reason a tight correlation is expected between decimetre radio spikes and the microwave continuum (the microwave continuum source becomes optically thick at decimetre wavelengths).

3.3.5 A change of the polarization sense in two consecutive spikes formed in one source was reported in Ref. [107] in interferometric observations with a high one-dimensional spatial resolution (the spatial properties are discussed in more detail in the next section). Here the first spike of lower intensity had a polarization of 26% in the O-mode and the next, more intensive spike had an 83% polarization in the X-mode. As noted in Ref. [107], the change of the polarization sense is most probably due to a change of the dominating emission mode and not to a magnetic field polarity change or spatially different sources of the spikes. Such behaviour is difficult to explain by plasma mechanisms, it is considered in Ref. [107] as supporting CME. It is not completely clear near which gyrofrequency harmonic, the first or the second, this radiation is emitted.

3.3.6 VLA observations with two-dimensional spatial resolution [108] have revealed that the radio spike polarization direction corresponds to an X-mode. However, the X-polarization in the associated type III burst is rather surprising.

An analysis of the group delay of the oppositely polarized components in partially polarized spikes [109] indicates a totally polarized emission forming in the source, and the appearance of a weaker mode in the observed radio emission is due to propagation effects of the primordial (100% polarized) radiation in the solar corona.

Therefore, the bulk of the observational data available supports the preferential X-polarization of the spikes, which can change to O-polarization. These properties are hard to explain by plasma models but they conform with the cyclotron maser mechanism. Nevertheless the published results remain somewhat contradictory, so further steps should be made to determine the dominating mode in the spike emission.

3.4 Spatial characteristics of the radio spikes
The spatial properties of the radio spikes include the following observational data: (a) the size and dynamics of the spike source (both for individual spikes and a group of spikes); (b) the dynamics of separate spike locations in the group; c) the location of the spike source relative to other sources of electromagnetic radiation in solar flare; (d) the dependence of the spike properties on the location on the solar disk and on the location relative to the local magnetic field direction.

3.4.1 As discussed above, an estimate of the source size in an individual spike can be obtained from the relative spectral

width and is 20-200 km ($0.03 - 0.3$ arcseconds). Microwave sources must be more compact than decimetre ones. However, these estimates are insufficient to make a choice between the currently available spike mechanisms and models. This could be done by direct measurements with a high angular resolution. In addition, the expected spike source size turns out to be much less than the existing angular resolution capability of instruments used in solar radioastronomy (VLA, WRST, SSRT). The required resolution can only be achieved using very long baseline interferometry (VLBI). Besides measurements of the source sizes, VLBI-systems with a high time resolution ($t_{int} < 1$ ms) can observe, for example, the dynamics of the size of a one-component radio source $D(t)$

$$D(t) \propto \sqrt{- \ln |V(t)|}, \qquad (91)$$

using the measured coherence coefficient value $V(t)$. This becomes possible if the burst duration is several times as long as t_{int}. VLBI-systems make it possible to watch the shift in the radio brightness phase centre of a multi-component radio source using observations of phase jumps and interferometric signal amplitude with an accuracy

$$\sigma(\phi) \approx \left(\frac{\lambda}{L} \right) \text{(noise/signal)}. \qquad (92)$$

For example, using a wavelength $\lambda = 50$ cm, a base length $L = 400$ km, and the signal-to-noise ratio $= 0.1$ it is possible to determine the radio brightness centre shift to an accuracy of 0.001 arcsec.

The first VLBI observations at a wavelength of 18 cm indicated the presence of a weak interferometric signal from a group of radio spikes with a size smaller than $0.07''$ [110, 111]. However, VLBI observations of 26 microwave ($f = 2.297$ GHz) bursts with a fine spike-like time structure performed in $1989 - 1990$ [113] did not reveal sources with sizes smaller than $0.09''$. One of the possible reasons for such a result could be the poor time resolution of the detectors, which did not allow individual spike sources to be watched [138]. The integration constant was 200 ms. Several spikes might have occurred during this time with sources spatially separated by an angular distance exceeding the interferometer beam width. In this case the interferometric picture is smeared. With such a time resolution the detector can register an interferometric signal only from a very compact source for the spike group, if any.

Another reason can be an increase in the observed spike source size due to radiowave scattering on the coronal plasma inhomogeneities [115]. However, at present coronal inhomogeneities are poorly studied. No systematic studies of the radio emission scattering in the solar corona at distances smaller than $2R_s$ have been made so far. Such observations are only being planned [138, 139]. If the turbulence has the same characteristics as assumed in Ref. [115], one should expect strong radiowave scattering and broadening of the visible sizes of spike sources up to $10 - 50$ arcseconds. In this case it is impossible to obtain an interferometric signal from the spike sources by conventional methods. The real source sizes can be estimated only by increasing the time and spectral resolution of the VLBI-detectors (e.g., the speckle-interferometry method).

There have been some experiments in which spikes were observed with short-base interferometers. From these observations, upper limits on the visible source sizes can be derived

and, given a sufficiently high time resolution, the spatial dynamics of the sources can be studied.

The first such observations were done with a two-element interferometer at a frequency of 2.8 GHz [107]. From an analysis of a burst with a time profile consisting of a continuum and spikes superimposed on it, an upper size to the spike source was derived to be 28″.

Observations conducted with the Siberian Solar Radiotelescope (SSRT) with an angular resolution of 15″ at a frequency of 5.57 GHz and a time resolution of 56 and 14 ms have shown that [112] the visible source size of bursts with a duration less than 1 s can lie within the range from 5 to 50″. The size of a long ('underlying') burst is usually larger than the size of the corresponding subsecond structures and varies from 5 to 65″. Apparently, these are the first observations that have enabled the visible source size of the centimetre radio emission superfine time structures to be determined.

Such large visible sizes are compatible with the assumption [115] of the significant role of scattering in the solar corona, which indicates once again the need for independent detailed observations of coronal inhomogeneities. The scale and amplitude of electron concentration inhomogeneities over an active region are needed not only for VLBI studies of solar flares, but also for solving other problems, such as studies of the nature of turbulence in the corona and of the particle dynamics in coronal magnetic loops, the interpretation of the observed radio emission spectra, etc.

3.4.2 Phase measurements on short-base interferometers provide us with information on the position change of separate spike sources in a group [see Eqn (91)]. In particular, it has been established that the source locations in one event can be strongly different.

The first two-dimensional VLA observations of the spikes [108] enabled the visible spike source size (70″) in metre wavelengths (333 MHz) to be measured for the first time. VLA has a low time resolution, $\Delta t = 1.667$ s, and the spike duration is of the order of 0.1 s, so during the size and location measurements more than one spike occurred (i.e. in fact the size and location of a group of spikes were measured). Nevertheless it was firmly established that the centre of mass of each of three registered spike groups (separated by $50″ - 130″$ and located at an altitude of $(4-5)10^{10}$ cm over the photosphere) slightly changed (by $10″ - 25″$) between the main intensity peaks (separated by $2-4$ s) on the time profile with a fine structure of a 15-s burst registered by the spectrometer.

Observations of solar spike-like bursts in decimetre wavelengths on a two-element (2 RT-7, $f = 540$ MHz) interferometer with a base of ~ 400 m using a specially designed correlator with a millisecond (down to 0.256 ms) time resolution [114] allowed the brightness centre shift to be observed in the source with the appearance of each new short pulse in the group. For the most intensive bursts the accuracy of the positional measurements in these experiments was $1-2″$. It was found that the position of separate peaks in multicomponent bursts sometimes differs significantly from others and this difference amounts to 15″ for peaks separated by $100-500$ ms.

Such spatial shifts of the spike sources can be interpreted as a confirmation of the model in which each spike is generated by a separate local source. On the other hand, the model based on the irregular regime of a nonlinear plasma mechanism predicts a more stable spike source position at a given frequency and thus does not explain the observed spatial shifts. The subsecond burst generation in separated local sources is possible both by plasma mechanisms (based on beam or loss-cone instability) and by CME. The fragmentation of the emission sources formed by the beam instability seems to arise only in the case of spatial fragmentation of the flare energy release and particle acceleration region. At the same time, the source fragmentation for the loss-cone instability can be due to secondary effects related to the presence of MHD-turbulence in the flare loop and the formation of local energetic electron traps (see Section 3.5).

3.4.3 According to Ref. [107], the position of the 'barycentre' of a spike source differs appreciably (by 25″) from the position of the corresponding microwave continuum burst observed at the same frequency. Additionally, an analysis of all subsecond spikes registered by SSRT [112] suggested that their source position in most cases coincided with the brightness centre of a long burst (within the error of the SSRT positional measurements of $5-8″$). Only during two events was a shift similar to that found in paper [107] observed. This provides additional evidence for the existence of physically different types of superfine time structure of solar radio emission, so the term 'radio spikes' for all subsecond features in the radio emission (which is sometimes met in the literature) is incorrect and can lead to a wrong burst classification and erroneous physical conclusions [112c].

