THE ASTROPHYSICAL JOURNAL, 391:854–864, 1992 June 1

A THERMAL/NONTHERMAL MODEL FOR SOLAR MICROWAVE BURSTS

STEPHEN G. BENKA
E. O. Hulburt Center for Space Research, Code 4171B, Naval Research Laboratory, Washington, DC 20375-5000
AND
GORDON D. HOLMAN
Laboratory for Astronomy and Solar Physics, NASA/Goddard Space Flight Center, Code 682.1, Greenbelt, MD 20771
Received 1991 April 9; accepted 1991 November 12

ABSTRACT

High-resolution spectra of microwave bursts from Owens Valley Radio Observatory show numerous departures from expectations based on simple thermal or nonthermal models. In particular, (1) ~80% of the events show more than one spectral peak; (2) many bursts have a low-side spectral index steeper than the maximum expected slope; and (3) the peak frequency stays relatively constant, and changes intensity in concert with the secondary peaks throughout a given event's evolution. We develop a theoretical formalism allowing both thermal and nonthermal particles to coexist in the flaring plasma. Within this formalism, we show that the observed spectral features can be accounted for through gyrosynchrotron radiation. Specifically:

1. The "secondary" components seen on the low-frequency side of many spectra are nonthermal enhancements superposed upon thermal radiation, occurring between the thermal harmonics.
2. A steep optically thick slope is accounted for by the thermal absorption of nonthermal radiation.
3. If the coexistence of thermal and nonthermal particles is interpreted in terms of electron heating and acceleration in current sheets, a changing electric field strength can account for the gross evolution of the microwave spectra.

We present our model distribution function, use it to calculate gyrosynchrotron spectra, and systematically analyze these spectra. We demonstrate the applicability of this approach by fitting an observed Owens Valley spectrum. Last, we offer our physical interpretation of the thermal/nonthermal distribution function as it relates to solar flares.

Subject headings: radiation mechanisms: cyclotron and synchrotron — Sun: flares — Sun: radio radiation

1. INTRODUCTION

Inpulsive microwave radiation carries a great deal of information about the physical conditions in the solar atmosphere immediately after the onset of a flare. The microwaves, at frequencies above 1 or 2 GHz, are widely agreed to be produced primarily by gyrosynchrotron radiation from mildly relativistic electrons spiraling along magnetic field lines in the flaring region. In principle, high-quality spectral observations of impulsive microwave emission, supplemented by spatial information, can provide information about not only the ambient magnetic field and various plasma parameters (e.g., emission measure and temperature for thermal models, electron energy distribution for nonthermal models) in the emission region but also the source size and geometry.

Such high-quality observations, as represented by those obtained from the frequency-agile interferometer at Owens Valley Radio Observatory (OVRO) (Stäli, Gary, & Hurford 1989, 1990, hereafter SGH1, SGH2), are now available. These observations are providing strong challenges to the conventional modeling of microwave bursts as being produced either by ultrahot ($T > 10^8$ K) isothermal plasmas or by simple power-law electron distributions (SGH1). In particular, the OVRO observations have shown that (1) ~80% of microwave bursts show complex spectra, containing two or more peaks in flux density; (2) slopes on the low-frequency side of the spectrum are frequently steeper than expected for nonthermal gyrosynchrotron radiation; and (3) the peak frequency of the spectrum does not shift as expected during flare development.

Some aspects of the spectra can be modeled as a superposition of spatially distinct source regions and/or of other plasma or magnetic field inhomogeneities. Simultaneous high-resolution microwave imaging of the bursts can help to clarify the spatial structures involved in complex events (Velusamy & Kundu 1982; Kundu et al. 1982; Holman, Kundu, & Kane 1989; Willson et al. 1990; Bogod et al. 1990; Gary & Hurford 1990). It is, however, difficult to see how inhomogeneities can produce steep, optically thick spectral slopes and relatively fixed peak frequencies without making very special ad hoc assumptions on a case-by-case basis.

In addition, many events show a coevolution in time of both the primary and secondary components (SGH1, SGH2), providing strong evidence that they have a common origin. In modeling the components independently, one is faced with a difficult theoretical conundrum. If the spectral components arise in the same location, how does the plasma support two noninteracting radiating populations of electrons for the duration of the burst? Conversely, if they arise in different locations, through what mechanism are these different locations coupled, strongly enough for them to coevolve throughout the course of the burst?

In this paper we develop a theoretical framework for modeling these spectra, which can account for the three above-mentioned observational challenges. We propose that the radiating electron population is neither simply thermal nor nonthermal but, in general, a combination of both. We employ the basic formulae of gyrosynchrotron radiation, and use for our electron distribution function a nonthermal tail joined to a bulk Maxwellian. We then systematically analyze the resultant spectra in terms of the interplay between thermal and nonthermal processes.

Hybrid distribution functions have been used before (e.g., Emslie & Vlahos 1980; Mok 1983) but in the context of very specific flare models and/or loop geometries. The effects of an ambient thermal plasma on a nonthermal population have been touched upon by various authors (e.g., Ramaty 1969; Klein 1987). The general approach taken here, independent of specific model assumptions, both uncovers important new effects and can serve to clarify some of the previous work.

In § 2 we briefly review the gyrosynchrotron radiation formulae, and in § 3 we introduce and discuss our model distribution function. Section 4 is devoted to a systematic analysis of the resultant theoretical spectra, and § 5 relates these results to observations and demonstrates the applicability of this approach. In § 6 we propose a possible physical origin of the thermal/nonthermal distribution function. Finally, in § 7 we summarize the results of our study.

2. GYROSYNCHROTRON FORMULAE

The following formalism is based on Ramaty (1969), as corrected by Trulsen & Fejer (1970). The reader is referred to Ramaty (1969) for further details.

The radiation intensity from a homogeneous source region is

$$I_\pm(\nu, \theta) = \frac{j_\pm}{\kappa_\pm}[1 - \exp(-\kappa_\pm D)] \text{ ergs (s sr Hz cm}^2)^{-1} . \quad (1)$$

Here j is the emissivity, κ is the absorption coefficient, D is the depth of the source, ν is the observation frequency, θ is the angle between the wave normal and the magnetic field direction $(0 \le \theta \le \pi)$, and $+$ $(-)$ refers to the ordinary (extraordinary) mode of polarization. Note that this is radiation intensity per unit *source* area. The total intensity is then $I = I_+ + I_-$.

For an electron distribution function of the form $N_r f(\gamma, \phi)$, where N_r is the number density of radiating electrons, $\gamma = (1 - \beta^2)^{-1/2}$ is the Lorentz factor with $\beta = v/c$, ϕ is the electron pitch angle, and

$$2\pi \int_1^\infty d\gamma \int_{-1}^1 d(\cos\phi) f(\gamma, \phi) = 1 ; \quad (2)$$

the emissivity and absorption coefficient are given by

$$j_\pm(\nu, \theta) = \frac{4\pi^2 e^2 \nu N_r}{|\cos\theta| c} \frac{1}{1 + a_{\theta\pm}^2} \sum_{s_{\min}}^{\infty} \int_{\gamma_1}^{\gamma_2} \Bigg\{ d\gamma f(\gamma, \phi_s)\beta^{-1} \times \left[\alpha_{\theta\pm}\left(\frac{\cot\theta}{n_\pm} - \beta\frac{\cos\phi_s}{\sin\theta}\right) J_s(x_s) - \beta \sin\phi_s J_s'(x_s)\right]^2 \Bigg\} , \quad (3)$$

$$\kappa_\pm(\nu, \theta) = \frac{4\pi^2 e^2 N_r}{mc\nu|\cos\theta|} \frac{1}{n_\pm(1 + a_{\theta\pm}^2)} \sum_{s_{\min}}^{\infty} \int_{\gamma_1}^{\gamma_2} d\gamma\, \beta^{-1} \times \left\{ -\beta\gamma^2 \frac{\partial}{\partial\gamma}\left[\frac{f(\gamma, \phi)}{\beta\gamma^2}\right] + \frac{n_\pm \beta\cos\theta - \cos\phi}{\beta^2\gamma \sin\phi} \frac{\partial}{\partial\phi} f(\gamma, \phi)\right\}_{\phi=\phi_s} \times \left[a_{\theta\pm}\left(\frac{\cot\theta}{n_\pm} - \beta\frac{\cos\phi_s}{\sin\theta}\right) J_s(x_s) - \beta\sin\phi_s J_s'(x_s)\right]^2 , \quad (4)$$

where

$$\cos\phi_s = \frac{1 - s\nu_b/\gamma\nu}{n_\pm \beta \cos\theta} , \quad x_s = \frac{s n_\pm \beta \sin\theta \sin\phi_s}{1 - n_\pm \beta\cos\theta\cos\phi_s} ,$$

$$s > \left(\frac{\nu}{\nu_b}\right)(1 - n_\pm^2 \cos^2\theta)^{1/2} ,$$

and

$$\gamma_{1(2)} = \frac{(s\nu_b/\nu)^2 + n_\pm^2 \cos^2\theta}{s\nu_b/\nu + (-)n_\pm|\cos\theta|[(s\nu_b/\nu)^2 + n_\pm^2\cos^2\theta - 1]^{1/2}} .$$

J_s is a Bessel function of order s, and the prime denotes differentiation with respect to the argument. These formulae have been rewritten from Ramaty (1969), in order to bring the summation over harmonics outside of the integral.

