Vol. 21 No. 3

PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF JAPAN

1969

Formation of Galaxies by Thermal Instability

Yoshiaki SOFUE

Department of Physics, Nagoya University, Nagoya

(Received April 16, 1969)

Abstract

It is found that the mass of a proto-galaxy is restricted to $10^6 M_\odot \lesssim \mathfrak{M} \lesssim 10^{12} M_\odot$, under the assumption that a density excess of gas in an expanding universe is initiated by the thermal instability due to radiative cooling and is followed by non-linear gravitational contraction, resulting in a gravitationally bound system such as a galaxy. Such contraction takes place at cosmic age $t \sim 10^7$ years.

The temperature of matter when the initial thermal instability occurs is required to be 10^6—10^{9}°K. This high temperature would be attained if relatively small kinetic fluctuation in cosmic background radiation causes turbulence in cosmic gas which heats the gas through its dissipation.

1. *Introduction*

It has been shown by LIFSHITZ (1946) that a growth of gravitational instability in a homogeneous expanding universe is too slow in a linear approximation to form a gravitationally bound system. GOLD and HOYLE (1959) have suggested that a thermal instability due to radiative cooling must have an important role in the formation of galaxies. The thermal instability has a larger growth rate than that of gravitational instability, so far as the amplitude of the density excess is sufficiently small. Recently, KIHARA (1967, 1968) has treated non-linear gravitational instability and has shown that a condensation with sufficiently large amplitude in density excess grows rapidly even in an expanding universe.

Therefore it is possible to consider that a fluctuation in the density of gas is created by thermal instability due to radiative cooling and grows by non-linear gravitational contraction after the density excess becomes sufficiently large.

It is the purpose of the present paper to investigate the dynamical process through which proto-galaxies are formed by thermal instability followed by gravitational contraction, and to search physical conditions for the formation of proto-galaxies through such a mechanism.

2. *Physical Conditions for the Formation of Proto-Galaxies*

We suppose that a gravitationally bound system in an expanding universe is produced by non-linear gravitational contraction, after the density fluctuation generated initially by thermal instability due to radiative cooling becomes sufficiently large. Under this assumption, the following conditions are required for the formation of proto-galaxies:

(A) The characteristic time scale of the growth of density fluctuation by thermal instability should be smaller than that of cosmological expansion. Otherwise the density excess will be diluted by the expansion of the universe.

(B) The energy density of background radiation should not be so large to obstruct the growth of instability.

(C) The temperature should smoothly decrease to the temperature at which non-linear gravitational contraction of the system is possible.

3. *Time Scale of the Growth of Thermal Instability*

We denote time scales of thermal instability and cosmological expansion by τ and $\tau_{\rm exp}$, respectively. The first condition (A) is written as

$$\tau < \tau_{\rm exp} \,. \tag{1}$$

By neglecting thermal conduction, the characteristic equation for the growth rate $\nu = 1/\tau$ of linear thermal instability with wave number k is reduced to (Field 1965; basic equations are shown in Appendix)

$$\nu^3 + c_s k_T \nu^2 + c_s^2 k^2 \nu + \frac{c_s^3 k^2}{\gamma}(k_T - k_\rho) = 0 \,. \tag{2}$$

Here, c_s, T, and γ represent the acoustic velocity, the temperature, and the ratio of specific heats of the gas, respectively. We assume that the cooling is in balance with heating for the background radiation and that the heating rate per unit mass and per unit time is constant. We denote the cooling rate per unit mass and per unit time by $\mathscr{L}$, then k_ρ and k_T are given by

$$k_\rho = \frac{\mu(\gamma-1)\rho \dfrac{\partial \mathscr{L}}{\partial \rho}}{\mathscr{R} c_s T} \,,$$

and

$$k_T = \frac{\mu(\gamma-1) \dfrac{\partial \mathscr{L}}{\partial T}}{\mathscr{R} c_s} \,,$$

where μ and $\mathscr{R}$ are the mean molecular weight and the gas constant. If we adopt an interpolation formula for the cooling rate

$$\mathscr{L} = \text{const} \times \rho T^\alpha \,, \tag{3}$$

k_ρ and k_T are reduced to

$$\left. \begin{aligned} k_\rho &= \mu / c_s \tau_{\rm cool} \,, \\ \text{and} \quad k_T &= \alpha\mu / c_s \tau_{\rm cool} \,. \end{aligned} \right\} \tag{4}$$

In the above expressions, $\tau_{\rm cool}$ represents the cooling time and is given by

$$\tau_{\rm cool}=\frac{1}{\gamma-1}\frac{\mathscr{R}T}{\mathscr{L}}\,. \tag{5}$$

Perturbation with wave number k grows if the characteristic equation has a root with positive real part.