3.4.4 In paper [10] no notable dependence of the microwave spike parameters on their position on the disk was found, although paper [106] reported on a decrease in the spike polarization from the centre of Sun to its limb. In Refs [107, 112] an increase in size of both the spike sources and the corresponding long bursts ('underlying' bursts of spikes) towards the limb is found. As noted above in Section 2, such behaviour of the source size and polarization can be explained by radio wave scattering on coronal plasma inhomogeneities.

3.4.5 Note that in papers [107, 112, 114] no parallel observations of broad-band dynamical spectra were made. So in spite of the short duration of the detected spike-like bursts, they cannot be unambiguously classified as spikes in the sense defined in Section 1. This is also confirmed by the fact that the subsecond bursts analysed in Ref. [112] are observed much more frequently (in 15% of all microwave bursts at $f = 5.7$ GHz) than spikes (Section 3.5). In reality most of these bursts may be narrow-band 'blip'-like bursts generated by the electron beams by the plasma mechanism [143] or broad-band type III bursts. So the question about the space-time fragmentation of the spike source in a group (or cluster) requires further studies including parallel observations with a high time, spatial, and spectral resolution.

Unfortunately, until now there have been no statistically complete studies aimed at measuring spike source positions relative to other sources (optical, X-ray, etc.). An interesting result was obtained recently in paper [145] where the position of two subsecond bursts at a frequency of 5.7 GHz relative to a long burst in the microwave and soft X-ray bands during the flare of 06.09.92 were analysed. The subsecond burst source was found to be located close to the top of the flare loop, which was characterized by a high plasma density, high temperature, and low magnetic field strength ($\omega_{pe}/\omega_{Be} \ll 1$). The authors came to the conclusion that these subsecond bursts were produced by the plasma mechanism. However in

this case the lack of spectral observations does not allow one to definitely conclude whether the bursts observed belong to the spike spectral class.

In contrast, in paper [146] the event of 14 May 1993 is analysed in which spectral observations suggest the rapid radio emission fluctuations to be radio spikes. In this case the radio spike source was found to lie at a distance of $17''$ (20 thousand km) from the main burst source. The spikes in this event were shown to form in a compact magnetic loop with a high magnetic field (~ 1600 G), and the plasma-to-gyrofrequency ratio was found to be $Y = 0.32$ (i.e. 10 times smaller than in the main burst source), which is a strong argument in favour of the maser cyclotron mechanism.

A spectral identification of the burst sources with interferometric observations seems to be very promising for studies of the nature of spikes. Such studies would allow the determination of the location (and distribution) of the spike sources inside the flare loop and, in particular, their position relative to the magnetic field direction near the spike source. This is important in order to choose the spike generation mechanism, since there are significant differences in the instability conditions for different types of generated waves (see Section 2.2).

Calculations [88, 93] have shown that there is a difference in the directions of X1- and X2-mode emission relative to the spike magnetic field. X2-waves are generated in a quasi-transverse direction for $0.1 \leqslant \omega_{pe}/\omega_{Be} \leqslant 0.9$. The X1-mode is effectively generated in a quasi-transverse direction only for very low Y. It is emitted in a direction practically parallel to the magnetic field for $\omega_{pe}/\omega_{Be} \geqslant 0.2$. Interferometric observations with differently oriented bases, which allow positional (phase) location of the brightness centres of the spike radiosources to be measured with an accuracy $\leqslant 1''$, combined with two-dimensional observations of flare loops (on such instruments like SSRT and VLA or on X-ray telescopes SXT/HXT onboard the YOHKOH satellite) could separate these modes. Such observations are convenient in enabling one to observe spike sources for any flare loop axis direction on the solar disk. The X1-mode should come from the flare loop parts which are seen by a small angle with respect to its longitudinal magnetic field, and inversely for the X2-mode.

It appears that simultaneous observations of the flare loops with a spatial resolution of $\sim 1 - 5''$, a time resolution of ~ 1 s in the X-ray and microwave bands, and radio interferometric measurements of the spike brightness centre positions using three radiotelescopes with differently oriented bases and a position measurement accuracy of $\sim 1''$, could be crucial for the choice between different spike generation mechanisms. Such observations would allow one not only to localize the part of the flare loop from which the spike radiation is emitted, but to determine the plasma density, magnetic field strength, and temperature inside this region.

Such complex experiments make it possible to measure one of the most important parameters $Y = \omega_{pe}/\omega_{Be}$ (as in paper [146]), as well as to determine the conditions of radiation escape from the spike source, which strongly depend on temperature.

3.5 Relation of radio spikes to other types of nonthermal electromagnetic radiation in solar flares

Radio spikes are usually observed at the impulse phase of a flare and are undoubtedly due to non-thermal fast electrons generated during the primary energy release. So it is natural to expect a noticeable correlation between the radio spike properties and characteristics of other secondary nonthermal emissions, such as the microwave continuum (gyrosynchrotron radiation), type III meter-wavelength bursts (plasma mechanism), and hard X-ray and (soft) gamma-ray emission (bremsstrahlung emission of fast electrons).

3.5.1 A lot of papers [116 – 122] are devoted to studies of correlation properties between dm-events with fine time structure and hard X-ray flare emission (HXR).

In paper [116] the relation between radio spikes, type III bursts, and HXR emission observed by ISEE-3 satellite was studied. A high correlation between these types of emission was found. In particular, during the event of 24.09.1980 (07:28 UT), four big HXR peaks were found to correlate better with groups of spikes than with type III bursts. The highest peak (with the hardest spectrum) coincides in time with two clearest type III bursts. An analysis of 17 events showed that HXR bursts connected with spike activity are the most impulsive among other events: their mean duration is ~ 144 s, while the average duration for all bursts is ~ 256 s. 71% of all spike-bursts were found to be connected with enhanced hard X-ray emission, while for the type III bursts the relation to HXR is found in 45% of cases.

Generally speaking, decimetre radio emission correlates much more with HXR than metre bursts (which are generated in higher coronal layers): 45% of all decimetre events occur during the hard X-ray emission peaks [117], whereas only 3% of metre type III bursts are associated with an enhanced HXR flux [118]. Nevertheless, in these cases a good time correlation between HXR and type III bursts is observed [118 – 120].

Various types of decimetre radio emission [14] demonstrate different degrees of correlation with HXR. In particular, all decimetre events (except for type III bursts) are related to HXR in 69% of cases [117]. This value is smaller for pulsations and is $20 - 35\%$ [121]. The continuum emission is usually related to a long HXR burst.

The most detailed studies of the relation and correlation degree between decimetre radio spikes and HXR were done in papers [122, 123]. Paper [122] focuses on time characteristics: durations, delays, correlated details. Data on HXR bursts registered by the SMM (Solar Maximum Mission) satellite with a sensitivity 3 times higher than ISEE-3 were used. Events satisfying the following criteria were selected in [122] among decimetre events with fine time structure: (a) the events form clusters from $10 - 10^4$ individual spikes irregularly distributed on the dynamical spectrum; (b) the cluster occupies a broad spectral range ($\geqslant 100$ MHz); (c) individual spikes have a narrow spectral line (a few %); (d) the individual spike duration is short and follows the dependence (84). Such an approach excludes from consideration not only other types of decimetre radio emission fine time structures, such as broad-band pulsations, groups of narrow-band decimetre type III bursts ('blips' [14]), but also metre type I bursts and also a fine structure observed near the initial frequency in type III bursts (at 300 MHz), which are called metre spikes. Spikes associated with both type III and IV bursts were included.

The full list of analysed events included 164 spike clusters registered from 24.12.1977 to 26.09.1989. Of them, 90% occurred during HXR observations by SMM. Only in two cases was no enhanced HXR flux detected. Nevertheless, in both these cases, weak continuum bursts in the microwave

wavelengths were observed. This allows us to suppose that in these two cases the hard X-ray emission was also generated but its intensity was lower than the detector sensitivity level.

In the statistical analysis, all 90 events (occurring over 44 solar flares) were separated into 4 groups according to the association rate with HXR, and into 4 groups according to the correlation degree with HXR (for events associated with HXR).

As a result, a very high association rate of the spike clusters with hard X-ray emission, of about 95%, was established. Only two events (2.2%) were not accompanied by enhanced X-ray emission and three (3.3%) occurred during long-term HXR flux increases with no pulse peaks on comparable time scales. Thus 98% of the spike clusters occur during enhanced X-ray emission flux and 94% are related to impulsive X-ray emission on the time scales comparable with radio emission.

3.5.2 Events connected with HXR demonstrate different correlation degrees between the average time profiles of radio- and X-ray emission. Of 84 radio-HXR pairs, a very tight correlation was found for 40 pairs, a smaller correlation (for example, there are uncorrelated details) was established for 26 cases, and in 18 cases HXR better correlates with type III burst time profiles. Therefore, in 75% of cases there is a statistically significant correlation between radio- and X-ray time profiles, and in 48% of cases the form of these emissions practically coincides. These statistical data show that milli-second radio spikes are mostly related to the hard X-ray emission generation among all types of dm-radio emission with a fine time structure. The hard X-ray event duration is usually longer than that of the corresponding radio spike cluster, with a linear relation between the two.

