THERMAL INSTABILITY AND THE FORMATION OF CLUMPY GAS CLOUDS

A. Burkert\textsuperscript{1} and D. N. C. Lin\textsuperscript{2}

\textit{Received 1998 October 22; accepted 2000 January 28}

Abstract

The radiative cooling of optically thin gaseous regions and the formation of a two-phase medium and of cold gas clouds with a clumpy substructure is investigated. We demonstrate how clumpiness can emerge as a result of thermal instability. In optically thin clouds, the growth rate of small density perturbations is independent of their length scale as long as the perturbations can adjust to an isobaric state. However, the growth of a perturbation is limited by its transition from isobaric to isochoric cooling when the cooling timescale is reduced below the sound crossing timescale across its length scale. The temperature at which this transition occurs decreases with the length scale of the perturbation. Consequently, small-scale perturbations have the potential to reach higher amplitudes than large-scale perturbations. When the amplitude becomes nonlinear, advection overtakes the pressure gradient in promoting the compression, resulting in an accelerated growth of the disturbance. The critical temperature for transition depends on the initial amplitude. The fluctuations that can first reach nonlinearity before their isobaric to isochoric transition will determine the characteristic size and mass of the cold dense clumps that would emerge from the cooling of an initially nearly homogeneous region of gas. Thermal conduction is, in general, very efficient in erasing isobaric, small-scale fluctuations, thus suppressing a cooling instability. A weak, tangled magnetic field, however, can reduce the conductive heat flux enough for low-amplitude fluctuations to grow isobarically and become nonlinear if their length scales are of order $10^{-2}$ pc. If the amplitude of the initial perturbations is a decreasing function of the wavelength, the size of the emerging clumps will decrease with increasing magnetic field strength. Finally, we demonstrate how a two-phase medium, with cold clumps being pressure confined in a diffuse hot residual background component, would be sustained if there is adequate heating to compensate the energy loss.

\textit{Subject headings:} globular clusters: general — instabilities — ISM: clouds

1. Introduction

The interstellar medium and cold gas clouds are characterized by a clumpy substructure and a turbulent velocity field (Larson 1981; Blitz 1993). As molecular clouds are the sites of star formation, their formation, internal structure, and dynamics determine the rate of star formation and the properties of young stars, such as their mass function or binarity. The understanding of the origin of cold clouds and their internal substructure is therefore of fundamental importance for a consistent theory of star formation and galactic evolution.

In the nearby clouds, the dispersion velocity inferred from molecular line width is often larger than the gas sound speed inferred from the line transition temperatures (Solomon et al. 1987). MHD turbulence may be responsible for the stirring of these clouds (Arons & Max 1975). This conjecture is supported by the polarization maps and direct measurements of field strength in some star-forming regions (Myers & Goodman 1988; Crutcher et al. 1993). Recent simulations of MHD turbulence, however, suggest that it dissipates rapidly (Gammie & Ostriker 1996; MacLow et al. 1998; MacLow 1999; Ostriker, Gammie, & Stone 1999). One possible source of energy supply is winds and outflows from young stellar objects (Franco & Cox 1983; McKee 1989). But in regions where star formation is inactive, clumpy structure with velocity dispersion is also observed. Thus, the origin and energy supply of clumpy cloud structure remains an outstanding issue.

2 Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany; burkert@mperia-hd.mpg.de.
3 UCO/Lick Observatory, University of California, Santa Cruz, CA 95064; lin@lick.ucolick.org.

On small scales, magnetic field pressure is important in regulating infall and collapse of protostellar clouds and the formation of low-mass stars (Mouschovias & Spitzer 1976; Nakano 1979; Shu et al. 1993). For clouds with subcritical masses, gravitational contraction is proceeded by ambipolar diffusion, which for typical cloud densities operates on a timescale $\tau_B \sim 10^{7-8}$ yr (Lizano & Shu 1989; Mouschovias 1991). In regions with intense star formation activities, such as the central region of Orion, $\tau_B$ for individual dense clumps is comparable to the typical age of the young stellar objects. But, the spread in stellar ages ($\Delta \tau_s \sim 10^6$ yr) appears to be considerably shorter than $\tau_B$ (Carpenter et al. 1997; Hillenbrand 1997). This coeval star formation history requires either a coordinated trigger mechanism for star formation within initially magnetically supported clumps or subcritical collapse, fragmentation, and star formation of a larger molecular cloud region in which the magnetic field plays a weak role.

A rapid and coordinated episode of star formation can also be inferred in globular clusters (Brown, Burkert, & Truran 1991, 1995; Murray & Lin 1992; Lin & Murray 1992). In some metal-deficient clusters, such as M92, the total amount of heavy elements corresponds to the yield of a few supernovae. If star formation has proceeded over a duration $\Delta \tau_s$ comparable to the expected life span ($\sim$ a few $10^6$ yr) of massive stars, a significant metallicity spread would be expected, in contrast to the observations (e.g., Kraft 1979). At least in these systems, $\Delta \tau_s < \tau_B$ and star formation may have proceeded through supercritical collapse. The dynamical timescale of most clusters at their half-mass radius is $\tau_s \approx 10^6$ yr. Any energy dissipation associated with the episode of star formation would imply an even longer dynamical timescale in the protocluster.
Although thermal instability proceeds faster than the collapse of the cloud, its growth rate is determined by the local cooling rate. During the initial linear evolution, variations in the initial overdensity (or undertemperature) might lead only to a weak dependence of the growth timescale on the wavelength. This paper shows, however, that there exist two important transitions that are very sensitively determined by the wavelengths of perturbations. (1) The growth of a perturbation is limited by its transition from isobaric to isochoric cooling, when the cooling timescale is reduced below the sound-crossing timescale across the wavelength of the perturbation. This transition occurs at a lower temperature, with correspondingly larger overdensity, for perturbations with smaller wavelengths. (2) For those perturbations that can become nonlinear before the isobaric to isochoric transition, advection overtakes the pressure gradient in promoting the compression and growth of the perturbed region at an accelerated rate. The fluctuations that can first reach nonlinearity would dominate the growth of all perturbations with longer wavelengths and homogenize disturbances with smaller wavelengths. Thus, they determine the characteristic size and mass of the cold dense clumps that would emerge from the cooling of an initially nearly homogeneous cloud. Thermal conduction could, in general, erase these fluctuations, thus suppressing the instability. Weak, tangled magnetic fields, however, would be efficient enough in reducing the conductive flux, thus allowing the medium to break up into cold clumps on the characteristic length scale.

