

THERMAL INSTABILITY

GEORGE B. FIELD

Princeton University Observatory

Received June 4, 1964; revised March 30, 1965

ABSTRACT

The stability of a dilute gas in mechanical and thermal equilibrium is studied with a view to applications to non-gravitational condensation phenomena in astronomy. It is shown that, under a wide range of conditions, thermal equilibrium is unstable and can result in the formation of condensations of higher density and lower temperature than are found in the surrounding medium. The instability criterion is shown to differ considerably from certain criteria proposed by previous authors. The modifications due to finite speed of sound, to thermal conduction, to a magnetic field, to rotation, to an external gravitational field, and to expansion of the medium are studied. Applications are made to the solar chromosphere and corona, to the interstellar medium in the galactic disk and halo, to planetary nebulae, and to intergalactic matter. It is shown that the principle of thermal instability is closely related to the formation of solar prominences, to condensations in planetary nebulae, and to condensation of galaxies from the intergalactic medium.

I. INTRODUCTION

a) Condensations Not Due to Gravitation

Most astronomical objects owe their existence to self-gravitation. However, there is a class of objects, including solar prominences, interstellar clouds, and condensations in planetary nebulae, whose existence cannot be explained in this way, as their calculated gravitational energies are far smaller than those due to internal pressure. In each case it appears reasonable to assume that internal pressure is being balanced by pressure in an external diffuse medium. It then appears likely that the objects in question are formed from the diffuse medium by some kind of condensation process not involving gravitation. The present paper aims at understanding such a condensation process. It is believed that the calculations presented will find a variety of astronomical applications, including those already cited. Perhaps the most exciting possibility is that the origin of galaxies may be explained along these lines.

In the remainder of this section we review the literature and present a physical description of the condensation process. Section II is a detailed analysis of condensation phenomena in an infinite uniform medium, while §§ III–VI explore the modifications introduced by a magnetic field, by rotation of the system, by an external gravitational field, and by expansion of the medium as a whole, respectively. Section VII considers several astronomical applications in detail.

b) Review of the Literature

Parker (1953) suggested that the condensation phenomenon in question (particularly that of solar prominences) can be understood as a consequence of instability in the thermal equilibrium of a diffuse medium. He reasoned that, if the thermal equilibrium of the medium is a balance between temperature-independent energy gains and temperature-dependent radiative losses, instability results if, near equilibrium, the losses increase with decreasing temperature, for then a cooler-than-average region cools more effectively than its surroundings, and its temperature rapidly drops below the initial equilibrium value.

Other authors, notably Zanstra (1955*a, b*), pointed out that the strong tendency toward pressure equilibrium in the medium would naturally result in compression of the cool regions and expansion of the hot ones; in this way one could understand the formation of cool condensations in a medium of high temperature. In applying this reasoning,

Zanstra used an instability criterion different from Parker's. Because of the incompleteness of both Parker's and Zanstra's reasoning, however, both arrived at incorrect criteria.

Weymann (1960*b*) appears to be the first to have given the correct criterion, in a paper devoted to chromospheric heating. He did not discuss his results in detail, and their importance appears to have been overlooked. Our first task, then, is to solve correctly and in detail the stability problem for a uniform medium, as considered by these authors. At the end of the paper we shall consider the implications of these results, and those found for a variety of stabilizing forces, for a group of astronomical applications already partially treated in the literature.

c) Physical Discussion of Instability

The correct instability criterion can be derived physically as follows. In an infinite, uniform, static medium of density ρ_0 and temperature T_0 , the continuity and force equations are automatically satisfied. We ignore self-gravitation for reasons given above, but it can be introduced easily at a later stage if required. The heat equation can be written,

$$\mathfrak{L}(\rho_0, T_0) = 0, \quad (1)$$

where $\mathfrak{L}$ is a generalized heat-loss function, defined as energy losses minus energy gains, per gram of material per second exclusive of thermal conduction (see Table 1). Note that $\mathfrak{L}$ can be written as a function of the local values of ρ and T if the gas is sufficiently diffuse to be optically thin, because under such circumstances the coupling of the gas to the radiation field is not complicated by radiative transfer effects. Equation (1) determines T_0 for any assigned value of ρ_0 . We introduce a perturbation of density and temperature such that some thermodynamic variable A (e.g., pressure) is held constant. The entropy of the material will change by an amount δS , and the heat-loss function, by an amount $\delta \mathfrak{L}$, both calculated at constant A . Since

$$\delta \mathfrak{L} dt = -T d(\delta S), \quad (2)$$

there is instability only if

$$\left(\frac{\partial \mathfrak{L}}{\partial S} \right)_A > 0. \quad (3)$$

In an isochoric perturbation, $T dS = C_v dT$, while in an isobaric one, $T dS = C_p dT$, so the corresponding instability criteria are

$$\left(\frac{\partial \mathfrak{L}}{\partial T} \right)_\rho < 0, \quad (\text{isochoric}), \quad (4a)$$

and

$$\left(\frac{\partial \mathfrak{L}}{\partial T} \right)_p = \left(\frac{\partial \mathfrak{L}}{\partial T} \right)_\rho - \frac{\rho_0}{T_0} \left(\frac{\partial \mathfrak{L}}{\partial \rho} \right)_T < 0, \quad (\text{isobaric}), \quad (4b)$$

the perfect gas law being assumed applicable. The isochoric criterion is that given by Parker on the basis of an analysis of the heat equation alone. It is quite apparent, however, that pressure variations accompanying isochoric temperature variations will in general drive motions which destroy the constancy of density, so that this criterion is incompatible with the force equation.

The isobaric criterion is consistent with the force equation, because the lack of pressure variations insures that no forces will destroy the postulated density field. We shall find from the detailed analysis that condensations are indeed governed by the isobaric criterion (4b). Since the ρ -derivative is often positive in astronomical applications, owing to the binary-collision nature of radiative losses, it is seen that it usually increases the tendency toward instability.

The condensation mode discussed above is quite different from a sound wave. It depends for its existence on the possibility of non-adiabatic motions, while waves persist as solutions to the equations even when all non-adiabatic phenomena vanish. But even sound waves, which are nearly isentropic in nature, grow in amplitude if an isentropic instability criterion is satisfied:

$$\left(\frac{\partial \mathfrak{L}}{\partial T}\right)_s = \left(\frac{\partial \mathfrak{L}}{\partial T}\right)_\rho + \frac{1}{\gamma - 1} \frac{\rho_0}{T_0} \left(\frac{\partial \mathfrak{L}}{\partial \rho}\right)_T < 0, \quad (\text{isentropic}). \quad (5)$$

If inequality (5) is satisfied, the near-isentropic oscillations in a sound wave are subject to a slight heating during the compression phase which tends to increase the amplitude of the wave. The amplification of sound waves is implicit in Weymann's treatment, but he did not point it out. We shall treat it on an equal footing with the growth of condensations, even though the stabilization introduced by the ρ -derivative makes it unlikely to arise in practice. The growth of sound waves is similar to the instability of stellar pulsations, energy flowing preferentially into the compressed regions from the external energy sources.

TABLE 1
SYMBOLS USED FREQUENTLY

Symbol	Equation Where First Used	Meaning
ρ, T, v, p	(1), (1), (6), (7)	Density, temperature, velocity, and pressure 0 indicates unperturbed value, 1 indicates perturbation, and * indicates base of atmosphere
γ, R, μ	(5), (9), (9)	Ratio of specific heats, gas constant, and molecular weight
$\mathfrak{L}, \mathfrak{L}_\rho, \mathfrak{L}_T$	(1), (13), (13)	Generalized heat-loss function and its ρ - and T -derivatives
c, a	(15), (47)	Speed of sound and of Alfvén waves
$K, K_{ }, K_{\perp}$	(8), (50), (50)	Thermal conductivity in absence of magnetic field, parallel to field, and perpendicular to field
k_ρ, k_T, k_K	(16), (16), (16)	Characteristic wavenumbers proportional to $\mathfrak{L}$, $\mathfrak{L}_T$, and $1/K$
k_c, k_m	(26), (45)	Critical wavenumber for instability, and wavenumber at which instability is maximum
k_T'	(55)	Value of k_T as modified by thermal conduction
$k_c(\theta)$	(60)	Critical wavenumber for disturbances making angle θ with magnetic field
k_R	(80)	Critical wavenumber in rotating system
$\sigma_\rho, \sigma_T, \sigma_K, \sigma_T'$	(17), (17), (17), (17)	Wavenumbers listed above, divided by k
n, k	(10), (10)	Growth rate and wave vector of perturbation
α, β	(27), (27)	Dimensionless ratios of characteristic wavenumbers
σ	(106)	Parameter that decreases with increasing thermal conductivity
N, M	(93), (146)	Exponents of temperature and density in cooling law

II. INSTABILITY IN A UNIFORM MEDIUM, INCLUDING EFFECTS OF THERMAL CONDUCTION

a) The Characteristic Equation

The basic equations are

$$\frac{d\rho}{dt} + \rho \nabla \cdot \mathbf{v} = 0, \quad (6)$$

$$\rho \frac{d\mathbf{v}}{dt} + \nabla p = 0, \quad (7)$$

$$\frac{1}{\gamma - 1} \frac{dp}{dt} - \frac{\gamma}{\gamma - 1} \frac{p}{\rho} \frac{d\rho}{dt} + \rho \mathfrak{L} - \nabla \cdot (K \nabla T) = 0, \quad (8)$$

and

$$p - \frac{R}{\mu} \rho T = 0, \quad (9)$$

where d/dt is $\partial/\partial t + \mathbf{v} \cdot \nabla$ and K is the coefficient of thermal conductivity. The first two terms of equation (8), the heat equation, are $\rho T(dS/dt)$, calculated for a perfect gas. Ionization energy is omitted, although it might be important for certain applications. The analysis is carried out for arbitrary ratio of specific heats, γ ; we shall use $\gamma = \frac{5}{3}$ in numerical examples.

The equilibrium state is characterized by $\rho = \rho_0$, $T = T_0$, $\mathbf{v} = 0$, and $\mathfrak{L}(\rho_0, T_0) = 0$. Assuming perturbations of the form

$$a(\mathbf{r}, t) = a_1 \exp(nt + i\mathbf{k} \cdot \mathbf{r}), \quad (10)$$

we find the linearized equations for the perturbations to be

$$n\rho_1 + \rho_0 i\mathbf{k} \cdot \mathbf{v}_1 = 0, \quad (11)$$

$$n\rho_0 \mathbf{v}_1 + i\mathbf{k} p_1 = 0, \quad (12)$$

$$\frac{n}{\gamma - 1} p_1 - \frac{n\gamma p_0}{(\gamma - 1)\rho_0} \rho_1 + \rho_0 \mathfrak{L}_\rho \rho_1 + \rho_0 \mathfrak{L}_T T_1 + K_0 k^2 T_1 = 0, \quad (13)$$

and

$$\frac{p_1}{p_0} - \frac{\rho_1}{\rho_0} - \frac{T_1}{T_0} = 0, \quad (14)$$

where K_0 and $\mathfrak{L}_\rho \equiv (\partial \mathfrak{L} / \partial \rho)_T$ and $\mathfrak{L}_T \equiv (\partial \mathfrak{L} / \partial T)_\rho$ are evaluated for the equilibrium state. Since $\mathbf{k} \times \mathbf{v}_1 = 0$ from equation (12), we are left with the four variables ρ_1 , p_1 , T_1 , and $\mathbf{k} \cdot \mathbf{v}_1$, which are specified by the four equations (11)–(14). The vanishing of the determinant of coefficients implies a characteristic equation for the growth rate, n , which can be written,

$$n^3 + n^2 c \left(k_T + \frac{k^2}{k_K} \right) + n c^2 k^2 + \frac{c^3 k^2}{\gamma} \left(k_T - k_\rho + \frac{k^2}{k_K} \right) = 0, \quad (15)$$

where c is the Laplacian speed of sound, $(\gamma p_0 / \rho_0)^{1/2}$, and where we have introduced the wavenumbers

$$k_\rho = \frac{\mu(\gamma - 1)\rho_0 \mathfrak{L}_\rho}{RcT_0}, \quad k_T = \frac{\mu(\gamma - 1)\mathfrak{L}_T}{Rc}, \quad k_K = \frac{Rc\rho_0}{\mu(\gamma - 1)K}. \quad (16)$$

Here k_ρ and k_T are wavenumbers of sound waves whose angular frequencies are numerically equal to the growth rates of isothermal and isochoric perturbations, respectively;

their signs depend upon those of the derivatives of $\mathfrak{L}$; k_K is the reciprocal of the mean free path of the conducting particles. The remainder of this section will be devoted to the study of the roots of equation (15) as functions of k , k_ρ , k_T , and k_K . Since these roots are often complex, the phrase "growth rate," will be taken to mean n , whether real or complex.

Equation (15) is a generalization of the equation for non-adiabatic propagation of sound given by Rayleigh (1945), who based his discussion on an 1851 paper by Stokes. Stokes considered forms of cooling ($\mathfrak{L}$) which could lead only to damping, which Rayleigh points out is consistent with the second law. The instabilities we shall find occur when the equilibrium state is far from thermodynamic equilibrium.

We shall see below that the three roots of equation (15) have a simple interpretation when the non-adiabatic effects are small, as has been understood for some time. Two roots, of course, correspond to sound waves, which are unstable if criterion (5) is satisfied. The other root corresponds to a mode in which density and temperature variations occur, but with opposite phase, so that pressure variations are negligible. The instability of such a thermal or condensation mode is governed by criterion (4b).

b) Domains of Instability

By introducing the non-dimensional quantities,

$$y = \frac{n}{kc}, \quad \sigma_\rho = \frac{k_\rho}{k}, \quad \sigma_T = \frac{k_T}{k}, \quad \sigma_K = \frac{k}{k_K}, \quad \sigma_{T'} = \sigma_T + \sigma_K, \quad (17)$$

we can write the characteristic equation in the form,

$$y^3 + y^2\sigma_{T'} + y + \gamma^{-1}(\sigma_{T'} - \sigma_\rho) = 0, \quad (18)$$

so that the parameters defining the character of the roots are reduced from four to two. We therefore seek the character of the roots throughout the $\sigma_{T'}$, σ_ρ -plane.

The reality of the roots is governed by the sign of

$$D = \frac{1}{4}(r - \frac{1}{3}pq + \frac{2}{27}p^3)^2 + \frac{1}{27}(q - \frac{1}{3}p^2)^3, \quad (19)$$

where p , q , and r are the last three coefficients of equation (18). The curve $D = 0$, along which there are three real roots (two of which are equal), is given by

$$\sigma_\rho = (1 - \frac{1}{3}\gamma)\sigma_{T'} \pm \frac{2}{27}\gamma\{(\sigma_{T'})^3 \pm [(\sigma_{T'})^2 - 3]^{3/2}\}, \quad (20)$$

and is plotted in Figure 1 for $\gamma = \frac{5}{3}$. $D < 0$ in the interiors of the thorn-like curves, and $D > 0$ elsewhere. Since $D > 0$ implies one real root and a complex conjugate pair, sound waves are present where D is positive; in particular, this is always true when k_ρ and k_T are small compared to k , so that cooling effects are small.

The regions $D < 0$ are characterized by three real unequal roots; the sound waves are overdamped for the large values of σ involved. The shortest waves for which this can happen occur at the points of the thorns: $\sigma_{T'} = \pm\sqrt{3}$, $\sigma_\rho = \pm(1 - \frac{1}{3}\gamma)\sqrt{3}$. As $\sigma_{T'} \rightarrow \pm\infty$, the branches of the curve $D = 0$ become asymptotic to $\sigma_\rho = \sigma_{T'}$ and $\sigma_\rho = \frac{4}{27}\gamma(\sigma_{T'})^3$.

We now investigate the signs of the roots, and hence the regions of instability. Denote the three real and unequal roots for $D < 0$ by y_1 , y_2 , and y_3 . Since

$$y_1 + y_2 + y_3 = -\sigma_{T'}, \quad y_1y_2y_3 = -\gamma^{-1}(\sigma_{T'} - \sigma_\rho), \quad (21)$$

the domain $D < 0$ is conveniently divided into four regions, depending on the signs of $\sigma_{T'}$ and of $\sigma_{T'} - \sigma_\rho$. When both are positive, all three roots are negative, and vice versa; if $\sigma_{T'}$ is positive and $\sigma_{T'} - \sigma_\rho$ is negative there are two negative roots and one positive

root, and vice versa. These results, derived from equation (21) together with the requirement of continuity with results derived below for $D > 0$, are indicated in Figure 1.

If $D > 0$, we are interested in the sign of the real root, (call it y_1), and in the sign of the real part, η , of the complex conjugate roots, $y_2 = \eta + i\psi$ and $y_3 = \eta - i\psi$. Since the product of the roots is equal to $y_1|y_2|^2$, y_1 has the same sign as $\sigma_\rho - \sigma_T'$. To obtain the sign of η , we derive the coefficients of the cubic equation it satisfies. The coefficient corresponding to r is

$$R = \frac{1}{8}[\sigma_T'(1 - \gamma^{-1}) + \sigma_\rho\gamma^{-1}], \quad (22)$$

so that by reasoning like that above, η has the sign of $-\sigma_T'(1 - \gamma^{-1}) - \sigma_\rho\gamma^{-1}$.