A large number of individual spikes in the event (up 10^4 and higher) was interpreted as evidence for the solar flare energy release fragmentation [12]: one spike was identified with an elementary act of energy release. So the discovery of noticeable time delays between radio- and X-ray emission in Ref. [122] turned out to be quite unexpected. On average, spike cluster peaks occur 4 s later than HXR peaks, and in some cases the delays can be as high as tens of seconds (there are even cases with negative delays).

This is even more surprising considering that the delays between HXR and type III outbursts do not usually exceed 1 s (in the cases with clear correlation) [16, 122]. Such appreciable 'HXR-radio spikes' delays are inconsistent with an interpretation of radio spikes as being direct indicators of the energy fragmentation in the flare.

In Ref. [122] more detailed time profile properties are also noted. The spikes most frequently start appearing after the beginning of the X-ray burst but before its maximum (i.e. at the HXR growth phase). The similarity of time profiles of both emissions are often observed on rather short time scales (of the order of several seconds).

Although radio spikes almost always are associated with hard X-ray emission, the reverse is not true: only about 2% of HXR bursts are associated with radio spikes. This unambiguously points to the need for some special (threshold?) requirements to be fulfilled for radio spikes to be generated. The following property, discovered in Ref. [122], seems to be important to finding these conditions: spikes in a cluster start appearing with an approximately constant amplitude after the HXR of a burst has begun, but their number per unit time increases. To the end of event, individual spikes do not

become weaker, but their number decreases. The spike clusters usually disappear earlier than the X-ray burst.

Such behaviour is interpreted in Ref. [122] as the system's passing above some critical threshold with the subsequent coming below the threshold. We return to discussing possible reasons for the discovered correlation properties of 'HXR-radio spikes' below, after having considered the role of spectral properties of the hard X-ray emission.

A more detailed study of the correspondence between time profiles of radio and hard X-ray emission is given in Ref. [123]. In this paper the hard X-ray time profile $f_x(t)$ is assumed to coincide with the time profile of the particle injection into a trap $f_i(t)$, and the observed radio signal $S_r(t)$ is thought to result from the trap response [described by a response function $g(t)$] to the electron injection:

$$S_r(t) - S_{r_0} = \int_{-\infty}^{t} \{ f_x(t') - f_{x_0} \} g(t - t') \, dt', \qquad (93)$$

where the background components of the X-ray (f_{x_0}) and radio emission (S_{r_0}) are explicitly omitted. The response function $g(t)$ was modelled by a Gaussian

$$g(t) = g_0 \exp \left\{ -\frac{(t - t_d)^2}{2\sigma^2} \right\}, \qquad (94)$$

where t_d corresponds to the time delay between the radio- and X-ray emission, the normalization

$$g_0 = q_f \left\{ \int_{-\infty}^{t} g(t - t') \, dt' \right\}^{-1} \qquad (95)$$

includes the threshold ratio of radio emission (expressed in sfu, 1 sfu = 10^{22} W/m^2 Hz) and X-ray emission (in units of counts per second), $q_f = F_r / F_x$.

During the analysis of observations, the free parameters t_d, σ, f_{r_0}, f_{r_0} varied such as to provide the best fit to the mean time profile of radio spikes. The radio data were averaged both in frequency (over the entire band occupied by the spike cluster or over some sufficiently broad intervals) and in time with an integration constant of 1 s, although non-averaged data (with a resolution of 0.1 s) were also considered.

The data list included 27 spike clusters which occurred during 14 X-ray outbursts. In no case was a negative delay t_d found. This delay was found to vary between 0.1 to 8 s, with most cases (20 of 27) lying within the range $2-5$ s. The dispersion of the delays obtained in Ref. [123] is smaller than in Ref. [122], where the relative time of the emission maxima (or their beginnings) were compared, since the data fitting by Eqns (91), (92) takes into account the entire time profile form and not only some 'characteristic' points like the beginning of a burst or its maximum (Fig. 17).

The fitting procedure described above is most adequate for the pulse growth and decay phases and is less accurate near the peak. In some cases a fraction of the X-ray peaks are not associated with the spikes, while others very tightly correlate with radio emission. Despite the similarity of the average time profiles, the radio emission demonstrates a much richer structure, with no peak correlation.

In Ref. [123] a linear correlation was found between the time delay t_d and the X-ray flux F_x

$$t_d \approx 2 + 6 \left(\frac{F_x}{1000 \text{ s}^{-1}} \right), \qquad (96)$$
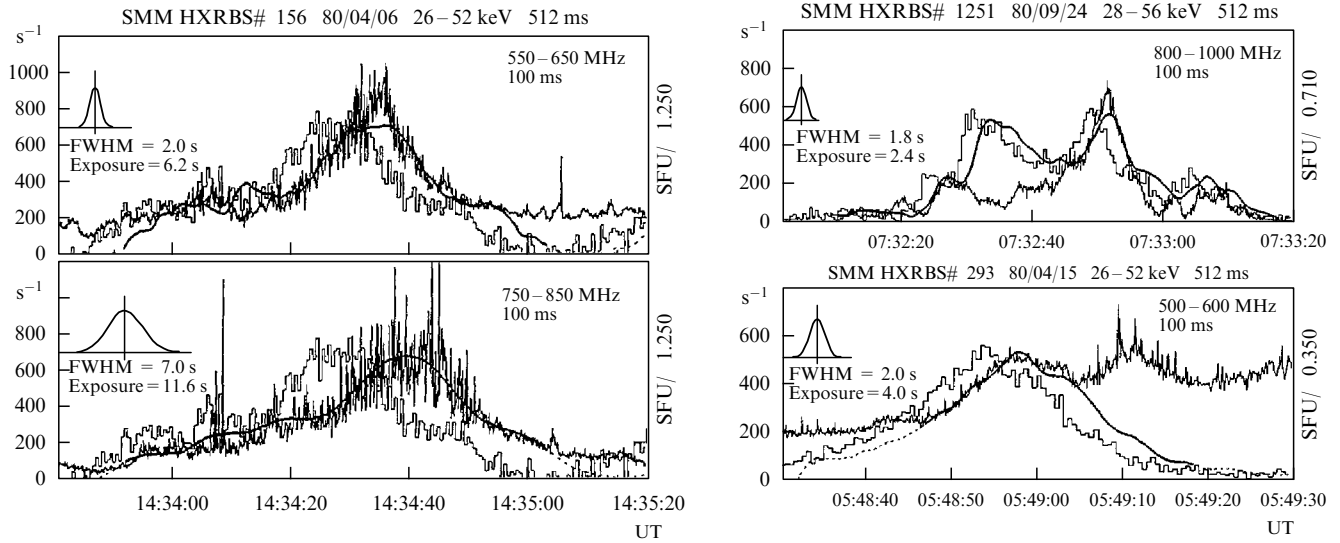
Figure 17. Relation between the spike cluster time profiles (the thin black line) and HXR emission (the thin grey line). The thick line shows the convolution of the average HXR profile with the response function shown in the upper left of each plot.
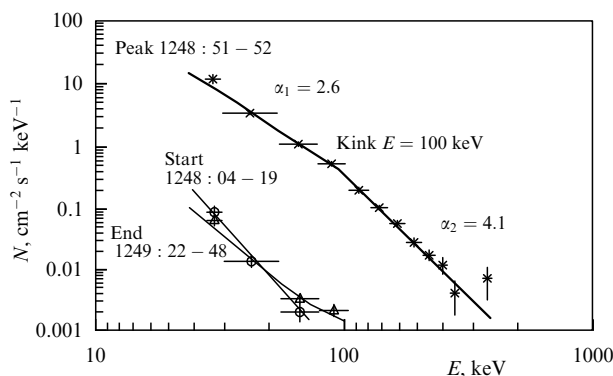
and a similar correlation was found between the X-ray and radiofluxes:

$$F_x(c^{-1}) \approx 100 + 0.6\left(\frac{F_r}{1 \text{ sfu}}\right). \tag{97}$$

3.5.3 The spectral evolution of the associated X-ray emission (and, correspondingly, of the fast electron spectrum) plays a fundamental role in understanding the radio spike generation process. As shown in Ref. [123], the time evolution of radio emission is mostly correlated with $25 - 50$ keV HXR.

The HXR spectra considered in most cases were best fitted by power laws, and in the most powerful outbursts a double power law with a break at 100 keV and a larger slope at high energies was observed (Fig. 18).

In the majority of outbursts (70%) the spectrum became harder with an increase of the HXR flux and then softer at the decay phase. In other cases no softening of the spectrum at the decay phase was noted.