(a) Case 1: $1/kc_s \ll \tau_{\rm cool}$. This is the case when the traveling time of acoustic wave through the wave length of perturbation is less than the cooling time. In such a case, the characteristic equation has the following solution,

$$\nu_1=-\frac{\mu(\alpha-1)}{\gamma}\cdot\frac{1}{\tau_{\rm cool}}+O\left(\frac{1}{kc_s\tau_{\rm cool}}\right), \tag{6}$$

$$\nu_{2,3}=\pm ikc_s-\frac{\mu\left(\alpha+\frac{1}{\gamma-1}\right)}{2}\cdot\frac{1}{\tau_{\rm cool}}+O\left(\frac{1}{kc_s\tau_{\rm cool}}\right). \tag{7}$$

Equation (6) gives a positive growth rate for $\alpha<1$, and equation (7) gives wave modes whose amplitudes damp away for $\alpha>-1/(\gamma-1)$.

(b) Case 2: $1/kc_s \gg \tau_{\rm cool}$. The characteristic equation has the following solution;

$$\nu_1=-\frac{1}{\tau_{\rm cool}}\{\alpha\mu-O(k^2c_s{}^2\tau^2_{\rm cool})\}\,, \tag{8}$$

$$\nu_{2,3}=kc_s\left[+i\sqrt{\frac{\alpha-1}{\gamma\alpha}}\left[1-5\frac{\{1+\alpha(\gamma-1)\}\{1+\alpha(\gamma/2-1)\}}{\alpha^3(1-\alpha)\mu\gamma^2}kc_s\tau_{\rm cool}\right]\right.$$
$$\left.-\frac{1+\alpha(\gamma-1)}{2\alpha^2\mu\gamma}kc_s\tau_{\rm cool}+O(k^2c^2{}_s\tau^2_{\rm cool})\right]\quad\text{for}\quad\alpha>1 \tag{9}$$

or

$$\nu_{2,3}=\pm\sqrt{\frac{1-\alpha}{\gamma\alpha}}kc_s+kc_s\left[\pm5\,\frac{\{1+\alpha(\gamma-1)\}\{1+\alpha(\gamma/2-1)\}}{\alpha^3(1-\alpha)\mu\gamma^2}-\frac{1+\alpha(\gamma-1)}{2\alpha^2\mu\gamma}\right]kc_s\tau_{\rm cool}$$
$$+O(k^2c^2{}_s\tau^2_{\rm cool})\quad\text{for}\quad\alpha<1\,. \tag{10}$$

Expression (8) gives a damping mode for any positive value of α. Expression (9) gives a wave mode whose amplitude damps away. When the value of α is less than unity, expression (10) gives a positive root ν_2 with the order of magnitude of kc_s, which means that a gas cloud condenses with acoustic velocity.

We take into account the cooling due to free-free radiation alone. Lyman α radiation is considered not to be effective for the cooling, because the gas cloud is optically thick for the Lyman α radiation at temperatures just above $10^4\,{}^\circ$K, and it is negligible compared with free-free radiation when the temperature becomes sufficiently high. If the temperature descends below $2\times10^4\,{}^\circ$K, the cooling rate due to free-free radiation becomes weak because ionized gas recombines to neutral one and the value of α exceeds unity not to cause a thermal instability. On the other hand, if the temperature exceeds $\sim10^9\,{}^\circ$K, the thermal instability would not occur, because at such a high temperature, the gas is in relativistic state and the temperature dependence of the cooling rate becomes so large that α exceeds unity. Cooling rate function $f(T)=m_{\rm H}\mathscr{L}/n$ for pure