For ν greater than the plasma cutoff frequencies, $s > 0$. These cutoffs are: the plasma frequency, $\nu_p = (N_{\rm tot} e^2/\pi m)^{1/2}$ for the ordinary mode (o-mode); and $\nu_x = \nu_b/2 + (\nu_p^2 + \nu_b^2/4)^{1/2}$ for the extraordinary mode (x-mode). The electron gyrofrequency is $\nu_b = eB/(2\pi mc)$, and $N_{\rm tot}$ is the plasma density, including nonradiating particles. The θ-polarization coefficient, a_θ, and index of refraction are given by

$$a_{\theta\pm}(\nu, \theta) = \frac{-2\nu(\nu_p^2 - \nu^2)\cos\theta}{-\nu^2\nu_b \sin^2\theta \pm [\nu^4\nu_b^2 \sin^4\theta + 4\nu^2(\nu_p^2 - \nu^2)^2\cos^2\theta]^{1/2}} , \quad (5)$$

$$n_\pm^2(\theta) = 1 + 2\nu_p^2(\nu_p^2 - \nu^2) \times \{\pm[\nu^4\nu_b^4 \sin^4\theta + 4\nu^2\nu_b^2(\nu_p^2 - \nu^2)^2\cos^2\theta]^{1/2} - 2\nu^2(\nu_p^2 - \nu^2) - \nu^2\nu_b^2\sin^2\theta\}^{-1} . \quad (6)$$

Note that we include the effects of Razin suppression at low frequencies.

Equations (3) and (4) are valid for $\theta \ne \pi/2$. When $\theta = \pi/2$, we obtain the following expressions for j and κ:

$$j_\pm\left(\nu, \frac{\pi}{2}\right) = \frac{4\pi^2 e^2 N_r}{c} \nu_b n_\pm \sum_{s=1}^{\infty} \beta_s^2 \int_{-1}^{1} d(\cos\phi) f(\gamma_s, \phi) Y_\pm(\phi) , \quad (7)$$

$$\kappa_\pm\left(\nu, \frac{\pi}{2}\right) = \frac{4\pi^2 e^2 N_r}{mc\nu^2} \nu_b \sum_{s=1}^{\infty} \beta_s^2 \int_{-1}^{1} d(\cos\phi) Y_\pm(\phi) \times \left\{ -\beta\gamma^2 \frac{\partial}{\partial\gamma}\left[\frac{f(\gamma, \phi)}{\beta\gamma^2}\right] - \frac{\cos\phi}{\beta^2\gamma\sin\phi} \frac{\partial}{\partial\phi} f(\gamma, \phi)\right\}_{\gamma=\gamma_s,\, \beta=\beta_s} , \quad (8)$$

where

$$Y_+(\phi) = [J_s(x_s)]^2 \cos^2\phi , \quad Y_-(\phi) = [J_s'(x_s)]^2 \sin^2\phi ,$$

$$\gamma_s = s\nu_b/\nu , \quad \beta_s = (1 - \gamma_s^{-2})^{1/2} ,$$

and

$$x_s = s\beta_s n_\pm \sin\phi .$$

3. THE MODEL

3.1. *Model Distribution Function*

The above formulae are valid for the general distribution function $f(\gamma, \phi)$. We are primarily interested in thermal versus nonthermal effects, and so to gain a certain measure of simplicity, we shall discard any ϕ-dependence in f. We therefore assume an isotropic electron distribution function having both bulk thermal and nonthermal power-law tail components:

$$f(\gamma) = \left(1 - \frac{N_{\rm nt}}{N}\right) f_{\rm th}(\gamma) + R\,\frac{N_{\rm nt}}{N} f_{\rm nt}(\gamma)\,, \qquad (9)$$

where N and $N_{\rm nt}$ are respectively the number densities of electrons in the entire plasma and in the nonthermal tail alone ($N = N_{\rm th} + N_{\rm nt}$, with $N_{\rm th}$ the density of thermal particles), and

$$f_{\rm th}(\gamma) = \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{mc^2}{kT}\right)^{3/2} \left(1 + \frac{15}{8}\frac{kT}{mc^2}\right)^{-1} \gamma(\gamma^2 - 1)^{1/2} \times \exp\left[\frac{-mc^2}{kT}(\gamma - 1)\right], \qquad (10)$$

$$f_{\rm nt}(\gamma) = (4\pi)^{-1}(\delta - 1)(\gamma_{\rm cr} - 1)^{\delta - 1}(\gamma - 1)^{-\delta}\,, \qquad (11)$$

$$\frac{N_{\rm nt}}{N_{\rm th}} = (32\pi)^{1/2} \left(\frac{mc^2}{kT}\right)^{3/2} \left(1 + \frac{15}{8}\frac{kT}{mc^2}\right)^{-1} \times \frac{\gamma_{\rm cr}(\gamma_{\rm cr} - 1)(\gamma_{\rm cr}^2 - 1)^{1/2}}{\delta - 1} \exp\left[-\frac{mc^2}{kT}(\gamma_{\rm cr} - 1)\right], \qquad (12)$$

$$\gamma_{\rm cr} = \left(\frac{p_{\rm cr}^2}{m^2c^2} + 1\right)^{1/2} = \left[\frac{(p_{\rm th}/mc)^2}{\epsilon - (p_{\rm th}/mc)^2(1 - \epsilon)} + 1\right]^{1/2}, \quad \text{with } \frac{kT}{mc^2} < \epsilon \lesssim 0.15\,, \qquad (13)$$

$$R = \begin{cases} 0 & \text{if } \gamma < \gamma_R = \gamma(p_R) \text{ with } p_R = p_{\rm cr} - \dfrac{p_{\rm th}}{2}\,, \\ \sin\left[\dfrac{(\pi/2)(\gamma - \gamma_R)^2}{(\gamma_{\rm cr} - \gamma_R)^2}\right] & \text{if } \gamma_R < \gamma < \gamma_{\rm cr}\,, \\ 1 & \text{if } \gamma > \gamma_{\rm cr}\,. \end{cases} \qquad (14)$$

Equations (10) and (11) are respectively the thermal and nonthermal components of the distribution, in the form used by Dulk & Marsh (1982). The thermal plasma has kinetic electron temperature T, k is Boltzmann's constant, $mc^2 = 511$ keV, and γ is the Lorentz factor. The nonthermal tail has spectral index δ and low-energy cutoff $E_{\rm cr} = (\gamma_{\rm cr} - 1)mc^2$. Equation (10) is valid for $T < 3 \times 10^9$ K. Relativistically, the thermal momentum is given by

$$\frac{p_{\rm th}}{mc} = \left(\frac{mc^2}{kT} - 1\right)^{-1/2}.$$

We shall refer to a distribution function of the type of equation (9) as a "TNT" (thermal/nonthermal) distribution, and we will call a spectrum generated by such a distribution function a TNT spectrum. Similarly, distributions and spectra from the first and second terms of equation (9) will be designated by "TH" and "NT" respectively. We emphasize that the physical TNT distribution which we expect in solar flares is a single, unified, smooth distribution of momenta, reflecting cotemporal heating and particle acceleration. The separation (eq. [9]) is a mathematical, and especially an analytical, convenience.

There are several things to notice about our TNT distribution function. First, since equations (10) and (11) are separately normalized to unity, the overall normalization of the TNT distribution is preserved (up to the negligible factor of R, which we will discuss shortly).

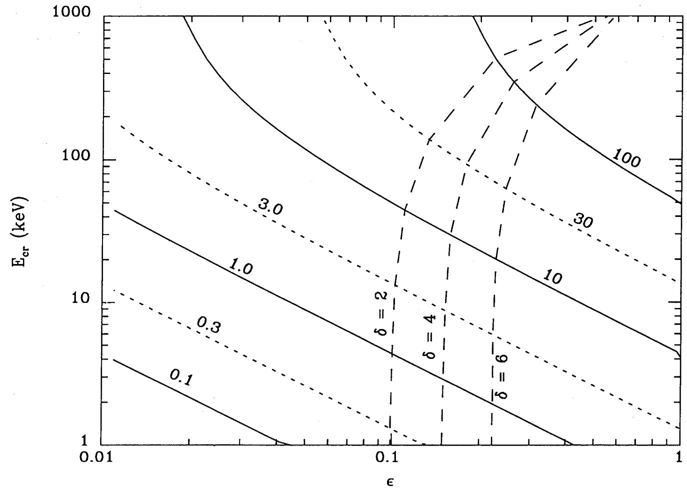

FIG. 1.—$E_{\rm cr}$ vs. ϵ for 10^6 K $< T < 10^9$ K. The alternating solid and dashed lines are for clarity only and are labeled in units of 10^7 K. The long-dashed lines show the suggested upper limits on ϵ for $\delta = 2, 4$, and 6 as discussed in the text.

Second, the critical Lorentz factor, $\gamma_{\rm cr}$ (or, alternatively, ϵ), acts in two roles. It sets the low-energy cutoff for $f_{\rm nt}$ as $E_{\rm cr} = (\gamma_{\rm cr} - 1)mc^2$, and also sets the normalization of $f_{\rm nt}$ with respect to $f_{\rm th}$ by demanding that

$$f_{\rm nt}(\gamma_{\rm cr}) = f_{\rm th}(\gamma_{\rm cr})\,, \qquad (15)$$

which in turn leads to equation (12). This method of joining the TH and NT distributions is reasonable, but not unique. In § 6 we interpret the TNT distribution function in terms of runaway electron acceleration and heating in current sheets, motivating this normalization. Figure 1 shows the conversion from ϵ to $E_{\rm cr}$, where the relativistic restriction on runaway production, $\epsilon > kT/mc^2$ (Connor & Hastie 1975), is clearly seen.