We study the cooling and fragmentation of gas using simplified power-law cooling functions. Since we are primarily interested in supernovae clouds, we neglect the effect of magnetic fields. Note that even a weak magnetic field could have an important destabilizing influence in thermal instability (Loewenstein 1990; Balbus 1995). In § 2, we obtain approximate analytic solutions that describe the evolution of a linear density perturbation in the isobaric and nearly isochoric regime. We show that the growth of overdensity in a thermally unstable fluctuation is limited by a transition from isobaric to isochoric evolution and that the limiting amplitude is a decreasing function of the length scale. We verify our analytic approximations with numerical, hydrodynamical calculations that are also used in § 3 to study the transition into the nonlinear regime. In § 4 we investigate the cooling of interacting perturbations and determine the critical length scale of clumps that emerge through thermal instability. The importance of thermal conduction is investigated in § 5. In § 6 we discuss the affect of heating processes and the formation of a stable two-phase medium. Finally, we summarize our results and discuss their implications in § 7.

2. THE INITIAL EVOLUTION OF THERMAL INSTABILITY

The dynamical evolution of the gas is described by the hydrodynamical equations

\[
\frac{\partial \rho}{\partial t} + \sum_{k=1}^{3} \frac{\partial \rho U_k}{\partial x_k} = 0 ,
\]

\[
\frac{\partial U_j}{\partial t} + \sum_{k=1}^{3} U_k \frac{\partial U_j}{\partial x_k} + R_g \frac{\partial}{\partial \rho} \sum_{k=1}^{3} U_k U_k \frac{\partial T}{\partial x_k} = 0 ,
\]

\[
\frac{\partial T}{\partial t} + \sum_{k=1}^{3} U_k \frac{\partial T}{\partial x_k} + (\Gamma - 1) T \frac{\partial}{\partial \rho} \sum_{k=1}^{3} U_k U_k \frac{\partial T}{\partial x_k} = -\frac{\rho A}{C_e} ,
\]
where \( j = 1, 2, 3 \) is the coordinate index, \( C_v = R_g/\mu (\Gamma - 1) \) is the heat capacity, \( R_g, \mu, \) and \( \Gamma \) are the gas constant, mean molecular weight, and adiabatic index, respectively.

In the unperturbed state, the gas remains at rest \((U_j = 0)\) and its density attains a constant value, \( \rho_0 \). The time-dependent energy equation (3) gives

\[
C_v \frac{\partial T_0}{\partial t} = - \rho_0 \Lambda ,
\]

where the cooling rate \( \Lambda = \Lambda_0 T_b^6 \). The power index is determined by the detailed atomic processes. Since we are interested primarily in the physical evolution of thermal instability, we adopt a simple constant \( \beta \) prescription. The cooling would be thermally unstable (with \( \tau_c = T_0/\xi T_0/\partial t \) as an increasing function of \( T_0 \)) in the isochoric region if \( \beta < 1 \) and in the isobaric region if \( \beta < 2 \). In the absence of external heating, the unperturbed gas temperature \( T_0 \) can be expressed as a function of the dimensionless time variable \( \tau \equiv t/\tau_c(0) \), such that

\[
T_0(t) = T_0(0)[1 - (1 - \beta)\tau]^{(1/1 - \beta)} ,
\]

where \( T_0(0) \) and \( \tau_c(0) \equiv C_v/\rho_0 \Lambda_0 T_0(0)^{6 - 1} \) are the initial (at \( t = 0 \)) temperature and cooling timescale, respectively.

### 2.1. The Perturbed Quantities

The evolution of the perturbed density \( (\rho_1 = \rho - \rho_0) \), temperature \( (T_1 = T - T_0) \), and velocities \( (U_j) \) are derived from the linearization of equations (1), (2), and (3):

\[
\frac{\partial}{\partial t} \frac{\partial U_j}{\partial \rho} = - \frac{3}{\partial j} \frac{\partial U_j}{\partial \xi_j} ,
\]

\[
\frac{\partial U_j}{\partial t} = - \frac{R_g T_0}{\mu} \frac{\partial}{\partial \xi_j} \left[ \frac{T_1}{T_0} + \frac{\rho_1}{\rho_0} \right] ,
\]



\[
\frac{\partial T_1}{\partial t} = \frac{\partial}{\partial \xi_j} \left[ \frac{3}{\partial j} \frac{T_1}{T_0} \right] \tau_c \left[ \frac{\rho_1}{\rho_0} + (\beta - 1) \frac{T_1}{T_0} \right] .
\]

where

\[
\tau_c(t) = \frac{T_0}{d T_0/d t} = \tau_c(0) - (1 - \beta)\mu
\]

is the characteristic cooling timescale at the instant of time \( t \).