Figure 19 demonstrates the dependence of the X-ray spectral index on time ($t = 0$ corresponds to an HXR peak) for different flares. The thick solid lines mark the intervals where spike clusters are observed. Clearly, the radio spikes appear during the X-ray emission with the hardest spectra.
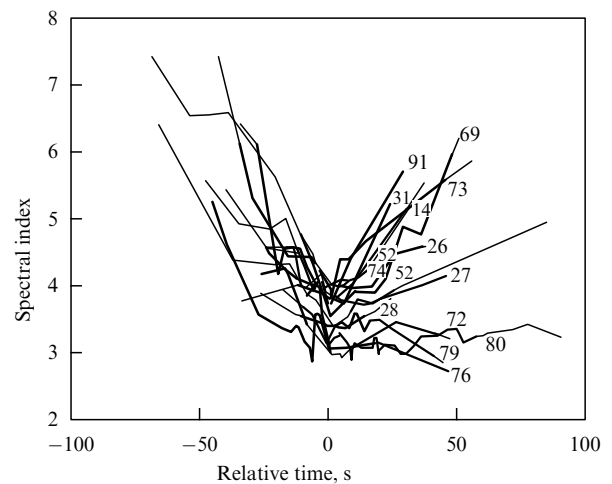
Figure 19. Dependence of the X-ray emission spectral index on time [123]. Intervals with cluster spikes are marked.

Paper [123] suggests several scenarios which could lead to the observed time delays, although it is noted that none of them explains the delays in a natural way. In our opinion, the key to understand the delays is the spectral evolution of fast electrons accumulated in the trap. Note that the plasma wave growth rates are most sensitive to the low-energy part of the fast electron spectrum, for instance, lowerhybrid wave growth rates (for $Y < 1$) increase with increasing spectral index ξ [93, 94] (Section 2.3, Fig. 8c). So if the spikes were generated by a plasma mechanism, they must have appeared with soft HXR spectra as well.

Oppositely, the transverse wave growth rates generated by CME decrease with increasing ξ. This is due to the negative

Figure 18. Example of an HXR-burst (spike cluster of 31.08.80) with the spectrum fitted by a double power law with a kink at 100 keV, steepening at higher energies [123].

contribution to the growth rate (79) of the distribution function derivative with respect to momentum modulus $p\partial f/\partial p = -\xi f$. For an instability to develop, the positive contribution from the derivative with respect to the angle $\partial f/\partial \mu$ should prevail over this negative contribution.

It is natural to consider a unique (but time-dependent due to continuous particle injection and their dynamics in the trap) electron energy spectrum to form during the flare [124]. The angular dependence of the distribution may vary strongly in different parts of the loop due to magnetic field fluctuations (local magnetic traps). In this case the instability conditions with a hard spectrum can be met in a large number of 'local sources', which increases the number of observed spikes (and the average radio flux).

There is some observational evidence for spectral hardening of electrons kept inside the flare loop. A comparison of the time profiles of HXR- and microwave (gyrosynchrotron) bursts [125–127] revealed a few second delay in the microwave emission, which suggests that the maximum in the number of high-energy (mildly relativistic) electrons in the trap occurs later than for low-energy electrons. A detailed study of the dynamics of the impulse bursts in centimetre and millimetre wavelengths based on the RAS data of the University of Bern [128, 129] showed that the frequency spectrum in the optically thin region becomes flatter during the total emission peak, and becomes especially 'hard' a few seconds after the maximum (Fig. 20). Since in cm-mm-wavelengths the emission is generated by the gyrosynchrotron mechanism, these results directly indicate the flattening of the energy spectrum of trapped electrons both in the intensity growth phase and in the after-maximum phase. Therefore, during these burst phases the conditions most favourable for spike generation by the cyclotron maser mechanism are created, in total accordance with observations [123].

More evidence for radio spike generation by the cyclotron maser mechanism comes from the presence of a kink at $E \sim 100$ keV in the HXR energy spectrum. As noted above in Section 2.3 [93, 94], the CME instability is suppressed by extended fast electron energy distributions. In the case shown in Fig. 18 the electron energy spectrum is hard at intermediate energies and rapidly decreases at high energies, which provides ideal conditions for CME with quasi-power-law fast electron distributions.

3.5.4 Such a radio spike generation model has some more observationally testable properties. The requirement of a sufficiently rapid fall in the electron spectrum at energies of several hundred keV (corresponding to the spectral kink in the HXR spectrum at $E \sim 100$ keV) suggests an absence (or anomalously low intensity) of gamma-ray emission in $0.3 - 2$ MeV energy range, which is generated by bremsstrahlung radiation of fast electrons (see also Section 2.3, Fig. 7). Indeed, no report is known in the literature on simultaneous observations of radio spikes and gamma-ray emission (we recall that almost all spike clusters are associated with hard X-ray emission).

The well-known event of 15 June 1991, which was accompanied by a very long gamma-ray emission, provides an exception. At present, only a preliminary analysis of the radio spikes in this event has been published [129a], based on which it is hard to make reliable conclusions on the relation between the spike generation and the prolonged gamma-ray emission.
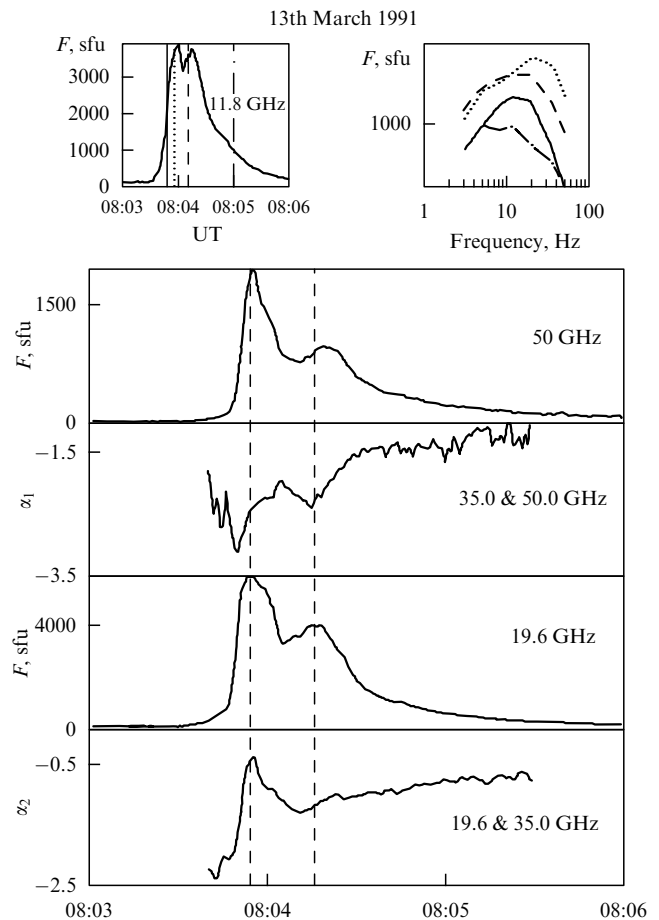
Figure 20. Radio emission spectrum flattening in cm-mm frequency range. α_1 and α_2 are spectral indices in the high-frequency region where the gyrosynchrotron source is optically thin, and near the maximum spectral frequency where the optical depth $\tau \leqslant 1$, respectively.

3.5.5 Along with generating gamma-ray emission, the fast electrons with mildly relativistic (and ultrarelativistic) energies manifest themselves in microwave continuum outbursts of the gyrosynchrotron radiation. In the framework of the model under discussion the events with radio spikes are expected to be accompanied by relatively weak microwave outbursts. As noted in Section 3.4 [107], the microwave radio spike source cannot emit the observed microwave continuum radiation at the same frequency. The question about the presence or absence of a correlation between decimetre ($f \leqslant 1$ GHz) radio spikes and the microwave continuum ($f \geqslant 1$ GHz) has not been studied in detail, although paper [130] argues that decimetre outbursts with a fine time structure are associated with weaker and more impulsive (with a smaller characteristic flux variability time) microwave outbursts.

3.6 Energetic characteristics of radio spikes
The brightness temperatures of individual spikes can, according to estimates [10, 12, 13], exceed 10^{15} K, which unambiguously suggests a coherent emission mechanism. In paper [123] a special algorithm is proposed to count the number of spikes in a cluster. This algorithm is applied to determine the characteristic flux density per spike. A comparison of X-ray and radio fluxes yields an (averaged) ratio of ~ 0.6 (counts/s)/

sfu, and the division of the radio flux by the appearance rate of spikes yields $0.3 - 30$ sfu/(spike/s). With a typical spike duration of ~ 100 ms, the characteristic radio flux in one spike is $3 - 300$ sfu. With a spectral width of ~ 10 MHz such a value amounts to an energy flux of $(3 - 300)10^{-12}$ erg$\times$cm^{-2}s^{-1}, which corresponds to an energy release in a single spike of the order of $10^{16} - 10^{18}$ erg (assuming isotropic emission). Such an energy is a tiny fraction $(< 10^{-10})$ of the total energy released in a solar flare.

3.7 A radio spike generation model

To find the most likely mechanism for decimetre and microwave radio spike emission, we assume them to represent a uniform group of phenomena and to be formed by one mechanism, which should be capable of explaining all the observational data described in detail above (we emphasize once again that we are considering short narrow-band radio spikes, and not the superfine structure of the entire radio emission). Various theoretical models based on plasma, maser cyclotron, and transition maser mechanisms were discussed in Section 2.