For sufficiently low $\gamma_{\rm cr}$ our TNT distribution function becomes unphysical, due to our choice of a power-law description for $f_{\rm nt}$. As $\gamma_{\rm cr}$ decreases, an ever-increasing fraction of the nonthermal particles are found at lower energies. At $\gamma_{\rm cr} = \gamma_{\rm th}$ most of the "nonthermal" particles have nearly thermal energies, and the concept of nonthermal particles loses its meaning.

The upper limit on ϵ in equation (13) serves to avoid this problem. It was determined by imposing the (somewhat arbitrary) condition $N_{\rm nt} < 0.25 N_{\rm th}$. Equation (12) then provides a maximum ϵ, for a given T and δ. This suggested $\epsilon_{\rm max}$ is plotted as dashed lines on Figure 1 for $\delta = 2, 4$, and 6.

We also point out that as $\gamma_{\rm cr} \to \infty$ ($\epsilon \to kT/mc^2$) equation (12) tells us that the nonthermal contribution to f vanishes.

Last, it can be seen from equations (9) and (15) that if we abruptly "switched on" the nonthermal component of f at $\gamma_{\rm cr}$, a discontinuity in f, on the order of $(N_{\rm nt}/N) f_{\rm th}(\gamma_{\rm cr})$, would result. The factor R, which "ramps up" the nonthermal tail from zero at some $\gamma_R < \gamma_{\rm cr}$ to its full value at $\gamma_{\rm cr}$, serves to smooth this out. Both the form of R and the value of γ_R are quite arbitrary, so long as R varies smoothly from 0 to 1 and is not too broad near $\gamma_{\rm cr}$. We have experimented with these quantities and have found no discernible differences in the resultant microwave spectra with changes in either R or γ_R. For completeness we nevertheless include it, choosing the form shown in equation (14).

3.2. *Model Calculations*

We have used a numerical code based on § 2, allowing for Bessel functions to be calculated out to order 1050 as needed. This allows the calculation of gyrosynchrotron spectra, with a 1% accuracy, from the plasma cutoff frequency out to about the 35th harmonic of the electron gyrofrequency. We will show that the most interesting phenomena occur at the lower harmonics ($s < 10$), and thus the code is quite inclusive. The integrals are performed via the Gaussian quadrature method, allowing the use of up to 256 points for estimating the value of the integral. If the desired accuracy cannot be achieved, the code uses the cautious, adaptive Romberg extrapolation method. Into this code we insert our model distribution function equations (9)–(14). We allow the Lorentz factor γ to range from very close to unity out to $\gamma = 10$. We cannot allow $\gamma = 1$ because equations (3) and (4) are singular for particles at rest. Particles having energies much in excess of 1 MeV contribute negligible amounts to the radiation we are studying. We have found that lowering the high-energy cutoff from $\gamma = 10$ has no effect on the resulting spectra until γ_{max} reaches ~ 2.5–3, when it starts to reduce the higher harmonics in the optically thin part of the spectrum. Our range of γ is thus quite inclusive. We have found excellent agreement with Ramaty (1969) for the NT case.

In order to compare with observations, we choose a depth, D, for the radiating source and assign the source an effective area A. This allows us to compute the radiation intensity in solar flux units (1 sfu = 10^{-19} ergs cm^{-2} s^{-1} Hz^{-1}) for each polarization mode.

There are eight necessary input parameters for computing a spectrum: T and N_{th}, which characterize the thermal contribution to the distribution; ϵ and δ, which characterize the nonthermal component; the magnetic field strength B, source depth D, and area A, which characterize the emitting region in which the single TNT distribution resides; and the viewing angle θ, between the wavevector and the magnetic field direction, characterizing our orientation with respect to the source region.

We can calculate three complete spectra for each set of input parameters, based separately on equations (9)–(11), i.e., spectra of type TNT, TH, and NT respectively. This allows us to assess the relative contributions of thermal and nonthermal particles at any frequency in a TNT spectrum.

In addition, we can separate out the individual harmonics of both the emission and absorption coefficients, in each of the two polarization modes, for all three types of spectra.

4. RESULTS/DISCUSSION

4.1. *General Results*

The spectrum of Figure 2 shows many of the features typically seen. Our plotting convection is the following: the solid line is the calculated TNT spectrum, the short-dashed line is NT, and the long-dashed line is TH. The input parameters are $T = 2 \times 10^8$ K, $N_{th} = 10^9$ cm^{-3}, $\delta = 3.0$, $\epsilon = 0.084$, $B = 200$ G, $D = 10^8$ cm, $A = 10^{16}$ cm^2, $\theta = 45°$. Figure 3 shows the first six harmonics of emissivity in each polarization mode, using the same plotting convention, with the x-axis now being harmonic number. Figure 3*a* shows the ordinary mode (s = 1–6), and Figure 3*b* shows the extraordinary mode (s = 2–7). Harmonics of the absorption coefficient are qualitatively very similar. The polarization is shown as the solid line in Figure 11.

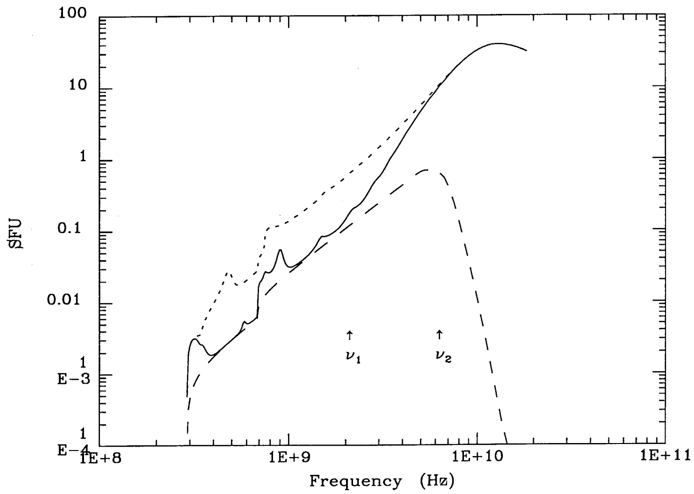

FIG. 2.—Typical TNT gyrosynchrotron spectrum as discussed in the text. Our plotting convention for this and all figures, unless explicitly stated otherwise, is the following: solid line = TNT spectrum; short-dashed line = NT; long-dashed line = TH. The transition frequencies are near ν_1 and ν_2. The input parameters are $T = 2 \times 10^8$ K, $N_{th} = 10^9$ cm^{-3}, $\delta = 3$, $\epsilon = 0.084$, $B = 200$ G, $D = 10^8$ cm, $\theta = 45°$, $A = D^2$.

We first point out some general features shared, to a greater or lesser extent, by all spectra of the TNT type.

1. At the lowest frequencies, the TNT spectrum closely follows a spectrum produced by purely thermal electrons, except in certain wave bands where the TNT spectrum can be enhanced over the TH spectrum.
2. At the highest frequencies, the TNT spectrum is indistinguishable from a purely NT spectrum, i.e., the bulk thermal population does not contribute to the highest energy radiation.
3. The transition between these two regions is characterized by two frequencies which we designate ν_1 and ν_2, with approximate locations shown on Figure 2. Between these two frequencies the TNT spectrum can have a steep positive slope.

All of these features are understood by noting that the expressions for the emission and absorption coefficients, equations (3) and (4), are linear in f, which itself is linear in f_{th} and f_{nt}. In other words, the radiative transfer equation (1) becomes

$$I(\nu, \theta) = \frac{j_{th} + j_{nt}}{\kappa_{th} + \kappa_{nt}} \{1 - \exp[-(\kappa_{th} + \kappa_{nt})D]\} \quad \text{ergs (s sr Hz cm}^2)^{-1} \quad (16)$$

for each polarization mode. This additivity is clearly seen in Figures 3*a* and 3*b*. The overall intensity spectrum in each polarization is determined, at each frequency, by whichever component (TH or NT) is dominating in each of emission and absorption at that frequency.

For example, the lower transition frequency, ν_1, is that at which the intensity of the TNT spectrum begins to exceed that of the TH spectrum, neglecting the enhancements. At frequencies below ν_1 thermal emission and absorption dominate; that is, we are in the Rayleigh-Jeans regime, and the slope on a log (flux density) versus log (frequency) plot is 2. Nonthermal radiation is completely suppressed by the thermal plasma, again neglecting the enhancements.