Since the perturbation equations are linear in \( x_j \), we adopt a local approximation in which the positional dependence of all the perturbed quantities is proportional to \( \exp(i k_j x_j) \) where \( k_j \) is the wave number in the \( j \)th direction. Substituting a dimensionless velocity variable \( V_j = i k_j \tau_c(0) U_j \), the perturbed equations reduce to

\[
\frac{\partial}{\partial t} \frac{\partial V_j}{\partial \rho} = - \frac{3}{\partial j} \frac{\partial V_j}{\partial \xi_j} ,
\]

\[
\frac{\partial V_j}{\partial t} = K_j [1 - (1 - \beta)\tau]^{(1/1 - \beta)} \left( \frac{P_j}{P_0} \right) ,
\]

where \( P_j/P_0 = T_j/T_0 + \rho_j/\rho_0 \) is the perturbed pressure, \( K_j \equiv \tau_c(0) k_j R_g T_0(0)/\mu \) is the ratio of the initial cooling to sound-crossing timescale over a characteristic wavelength

\[
2\pi/k_j ,
\]

and

\[
\frac{\partial}{\partial t} \frac{T_1}{T_0} = - (1 - \beta) \sum_{j=1}^{3} V_j - \frac{\tau_c(0)}{\tau_c} \left[ \frac{\rho_1}{\rho_0} + (\beta - 1) \frac{T_1}{T_0} \right] .
\]

For a perfect gas, the unperturbed pressure \( P_0 = R_g \rho_0 T_0/\mu \) decreases at the same rate everywhere. We find from equations (9) and (11) that the amplitude of the perturbed pressure is

\[
\frac{\partial}{\partial t} \frac{P_1}{P_0} = \frac{\partial}{\partial t} \left( \frac{\rho_1}{\rho_0} + \frac{T_1}{T_0} \right) = - \beta \sum_{j=1}^{3} V_j \left[ 1 - (1 - \beta)\tau \right]
\]

\[
\times \left[ (2 - \beta) \frac{\rho_1}{\rho_0} + (1 - \beta) \frac{P_1}{P_0} \right].
\]

### 2.2. The Initially Isochoric Regime: \( K \leq 1 \)

For computational simplicity, we now consider a one-dimensional limit treatment in which the initial (at \( \tau = 0 \)) amplitude of \( \rho_1 \) equals to a finite value \( \rho_a \) with that of \( V_1 \) and \( P_1 \) equal to zero. These conditions correspond to an initially almost homogeneous, hot region of gas in pressure equilibrium. To third order in \( \tau \), equations (9), (10), and (12) give the following approximate solution

\[
\frac{\rho_1}{\rho_0} \approx \frac{\rho_a}{\rho_0} \left[ 1 + \frac{K^2}{6} (2 - \beta)\tau^3 \right] ,
\]

\[
V \approx \frac{(1 - 2 K^2)}{6} \frac{\rho_a}{\rho_0} (3 \tau^2 + 2 \beta \tau^3) ,
\]

\[
P \approx (\beta - 2) \frac{\rho_a}{\rho_0}
\]

\[
\times \left[ \tau^2 + (1 - \beta)\tau^3 + \left( 4 - \frac{7}{3} \beta + \beta^2 - \frac{\Gamma K^2}{6} \right) \tau^3 \right] .
\]

Figure 1 compares this solution with a numerical integration of the complete nonlinear hydrodynamical equations (1), (2), and (3) for \( K = 1 \) and \( K = 0.5 \). We use a one-dimensional version of the second-order Eulerian hydro code, which is described in Burkert & Bodenheimer (1993). The agreement between the numerical results (solid lines) and analytical solution (dots) is excellent, even for large values of \( \tau \approx 1 \), where the basis of the analytic approximation is no longer valid.

Because of slightly more efficient cooling within the density perturbation, a small pressure gradient builds up. In an attempt to maintain pressure balance, the slightly warmer gas in the low-density regions continually compresses the more dense and cooler parts. Consequently, the overdensity in the perturbed region increases as the gas cools. Figure 1 and equations (13), (14), and (15) show that for small values of \( K \leq 1 \), the growth rate of the density enhancement depends on the size of the perturbation and increases with increasing values of \( K \) or decreasing wavelength. However, because of its long sound-crossing timescale, the perturbation cannot be compressed significantly while cooling; it cools almost isochorically (Parker 1953). After one cooling timescale, the gas temperature has
reached its minimum value with the density enhancement still in the linear regime. Now, the pressure gradient reverses, erasing the fluctuation.

2.3. The Initially Isobaric Regime: $K \geq 1$

The solid lines in Figure 2 show a numerical calculation of the evolution of a density perturbation with $K = 200$, $\beta = 0$, and $\Gamma = 5/3$. For $K \gg 1$, perturbations can react quickly on any pressure gradients because of the short sound-crossing timescale, relative to the cooling timescale. The simulations indicate a solution that consists of a fast oscillatory part and a slowly growing part. Linearizing the slowly growing part, we find from equations (9), (10), (11), and (12) the following approximate solution:

$$\frac{P_1}{P_0} \approx \frac{\rho_a}{\rho_0} \left( \frac{i \omega}{i \omega \Gamma - 2 + \beta} \right) \left[ 1 + \frac{2 - \beta}{\Gamma} \left( \tau + i \omega e^{i \omega \tau} \right) \right],$$  

(16)

$$V \approx \frac{2 - \beta}{\Gamma} \frac{\rho_a}{\rho_0} \left( \frac{i \omega}{i \omega \Gamma - 2 + \beta} \right) \left( 1 + i \omega \tau - e^{i \omega \tau} \right),$$  

(17)

and

$$\frac{P_1}{P_0} \approx \frac{2 - \beta}{\Gamma} \frac{\rho_a}{\rho_0} \left( \frac{i \omega}{i \omega \Gamma - 2 + \beta} \right) \times \left[ 1 + (1 - \beta) \tau - e^{i \omega \tau} \right].$$  

(18)

The characteristic frequency is determined by a cubic dispersion relation

$$\omega^3 + i(1 - \beta) \omega^2 - K^2 \omega - i(2 - \beta) K^2 = 0.$$  

(19)

A similar relation was discussed by Balbus (1995). Note that the linearized solution is also valid for $\tau > 1/\omega$ as long as $\tau \ll 1$. As $K \gg 1$, the two dominant real roots ($\omega \approx \pm \sqrt{\Gamma K}$) of the dispersion equation (19) yield oscillatory parts in $\rho_1/\rho_0$, $V$, and $P_1/P_0$. The upper panels of Figure 2 show that for $\tau < 0.1$ the analytical approximation is in good agreement with the numerical integration of the non-linear hydrodynamical equations.