Let us call a real positive root or a complex root with positive real part a "positive root." The former corresponds to instability and the latter to overstability. Then we

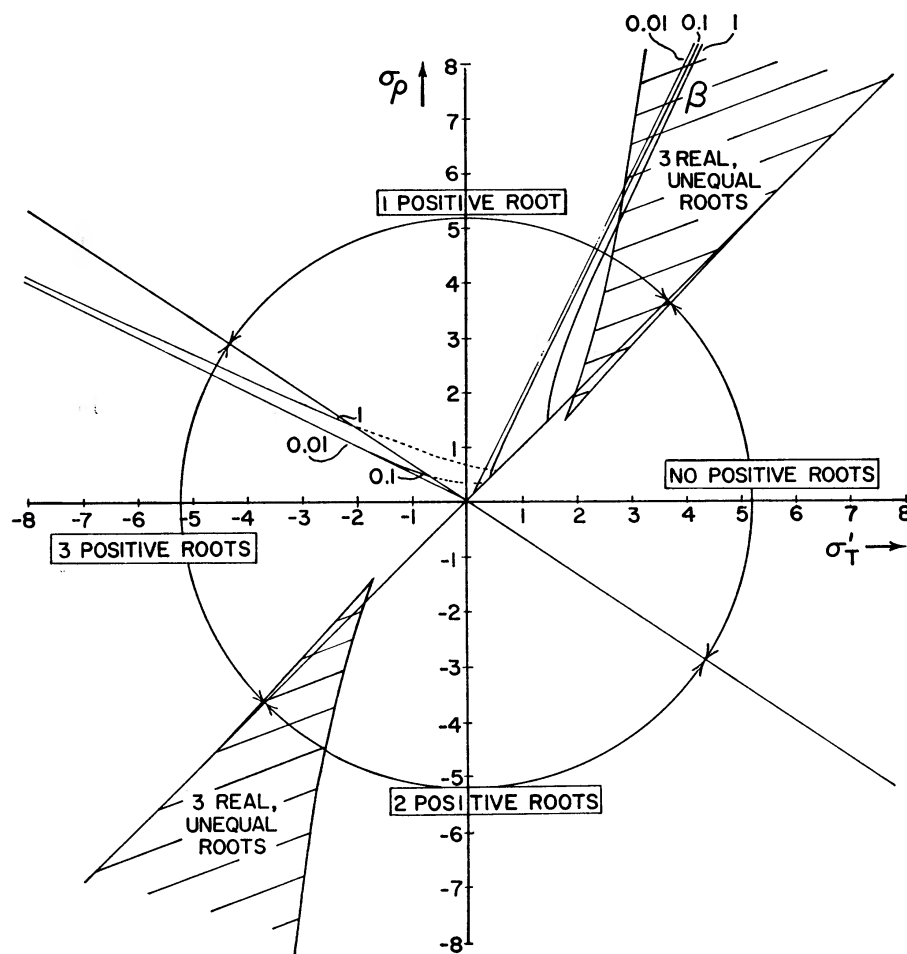


FIG. 1.—The domains of instability for a uniform gas, indicated by the presence of positive roots. The abscissa is the normalized form of the temperature derivative of $\mathfrak{L}$, as modified by thermal conduction in eq (17); the ordinate is related to the density derivative of $\mathfrak{L}$. The gas is neutral for isobaric disturbances along the straight line of slope $+1$ and for isentropic disturbances along the line of slope $-\frac{2}{3}$ according to eq. (33). Near the origin the cooling times exceed λ/c and the acoustic mode is oscillatory. In the shaded regions the acoustic mode is non-oscillatory, but its stability is governed by the same criteria as elsewhere. The non-radial lines represent the effects of changing the wavelength in a given medium. Conductive stabilization occurs where these lines encounter the neutral lines; this happens at longer wavelengths for larger values of β , the normalized thermal conductivity according to eqs. (26) and (27).

note that the number of positive roots is determined by the signs of $\sigma_T' - \sigma_\rho$ and of $\sigma_T' + \sigma_\rho/(\gamma - 1)$. When both are positive all roots are negative, and vice versa. When $\sigma_T' - \sigma_\rho$ is positive and $\sigma_T' + \sigma_\rho/(\gamma - 1)$ is negative, there are one positive root and two negative roots, and vice versa. This is shown in Figure 1. Near the origin, where k_ρ and k_T are small, and in the case of vanishing K , we therefore see that there is one real positive root whenever

$$T_0 \mathfrak{L}_T - \rho_0 \mathfrak{L}_\rho < 0. \quad (23)$$

We recognize this as the isobaric instability criterion (4b). The root therefore corresponds to the formation of condensations under constant pressure conditions, as described in § I, and we shall refer to it as the condensation mode. There are two complex conjugate roots with positive real part whenever

$$T_0 \mathfrak{L}_T + \frac{1}{\gamma - 1} \rho_0 \mathfrak{L}_\rho < 0. \quad (24)$$

We recognize this as the isentropic instability criterion (5). These roots correspond to growing sound waves, and we shall refer to them collectively as the wave mode.

Moving radially away from the origin in certain directions changes the reality of these roots as we cross the curve $D = 0$, but the number of positive roots in the sense defined above remains the same. Hence varying the wavelength (still in the absence of thermal conduction) does not affect the instability criteria, but only determines whether the wave mode is underdamped or overdamped.

When conduction is not negligible, the instability criteria are

$$\sigma_T - \sigma_\rho < -\sigma_K \quad \text{and} \quad \sigma_T + \sigma_\rho/(\gamma - 1) < -\sigma_K \quad (25)$$

for the condensation and wave modes, respectively. Since σ_K is positive-definite, the effect of conduction is always to reduce instability. The dimensional forms of inequalities (25) show that whenever the instability criteria in the absence of conduction are satisfied, there is a critical wavenumber such that larger wavenumbers are stabilized through thermal conduction. These critical wavenumbers are

$$k_c = [k_K(k_\rho - k_T)]^{1/2} \quad \text{and} \quad k_c = \{k_K[-k_T - k_\rho/(\gamma - 1)]\}^{1/2} \quad (26)$$

for the condensation and wave modes, respectively. The critical wavenumber for the condensation mode was derived by Parker (1953) without the k_ρ term.

It is instructive to follow the stability of a given medium (k_ρ , k_T , k_K fixed) as k is varied. Define

$$a = k_T/k_\rho = \sigma_T/\sigma_\rho = T_0 \mathfrak{L}_T / \rho_0 \mathfrak{L}_\rho \quad (27)$$

and

$$\beta = k_\rho/k_K = \sigma_\rho \sigma_K = \frac{\mu^2 (\gamma - 1)^2 K \mathfrak{L}_\rho}{R^2 c^2 T_0}.$$

In terms of these quantities, which are functions of the medium alone,

$$\sigma_T' = a\sigma_\rho + \beta/\sigma_\rho = ak_\rho/k + \beta k/k_\rho. \quad (28)$$

Curves of equation (28) for $a = \frac{1}{2}$ and -2 , and $\beta = 0.01, 0.1$, and 1 , are plotted in Figure 1. The first term of equation (28) dominates for large σ_ρ , and a given medium is represented by a radial line; the instability is unaffected by k . As k increases, the second term pushes the curves to the right; the larger β (hence conductivity), the smaller the value of k at which this happens. The result is to encounter one of the radial lines representing marginal instability when $k = k_c$, defined by equation (26).

c) *Growth Rates*

In the present section we calculate growth rates in various parts of the σ_T', σ_ρ -plane, by solution of the characteristic equation (18). First we give the results of a series expansion valid near the origin in the case $K = 0$.

$$y_1 = (\sigma_\rho - \sigma_T)/\gamma - (\sigma_\rho - \sigma_T)^2[\sigma_\rho + (\gamma - 1)\sigma_T]/\gamma^3 + O(\sigma^5), \quad (29)$$

and

$$y_{2,3} = \pm i - \frac{\sigma_\rho + (\gamma - 1)\sigma_T}{2\gamma} \pm i \frac{[\sigma_\rho + (\gamma - 1)\sigma_T][3\sigma_\rho - (3 + \gamma)\sigma_T]}{8\gamma^2} \\ + (\sigma_T - \sigma_\rho)^2 \frac{\sigma_\rho + (\gamma - 1)\sigma_T}{2\gamma^3} + O(\sigma^4). \quad (30)$$

These expressions are valid as $k \rightarrow \infty$ when $K = 0$; if $K \neq 0$, they apply with the substitution of σ_T' for σ_T , but only if the additional condition, $k \ll k_e$, is fulfilled. In dimensional form, the growth rates are

$$n_1 = -(\gamma - 1)\mu(T_0\mathfrak{L}_T - \rho_0\mathfrak{L}_\rho)/\gamma RT_0, \quad (31)$$

and

$$n_{2,3} = \pm i k c - (\gamma - 1)^2 \mu [T_0\mathfrak{L}_T + \rho_0\mathfrak{L}_\rho/(\gamma - 1)]/2\gamma RT_0. \quad (32)$$

The growth rate of the condensation mode is independent of k as $k \rightarrow \infty$, and is equal to $-(\delta\mathfrak{L})_p$ divided by $T(\delta S)_p$, in accordance with the discussion of § I. Similarly, the growth rate of the wave mode is proportional to $-(\delta\mathfrak{L})_S$.

An expansion valid as $\sigma \rightarrow \infty$, if σ_ρ is of the same order as σ_T , gives the three roots as

$$y_1 = -\sigma_T + [\sigma_\rho + (\gamma - 1)\sigma_T]/\gamma\sigma_T^2 \\ + [\sigma_\rho + (\gamma - 1)\sigma_T][2\sigma_\rho + (\gamma - 2)\sigma_T]/\gamma^2\sigma_T^5 + O(\sigma^{-5}), \quad (33)$$

and

$$y_{2,3} = \pm [(\sigma_\rho - \sigma_T)/\gamma\sigma_T]^{1/2} - [\sigma_\rho + (\gamma - 1)\sigma_T]/2\gamma\sigma_T^2 \\ \pm 5[(\sigma_\rho - \sigma_T)/\gamma\sigma_T]^{-1/2}[\sigma_\rho + (\gamma - 1)\sigma_T][\sigma_\rho + (\frac{1}{2}\gamma - 1)\sigma_T]/\gamma^3\sigma_T^4 \\ + O(\sigma^{-3}). \quad (34)$$

If $|\sigma_T| \ll |\sigma_\rho|^{1/3}$, all three roots are contained in the expression

$$y_{1,2,3} = (\sigma_\rho/\gamma)^{1/3} + O(\sigma_\rho^{-1/3}). \quad (35)$$

Equations (33)–(35) indicate a variety of ways in which n or its real part depend on k as $k \rightarrow 0$ for the unstable modes. In each case the dependence is on k to some power (call it m), and m depends on the signs of σ_T , $\sigma_T - \sigma_\rho$, and $\sigma_T(1 - \gamma^{-1}) + \gamma^{-1}\sigma_\rho$ as shown in Table 2.

The specific example given before ($\alpha = \frac{1}{2}$) lies in region b , and the growth rate as $\sigma \rightarrow \infty$ is

$$n = [(\sigma_\rho - \sigma_T)/\gamma\sigma_T]^{1/2} k c = \gamma^{-1/2} k c, \quad (36)$$

which approaches 0 with k . In this region $\mathfrak{L}_T > 0$, and instability requires that $\mathfrak{L}_\rho$ be large and positive, and that the medium has the tendency to increase the density in the growing condensations. It therefore seems natural that the growth rate is limited by the finite speed of sound in the long-wavelength limit, as indicated by equation (36). Correspondingly, we note that when $\mathfrak{L}_T < 0$, as in regions d , e , and f , the growth rate is independent of k in the long-wavelength limit, as the temperature perturbations can then grow even when the density perturbations are small. For example, when $\alpha = -2$, we are in region e , and

$$n = -k_T c = -\mu(\gamma - 1)\mathfrak{L}_T/R. \quad (37)$$

Therefore, in the long-wavelength limit, Parker's isochoric criterion applies in the sense that, unless $\mathfrak{L}_T < 0$, the condensation mode grows at a rate which vanishes with k . However, this limit, which corresponds to considerable cooling in a small fraction of a sound-wave period, is rarely met in practice.

d) Scale of Maximum Instability

Except for the condensation mode when $\sigma_T < 0$, the growth rate approaches zero as k decreases. On the other hand, the growth rate falls to zero at a sufficiently large value of k , k_c . This implies that there exists a wavenumber at which the positive growth rate attains a maximum, denoted by k_m , which is of interest because it defines the scale of disturbances that will grow the fastest and hence attain observable amplitude. The

TABLE 2
DEPENDENCE OF GROWTH RATE OF UNSTABLE MODES ON k , AS $k \rightarrow 0$

Region	σ_T	$\sigma_T - \sigma_\rho$	$\frac{\sigma_T(1-\gamma^{-1})}{+\gamma^{-1}\sigma_\rho}$	Behavior of Positive Roots*
<i>a</i>	+	+	+	None
<i>b</i>	+	—	+	One real (1)
<i>c</i>	0	—	+	One real (2/3)
<i>d</i>	—	—	+	One real (0)
<i>e</i>	—	—	—	One real (0), two complex (2)
<i>f</i>	—	+	—	Two real (0, 1)
<i>g</i>	0	+	—	Two complex (2/3)
<i>h</i>	+	+	—	Two complex (2)

* Value of m in parentheses

value of k_m is obtained by viewing the characteristic equation (15) as a quadratic in k^2 . By introducing the new non-dimensional quantities

$$u = (k/k_\rho)^2, \quad x = n/k_\rho c, \quad (38)$$

this quadratic can be written

$$u^2\beta/\gamma + u[\beta x^2 + x + (a-1)/\gamma] + x^3 + ax^2 = 0, \quad (39)$$

which is useful if n is real. If n is complex, the cubic for η can be written

$$u^2[\frac{1}{4}\beta^2\xi + \frac{1}{8}(1-\gamma^{-1})\beta] + u[\beta\xi^2 + \frac{1}{4}\xi + \frac{1}{2}a\beta\xi + \frac{1}{8}(1-\gamma^{-1})a + \frac{1}{8}\gamma^{-1}] \\ + \xi^3 + a\xi^2 + \frac{1}{4}a^2\xi = 0, \quad (40)$$

where $\xi = \text{Re}(x)$. This procedure can be modified if $k_\rho = 0$ by substituting k_T into equation (38), since non-triviality requires that not both k_T and k_ρ vanish. The roots of equations (39) and (40) for the values of a and β considered previously are plotted in Figures 2 and 3; the existence of a maximum growth rate is confirmed, and the decrease at small k and the cutoff at large k are evident. Note the linear and quadratic dependence for the cases $a = \frac{1}{2}$ and -2 , respectively, as stated in Table 2. Data pertaining to k_m and to the values of x or ξ at $k = k_m$ (which are denoted x_m or ξ_m) are plotted in Figures 4 and 5, along with asymptotic results derived below for $\beta \rightarrow 0$.

Since β is small in all the applications discussed in § VII, it is useful to have approximate formulae for k_m and n (k_m) in that limit. Here n (k_m) is given by the condition, $B^2 = 4AC$, where A , B , and C are the coefficients (functions of x or ξ) in equations (39) and (40). The corresponding values of k_m are given by the relation $u_m = -B/2A$, where A and B are evaluated at x_m (or ξ_m). This equation provides one relation between u_m and x_m (or ξ_m), while equations (39) and (40) provide another. Elimination of x_m (or

ξ_m) gives an equation satisfied by u_m , whose coefficients depend only on the parameters α , β , and γ . The resulting equation is quartic for n real and cubic for n complex, and can be expressed in the form

$$\sum_{j,k} a_{jk}(\alpha, \gamma) \beta^j u_m^k = 0. \quad (41)$$

The non-vanishing coefficients are easily derived but are unnecessary to give here. In principle, equation (41) determines u_m as a function of the properties of the medium,

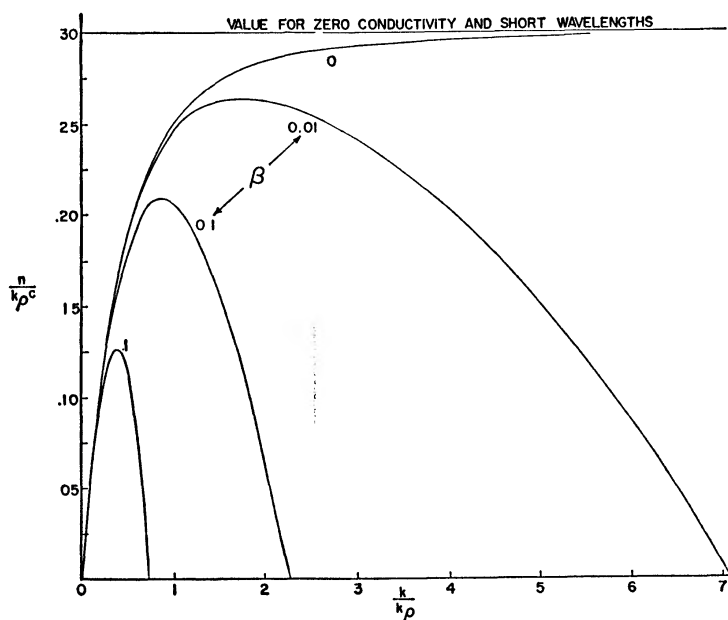


FIG. 2.—Normalized growth rate of the condensation mode as a function of normalized wavenumber for $\alpha = \frac{1}{2}$ and for four values of β . Note that large thermal conductivity and long wavelength enhance stability. Compare with non-radial lines in first quadrant of Fig. 1.