At frequencies near ν_1 the nonthermal emissivity begins to contribute, although absorption is still dominated by the thermal electrons. Above ν_1 the nonthermal emissivity con-

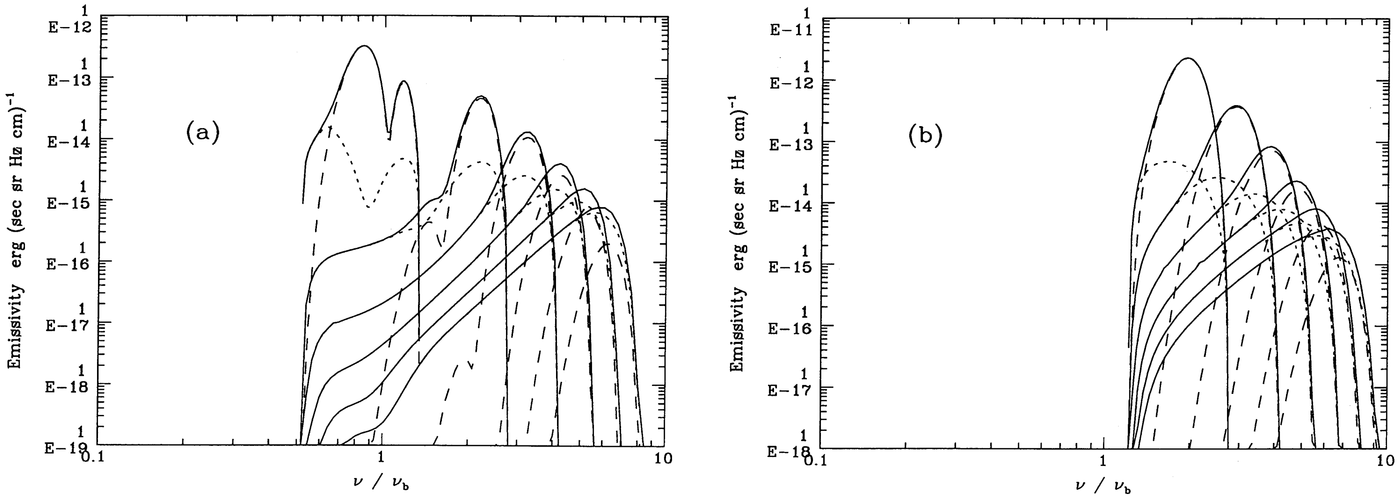

FIG. 3.—First six harmonics of the emissivity for the spectrum of Fig. 2, using the same plotting convention. The x-axis is now harmonic number. (a) Ordinary mode. (b) Extraordinary mode.

tinues to grow relative to the thermal, eventually dominating completely at some frequency between ν_1 and ν_2. As the frequency is increased to approach the upper transition frequency ν_2, thermal absorption contributes less and less relative to that due to the nonthermal electrons. Since nonthermal emission is already dominating thermal emission, above ν_2 the spectrum is produced completely by nonthermal processes.

Between ν_1 and ν_2 the interesting situation arises in which the nonthermal particles are emitting more radiation than their thermal counterparts ($j_{\rm nt} > j_{\rm th}$), but the thermal particles are able to absorb much more of it than nonthermal self-absorption alone can account for ($\kappa_{\rm th} > \kappa_{\rm nt}$). This additional thermal absorption of nonthermal radiation explains the steep slope seen in this region of the spectrum. In Figure 2 the slope from 2.5 to 5.5 GHz is $\sim$4.3. Optically thick TNT slopes as steep as 7 have been seen in the course of this study, and it is likely that steeper slopes are possible. Note, as in Figure 4b, that although steep slopes are naturally accounted for in this model, they are not inevitable, especially at very high temperatures.

The index of refraction for the plasma of Figure 2 is $\sim$0.8 at 4×10^8 Hz, and decreases rapidly for lower frequencies. This is the regime of Razin suppression, close to the plasma frequency cutoff.

Some additional discussion of point 1 above is now in order, the enhancements mentioned there being of prime observational importance. The lowest harmonics in a gyrosynchrotron spectrum do not overlap and blend as much as the higher ones, as can be seen in Figures 3a and 3b. This is especially evident for thermal radiation, and the harmonics become still narrower for high viewing angles. The abundance of high-energy electrons in a nonthermal distribution effectively creates a low-frequency broadening of each harmonic in the NT spectrum, since an energetic electron resonates with a Doppler-shifted gyrofrequency, ν_b/γ. As discussed above and shown in Figure 3, for $\nu < \nu_1$ the NT harmonics will have lower peak intensities than the TH harmonics; the spectrum is in the Rayleigh-Jeans regime. Nevertheless, the Doppler-broadened NT harmonics can contribute in the frequency bands between the TH harmonics. Within these bands, not only does the NT emissivity exceed the TH emissivity, but the TH absorption is also too low to suppress the NT radiation. It is this limited NT contribution, *between* the TH harmonics, which produces the enhancements seen in the theoretical spectra. Since TH harmonics become narrower with either decreasing temperature or increasing viewing angle, it is evident that these NT enhancements can occur between any of the higher TH harmonics that are not sufficiently blended together (cf. Figs. 4a and 10).

4.2. *Variations with Parameters*

Figures 4–10 show how a basic TNT microwave spectrum (shown in the figures as a dot-dash line), similar to that of Figure 2, will typically vary as each of the input parameters is varied in turn. These displayed variations consist of seven mutually orthogonal excursions from a single point in seven-dimensional parameter space. The general trends depicted will certainly hold throughout much of parameter space. We do not display variations with area, A, since this factor is a premultiplier and thus simply raises or lowers the intensity of the spectrum *in toto*, leaving its shape unaffected.

We will use the following notation in this discussion: $I_{\rm peak}$ is the intensity at $\nu_{\rm peak}$, the peak frequency. RJ refers to the low-frequency (Rayleigh-Jeans, thermal) portion of the spectrum, having $\nu < \nu_1$, whether or not enhancements are present. PC refers to the primary component (in the nomenclature of SGH1 and SGH2), i.e., all frequencies $\nu > \nu_1$. The enhancements that appear on the optically thick side we will designate H-n, with $n = 1$, 2, etc. H-1 will always refer to the o-mode nonthermal enhancement below the gyrofrequency ($\nu_b = 1.12 \times 10^9$ Hz for the basic spectrum). Recall that these features are *not* harmonics, but rather are produced between the thermal harmonics. In § 5 we will identify one or more of the H-n with the secondary components of SGH2.

The basic spectrum which we vary has these input parameters: $T = 7 \times 10^7$ K, $N_{\rm th} = 10^9$ cm^{-3}, $\delta = 4.0$, $\epsilon = 0.06$, $B = 400$ G, $D = 10^8$ cm, $\theta = 45°$, $A = D^2 = 10^{16}$ cm^2. Note that the vertical scales of Figures 4–10 vary.

4.2.1. *Temperature T*

Figures 4a and 4b show that as the temperature rises, the TNT spectrum varies in four major ways: (1) The underlying thermal spectrum is raised in intensity (and shifts to a slightly higher peak frequency). This shifts the RJ part of the TNT spectrum upward in intensity. (2) $I_{\rm peak}$ increases. (3) The $\nu_{\rm peak}$

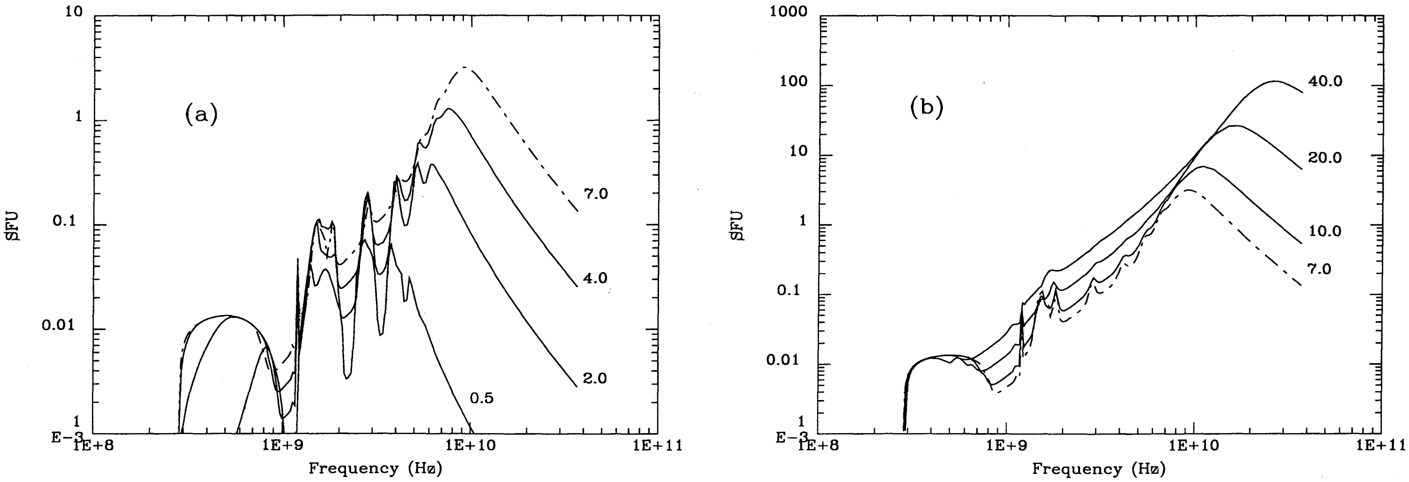

FIG. 4.—Variations with temperature of our basic TNT spectrum (*dot-dash line*). The basic spectrum gas input parameters $T = 7 \times 10^7$ K, $N_{\text{th}} = 10^9$ cm^{-3}, $\delta = 4.0$, $\epsilon = 0.06$, $B = 400$ G, $D = 10^8$ cm, $\theta = 45°$, $A = 10^{16}$ cm^2). The spectra are labeled in units of 10^7 K. (*a*) $T = 5 \times 10^6$ and 2, 4, and 7×10^7 K. (*b*) $T = 7 \times 10^7$ and 1, 2, and 4×10^8 K. Note the different vertical scales in (*a*) and (*b*).

moves to higher values. (4) The H-*n* features become less prominent owing to the increased Doppler broadening of the thermal harmonics. At the highest temperatures, the optically thick spectrum is quite smooth, changing character gradually from classic thermal (slope ~ 2) to nonthermal (slope ~ 3) (Dulk & Marsh 1982; SGH1). At intermediate temperatures, the H-*n* appear and the transition is more abrupt, having an optically thick slope greater than 3. As the temperature is lowered, the H-*n* become more apparent, because of the narrowing of the thermal harmonics, i.e., the wave bands between the thermal harmonics are wider.