None of the plasma mechanism modifications can explain all the observational data. First of all this concerns the harmonic ratios of radio spikes: the existing plasma models predict either no harmonic structure (as in the transition maser mechanism), or the frequency ratio 1:2 (wave generation near the plasma and twice the plasma frequency). The preferential spike polarization corresponds to X-waves while plasma mechanisms predict the opposite polarization. A nonlinear plasma mechanism seems to be attractive, which could provide a natural explanation to sometimes observed transitions from radio spikes to pulsations. However in this case it is difficult to understand the observed change in individual time profiles of the spikes and the shift of the radio emission source from spike to spike (in addition to the above difficulties in interpreting harmonic structure and polarization). It also remains unclear whether the plasma mechanism can explain the radio spike properties derived from studies of their association with hard X-ray emission.

Despite the fact that several modified plasma models have been proposed (mainly to explain the noninteger harmonic ratios) [131, 132], we should point out serious (and apparently insurmountable) difficulties in interpreting all the variety of radio spike properties with a plasma mechanism.

Oppositely, cyclotron maser emission can explain all the observational data. Indeed, CME mostly generates X-waves but allows polarization sign changing (Section 3.3). The spike time profiles with a Gaussian growth phase naturally result from any initially unstable distribution of fast electrons (Section 3.1). Some observational data can be understood only for a specific primary electron distribution, viz. an anisotropic function with a power-law dependence on momentum. The formation of such non-thermal distributions is a natural consequence of a strong energy release in solar flare and the accelerated electron motion in inhomogeneous magnetic fields.

A model based on a power-law electron distribution naturally produces non-integer harmonic ratios with sufficiently narrow emission spectral lines (section 3.2), and also gives a relation of radio spikes with hard X-ray emission, including time delays and the dependence on the HXR spectral hardness.

Using the latter dependence we can definitely conclude the angular form of the electron distribution function. Note (see

Fig. 19) that the observed radio spikes begin appearing only for sufficiently hard electron spectra with a power-law index $\xi \leqslant \xi_0 \sim 10$ (the power-law index of HXR $\alpha_X \leqslant 6$). The Gaussian angular dependence (78) used in Refs [93 – 95] yields relatively small derivatives with respect to angle and is apparently a poor approximation to the real anisotropy since the corresponding growth rates for $\xi > 8$ are small or negative.

However, an angular dependence like Eqn (46) provides an order of magnitude higher derivatives with respect to angle and can lead to cyclotron instability over the entire range of $\xi \leqslant \xi_0$. For this reason it is necessary to calculate growth rates of the eigen modes of a magnetoactive plasma for a power-law momentum distribution with anisotropy like Eqn (46).

Understanding the spatial characteristics of radio spikes (Section 3.4) requires not only the microscopic emission mechanism to be known, but also a certain knowledge of the plasma emission region itself.

The observed random walk of the radiosource barycentre from spike to spike, along with the fairly broad frequency range occupied by the spike cluster, indicates separate spike sources (with characteristic size $l < 200$ km and $\Delta B/B \leqslant \Delta f/f \sim 0.01 - 0.08$) being distributed (at random in time and space) inside a more extended plasma region (a global trap), which constitutes the spike cluster source. In such a picture, each local source generates a single spike at the moment when favourable conditions for the cyclotron instability development have appeared at the given site. What can these conditions be related to?

In our opinion, a natural consequence of a strong energy release in plasma (for example, in a solar flare) is the formation of chaotic motion (turbulence) leading to plasma parameter fluctuations, in particular, of the magnetic field. It is these magnetic field inhomogeneities that can play the role of the 'local traps'. When particles move in coronal loops with such traps, the angular dependence of their distribution will strongly depend on the local conditions (with a practically unchanged momentum dependence) and in some places the anisotropy can be sufficient for the cyclotron instability to begin.

The harder the electron energy spectrum, the higher the probability of meeting these instability conditions. Note that the 'local trap' model, suggested here to describe the radio spike source spatial distribution, also naturally explains the (apparently independent) increase in the spike number during the hardest X-ray emission phase.

Another problem, posed as early as in the first theoretical papers on CME [76 – 79] in solar corona, deals with the formation of an unstable 'loss-cone' fast electron distribution in a coronal arch. The point is that the typical time of formation of a loss-cone quasi-stationary distribution, due to multiple electron reflections from magnetic mirrors in a trap, exceeds by many orders the characteristic quasi-linear relaxation times of the loss-cone distribution (and the spike duration, respectively). To overcome this difficulty, a single reflection of electrons from the magnetic mirror and the formation of a one-sided (upward-directed) loss-cone were proposed in Ref. [76]. In this case the spike cluster source would be more compact (localized in the lower corona near the foot point) than actually observed.

Another approach to the anisotropy problem leading to the loss-cone instability was proposed by Robinson [133]. In his model, the loss-cone is assumed to form in the arch from an initially isotropic particle distribution. At the beginning,

the anisotropy is small, as are the transverse wave growth rates. The growth rates increase with anisotropy, and the relaxation process starts only after the waves have been amplified to a significant level. The characteristic time of this process is of the order of the mean geometric value between the (big) formation time of a stationary loss-cone (with no account for relaxation) and the (small) relaxation time of the unstable distribution formed. However, this scenario contradicts the observed time profiles of the CME growth phase: calculations [133] imply an over-exponential radio signal growth $s(t) \propto \exp(at^2)$, whereas a slower increase $s(t) \propto \exp[-a(t - t_0)^2]$ is observed [99, 100].

According to the local trap concept, the distribution anisotropy forms on appreciably smaller scales ($\leqslant 10^7$ cm) than the coronal arch size, which takes the time $t_f < 1$ ms. Thus the problem of the 'pumping' time for cyclotron maser emission is also solved here in a natural way.

The main argument against radio spike generation by the cyclotron maser mechanism is based on estimates of the optical depth in gyroresonance layers. The point is that to escape from the solar corona, radio emission should pass through plasma regions in which $B = B_0/s, s = 2, 3, \ldots$, where B_0 is the magnetic field in the radio source. For a Maxwellian plasma, the optical depth can be high even inside the third gyrolayer ($s = 3$), not to mention the second. So even if the CME is generated near the first gyrofrequency harmonic, the corresponding radio emission has been thought to be absorbed in the corona [76, 79] and not be observed on Earth. Such a conclusion appears to be too straightforward. As was noted in one of the first papers [79], in the gyrolayers an addition of fast electrons can be present, which can reduce the absorption coefficient there (or even change its sign). We recall that the solar flare phenomenon (energy release, particle acceleration) occupies a much more extended volume than a radio spike source, so the assumption of non-equilibrium (anisotropically distributed) non-thermal particles being present in the gyrolayers seems to be quite natural.

In addition, a strong wave absorption can be avoided if they propagate at a small angle to the magnetic field [75] (through a longitudinal transparency 'window' [134]). We recall that for $Y \sim 0.3$ the maximum growth rates of X1-waves are reached just along longitudinal directions [87, 93, 94].

In paper [133] it was found that the longitudinal transparency window size (and the probability of exit from the corona) increases (up to 10 times) if strong, of the order of a few percent of the regular field, magnetic field inhomogeneities due to MHD turbulence are present (the same inhomogeneities form the local traps). Evidence for the relation of radio spike characteristics in the $1 - 2$ GHz frequency range with MHD turbulence was found in Ref. [102].

In paper [136] numerical calculations of the CME propagation through the second gyroresonance layer were done. A small fraction (~ 1%) of the emission incident to this layer was found to pass through it due to non-linear effects (non-linear absorption), although the optical depth, estimated in the linear approximation, exceeded 10^3.

We see that the CME generated near the first harmonic (not to mention higher ones) can well pass through the gyroresonance layers and be observed on Earth as radio spikes.

The partial (and not 100%) polarization of radio spikes is apparently also related to effects of the radio wave propagation in corona. An analysis of group delays between radio spike components with opposite circular polarization [109] shows that the wave in the source has 100% polarization, and a partial conversion into a wave with other polarization occurs during its propagation. As shown in Ref. [135], the probability of such a conversion increases (typically by 10 times) if magnetic field inhomogeneities are present in the second gyrolayer.

Thus we conclude that the model of radio spike generation in coronal arches with sufficiently strong magnetic field inhomogeneities by the cyclotron maser mechanism near the 1st and 2nd local gyrofrequency harmonics, assuming a sufficiently hard kinked power-law energy spectrum of nonthermal electrons, adequately describes all the observational data and should be taken as the basis in further theoretical and experimental studies of decimetre and microwave radio spikes.

4. Solar radio spike problems

In spite of much evidence for the cyclotron maser mechanism of solar radio spike emission in the frame of the afforded source model, there are a lot of important unsolved problems.

4.1 Why are only 2% of X-ray bursts associated with radio spikes? The proposed model suggests three possibilities for this: (1) the condition $Y \leqslant 0.3$, which requires a sufficiently strong magnetic field inside the loop, is rarely met; (2) a hard spectrum of electrons with a kink in the range $200 - 300$ keV forms rarely; (3) sufficiently strong MHD turbulence, which forms the local electron distribution anisotropy, also appears rarely. It is unclear at present which of these conditions is determining.