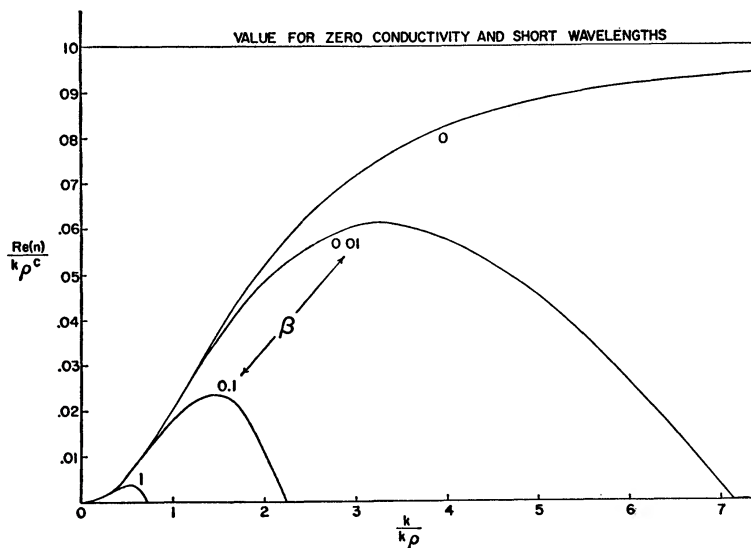


FIG. 3.—Growth rate of the wave mode as a function of wavenumber for $\alpha = -2$. Compare with non-radial lines in second quadrant of Fig. 1.

α , β , and γ . This solution is readily determined in the limit of small β by assuming a solution of the form

$$u_m = C\beta^l, \quad (42)$$

in which case the term (j, k) in equation (41) is proportional to β^{j+lk} . The power $q(j, k, l) = j + lk$ is then plotted against l for each (j, k) ; the result is a diagram which shows at once the ordering of all of the terms in equation (41) as a function of l . A solu-

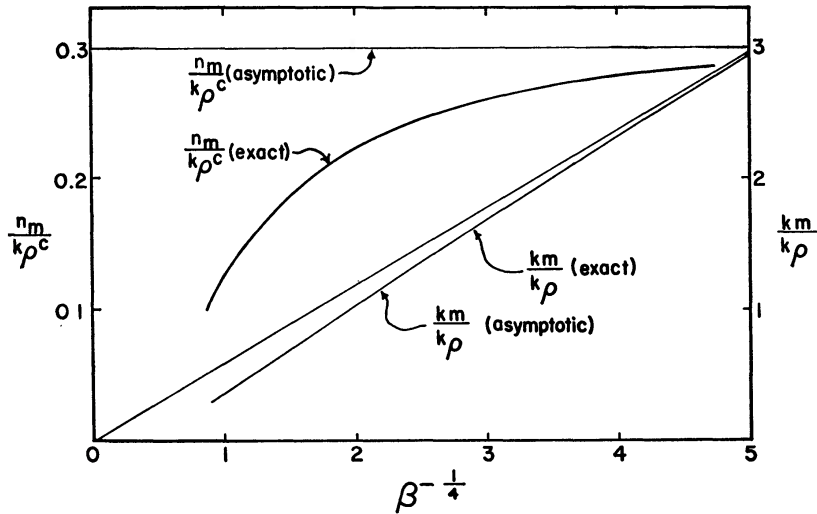


FIG. 4.—Normalized growth rates and normalized wavenumbers at the maxima of the curves in Fig. 2. Asymptotic values are from eqs. (29) and (45). Thermal conductivity decreases to the right.

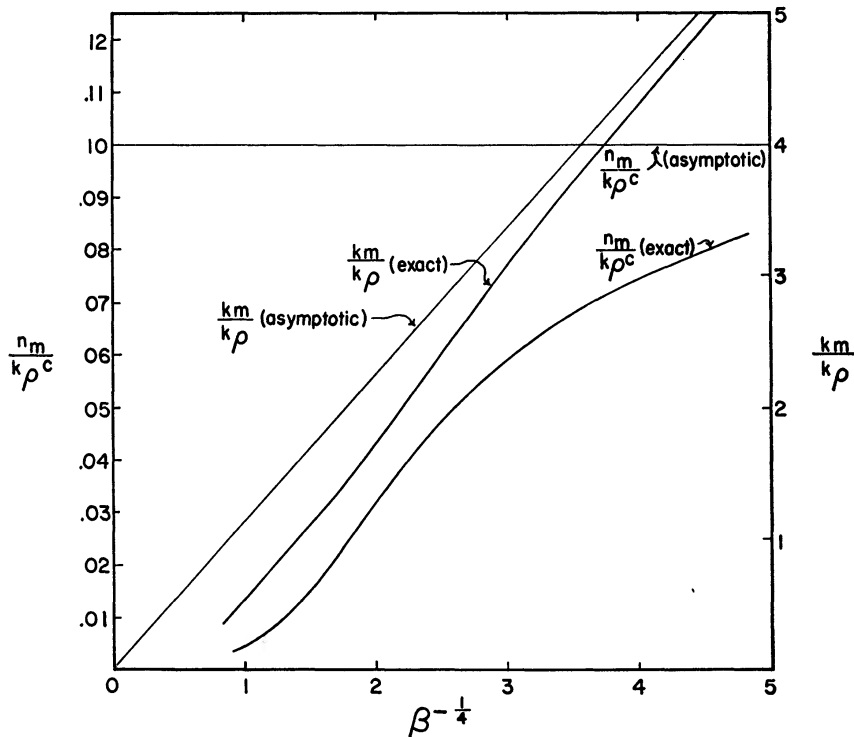


FIG. 5.—Growth rates and wavenumbers at the maxima of the curves in Fig. 3. Asymptotic values are from eqs. (30) and (45).

tion can only occur for a value of l at which $q(j, k, l) = q(j', k', l)$ —for only then are the corresponding two terms of the same order; such a value of l is represented by an intersection on the plot. In order to assure that the intersection represents a balance between dominant terms, one requires that q at the intersection is the lowest at that value of l . In this way, one finds that

$$l = -\frac{1}{2}, \quad (43)$$

with a balance between (0, 0) and (1, 2) terms for both condensation and wave modes, so that the solution is of the form (42) and

$$u_m = \{[a(\gamma - 1) + 1]/\beta\gamma^2\}^{1/2}|a - 1|, \\ \{[a(1 - \gamma) - 1]/[\beta(\gamma - 1)]\}^{1/2}|a - 1|/\gamma \quad (44)$$

for the condensation and wave modes, respectively. It is readily verified that the radicals are real in the unstable domains of a . The dominant terms of equation (41) are of zero order, while the next larger term, (1, 1), is of order $\frac{1}{2}$. Hence the relative accuracy of equation (44) is of order $\beta^{1/2}$. Equation (44) can be written

$$k_m = D(a, \beta)\beta^{-1/4}k_\rho, \quad (45)$$

where $D(\frac{1}{2}, \frac{5}{3}) = 0.588$ and $D(-2, \frac{5}{3}) = 1.128$. These results are compared with exact calculations in Figures 4 and 5. It appears that for $\beta < 0.01$ ($\beta^{-1/4} > 3.16$), the approximate formula is accurate to better than 10 per cent. Also shown are the asymptotic growth rates, calculated for short wavelengths in the limit $K = 0$. The accuracy is of the same order here also for the condensation mode, but only of order 30 per cent for the wave mode. Using equation (26) k_m can be written

$$k_m = [(1 - a)^2/\gamma^2 + a(1 - a)/\gamma]^{1/4}(k_\rho k_c)^{1/2}, \quad |(a - 1)/\gamma|^{1/2}(k_\rho k_c)^{1/2} \quad (46)$$

for the condensation and wave modes, respectively, showing that k_m is essentially the geometric mean of the wavenumber at which the growth rate drops because of finite speed-of-sound effects, and the wavenumber at which it vanishes because of conduction effects. The coefficients in equation (46) are close enough to unity (0.7 for $a = \frac{1}{2}$, 1.3 for $a = -2$) that one may use the approximation $k_m = (k_\rho k_c)^{1/2}$ in practice.

III. EFFECT OF A MAGNETIC FIELD

a) Physical Considerations

A uniform magnetic field penetrating the diffuse medium will have at least three effects if the electrical conductivity is high. First, two additional wave modes, corresponding to the two polarizations of Alfvén waves for propagation along the field, will be present. Second, the condensation mode will be unaffected when $\mathbf{k}$ is parallel to $\mathbf{B}_0$, but when $\mathbf{k}$ is normal to $\mathbf{B}_0$, the instability will be inhibited by the necessity that gas pressure in the rarified regions becomes large enough to overcome the increased magnetic pressure in the condensed regions. Third, thermal conduction across the magnetic field will be greatly reduced by the spiraling of electrons between collisions; this will permit fine irregularities to develop in a direction approximately normal to the field.

The additional wave modes differ in their relation to the non-adiabatic effects considered in this paper. As these modes propagate at successively larger angles to the field, the one which is polarized normal to $\mathbf{k}$ and to $\mathbf{B}_0$ remains purely transverse. Since the pressure forces associated with the instability are irrotational, they cannot feed energy into this wave, and the wave should be neutrally stable. The other polarization, together with the wave which is purely acoustic for propagation along the field, experiences non-vanishing compression for propagation at arbitrary angles to the field, and so both should be subject to amplification. One expects the instability criterion to remain of the form of equation (5).

One can easily derive the new instability criterion for condensations across the magnetic field. The pressure in the isobaric criterion must be replaced by P , the gas pressure plus magnetic pressure. Since the field is frozen into the gas, the coefficient of the $\mathfrak{L}_\rho$ -term in inequality (4b) is replaced by

$$\left(\frac{\partial \rho}{\partial T}\right)_P = -\frac{\rho_0}{T_0(1 + B_0^2/4\pi p_0)} = -\frac{\rho_0}{T_0(1 + \gamma a^2/c^2)}, \quad (47)$$

where a equals $B_0/(4\pi\rho_0)^{1/2}$, the Alfvén velocity, so the new instability criterion is

$$T_0\mathfrak{L}_T - (1 + \gamma a^2/c^2)^{-1}\rho_0\mathfrak{L}_\rho < 0 \quad (48)$$

for $\mathbf{k} \perp \mathbf{B}_0$. The effect is to increase the slope of the line of marginal instability, $\sigma_\rho = \sigma_T$, in Figure 1, from unity to $1 + a^2/c^2$, approaching the isochoric line, $\sigma_T = 0$, for $a \gg c$. This has different effects in different areas of the σ_T, σ_ρ -plane. If σ_ρ and σ_T are positive with $\sigma_\rho > \sigma_T$, the situation is ordinarily unstable because the increase of cooling rate caused by the higher density in a condensation is greater than the decrease because of lower temperature. For a sufficiently large magnetic-field strength,

$$B_0 = c[4\pi\rho_0(k_\rho - k_T)/\gamma k_T]^{1/2}, \quad (49)$$

the required increase in density is prevented by magnetic pressure and the situation is stabilized. When σ_ρ is positive and σ_T is negative, the condensation mode is unstable both isochorically and isobarically, and the magnetic field does not affect stability. When σ_ρ and σ_T are both negative with $\sigma_\rho < \sigma_T$, the situation is ordinarily stable, because the increase of cooling rate due to the lower temperatures in a condensation is smaller than the decrease due to lower density. In this case, a magnetic-field strength exceeding B_0 destabilizes the situation by permitting the temperature decrease without the accompanying density increase. Inasmuch as k_ρ is positive in most applications, however, this latter situation does not usually arise, and the usual effect of a sufficiently strong magnetic field is to stabilize the condensation mode.

The quantitative effect of the suppression of thermal conduction normal to the field depends on the state of the matter. We shall consider particularly a fully ionized gas, for which the reduction in conductivity is

$$K_\perp/K_\parallel = 9.4 \times 10^{-3}(\omega_{ci}t_{ci})^{-2} = 4.8 \times 10^{-10}n_e^2/B_0^2T_0^3, \quad (50)$$

according to Spitzer (1962), where ω_{ci} is the ion gyrofrequency, t_{ci} is the ion "self-collision time," and the second formula applies for pure hydrogen of density $n_e \text{ cm}^{-3}$. This can range from 10^{-1} in the chromosphere to 10^{-20} in the galactic halo, so that the effects can be either minor or extraordinarily large, depending on circumstances.

b) Modification of the Characteristic Equation

The basic equations are modified as follows. The magnetic force, $\nabla(B^2/8\pi) - \mathbf{B} \cdot \nabla \mathbf{B}/4\pi$, is added to the left-hand side of equation (7), and the conduction term in the heat equation is replaced by $-\nabla \cdot [K_\parallel \nabla_\parallel T + K_\perp \nabla_\perp T]$, where $\parallel$ and $\perp$ refer to angles $\theta = 0$ and $\pi/2$ between the direction concerned and the direction of $\mathbf{B}_0$. To specify the development of $\mathbf{B}$ we assume the form of the induction equation valid for infinite electrical conductivity,

$$d\mathbf{B}/dt + \mathbf{B} \nabla \cdot \mathbf{v} - (\mathbf{B} \cdot \nabla) \mathbf{v} = 0. \quad (51)$$

The perturbation equations are modified as follows. The magnetic force terms, $i(\mathbf{B}_0 \cdot \mathbf{B}_1)\mathbf{k}/4\pi - ik_\parallel B_0 \mathbf{B}_1/4\pi$, are added to equation (12). The term $K_0 k^2 T_1$ is replaced by $(K_\parallel k_\parallel^2 + K_\perp k_\perp^2)T_1$. The perturbed induction equation is

$$n\mathbf{B}_1 + i\mathbf{B}_0(\mathbf{k} \cdot \mathbf{v}_1) - ik_\parallel B_0 \mathbf{v}_1 = 0. \quad (52)$$

We introduce the coordinate system $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$ specified by

$$\mathbf{u}_z = \mathbf{B}_0/B_0, \quad \mathbf{u}_y = \mathbf{B}_0 \times \mathbf{k} / |\mathbf{B}_0 \times \mathbf{k}|, \quad \mathbf{u}_x = \mathbf{u}_y \times \mathbf{u}_z. \quad (53)$$

One then finds that v_y is uncoupled from the other velocities, and that its frequency is that of a pure transverse Alfvén wave, ka , unaffected by the heat equation. Thus the pure Alfvén wave is neutral; we shall not consider it further.

The motions in the other modes are constrained to the x - z -plane, and are governed by the characteristic equation,

$$\begin{aligned} n^5 + n^4 ck_T' + n^3 k^2 (c^2 + a^2) + n^2 k^2 [c^3 (k_T' - k_\rho) / \gamma + a^2 ck_T'] \\ + nc^2 a^2 k^4 \cos^2 \theta + (k_T' - k_\rho) c^3 a^2 k^4 \cos^2 \theta / \gamma = 0, \end{aligned} \quad (54)$$

where θ is the angle between $\mathbf{k}$ and $\mathbf{B}_0$, and where

$$k_T' = k_T + k_{||}^2 / k_{K||} + k_{\perp}^2 / k_{K\perp} = k_T + k^2 (\cos^2 \theta / k_{K||} + \sin^2 \theta / k_{K\perp}) \quad (55)$$

is defined as the value of k_T modified by conduction effects. The terms $k_{K||}$ and $k_{K\perp}$ are defined as in equation (16), with $K_{||}$ and $K_{\perp}$, respectively.

c) Growth Rates in Limiting Cases

We shall not attempt a complete analysis of the unstable domains associated with equation (54), but shall limit ourselves to the region where k_T' and k_ρ are small compared to k (corresponding to the region around the origin in Fig. 1), as this is the usual condition in practice.

If k_T' and k_ρ are identically zero, the five roots of equation (54) are given by zero for the condensation mode and

$$n^2 = \frac{1}{2} k^2 \{ -c^2 - a^2 \pm [(c^2 + a^2)^2 - 4a^2 c^2 \cos^2 \theta]^{1/2} \} \quad (56)$$

for the wave modes. The wave modes are recognized as the fast and slow hydromagnetic waves.

When k_T' and k_ρ are small, we assume an expansion of the form $n = n_0 + n_1 + \dots$, where the subscript refers to the order in k_T' and k_ρ , so that n_0 is that given previously. It is found that the form of the expansion differs depending on whether or not $\theta = \pi/2$. If $\theta \neq \pi/2$, the growth rates are

$$n_1 = \frac{c(k_\rho - k_T')}{\gamma}, \quad - \frac{c^3(n_0^2 + a^2 k^2 \cos^2 \theta)[k_\rho + (\gamma - 1)k_T']}{2\gamma(c^2 + a^2)[n_0^2 + 2c^2 a^2 k^2 \cos^2 \theta / (c^2 + a^2)]} \quad (57)$$

for the condensation and wave modes, respectively, where n_0 is given above. Equation (57) is identical, except for notation and for the modified definition of k_T' , with the leading term of equation (29); the growth rate of the condensation mode for $\theta \neq \pi/2$ is therefore independent of magnetic-field strength in the limit of small cooling rate. Similarly, the growth rate of the wave modes differs only by a multiplicative non-negative factor from the corresponding term in equation (30).