For all length scales, the ratio of sound propagation to cooling timescale

$$Q \equiv \tau_s(t) \sqrt{R_g T_0(t)/\mu} = K[1 - (1 - \beta) \tau]^{(3 - 2\beta)/2 - 2\beta}$$  

(20)

decreases during the subsequent evolution. Provided $Q \gg 1$, equation (18) implies that the magnitude of $P_1/P_0$ is much smaller than both $V$ and $\rho_1/\rho_0$. That is, the fluctuation reacts isobaric. Adopting $P_1/P_0 \approx 0$ and neglecting the oscillatory term, the equations (9) and (12) can be combined to

$$-V = \frac{\partial}{\partial \tau} \left( \frac{\rho_1}{\rho_0} \right) = \frac{(2 - \beta)}{\Gamma[1 - (1 - \beta) \tau]} \frac{\rho_1}{\rho_0},$$  

(21)

with the solution

$$\frac{\rho_1}{\rho_0} = \rho_0 \left[ 1 - (1 - \beta) \tau \right]^{-(2 - \beta)(1 - \beta) \Gamma}$$  

(22)

and

$$V = -\frac{(2 - \beta)}{\Gamma} \frac{\rho_a}{\rho_0} \left[ 1 - (1 - \beta) \tau \right]^{-2(1 - \beta) \Gamma}.$$  

(23)

In contrast to fluctuations with $K < 1$, the evolution of isobaric fluctuations is independent of $K$. For $(1 - \beta)^{-1} \gg \tau > \omega^{-1}$, solutions for $\rho_1/\rho_0$ in equations (16) and (22) are in agreement to first order in $\tau$. The lower panels of Figure 2 show that within a cooling time both $\rho_1/\rho_0$ and $V$ are amplified to very large values. The agreement between the numerical calculation and the analytical prediction (eqs. [22] and [23]) is excellent. The opposite signs of $\rho_1/\rho_0$ and $V$ confirm that mass is being pushed into the cool dense regions. For $\tau \approx 1$, the numerically derived density enhancement falls below the predicted values as the
fluctuation becomes isochoric and contributions from the perturbed pressure cannot be neglected anymore.

2.4. Transition to Isochoric Evolution and the Emergence of Small-Scale Perturbations

Figure 3 shows the density evolution of initially isobaric fluctuations with different ratios of cooling to sound-crossing times $K$ as determined from the numerical calculations. The initially isobaric growth of the density fluctuations is independent of wavelength and $K$ and in excellent agreement with equation (22) (dashed curve). The perturbations transform, however, to the isochoric solution for the epoch after $Q$ has declined below unity. Thereafter, gas in the perturbed region cools off faster than it can adjust to a pressure equilibrium with the surrounding region. Subsequently, the overdensity of the perturbed region is slowly modified by the inertial motion $V_{\text{trans}}$ of the gas at the time of the transition and its growth stalls. In Figure 3, the transition into the isochoric regime is indicated by the overdensity falling below the expected value shown by the dashed thick line.

Although the growth of the perturbed quantities does not explicitly depend on the wavelength $k$, the critical transition time when $Q \approx 1$

$$\tau_{\text{trans}} = \frac{1 - K^{2\beta - 2/3 - 2\beta}}{1 - \beta}$$

is a function of $K$ (and $k$). At this transition point, the overdensity in the perturbed region is

$$\frac{\rho_{\text{trans}}}{\rho_0} = \frac{\rho_2}{\rho_0} K^{(4 - 2\beta)/3 - 2\beta(1 - \beta)(1 - \beta)}$$

and the velocity is

$$V_{\text{trans}} = -\frac{2 - \beta}{\Gamma} \frac{\rho_2}{\rho_0} K^{(2 - 2\beta)/3 - 2\beta(1 + 2 - \beta)(1 - \beta)}.$$  

In the isochoric regime, the amplitude of the perturbed density increases as

$$\frac{\rho_1}{\rho_0} = \frac{\rho_{\text{trans}}}{\rho_0} - V_{\text{trans}} \ast (\tau - \tau_{\text{trans}}),$$
the evolution numerically, solving the complete nonlinear hydrodynamical equations. The simulations shown in Figure 3 assumed $\beta = 0$, $\Gamma = 5/3$, and $\rho_d/\rho_0 = 10^{-3}$. For these values, the simple approximation in equation (28) predicts $K_{\text{crit}} = 5600$, which is roughly in agreement with the numerical results where the transition into the nonlinear regime occurs more smoothly between $K = 1000$ and $K = 5000$. Equation (28) somewhat overestimates $K_{\text{crit}}$ because nonlinear effects actually become important earlier, when the overdensity is in the range $0.1 < \rho_1/\rho_0 < 1$.

Figure 4 shows the structure and evolution of a nonlinear fluctuation. During the early isobaric evolution, the pressure gradient (lower right panel) is negligible. A small pressure gradient builds up in the nonlinear regime, where the profiles cannot be approximated anymore by sinusoidal functions but instead become strongly peaked toward the center. Nonlinear fluctuations grow fast, with the density and temperature reaching their maximum and minimum values, respectively, at a time $\tau_{\text{crit}} < 1$, which is shorter than a cooling time. Because of the fast growth in the nonlinear regime, $\tau_{\text{crit}}$ is roughly given by the time when $\rho_1/\rho_0 = 1$. From equation (22), we find

$$\tau_{\text{crit}} = \frac{1}{1 - \beta} \left[ 1 - \left( \frac{\rho_a}{\rho_0} \right)^{1/(1 - \beta) / (2 - \beta)} \right].$$  \hspace{1cm} \text{(29)}

At $t = \tau_{\text{crit}}$, the dimensionless velocity (see eq. [23])

$$V_{\text{crit}} = V(\tau_{\text{crit}}) = \frac{\beta - 2}{\Gamma} \left( \frac{\rho_a}{\rho_0} \right)^{-1/(1 - \beta / 2 - \beta)}$$  \hspace{1cm} \text{(30)}

is much larger than unity for perturbations with small initial amplitudes, such that contributions caused by nonlinear advection (such as $U_2 \partial \rho / \partial x_0$, $U_2 \partial U_1 / \partial x_0$, and $U_2 \partial T / \partial x_0$) would exceed the linear contributions contained in the perturbed equations (6), (7), and (8) before the overdensity $\rho_1$ has become comparable to $\rho_0$ (see above). Advection generally enhances the effect of compression and promotes the growth of density contrast at an accelerated rate.