The role of the second conditions could be tested using already existing data on X-ray bursts. In spite of the systematic differences between X-ray bursts, which may or may not be associated with radio spikes, were established long ago (radio spikes relate to shorter than average X-ray bursts [116]), the role of spectral properties of X-ray emission has been studied quite superficially. In particular, the principal conclusion that the spikes appear on the harder X-ray burst phase is made in Ref. [123] in a figure caption and is not discussed at all in the text. In this connection it seems necessary to make a dedicated study of the differences and the probability of the spike appearance on the decay phase of HXR and gamma-ray emissions separately for bursts characterized by spectral softening and hardening in the decay phase (the well-known 'soft-hard-soft' and 'soft-hard-harder' spectral behaviour). It is quite obvious that to solve this problem, a comparative analysis of the spikes and continuum microwave emission in optically thin frequency range (cm-mm wavelengths), which are known to reflect adequately the spectral behaviour of electrons with energies of several hundred keV captured in the flare loop, would be very desirable.

One more possibility for the spike observation probability decrease cannot be excluded. This comes from the CME-model assuming the presence of some special directions of electromagnetic emission relative to the flare loop magnetic field. Under some conditions the beaming can be very narrow. This strongly decreases the observation probability of spike sources located in the loop such that the line of sight is at an unfavourable angle with respect to the magnetic field. The spike emission beaming could be derived and studied

from VLBI observations coordinated with other high angular resolution observations of flare loops (X-ray, UV, optical, and microwave continuum).

4.2 Another principally important question is whether the solar radio spikes form a homogeneous group. We recall that the radio spike properties discussed in Section 3 were obtained from the analysis either of a part of the existing data (for example, 20 – 30 clusters of more than 150 registered in Zurich), or of individual events, as in the case of observations with a high spatial resolution. Thus most events have not been systematically analysed for some reason. So we believe that at present the possibility of division of the spikes into subclasses cannot be excluded. In this case the radio spikes from some subclasses could be generated by some other emission mechanisms, along with the cyclotron maser mechanism, for example, by a plasma or transition maser mechanism.

So further theoretical studies of all three emission mechanisms appear very important. CME studies must use distribution functions in a form as close as possible to reality. While its dependence on momentum is clear from hard X-ray spectra (it is a kinked-power-law function), the functional form of its angular dependence from independent studies is unknown. It must be taken in accordance with the dependence of the spike appearance on spectral hardness (see Fig. 19). On the other hand, the angular dependence of a non-thermal electron distribution could be found theoretically if the magnetic turbulence structure in the flare region is reliably established. So the development of an adequate turbulence theory in solar flares (based upon observations) is necessary to better understand the process of radio spike generation.

The stabilization of cyclotron instability by quasi-linear relaxation must be accurately studied with the use of realistic initial distribution functions. A non-thermal particle source, as well as non-linear wave interactions [27], must necessarily be taken into account in these calculations (the injection function time profile should be determined using hard X-ray emission data). Note that the observed negative frequency drift of the spikes may be due to stimulated wave scattering on plasma particles leading to energy transport towards lower frequencies.

Studies of the non-linear plasma mechanism must also be continued. The irregular (spike-like) pulsations and their transition to normal periodical oscillations (and vice versa) are of special interest here. In this problem, the finite duration of particle injection and the wave back reaction on particles can produce a notable effect. Even if all radio spikes are generated by the same cyclotron maser mechanism, the non-linear plasma mechanism provides some other types of superfine structure of solar radio emission, so studies of it are important independently of whether the spikes constitute a homogeneous group or not.

4.3 To date, not all the possibilities of processing and analysis of existing data have been used. We wish to emphasize several important points that could be clarified without involving additional observations (i.e. before the next solar activity maximum).

As noted above (Section 3.2), the emission line width does not depend on the macroscopic parameters of the emission region (e.g., on the emission frequency). The CME theory predicts this value to be mostly related to the electron distribution form. So it is natural to expect that a correlation between the radio line width and hard X-ray emission

properties (in particular, the spectral ones) exists. The effect of wave scattering on small-scale inhomogeneities along the wave propagation direction towards the observer on the spectral line width should cause the value of Γ to vary from the disk centre to the limb and also to correlate with the spike polarization degree. Unfortunately, such an analysis has not been made as yet.

4.4 Properties of harmonic emission bands of the spikes (in the events with harmonic structure) are not known in detail, in particular, there are no answers to the following questions. What is the polarization of different harmonics? What are their characteristic durations, do they satisfy the general law (84), (85)? Could different harmonics in one event have opposite polarizations? Are there differences in the disk distributions of events with and without harmonic structure (i.e. is there correlation between the harmonics appearing in emission with a certain direction of the magnetic field with respect to the line of sight)? Do the harmonic structure events occur in microwave wavelengths?

4.5 The correlation between the radio spikes and microwave continuum and soft gamma-ray emission has been much less analysed than in the case with hard X-ray emission. We stress once again that the spectral properties of continuum emissions (related to non-thermal electron spectra) play a very important, if not the key role here.

4.6 Nevertheless we assume that 'crucial' experiments will be done at a new technical level during the coming 23rd solar activity maximum. Radio observations can provide us with essentially new information if conducted simultaneously with high spatial, frequency, and time resolution. This goal can be achieved both by modifying observational methods on individual instruments and by conducting coordinated simultaneous observations on different instruments. Presently, the first of these possibilities is under study at several observatories [137 – 142].

Big progress has been made in modifying the possibilities of the Siberian solar radiotelescope (SSRT, Irkutsk). A multiprocessor registration system for rapid bursts has been created, which enables one-dimensional scans of radio bursts at a frequency of 5.7 GHz with a spatial resolution of $\sim 15''$ and a time resolution up to 14 ms to be obtained. At present, observations accompanied by constructing two-dimensional images of radio sources are under way [112]. The construction of a spectropolarimeter in the range 4 – 8 GHz as a part of the project of the SSRT instrument complex modification [142] is also important for the radio spike problem. The range 4 – 8 GHz is planned to be subdivided into 120 MHz channels, and an acoustic optical spectral analyser is proposed to be used in the SSRT registration band.

In the Scientific Research Institute of Radio Physics (the city of Nizhniĭ Novgorod), interferometric observations of solar flares in decimetre wavelengths with a time resolution of 0.256 ms [138, 139] have begun. The interferometric system includes large Russian dishes (RT-64, Medvezh'i Ozera and Kalyazin; RT-22, Pushchino; RT-70, Ussuriĭsk), as well as small radio telescopes of Scientific Research Institute of Radio Physics at the radio observatories 'Staraya Pustyn'' (2 RT-14, 2 RT-7) and 'Znamenki' (2 RT-15). The distances between the radio telescopes provide practically any angular resolution necessary for spike studies, from fractions of arcseconds to tens of arcminutes. The bases of this system

have different orientations, which allows, in the case of simultaneous observations, the source sizes in at least two directions to be obtained and two-dimensional positional measurements to be made. In addition, the program envisages plasma inhomogeneity studies at low altitudes in the solar corona $\leqslant 2R_s$, which are important for the spike emission scattering problem. Note in this connection that for small elongations the VLBI method has more advantages than the widely used scintillation method [139].

The transformation of the biggest Russian radiotelescope RATAN-600 into a multiwave radio heliograph with a two-dimensional resolution of several arcseconds and a time resolution down to a few milliseconds is beginning. For this purpose, a non-sinphase irradiation regime is proposed when the main beam is divided into several hundred independent beams almost randomly distributed in space [140, 141]. A regime of simultaneous multi-channel registration at several frequencies separated by a few MHz is planned. Model calculations show that a special correlation analysis of multi-beam observations yields a sharp image of a point-like source. The method allows one to obtain almost instantaneous radio images at several frequencies currently used on RATAN-600.

The corresponding observations with a high spatial, spectral, and time resolution will allow detailed study of the form of spectral lines (including harmonic emission bands), their dynamics, and the time profiles of spikes with different intensity at different frequencies, the determination of the location, size, dynamics, and proper motion of the spike source, the determination of spike and the corresponding continuum distribution in the flare loop (loops), etc. Direct measurements of the brightness temperature of radio spikes and the corresponding continuum can be obtained in this way. It is not easy but very important to determine ω_{pe} and ω_{Be} in the flare loop by several independent means.

Very valuable information can be obtained from simultaneous radio and hard X-ray high-resolution observations. As noted in paper [123], in many events some X-ray peaks very tightly correlate with the spike radio emission, whereas other peaks are not associated with radio emission at all. From here a conclusion was drawn that non-correlated X-ray peaks are generated in other loops, however no direct evidence for such a model has been found. It is not excluded that the reason for this lies not in the different sources, but in the nature of the radio spike generation process itself. High-resolution HXR observations will enable us to make a choice between these two possibilities.

We arrive at the conclusion that the radio spike generation problem and the emission source model can be definitely found in the near future, which will allow us to use the millisecond radio spike phenomenon as a sensitive tool for plasma diagnostics both inside the source of emission and along its propagation in the corona.

5. Conclusions

In the present review, we have tried to analyse all currently known observational data on millisecond solar radio spikes and to compare them with detailed predictions of the existing theoretical models. The analysis enabled us to significantly decrease the number of models capable of reproducing all variety of these properties.