How are we to explain these somewhat surprising results? It is found from the basic equations that in the condensation mode,

$$v_{\perp}/v_{||} = n^2 \tan \theta / (n^2 + a^2 k^2), \quad (58)$$

which from equation (57) behaves like

$$c^2(k_\rho - k_T')^2 / a^2 k^2 \quad (59)$$

as k_ρ and k_T' become much smaller than k . Hence the motion in the condensation mode is predominantly along the magnetic field whenever $k_\rho - k_T' \ll ak/c$. This means that a magnetic-field strength which approaches equipartition ($a \simeq c$) is sufficient to

dominate the motions in the slowly growing condensation mode. The resulting motions do not change the magnetic field, so that there are no magnetic pressures to be overcome, and the instability is essentially governed only by the requirement of constant pressure along the field lines. The result is that the instability criterion is still the isobaric one, inequality (4b).

The wavenumber required for conductive stabilization of the condensation mode of course depends on θ , and satisfies

$$k_c(\theta) = k_c(\cos^2 \theta + \sin^2 \theta K_{\perp}/K_{\parallel})^{-1/2}, \quad (60)$$

where k_c is the critical wavenumber defined in equation (26). It follows that for instability the conditions $k_{\parallel} < k_c$ and $k_{\perp} < k_c(K_{\parallel}/K_{\perp})^{1/2}$ must be satisfied, showing that indeed much larger wavenumbers are permitted normal to the direction of the field. It seems likely that k_m also increases with θ .

Similarly, the growth of the wave modes is governed by the same isentropic instability criterion as before, corresponding to the fact that growth results from energy input during the compression phase of the wave. Conductive stabilization occurs at a wavenumber given by equation (60), where now k_c is calculated from equation (26).

The case $\theta = \pi/2$ is distinguished by the fact that all motions must displace field lines. It is governed by the equation

$$n^5 + n^4 ck_T' + n^3 k^2 (c^2 + a^2) + n^2 [c^3 (k_T' - k_{\rho})/\gamma + a^2 ck_T'] k^2 = 0. \quad (61)$$

Two roots of this equation vanish, corresponding to the zero velocity of the slow wave for propagation across the field. The fast wave (now the magnetosonic mode) has the frequency, $\omega_0 = k(a^2 + c^2)^{1/2}$, and the growth rate given by

$$n_1 = -c^3 [k_{\rho} + (\gamma - 1)k_T'] / 2\gamma(c^2 + a^2), \quad (62)$$

which are both given by the previous analysis for the particular case $\theta = \pi/2$.

The growth of the condensation mode, however, is governed by the condition that the last term of equation (61) be negative, which in the absence of conduction is equivalent to the modified isobaric criterion (46), showing that magnetic pressure tends to stabilize condensations across the magnetic-field lines. The critical wavenumber for conductive stabilization is given by

$$k_{c\perp} = k_{K\perp}^{1/2} [k_{\rho}/(1 + \gamma a^2/c^2) - k_T]^{1/2}. \quad (63)$$

The growth rate for small k_{ρ} and k_T' is

$$n_1 = \frac{c(c^2 + \gamma a^2)[k_{\rho}/(1 + \gamma a^2/c^2) - k_T]}{\gamma(c^2 + a^2)}. \quad (64)$$

IV. EFFECT OF ROTATION

a) Physical Considerations

Consider the growth of condensations in a rotating, axially symmetric system, in which the major force balance is between gravity and centrifugal force. Instability will be unaffected by rotation as long as the condensation involves motions only along the azimuthal direction, since the gas can stream slowly without change of angular momentum to new azimuthal positions.

Radial motions will be greatly inhibited, however, if we make the reasonable assumption that the gravitational field decreases more slowly than ϖ^{-3} , for under such conditions a particle moving outward with constant angular momentum experiences a greater attractive gravitational field than a repulsive centrifugal force and tends to be restored to its initial radius. In order for radial motions to occur in spite of this effect, pressure gradients must be built up. This is more easily accomplished in a small-scale perturba-

tion, and so we expect that small scales will be governed by the same criterion as before, while large scales will be stabilized.

To consider the main features of this problem, we examine the stability of a cylindrical rotating system of infinite extent in the axial direction, in which the initial distribution of density and temperature is uniform. The gravitational field is axially symmetric but otherwise arbitrary. Admittedly, such a system differs from any astronomical example, but the effects derived are useful in visualizing what may happen in actual cases. Because of what has been said, attention is directed to radial perturbations alone.

b) Modification of the Characteristic Equation

The basic equations are modified only by the addition of $-\rho g$ to equation (7). If we express these equations in cylindrical coordinates ϖ , θ , and z , and take all quantities to vary with ϖ alone, we find that equations for the unperturbed state are satisfied by $\rho = \rho_0$ and $T = T_0$ (constants) if

$$\mathfrak{L}(\rho_0, T_0) = 0 \quad (65)$$

and

$$v_\theta = V(\varpi) = [\varpi g(\varpi)]^{1/2}; \quad (66)$$

$g(\varpi)$ is the inward gravitational field. The equations for the perturbations are

$$n\rho_1 + \rho_0 \frac{dv_\varpi}{d\varpi} + \rho_0 \frac{v_\varpi}{\varpi} = 0, \quad (67)$$

$$n v_\varpi - \frac{2V(\varpi)}{\varpi} v_\theta + \rho_0^{-1} \frac{dp_1}{d\varpi} = 0, \quad (68)$$

$$n v_\theta + v_\varpi \frac{dV(\varpi)}{d\varpi} + v_\theta \frac{V(\varpi)}{\varpi} = 0, \quad (69)$$

$$\frac{np_1}{\gamma - 1} - \frac{n\gamma p_0 \rho_1}{\rho_0(\gamma - 1)} + \rho_0 \mathfrak{L}_\rho \rho_1 + T_0 \mathfrak{L}_T T_1 - \frac{K_0}{\varpi} \frac{d}{d\varpi} \left(\varpi \frac{dT_1}{d\varpi} \right) = 0, \quad (70)$$

and equation (14). The subscript "1" on the perturbation quantities v_ϖ and v_θ is omitted.

Equations (68) and (69) combine to give

$$(n^2 + f^2) v_\varpi + \frac{n}{\rho_0} \frac{dp_1}{d\varpi} = 0, \quad (71)$$

where

$$f^2 = 2V(\varpi)\varpi^{-1}[dV(\varpi)/d\varpi + V(\varpi)/\varpi] \quad (72)$$

is the square of the classical epicyclic frequency. Equation (71) shows that radial motions are governed by $n = \pm if$ in a pressure-free system, in accord with the results of classical theory. This oscillation results from the gravitational restoring force as explained above.

The velocity v_ϖ is eliminated in favor of the density by use of the continuity equation (67), and we find that

$$\rho_1 - \frac{1}{\varpi} \frac{d}{d\varpi} \left(\frac{\varpi dp_1/d\varpi}{n^2 + f^2} \right) = 0. \quad (73)$$

Another relation between ρ_1 and p_1 is furnished by eliminating T_1 from the heat equation (70) by the use of equation (14). With the abbreviation $D = d/d\varpi$, this can be written,

$$[n + ck_T - c(k_K\varpi)^{-1}D(\varpi D)]p_1 - c^2[n + c(k_T - k_\rho)/\gamma - c(\gamma k_K\varpi)^{-1}D(\varpi D)]\rho_1 = 0, \quad (74)$$

where the various k 's are defined in § II, and are independent of ϖ . The characteristic equation which results from the combination of equations (73) and (74) is

$$\{n + ck_T - c(k_K\varpi)^{-1}D(\varpi D) - c^2[n + c(k_T - k_\rho)/\gamma - c(\gamma k_K\varpi)^{-1}D(\varpi D)]\varpi^{-1}D \\ [\varpi/(n^2 + f^2)]D\}p_1 = 0. \quad (75)$$

c) Solid-Body Rotation

To go further we must specify the rotation law, $V(\varpi)$, and hence $f(\varpi)$. The simple example provided by solid-body rotation,

$$V = \omega_0\varpi, \quad f = 2\omega_0, \quad (76)$$

will be considered at this point. The characteristic equation is

$$\{(n^2 + 4\omega_0^2)(n + ck_T) - [(n^2 + 4\omega_0^2)/ck_K + n + c(k_K - k_\rho)/\gamma]c^2L + c^3(\gamma k_K)^{-1}L^2\}p_1 = 0, \quad (77)$$

where L is the operator, $\varpi^{-1}D(\varpi D)$.

The solutions of equation (77) which are finite at the origin are of the form, $p_1 = J_0(k\varpi)$. Since these functions (with k arbitrary) form a complete set, we need consider only one of them. The growth rate for the wavenumber k is a solution of the cubic,

$$n^3 + n^2c(k_T + k^2/k_K) + n(c^2k^2 + 4\omega_0^2) + c[c^2k^2(k_T - k_\rho + k^2/k_K)/\gamma \\ + 4\omega_0^2(k_T + k^2/k_K)] = 0. \quad (78)$$

This equation is very similar to equation (15) for a static medium, and may be analyzed in a similar manner. One finds that the instability criterion for condensations in the absence of conduction is

$$T_0\mathcal{Q}_T - (1 + 4\gamma\omega_0^2/k^2c^2)^{-1}\rho_0\mathcal{Q}_\rho < 0, \quad (79)$$

whose appearance is similar to the criterion for condensation across a magnetic field, inequality (48). As in that case, the effect of large rotation is to increase the slope of the line of marginal instability in Figure 1 (see above), and consequently to stabilize situations in the upper part of the first quadrant and to destabilize situations in the lower part of the third quadrant.

A major difference from the magnetic case is the dependence on k , which arises because the gravitational restoring forces increase with scale, while pressure driving forces decrease with scale. For sufficiently large k , the rotation has no effect. Stabilization (for the usual case, k_ρ positive) occurs for wavenumbers less than a critical value,

$$k_R = \frac{2\omega_0}{c} \left(\frac{\gamma k_T}{k_\rho - k_T} \right)^{1/2}. \quad (80)$$

When conduction is present, the vanishing of the last term of equation (78) (which determines marginal stability) requires solution of a quadratic in k^2 . If, however, k_c as defined by equation (26) is much larger than k_R , as is often the case, equation (80) for k_R remains the same and k_c is given by the usual expression. The instability of the wave mode is governed by the isentropic criterion, independent of the rotation, as is the case with a magnetic field present.

Growth rates may be obtained by an expansion valid for small cooling rates. We find that the terms of first order in k_ρ and k_T are

$$n_1 = - \frac{c[c^2k^2 + 4\gamma\omega_0^2]k_T' - k^2c^2k_\rho}{\gamma(c^2k^2 + 4\omega_0^2)}, \quad - \frac{c^3k^2[(\gamma - 1)k_T' + k_\rho]}{2\gamma(c^2k^2 + 4\omega_0^2)} \quad (81)$$

for the condensation and wave modes, respectively. The wave mode has the frequency, $(4\omega_0^2 + c^2k^2)^{1/2}$, indicating the modification of sound waves by rotation.

d) Other Rotation Laws

The characteristic equation (75) is fourth order; for solid-body rotation it reduces to a quadratic algebraic equation in k^2 that shows how conduction modifies the instability for large values of k , and how rotation modifies the instability for small values of k . For more general rotation laws it is helpful to ignore conduction. The resulting equation for p_1 is only of second order, and can be written

$$\{D^2 + [\varpi^{-1} - D \ln(n^2 + f^2)]D + \chi(n^2 + f^2)/c^2\} p_1 = 0, \quad (82)$$

where

$$\chi = \frac{ck_T + n}{[c(k_\rho - k_T)/\gamma] - n} \quad (83)$$

is independent of ϖ . This equation contains the information we are seeking concerning the behavior of long-wave disturbances that are most affected by rotation. Examination shows that, even when f is variable, equation (82) admits WKB solutions of the form

$$p_1 = [k(\varpi)/\varpi]^{1/2} \exp \pm i \int k(\varpi) d\varpi, \quad (84)$$

provided that $k\varpi \gg 1$ and that $V \gg c$; $k(\varpi)$ is of the form,

$$k(\varpi) = \{\chi[n^2 + f^2(\varpi)]\}^{1/2}/c. \quad (85)$$

In the limit of small cooling rate, the wave mode is characterized by $\chi = -1$, and hence

$$k(\varpi) = [\omega^2 - f^2(\varpi)]^{1/2}/c, \quad (86)$$

where $\omega = in$. This shows that the modified sound wave can propagate only over those values of ϖ such that $\omega^2 > f^2(\varpi)$.

The marginally unstable condensation mode, with $n = 0$, is characterized by

$$\chi = \chi_0 = \gamma k_T/(k_\rho - k_T), \quad k_R(\varpi) = f(\varpi)c^{-1}[\gamma k_T/(k_\rho - k_T)]^{1/2}. \quad (87)$$

In the usual case ($k_\rho > 0$), k must exceed k_R for instability; for $n > 0$, the WKB wavenumber calculated from equations (83) and (85) exceeds k_R everywhere. In terms of the ϖ -independent quantity, $h = f^2(\varpi)/c^2 k^2(\varpi)$,

$$n = c\gamma^{-1}(k_\rho - k_T)(1 - h\chi_0)/(1 - h), \quad (88)$$

showing that as $h \rightarrow 0$ ($k \gg k_R$), n rapidly approaches the value given by equation (29), valid in the absence of rotation.

There are exact solutions for the marginally unstable condensation mode whenever $V(\varpi)$ and $f(\varpi)$ vary as a power of ϖ . Thus, if

$$V(\varpi) = W[2(s+1)]^{-1/2}\varpi^s, \quad f(\varpi) = W\varpi^{s-1}, \quad (89)$$

where $s > -1$, the solutions are of the form

$$p_1 = \varpi^{s-1} Z_{1-s-1}(W\chi_0^{1/2}s^{-1}c^{-1}\varpi^s), \quad (s \neq 0) \quad (90)$$

and

$$p_1 = \exp \{ \ln \varpi [-1 \pm \sqrt{(1 - \chi_0 W^2/c^2)}] \}, \quad (s = 0). \quad (91)$$

If $s = 1$, we recover the results of the previous section for solid-body rotation; $s = -\frac{1}{2}$ and $W = (GM)^{1/2}$ corresponds to Keplerian motion. Equation (90) then shows that the marginally unstable mode is an oscillatory function of ϖ if and only if $\chi_0 > 0$; since this condition occurs only in the upper half of the first quadrant and in the lower half of the third quadrant of the σ_T, σ_ρ -plane, we conclude that stabilization and destabilization occur at finite wavenumbers, independent of the ratio of $V(\varpi)$ to c . This conclusion agrees with the WKB results, which are, however, valid only for $V(\varpi) \gg c$. The case

$s = 0$, $W = V\sqrt{2}$ corresponds to V independent of ϖ ; it is of interest because Mestel (1963) has suggested that it provides a theoretically satisfactory representation of galactic rotation laws. In this case, equation (91) indicates that there are oscillations in the marginally unstable mode if and only if

$$(1 + 2 \gamma V^2/c^2)k_T > k_\rho > k_T. \quad (92)$$

In the limit $V \gg c$, this region is that described previously where stabilization (or destabilization) occurs at a finite wavenumber. As $V/c \rightarrow 0$, however, the region shrinks to zero, indicating that under these conditions, arbitrarily long wavelength disturbances are unstable (or stable if $k_\rho < 0$). This conclusion, which could not be derived by the WKB analysis (valid only for $V/c \rightarrow \infty$), is in accord with the fact that as $V/c \rightarrow 0$, centrifugal forces are dominated by pressure forces.

V. EFFECT OF DENSITY STRATIFICATION

a) *Physical Considerations*

In § II we considered the stability of a static uniform medium. In practice, a gravitational field is required to keep the gas motionless in spite of its internal pressure, and this field stratifies the gas into layers of varying density. If disturbances are much shorter than a scale height, one would expect the stability to be governed by the considerations of § II. Larger scales, however, will necessarily encompass regions of differing densities (and temperatures); to compensate these differences and their effects on growth rate by conduction effects, the scale must vary with position, and the normal modes will therefore differ from sine waves. Another feature is that gravitational forces acting on density perturbations are available to compensate pressure perturbations, and one finds that the pressure varies with height in such a way that it remains constant for each particle.

It is no longer a trivial matter to satisfy hydrostatic equilibrium and thermal equilibrium in the unperturbed state. Some attention must therefore be given to the development of reasonable atmospheres whose stability may then be studied. An interesting class results from the adoption of a heat-loss function of the form,

$$\mathcal{Q} = \rho(c_1 T^N - c_2), \quad (93)$$

where N , c_1 , and c_2 are arbitrary positive constants, which would be appropriate for cooling by inelastic collisions and heating either by superelastic collisions (which draw their energy from photo-ionization or from cosmic-ray ionization), or by dissipation of some type of waves at a rate proportional to ρ^2 per unit volume. In either case, the heating is assumed to be independent of the temperature of the medium, although this restriction is not essential.

b) *Equilibrium State: Isothermal Atmosphere*

The basic equations are those of § I, with the term $-\rho g$ added to equation (7). They can be written in the form,

$$\frac{R}{\mu} \frac{d}{dz} (\rho_0 T_0) + \rho_0 g = 0 \quad (94)$$

and

$$\rho_0 \mathcal{Q}(\rho_0, T_0) - \frac{d}{dz} \left[K(\rho_0, T_0) \frac{dT_0}{dz} \right] = 0, \quad (95)$$

if we assume that all variables depend only on z . We assume that $K = \eta T^{5/2}$, appropriate for fully ionized hydrogen in the absence of a magnetic field, where $\eta = 1.9 \times 10^{-5}$ /in Λ cgs according to Spitzer (1962, p. 144).