4.2.2. *Density N_{th}*

Figure 5 shows variations with density of the basic TNT spectrum. The main changes as the density increases are the following: (1) The plasma frequency cutoff moves to higher frequencies, reaching 1 GHz at $N_{\text{th}} \sim 10^{10}$ cm^{-3}. The total number of thermal and nonthermal particles increases, since $N_{\text{nt}} \propto N_{\text{th}}$, all else remaining unchanged (eq. [12]). Thus (2) I_{peak} increases relative to RJ, and (3) ν_{peak} increases, both resulting from the increasing optical depth of the source. (4) As the plasma frequency increases, the H-*n* structures are successively lost below the cutoff.

4.2.3. *Spectral Index δ*

Hardening the spectral index increases the nonthermal nature of the TNT spectrum (Fig. 6). Specifically, (1) I_{peak} increases dramatically with respect to RJ; (2) ν_{peak} increases; (3) the H-*n* features become increasingly pronounced, consistent with our discussion at the end of § 4.1; (4) the lower transition frequency ν_1 is shifted downward, i.e., the TNT spectrum departs from TH sooner; and (5) the width of the PC increases for two reasons: as the NT contribution increases, more of the optically thick spectrum is at frequencies above ν_2, and has a slope of ~ 3 (Dulk & Marsh 1982; SGH1), whereas the slope between ν_1 and ν_2 is steeper than this; the optically thin slope flattens with harder δ, which further widens the PC.

The dashed line in Figure 6 is the associated purely thermal spectrum, and is closely approached by $\delta = 8$.

4.2.4. *Low-Energy Cutoff ϵ*

We see in Figure 7 that as ϵ rises, i.e., as the nonthermal tail "climbs up" the Maxwellian to lower energies, the dominant

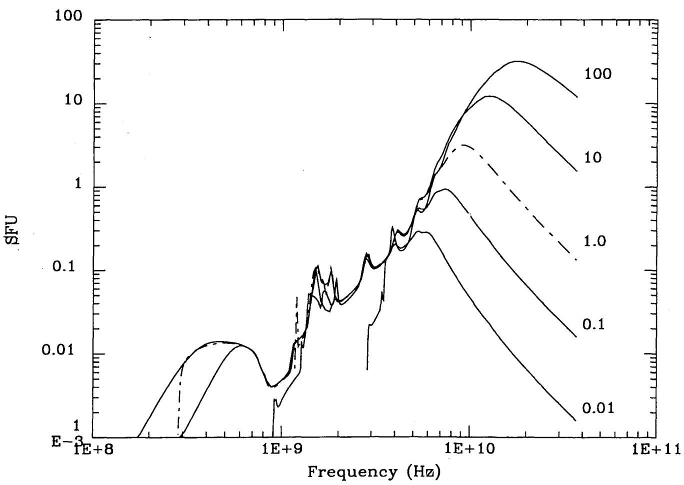

FIG. 5.—Variations with density of our basic spectrum. Here, 10^7 cm^{-3} < N_{th} < 10^{11} cm^{-3}. The spectra are labeled in units of 10^9 cm^{-3}.

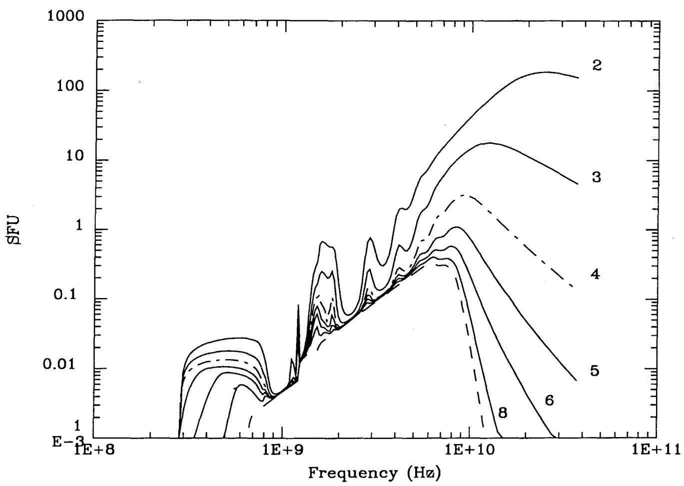

FIG. 6.—Variations with electron spectral index $2 < \delta < 8$. The dashed line is the associated TH spectrum.

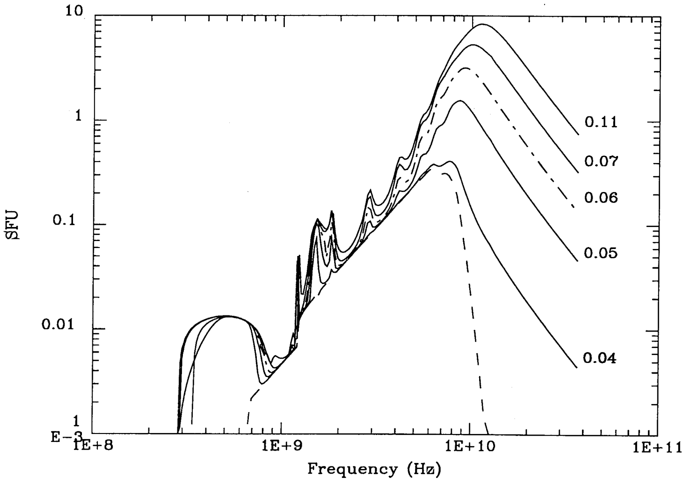

FIG. 7.—Variations with nonthermal parameter, $0.04 < \epsilon < 0.11$. The dashed line is the associated TH spectrum.

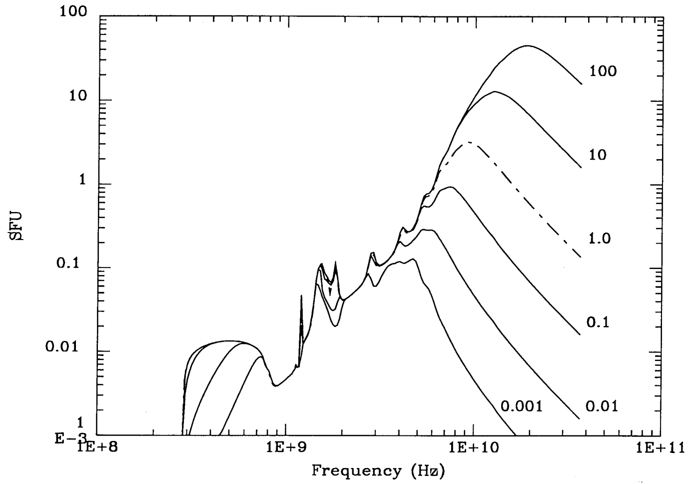

FIG. 9.—Variations with source depth. The spectra are labeled in units of 10^8 cm. $D = 10^5, 10^6, 10^7, 10^8, 10^9$, and 10^{10} cm.

changes are (1) to the PC, which increases in intensity relative to RJ, and (2) to the H-*n*, which also rise relative to RJ, consistent with our previous discussion. The PC drifts slowly to higher $\nu_{\rm peak}$. Decreasing ϵ below what is shown renders the TNT spectrum indistinguishable from the associated TH spectrum (*dashed line*), until at some very low intensity level the optically thin thermal slope of ~8–10 will change to $\sim 1.2 - 0.9\delta$ (§ 5.1), when the nonthermal particles are finally able to contribute to the radiation.

4.2.5. *Magnetic Field B*

The intensity of gyrosynchrotron radiation is very sensitive to magnetic field strength, increasing greatly with B. To compensate for this, the three spectra in Figure 8 have been multiplied by $400/B$ for clarity.

Increasing the magnetic field strength from B_1 to B_2 also translates any spectral feature to a higher frequency. This effect is nearly linear, and the frequency translation can often be approximated by a factor of $\sim B_2/B_1$, since a given feature is found near a particular multiple of the gyrofrequency $\nu_b \sim 2.8 \times 10^6 B$. This rough scaling of frequency with B holds as long as the feature at the lower field strength is not too close to ν_p, the plasma frequency.

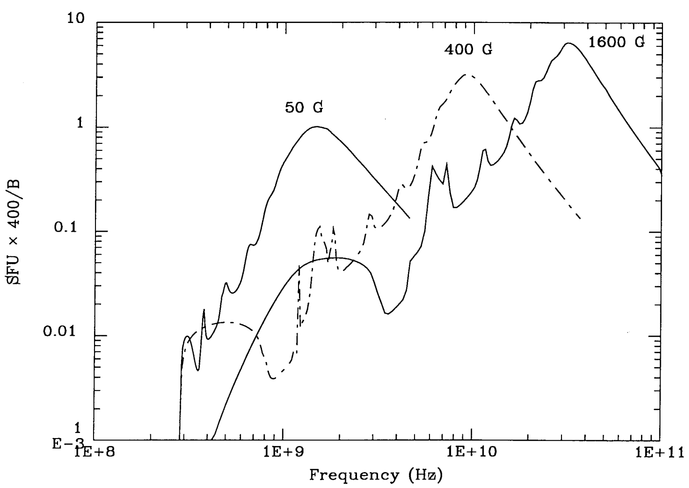

FIG. 8.—Variations with magnetic field strength 50 G < B < 1600 G. Note that the vertical scale is different for each of the three displayed spectra.

