A cyclotron resonance mechanism for very-low-frequency whistler-mode sideband wave radiation. I. Introduction and description of external resonances

L. A. D. Sá
Space, Telecommunications and Radioscience Laboratory, Stanford University, Stanford, California 94305

(Received 8 December 1988; accepted for publication 23 June 1989)

The first part of the paper introduces the reader to the physical and mathematical aspects of the problem of sideband formation associated with very-low-frequency whistler-mode waves propagating in ducts in the magnetosphere. It starts with an introduction to the general problem of charge distribution in the duct and looks at the effect of monochromatic electromagnetic waves acting on such distributions. It describes sidebands as due to wave-wave interactions occurring in the magnetosphere through electron cyclotron resonances, and concentrates on sideband radiation coming from resonances located outside the wave potential wells (external resonances). It also presents a general solution to the electron equation of motion, based on the Kolmogorov–Arnold–Moser theorem, that lists the possible sideband radiation frequencies. Both an analytical treatment, based on Lie transform perturbation theory, and a numerical treatment, based on phase plots, are used in the detailed part of the analysis. Effects of the inhomogeneity of the magnetic field of the Earth are mainly neglected in the first part of the paper.

I. INTRODUCTION

The purpose of this three-part paper is to provide a comprehensive description of the processes leading to the formation of sideband waves associated with discrete frequency transmissions originated in the very-low-frequency (VLF) transmitter located at Siple Station, Antarctica and received at Lake Mistissini, Canada after ducted wave propagation. Those sidebands are found both in single- and double-frequency transmissions. In a recent paper,[1] we have shown that "single-frequency" sidebands are due to the combined effect of a transmitted wave from Siple and radiation present in the duct at multiples of 60 Hz probably, but not certainly, coming from the Canadian power distribution system. This allowed us to understand the VLF sideband process as being always caused by the effect on the magnetospheric electrons of several discrete frequency carriers that, disturbing the electron motion, change the electron distribution in such a way that radiation always occurs at certain well-defined frequencies, given by a simple formula [see Eq. (34)].

This paper relies on nonlinear mechanics formalism and methods that may be unfamiliar to the reader. They are, however, essential to the problem we will treat, as can be shown by some simple considerations: The sideband process in the magnetospheric plasma is a combined effect of accelerated charge radiation, regulated by Maxwell's equations, and charge acceleration, described by the Lorentz force equation. Maxwell's equations are linear and predict radiation only at the frequencies at which charges oscillate. Sidebands clearly contain radiation at frequencies different from the incoming wave frequencies. Nonlinear forces, therefore, must be at play, producing accelerations at frequencies different from those present in the incoming wave field. The solution of the Lorentz force equation that describes such an effect is an exercise in mechanics, and is part of the effort of this paper. An understanding of the radiation coming from the accelerated charge distributions complements it. The results are fruitful: the main aspects of the radiation described by Sá and Helliwell[1] are easily obtained. The reader that wants an overview of nonlinear oscillator systems together with references to recent research on the subject should consult Lichtenberg and Lieberman.[2]

In Sec. II of this paper we will show how some simple symmetry properties of the duct in the absence of an external wave field imply the existence of a nonradiating electron distribution. In Sec. III we show how an injected electromagnetic field, breaking the symmetry, creates radiation. In Sec. IV, the equations governing the electron motion under the influence of one or more right-hand circularly polarized (RHCP) whistler-mode waves are introduced and an interpretation of their solutions is made. In Sec. V, the equations of motion are solved numerically and the solutions are displayed with the help of phase plots. Emphasis is put on effects occurring outside the potential well of each incoming wave and, more specifically, on the occurrence of resonances of the electron motion at the frequencies where radiation is observed (external resonances). An example of "bunching" produced by those resonances is presented. Section VI begins the analytical treatment of the wave interaction process, with the application of the Kolmogorov–Arnold–Moser (KAM) theorem for nonlinear systems to the interaction we are describing. As a result we obtain an extremely general and simple expression for the possible frequencies of the sideband waves. Section VII describes a perturbative method based on Lie transforms which is used for a detailed study of the electron motion. Section VIII applies the Lie perturbation method to nonresonant electrons. Section IX describes the lowest-order resonance, the half harmonic, a second-order effect produced by the interaction of any wave pair in the incoming wave system. Section X describes third-order effects associated with two or three interacting waves. Section

 0021-8979/89/203482-13$02.40

XI contains a description of higher-order effects, and Sec. XII the conclusions.

II. DUCT SYMMETRIES AND NONRADIATING ELECTRON CONFIGURATIONS

We will assume that in the absence of outside incoming radiation the electron distribution is so stirred up that all points in a perpendicular cross section of the duct will have the same electron density. Moreover, we will assume that this local electron distribution is isotropic, that is, the number of electrons observed moving in any direction is the same, independently of the direction chosen. At each point, the observed electrons will also have different velocities of motion in the direction parallel to the axis of the duct, $v_\parallel$. We will have to assume that the distribution is homogeneous and isotropic for each $v_\parallel$, otherwise by selecting a specific $v_\parallel$ we would be able to define a specific direction in the duct, contrary to the assumption that all directions are equally important. This assumption of homogeneity and isotropy will provide us with the means to build the electron distribution in the duct.

Figure 1 shows a perpendicular cross section of the duct with some electrons singled out. Those electrons will be moving in a helical orbit due to the presence of the magnetic field of the Earth. Let the projected motion of all electrons be counterclockwise. We will show that every electron in the distribution belongs to a direct current (dc) loop. Suppose that at point P_1 we consider a volume element $dA\ dl$ of electrons moving in the direction and sense defined by the arrow in the picture. Let the circle in the picture be the projection of their helical orbit on the plane of the cross section. At another point P_2, arbitrarily located on that circumference, we draw a volume element identical to $dA\ dl$. Due to the homogeneity of the duct the number of electrons in this volume, dN_2, will be equal to dN_1. If we now draw a third volume element $dA\ dl$ with dl tangential to the circumference, the number of electrons, dN_3, moving in it and having the circumference as their projected trajectory will be equal to dN_2 due to the isotropy of the duct. Therefore, $dN_1 = dN_3$. Since P_2 is arbitrary, we conclude that the circumference we have drawn is uniformly populated by electrons, and moves as a rigid body in the direction of the Earth's magnetic field with velocity $v_\parallel$.

If we assume that the duct is in stationary equilibrium (the distribution is translationally invariant in time), the plane of the cross section will be continually crossed by those dc loops, identical to each other, indicating that the distribution is in reality made up of cylinders with constant charge density on their surfaces and moving with velocity $v_\parallel$ as in Fig. 2. Such a charge configuration will not radiate due to its rotational symmetry. Symmetry considerations, however, cannot constrain the value of the surface charge density to be the same for cylinders with different $v_\parallel$, or different radii. In reality, in a real electron distribution, the cylinders will have surface charge varying as a function of their radii and $v_\parallel$. This is a very important fact since we will see that a gradient in the surface charge density as a function of $v_\parallel$ is necessary for the radiation due to the sum of all cylinders to be different from zero when an electromagnetic wave interacts with the electron distribution. As electromagnetic radiation reaches each cylinder, the isotropy of the forces will be broken by the vector fields of the wave, and the electron distribution will lose its isotropy and radiate.

III. EFFECT OF RADIATION ON THE ELECTRON DISTRIBUTION

To simplify our analysis of the radiation generation process, we assume that the interaction between the electron distribution and the incoming waves occur in a region with sharply defined boundaries, and that the incoming electromagnetic radiation is a superposition of monochromatic whistler-mode waves. The region of interaction, located around the equator, will have dimensions such that inside it, the inhomogeneity of the magnetic field of the Earth can be neglected. It will be extremely convenient to observe the interaction from a reference system with no translational motion, rotating around the axes of the charge cylinders with an angular velocity ω_1, equal to the frequency of one of the waves, and having its z axis pointing in the direction opposite to $v_\parallel$, the velocity of the electrons resonating with the electromagnetic field.

Figure 3 shows the described situation, and our approach to the analysis of the problem: We imagine each cylinder to be composed, outside the interaction region, of a sequence of dc loops. As each of those current loops enters and propagates through the interaction region, it will be deformed by the forces of the incoming electromagnetic waves acting on it. Since vector fields define preferred directions in space, the distortions will be such that the isotropy of the cylindrical charge distribution will be lost. As a consequence, each cylinder will now be made up of a sequence of alternating current (ac) loops, and radiation will start to be emitted. (An accurate analysis shows that, due to the way

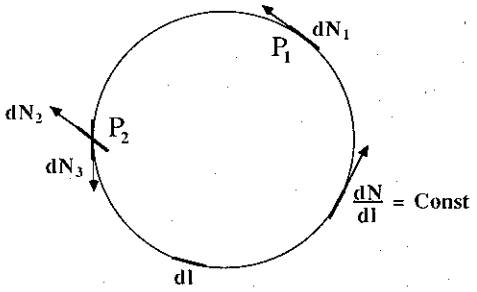

FIG. 1. Duct isotropy and homogeneity. Due to symmetries in the duct geometry, the electron distribution in a perpendicular cross section of the duct can be broken down into a set of dc current loops.

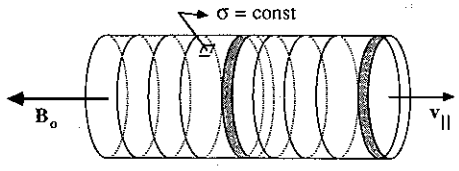

FIG. 2. Undisturbed electron distribution. Because the duct is in stationary equilibrium, each loop belongs to a cylinder with constant surface charge density, σ, which propagates along $-z$ with velocity $v_\parallel$.