Equations (94) and (95) constitute a third-order system for a plane-parallel atmosphere in hydrostatic and thermal equilibrium, for which the boundary conditions can be taken as

$$\frac{dT_0}{dz}(0) = 0, \quad T_0(\infty) < \infty, \quad \text{and} \quad \rho_0(0) = \rho_*. \quad (96)$$

The first condition implies no conductive heat flux through the lower boundary, while the second eliminates unphysical heat sources at infinity.

If $T_0(0)$ assumes the particular value $T_* = (c_2/c_1)^{1/N}$, the solution is an isothermal atmosphere with the temperature equal to this value at all z . In this atmosphere, the net heat loss vanishes everywhere, the temperature required for this being independent of the density. As the equations are non-linear, it is not obvious that solutions based on any other assumed value of $T_0(0)$ would contradict the second requirement (96), so this solution may not be unique. The density is given by

$$\rho_0(z) = \rho_* e^{-z/H} = \rho_* \zeta(z), \quad (97)$$

where the scale height is $RT_*/\mu g$, and the pressure is given by

$$p_0 = RT_* \rho_0/\mu = R\rho_* T_* \zeta(z)/\mu = p_* \zeta(z). \quad (98)$$

c) Characteristic Equation for Condensation Mode

The general treatment of the linearized equations would include both sound waves and condensations with various wavenumbers in the x - y -plane. In order to simplify the analysis, we eliminate the wave modes by omitting the acceleration terms. The resulting equations are adequate for studying the condensation mode in the limit where the growth rates are small compared to the speed of sound divided by the scale height; this limit is equivalent to taking σ_T and σ_ρ small in § II. Furthermore, we study only the modes with no horizontal motions. By so doing we still gain some insight into the effects of gravitation, which will be felt primarily in the vertical direction.

The perturbation equations are

$$n\rho_1 + \frac{d}{dz}(\rho_0 v_z) = 0, \quad (99)$$

$$\frac{dp_1}{dz} + g\rho_1 = 0, \quad (100)$$

$$\begin{aligned} \frac{n\rho_1}{\gamma-1} - \frac{n\gamma p_0 \rho_1}{\rho_0(\gamma-1)} + g\rho_0 v_z - c_1(N-1)\rho_0 T_0^N \rho_1 \\ + c_1 \rho_0^2 p_0^{-1} N T_0^N p_1 - \eta T_0^{5/2} \frac{d^2 T_1}{dz^2} = 0, \end{aligned} \quad (101)$$

and equation (14), where $T_0 = T_*$, and ρ_0 and p_0 are given by equations (97) and (98). The fourth-order system is readily reduced to second order. Equations (99) and (100) show that

$$p_1(z) = \left[g\rho_* T_*^{-1} \int_0^z T_1 dz + a_1 \right] e^{-z/H} = g\rho_* n^{-1} v_z(z) e^{-z/H} + a_2. \quad (102)$$

If we now impose the boundary conditions (96) and require that $v_z(0) = 0$ and $v_z(\infty) < \infty$ (which follow by eliminating unbounded solutions and from the assumption of a rigid, non-conducting bottom boundary), we conclude from equation (102) that $p_1(\infty) = 0$ and that $a_2 = 0$. Equation (102) implies that $p_1(0) = 0$, so that $a_1 = 0$. Hence

$$p_1(z) = g\rho_* T_*^{-1} e^{-z/H} \int_0^z T_1 dz = g\rho_* n^{-1} v_z(z) e^{-z/H}. \quad (103)$$

By use of this expression v_z may be eliminated from the heat equation, with the result

$$n\gamma(\rho_0 p_1 - p_0 \rho_1)/\rho_0(\gamma - 1) + Nc_2\rho_0(\rho_0 p_1 - p_0 \rho_1)/p_0 - \eta T_0^{5/2} d^2 T_1/dz^2 = 0, \quad (104)$$

where we have replaced $c_1 T_0^N$ by c_2 . In view of equation (14), this equation contains T_1 only, and may be written

$$\zeta d^2 T_1/d\zeta^2 + \zeta dT_1/d\zeta + \sigma^2 \zeta(\zeta - \nu) T_1 = 0, \quad (105)$$

where

$$\sigma = Hk_{c*}, \quad \nu = -\gamma n/ck_{T*}. \quad (106)$$

k_{T*} is the value of k_T calculated at $z = 0$, and k_{c*} is the critical wavenumber $(-k_{K*}k_{T*})^{1/2}$, which agrees with equation (26) since $k_{\rho*} = 0$. Also σ is 2π times the number of critical wavelengths in a scale height; it is inversely proportional to the thermal conductivity; ν is the growth rate normalized with respect to that which would obtain for the condensation mode in a homogeneous atmosphere having density ρ_* and temperature T_* .

d) Eigenfunctions

The characteristic equation (105) is to be solved for arbitrary σ ; ν is then determined by imposing the conditions $T_1(\infty) < \infty$ and $T_1'(0) = 0$. The solution is

$$T_1(\zeta, \nu, \sigma) = e^{i\sigma\zeta} M(\tfrac{1}{2} + \tfrac{1}{2}i\nu\sigma, 1, -2i\sigma\zeta), \quad (107)$$

where M is the confluent hypergeometric function (Jahnke and Emde 1945). The term $\nu(\sigma)$ is defined by the condition that

$$\frac{dT_1}{d\zeta}(1, \nu, \sigma) = 0. \quad (108)$$

Let us first look at the marginally unstable modes ($\nu = 0$). From equation (107) these can be written

$$T_1(\zeta, 0, \sigma) = J_0(\sigma\zeta), \quad (109)$$

so that equation (108) implies that σ is a root of

$$\frac{dT_1}{d\zeta}(1, 0, \sigma) = \sigma J_1(\sigma) = 0; \quad (110)$$

these roots are $\sigma = 0, 3.83, 7.02, 10.17$, etc. Now consider what happens as k_{T*} moves through zero from positive to negative values. When $k_{T*} > 0$, σ is imaginary and there are no roots of equation (110). When $k_{T*} = 0$ we must reformulate the problem in view of definition (106). The solution is readily obtained as

$$T_1(\zeta, k_{T*} = 0) = I_0[2H(k_{K*}n\zeta/c)^{1/2}], \quad (111)$$

a function whose derivative at $\zeta = 1$ vanishes if and only if $n = 0$ and $T_1 = \text{const}$. It follows that a necessary condition for instability is $k_{T*} < 0$ or $N < 0$; this criterion is in agreement with condition (4b) since in this case $k_\rho = 0$. In the rest of this section we assume that this condition is fulfilled, and examine the effects of thermal conduction.

Consider what happens as conductivity drops from ∞ , or σ increases from 0. For small values of σ the expansion of equation (107) yields

$$\begin{aligned} T_1 = & 1 + \nu_1 \zeta_1 + (2!)^{-2}(\nu_1^2 - 1)\zeta_1^2 + (3!)^{-2}(\nu_1^3 - 5\nu_1)\zeta_1^3 \\ & + (4!)^{-2}(\nu_1^4 - 14\nu_1^2 + 9)\zeta_1^4 + (5!)^{-2}(\nu_1^5 - 30\nu_1^3 \\ & + 89\nu_1)\zeta_1^5 + (6!)^{-2}(\nu_1^6 - 55\nu_1^4 + 439\nu_1^2 - 225)\zeta_1^6 + \dots, \end{aligned} \quad (112)$$

where

$$\zeta_1 = \sigma\zeta, \quad \nu_1 = \sigma\nu. \quad (113)$$

If we subject this solution to condition (108) we find

$$\nu = \frac{1}{2} \left(1 + \frac{\sigma^2}{4!} - \frac{\sigma^4}{6!} + \dots \right), \quad (114)$$

so that as $\sigma \rightarrow 0$, $\nu \rightarrow \frac{1}{2}$, and $T_1 \rightarrow \text{constant}$. This corresponds to the simultaneous cooling or heating of the whole atmosphere. The growth rate is

$$n = -ck_{T*}/2\gamma = -(\gamma - 1)\mu c_1 \rho_* T_*^{N-1} N/2\gamma R. \quad (115)$$

The pressure and density perturbations are

$$p_1 = p_* T_1 z (HT_*)^{-1} e^{-z/H}, \quad \rho_1 = -\rho_* T_1 T_*^{-1} (1 - z/H) e^{-z/H} \quad (116)$$

as shown in Figure 7 (see below). If $T_1 < 0$, so that the atmosphere cools everywhere, the material moves downward and condenses at the bottom. The pressure decreases everywhere, with the largest change at $z = H$. The resulting pressure forces are downward in the upper regions and upward in the lower regions; these forces are compensated by gravitation acting on the perturbed density. Of course, if $T_1 > 0$, the atmosphere expands and everything is reversed.

We conclude that this idealized atmosphere exhibits considerably different behavior from that of a uniform medium as the conductivity becomes large. In the latter case, disturbances having positive growth rate necessarily had very large scale so that the temperature variations necessary to balance the density variations in an isobaric condensation would not be destroyed by thermal conduction. In the stratified atmosphere, temperature variations are no longer necessary, because the pressure variations are compensated by gravity, so disturbances in density having finite scale grow even when the conductivity is high. Since equation (103) may be written $dp_1/dt = 0$, however, the perturbations are isobaric as seen by an observer moving with the gas.

As σ is increased, the temperature perturbation in the fundamental mode is more concentrated to the bottom of the atmosphere. Values of T_1 can be calculated from equation (112) for small values of σ and by numerical integration for intermediate values of σ . For the large values of σ (small conductivity), T_1 can be calculated by a variational technique. The corresponding value of ν is given by

$$\nu = 1 - \frac{3}{2}\pi^{-1/3} \sigma^{-2/3} + \pi^{-2/3} \sigma^{-4/3} + O(\sigma^{-2}). \quad (117)$$

Curves of T_1 (fundamental mode) are presented in Figure 6. Corresponding curves for the density are in Figure 7, and $\nu(\sigma)$ is shown in Figure 8. Evidently, as $\sigma \rightarrow \infty$, n approaches the value appropriate for a homogeneous atmosphere whose properties coincide with those of the bottom layers.

As σ increases, one encounters marginal instability for higher modes, in accordance with equation (110), and as σ increases further, these higher modes become truly unstable. For example, Figure 9 shows the temperature in the fourth marginally unstable mode. The amplitude increases with z because of conductivity smoothing of the finer scales at the lower heights. If we number the modes $j = 0, 1, 2, \dots$, the condition that the j th mode is unstable can be written as

$$\sigma_j < Hk_{c*} \quad (118)$$

where σ_j is the j th root of equation (110). Since σ_j/H is approximately the vertical wave-number in the j th mode, this relation is similar to the condition $k < k_c$ encountered in § II.

VI. EFFECT OF EXPANSION

a) Physical Considerations

Intuitively one would expect that formation of condensations in an expanding medium would be particularly difficult, since in the absence of disturbances, the density tends

to decrease everywhere. This question is relevant to condensation phenomena in nova shells, stellar winds, planetary nebulae, and the expanding universe. Actually, if the zero-order state is a uniform expansion with an infinite heat sink (which might approximate the cooling of a gas which has been heated and expanded by a radial shock wave), one finds that for certain cooling laws there is still a condensation instability for short wavelengths. What happens is that at certain epochs in the expansion, the inward compression of a condensation by outlying hot gas proceeds fast enough to overcome and reverse the original expansion effect. In all cases, however, this ceases at a certain finite time, owing to the general decrease of density and consequent reduction in cooling rate.

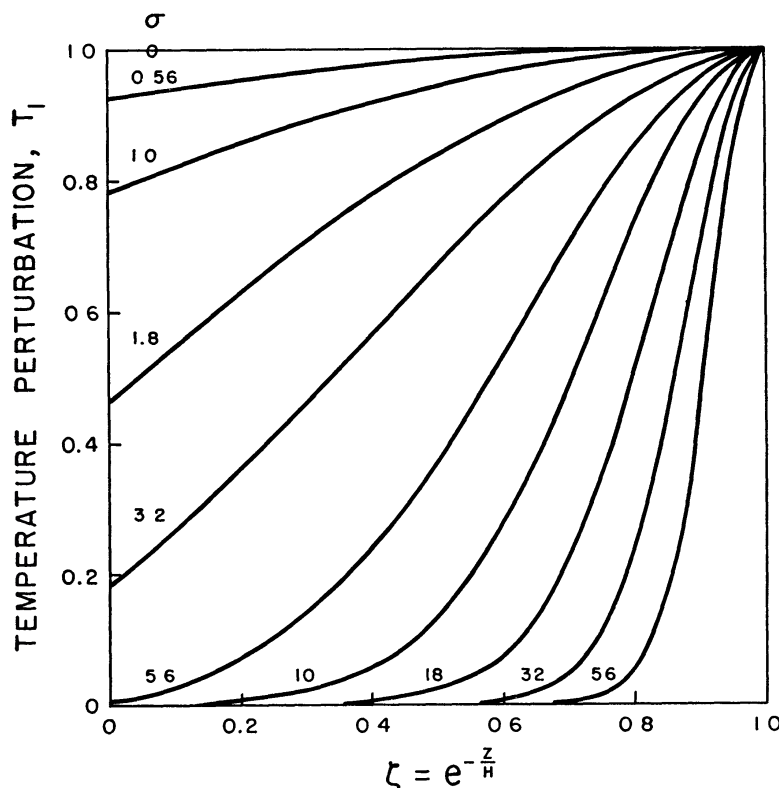


FIG. 6.—Solutions of equation (105) for the fundamental mode of the temperature perturbation in an isothermal atmosphere, normalized at $z = 0$. The parameter σ , defined by eq. (106), increases with decreasing thermal conductivity.

In a larger variety of cases it is found that condensations cool faster and decrease in density more slowly than the surrounding material, even for arbitrarily large times. The ultimate fate of such regions, which become relatively more and more dense compared to their surroundings, cannot be described by linear theory. Conduction limits the growth of small condensations, as usual. However, if the thermal conductivity decreases as the temperature does, owing to the expansion, this effect ceases after a certain time, and instability may proceed if the proper conditions in the absence of thermal conduction are satisfied.

In order to make progress in view of the complexities introduced by a time-dependent zero-order state, we restrict attention to a uniformly expanding gas. A particular problem of interest is the formation of galaxies in a uniformly expanding gas supposed to constitute the original material of the universe. As Stephan Hawking, of Cambridge University, has pointed out, our idealization is inappropriate for one of the closed ex-

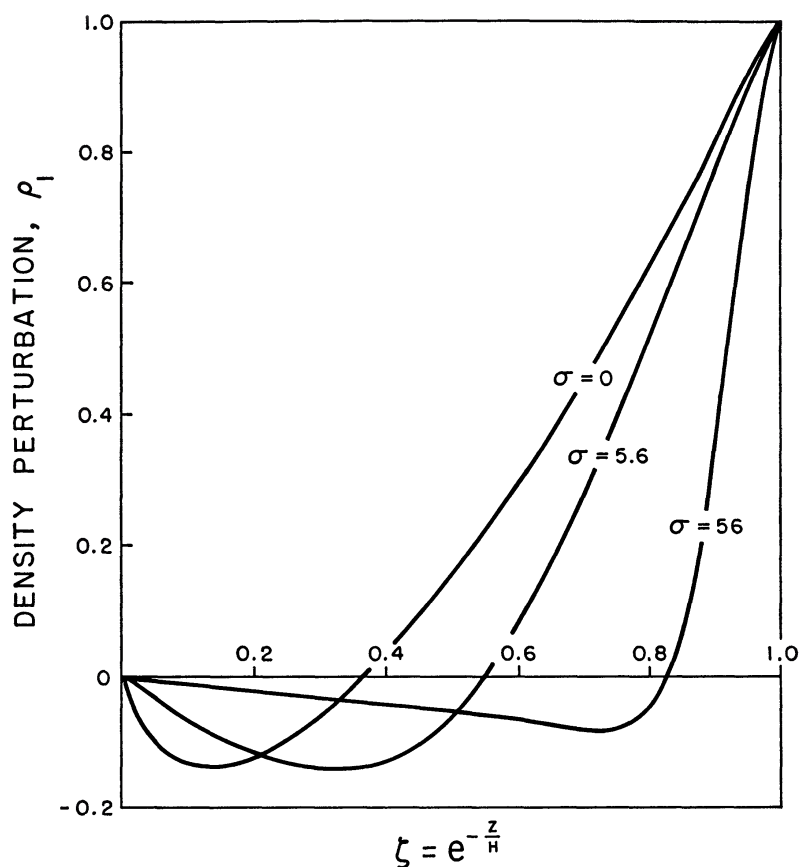


FIG. 7.—Density perturbations in an isothermal atmosphere, corresponding to the temperature perturbations of Fig. 6, normalized at $z = 0$. The density at $z = 0$ increases at the expense of the density at higher levels.