4.2.6. *Source Depth D*

Figure 9 shows the change with depth. Increasing the source depth is commensurate with increasing its optical depth, since the total number of particles along the line of sight increases. Thus we see the PC rise in intensity and shift to higher $\nu_{\rm peak}$. The H-*n* are unaffected over a wide range of source depths, being reduced for the smallest *D*'s.

4.2.7. *Observation Angle θ*

As is known, the harmonic structure of gyrosynchrotron radiation is much more pronounced when viewed at high angles to the magnetic field (cf. Takakura 1960; Takakura & Scalise 1970). This results from all harmonics having both a greater contribution to the spectrum and a narrower bandwidth. Of course, the nonthermal harmonics are still Doppler-broadened compared with the thermal harmonics.

The observed consequences are that $I_{\rm peak}$ increases and shifts to higher $\nu_{\rm peak}$, while the individual enhancements become much more apparent. This is borne out in Figure 10. In addition, we find that at low angles, the PC narrows and can become quite sharply peaked.

4.3. *Polarization*

In Figure 11 we show the fractional polarization

$$\Pi = \frac{I_+ - I_-}{I_+ + I_-} \tag{17}$$

for the TNT spectrum of Figure 2 (*solid line*) as well as for our basic spectrum (*dot-dash line*). These polarization characteristics are typical for TNT spectra in general: (1) the optically thin emission is *x*-mode polarized, and can exceed 50%; (2) neglecting enhancements, the average optically thick polarization is ~5%–25% in the *o*-mode; (3) the *o*-mode polarization of the enhancements can exceed 50%; (4) where pure thermal emission occurs between the H-*n*, it is unpolarized; (5) below the *x*-mode cutoff, all of the emission is *o*-mode. These results are consistent with what is expected from one or the other of thermal and nonthermal particle distributions (Ramaty 1969).

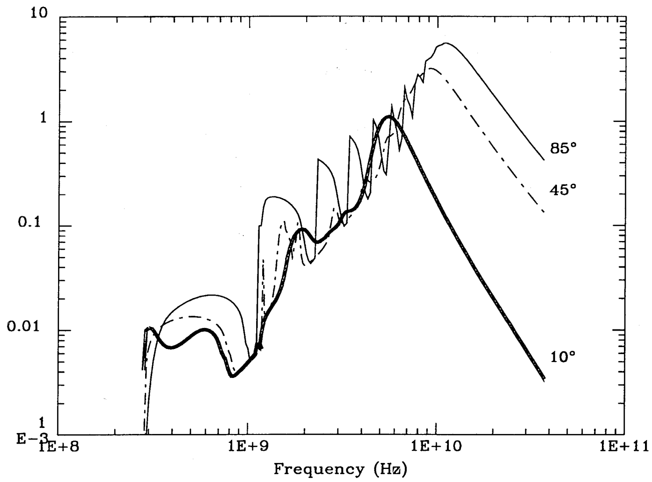

FIG. 10.—Variations with observation angle between the line of sight and the magnetic field, $10° < \theta < 85°$. The bold line for $\theta = 10°$ is for clarity.

We stress that the polarization discussed here is as it leaves a homogeneous source, with no consideration given to an intervening plasma. It is known that polarization can be reduced by propagation effects, and even reversed if the component of magnetic field along the line of sight goes to zero between the source and observer (Kundu 1965; Zheleznyakov 1970; Bandiera 1982). Observational examples of these effects are numerous (Tanaka & Kakinuma 1959; Kundu et al. 1977; Alissandrakis & Kundu 1984; Kundu & Alissandrakis 1984; Bogod et al. 1990; Brosius et al. 1992). We advise caution in drawing conclusions based on polarization data without an investigation of the intervening plasma, especially at the lower frequencies.

4.4. *The Lowest Harmonic*

The lowest harmonic which contributes to a spectrum is invariably double-peaked, whether the gyrofrequency is above or below the plasma frequency. It is always *o*-mode (below the *x*-mode cutoff) and can be any harmonic of the gyrofrequency (it is $s = 1$ only for $\nu_p \leq \nu_b$). It is possible to observe this harmonic if it corresponds to $s \geq 3$ (cf. § 5).

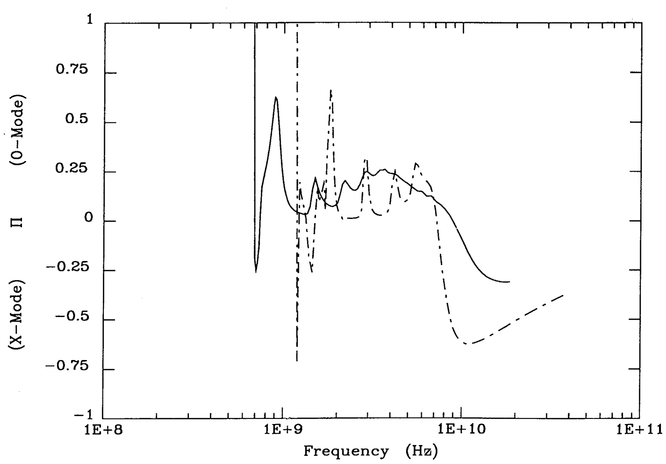

FIG. 11.—Fractional polarization (eq. [17]) for the TNT spectrum of Fig. 2 (*solid line*) and for the basic spectrum of Figs. 4–10 (*dot-dash line*). Below their respective *x*-mode cutoffs, the spectra are 100% *o*-mode polarized.

To understand this structure, we surmise from Figure 3 that the existence of twin peaks is independent of the particular electron distribution function used. This leads us to choose a δ-function distribution, $f \sim \delta(\gamma - \gamma_0)$, insertion of which into equation (3) for a single fixed harmonic $s = S$ leads us to the relatively simple equation

$$j_{\pm} = \frac{4\pi^2 e^2 N}{|\cos\theta| c} \frac{1}{(1 + a_{\theta\pm}^2)} \frac{\nu}{\beta_0} \times \left[a_{\theta\pm}\left(\frac{\cot\theta}{n_{\pm}} - \beta_0 \frac{\cos\phi_S}{\sin\theta}\right) J_S(x_S) - \beta_0 \sin\phi_S J'_S(x_S)\right]^2 . \quad (18)$$

In Figure 12 we plot the *o*-mode emissivity for $S = 1$ for various values of γ_0, where the different line types are used for clarity only and do not correspond to our usual plotting convention. The double-peaked structure is confirmed for the *o*-mode, and persists but is modified for higher S (not shown). Note that higher energy electrons radiate at lower Doppler-shifted frequencies, as expected.

This structure is related to the variability with frequency of the θ-polarization coefficient, equation (5). When we set a_θ equal to a constant, the two peaks do not appear. Equation (5) for $a_{\theta\pm}$ is a monotonically decreasing function of frequency, but differs markedly in the two polarization modes. In the *o*-mode, $a_{\theta+}$ ranges from values very much greater than unity at the lowest frequencies to asymptotically approach unity at high frequencies. On the other hand, $a_{\theta-}$ decreases only gradually, from 0 to -1. The changing interplay between the two terms in square brackets in equation (18) (and the analogous terms in eqs. [3] and [4]) produces the observed structure in the *o*-mode harmonic. That is, at the higher frequencies within this harmonic, the two terms generate the usual rising and falling harmonic structure. As the frequency is decreased, however, at some point the increasing value of $a_{\theta+}$ gives additional weight to the first term so that it again dominates. Finally, at the lowest frequencies, either Razin suppression or a lack of energetic particles (as in a Maxwellian) cuts off this "anomalous" harmonic entirely. Although this effect can persist for higher harmonics, the fact that $a_{\theta+}$ is approaching unity means that, for the higher harmonics, it is less apparent

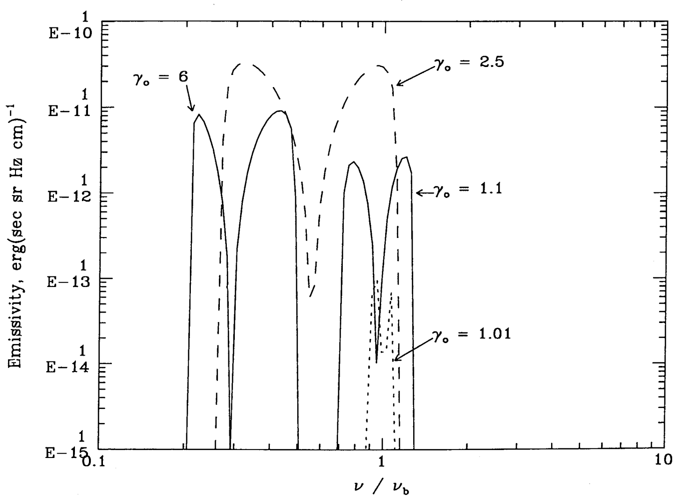

FIG. 12.—First contributing harmonic (in this case $s = 1$) of the emissivity from a δ-function electron distribution, as discussed in the text, for $\gamma_0 = 1.01$, 1.1, 2.5, 6. In this figure the different line types are for clarity only and do not correspond to our usual plotting convention.

(the first term in the square brackets is not as heavily weighted). We remark that this *o*-mode structure has been noticed before (Tamor 1978), but the explanation of it was incorrect (Tamor 1979).