Although the time of maximum compression for a fluctuation with $K > K_{\text{crit}}$ does not depend explicitly on the length scale, it is determined by the initial amplitude $\rho_d/\rho_0$ of the perturbation, which may be a function of the wavelength. For an initial power-law perturbation, in which $\rho_d/\rho_0 = A_0(k/k_0)\eta$,

$$\tau_{\text{crit}} = \frac{k}{k_0} \left[ 1 - \left( \frac{A_0}{k_0} \right)^{1/(1 - \beta \epsilon / 2 - \beta)} \right].$$  \hspace{1cm} \text{(31)}

If the amplitude of the initial perturbation is an increasing function of the wavelength (which corresponds to a negative $\eta$), $\tau_{\text{crit}}$ would be an increasing function of $k$ in the thermally unstable region with $\beta < 1$. In this case, nonlinearity would be first reached on the largest length scale with $K > K_{\text{crit}}$. If, however, the amplitude of the initial perturbation is a decreasing function of the wavelength (i.e., $\eta > 0$), $\tau_{\text{crit}}$ would be a decreasing function of $k$ and nonlinearity would be reached on the smallest scale first. Note that for $1 < \beta < 2$, the dependence of $\tau_{\text{crit}}$ on $\eta$ and $k$ is reversed.

In Figure 3, a temperature-independent cooling function has been used. To determine the dependence on the specific form of the cooling function, additional simulations have been performed, adopting a more realistic cooling function (Dalgarno & McCray 1972) that assumes solar element abundance and collisional equilibrium ionization. Note

which increases much less steeply than the isobaric fluctuations (eq. [22]).

For thermally unstable clouds, $\beta < 1$ such that $\rho_{\text{trans}}/\rho_0$ is an increasing function of $K$ or a decreasing function of the wavelength ($\lambda$) of the perturbations. Despite the independence of the rate of change of $\rho_1/\rho_0$ on $\lambda$, equation (24) shows that for $\beta < 1$ the short length scale disturbances undergo isobaric to isochoric transition at a later time and therefore acquire a greater limiting amplitude than the long length scale disturbances. Thus, the short length scale disturbances would emerge to dominate the structure of the cloud unless the initial perturbation amplitude $\rho_d/\rho_0$ increases with $K$ more rapidly than $K^{(2\beta - 3)/(3 - 2\beta)}$. This evolution is physically equivalent to the fragmentation process in which the contrast between the enhanced density in a disturbance and the average cloud density becomes most pronounced on the smallest scales.

### 3. THE TRANSITION INTO THE NONLINEAR REGIME

In Figure 3 fluctuations with very large values of $K$ show yet another evolution: for later times, the overdensity rises faster than predicted by equation (22). These fluctuations become nonlinear with $\rho_1/\rho_0 > 1$ before the transition into the isochoric regime. The critical value of $K$ for this evolution can be estimated from equation (25) assuming $\rho_{\text{trans}}/\rho_0 = 1$:

$$K_{\text{crit}} = \left( \frac{\rho_a}{\rho_0} \right)^{(2\beta - 3)/(4 - 2\beta)}. \hspace{1cm} \text{(28)}$$

For $K > K_{\text{crit}}$, the analytical approximations discussed previously are not valid anymore and we have to investigate

![Figure 3](image-url)
that for temperatures \( T > 10^4 \, \text{K} \), the cooling rate is several orders of magnitudes larger than for \( T < 10^4 \, \text{K} \), defining two different temperature regimes with very different cooling timescales. The simulations show that the previous results remain valid for each of these temperature regimes. Starting in the low-temperature regime, a fluctuation will become nonlinear for \( K > K_{\text{crit}} \) and cool down to the lowest allowed temperature. The same is true for fluctuations that start in the high-temperature regime. Nonlinear fluctuations in this regime, however, do stop cooling efficiently at \( T \approx 10^4 \, \text{K} \), leading to high-density clumps with such a temperature.

4. INTERACTING FLUCTUATIONS AND THE EMERGENCE OF SUBSTRUCTURE WITH A CRITICAL WAVELENGTH

Up to now, we have investigated the evolution of isolated fluctuations. In reality, however, a cooling gaseous region consists of a superposition of fluctuations with different wavelengths and amplitudes. As indicated in § 3, the outcome of the thermal instability may be determined by the wavelength dependence in the initial amplitude of the perturbations.