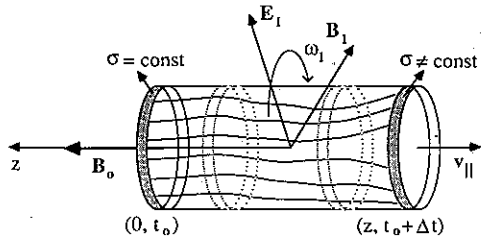

FIG. 3. Effect of electromagnetic field on electron distribution. As each cylinder penetrates the interaction region, forces due to the electromagnetic field change its charge distribution, breaking its rotational symmetry and creating radiation. A ring obtained by slicing the cylindrical charge at $z=0$ will have constant charge density, another one at $z\neq 0$ will not. The wavy horizontal lines represent schematically the electron streams into which the cylindrical charge distribution can also be decomposed. With the conventions used in this paper, z is negative in the picture.

the electrons and the electromagnetic field interact, each loop is distorted in the z direction as it propagates along the interaction region. The ac loops that make up the distorted cylinder are not simply individually distorted dc loops, but a combination of parts of several of them, each one having entered the interaction region at a slightly different time.)

If the incoming radiation is monochromatic, with frequency ω_1, and if the reference system rotates with the same frequency ω_1, the whistler-mode wave will have a static configuration. All current loops will see the same forces and will have the same distortions as they reach the same points in the interaction region. The distorted charge configuration will be independent of time, will behave as a rigid body with an asymmetrical charge distribution rotating with angular velocity ω_1, and will radiate at the same frequency as the incoming monochromatic radiation. The emitted radiation will in general have variable phase and variable amplitude (especially in the initial transient stage when the wave first meets the electron distribution). This might give rise to some spectral broadening, but no well-defined sidebands.

For two or more waves, even after we adjust our reference system to rotate with one of them, the resultant electromagnetic field will be explicitly time dependent. Each loop as it penetrates the interaction region will see a different electromagnetic field configuration. At each point the charge densities on the surface of the cylindrical electron distribution will be time dependent, and as the cylinder rotates, frequencies different from ω_1 will be generated. The problem of sideband formation will be reduced to the study of radiation created by the rotation of this time-dependent cylindrical charge distribution. For that purpose a cylinder is divided into loops fixed in space (at a constant z), and the radiation that comes from each of them is studied. The total radiation is obtained by integrating all loops over different values of $\mathbf{v}_\parallel$, for the same z, then integrating over different z values, over cylinders with different radii, and over different locations of the symmetry axis. The charge configuration in each loop is obtained from the electron's equation of motion under the influence of the incoming electromagnetic field. The solution of this equation, a function of $(z,\mathbf{v}_\parallel)$, can be obtained either numerically through the use of phase plots or analytically with the help of perturbation theory.

IV. EQUATIONS FOR ELECTRON MOTION UNDER THE INFLUENCE OF SEVERAL MONOCHROMATIC WAVES

Figure 4 shows an electron at a point P, moving with velocity $\mathbf{v}=v_\parallel\mathbf{u}_\parallel+v_\perp\mathbf{u}_\perp$. $\mathbf{u}_\parallel$ is a unit vector pointing in the direction opposite to $\mathbf{B}_0$, the Earth's magnetic field, and $\mathbf{u}_\perp$ is a unit vector pointing along the projection of the electron velocity on the plane (x,y) perpendicular to $\mathbf{B}_0$. $(\mathbf{E}_1,\mathbf{B}_1)$ represents the electric and magnetic fields of one of the N waves in the incoming radiation singled out as a reference for the electron motion. If we define

$$\mathbf{E}_t=\sum_{i=1}^{N}\mathbf{E}_i, \tag{1}$$

$$\mathbf{B}_t=\mathbf{B}_0+\mathbf{B}_w=\mathbf{B}_0+\sum_{i=1}^{N}\mathbf{B}_i. \tag{2}$$

The equation of motion can be written as

$$\dot{\mathbf{v}}=\dot{v}_\parallel\mathbf{u}_\parallel+\dot{v}_\perp\mathbf{u}_\perp+v_\perp\dot{\mathbf{u}}_\perp=-(e/m)(\mathbf{v}\Lambda\mathbf{B}_t+\mathbf{E}_t). \tag{3}$$

It is possible to show that

$$\dot{\mathbf{u}}_\perp=(\dot{\theta}+\omega_1+k_1v_\parallel)\mathbf{u}_n, \tag{4}$$

where ω_1 and k_1 are the angular frequency and wave number of the reference whistler-mode wave, and $\dot{\theta}$ is the angular velocity of the electron referred to the same wave field. It should be noted that $\dot{\theta}$ is a convective derivative. Its value takes into account the change in relative wave-electron position due not only to the electron rotation as it moves along its path, but also to the different orientations of the wave field at different points in space and different instants of time. From the last two equations we get, neglecting terms important only for very small pitch angles,

$$\dot{\theta}=-\omega_1-k_1v_\parallel+\frac{\dot{\mathbf{v}}\cdot\mathbf{u}_n}{v_\perp}$$

$$=-k_1v_\parallel-\omega_1+\frac{eB_0}{m}=-k_1(v_\parallel-v_{\rm res}). \tag{5}$$

If $v_\parallel=v_{\rm res}$, we say that the electron is in resonance with the wave. This implies $\dot{\theta}\equiv 0$.

From (3), the Lorentz force equation, we get

$$\dot{v}_\parallel=\dot{\mathbf{v}}\cdot\mathbf{u}_\parallel=(ev_\perp/m)(\mathbf{u}_n\cdot\mathbf{B}_w), \tag{6}$$

which can be simplified to

$$\dot{v}_\parallel=\frac{ev_\perp}{m}\sum_{i=1}^{N}B_i\sin(\theta+\Delta\theta_i), \tag{7}$$

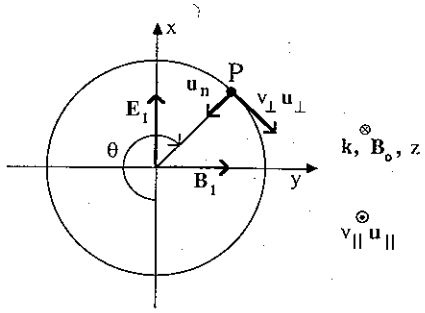

FIG. 4. Diagram of relevant field and electron variables used in this paper. Notice that the waves propagate in the $+z$, and the electrons in the $-z$ direction.

with

$$\Delta\theta_i = \Delta k_i z - \Delta\omega_i t + \Delta\phi_i, \tag{8}$$

where the Δ operation means taking the difference between a parameter of wave i and the same parameter of reference wave 1. Putting in $z = -v_\parallel (t - t_0)$, where t_0 is the time the electron enters the interaction region, we get after some manipulation,

$$\Delta\theta_i = -\Omega_i t + \phi_i, \tag{9}$$

with

$$\Omega_i = [(v_\parallel / v_g) + 1]\Delta\omega_i \tag{10}$$

and

$$\phi_i = (\Omega_i - \Delta\omega_i) t_0 + \Delta\phi_i, \tag{11}$$

where v_g is the wave group velocity at the average radiation frequency, and Ω_i is the Doppler-shifted frequency difference between waves i and 1.

Finally, from (5), (7), and (9),

$$\ddot{\theta} = -\sum_{i=1}^{N} \Omega_{ti}^2 \sin(\theta - \Omega_i t + \phi_i), \tag{12}$$

where $\Omega_{ti}^2 = ev_\perp k_\perp B_i / m$ is the square of the trapping frequency associated with wave i. Reminding ourselves that

$$\frac{d^2}{dt^2} = v_\parallel^2 \frac{d^2}{dz^2}, \tag{13}$$

we can also write

$$\theta'' = -\sum_{i=1}^{N} \left(\frac{\Omega_{ti}}{v_\parallel}\right)^2 \sin\left[\theta + \left(\frac{\Omega_i}{v_\parallel}\right) z - \Delta\omega_i t_0 + \Delta\phi_i\right]. \tag{14}$$

The last equation emphasizes that the variation in angular position of the electron depends on how far it has traveled inside the interaction region, that the electron oscillations to be discussed in the following sections do not occur at a fixed point in space as function of time but are spread in space as the electron travels along the z axis, and that their periods do not modulate the electromagnetic waves in any direct way. The solution to the equation will be

$$\theta = \theta(\theta_0, \theta_0', z, t_0). \tag{15}$$

The trajectory described by the electron will depend on θ_0, its initial orientation, θ_0', which is related to the initial value of $v_\parallel$, and on t_0, the time the electron arrives at the interaction region. If only one wave is present, all $\Delta\omega_i$ will be equal to zero, there will be no dependence on t_0, and the trajectories will be the same for all electrons starting with the same initial angular positions and same $v_\parallel$. As stated in Sec. III, the charge distribution will behave as a rotating rigid body radiating at the single wave frequency. If more than one wave is present, the equation predicts the existence of a set of trajectories, with initial point and derivative given by θ_0 and θ_0', which are slowly (compared with ω_1, the frequency of the reference wave) being distorted in time. The electron distribution can be decomposed into a sum of loops rotating with velocity ω_1 and having time-dependent charge distributions. This charge distribution can be obtained by studying the solution θ at a fixed point in space, z_0, as a function of the initial angular position, initial $v_\parallel$, and moment of entrance in the interaction region.