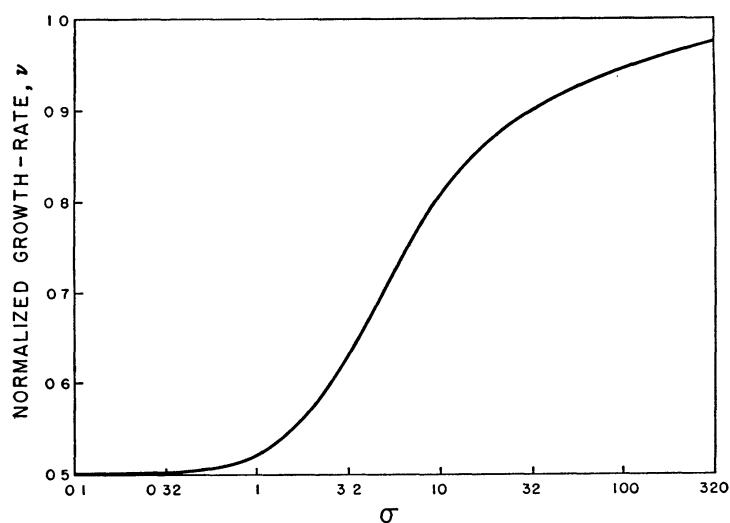


FIG. 8.—Normalized growth-rate of the fundamental condensation mode in an isothermal atmosphere as a function of the normalized thermal conductivity (decreasing to the right).

panding universes described by general relativity, since radiation emitted by the gas is trapped in the universe by the gravitational field and cannot escape from the system as we have assumed. Hence our calculations can have relevance only if the system of galaxies is both finite and open, i.e., a classical astronomical system which would violate the cosmological principle. The observations in favor of a closed universe and the cosmological principle are so inconclusive that it is useful to consider such systems.

It is easily verified (Bondi 1960), that the continuity and momentum equations (6) and (7) are satisfied by the velocity and density fields,

$$v_0(r, t) = r/t, \rho_0(t) = \rho_* t_*^3/t^3, \quad (119)$$

where ρ_* is the (uniform) density at $t = t_*$, provided that the temperature depends only on the time through the heat equation,

$$\frac{R\rho_0^{\gamma-1}(t)}{\mu(\gamma-1)} \frac{d}{dt} \left[\frac{T_0(t)}{\rho_0^{\gamma-1}(t)} \right] - \mathfrak{L}[\rho_0(t), T_0(t)] = 0. \quad (120)$$

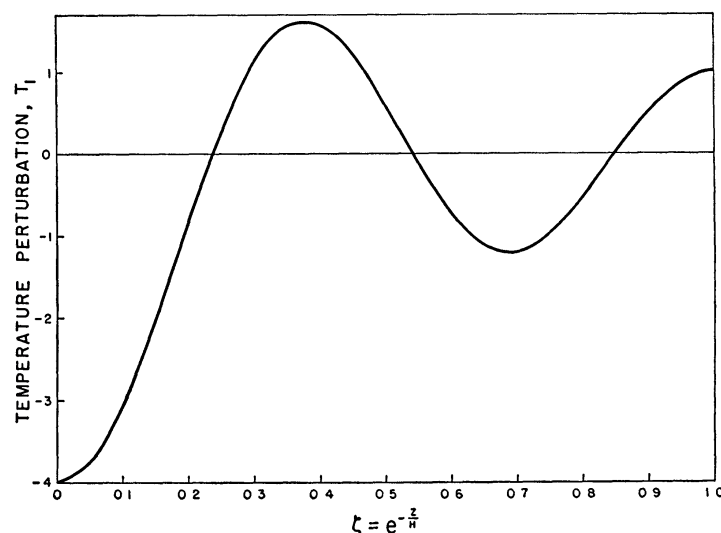


FIG. 9.—The fourth marginally unstable temperature perturbation in an isothermal atmosphere, normalized at $z = 0$; $\sigma = 10.17$.

This solution represents an infinite uniform distribution of gas in which pressure forces and gravity vanish, so that v_0 is constant for each particle. Such a system would approximate the local behavior of a low-density model of the expanding universe, although its implications become unphysical as v_0 approaches the speed of light.

The equation for T_0 is readily solved in the special case when the heat-loss function is of the form (93) with $c_2 = 0$, appropriate for collisional cooling and no energy input. If $\gamma = \frac{5}{3}$, the solution for N not equal to zero or one is

$$T_0(t) = A^{1/(1-N)} t^{-2} (1 + b_N t^{-2N})^{1/(1-N)}, \quad (121)$$

where A is an arbitrary positive constant and

$$b_N = (1 - N) \mu c_1 \rho_* t_*^3 / 3 N R A. \quad (122)$$

When b_N is small, signifying a low cooling rate, T_0 is proportional to t^{-2} and therefore to $\rho_0^{2/3}$, as in an adiabatic expansion. A evidently labels the adiabat that would apply in the absence of cooling.

The effects of a finite cooling rate are quite different, depending on the value of N . If $N < 0$, $b_N < 0$, and T_0 drops to zero at the finite time, $(-b_N)^{1/2N}$, as a consequence of the increased cooling rate as T_0 falls. The behavior is adiabatic for $t \ll (-b_N)^{1/2N}$. When $N = 0$,

$$T_0(t) = At^{-2} (1 + b_0 \ln t), \quad (123)$$

so that T_0 approaches zero slowly at the finite time e^{-1/b_0} , where

$$b_0 = -2\mu c_1 \rho_* t_*^3 / 3RA. \quad (124)$$

When $0 < N < 1$, $b_N > 0$, and $b_N^{1/2N}$ is a time such that cooling is important for earlier times and negligible for later times. The adiabat approached asymptotically depends on the value of A and on the effectiveness of cooling. When $N = 1$,

$$T_0(t) = At^{-2} e^{b_1/t^2}, \quad (125)$$

where

$$b_1 = \mu c_1 \rho_* t_*^3 / 3R. \quad (126)$$

This solution is adiabatic for times later than b_1 , but decreases very rapidly from ∞ as t increases from 0. Finally, when $N > 1$, $b_N < 0$, and $(1 - N)^{-1} < 0$, so that the temperature is infinite at the finite time $(-b_N)^{1/2N}$; the solution is meaningful only at later times. The behavior is nearly adiabatic after several such periods have passed.

b) Condensations: Zero Conductivity

The results of § II showed that condensations of sufficiently small scale developed in an isobaric manner; we could have derived the correct instability criterion, as well as the growth rate for small scales, simply by placing $p_1 = 0$ in the perturbed heat equation. This procedure will be used here to avoid the necessity of solving the perturbed continuity and momentum equations, and should yield the correct instability criterion for the condensation mode, although of course the wave mode is lost. We shall confine attention to radial perturbations and to the cases $N \neq 0$ or 1. The equation for the "relative temperature," $\tau = T_1/T_0$, is

$$\frac{\partial \tau}{\partial t} + \frac{r}{t} \frac{\partial \tau}{\partial r} - \frac{\delta(t) \tau}{t} = \frac{r_1^2(t)}{tr^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \tau}{\partial r} \right), \quad (127)$$

where

$$\delta(t) = \delta_N(1 + t^{2N}/b_N), \quad \delta_N = 6N(2 - N)/5(1 - N), \quad r_1^2(t) = 2\mu t K_0(t)/5R\rho_0(t). \quad (128)$$

Equation (127) indicates that the time-dependence of τ is not exponential, as in previous sections, because of the time variation of the unperturbed state. Equation (127) is simplified by the introduction of the Lagrangian variable $v (= v_0)$ to

$$\frac{\partial \tau}{\partial t} - \frac{\delta(t) \tau}{t} = \frac{r_1^2(t)}{t^3 v^2} \frac{\partial}{\partial v} \left(v^2 \frac{\partial \tau}{\partial v} \right). \quad (129)$$

Further, if we write

$$\tau(t, v) = g(t, v) \exp \int^t t^{-1} \delta(t) dt = g(t, v) (1 + b_N t^{-2N})^{-\delta_N/2N}, \quad u(t) = \int^t t^{-3} r_1^2 dt, \quad (130)$$

the equation becomes

$$\frac{\partial g}{\partial u} = \frac{1}{v^2} \frac{\partial}{\partial v} \left(v^2 \frac{\partial g}{\partial v} \right), \quad (131)$$

which is just the diffusion equation for a radial disturbance, with u playing the role of time, and v , the role of space.

If the thermal conductivity is zero, equation (131) shows that g is an arbitrary function of v , so that

$$\tau(u, v) = g(v)(1 + b_N t^{-2N})^{-\delta_N/2N} \quad (132)$$

$$T_1(u, v) = g(v)A^{1/(1-N)}t^{-2}(1 + b_N t^{-2N})^{1/(1-N)-\delta_N/2N}, \quad (133)$$

and

$$\rho_1(u, v) = -g(v)\rho_* t_*^3 t^{-3}(1 + b_N t^{-2N})^{-\delta_N/2N}. \quad (134)$$

Each expression contains a form factor, $g(v)$ and a time-dependent amplitude. Our concern is with the absolute value of the amplitude. If the amplitude of T_1 or ρ_1 increases with time, we may speak of *absolute growth* in the sense that the absolute value of the temperature or density perturbations grows with time. If the τ -amplitude increases with time, but the other amplitudes decrease, we speak of *relative growth*; the contrast between the perturbed and average values increases. In such a situation, the high-density regions (for example) become less dense with time, but at a slower rate than the average. Hence, the most that can result from relative growth is that the density or temperature of a region levels off, rather than changing at the same rate as the average.

Equations (132)–(134) permit one to calculate the signs of the time derivatives of τ , T_1 , and ρ_1 , as functions of N and of t . Here $\tau' > 0$ for $N < 2$ and $\tau' < 0$ for $N > 2$, for all t . $T_1' < 0$ if $N > 0$; if $N < 0$, $T_1' < 0$ for $0 < t < t_T$, $T_1' > 0$ for $t_T < t < t_0$, where t_T is defined by the equation,

$$(RA/\mu c_1 \rho_* t_*^3) t_T^{2N} = -(3N^2 - 6N + 5)/15N, \quad (N < 0), \quad (135)$$

and t_0 by

$$(RA/\mu c_1 \rho_* t_*^3) t_0^{2N} = -(1 - N)/3N, \quad (1 < N < 0). \quad (136)$$

One finds that $t_T < t_0$ (the moment when $T_0 = 0$ for $N < 0$) so that there is always some growth of temperature fluctuations in the later stages if $N < 0$.

The behavior of ρ_1' is more complex, and depends on the value of N relative to 0, 1, 2, and N_- , where the latter is defined as the smaller root of $2N^2 - 9N + 5 = 0$, or $\frac{1}{4}(9 - \sqrt{41}) = 0.65$. It also depends on the time relative to t_0 (equation 136) and to t_ρ , defined by

$$(RA/\mu c_1 \rho_* t_*^3) t_\rho^{2N} = -(2N^2 - 9N + 5)/15N, \quad (N < 0 \text{ or } N_- < N < N_+), \quad (137)$$

where N_+ is the root complementary to N_- , 3.85.

We may summarize the growth of condensations as follows. Relative growth occurs if and only if $N < 2$. Absolute growth occurs if and only if $N < 0$ or $N_- < N < 2$, but only for specified intervals of time. If $N < 0$, growth takes place only between $t = t_\rho$ and $t = t_0$. If $N_- < N < 2$, the ρ_1 -amplitude grows from the first moment for which the solution for T_0 makes sense, up to the time t_ρ , but decreases thereafter.

These results may be compared with those which would apply to a static medium cooling according to equation (93), with no heat input ($c_2 = 0$). In that case one finds that ρ_1' is always positive for $N < 2$, so that growth occurs up to t_0 for $N < 0$ and for $t \rightarrow \infty$ for $0 < N < 2$. It appears that as compared with this case, expansion indeed stabilizes the situation by eliminating absolute growth for $0 < N < N_-$, and by confining growth for other values of N to periods earlier than t_ρ . The latter effect may be understood as the result of the general decrease in ρ in the expansion, resulting in a decreasing cooling efficiency.

We may also compare the situation with that which obtains for a static gas in thermal equilibrium, with an energy input independent of ρ and T . The isobaric criterion is $N < 1$ for this case, so that a system which is cooling in the unperturbed state is more unstable than one in thermal equilibrium. Our results can be understood, therefore, as an enhancement of instability by lack of heat input, coupled with a suppression of instability by the effects of expansion.

c) *Conduction Effects*

We seek separable solutions to equation (131), of the form

$$g(u, v; \alpha) = v^{-1} e^{-\alpha^2 u} \sin \alpha v, \quad (138)$$

where the separation constant, α , is $2\pi/\Delta v$, with Δv a wavelength in velocity space. To go further, one must assume more about the properties of the medium; we shall consider ionized hydrogen, with K that given in § Vb, and cooling by free-free emission ($N = \frac{1}{2}$). In this case, only relative growth of condensations can be expected; we seek here the further limitations imposed by conduction.

Introducing equations (138) and (128) into equation (130), we find that

$$\tau(t, v; \alpha) = v^{-1} \sin \alpha v (1 + b_{1/2} t^{-1})^{-9/5} \exp \left[-d \alpha^2 \int^t t^{-4} (1 + b_{1/2} t^{-1})^5 dt \right], \quad (139)$$

where

$$d = 2\mu\eta A^5 / 5R\rho_* t_*^3. \quad (140)$$

Since the integral in the exponential increases with time, it is evident that conduction inhibits the relative growth of density perturbations, and that this inhibition is stronger for larger wavenumbers, as expected. However, the damping term levels off after a certain time, owing to the drop in thermal conductivity, and the growth represented by the first term takes over. One finds that τ' changes from negative to positive at the time t_K given by

$$(b_{1/2} t_K^{-1})^2 (1 + b_{1/2} t_K^{-1})^6 = \zeta = 9b_{1/2}^3 / 5d\alpha^2 = \mu^2 c_1^3 \rho_*^4 t_*^{12} / 6R^2 A^8 \eta \alpha^2. \quad (141)$$

One finds that if $\zeta \ll 1$, so that $t_K \gg b_{1/2}$,

$$t_K = (2\eta/3c_1)^{1/2} A^3 \alpha / \rho_* t_*^3. \quad (142)$$

If $\zeta \gg 1$, $t_K \ll b_{1/2}$, so that there is a long interval after conductive stabilization ceases and before cooling diminishes in which relative growth may proceed. In this case,

$$t_K = 3^{-1} (6\eta)^{1/8} (\mu/R)^{1/4} c_1^{5/8} \rho_*^{1/2} t_*^{3/2} \alpha^{1/4}, \quad (143)$$

independent of the value of A . As A is not known from observation in the case of the expanding universe, this is a fortunate result. In other words, the phase of the expansion at which conduction becomes ineffective can be calculated (for any given scale) from a knowledge of the present density alone, provided that the phase in question occurred before the cessation of cooling.

VII. APPLICATIONS

a) *Introduction*

In this section we consider several of the calculations of thermal instability for various situations which have appeared in the literature. In his original paper, Parker (1953) attributed solar flares and prominences to this phenomenon; prominences in particular have been studied by several other authors from this standpoint. Athay and Thomas (1956) applied the concept of thermal instability to the determination of the chromospheric temperature profile. Zanstra (1955*a, b*) pointed out that condensations in planetary nebulae could form in this way, while Spitzer (1956) suggested that spiral arms might condense out of the galactic halo in a similar manner. Gold and Hoyle (1959) argued that galaxies could similarly condense from the intergalactic medium, while Field (1962*a*) applied the idea to the condensation of interstellar clouds.

b) *Chromospheric Temperature Profile*

Athay and Thomas (1956) subjected the chromospheric temperature to the condition $\mathfrak{L}_T > 0$, which we recognize as the isochoric stability criterion. Two objections can be

made to this procedure. First, as in § V, conduction should be taken into account by analyzing a proposed steady-state atmosphere by the normal-mode method. Estimates made by Kopp (1963) suggest that the conductive damping in the chromosphere may not be negligible.

Another objection is that, even if conduction is omitted, the proper criterion is not the isochoric but the isobaric criterion. Since they considered $\mathfrak{L}$ to be of the form,

$$\mathfrak{L} = c_1 \rho L(T) - c_2 G(\rho), \quad (144)$$

where L represents a collisional heat loss and G a gain from wave dissipation, the isobaric stability criterion is

$$\frac{d \log L}{d \log T} > 1 + \frac{d \log G}{d \log \rho}. \quad (145)$$

If G is proportional to ρ , this agrees with the isochoric criterion, but in the more likely event that it is independent of ρ , the L -derivative must exceed unity for stability. It happens that in the case considered by Athay and Thomas, it tends to exceed unity whenever it exceeds zero, so that the practical effect of this change is small.

c) Solar Corona: Solar Wind and Prominences

The heat-loss function of the corona may be taken in the form (144) with $L = T^N$ and $G = \rho^M$, where the exponent N of the collisional-loss term is ≈ 0 according to Weymann (1960a) and Orrall and Zirker (1961). The value of M would have to be unity in order to apply the results of § V; if the input is by wave dissipation, a smaller value (approaching zero) appears more likely. The mean of Weymann's and Orrall and Zirker's determination of c_1 is 5×10^{24} cgs units. If we take the bottom of the corona to be characterized by $n_e = 2 \times 10^8 \text{ cm}^{-3}$ and $T = 10^6 \text{ }^\circ \text{K}$, and compute the thermal conductivity from § Vb, we derive the following parameters:

$$\begin{aligned} k_{\rho*} - k_{T*} &= 7 \times 10^{-13}(1 - M - N)\text{cm}^{-1}, & k_{K*} &= 7 \times 10^{-10} \text{ cm}^{-1}, \\ k_{e*} &= 2 \times 10^{-11}(1 - M - N)^{1/2} \text{ cm}^{-1}, & H &= 5 \times 10^9 \text{ cm}^{-1}, \\ \sigma &= 0.1(1 - M - N)^{1/2}, \end{aligned} \quad (146)$$

where we have used the isothermal-atmosphere approximation.