To illustrate various competing effects, such as isobaric to isochoric transition and the onset of nonlinear growth, we present in Figure 5 a series of models with \( \beta = 0 \) and \( \Gamma = 5/3 \), where the initial density distribution consists of the superposition of two fluctuations with ratios of wavelengths \( \lambda_1/\lambda_2 = 20 \) and amplitude ratios \( \rho_{a,1}/\rho_{a,2} = 2 \), which corresponds to \( \eta = -0.23 \). Four values of \( \lambda_1 \) were chosen, and they correspond to \( K_1 = 1, 10, 100, \) and \( 1000 \), respectively. The \( K_2 \) values for the smaller perturbation are always a factor of 20 larger. Since the initial overdensity \( \rho_{a,1}/\rho_0 = 0.01 \), the critical value of \( K \) for nonlinear evolution is \( K_{\text{crit}} = 316 \), according to equation (28). We show the density distribution after \( 1 \, \tau_c(0) \). In the upper left-hand panels of Figure 5, the fluctuations have values of \( K_1 = 1 \) and \( K_2 = 20 \), which are small compared to \( K_{\text{crit}} \). Their growth therefore stalls because of the transition into the isochoric regime and the overdensity after a cooling time is still linear. The smaller fluctuation dominates at the end because its isochoric transition occurs later and at a higher overdensity than for the larger perturbation. In the upper right panels, with \( K_1 = 10 \) and \( K_2 = 200 \), the smaller perturbation is again dominating after \( \tau = 1 \); although, now,
the density distribution is also affected by the underlying larger perturbation. In both cases, the density within the density peaks does not decrease much with respect to its initial value. The situation is different in the lower left panel with $K_1 = 100$ and $K_2 = 2000$. Here the smaller perturbation has become nonlinear, generating small dense clumps that stand out against the larger perturbation. Up to now, the smaller perturbation was always dominating the density distribution after a cooling time. The situation, however, is different in the lower right-hand panel, where $K_1 > K_{\text{crit}}$. Now, the larger perturbation becomes nonlinear and advection drives all the gas and its small-scale fluctuations into one very dense, cold clump that is embedded in a hot diffuse environment, thus erasing smaller scale fluctuations.

The dependence of structure formation on the initial power-law perturbation index $\eta$ is illustrated in Figure 6, which shows the initial and final density distribution of two interacting perturbations with $\lambda_1/\lambda_2 = 10$, $K_1 = 2 \times 10^4$, $\rho_{a,1}/\rho_{a,2} = 10$ ($\eta = -1$) in the left-hand panels and $\rho_{a,1}/\rho_{a,2} = 0.1$ ($\eta = 1$) in the right-hand panels. As expected, in the case of $\eta = -1$, the larger perturbation becomes nonlinear first, leading to one massive density peak after a cooling time. For $\eta = 1$, the small length scale perturbations begin to dominate after a cooling time, breaking the region up into dense clumps on the smallest scale. More specifically, if the amplitude of the initial perturbation increases with increasing wavelength ($\eta < 0$), clumps will form with length scales $\lambda \approx \lambda_{\text{crit}}$. Otherwise, the sizes of the fastest growing perturbations will be $\lambda \approx \lambda_{\text{crit}}$. In this case, the clump sizes should decrease with increasing magnetic field strength.

5. THE IMPORTANCE OF THERMAL CONDUCTION

During the growth of linear density perturbations in the isobaric regime, the resulting temperature gradient will induce conductive heating of the fluctuations.

5.1. Thermal Conduction in the Absence of Magnetic Fields

Several studies (e.g., McKee & Begelman 1990; Ferrara & Shchekinov 1993) have demonstrated that thermal conduction could stabilize and even erase a density perturbation if its scale is smaller than the Field length (Field 1965).
Figure 6.—The left-hand upper and lower panels show, respectively, the initial ($\tau = 0$) and final ($\tau = 1$) density distribution of two nonlinear interacting density perturbations with $\lambda_1/\lambda_0 = 10$ and $\eta = -1$. The initial and final density distributions for interacting nonlinear fluctuations with $\eta = 1$ are shown, respectively, in the upper and lower right-hand panels.

In the early stages of cooling ($\tau \ll 1$), fluctuations will therefore be erased by thermal conduction if their wavelengths are

$$\lambda < \lambda_\kappa = \frac{2\pi}{(2 - \beta)^{1/2}} \frac{\lambda_T}{\lambda_F}.$$  

Figure 7 shows the evolution of an initially isobaric, one-dimensional fluctuation with a ratio of cooling–to–sound-crossing time $\kappa = 200$ and $\beta = 0$. The one-dimensional, nonlinear hydrodynamical equations (1), (2), and (3) are solved numerically, including thermal conduction. The solid line shows the evolution of the density contrast $\rho_1/\rho_0$ as predicted by the analytical model (eq. [22]), which is in excellent agreement with the numerical result (filled points) for $\lambda > \lambda_\kappa$. For $\lambda = \lambda_\kappa$ (upper dashed line), conductive heating cannot be neglected anymore and the fluctuation grows less fast. For $\lambda < \lambda_\kappa$, the growth of fluctuations is suppressed by thermal conduction.

In summary, thermal conduction can play a significant role in regulating the breakup of a radiatively cooling gaseous medium. Small-scale substructure can only emerge
5.2. Thermal Conduction, Including Magnetic Fields

In general, the interstellar medium is penetrated by magnetic fields. In most situations, the electron mean free path $\lambda_e$ is large compared to the length scale at which the resistive destruction of the magnetic field is significant. A tangled magnetic field can then develop, concentrated on scales $l_B$ that are smaller than $\lambda_e$ (Chandran & Cowley 1998). When the gyroradius

$$a = \frac{v_T m_e c}{e B} \approx 2.2 \times 10^8 \sqrt{T/10^8 K} \left( \frac{\mu G}{B} \right) \text{ cm}$$

(38)

of thermal electrons with typical velocities $v_T = (kT/m_e)^{1/2}$ is much smaller than $l_B$ or $\lambda_e$, the magnetic field controls the
Fig. 9.—Evolution of a nonlinear perturbation with $K = 1000$ and $\beta = 0$. Upper panels show the evolution of the maximum density and temperature (solid line) and of the minimum density and temperature (dashed lines) for two cooling timescales. Shortly after one cooling timescale, the low-density regions cool down to the minimum temperature too and the reversed pressure gradient erases the density contrast. The lower panels show the evolution of the density and temperature during the short epoch at $\tau \approx 1$, when a two-phase medium has formed.