The energy source for radio spike emission is non-thermal mildly relativistic electrons with a hard power-law spectrum

in the energy range from tens to several hundred keV and a drop (steepening) at $E \geqslant 300$ keV. The angular part of the distribution function (whose exact functional dependence has so far been unknown) fluctuates due to electron motions in a coronal arch with magnetic field inhomogeneities. In some regions of the arch (local traps), the distribution anisotropy proves to be sufficient for a cyclotron instability to develop. The harder the electron energy spectrum, the higher the probability of meeting the instability conditions, so most spikes are generated on the hardest particle spectrum phase (and, correspondingly, of the X-ray emission). Thus the radio spikes appear not due to the fragmentation of a primary energy release, as was thought earlier [12], but due to the 'secondary' fragmentation, i.e. due to the presence of sufficiently strong magnetic field inhomogeneities in the loop, although these inhomogeneities themselves seem to directly result from the primary energy release.

It has been widely recognized that the radio spikes have a high diagnostic potential. Nevertheless the realization of such a potential is not an easy task. As an example we note that the CME instability regions may be strongly shifted relative to the lowest gyrofrequency harmonics (ω_{Be} and $2\omega_{Be}$), so even for events with harmonic structure it appears difficult to find an exact value of the magnetic field inside the spike source (but is possible within a factor of order 2).

Theoretical and observational problems to be solved in the near future are described in detail in the previous section. Undoubtedly, these studies will significantly advance the possibilities of solar plasma diagnostics using observed radio spike parameters, such as the frequency, width and form of spectral lines; sizes, locations, dynamics of the source; emission time profiles, etc.

Note that a detailed study of the superfine structure of solar radio emission is also important in a broader astrophysical aspect. The fine structure of radio emission is observed for other stars, in particular, for eruptive red dwarfs [144]. Since much less is directly known about remote stellar atmospheres than that of the Sun, sun-stellar analogues can in some cases be very fruitful in constructing models for the rapidly varying component of stellar radio emission and the subsequent plasma diagnostics in stellar coronas.

References

1. Droge F, Riemann P *Int. Bull. Solar Radio Obs. Europe* **8** 6 (1961)
2. De Groot T *Int. Bull. Solar Radio Obs. Europe* **9** 3 (1962)
3. Elgaroy O *Astrophys. Norv.* **7** 235 (1961)
4. Malville J M, Aller H D, Jensen C J *Astrophys. J.* **147** 711 (1967)
5. Droge F *Z Astrophys.* **66** 176 (1967)
6. Barrow C H, Saunders H *Astrophys. Lett.* **12** 211 (1972)
7. Droge F *Astron. Astrophys.* **57** 285 (1977)
8. Slottje C *Nature* (London) **275** 520 (1978)
9. Slottje C *Atlas of fine structures of dynamical spectra of solar type IV-dm and some type II radio bursts* (Dwingeloo Observatory, 1981)
10. Stali M, Magun A *Solar Phys.* **104** 117 (1986)
11. Benz A O et al. *Astron. Astrophys. Suppl.* **93** 539 (1992)
12. Benz A O *Solar Phys.* **96** 357 (1985)
13. Benz A O *Solar Phys.* **104** 99 (1986)
14. Gudel M, Benz A O *Astron. Astrophys. Suppl.* **75** 243 (1988)
15. Ginzburg V L Zheleznyakov V V *Astron. Zh.* **35** 694 (1958)

16. Aschwanden M J et al. *Solar Phys.* **130** 39 (1990)
17. Benz A O et al. *Solar Phys.* **141** 335 (1992)
18. Genkin L G , Erukhimov L M, Levin B N *Pis'ma Astron. Zh.* **15** 167 (1989) [*Sov. Astron. Lett.* **15** 71 (1989)]
19. Genkin L G, Erukhimov L M, Levin B N *Solar Phys.* **128** 423 (1990)
20. Tsytovich V N *Nelineĭnye Effekty v Plazme* (Nonlinear Effects in Plasma) (Moscow: Nauka, 1967) [Translated into English (New York: Plenum Press, 1970)]
21. Pustovalov V V, Silin V P *Trudy FIAN* **61** 42 (1972)
22. Kadomtsev B B *Kollektivnye Yavleniya v Plazme* (Collective processes in plasma) (Moscow: Nauka, 1988)
23. *Elektrodinamika Plazmy* (Plasma Electrodynamics) (Ed. A I Akhiezer) (Moscow: Nauka, 1974) [Translated into English (International Series of Monographs in Natural Philosophy, Vol. 68) (Oxford: Pergamon Press, 1975)]
24. Sizonenko V L, Stepanov K N *Zh. Eksp. Teor. Fiz.* **49** 1197 (1965) [*Sov. Phys. JETP* **22** 832 (1966)]
25. Rowlands J, Shapiro V D, Shevchenko V I *Zh. Eksp. Teor. Fiz.* **50** 979 (1966) [*Sov. Phys. JETP* **23** 651 (1966)]
26. Hamilton R J, Petrosian V *Astrophys. J.* **365** 778 (1990)
27. Fleishman G D *Solar Phys.* **153** 367 (1994)
28. Zheleznyakov V V *Radioizluchenie Solntsa i Planet* (Radio Emission of the Sun and Planets) (Moscow: Nauka, 1964) [Translated into English (International Series of Monographs in Natural Philosophy, Vol. 25) (Oxford: Pergamon Press, 1969)]
29. Zheleznyakov V V, Zaitsev V V *Astron. Astrophys.* **39** 197 (1975)
30. Vlahos L, Sharma R R, Papadopoulos K *Astrophys. J.* **275** 374 (1983)
31. Melrose D B *Astrophys. J.* **380** 256 (1991)
32. Zaitsev V V, Stepanov A V *Studies on geomag., aeron., and physiscs of Sun* **37** 11 (1975)
33. Zaitsev V V *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* **13** 837 (1970)
34. Zaitsev V V, Stepanov A V *Solar Phys.* **88** 297 (1983)
35. Stepanov A V *Studies on geomag., aeron., and physiscs of Sun* **71** 210 (1985)
36. Bardakov V M, Stepanov A V *Pis'ma Astron. Zh.* **5** 460 (1979) [*Sov. Astron. Lett.* **5** 247 (1979)]
37. Stepanov A V *Studies on geomag., aeron., and physiscs of Sun* **50** 191 (1980)
38. Stepanov A V *Studies on geomag., aeron., and physiscs of Sun* **54** 141 (1980)
39. Mazur V A, Stepanov A V *Publ. Debrecen Heliophys. Obs.* **5** 419 (1983)
40. Rozenraukh Yu M, Stepanov A V *Astron. Zh.* **65** 300 (1988) [*Sov. Astron.* **32** 153 (1988)]
41. Aschwanden M J, Benz A O *Astrophys. J.* **332** 466 (1988)
42. Aschwanden M J *Astrophys. J.* **416** 857 (1993)
43. Aschwanden M J *Solar Phys.* **111** 113 (1987)
44. Korsakov V B, Fleishman G D *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* **41** 46 (1988)
45. Stepanov A V *Astron. Zh.* **50** 1243 (1973)
46. Zaitsev V V, Stepanov A V *Astron. Astrophys.* **45** 135 (1975)
47. Benz A O, Kuijpers J *Solar Phys.* **46** 275 (1976)
48. Fleishman G D, Stepanov A V, Yurovsky Yu F *Solar Phys.* **153** 403 (1994)
49. Fleishman G D, Stepanov A V, Yurovsky Yu F *Space Sci. Rev.* **68** 205 (1994)
50. Harris E G *Phys. Rev. Lett.* **2** 34 (1959)
51. Sagdeev R Z, Shafranov V Ts *Zh. Eksp. Teor. Fiz.* **39** 181 (1960) [*Sov. Phys. JETP* **12** 130 (1961)]
52. Benediktov E A et al. *Kosmich. Issled.* **3** 614 (1965)
53. Croley D R, Mizera P F (Jr.), Fennel J F *J. Geophys. Res.* **83** 2701 (1978)
54. Wu C S, Lee L C *Astrophys. J.* **230** 621 (1979)
55. Lee L C, Kan J R, Wu C S *Planet. Space Sci.* **28** 703 (1980)
56. Omidi N, Gurnett D A *J. Geophys. Res.* **87** 2377 (1982)
57. Melrose D B, Ronnmark K G, Hewitt R G *J. Geophys. Res.* **87** 5140 (1982)
58. Dusenbery P B, Lyons L R *J. Geophys. Res.* **87** 7467 (1982)
59. Hewitt R G, Melrose D B, Dulk G A *J. Geophys. Res.* **88** 10065 (1983)
60. Pritchett P L, Strangeway R J *J. Geophys. Res.* **90** 9650 (1985)
61. Winglee R M, Pritchett P L *J. Geophys. Res.* **91** 13531 (1986)
62. Pritchett P L, Winglee R M *J. Geophys. Res.* **94** 129 (1989)
63. Calvert W *Geophys. Res. Lett.* **8** 919 (1981)
64. Benson R F *Geophys. Res. Lett.* **9** 1120 (1982)
65. Gurnett D A, Shawhan S D, Shaw R R *J. Geophys. Res.* **88** 329 (1983)
66. Wu C S *Space Sci. Rev.* **41** 215 (1985)
67. Zarka P, Le Queau D, Genova P *J. Geophys. Res.* **91** 13542 (1986)
68. Louarn P et al. *J. Geophys. Res.* **95** 5983 (1990)
69. Ryabov B P, Gerasimova N N *Dekametrovoe Sporadicheskoe Radioizluchenie Yupitera* (Decametre Sporadic Emission from Jupiter) (Kiev: Naukova Dumka, 1990)
70. Ladreiter H P, Leblanc Y *J. Geophys. Res.* **95** 6423 (1990)
71. Galopeau P, Zarka P, Le Queau D *J. Geophys. Res.* **94** 8793 (1989)
72. Winglee R M, Menietti J D, Wong H K *J. Geophys. Res.* **97** 17131 (1992)
73. Pedersen B M et al. *J. Geophys. Res.* **97** 19201 (1992)
74. Stepanov A V *Pis'ma Astron. Zh.* **4** 193 (1978) [*Sov. Astron. Lett.* **4** 103 (1978)]
75. Holman G D, Eichler D, Kundu M R, in *Radiophysics of the Sun* (Eds M R Kundu, T E Gergely) (Dordrecht: Reidel, 1980) p. 457
76. Melrose D B, Dulk G A *Astrophys. J.* **259** 844 (1982)
77. Sharma R R, Vlahos L, Papadopoulos K *Astron. Astrophys.* **112** 377 (1982)
78. Ladreiter H R *Astrophys. J.* **370** 419 (1991)
79. Sharma R R, Vlahos L *Astrophys. J.* **280** 405 (1984)
80. Dory R A, Guest G E, Harris E G *Phys. Rev. Lett.* **14** 131 (1965)
81. Winglee R R, Dulk G A *Solar Phys.* **104** 93 (1986)
82. Louarn P, Le Queau D, Roux A *Solar Phys.* **111** 201 (1987)
83. Winglee R R, Dulk G A, Pritchett P L *Astrophys. J.* **328** 809 (1988)
84. Winglee R M *J. Geophys. Res.* **90** 9663 (1985)
85. Hewitt R G, Melrose D B, Ronnmark K G *Aust. J. Phys.* **35** 447 (1982)
86. Melrose D B, Hewitt R G, Dulk G A *J. Geophys. Res.* **89** 897 (1984)
87. Aschwanden M J, Benz A O *Astrophys. J.* **332** 447 (1988)
88. Aschwanden M J *Astron. Astrophys. Suppl.* **85** 1141 (1990)
89. Li H W *Solar Phys.* **104** 131 (1986)
90. Wu C S, Zhou G C, Gaffey J D *Phys. Fluids* **28** 846 (1985)
91. Wu C S et al. *Astrophys. J.* **248** 384 (1981)
92. Pritchett P L *Phys. Fluids* **29** 2919 (1986)
93. Fleĭshman G D, Yastrebov S G *Astron. Zh.* **71** 531 (1994) [*Astron. Rep.* **38** 468 (1994)]
94. Fleishman G D, Yastrebov S G *Solar Phys.* **153** 389 (1994)
95. Fleishman G D, Yastrebov S G *Solar Phys.* **154** 361 (1994)
95a. Arzner K, Fleishman G D *Solar Phys.* (1999) (in preparation)
96. Platonov K Yu, Fleĭshman G D *Zh. Eksp. Teor. Fiz.* **106** 1053 (1994) [*JETP* **79** 572 (1994)]
97. Platonov K Yu, Fleĭshman G D *Zh. Eksp. Teor. Fiz.* **108** 1942 (1995) [*JETP* **81** 1059 (1995)]
98. Platonov K Yu, Fleĭshman G D *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* **40** 941 (1997)
99. Gudel M, Benz A O *Astron. Astrophys.* **231** 202 (1990)
100. Mel'nikov V F et al. *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* (in press) (1998)
100a. Isliker H *Solar Phys.* **141** 325 (1992)
101. Csillaghy A, Benz A O *Astron. Astrophys.* **274** 487 (1993)
102. Karlicky M, Sobotka M, Jiricka K *Solar Phys.* **168** 375 (1996)
103. Benz A O, Gudel M *Solar Phys.* **111** 175 (1987)
104. Gudel M *Astron. Astrophys.* **239** L1 (1990)
105. Krucker S, Benz A O *Astron. Astrophys.* **285** 1038 (1994)
106. Gudel M, Zlobec P *Astron. Astrophys.* **245** 299 (1991)
107. Gary D E, Hurford G J, Flees D J *Astrophys. J.* **369** 255 (1991)
108. Krucker S et al. *Astron. Astrophys.* **302** 551 (1995)
109. Benz A O, Pianezzi P *Astron. Astrophys.* **323** 250 (1997)
110. Tapping K F et al. *Astron. Astrophys.* **122** 177 (1983)
111. Tapping K F *Solar Phys.* **104** 199 (1986)
112a. Altyntsev A T et al. *Astron. Astrophys.* **303** 249 (1995)
112b. Altyntsev A T et al. *Astrophys. J.* **469** 976 (1996)
112c. Benz A O *Abstract of CESRA workshop* (Helsinki: Univ. Techn., 1998) p. 16
113. Benz A O et al. *Astron. Astrophys.* **305** 970 (1996)
114a. Alekseev V A et al. *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* **40** 1063 (1997)