5. MODELING OBSERVED SPECTRA

5.1. *A "Typical" Spectrum*

It is beyond the scope of this paper to model any flare in detail, which would ideally involve modeling a time sequence of spectra, throughout the event, together with simultaneous observations in other wavelengths, especially X-rays. We have, however, tested the thermal/nonthermal approach against observations by fitting a "typical" spectrum from SGH1. Those authors state: "More than 80% of the microwave events have complex spectra consisting of more than one component. About 80% of the minor spectral components occur on the low frequency side of the main peak in the spectrum, i.e. at frequencies lower than the peak frequency." In Figure 13 we show our fit to a spectrum (SGH1, their Fig. 2*d*) which is representative of this large class of events. The observations are shown by the crossed bars. The horizontal bars represent the fixed 50 MHz channel width of the receivers. The vertical error bars, centred on the observed flux intensity, are ± 1 $\sigma = 1.0 + 0.1 \times$ flux (G. J. Hurford 1991, private communication). The deduced fitting parameters are $T = 7.6 \times 10^7$ K, $N_{th} = 6 \times 10^9$ cm^{-3}, $\delta = 4.5$, $\epsilon = 0.078$, $B = 340$ G, $D = 5.0 \times 10^6$ cm, $A = 5.0 \times 10^{17}$ cm^2, $\theta = 57°$. The reduced χ^2 statistic is 1.449, giving a χ^2 probability function of 99.3%. For this spectrum, $E_{cr} = 48.7$ keV.

It is seen that our fit is very good for $\nu \gtrsim 2$ GHz, and less acceptable at lower frequencies. This is to be expected, for the following reason. The first and second harmonics of gyrosynchrotron radiation have difficulty escaping from the solar corona, because of the likelihood of their being reabsorbed as second and third harmonics of a weaker magnetic field. The reabsorption is strongest for the gyrofrequency and diminishes with successive harmonics. We thus expect that gyrosynchrotron radiation from a homogeneous source will provide a good

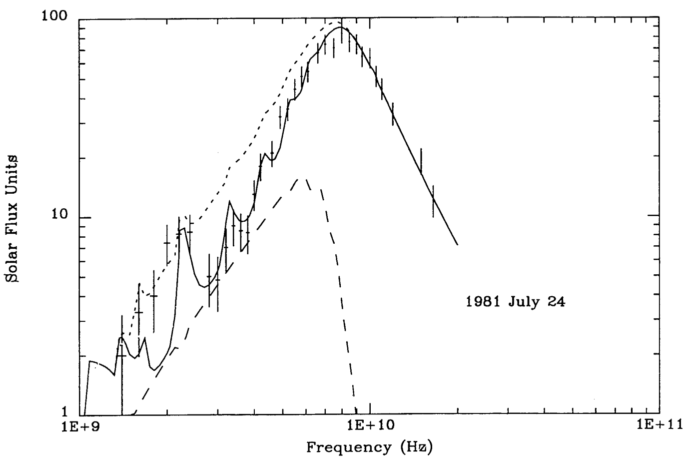

FIG. 13.—Theoretical fit to Fig. 2*d* of Stäli, Gary, & Hurford 1989. The flare occurred on 1981 July 24. The emission is assumed to be entirely gyrosynchrotron radiation from a TNT distribution function. Our plotting convention of Figs. 2–10 is used. This complex observed spectrum is well fitted except at the lowest frequencies, as discussed in the text. The derived parameters are $T = 7.6 \times 10^7$ K; $N_{th} = 6 \times 10^9$ cm^{-3}; $\delta = 4.5$; $\epsilon = 0.078$; $B = 340$ G; $D = 5.0 \times 10^6$ cm; $A = 5.0 \times 10^{17}$ cm^2; $\theta = 57°$.

fit only at frequencies $\nu \gtrsim 3\nu_b$. In other words, we generally expect the theoretical H-1 and H-2 enhancements to be largely unobservable, while the observed H-3 feature will be somewhat modified. This is consistent with what we see in Figure 13, where the "secondary component" near 2 GHz corresponds to H-3 ($3\nu_b = 2.8$ GHz). We have also fitted the spectrum assuming this enhancement to be both H-2 and H-4, but the lowest χ_ν^2 results for H-3. At frequencies lower than $3\nu_b$, external absorption and/or other emission mechanisms may be important.

A major result of the TNT model is the natural production of low-frequency enhancements. The presence, both observationally and theoretically, of these features provides a powerful diagnostic for flare magnetic field strengths and plasma temperatures.

From the discussion at the end of § 4.1, it follows that the dips in the spectrum between the enhancements mark the harmonics of the thermal gyrosynchrotron spectrum. The frequencies of these local minima are then related to the magnetic field by $\nu_{lm} \sim 2.8 \times 10^6 sB$, with s a positive integer. The separation in frequency between successive minima changes with the s chosen, and so observation of two or more such minima can rapidly lead to a determination of the field strength in the radiation-producing region, to within ~ 10 G or less. This is how the magnetic field strength was deduced for the spectrum of Figure 13.

The intensity of the radiation at these local minima also provides a diagnostic for the temperature in the region. Deep minima imply low-intensity thermal emission and thus a low temperature (cf. Fig. 4). The converse applies to shallow minima. Using the first observed minimum at or above $3\nu_b$ provides the best estimate for the temperature, since higher frequencies are increasingly influenced by the nonthermal particles. The temperature deduced for Figure 13 is hotter than the typical flare soft X-ray plasma and consistent with the superhot component seen in hard X-rays by Nitta, Kiplinger, & Kai (1989).

Since ν_{peak} often occurs near or above the 10th harmonic, the negative slope of the optically thin part of the spectrum provides a diagnostic for the nonthermal electron spectral index, through the empirical approximation (slope) $\sim 1.22 - 0.9\delta$ (Dulk & Marsh 1982). If brightness temperature is measured instead of flux density, the relation is (T_B slope) $\sim -0.78 - 0.9\delta$. These expressions are approximately valid for $\delta < 5$ and $20° < \theta < 80°$. For θ near 20°, the range of δ can be extended to ~ 7.

The presence of fine structure along the steep optically thick slope between ν_1 and ν_2 indicates that the higher order H-*n* structures are not completely blended. This in turn can provide a diagnostic for the viewing angle θ, to which these structures are very sensitive (cf. Fig. 10).

The success of the eight-parameter fit for a spectrum as complex as that of Figure 13 is encouraging. For instance, only one "primary" and one "secondary" component would require two peak intensities, two peak frequencies, and four associated slopes for a two-triangle description. We also point out that OVRO samples up to 40 frequencies between 1 and 18 GHz, a number much greater than our eight fitting parameters.

In practice, the TNT model usually provides a *seven*-parameter fit, because the optical depth is determined by the product *ND*, density entering independently only at the lowest frequencies (near ν_p).

As discussed above, the most easily determined parameters are magnetic field strength B, temperature T, and spectral index δ. Having determined three parameters using gross features of the spectrum, one is faced with an iterative process for determining four more. ND determines in part both the peak frequency and the relative intensities of the main peak and enhancements. The shapes of the main peak and enhancements can be subtly influenced by ϵ, as well as their relative intensities and to some extent ν_{peak}. If some fine structure is seen on the steep, positive, optically thick slope, the viewing angle θ can be determined very reliably for each iterative choice of ND and ϵ. The best-fit area A is uniquely determined for each combination of ND, ϵ, and θ. Having iteratively achieved an acceptable fit above $\sim 3\nu_b$, one can adjust N holding ND constant, to improve the fit at lower frequencies. This in turn may require a readjustment of θ, since the enhancements and positive-slope fine structure can be subtly affected by changes in N. As discussed above, however, achieving a "better fit" at frequencies below $\sim 3\nu_b$ is not necessarily the same as determining N to higher accuracy, since the role of external absorption and other emission mechanisms is unclear at these low frequencies.

Given our identification of the enhancement near 2 GHz as H-3, we believe that the parameters deduced for Figure 13 are each determined to within 10% accuracy, and some of them (B, δ, T, θ) to better than 5%. Because of the ambiguity in N and D mentioned above, this accuracy applies only to the product ND.

Simultaneous observations with other instruments can greatly facilitate the fitting process. For instance, the high spatial resolution of the Very Large Array makes it useful for constraining the source size and geometry. X-ray observations can help constrain the temperature and/or spectral index. The extension of this thermal/nonthermal approach to the X-ray regime is currently underway.

It is interesting that the deduced value for ND implies, for coronal densities, a very small source depth—on the order of 100 km. Combining this with the deduced area, one is led to the possibility that the observed gyrosynchrotron emission originates in a thin "disk" somewhere in the leg of a magnetic loop, i.e., footpoint emission, rather than throughout the loop volume as generally supposed. More detailed modeling of complete events is required to confirm or disprove this suggestion. If we do in fact observe such a compact emission region, this could help explain why the TNT model for a *homogeneous* source can work as well as it does.

5.2. *The Peak Frequency*

It is known that the peak frequency of a microwave burst changes slowly if at all during the burst's evolution, in stark contrast to the predictions of the simpler models (SGH1). In addition, the enhancements are seen to rise and fall together with the main peak (SGH2).