Fig. 10.—Evolution of the maximum and minimum density (left-hand panel) and maximum and minimum temperature (right-hand panel) of a nonlinear perturbation, including a cooling term with $\beta = 0$ and a heating term that dominates for $\rho < \rho_c = 0.1 \rho_0$, where $\rho_0$ is the initial average density. The density perturbation grows and becomes nonlinear within a cooling timescale. As gas is pushed into the high-density regions, the density in the interclump region decreases below the critical value where heating begins to dominate and the interclump gas heats up. A stable two-phase medium forms.
motion of individual electrons. This condition is satisfied in many astrophysical plasmas even if the magnetic field is too weak to be hydrodynamically important.

If \( a \) is small compared to the length scale \( \lambda \) of a fluctuation, heat is conducted according to the classical thermal conduction equation (Spitzer 1962); however, with a thermal conductivity \( \kappa_B \) that is reduced from the classic Spitzer value \( \kappa \) as a result of the tangled magnetic field by (Chandran & Cowley 1998)

\[
\kappa_B \approx \frac{0.1}{\ln(l_B/a)} \kappa .
\]  

(39)

If we normalize \( B \) to its value for magnetic-to-thermal energy equipartition \( B_T = (24\pi\rho R_g T)^{1/2} \), we find

\[
\ln\left(\frac{l_B}{a}\right) = -3.1 + \ln\left(\frac{B}{B_T}\right) + 2\ln\left(\frac{T}{K}\right) - 0.5\ln\left(\frac{n}{cm^{-3}}\right) + \ln\left(\frac{l_B}{\lambda_c}\right) .
\]  

(40)

For weak magnetic fields \( B \approx 0.01B_T \) and length scales \( lB \) of the order of the electron mean free path, equation (40) leads to \( \ln(l_B/a) \geq 10 \), thus reducing \( \kappa_B \) by 2 orders of magnitude and the length scale of thermal conduction by 1 order of magnitude. As an example, the shaded area in Figure 8b shows the wavelength regime \( \lambda_c \leq \lambda \leq \lambda_{crit} \), where linear fluctuations with \( \rho_{in}/\rho_0 = 10^{-2} \) could grow as a result of cooling, assuming a pressure of \( P/k_B = 10^4 \) K cm\(^{-3}\) and \( l_B = 0.1\lambda_c \). The presence of a weak magnetic field can suppress thermal conduction efficiently, allowing small-scale structure with wavelengths \( \lambda \approx 10^{-2} \) pc to emerge as a result of cooling.

6. THE IMPORTANCE OF HEATING AND THE EMERGENCE OF A STABLE TWO-PHASE MEDIUM

The calculations in §§ 3 and 4 showed that cooling gas clouds with small initial perturbations break up on a critical wavelength \( \lambda_{crit} \) below which overdensity first becomes nonlinear. If the initial amplitude is a decreasing function of \( \lambda \), the clouds would break up on the smallest length where the local radiative cooling law remains valid and fluctuations are not destroyed by conduction. But, if the initial amplitude is an increasing function of the wavelength,

\[
\lambda_{crit} = \frac{2\pi\tau_c(0)}{K_{crit}} \sqrt{\frac{R_g T_0(0)}{\mu}} ,
\]  

(41)

small perturbations on scales \( \lambda < \lambda_{crit} \) are erased when the gas accumulates in the center of fluctuations with \( \lambda = \lambda_{crit} \). Perturbations with \( \lambda > \lambda_{crit} \) do not become nonlinear, but they break up into substructures with \( \lambda = \lambda_{crit} \).

After a cooling time, the dense, cold, nonlinear perturbations are embedded in a warmer, diffuse environment (see Fig. 9). However, the example in Figure 9 shows that the cooling timescale of the interclump gas remains short compared with the initial cooling timescale. This gas therefore cools to a ground state temperature \( T_{min} \) shortly after \( \tau = 1 \). Subsequently, the reversed pressure gradient would remove the fluctuations unless they are gravitationally bound.

To maintain a stable two-phase medium, a heating term must be included (Field, Goldsmith, & Habing 1969). Here we assume a power-law dependence of the heating rate

\[
\Gamma_h = \Gamma_0 \rho^\gamma ,
\]  

(42)

where \( \Gamma_0 \) and \( \gamma \) are constants. In general, the size of the whole cooling region is large compared with \( \lambda_{crit} \), such that it cannot establish pressure equilibrium with the surrounding confining medium during a cooling timescale. In this case, the region cools isochorically and breaks up into substructures on scales of \( \lambda_{crit} \) before establishing pressure equilibrium with the environment. If the average gas density \( \rho_0 \) is smaller than a critical value

\[
\rho_T = \left(\frac{\Gamma_0}{L_0} T^{-\beta}\right)^{(1/2-\gamma)} ,
\]  

(43)

heating would dominate everywhere and the region would adjust to a thermal equilibrium state where heating is balanced against cooling. If \( \rho_0 > \rho_T \), cooling dominates and the density fluctuations would grow and become nonlinear as discussed in the previous sections. Eventually, after a cooling time, the region would break into cold, high-density condensations, which are separated by warm gas with densities \( \rho_{min} < \rho_0 \). If \( \rho_{min} > \rho_T \), this interclump medium would cool as shown in Figure 9 and the density fluctuations would be erased. Figure 10 shows a situation with \( \rho_{min} < \rho_T \). Heating dominates in the interclump region, where the gas temperature and gas pressure rise, until pressure equilibrium is established. A stable two-phase medium has formed with cold clouds of minimum temperature embedded in a hot interclump medium with a temperature that is determined by the balance of cooling and heating.

7. DISCUSSION

The discussions in this paper focused on the emergence of small-scale perturbations. We have assumed the preexistence of small initial perturbations, which is a reasonable assumption for dynamically evolving systems like the interstellar medium in galaxies or in galactic clusters. We have limited our analysis to the optically thin regime such that radiation transfer is caused solely by optically thin local radiative processes. This approximation is appropriate for the collapse of supercritical clouds where the effect of a magnetic field is dominated by thermal processes. Such a situation may be particularly relevant for the formation of stellar clusters and first-generation stars in galaxies. Provided that the density of the progenitor clouds is relatively small, the local cooling approximation is adequate. We also neglected the interaction and merging of clumps. These processes become important for the subsequent evolution and they will be considered in subsequent papers.