For actual calculations, it will be more convenient to use the time-dependent form of the equation of motion, keeping in mind that observation of the motion at a fixed position in space is equivalent to an observation at a time $t_{z_0} = |z_0|/v_\parallel + t_0$, a function of the electron travel time and instant of arrival at the interaction region. We assume in that expression that $v_\parallel$ is approximately constant over the electron trajectory.

V. PHASE PLOTS

To study the evolution of the electron distribution, we assume initially that all electrons arrive at time t_0 at the interaction region, having a random distribution in θ_0 and $\dot{\theta}_0$ with constant probability density in each variable. This distribution will not give rise to radiation because its gradient as a function of $\dot{\theta}_0$ (initial $v_\parallel$) is zero. However, if we follow the evolution of each electron trajectory in space and plot its projection in the $(\theta,\dot{\theta})$ plane, we will arrive at a clear qualitative understanding of the radiation process.

Figure 5 shows the resulting phase plots for the case of two equal amplitude waves separated by a frequency Ω. Different values for t_0 were chosen in each plot so that the waves in the beginning of the interaction region are rotated relatively to each other by an additional 60° from one plot to another. They are in phase in the first plot.

To create the plots, for a given initial time t_0, several trajectories are started at random values of $(\theta_0,\dot{\theta}_0)$. The time evolution of each trajectory is calculated and the electron position $(\theta,\dot{\theta})$ is plotted as a point in the plane at times

$$t = t_0 + (2n\pi/\Omega), \quad n = 1,2,\ldots\infty. \tag{16}$$

This must be done because, as seen by looking at the different plots, all trajectories possess a continuous upwards motion which would blur the pictures if all moments in time were plotted. The choices of t given by the above equation are such that the waves' relative angular position is always the same and the trajectories are always at the same position. A complete description of the motion would consist of the set of all phase plots for all values of t_0.

The plots show various resonances. Those are sets of electron trajectories which have been strongly affected by the presence of the two waves. The trajectories have been changed from vertical straight lines to sets of concentric distorted ovals, defining regions from which we should expect most of the radiation to come. There is one resonance at each incoming wave frequency, one halfway between them (half harmonic), two at one third of the distance from one wave to another, and many more that would be seen if the plots were more accurate.

Figure 6 shows a phase plot of the same two waves after their amplitudes are increased by a factor of 12.5 (22-dB growth). Due to the greater amplitude values, the previously seen resonances have been destroyed, chaotic electron motion being present, instead. However, outside the chaotic region, two resonances are present. They are separated from the main waves by a frequency Ω and will radiate at their first harmonic frequencies.

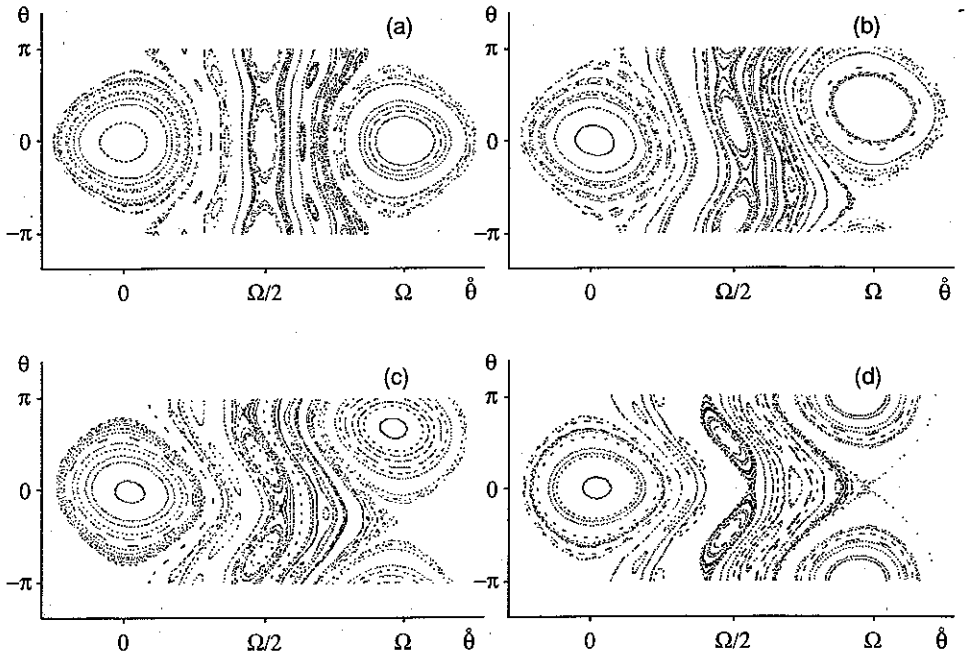

FIG. 5. Resonances created by two equal intensity carriers. The presence of each carrier resonance at $\dot{\theta}=0, \Omega$ is readily apparent, together with additional resonances at $\dot{\theta}=\Omega/3, \Omega/2$, and $2\Omega/3$. As the carriers rotate relatively to each other, the additional resonances are seen to rotate and radiate at frequencies $\omega/3$, $\omega/2$, and $2\omega/3$. The distortions they undergo during rotation, will frequency modulate their radiations, adding weaker lines to the spectrum: (a) Relative carrier phase $=0°$. (b) Relative carrier phase $=60°$. (c) Relative carrier phase $=120°$. (d) Relative carrier phase $=180°$.

Figure 7 shows plots for a more realistic situation. The region of phase space chosen for observation is located, as shown in Fig. 7(a), around the half harmonic resonance. The chosen electron distribution is constant in θ at $z=0$ but has a constant gradient in $\dot{\theta}$ (that is, in $v_{\parallel}$), as shown in Fig. 7(b). Figure 7(c) shows the effect of the waves on the electron distribution at a point z_0 well into the interaction region. The gradient in $\dot{\theta}_0$ has been partially transformed into a gradient in θ. The gradient in θ is equivalent to an inhomogeneous charge loop at z_0, rotating with the average frequency of the two waves, and giving rise to radiation at the same frequency. We see that not all bunching comes from inside the resonance. Figure 7(d) shows the wave effect on a distribution with zero gradient. In this case there is no bunching, indicating that an initial density gradient is necessary for radiation to be generated.

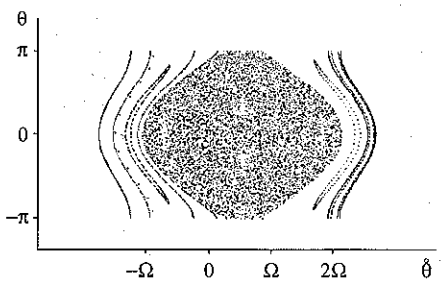

FIG. 6. Chaotic motion and first harmonics created by two strong waves. The carriers are at 0 and Ω, and their relative phase is 0°. Their common trapping frequency is $\Omega/2$. The first harmonics are at $-\Omega$ and 2Ω. Due to the high intensity of the carriers, organized electron motion in a large region containing the carrier potential wells is vompletely destroyed and replaced by chaotic motion, where wave growth cannot take place. The first harmonic resonances would be destroyed only if they were wide enough to overlap the chaotic region. Chaotic motion is believed by the author to be one of the main cuases of triggered emissions observed in magnetospheric data.

VI. GENERAL SOLUTION OF THE EQUATION OF MOTION—KAM THEOREM

When writing the equation of motion

$$\ddot{\theta}=\sum_{i=1}^{N}\Omega_{ti}^2\sin(\theta-\Omega_i t+\phi_i), \tag{17}$$

we have assumed that the reference wave has been wave 1 and that all $\Delta\omega_i$ are measured relative to it. It will be convenient in our next considerations, to measure frequencies from an arbitrary origin. With that in mind, we will shift our origin to an infinitesimally weak wave having an arbitrary frequency, ω_r, which will serve only as a frequency reference, being ignorable for all other purposes. This will imply $\Delta\omega_1\neq 0$ and that all $\Delta\omega_i$ will be defined up to an overall additive constant.

The above equation of motion is equivalent to the Hamiltonian

$$H(q,p,t)=\frac{p^2}{2}-\sum_{i=1}^{N}\Omega_{ti}^2\cos(q-\Omega_i t+\phi_i), \tag{18}$$

as can be seen by Hamilton's equations:

$$\dot{q}=\frac{\partial H}{\partial p}=p, \tag{19}$$

$$\dot{p}=-\frac{\partial H}{\partial q}=-\sum_{i=1}^{N}\Omega_{ti}^2\sin(q-\Omega_i t+\phi_i). \tag{20}$$

The system we are studying has $N+1$ degrees of freedom: the phase of the N waves, and the electron angular position. This is explicitly shown by the fact that it can be described by the $N+1$ dimensional Hamiltonian,

$$H(q_i,p_i)=\frac{p_{N+1}^2}{2}+\sum_{i=1}^{N}\Omega_i p_i-\sum_{i=1}^{N}\Omega_{ti}^2\cos(q_{N+1}-q_i+\phi_i), \tag{21}$$

which has the following equations of motion:

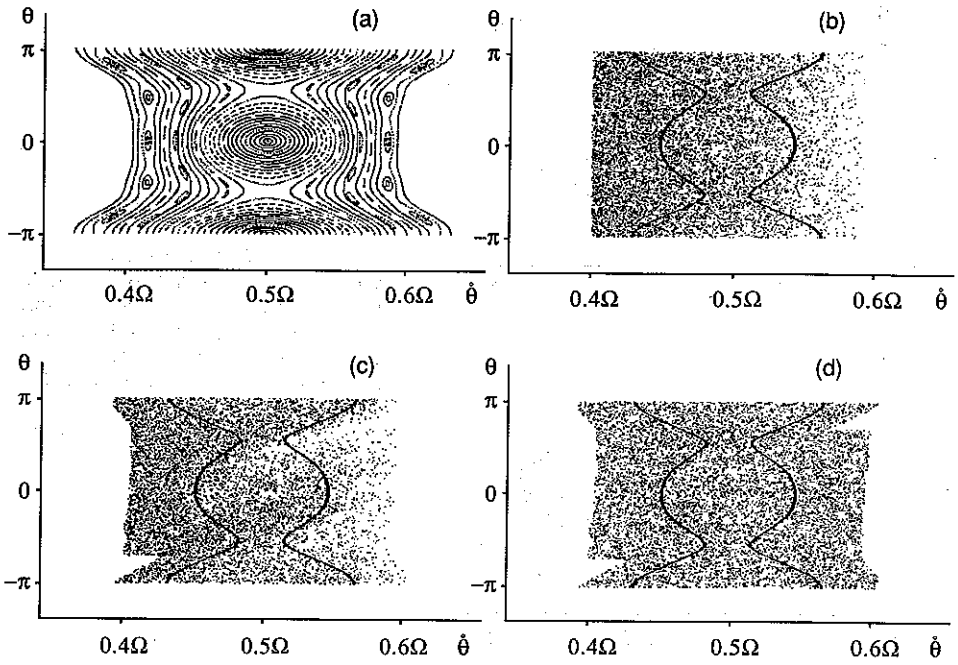

FIG. 7. "Bunching" of electron distributions created by half harmonic resonance. The carriers are at frequencies 0 and Ω. The dark lines in (b), (c), and (d) are the approximate two branches of the separatrix of resonance (a): (a) Phase plot of half harmonic resonance which produces the "bunching" effect in the following pictures. (b) An initial electron distribution with constant gradient in the $\dot{\theta}$ direction. (c) Effect of resonance (a) on electron distribution (b) after it has propagated for some distance along the interaction region. It can be seen that "bunching," when it occurs, is as pronounced outside as inside a resonance. The charge rarefactions at $\approx \pi/2$ and $\approx -\pi/2$ are slightly different from each other due to distortions in the resonance, as mentioned in Sec. IX. Because of it, a net amount of radiation will be created. (d) Effect of resonance (a) on an electron distribution with zero gradient. It is null, as shown by this picture made at the same z position as in (c).

$$\dot{q}_{N+1} = \frac{\partial H}{\partial p_{N+1}} = p_{N+1}, \tag{22}$$

$$\dot{p}_{N+1} = -\frac{\partial H}{\partial q_{N+1}} = -\sum_{i=1}^{N} \Omega_{ti}^2 \sin(q_{N+1} - q_i + \phi_i), \tag{23}$$

$$\dot{q}_i = \frac{\partial H}{\partial p_i} = \Omega_i, \tag{24}$$

and

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} = \Omega_{ti}^2 \sin(q_{N+1} - q_i + \phi_i). \tag{25}$$

The p_i are dummy variables and can be ignored. The equations for the q_i give

$$q_i = \Omega_i t. \tag{26}$$

Substituting those values in Eqs. (22) and (23), eliminating p_{N+1}, and making $q_{N+1} = \theta$, we get the original equation of motion. If we now define

$$H_0(p_i) = \frac{p_{N+1}^2}{2} + \sum_{i=1}^{N} \Omega_i p_i \tag{27}$$

and

$$H_1(q_i,p_i) = -\sum_{i=1}^{N} \Omega_{ti}^2 \cos(q_{N+1} - q_i + \phi_i), \tag{28}$$

we can write

$$H_1(q_i,p_i) = H_0(p_i) + H_1(q_i,p_i), \quad i = 1...(N+1), \tag{29}$$

where $H_1(q_i,p_i)$ is periodic in all q_i, and $H_0(p_i)$ depends only on the p_i. Under those circumstances, the KAM theorem for nonlinear systems states that the motion described by H_0 (free-electron motion) will be appreciably affected by H_1 only if the following relationship holds for the variables describing the free motion:

$$\sum_{i=1}^{N+1} n_i \dot{q}_i = 0, \tag{30}$$

n_i being arbitrary integers. Making $\dot{q}_{N+1} = \dot{\theta} = \Omega$, we will have

$$n\Omega = \sum_{i=1}^{N} n_i \Omega_i. \tag{31}$$

Now, the interaction between the electron and the system of electromagnetic waves should not be dependent on the chosen origin of coordinates. Therefore, if we make a shift $\delta\Omega$ in the coordinate system, the new equation obtained should also be a valid description of the interaction:

$$n(\Omega + \delta\Omega) = \sum_{i=1}^{N} n_i(\Omega_i + \delta\Omega). \tag{32}$$

Since $\delta\Omega$ is arbitrary, the two equations can be true only if

$$n = \sum_{i=1}^{N} n_i. \tag{33}$$

Since $\omega_i = \omega_r + \Delta\omega_i = \omega_r + \Omega_i(v_\parallel + v_g)/v_g$, the formula that gives the frequencies for strong electron-wave interaction, that is, the resonant frequencies, is

$$\omega = \sum_{i=1}^{N} n_i\omega_i \Big/ \sum_{i=1}^{N} n_i, \tag{34}$$

where $\omega = \omega_r + \Omega(v_\parallel + v_g)/v_g$ is both the frequency of a wave resonant with an electron of $\dot{\theta} = \Omega$, and the frequency at which such an electron would radiate. The allowed resonance frequencies are just the weighted averages, with arbitrary integer weights, of the incident wave frequencies. This simple formula, independent of the detailed form of the Hamiltonian describing the motion, and mentioned in the Introduction of this paper, was the one successfully used by Sá and Helliwell[1] to describe the Siple sideband spectra.

VII. ANALYTICAL SOLUTION—LIE TRANSFORM PERTURBATION METHOD

To study specific resonances, a more sophisticated mathematical treatment is required. We choose the Lie

transform perturbation method, applicable to any Hamiltonian system, which has the following advantages: It does not use mixed variables in its formalism, allowing the perturbation expansion to be extended to arbitrarily high orders. It does not generate secular terms or small denominators if a judicious choice of the associated generating function is made. Its formalism, being based on Poisson brackets series, is canonical and invariant in form under any canonical change of coordinates.

The method has been described in several places. Here we follow the notation used by Cary[3] (Deprit perturbation series), and summarize it as follows:

We assume that the known Hamiltonian, h, is expressible as a power series for a small quantity ϵ:

$$h(q,p,t) = \sum_{n=0}^{\infty} \epsilon^n h_n(q,p,t), \tag{35}$$

and we look for a new and simpler Hamiltonian, K:

$$K(Q,P,t) = \sum_{n=0}^{\infty} \epsilon^n K_n(Q,P,t) \tag{36}$$

through a change in coordinates

$$Q = T(q,p,t)q, \tag{37}$$

$$P = T(q,p,t)P, \tag{38}$$

where T is an unknown operator expressible in powers of ϵ:

$$T(q,p,t) = \sum_{n=0}^{\infty} \epsilon^n T_n(q,p,t). \tag{39}$$

The solution is found with the help of a generating function

$$W(q,p,t) = \sum_{n=0}^{\infty} \epsilon^n W_{n+1}(q,p,t) \tag{40}$$

and a sequence of operators L_n defined by

$$L_n f(q,p,t) = \{W_n, f\}, \quad n = 1 \ldots \infty, \tag{41}$$

where the curly brackets represent the Poisson bracket operation.

The new Hamiltonian, expressed in the new variables (Q,P), can be obtained up to third order from the following set of equations:

$$K_0 = h_0, \tag{42}$$

$$\frac{\partial W_1}{\partial t} + \{W_1,h_0\} = K_1 - h_1, \tag{43}$$

$$\frac{\partial W_2}{\partial t} + \{W_2,h_0\} = 2(K_2 - h_2) - L_1(K_1 + h_1), \tag{44}$$

$$\frac{\partial W_3}{\partial t} + \{W_3,h_0\} = 3(K_3 - h_3) - L_1(K_2 + 2h_2) - L_2(K_1 + h_1/2) - L_1^2 h_1/2, \tag{45}$$

and the inverse transformation, to return to the original (q,p) pair is

$$q = T^{-1}Q, \tag{46}$$

$$p = T^{-1}P, \tag{47}$$

$$T^{-1}(Q,P,t) = \sum_{n=0}^{\infty} \epsilon^n T_n^{-1}(Q,P,t), \tag{48}$$

where

$$T_0^{-1} = I, \tag{49}$$

$$T_1^{-1} = L_1, \tag{50}$$

$$T_2^{-1} = L_2/2 + L_1^2/2, \tag{51}$$

$$T_3^{-1} = L_3/3 + L_1L_2/6 + L_2L_1/3 + L_1^3/6. \tag{52}$$

We see that the set of functions W_n defines the change of coordinates completely. Since the set of equations defining the new Hamiltonian has $2n$ unknowns, up to order n, we are free to choose the W_n as we please, defining the change of coordinates arbitrarily. After a choice is made, the K_n can be calculated together with the T_n^{-1} that define the inverse transformation. The choice of the W_n should be made judiciously, however, since they have to be bounded for the perturbation expansion to be valid. The equation for the W_n can be written in the form

$$\frac{\partial W_n}{\partial t} + \{W_n,h_0\} = nK_n(Q,P,t) - f_n(Q,P,t) = g_n(Q,P,t). \tag{53}$$

Assuming we are solving the equations recursively, f_n is known because it depends on $h_n \ldots h_1$, which are given, and on $W_{n-1} \ldots W_1$ and $K_{n-1} \ldots K_1$, which are found at a lower order. The solution to such an equation is given by Cary[3]:

$$W_n = \int^t d\tau\, g_n[Q + \Omega(\tau - t),P,\tau], \tag{54}$$

with

$$\Omega = \frac{\partial h_0}{\partial P}. \tag{55}$$

It seems clear that if g_n has terms which are slowly varying in time as, for example, a constant or a resonance for a certain value of P, W_n will vary approximately linearly in time, will be unbounded, and the perturbation expansion will fail. This gives us the guidelines for choosing K_n, and implicitly, the desired change of coordinates: K_n must absorb all constant terms present in f_n together with any term representing a resonance in the neighborhood of the P we are interested in. This will make W_n a bounded function and, at the same time, will transfer to K_n the resonances we intend to study.