An examination of Figures 4–10 indicates that, holding all other parameters constant, a varying ϵ is a likely candidate to explain these gross evolutionary changes, perhaps in concert with small changes in δ. Both the "primary" and "secondary" components vary in unison with ϵ, and ν_{peak} stays relatively fixed. The case for ϵ can be made more plausible by recognizing that ν_{peak} is determined by the optical depth, i.e., absorption coefficient. This in turn depends on the derivative of f with respect to γ, and all ϵ-dependence in f can be factored out of the derivative and indeed out of the summation in equation (4). This is not true for any of T, δ, B, or θ, while N and D directly alter the optical depth by changing the number of particles along the integrated line of sight. In the next section we will relate ϵ to the electric field strength in the acceleration region, which is expected to be a naturally evolving physical quantity in a flare.

The constancy of ν_{peak} and the coevolution of the components could also be due to a changing source area, which as mentioned earlier, is a premultiplier. The area change would have to be independent of other physical quantities, however, so as not to alter the optical depth of the source. Microwave imaging at multiple frequencies will help resolve this point.

We note that the nonthermal approximation for ν_{peak} (Dulk & Marsh 1982), within the range of δ discussed above, must be amended for this model by a factor of

$$\left(\frac{E_{\text{cr}}}{10\ \text{keV}}\right)^{-0.32+0.35\delta-0.03\delta^2},$$

since E_{cr} can be much higher than was assumed by those authors. This correction is valid for E_{cr} up to at least 150 keV.

6. A PHYSICAL INTERPRETATION FOR f

The model we have presented is complete but largely descriptive. Virtually all flare energization mechanisms one might think of will produce both heating and acceleration. The simplest physical mechanism we can think of to accomplish this involves large-scale DC electric fields (current sheets). This provides a plausible and coherent physical framework in which to place our model TNT distribution function, thereby gaining physical insight into the flaring process.

Large-scale DC electric fields are likely to be generated by the magnetic instabilities which are widely associated with flare onset (cf. Tandberg-Hanssen & Emslie 1988 and references therein). They are also consistent with the observed association of flares with nonpotential (current-carrying) magnetic fields (e.g., Hagyard, Moore, & Emslie 1984). Electric fields have received increasing attention in recent years (e.g., Holman 1985; Moghaddam-Taaheri & Goertz 1990). The presence of current sheets simultaneously accelerates charged particles via the runaway process, and heats the ambient plasma via Joule dissipation (Holman 1985). The currents are expected to be aligned with the magnetic field, or in magnetic neutral sheets, so little gyrosynchrotron radiation will come from the TNT current in the sheets themselves. The presence of some pitch-angle scattering in the current sheets (e.g., by plasma turbulence) will introduce nonthermal particles into the heated ambient plasma in sufficient numbers to produce the observed microwave burst. It is also possible to account for the observed hard X-ray emission from flares within this framework. This will be treated in a future paper.

Placed in this context, our nonthermal parameter ϵ becomes the ratio E/E_D, where E is the DC electric field strength and E_D, the Dreicer field, is that field strength for which all of the electrons in the sheets would experience runaway acceleration; that is, none of the thermal particles would be retarded by collisions. In this interpretation, p_{cr} is the critical momentum for electron runaway. The reader is referred to Holman (1985, with v replaced by $p/m\gamma$) for further details. This p_{cr} provides the low-energy cutoff for f_{nt} and motivates our normalization equation (15). Electrons having $p < p_{\text{cr}}$ serve primarily to heat the ambient plasma, and are not part of the nonthermal tail.

If the gyrosynchrotron emission were produced in the current sheets themselves, equation (15) would not hold and our TNT distribution would have to be constructed differently. The bulk population would then be drifting (thermal current) and would become enhanced at lower momenta than p_{cr} through upscattering into the runaway regime in the presence of the electric field. These processes are relevant to the hard X-ray production.

To complete the physical interpretation that we propose, an understanding of the production of runaway electrons in a finite plasma is required. This work is currently underway. We expect that the electron spectral index δ will be related to the length of the acceleration region, as well as to the strength of the electric field in that region. The origins of the postulated DC fields are currently an area of active research (Syrovatskii 1981; van Ballegooijen 1985; Mikić, Schnack, & Van Hoven 1989; Wolfson 1989; Karpen, Antiochos, & DeVore 1990).

7. SUMMARY

We have presented a new framework within which to model solar microwave bursts, which provides new physical insight into high-resolution microwave spectra. The approach taken combines elements of the previously disparate thermal and nonthermal models, allowing both of these populations to coexist during a flare. Our perspective is thus shifted, from asking whether flare emissions are produced by thermal or by nonthermal particles to asking how much of a contribution each population makes. The assumption of a thermal/nonthermal electron distribution is physically reasonable and also has important observational consequences.

For the large class of observed flares showing enhancements on the low-frequency side of the spectral peak, we find that the thermal/nonthermal approach can successfully address three observational challenges:

1. The enhancements themselves are explained as nonthermal radiation produced between the lowest of the thermal harmonics. Within these wave bands the otherwise dominant thermal processes are incapable of suppressing the nonthermal emission.

2. The appearance of steep slopes on the optically thick side of the spectrum is accounted for by the thermal absorption of nonthermally produced radiation.

3. The relative constancy of the peak frequency throughout a burst, and the coevolution of the low-frequency enhancements, is made plausible if the major contributor to the evolution of the electron distribution is a changing electric field strength.

Physically, we place this model within a plausible and internally consistent framework, by assuming the presence of current sheets in the flaring plasma. The current sheets provide both primary heating of the plasma and nonthermal acceleration of particles.

We have demonstrated the applicability of this model by fitting a typical high-resolution microwave spectrum, observed at the Owens Valley Radio Observatory. The sophistication of such present-generation observations shows the limitations of simpler approaches. The approach presented here can be used to model such observations directly, and can provide us with a more unified perspective in which to study solar flares.

We gratefully acknowledge T. J. Kelly and Mark Pesses for developing the code for equations (3) and (4). We thank Gordon Hurford for providing the error bars in Figure 13. This work was done while S. G. B. was a NASA Graduate Studies Fellow, in residence at the Laboratory for Astronomy and Solar Physics at Goddard Space Flight Center, supported by NASA grant NGT-50337. G. D. H. acknowledges support from NASA RTOP 170-38-53-16.

REFERENCES

Alissandrakis, C. E., & Kundu, M. R. 1984, A&A, 139, 271
Bandiera, R. 1982, A&A, 112, 52
Bogod, V. M., Korzhavin, A. N., Akhmedov, SH. B, Auras, H., Hildebrandt, J., & Krüger, A. 1990, Sol. Phys., 129, 351
Brosius, J. W., Willson, R. F., Holman, G. D., & Schmelz, J. T. 1992, ApJ, 386, 347
Connor, J. W., & Hastie, R. J. 1975, Nucl. Fusion, 15, 415
Dulk, G. A., & Marsh, K. A. 1982, ApJ, 259, 350
Emslie, A. G., & Vlahos, L. 1980, ApJ, 242, 359
Gary, D. E., & Hurford, G. J. 1990, ApJ, 361, 290
Hagyard, M. J., Moore, R. L., & Emslie, A. G. 1984, Adv. Space Res., 4, 71
Holman, G. D. 1985, ApJ, 293, 584
Holman, G. D., Kundu, M. R., & Kane, S. R. 1989, ApJ, 345, 1050
Karpen, J. T., Antiochos, S. K., & DeVore, C. R. 1990, ApJ, 356, L67
Klein, K.-L. 1987, A&A, 183, 341
Kundu, M. R. 1965, Solar Radio Astronomy (New York: Wiley)
Kundu, M. R., & Alissandrakis, C. E. 1984, Sol. Phys., 94, 249
Kundu, M. R., Alissandrakis, C. E., Bregman, J. D., & Hin, A. C. 1977, ApJ, 213, 278
Kundu, M. R., Schmahl, E. J., Velusamy, T., & Vlahos, L. 1982, A&A, 108, 188
Mikić, Z., Schnack, D. D., & Van Hoven, G. 1989, ApJ, 338, 1148
Moghaddam-Taaheri, E., & Goertz, C. K. 1990, ApJ, 352, 361
Mok, Y. 1983, ApJ, 275, 901
Nitta, N., Kiplinger, A. L., & Kai, K. 1989, ApJ, 337, 1003
Ramaty, R. 1969, ApJ, 158, 753
Stäli, M., Gary, D. E., & Hurford, G. J. 1989, Sol. Phys., 120, 351 (SGH1)
———. 1990, Sol. Phys., 125, 343 (SGH2)
Syrovatskii, S. I. 1981, ARA&A, 19, 163
Takakura, T. 1960, PASJ, 12, 352
Takakura, T., & Scalise, E. 1970, Sol. Phys., 11, 434
Tamor, S. 1978, Nucl. Fusion, 18, 229
———. 1979, Nucl. Fusion, 19, 455
Tanaka, H., & Kakinuma, T. 1959, in Paris Symposium on Radio Astronomy, ed. R. N. Bracewell (Stanford: Stanford Univ. Press), 215
Tandberg-Hanssen, E., & Emslie, A. G. 1988, The Physics of Solar Flares (Cambridge: Cambridge Univ. Press)
Trulsen, J., & Fejer, J. A. 1970, J. Plasma Phys., 4, 825
van Ballegooijen, A. A. 1985, ApJ, 298, 421
Velusamy, T., & Kundu, M. R. 1982, ApJ, 258, 388
Willson, R. F., Klein, K.-L., Kerdraon, A., Lang, K. R., & Trottet, G. 1990, ApJ, 357, 662
Wolfson, R. 1989, ApJ, 344, 471
Zheleznyakov, V. V. 1970, in Radio Emission of the Sun and Planets (Oxford: Pergamon)