In the context of our approximations, we have shown that thermal instability can lead to the breakup of large clouds into cold, dense clumps with a characteristic length scale, which is given by \( \lambda_{crit} \) in equation (41) or by the smallest unstable wavelength that is not erased by thermal conduction, depending on whether the amplitude of the initial perturbation is an increasing or decreasing function of wavelength. For linear perturbations with overdensities \( \rho_B/\rho_0 \approx 0.01 \), the critical wavelength lies in the regime of \( 10^{-3} \) to \( 10^{-1} \) pc, depending on the initial temperature. The emergence of small-scale dense subcondensations is equivalent to fragmentation. As in a thermally unstable region, the
cooling timescale is shorter than the dynamical timescale, gravity has no time to play an important role during this fragmentation process. \( \lambda_{\text{crit}} \) may be either smaller or bigger than the Jeans' length. In the latter case, gravity becomes important eventually. In general, however, thermally induced fragmentation of clouds with small initial density fluctuations proceeds the onset of gravitational instability of their individual clumps.

In our analyses, we adopted an idealized power-law cooling function. In reality, the cooling efficiency would terminate when the main cooling agents reach their ground state or establish an equilibrium with some external heating source. The latter is necessary for the clouds to attain a two-phase medium. Interaction between these two phases may determine the pressure, density, and infall rate of the cloud complex as well as the dynamical evolution and size distribution of cloudlets and subcondensations. The analysis of this interaction will be presented elsewhere.

We would like to thank referee Andrea Ferrara for helpful suggestions and for pointing out the importance of thermal conduction. A. Burkert thanks the staff of UCO/Lick Observatory for the hospitality during his visits. We thank S. D. Murray for useful conversation. This research was supported by NASA through NAG 5-3056 and an astrophysics theory program that supports a Joint Center for Star Formation Studies at NASA-Ames Research Center, University of California, Berkeley, and University of California, Santa Cruz.

REFERENCES

Arons, J., & Max, C. E. 1975, ApJ, 196, L77
Balbus, S. A. 1986, ApJ, 303, L79
———. 1995, in ASP Conf. Proc. 80, The Physics of the Interstellar Medium and Intergalactic Medium, ed. A. Ferrara, C. F. McKee, C. Heiles, & P. R. Shapiro (San Francisco: ASP), 328
Blitz, L. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine (Tucson: Univ. of Arizona Press), 125
Brown, J. H., Burkert, A., Truran, J. W. 1991, ApJ, 376, 115
———. 1995, ApJ, 440, 666
Burkert, A., Bate, M., & Bodenheimer, P. 1997, MNRAS, 289, 497
Burkert, A., & Bodenheimer, P. 1993, MNRAS, 264, 798
———. 1996, MNRAS, 280, 1190
Carpenter, J. M., Meyer, M. R., Dougados, C., Strom, S. E., & Hillenbrand, L. A. 1997, AJ, 114, 198
Chandran, B. D. G., & Cowley, S. C. 1998, Phys. Rev. Lett., 80, 3077
Cowie, L. L., & McKee, C. F. 1977, ApJ, 211, 135
Crutcher, R. M., Troland, T. H., Goodman, A. A., Heiles, C., Kazes, I., & Myers, P. C. 1993, ApJ, 407, 175
Dalgarno, A., & McCray, R. A. 1972, ARA&A, 10, 375
Ferrara, A., & Shchekinov, Y. 1993, ApJ, 417, 595
Field, G. B. 1965, ApJ, 142, 531
Field, G. B., Goldsmith, D. W., & Habin, H. J. 1969, ApJ, 155, L49
Franco, J., & Cox, D. P. 1983, ApJ, 273, 343
Gammie, C. F., & Ostriker, E. C. 1996, ApJ, 466, 814
Hillenbrand, L. A. 1997, AJ, 113, 1733
Hunter, C. 1962, ApJ, 136, 594
Kraft, R. P. 1979, ARA&A, 17, 309
Larson, R. B. 1981, MNRAS, 194, 809
Lin, D. N. C., & Murray, S. D. 1992, ApJ, 394, 523
Lizano, S., & Shu, F. H. 1989, ApJ, 342, 834
Loewenstein, M. 1990, ApJ, 349, 471
MacLow, M. M. 1999, ApJ, 524, 169
MacLow, M. M., Klessen, R. S., Burkert, A., Smith, M. D., & Kessel, O. 1998, Phys. Rev. Lett., 80, 2754
McKee, C. F. 1989, ApJ, 345, 782
McKee, C. F., & Begelman, M. C. 1990, ApJ, 358, 392
Mouschovias, T. C. 1991, ApJ, 373, 169
Mouschovias, T. C., & Spitzer, L. H. 1976, ApJ, 210, 326
Murray, S. D., & Lin, D. N. C. 1992, ApJ, 400, 265
Myers, P. C., & Goodman, A. A. 1988, ApJ, 326, L27
Nakano, T. 1979, PASJ, 31, 697
Ostriker, E. C., Gammie, C. F., & Stone, J. M. 1999, ApJ, 513, 259
Parker, E. N. 1953, ApJ, 117, 431
Schwarz, J., McCray, R., & Stein, R. 1972, ApJ, 175, 673
Shu, F., Najita, J., Gulli, D., & Ostriker, E. C. 1993, in Protostars and Planets III, ed. E. H. Levy & J. I. Lunine (Tucson: Univ. of Arizona Press), 3
Solomon, P. M., Rivolo, A. R., Barrett, J., & Yahil, A. 1987, ApJ, 319, 730
Spitzer, L. 1962, Physics of Fully Ionized Gases (New York: Wiley)
———. 1978, Physical Processes in the Interstellar Medium (New York: Wiley)