The process of solution of our equations is therefore defined: Starting with K_0, and proceeding to successively higher orders, we put in the K_n all constants present in the functions f_n of our equations. At the same time, we inspect each f_n for a resonance of interest. If we do not find it, we calculate W_n and proceed to a higher order. If we do find it, we add it to K_n and stop the search. The new Hamiltonian, K, will be composed of constant terms plus a single, higher-order term describing a resonant process, and will be valid for all values of P away from resonances skipped and incorporated in the lower-order W_n's. Expressions for the (q,p) can be obtained by calculating the operator T^{-1} and applying it to (Q,P).

VIII. ELECTRON MOTION FOR NONRESONANT VALUES OF $V_\parallel$

If we use the explicitly time-dependent form of the Hamiltonian, and put $\Omega_{ti}^2 = \epsilon A_i$, we have

$$h(q,p,t) = h_0(q,p,t) + \epsilon h_1(q,p,t), \tag{56}$$

with

$$h_0=\frac{p^2}{2},\quad h_1=-\sum_{i=1}^{N}A_i\cos(q-\Omega_i t+\phi_i). \tag{57}$$

To study nonresonant terms, we ignore all resonances, and put in the K_n only constant terms. To get the simplest nontrivial transformation of coordinates we have to go up to second order in the perturbation expansion. Doing this, we get $K_0=h_0$ and $K_1=0$ (because h_1 has no constant terms).

This implies

$$W_1=-\int^t d\tau\, h_1[Q+\Omega(\tau-t),P,\tau]$$

$$=\sum_{i=1}^{N}\frac{A_i}{(P-\Omega_i)}\sin(Q-\Omega_i t+\phi_i), \tag{58}$$

since $\Omega=P$ for our h_0.

We then choose

$$2K_2=\langle L_1 h_1\rangle=\langle\{W_1,h_1\}\rangle, \tag{59}$$

where the angle bracket operation means that the constant part of the function enclosed in brackets should be taken. A straightforward calculation shows that

$$\{W_1,h_1\}=-\frac{\partial W_1}{\partial P}\frac{\partial h_1}{\partial Q}$$

$$=-\sum_{i,j=1}^{N}\frac{A_iA_j}{2(P-\Omega_i)^2}$$

$$\times\{\cos[2Q-(\Omega_i+\Omega_j)t+\phi_i+\phi_j]$$

$$-\cos[(\Omega_i-\Omega_j)t-(\phi_i-\phi_j)]\}. \tag{60}$$

Therefore,

$$K_2=\frac{1}{2}\langle\{W_1,h_1\}\rangle=\frac{1}{4}\sum_{i=1}^{N}\frac{A_i^2}{(P-\Omega_i)^2}. \tag{61}$$

and

$$K=\frac{P^2}{2}+\frac{\epsilon^2}{4}\sum_{i=1}^{N}\frac{A_i^2}{(P-\Omega_i)^2}. \tag{62}$$

The transformation back to (q,p) is given by

$$q=Q+\epsilon\{W_1,Q\}+(\epsilon^2/2)\{W_1,\{W_1,Q\}\}+\ldots, \tag{63}$$

$$p=P+\epsilon\{W_1,P\}+(\epsilon^2/2)\{W_1,\{W_1,P\}\}+\ldots. \tag{64}$$

Keeping only first-order terms and constant second-order terms,

$$q=Q-\epsilon\frac{\partial W_1}{\partial P}-\frac{\epsilon^2}{2}\left\langle\left\{W_1,\frac{\partial W_1}{\partial P}\right\}\right\rangle, \tag{65}$$

$$p=P+\epsilon\frac{\partial W_1}{\partial Q}+\frac{\epsilon^2}{2}\left\langle\left\{W_1,\frac{\partial W_1}{\partial Q}\right\}\right\rangle, \tag{66}$$

and

$$q=Q+\epsilon\sum_{i=1}^{N}\frac{A_i}{(P-\Omega_i)^2}\sin(Q-\Omega_i t+\phi_i), \tag{67}$$

$$p=P-\frac{\epsilon^2}{2}\sum_{i=1}^{N}\frac{A_i^2}{(P-\Omega_i)^3}+\epsilon\sum_{i=1}^{N}\frac{A_i}{(P-\Omega_i)}$$

$$\times\cos(Q-\Omega_i t+\phi_i). \tag{68}$$

Hamilton's equations of motion for K can be solved:

$$\dot{P}=-\frac{\partial K}{\partial Q}=0, \tag{69}$$

$$\dot{Q}=\frac{\partial K}{\partial P}=P-\frac{\epsilon^2}{2}\sum_{i=1}^{N}\frac{A_i^2}{(P-\Omega_i)^3}=\Omega. \tag{70}$$

The solution is

$$P=P_0\approx\Omega, \tag{71}$$

$$Q=\Omega t+Q_0. \tag{72}$$

The meaning of Ω (and Q_0) can be obtained by noting that $Q\to q$ and $\epsilon\to 0$. For a free electron, q represents the angular position of the electron measured from the reference wave as the electron propagates along the $-z$ direction. We will consider electrons with a $v_\parallel$ such that its motion is in resonance with an infinitesimally weak wave of frequency $\Delta\omega$ above the reference wave frequency (this imposes no constraint on the motion). If the electron is in resonance with such a wave, their relative angular position will be constant, and the electron's angular position with respect to the reference wave will be the difference in phase between the two waves, $\Delta\theta$. This expression is known from Eqs. (9) and (11), allowing us to write

$$Q=q=\Delta\theta=\Omega(t-t_0)+\Delta\omega t_0-\Delta\phi=\Delta\omega t_0+q_{z_0}. \tag{73}$$

We see that Ω is the Doppler-shifted frequency difference of the two waves seen by the electron as it moves along. $(t-t_0)$ is the constant time it takes each electron to go from the beginning of the interaction region to the point of observation, z_0. Substituting those values in the expressions for q and p,

$$q=\Delta\omega t_0+q_{z_0}+\epsilon\sum_{i=1}^{N}\frac{A_i}{(\Omega-\Omega_i)^2}$$

$$\times\sin[(\Delta\omega-\Delta\omega_i)t_0+\phi_{z_0}], \tag{74}$$

$$p=\Omega+\epsilon\sum_{i=1}^{N}\frac{A_i}{(\Omega-\Omega_i)}\cos[(\Delta\omega-\Delta\omega_i)t_0+\phi_{z_0}], \tag{75}$$

where $\phi_{z_0}=(\Omega-\Omega_i)(t-t_0)-(\Delta\phi-\Delta\phi_i)$ and q_{z_0} are arbitrary but constant phases at the point of observation z_0. The equations state that a stream of electrons starting with a constant phase relative to the infinitesimally weak wave of frequency $\Delta\omega$ will in the limit $\epsilon\to 0$, have an intersection with the plane $z_0=-v_\parallel(t-t_0)$ which turns, as t_0 varies, with angular velocity $\omega=\omega_r+\Delta\omega$. If those streams are initially equally distributed in phase, the resulting intersection will form a dc current loop and will not radiate as expected. If $\epsilon>0$, the equations predict conditionally periodic changes away from constant frequency rotation. Those distortions are the first-order approximation to the bunching created by the incoming waves outside their potential wells. Each term in the summation represents the effect on the electron of one wave, and will produce radiation at the perturbing wave frequency. Because we have kept only the first-order terms, the effects add up linearly, and the bunching from one wave does not affect the bunching created by the others.

An important effect, here due to the adiabatic invariance of P, must also be noted: If we average away the oscilla-

tory terms, and assume that a certain particle has $p = P = \Omega_0$ for $\epsilon = 0$, and that the wave fields are adiabatically turned on, then P will remain constant, and p will vary as

$$p = \Omega_0 - \frac{\epsilon^2}{2} \sum_{i=1}^{N} \frac{A_i^2}{(\Omega_0 - \Omega_i)^3}. \tag{76}$$

The particle will be shifted due to the presence of the waves, the forces being such that each wave tends to pull the particle towards it. The combination of such an effect with the natural fall-off of the electron distribution density with increasing $v_\parallel$ (that is, increasing ω) can explain why sideband growth is more pronounced on the upper frequency side of a carrier. Figure 8 shows the effect of a single wave on the electron distribution. The electron shift flattens it on the low-frequency side, and steepens it on the high-frequency side. Since wave growth increases with the electron density gradient, growth will be expected to be more noticeable on the high-frequency side of the carrier as is seen in most Siple data.

IX. HALF HARMONIC RESONANCE

If we look for resonances using the perturbation expansion up to second order, first we will get

$$K_0 = h_0. \tag{77}$$

We will have then to put

$$K_1 = 0 \tag{78}$$

because h_1 has resonances only for $\dot{Q} \approx \Omega_i$. Those are resonances at the incoming waves' frequencies which we do not intend to study now. They will be focused in the second part of this paper with the help of different canonical variables.