We note that local instability occurs in the absence of conduction if $1 - M - N > 0$. This inequality provides a basis for reconciling Kiepenheuer's (1953) statement that since free-free and free-bound losses are proportional to density they should lead to instability, and Athay and Thomas' (1956) statement that since $N = \frac{1}{2}$ for free-free emission, it should lead to stability. Both statements are partially correct in the sense that the ρ -dependence of the collisional loss leads to the $+1$ in the inequality, while a positive value of $N (= \frac{1}{2})$ in fact reduces the value of the left-hand side. Thus, the correct criterion takes *both* into account, along with a stabilizing term due to the ρ -dependence of the gain term. For example, if $M \approx 0$, free-free emission is unstable, and free-bound emission is even more so.

Although σ has well-defined meaning only for an isothermal atmosphere ($M = 1$), it seems likely that its small magnitude compared to the 3.83 necessary for instability of higher-order modes implies conductive stabilization of all but the fundamental mode in the absence of a magnetic field. If we accept Orrall and Zirker's result that $N \approx 0$, the fundamental mode is unstable if and only if $M < 1$; although the value of M is very uncertain, this inequality seems likely to hold. We therefore expect the fundamental mode to be unstable, with a growth rate given by

$$n = c(k_{\rho*} - k_{T*})/2\gamma = 3 \times 10^{-6}(1 - M - N) \text{ sec}^{-1}, \quad (147)$$

corresponding to an e -folding time of $4(1 - M - N)^{-1}$ days. The instability should lead to gross expansion or contraction of the atmosphere, with temperature and density perturbations like those found in § V. According to equation (103), the perturbation velocity may be written,

$$v_z(z) = nH p_1(z)/p_0(z), \quad (148)$$

so that if the perturbation pressure should grow to be comparable with the ambient pressure, the velocity could approach the value $nH = 1.5 \times 10^4 (1 - M - N)$ cm/sec. The corresponding electron flux at about 1 scale height could attain $F = 3 \times 10^{12} (1 - M - N)$ cm⁻² sec⁻¹. The linear theory cannot tell us whether the corona will expand or contract. If it expands, it could lead to an electron flux at the Earth of $F_{\oplus} = 8 \times 10^7 (1 - M - N)$ cm⁻² sec⁻¹, a value which may be compared with observed values of about 6×10^8 cm⁻² sec⁻¹ in the solar wind (Parker 1963). The explanations advanced by Parker for the solar wind turn on the impossibility of a static equilibrium. The above considerations suggest that, even if a static equilibrium were possible, it could well be unstable, tending toward a dynamic equilibrium having properties rather like those observed.

Turning to the formation of small structures such as prominences, we note that they cannot be attributed directly to thermal instability unless the conductivity is suppressed. The most reasonable agency to do this is a magnetic field, so prominences should form preferentially in magnetic regions, in agreement with observation. Let us consider the shape of the prominence in relation to the geometry of the magnetic field. According to § III, the half-wavelength parallel to $\mathbf{B}$ must satisfy

$$\frac{1}{2}\lambda_{||} > \pi/k_{c*} = 1.4 \times 10^{11} (1 - M - N)^{-1/2} \text{ cm}, \quad (149)$$

equal to roughly 1 solar diameter. (We assume here that results for a uniform gas are approximately valid, since the vertical extent is ≤ 1 scale height.) Since many prominences are observed to be much shorter than this, we conclude that, if prominences are to be interpreted as condensations in the corona, the material within them must be initially isothermal along the field lines, up to the points where the field enters the chromosphere. (At the latter points the conductivity drops sharply and temperature gradients are permitted.) The identification of prominences with fairly uniform regions extending along magnetic lines of force also agrees with observation.

Since no temperature variations are permitted along the coronal lines of force, the instability must develop normal to $\mathbf{B}$, and is therefore governed by equations (61), (63), and (64). With the adopted heat-loss function, the modified isobaric instability criterion (48) applicable here can be written

$$(1 - M)(1 + B_0^2/4\pi p_0)^{-1} > N, \quad (150)$$

so that for $N = 0$, all values of $M < 1$ imply instability, independent of the magnetic-field strength. Since this condition on M seems likely to be fulfilled, one would expect instability to occur wherever the field is strong enough to suppress conduction, if Orrall and Zirker's estimate that $N = 0$ is correct.

On the other hand, if N is positive, equation (49) implies that instability can occur only when B_0 is less than the value

$$B_0 = [4\pi p_0(1 - M - N)/N]^{1/2}, \quad (151)$$

which is about 1 gauss if $(1 - M - N)/N$ is near unity ($p_0 = 0.06$ dyne/cm²). Magnetic stabilization would then be expected to occur throughout large parts of a magnetic region, and only those areas having magnetic fields large enough to suppress conduction, but not large enough to exert significant pressures, would be expected to be suitable for the formation of prominences.

The discussion so far pertains only to the initial growth of the prominence, when the linear equations apply. The fully developed prominence has a density about 100 times larger than that of coronal material, and if condensation occurs across the lines of force as we have suggested, the magnetic-field strength should be in the same ratio and the magnetic pressures should be in the ratio 10^4 . This means that, if the internal magnetic pressure is never to exceed the external gas pressure, the initial magnetic-field strength must be less than 10^{-2} gauss. The observed magnetic-field strength in quiescent prominences (Zirin 1962) is about 50 gauss, corresponding to initial values of about $\frac{1}{2}$ gauss. Since $B^2/8\pi$ is much greater than the gas pressure for such fields, the field must be almost force-free in the fully developed prominence, because of some mechanism not considered in our theory. One suggestion is that the lines of force are constantly being twisted by motions in the photosphere (Danielson 1964). In any case, if the field continually adopts a force-free geometry because of external circumstances, the growth of a condensation can proceed as if the dynamical effects of the field were not present, although magnetic inhibition of conduction would still be important. In this event, one readily derives that

$$\frac{1}{2}\lambda_{c1*} = 7 \times 10^5 (1 - M)^{1/2} \text{ cm}, \quad \frac{1}{2}\lambda_{m1*} = 2 \times 10^9 (1 - M)^{1/4} \text{ cm} \quad (152)$$

for $N = 0$, $B_0 = 1$ gauss. The latter corresponds to a diameter of $2000 (1 - M)^{1/4}$ km in the condensed state, compared to Orrall and Zirker's quoted observational value, $\lesssim 1000$ km. The cooling rate is still given by equation (147).

What is the relation of the picture presented here to proposals by other authors? Kleczek (1957) derived an isobaric cooling time of 10^7 sec (compared to our $3 \times 10^5 [1 - M - N]^{-1}$ sec), using different calculations of the heat-loss function. He disregarded the magnetic-pressure problem (the field strengths not having been determined at that time). Later (Kleczek 1958) he pointed out that the time scales of certain AS prominences are too short (10^2 – 10^4 sec) to be attributable to thermal instability, and instead postulated compression by magnetic forces, followed by rapid cooling at high density. Kiepenheuer (1959) indicated the type of magnetic forces which could bring about the compression. Lüst and Zirin (1960) also assumed magnetic forces were active and showed that conductive cooling into the chromosphere could reduce the time scale. Uchida (1963) proposed a specific model of the magnetic field, according to which prominences in regions from which the magnetic field, ordinarily present, has been removed by a cloud of cosmic rays. Subsequently field-free matter runs into the cavity, is compressed by the surrounding magnetic stresses, and cools to form a prominence.

Evidently the key point in the theory of prominences is the role of the magnetic field. If Zirin is correct in postulating that the magnetic field is force-free in the fully developed prominence, it may be reasonable to assume that it is during the development phase, at least in quiescent prominences. Our calculations then show that gas pressure forces alone can lead to the formation, through thermal instability, of quiescent prominences which are similar to those observed in respect to density, dimensions, shape (in relation to field structure), and time scale. It still appears necessary to assume compression by magnetic forces to account for the short time scales of active prominences, however.

d) Galactic Halo

Since the galactic halo, or corona, proposed by Spitzer (1956) is rather similar to the solar corona in some respects, we may proceed as in the previous subsection. Taking $n_e = 5 \times 10^{-4} \text{ cm}^{-3}$ and $T = 10^6$ °K, we find that

$$\begin{aligned} k_{\rho*} - k_{T*} &= 2 \times 10^{-24} (1 - M - N) \text{ cm}^{-1}, & k_{K*} &= 3 \times 10^{-21} \text{ cm}^{-1}, \\ k_{c*} &= 7 \times 10^{-23} (1 - M - N)^{1/2} \text{ cm}^{-1}, & H &= 2 \times 10^{22} \text{ cm}, \\ \sigma &= 2(1 - M - N)^{1/2}. \end{aligned} \quad (153)$$

Again, equation (147) applies, showing that, if M is nearly zero, it is not necessary that $N < 0$ for instability to occur, as Spitzer suggested. Rather, both inelastic electron-ion collisions ($N = -\frac{1}{2}$) and free-free emission ($N = \frac{1}{2}$) would imply instability.

The somewhat larger value of σ is still too small to permit anything but the fundamental mode to develop in the absence of a magnetic field, although a slight change in the assumed parameters would permit the first few harmonics to grow. In any case, it is possible that a gross expansion or contraction of the halo would result from thermal instability. It would have a growth rate $n = 7 \times 10^{-18} (1 - M - N) \text{ sec}^{-1}$ corresponding to an e -folding time of $4 \times 10^9 (1 - M - N)^{-1}$ years, in rough agreement with Spitzer's "cooling times" computed with a somewhat different heat-loss function.

Small-scale disturbances can grow only if conduction is suppressed by a magnetic field. Although equation (150) implies that growth across the field lines will occur if $N \approx 0$ (Orrall and Zirker) or $N \approx -\frac{1}{2}$ (Spitzer), independent of the value of $B_0^2/8\pi p_0$, we have seen in § III that the instability does not continue into the non-linear regime, and cannot lead to interesting density changes if the field strength is large. Since the value of $(8\pi p_0)^{1/2} = 1.8 \times 10^{-6}$ gauss in Spitzer's model, while the radio-astronomical evidence points to a field strength of 6×10^{-6} gauss (Woltjer 1962a), it appears that condensation across the field lines will be very limited, unless, as in the case of prominences, we postulate that the field is twisted (Ireland 1961; Woltjer 1962b) and force-free in the region of interest.

The analogue of equation (149) is

$$\frac{1}{2}\lambda_{||*} > \pi/k_{e||*} = 15 (1 - M - N)^{-1/2} \text{ kpc} . \quad (154)$$

Hence condensations of this length or greater will be permitted parallel to the field. Spitzer (1956) suggested that spiral arms form from the halo in this way. The difficulty is that the minimum length is so large compared with the circumference of the Galaxy (which may be taken as an estimate of the length of a tube of force) that the amount of material available would be insufficient to increase the density 1000-fold, as required to build a spiral arm from halo material.

The effects of rotation may be briefly examined, using the approximate rotation-law, $V(\varpi) = \text{constant} = 210 \text{ km/sec}$; this appears to agree with the observed law within 10 per cent for $\varpi = 1.3\text{--}12 \text{ kpc}$, according to Oort (1962). According to equation (92), there are oscillations in the marginally unstable radial condensation mode if and only if $N < 1 - M < 4N$, which for the case $M = 0$ gives $\frac{1}{4} < N < 1$. Since the value of N may or may not actually be in the range specified, it is uncertain as to whether rotational stabilization will limit the growth of waves exceeding a certain length or not; if $N = 0$ or $-\frac{1}{2}$, it will not. This behavior can be traced to the dominance of pressure over rotation for the assumed values of V (210 km/sec) and c (145 km/sec).

Field (1962a) suggested that interstellar clouds may be formed by the operation of thermal instability in a dilute intercloud medium of neutral hydrogen heated by cosmic-ray ionization and cooling by electron-ion collisions. Adopting the values $n_H = 8 \times 10^{-2} \text{ cm}^{-3}$ and $T = 10^4 \text{ }^\circ\text{K}$, we find that

$$\begin{aligned} k_\rho - k_T &= 2 \times 10^{-23} \text{ cm}^{-1}, & k_K &= 5 \times 10^{-16} \text{ cm}^{-1}, & k_e &= 9 \times 10^{-20} \text{ cm}^{-1}, \\ k_m &= 10^{-21} \text{ cm}^{-1}, \end{aligned} \quad (155)$$

using the data given in Field's paper. The growth rate is $n = 6 \times 10^{-18} \text{ sec}^{-1}$ implying an e -folding time of 6×10^9 years. This is much too long to be of interest under the assumed conditions. The regions of maximum instability would have initial diameters of order π/k_m , or 800 pc, and final diameters of 160 pc—an order of magnitude greater than observed. It appears that the postulated intercloud medium of neutral hydrogen would not condense into clouds of the required dimensions in the required time. If the medium were ionized rather than neutral, the e -folding time would be decreased to about

10^6 years, but then the problem of maintaining the temperature of the medium against radiation loss would be much more severe.

Rotation has more influence in this case than for the halo, because of the lower temperature. Since $c = 10$ km/sec, $f = 2^{1/2} V/\bar{\omega}$, and $k_T/(k\rho - k_T)$ is about 0.2, equation (87) implies a limiting wavenumber, $k_R = 17/\bar{\omega}$, corresponding to a limiting wavelength $\lambda_R = 0.4\bar{\omega}$, or about 3 kpc at the orbit of the Sun. This is still large enough so that the upper limit on scale will be imposed by the finite velocity of sound in practice, as previously indicated.

e) *Planetary Nebulae*

Zanstra (1955a, b) attributed the condensations observed in the planetary nebula NGC 7293 to the operation of thermal instability associated with the cooling due to collisions of electrons with N^+ and N^{++} ions. Since the former possesses a level which is readily excited by electron collisions, the shift from N^{++} to N^+ accompanying a drop in temperature would lead to an increase in cooling rate and hence a negative value of $\mathfrak{L}_T$. It is necessary to re-examine Zanstra's argument, however, since the instability criterion he adopted does not agree with ours. This raises a question as to whether his numerical results and the more detailed ones of Daub (1963) on the same problem are valid.

Zanstra adopted the instability criterion

$$\left(\frac{d \log p}{d \log \rho}\right)_{\mathfrak{L}=0} < 0. \quad (156)$$

Since it is readily shown that

$$\left(\frac{d \log p}{d \log \rho}\right)_{\mathfrak{L}=0} = \left(\frac{\partial \mathfrak{L}}{\partial T}\right)_p / \left(\frac{\partial \mathfrak{L}}{\partial T}\right)_\rho, \quad (157)$$

his criterion agrees with the correct one (eq. [4b]) if and only if

$$\left(\frac{\partial \mathfrak{L}}{\partial T}\right)_\rho > 0, \quad (158)$$

that is, if the gas is isochorically stable. As a counterexample to Zanstra's criterion, consider $\mathfrak{L}$ of the form $f_1(\rho)[f_2(T) - f_3(T)]$, a form which actually applies under many conditions considered by Spitzer (1948, 1949) and Spitzer and Savedoff (1950), where both heating and cooling depend in the same way on density. In this case T_0 is established by the equality of f_2 and f_3 . The value of T_0 is therefore independent of density and the equilibrium pressure is proportional to density, implying instability by Zanstra's criterion. The correct stability criterion can be written $f_2'(T_0) > f_3'(T_0)$ since $\mathfrak{L}_\rho = 0$, and this is seen to be equivalent to the condition (158) that the two criteria agree. If this stability criterion is violated, the situation is in fact unstable, contrary to the Zanstra criterion.