114b. Alekseev V A et al. *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* (in press) (1998)

115. Bastian T S *Astrophys. J.* **426** 774 (1994)

116. Benz A O, Kane S R *Solar Phys.* **111** 179 (1987)

117. Aschwanden M J, Wiehl H J, Benz A O, Kane S R *Solar Phys.* **97** 159 (1985)

118. Kane S R *Astrophys. J.* **247** 1113 (1981)

119. Kane S R, Pick M, Raoult A *Astrophys. J.* **241** L113 (1980)

120. Kosugi T *Solar Phys.* **71** 91 (1980)

121. Aschwanden M J, Benz A O, Kane S R *Astron. Astrophys.* **229** 206 (1990)

122. Gudel M, Aschwanden M J, Benz A O *Astron. Astrophys.* **251** 285 (1991)

123. Aschwanden M J, Gudel M *Astrophys. J.* **401** 736 (1992)

124. Vilmer N, Kane S R, Trottet G *Astron. Astrophys.* **108** 306 (1982)

125. Lu E T, Petrosian V *Astrophys. J.* **327** 405 (1988)

126. Bruggmann G, Vilmer N, Klein K-L, Kane S R *Solar Phys.* **149** 171 (1994)

127. Takakura T *Solar Phys.* **104** 363 (1986)

128. Mel'nikov V F, Magun A *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* **39** (1456) (1996)

129. Melnikov V F, Magun A *Solar Phys.* **178** 153 (1997)

129a. Zlobec P, Karlicky M *Abstract of CESRA workshop* (Helsinki: Univ. Techn.,1998) p. 69

130. Mel'nikov PhD Thesis (Gorky: NIRFI, 1990)

131. Wills A J, Robinson P A *Astrophys. J.* **467** 465 (1996)

132. Stepanov A V et al. *Solar Phys.* **176** 147 (1997)

133. Robinson P A *Solar Phys.* **134** 299 (1991)

134. Alikaev V V et al., in *Vysokochastotnyĭ Nagrev Plasmy. Materialy Vsesoyuzn. Soveshch.* (High-frequency plasma heating. Proc. All-Union Conf.) (Ed. A G Litvak) (Gorky, 1983) p. 6

135. Robinson P A *Solar Phys.* **136** 343 (1991)

136. McKean M E, Winglee R R, Dulk G A, in *Proc. of 3th Workshop Max'91 Solar Flares* (Eds. R R Winglee, A L Kiplinger) (Colorado: Boulder, 1990) p. 29

137. Altyntsev A T et al. *Solar Phys.* **168** 145 (1996)

138. Alekseev V A et al. *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* **38** 1046 (1995)

139. Alekseev V A et al., in *Problems of Modern Radioastronomy. Proc. 27th radioastron. conf.* Vol. 2 (S.-Petersbourgh, 1997) p. 178

140. Bogod V G et al. *Izv. Vyssh. Uchebn. Zaved. Radiofiz.* **39** 527 (1996)

141. Bogod V G et al., in *Problems of Modern Radioastronomy. Proc. 27th radioastron. conf.* Vol. 1 (S.-Petersbourgh, 1997) p. 31

142. Altyntsev A T, Zandanov V G, Rosenraukh Yu M, in *Problems of Modern Radioastronomy. Proc. 27th radioastron. conf.* Vol. 3 (S.-Petersbourg, 1997) p. 168

143. Levin B N, Melnikov V F *Solar Phys.* **148** 325 (1993)

144. Bastian T S et al. *Astrophys. J.* **353** 265 (1990)

145. Altyntsev A T, Grechnev V V, Hanaoka Y, Preprint NRO No.97-07 (1997) Mitaka, Tokyo 181, Japan, 1997

146. Fu Q J, Huang G L *Abstract of CESRA workshop* (Helsinki: Univ. Techn., 1998) p. 75