With those choices, W_1 will be the same as for the nonresonant case. The possible second-order resonances will come from the term $L_1 h_1$. This term was already calculated [Eq. (60)] and is made up of two sums. The second sum contains, among other things, a constant term which we must include in K_2. The first has several resonances all of the same form. We will study one such resonance for a pair of waves $(i,j), i \neq j$. (For $i = j$ the resonance again falls on one of the incoming carriers.) By choosing

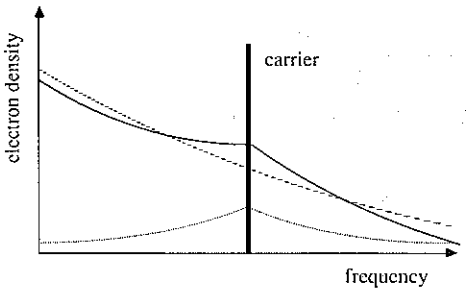

FIG. 8. Distortion of electron distribution by electromagnetic wave. The dotted line shows the effect of the carrier on an otherwise constant electron distribution. The dashed line represents a typical unperturbed distribution. The continuous line shows the resulting flattening at lower frequencies and steepening at the higher frequencies present on the final electron distribution as a consequence of the wave forces.

$$2K_2 = \frac{1}{2} \sum_{i=1}^{N} \frac{A_i^2}{(P-\Omega_i)^2} - \left(\frac{A_i A_j}{2(P-\Omega_i)^2} + \frac{A_i A_j}{2(P-\Omega_j)^2} \right) \times \cos[2Q - (\Omega_i + \Omega_j)t + \phi_i + \phi_j], \tag{79}$$

and defining $\Omega_{1/2} = (\Omega_i + \Omega_j)/2$, we get

$$K = \frac{P^2}{2} + \frac{\epsilon^2}{4} \sum_{i=1}^{N} \frac{A_i^2}{(P-\Omega_i)^2} - \frac{\epsilon^2 A_i A_j}{4} \times \left(\frac{1}{(P-\Omega_i)^2} + \frac{1}{(P-\Omega_j)^2} \right) \times \cos(2Q - 2\Omega_{1/2} t + \phi_i + \phi_j). \tag{80}$$

Hamilton's equations applied to K will give

$$\dot{Q} = \frac{\partial K}{\partial P} = P - \frac{\epsilon^2}{2} \sum_{i=1}^{N} \frac{A_i^2}{(P-\Omega_i)^3} + \mathcal{O}(P - \Omega_{1/2}), \tag{81}$$

$$\dot{P} = -\frac{\partial K}{\partial Q} = -\frac{\epsilon^2 A_i A_j}{2} \left(\frac{1}{(P-\Omega_i)^2} + \frac{1}{(P-\Omega_j)^2} \right) \times \sin(2Q - 2\Omega_{1/2} t + \phi_i + \phi_j). \tag{82}$$

Since we are considering only points close to the resonance,

$$P \approx \dot{Q} \approx \Omega_{1/2}, \tag{83}$$

and the equation of motion will be

$$\Delta\ddot{Q} = -\frac{8\epsilon^2 A_i A_j}{(\Omega_i - \Omega_j)^2} \sin(\Delta Q), \tag{84}$$

with

$$\Delta Q = 2Q - 2\Omega_{1/2} t + \phi_i + \phi_j. \tag{85}$$

This is the familiar pendulum equation, describing a resonance region of width (in the variable Q)

$$\Gamma_{1/2} = \frac{4\sqrt{2}\epsilon\sqrt{A_i A_j}}{|\Omega_i - \Omega_j|}. \tag{86}$$

The equilibrium points are given by $\Delta Q = 2n\pi$, or

$$Q = \Omega_{1/2} t - (\phi_i + \phi_j)/2 + n\pi, \quad n = 0,1. \tag{87}$$

Substituting ϕ_i and ϕ_j for their values, and defining

$$\Delta\omega_{1/2} = (\Delta\omega_i + \Delta\omega_j)/2, \tag{88}$$

$$\phi_{ij} = \Omega_{1/2}(t - t_0)(\Delta\phi_i + \Delta\phi_j)/2, \tag{89}$$

we have

$$Q = \Delta\omega_{1/2} t_0 + \phi_{ij} + n\pi, \quad n = 0,1 \tag{90}$$

ϕ_{ij} is a constant phase, dependent, however, on the point of observation z_0. The equation describes two resonances oppositely located in phase, which rotate in the Q direction with velocity $\Delta\omega_{1/2}$. Because the two branches are oppositely phased in the (Q,P) plane, we see that to obtain net radiation we will have to depend on the existence of different electron populations in each resonance branch. Expressed in terms of the original variables,

$$q = \Delta\omega_{1/2} t_0 + \frac{\Delta Q}{2} + \phi_{ij} + \epsilon \sum_{k=1}^{N} \frac{A_k}{(\Omega_{1/2} - \Omega_k)^2} \times \sin\left((\Delta\omega_{1/2} - \Delta\omega_k) t_0 + \frac{\Delta Q}{2} + \phi_{ijk} \right), \tag{91}$$

$$p=\Omega_{1/2}+\frac{\Delta\dot{Q}}{2}+\epsilon\sum_{k=1}^{N}\frac{A_k}{(\Omega_{1/2}-\Omega_k)}$$
$$\times\cos\left((\Delta\omega_{1/2}-\Delta\omega_k)t_0+\frac{\Delta Q}{2}+\phi_{ijk}\right) \quad (92)$$

where ϕ_{ijk} does not depend on time. The $\Delta Q/2$ term in q shows that the relative position of any two electrons can be changed by as much as π, making bunching effects appreciable.

The radiation from each charge can be obtained if we go to configuration space. Each charge is obtained by the intersection of the electron stream with the plane $z=z_0$. In that plane, each charge will be rotating with an angular velocity, ω, given by

$$\omega=\frac{\partial q}{\partial t_0}+\omega_r. \quad (93)$$

(Note that p is irrelevant for this calculation, since it describes each individual electron motion as it moves in the $-z$ direction, and we are interested in the stream at a fixed point.) The radiation will be circularly polarized. To obtain the frequency spectrum, we will project the motion in the (x,z) plane. If we define

$$\omega_{1/2}=\omega_r+\Delta\omega_{1/2}=(\omega_i+\omega_j)/2, \quad (94)$$
$$\Delta\omega_{ijk}=\Delta\omega_{1/2}-\Delta\omega_k, \quad (95)$$
$$b_k=\frac{A_k}{(\Omega_{1/2}-\Omega_k)^2}, \quad (96)$$

and

$$\theta_1=\sum_{k=1}^{N}b_k\sin(\Delta\omega_{ijk}t_0+\eta_k), \quad (97)$$

we can write

$$\theta=\omega_{1/2}t_0+(\Delta Q/2)+\phi_{ij}+\epsilon\theta_1. \quad (98)$$

To first order the electron's acceleration along x will be

$$\ddot{x}=-r\omega_{1/2}^2\cos\theta$$
$$=-r\omega_{1/2}^2\cos[\omega_{1/2}t_0+(\Delta Q/2)+\phi_{ij}]$$
$$+\epsilon\theta_1 r\omega_{1/2}^2\sin(\omega_{1/2}t_0) \quad (99)$$

or

$$\ddot{x}=-r\omega_{1/2}^2\,t_0\cos[\omega_{1/2}t_0+(\Delta Q/2)+\phi_{ij}]$$
$$+\frac{\epsilon r\omega_{1/2}^2}{2}\sum_{k=1}^{N}b_k\{\sin[(\omega_i+\Delta\omega_j-\Delta\omega_k)t_0+\eta_k]$$
$$+\sin[\omega_k t_0-\eta_k]\}. \quad (100)$$

The equation shows that radiation will be mainly produced at the frequency $\omega_{1/2}$, as should be expected by the location of the resonance in frequency space, but that there are smaller contributions at the frequencies ω_k and $\omega_i+\Delta\omega_j-\Delta\omega_k$, $k=1...N$. Those are radiation on top of wave k and at the mirror image position of k in the interval defined by i and j. We will see that radiation at those frequencies is generated by other resonances in a more efficient way, and will therefore neglect those small corrections. With this approximation, we can consider all radiation coming from the half harmonic resonance as being located at the frequency $\omega_{1/2}$.

If both resonance branches are equally populated in the (Q,P) plane, for each particle with position ΔQ it is always possible to find an equivalent particle with position $\Delta Q+\pi$. Those particles have opposite accelerations and cancel each other's radiation in the z direction. However, at the beginning of the interaction region the two branches will be located at slightly different positions in the (q,p) plane due to the distortion effects created by the incoming waves. If the electron distribution has a gradient in p different from zero, the population of the two branches will not be the same and a net amount of radiation will be produced [see Fig. 7(c)]. We note also that all points in the resonance radiate at the same frequency $\omega_{1/2}$, despite the resonance finite width, and that the two-branch resonance structure is inconsistent with the presence of radiation. As soon as radiation starts, it will form an additional resonance, rotated by $-\pi/2$ from the radiating resonance, changing the initial electron motion and resonance configuration.