We therefore shall consider in what follows the range of validity of condition (158) for the situation considered by Zanstra. His heat-loss function is of the form,

$$\mathfrak{L}(\rho, T) = \alpha(T)[c_1 \rho^2 T^{-1/2} e^{-1/T} - c_2 \rho(\bar{T} - \beta T)], \quad (159)$$

$\alpha(T)$ being the temperature dependence of the recombination of N^{++} to form N^+ , the first term being losses through superelastic collisions of electrons with N^+ ions, and the second term being gains through superelastic collisions of electrons with N^{++} ions. The electron captures take place at a mean energy, $\frac{3}{2}\beta kT$, and the mean excess ionization energy is $\frac{3}{2}k\bar{T}$. (All temperatures in equation [159] are in units of $\chi/k = 21800^\circ \text{K}$, where χ is the excitation energy of the N^+ level.) One finds that

$$T_0 \mathfrak{L}_T = c_1 \rho^2 T^{-1/2} \alpha(T) e^{-1/T} [\frac{3}{2}\beta T - (\frac{1}{2}\bar{T} + \beta) + \bar{T}/T](\bar{T} - \beta T)^{-1} \quad (160)$$

and

$$\rho_0 \mathfrak{L}_\rho = c_1 \rho^2 T^{-1/2} \alpha(T) e^{-1/T}, \quad (161)$$

where the subscript has been omitted from T_0 . One readily shows that condition (158) is satisfied if and only if

$$0.202\beta = (10 - 4\sqrt{6})\beta < \bar{T} < (10 + 4\sqrt{6})\beta = 19.798\beta \quad (162)$$

or

$$0.135 < \bar{T} < 13.199 \quad (163)$$

for Zanstra's assumed value of $\beta (= \frac{2}{3})$. Hence his criterion agrees with ours for all his calculations except those for $\bar{T} = 20$ (432000°K), which are summarized in his Figure 4 (Zanstra, 1955*b*). A curious feature of his results is a region where $(dp/d\rho)_{\mathfrak{R}=0}$ is positive within a larger region where it is negative. According to Zanstra's criterion, a nebula would be stable in this region, extending between $T = 2.4$ and 8.2 . Equation (160) however, shows that condition (158) is violated when T is in the interval

$$\frac{1}{3}[1 + \bar{T}/2\beta - (1 - 5\bar{T}/\beta + \bar{T}^2/4\beta^2)^{1/2}] < T < \frac{1}{3}[1 + \bar{T}/2\beta + (1 - 5\bar{T}/\beta + \bar{T}^2/4\beta^2)^{1/2}] \quad (164)$$

whenever condition (162) is violated, as in the present example. This interval is $2.428 < T < 8.2395$ for $\beta = \frac{2}{3}$ and $\bar{T} = 20$, in agreement with his Figure 4. It follows that a case in this little region, which probably had escaped notice, is actually unstable according to our criterion even though it would have been pronounced stable if one had applied Zanstra's criterion.

From equations (160) and (161), and the correct instability criterion (4b), one finds that a nebula is unstable if

$$\begin{aligned} \frac{1}{5}[1 + 3\bar{T}/2\beta - (1 - 7\bar{T}/\beta + 9\bar{T}^2/4\beta^2)^{1/2}] < T < \frac{1}{5}[1 + 3\bar{T}/2\beta \\ + (1 - 7\bar{T}/\beta + 9\bar{T}^2/4\beta^2)^{1/2}]. \end{aligned} \quad (165)$$

For $\beta = \frac{2}{3}$ and $\bar{T} = 4.5$, for example, the range of T corresponding to instability is 0.7248 – 3.7254 , in good agreement with Zanstra's Figure 2. The range of instability shrinks to zero when $\bar{T} = (14 + \sqrt{160})\beta/9 = 2.9610\beta$ or 1.974 for $\beta = \frac{2}{3}$. This value agrees well with Zanstra's calculated value, 1.968 .

According to Spitzer's (1948) calculations, a more realistic value of β would be 0.55 in the relevant range of temperature. The revised upper limit of $\bar{T}$ for agreement between the two criteria is 10.9 , or 238000°K . The range of T corresponding to instability would be changed to 0.72 – 4.59 for $\bar{T} = 4.5$, and the range of instability would vanish at $\bar{T} = 1.63$.

Daub (1963) pointed out that Zanstra's calculations were unrealistic in taking account of only one source of heat and one cooling agent. In a real nebula, various ions of various elements will be phased in and out as the temperature changes, and it is quite conceivable that the increase in cooling rate as T drops and N^+ grows in abundance in Zanstra's example would be offset by a decrease in the abundance of some ion of higher ionization potential. Daub therefore included all relevant ions in a much more detailed study of the same mechanism. Unfortunately, he adopted Zanstra's instability criterion, and therefore focused attention on the sign of $(dp/d\rho)_{\mathfrak{R}=0}$. As we have seen, this procedure is valid only when $\mathfrak{R}_T > 0$.

Daub found that the inclusion of other ions keeps the nebular temperature almost constant as the density is varied. For example, over the entire range of n_e from 10^2 cm^{-3} to 10^6 cm^{-3} , T varies only between a minimum of $1.2 \times 10^4^\circ \text{K}$ and a maximum of $2.7 \times 10^4^\circ \text{K}$, if the stellar temperature is 138000°K . (The comparable results from Zanstra, with $\bar{T} = 4.5$, corresponding to a stellar temperature of 147000°K , range over a factor of 20 in temperature for a 35-fold variation in density.) As a consequence, Daub's results for p_0 (ρ_0) differ little from isotherms, which are stable according to the Zanstra criterion. Daub therefore concluded that planetary nebulae are stable. We have shown above, however, that even isothermal equilibria are not necessarily stable, so that Daub's results do not in fact imply stability. Unfortunately, Daub does not present

data on the variation of ξ with T for constant ρ , needed to ascertain the sign of ξ_T and hence the correct range of instability. However, he does present a curve showing the rise and fall of ξ as the density is varied. Presumably this behavior reflects the rise and fall in abundance of various ions, and should therefore be present in the variation of ξ with T . If so, it appears possible that nebulae are in fact unstable toward formation of condensations for limited ranges of temperature. It would be very interesting to see Daub's calculations repeated with the correct instability criterion in mind. One might find that the instability sets in at the temperatures and densities corresponding to definite radii in the diverging flow within a planetary nebula, and that this could account for the apparent onset of the condensation phenomenon at a certain radius in NGC 7293 and its subsequent growth at greater radii.

f) Formation of Galaxies

Hoyle (1958) and Gold and Hoyle (1959) proposed that galaxies could be formed by the condensation of a diffuse intergalactic medium, initiated by thermal instability. The necessity for some such method of forming galaxies in a steady-state universe has increased since Harwit's (1961) proof that gravitational condensation is too weak to do so. Gold and Hoyle derived a characteristic cooling time, based on free-free emission, equal to

$$\tau = kT^{1/2}/\mu(\gamma - 1)Fn_e, \quad (166)$$

where $F = 1.4 \times 10^{-27}$ cgs units is the ratio of the emission coefficient to $n_e^2 T^{1/2}$. If we follow the isobaric cooling of an element with no heat input, we find from equation (8) an e -folding time which is γ times larger than equation (166), traceable to the necessity of radiating away the work done in compressing the gas. Our result is $n^{-1} = 4.8 \times 10^{11} T^{1/2}/n_e$ sec, 11.4 times greater than the numerical value of Gold and Hoyle, apparently due to an error of a factor of 7 in their calculation. If, following Gold and Hoyle, one assumes isobaric conditions throughout intergalactic space, with $n_e T = \text{constant} = 10^4 \text{ cm}^{-3} \text{ }^\circ\text{K}$, we find an e -folding time of $n^{-1} = 4.8 \times 10^7 T^{3/2}$ sec. Gold and Hoyle chose the temperature by the requirement that $n^{-1} = \frac{1}{3}H^{-1} = 1.3 \times 10^{17}$ sec, the characteristic time of the expansion of the universe. In this way we obtain $T = 2 \times 10^6 \text{ }^\circ\text{K}$ and $n_e = 5 \times 10^{-3}$. These values were supposed by Gold and Hoyle to apply to "primary condensations," which are formed in some unspecified manner from a general intergalactic medium having $T = 10^9 \text{ }^\circ\text{K}$ and $n_e = 10^{-5} \text{ cm}^{-3}$; the latter corresponds to the mass density $3H^2/8\pi G$.

The model as proposed has now been disproved by recent X-ray observations. Absence of substantial X-ray emission has been shown by Gould and Burbidge (1963) to imply that the temperature of the intergalactic medium is less than $10^7 \text{ }^\circ\text{K}$, and that $n_e T$ is therefore less than 100, rather than the proposed value, 10^4 . This comes at a time when we have shown that the "primary condensations" must be five times denser than proposed by Gold and Hoyle if they are to cool rapidly enough; the new value of the pressure then implies that the temperature of the condensations is less than $9 \times 10^4 \text{ }^\circ\text{K}$. Such low temperatures should make the condensations visible in the optical range and thereby subject the matter to observational test. At these lower temperatures, free-bound cooling would reduce the cooling time.

Turning now to evolutionary models, we may employ the results of § VI. Using the assumption of free-free cooling, following Gold and Hoyle, we find that $c_1 = F/m_H^2 = 5.4 \times 10^{20}$ cgs. Cooling is important before the epoch $t = b_{1/2}$, where

$$b_{1/2} = \mu c_1 \rho_* t_*^3 / 3RA \quad (167)$$

from equation (122); A is the asymptotically constant value of $T_0^{1/2} t$, and hence depends on the unknown value of T_0 at the present time, T_* . We take $\rho_* = 1.7 \times 10^{-29} \text{ gm cm}^{-3}$ and $t_* = 3 \times 10^{17}$ sec as typical values, although they would imply considerable space

curvature according to general relativity, in contradiction with our hypothesis of essentially flat space (see § VIa). With $T_* < 10^7$ °K, $b_{1/2} > 1.7 \times 10^7$ years.

Can galaxies have been formed by thermal instability at such an early phase in the expansion? We have seen that $N = \frac{1}{2}$ (free-free emission) is characterized by relative instability; ρ_1 decreases with time, but more slowly than ρ_0 . Density contrast therefore increases. Thermal conduction stabilizes even relative growth of condensations in the period $t < t_K$, and hence effective relative growth proceeds only in the interval $t_K < t < b_{1/2}$. Now t_K is proportional to $a^{1/4}$ in this situation according to equation (143) and therefore increases as the velocity wavelength, Δv , decreases. The absolute minimum velocity scale which is unstable, $\bar{\Delta v}$, is characterized by $t_K = b_{1/2}$, or $\zeta = 64$, from equation (141), and one finds that

$$T_* = (\rho_* t_*/4)(\mu/6\pi R)^{1/2}(c_1^3/\eta)^{1/4}(\bar{\Delta v})^{1/2} = 4.6(\bar{\Delta v})^{1/2} \text{ °K} . \quad (168)$$

If one identifies the smallest structures owing their existence to thermal instability as clusters of galaxies, following Gold and Hoyle, Δv is the velocity corresponding to 10 Mpc, about 10^8 cm/sec, yielding an estimate, $T_* = 4.6 \times 10^4$ °K, for the present temperature of the intergalactic medium. This estimate is consistent with the X-ray results (Field and Henry 1964) and would account naturally for the absence of observable 21-cm absorption (Field 1962*b*) and emission (Goldstein 1963). This estimate gives $b_{1/2} = 2.5 \times 10^8$ years by equation (167). At this time the density would be 10^{-24} gm cm $^{-3}$ and the temperature, 7.4×10^7 °K. The present density in clusters is of the order of 4×10^{-27} gm cm $^{-3}$, intermediate between the mean density when the instability set in, and that which is now prevalent, as expected from the principle of relative growth.

The question arises as to why larger structures did not grow in the period before $t = b_{1/2}$, as they would be permitted to by equation (143), which states that the minimum velocity scale is given by $\Delta v_{\min} = \text{const.} \times t^{-4}$. The answer may be related to the effects of finite speed of sound, which were not considered in § VI. We know from § II that when $\mathcal{R}_T > 0$, as it is here, isobaric instability depends on the ability of the medium to increase the density in cool regions, and thereby the rate of cooling. Very large regions will be characterized by cooling times shorter than the time for pressure equilibrium to be effected, and hence will have limited growth rates (Table 2). This effect can be estimated from equations (29) and (132) to seriously limit the instability of regions whose wavenumber is less than $1.6 n/c$ where n is the growth rate of the relative density in the present example, namely, $1.8/t$. Since the velocity wavenumber is given by $a = kt$, finite sound-speed effects limit the value of Δv to $\Delta v_{\max} \approx 2c$. In the previous example, $c = 1.5 \times 10^8$ cm/sec at $t = b_{1/2}$, so Δv is confined to the range 10^8 to 3×10^8 cm/sec. This might serve to explain the relative lack of much larger systems originating at $t \approx b_{1/2}$. Furthermore, this result together with equation (121) implies that $\Delta v_{\max}$ is proportional to t^{-2} , while $\Delta v_{\min}$ is proportional to t^{-4} . Hence, at much earlier times, $\Delta v_{\min}$ exceeds $\Delta v_{\max}$, and no scales can be unstable with reasonable growth rates, since they are either too large for pressure equilibrium to be effected, or too small to avoid conductive stabilization.

The conclusion of this subsection is that the masses of clusters of galaxies can be understood if they originated in thermal instability some 2.5×10^8 years after the beginning of the expansion of the universe. If this picture is correct, the present intergalactic temperature is about 50000° K. In the steady-state picture, thermal instability as such can play no role in the formation of such systems from the intergalactic medium, since the present density is much too low for the growth rate to be comparable to $3H$, although if they are compressed by some unspecified mechanism to 500 times the mean density, they can cool isobarically to form more compact systems. In either picture it is assumed that gravitational contraction takes over when the radius of the system approaches GM/c^2 . A shortcoming that permeates our discussion of the formation of galaxies is the

adoption of an expanding universe in which $v_0 \rightarrow \infty$ and relativistic effects are ignored. It may be hoped, however, that the results concerning local effects may be approximately valid as long as Δv is much less than the velocity of light.

This work was carried out largely during three periods with the Summer Study Program in Geophysical Fluid Dynamics at the Woods Hole Oceanographic Institution. The author is indebted to the Program for support in 1961. Many stimulating discussions with Program participants, particularly Drs. E. A. Spiegel, G. Veronis, W. V. R. Malkus, D. W. Moore, D. E. Osterbrock, and L. Mestel, played a major role in the work reported here. Dr. L. Spitzer, Jr., kindly read the manuscript and offered several suggestions for its improvement. The work was partially supported by the National Aeronautics and Space Administration under research grant NsG-414.

REFERENCES

- Athay, R. G., and Thomas, R. N. 1956, *Ap. J.*, **123**, 299.
 Bondi, H. 1960, *Cosmology* (Cambridge: Cambridge University Press), eqs. (9.5) and (9.14).
 Danielson, R. E. 1964, private communication.
 Daub, C. T. 1963, *Ap. J.*, **137**, 184.
 Field, G. B. 1962a, *Interstellar Matter in Galaxies*, ed. L. Woltjer (New York: W. A. Benjamin, Inc.), p. 183.
 ———. 1962b, *Ap. J.*, **135**, 684.
 Field, G. B., and Henry, R. C. 1964, *Ap. J.*, **140**, 1002.
 Gold, T., and Hoyle, F. 1959, *Paris Symposium on Radio Astronomy*, ed. R. N. Bracewell (Stanford, Calif.: Stanford University Press), paper 104.
 Goldstein, S. J., Jr. 1963, *Ap. J.*, **138**, 978.
 Gould, R. J., and Burbidge, G. R. 1963, *Ap. J.*, **138**, 969.
 Harwit, M. 1961, *M.N.*, **122**, 47.
 Hoyle, F. 1958, *Proceedings of the 11th Solvay Congress* (Brussels: Stoops).
 Ireland, J. G. 1961, *M.N.*, **122**, 461.
 Jahnke, E., and Emde, F. 1945, *Tables of Functions* (4th ed.; New York: Dover Publications).
 Kiepenheuer, K. O. 1953, *The Sun*, ed. G. P. Kuiper (Chicago: University of Chicago Press), chap. vi, p. 430.
 ———. 1959, *Zs. f. Ap.*, **48**, 290.
 Kleczek, J. 1957, *Bull. Astr. Inst. Czechoslovakia*, **8**, 120.
 ———. 1958, *ibid.*, **9**, 115.
 Kopp, R. A. 1963, *Thermal Instabilities in the Transition Region of the Solar Atmosphere* (Cambridge, Mass.: Harvard College Observatory).
 Lüst, R., and Zirin, H. 1960, *Zs. f. Ap.*, **49**, 8.
 Mestel, L. 1963, *M.N.*, **126**, 553.
 Oort, J. H. 1962, *Interstellar Matter in Galaxies*, ed. L. Woltjer (New York: W. A. Benjamin, Inc.), p. 3.
 Orrall, F. Q., and Zirker, J. B. 1961, *Ap. J.*, **134**, 72.
 Parker, E. N. 1953, *Ap. J.*, **117**, 431.
 ———. 1963, *Interplanetary Dynamical Processes* (New York: Interscience Publishers), p. 15.
 Rayleigh, Lord, 1945, *The Theory of Sound* (New York: Dover Publications), Section 247.
 Spitzer, L., Jr., 1948, *Ap. J.*, **107**, 6.
 ———. 1949, *ibid.*, **109**, 337.
 ———. 1956, *ibid.*, **124**, 20.
 ———. 1962, *Physics of Fully Ionized Gases* (2d ed.; New York: Interscience Publishers).
 Spitzer, L., Jr., and Savedoff, M. P. 1950, *Ap. J.*, **111**, 593.
 Uchida, Y. 1963, *Pub. Astr. Soc. Japan*, **15**, 65.
 Weymann, R. 1960a, *Ap. J.*, **132**, 380.
 ———. 1960b, *ibid.*, p. 452.
 Woltjer, L. 1962a, *Interstellar Matter in Galaxies*, ed. L. Woltjer (New York: W. A. Benjamin, Inc.), p. 88.
 ———. 1962b, *ibid.*, p. 247.
 Zanstra, H. 1955a, *Gas Dynamics of Cosmic Clouds*, ed. J. M. Burgers and H. C. van de Hulst (Amsterdam: North-Holland Publishing Co.), chap. xiii.
 ———. 1955b, *Vistas in Astronomy*, **1**, ed. A. Beer (New York: Pergamon Press), p. 256.
 Zirin, H. 1962, *Soviet Astr.*, **5**, 660.