X. THIRD-ORDER RESONANCES

A. General expressions

To look for third-order resonances, we choose K_0, K_1, W_1, and K_2 the same as in the nonresonant case. This implies

$$W_2=\sum_{i,j=1}^{N}\frac{A_iA_j}{2(P-\Omega_i)^2[2P-(\Omega_i+\Omega_j)]}$$
$$\times\sin[2Q-(\Omega_i+\Omega_j)t+\phi_i+\phi_j]$$
$$-\sum_{\substack{i,j=1\\ i\neq j}}^{N}\frac{A_iA_j}{2(P-\Omega_i)^2(\Omega_i-\Omega_j)}$$
$$\times\sin[(\Omega_i-\Omega_j)t-(\phi_i-\phi_j)]. \quad (101)$$

Among the third-order terms, L_1K_2 can be ignored for it contains resonances only at the incoming carriers' frequencies. Third-order resonances must come, then, from the terms L_2h_1 and $L_1^2h_1$. After a straightforward but rather lengthy calculation they are found to be

$$L_2h_1=\sum_{i,j,k=1}^{N}\frac{b_{ijk}}{2}$$
$$\times\cos[3Q-(\Omega_i+\Omega_j+\Omega_k)t+\phi_i+\phi_j+\phi_k]$$
$$+\sum_{i,j,k=1}^{N}\frac{a_{kji}-a_{ikj}-b_{ijk}}{2}$$
$$\times\cos[Q-(\Omega_i+\Omega_j-\Omega_k)t+\phi_i+\phi_j-\phi_k], \quad (102)$$

$$L_1^2h_1=\sum_{i,j,k=1}^{N}\frac{c_{ijk}-d_{ijk}}{2}$$
$$\times\cos[3Q-(\Omega_i+\Omega_j+\Omega_k)t+\phi_i+\phi_j+\phi_k]$$
$$+\sum_{i,j,k=1}^{N}\frac{c_{ijk}-c_{kji}-c_{ikj}+d_{ijk}}{2}$$
$$\times\cos[Q-(\Omega_i+\Omega_j-\Omega_k)t+\phi_i+\phi_j-\phi_k], \quad (103)$$

with

$$a_{ijk} = \frac{A_i A_j A_k}{(P-\Omega_i)^3(\Omega_i-\Omega_j)}, \tag{104}$$

$$b_{ijk} = -\frac{A_i A_j A_k [3P-(2\Omega_i+\Omega_j)]}{(P-\Omega_i)^3[2P-(\Omega_i+\Omega_j)]^2}, \tag{105}$$

$$c_{ijk} = \frac{A_i A_j A_k}{(P-\Omega_i)^3(P-\Omega_k)}, \tag{106}$$

$$d_{ijk} = \frac{A_i A_j A_k}{(P-\Omega_i)^2(P-\Omega_k)^2}. \tag{107}$$

K_3 will have to be chosen from the terms in

$$\frac{L_2 h_1 + L_1^2 h_1}{6} = \sum_{i,j,k=1}^{N} \alpha_{ijk} \cos[3Q-(\Omega_i+\Omega_j+\Omega_k)t + \phi_i+\phi_j+\phi_k] + \sum_{i,j,k=1}^{N} \beta_{ijk} \cos[Q-(\Omega_i+\Omega_j-\Omega_k)t + \phi_i+\phi_j-\phi_k]. \tag{108}$$

For $i=j=k$ all terms in the above expression are resonant at the incoming carriers' positions and are not of interest. We must pick terms with at least one index different from the other remaining two. They will provide the third-order resonances we are looking for.

B. Two-wave resonances

1. One-third subharmonic

For $i=j\neq k$, the term symmetrical under permutation of the three indices gives

$$K_3 = (\alpha_{iik}+\alpha_{iki}+\alpha_{ikk}) \times \cos[3Q-(2\Omega_i+\Omega_k)t+2\phi_i+\phi_k]. \tag{109}$$

Defining $\Omega_{2/3} = (2\Omega_i+\Omega_k)/3$, and calculating the α's:

$$K_3 = -\frac{243 A_i^2 A_k}{16(\Omega_i-\Omega_k)^4} \cos(3Q-3\Omega_{2/3}t+2\phi_i+\phi_k). \tag{110}$$

The Hamiltonian will be

$$K = \frac{P^2}{2} + \frac{\epsilon^2}{4}\sum_{i=1}^{N}\frac{A_i^2}{(P-\Omega_i)^2} - \epsilon^3 \frac{243 A_i^2 A_k}{16(\Omega_i-\Omega_k)^4} \times \cos(3Q-3\Omega_{2/3}t+2\phi_i+\phi_k), \tag{111}$$

which will give the equation of motion:

$$\Delta\ddot{Q} = -\epsilon^3 9 \frac{243 A_i^2 A_k}{16(\Omega_i-\Omega_k)^4} \sin(\Delta Q), \tag{112}$$

with

$$\Delta Q = 3Q - 3\Omega_{2/3}t + 2\phi_i + \phi_k. \tag{113}$$

The motion will have three points of equilibrium for $\Delta Q = 2n\pi$, indicating a three-lobed resonance in Q:

$$Q = \Omega_{2/3}t - \frac{2\phi_i+\phi_k}{3} + \frac{2n\pi}{3}, \quad n=0,1,2. \tag{114}$$

In terms of t_0

$$Q = \Delta\omega_{2/3}t_0 + \Omega_{2/3}(t-t_0) - \frac{2\Delta\phi_i+\Delta\phi_k}{3} + \frac{2n\pi}{3}, \quad n=0,1,2, \tag{115}$$

with $\Delta\omega_{2/3} = (2\Delta\omega_i + \Delta\omega_k)/3$. Observed at a fixed point in space, the resonance will rotate and radiate at

$$\omega_{2/3} = \omega_r + \Delta\omega_{2/3} = (2\omega_i+\omega_k)/3, \tag{116}$$

and will create a line in the frequency axis $\frac{2}{3}$ of the way from wave i towards j. Since the labeling of the waves is arbitrary, there will be another resonance, identical to this one but with wave indices permutated. It will be a resonance creating a sideband wave at a frequence $\frac{1}{3}$ of the way from i to j. The full width of the first resonance in the (Q,P) plane will be

$$\Gamma_{2/3} = 9\sqrt{3}\frac{\epsilon^{3/2}A_i\sqrt{A_k}}{(\Omega_i-\Omega_k)^2} \tag{117}$$

and of the second one

$$\Gamma_{1/3} = 9\sqrt{3}\frac{\epsilon^{3/2}A_k\sqrt{A_i}}{(\Omega_i-\Omega_k)^2}. \tag{118}$$

The different dependence of each resonance on the variables A_k and A_i is clearly very important when one of the carriers in the interacting pair is much weaker than the other.

2. First harmonic resonance

Again assuming $i=j\neq k$ in Eq. (108), we find from the term asymmetrical in i,j,k the following resonance:

$$K_3 = \beta_{iik}\cos[Q-(2\Omega_i-\Omega_k)t+2\phi_i-\phi_k]. \tag{119}$$

If we define $\Omega_{1h} = (2\Omega_i-\Omega_k)$, and calculate β_{iik},

$$K_3 = -\frac{A_i^2 A_k}{16(\Omega_i-\Omega_k)^4}\cos(Q-\Omega_{1h}t+2\phi_i-\phi_k). \tag{120}$$

The full Hamiltonian will be:

$$K = \frac{P^2}{2} + \frac{\epsilon^2}{4}\sum_{i=1}^{N}\frac{A_i^2}{(P-\Omega_i)^2} - \epsilon^3\frac{A_i^2 A_k}{16(\Omega_i-\Omega_k)^4} \times \cos(Q-\Omega_{1h}t+2\phi_i-\phi_k). \tag{121}$$

In a way similar to the one-third subharmonic, we get the equation of motion:

$$\Delta\ddot{Q} = -\epsilon^3\frac{A_i^2 A_k}{16(\Omega_i-\Omega_k)^4}\sin(\Delta Q), \tag{122}$$

with

$$\Delta Q = Q - \Omega_{1h}t + 2\phi_i - \phi_k. \tag{123}$$

There is only one equilibrium point, $\Delta Q = 0$, and only one lobe in this resonance (making it a more efficient radiation source). For the equilibrium point, as a function of t_0, we have

$$Q = \Delta\omega_{1h}t_0 + \Omega_{1h}(t-t_0) - (2\Delta\phi_i - \Delta\phi_k), \tag{124}$$

where $\Delta\omega_{1h} = (2\Delta\omega_i - \Delta\omega_k)$. The resonance has width

$$\Gamma_{1h} = \frac{\epsilon^{3/2}A_i\sqrt{A_k}}{(\Omega_i-\Omega_k)^2} \tag{125}$$

and rotates with frequency

$$\omega_{1h} = \omega_r + \Delta\omega_{1h} = \omega_i + (\omega_i - \omega_k). \tag{126}$$

The radiation will be a first harmonic of the two-wave frequency separation, located next to wave i. Again, because the labeling is arbitrary, there will be another first harmonic resonance next to wave k, rotating with frequency $\omega_k + (\omega_k - \omega_i)$ and having width $\epsilon^{3/2}A_k\sqrt{A_i}/(\Omega_i - \Omega_k)^2$.

C. Three-wave effects

1. Arithmetic mean

Assuming $i \neq j, j \neq k, k \neq i$ we get from the symmetrical term a single type of interaction:

$$K_3 = (\alpha_{ijk} + \alpha_{jik} + \alpha_{ikj} + \alpha_{kij} + \alpha_{kji} + \alpha_{jki}) \times \cos[3Q - (\Omega_i + \Omega_j + \Omega_k)t + \phi_i + \phi_j + \phi_k], \tag{127}$$

which can be calculated to give:

$$K_3 = -\frac{3^5}{2}\frac{A_iA_jA_k(\Delta_{ij}^2 + \Delta_{ij}\Delta_{jk} + \Delta_{jk}^2)}{(2\Delta_{ij} + \Delta_{jk})^2(2\Delta_{jk} + \Delta_{ij})^2(\Delta_{ij} - \Delta_{jk})^2} \times \cos[3Q - (\Omega_i + \Omega_j + \Omega_k)t + \phi_i + \phi_j + \phi_k], \tag{128}$$

where $\Delta_{ij} = \Omega_i - \Omega_j$ and $\Delta_{jk} = \Omega_j - \Omega_k$. This resonance has a three-lobe structure similar to the one-third subharmonic, and radiates at the frequency

$$\omega = (\omega_i + \omega_j + \omega_k)/3, \tag{129}$$

that is, at the average frequency of the three waves. The resonance width will be

$$\Gamma_{\text{arith}} = 2\sqrt{2}\epsilon^{3/2} \times \left(\frac{3^5A_iA_jA_k(\Delta_{ij}^2 + \Delta_{ij}\Delta_{jk} + \Delta_{jk}^2)}{(2\Delta_{ij} + \Delta_{jk})^2(2\Delta_{jk} + \Delta_{ij})^2(\Delta_{ij} - \Delta_{jk})^2}\right)^{1/2}. \tag{130}$$

The denominator is divergent, effectively increasing the resonance width up to second order when one of the two waves is approximately halfway between the other two. The perturbation expansion will break down (due to the formation of chaotic motion) when the resonances associated with any two waves begin to overlap. This imposes a minimum displacement of the sideband location away from wave j: $P \approx \Omega_j + 2\sqrt{\epsilon A_j}$. This implies $\Delta_{ij} - \Delta_{jk} \approx 6\sqrt{\epsilon A_j}$, and $|\Delta_{ij}| \approx |\Delta_{ik}|/2$. The resonance width will be under those conditions approximately equal to

$$\Gamma_{\text{arith}} \approx 2\epsilon\sqrt{2A_iA_k}/|\Delta_{ik}|, \tag{131}$$

which is of the order of the half harmonic resonance width.

2. Intermodulation effects

For $i \neq j, j \neq k, k \neq i$, we get from the asymmetrical term in Eq. (108):

$$K_3 = (\beta_{ijk} + \beta_{ijk})\cos[Q - (\Omega_i + \Omega_j - \Omega_k)t + \phi_i + \phi_j - \phi_k], \tag{132}$$

which gives, when the coefficients are calculated,

$$K_3 = -\frac{1}{2}\frac{A_iA_jA_k}{2\Delta_{ik}\Delta_{jk}(\Delta_{ik} + \Delta_{jk})^2} \times \cos[Q - (\Omega_i + \Omega_j - \Omega_k)t + \phi_i + \phi_j - \phi_k]. \tag{133}$$

The above term creates three different resonances, depending on the wave that is singled out as wave k. All resonances have a single lobe, and therefore, are efficient radiators. Figure 9 shows the three possible radiation frequencies. The effect is equivalent to the intermodulation of one wave by the frequency difference of the other two, either sign being possible. The frequencies of radiation will be

$$\omega = \omega_i + (\Delta\omega_j - \Delta\omega_k) = \omega_i + (\omega_j - \omega_k), \tag{134}$$

and the resonance widths will be

$$\Gamma_{\text{int}} = \frac{2\epsilon^{3/2}\sqrt{2A_iA_jA_k}}{\sqrt{|\Delta_{ik}||\Delta_{jk}||\Delta_{ik} + \Delta_{jk}|}}. \tag{135}$$

When wave k is approximately halfway between the other two interacting waves, the resonance width increases without bounds. If wave k is specially weak, the observed phenomenon will be similar to an amplification of the said wave with a slight shift in its frequency. The limiting position of the resonance is given by $P \approx \Omega_k - 2\sqrt{\epsilon A_k}$. This implies $\Delta_{ik} + \Delta_{jk} \approx -2\sqrt{\epsilon A_k}$, and $\Delta_{ik} \approx \Delta_{kj} \approx \Delta_{ij}/2$.

The resonance width will be approximately

$$\Gamma_{\text{int}} \approx 2\epsilon\sqrt{2A_iA_j}/|\Delta_{ij}|, \tag{136}$$

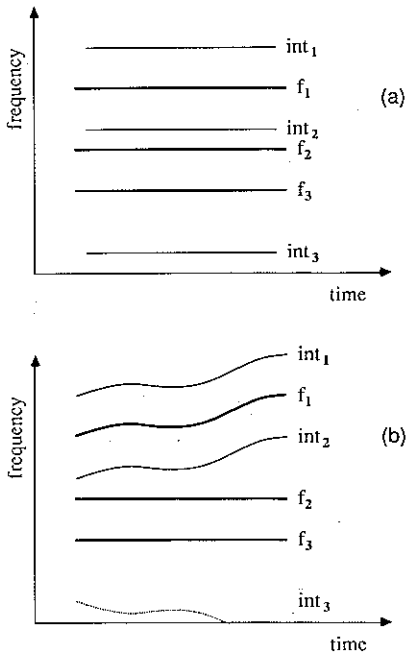

FIG. 9. Frequency distribution of sidebands created by third-order intermodulation process: (a) f_1, f_2, and f_3 are incoming carriers and int_1, int_2, and int_3 are the three different sideband waves produced by the intermodulation process. For the spacings shown in the picture, int_2 is always greater in amplitude than int_1 and int_3. (b) Intermodulation is a nonlocal effect. If f_1 has variable frequency, the sideband waves created by the interaction are able to follow its shift in frequency, despite the fact that f_2 and f_3 remain fixed. int_3 is represented by a dotted line because it weakens as f_1 moves up.

which is of the order of the half harmonic resonance width. Under those conditions, and due partly to its single-lobe structure, this resonance will be the most efficient radiator among all we have seen in this paper.

XI. HIGHER-ORDER EFFECTS

It can be seen from the lower-order terms and from the way they are generated that the general expression for the resonance frequencies of order n will be given by Eq. (34) with $\Sigma_{i=1}^{N}|n_i| = n$ as a constraining condition. This shows that an nth-order effect can involve at most n different waves.

The expression for the frequency of the second harmonic resonance of a pair of waves (i,j) is $\Omega = \Omega_i + 2(\Omega_i - \Omega_j)$. The order of this term will be $n = |3| + |-2| = 5$. The second harmonic is, therefore a fifth-order effect, and very likely will not produce directly observable radiation. Its creation through the present mechanism can be possible only through a two-step process in which the first harmonic is created, grows, and then interacts with one of the carriers to produce another first harmonic which is the second harmonic we are looking for.

XII. CONCLUSION

As main features of the sidebands just studied, we can quote the following:

(1) Sidebands are created by a line-line interaction process. No single-line sidebands appear in the formalism.

(2) Sidebands can be created only at certain specific frequencies given by

$$\omega = \sum_{i=1}^{N} n_i\omega_i \Big/ \sum_{i=1}^{N} n_i. \tag{34}$$

(3) Sidebands can be created only if the electron distribution has a nonzero gradient in $v_\parallel$.

(4) Each sideband wave is associated with a resonance of the electron motion which, in an adequate system of coordinates, is always described by the pendulum equation.

(5) Radiated sideband frequencies do not depend on $v_\perp$, only the amplitudes do. Integration over different pitch angles will not smear out the resonances increasing the radiation linewidth.

(6) The frequency of individual electron oscillations ("trapping" oscillations) does not directly affect the observed sideband frequencies in any way. Those oscillations are convective, occurring as the electrons, as part of a stream, move along the interaction region. Although each electron oscillates, the stream as a whole constitutes a dc current line and will not radiate unless distorted or accelerated by some additional cause. The overall emitted radiation will be a combined result from the acceleration of all such current lines, and its spectrum will reflect the frequency structure of the externally accelerating forces.

(7) Growth will affect line amplitudes appreciably. The resonance widths calculated in this paper will give only an order of magnitude estimate of the observed relative radiation intensities.

(8) Due to its recurrent law of formation, the sideband system has a fractal structure, exhibiting many properties associated with randomness. Only for cases where the recurrence process of new wave formation stops due to limited wave growth do we obtain spectra with a clearly recognizable order in the frequency distributions.

(9) Chaos, present when the potential wells of different waves start to overlap, as seen in Fig. 6, is an important factor in the formation of the final radiation spectra. This author believes that, combined with inhomogeneity effects, chaos is the cause for the triggered emissions frequently associated with Siple and other narrow bandwidth signals as they propagate through the magnetosphere.

In this first part of the paper we have focused on resonances created outside the potential wells of a set of single-frequency carriers interacting with the magnetospheric plasma. Those resonances are narrower than the ones directly associated with each carrier, and are therefore more susceptible to disruption by forces coming from the inhomogeneity of the Earth's magnetic field. In the second part of this paper, we will see that radiation, at the same frequencies as the ones we have studied in this first part, can also be created by resonances inside each carrier potential well. Those resonances are more protected from inhomogeneity effects and should produce radiation over a longer interaction length than the ones considered here.

ACKNOWLEDGMENTS

The author thanks his StarLab colleagues for helpful discussions. This work was sponsored by the Division of Polar Programs of the National Science Foundation under Grant No. DPP-86-13783.

[1] L. A. D. Sá and R. A. Helliwell, J. Geophys. Res. **93**, 1987 (1988).
[2] A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion* (Springer, New York, 1983).
[3] J. R. Cary, Phys. Rep. **79**, 129 (1981).