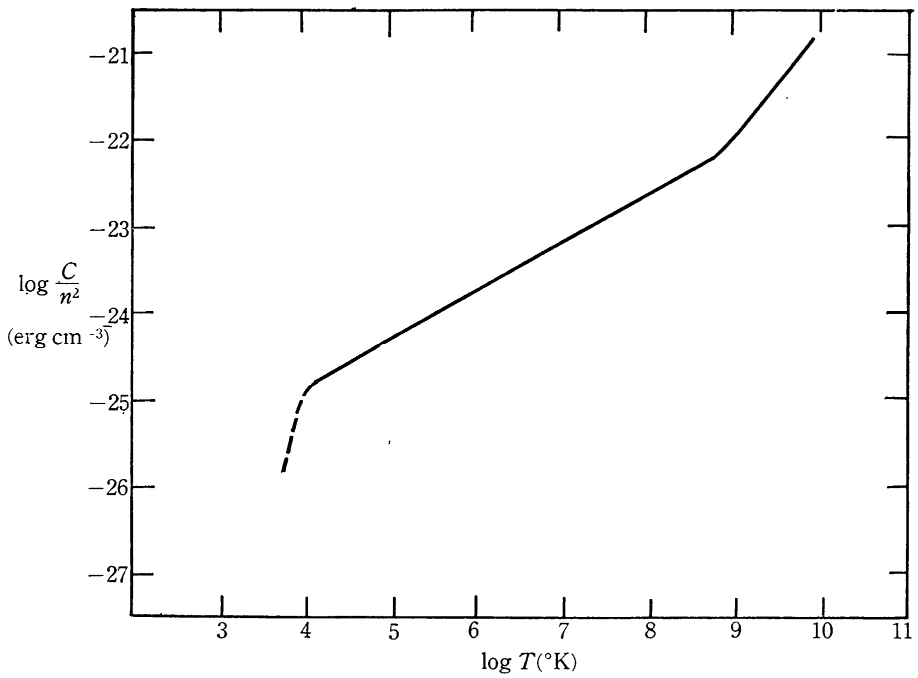

FIG. 1. Cooling rate function $m_H\mathscr{L}/n=C/n^2$, where C denotes the cooling rate per unit volume and per unit time.

hydrogen gas due to free-free radiation is shown in Figure 1. Here, m_H and n denote the mass of hydrogen atom and the number density of gas, respectively. The thermal instability occurs in temperature range $2\times10^4\,°\mathrm{K}<T<10^9\,°\mathrm{K}$, since the value of α is less than unity in this temperature range.

Hence, the time scale τ of the growth of thermal instability due to free-free radiation is approximately given as

$$\tau\approx\frac{10}{3}\cdot\tau_{\mathrm{cool}}=5\frac{\mathscr{R}T}{\mathscr{L}}\,,\qquad \text{if}\quad \frac{1}{kc_s}<\tau_{\mathrm{cool}}\,, \tag{11}$$

and

$$\tau\approx\sqrt{5/3}\,\frac{1}{kc_s}\,,\qquad \text{if}\quad \frac{1}{kc_s}>\tau_{\mathrm{cool}}\,, \tag{12}$$

by putting as $\alpha=1/2$ and $\gamma=5/3$.

4. *Time Scale of Cosmological Expansion*

As the time scale of the cosmological expansion, the following expression is used;

$$\tau_{\exp}=\frac{R^3}{\frac{d}{dt}(R^3)}=\frac{1}{2}t\,, \tag{13}$$

where t and R are the cosmic time and expansion factor of the universe. We used here an approximation that the cosmological expansion obeys $R/R_0=(t/t_0)^{2/3}$, where subscript 0 denotes the value of the present epoch and $t_0=10^{10}$ years.

5. *Effect of Background Radiation*

We may restrict the epoch of formation of galaxies due to thermal instability to the stage after the separation of matter and radiation, that is, $t \gtrsim 10^5$ years, by taking the present value of cosmic background radiation as 3°K. Before $t \sim 10^5$ years, the radiation and the matter are considered to be in isothermal and just after the separation the radiation drag may obstruct the growth of thermal instability of gas. Here we represent the second condition (B) by $n<10^5\,\mathrm{cm}^{-3}$ approximately, where n denotes the initial density when thermal instability takes place.

6. *Evolutionary Path of Growth of Thermal Instability*

In Figures 2—4, the regions, where inequalities $\tau<\tau_{\mathrm{exp}}$ and $n<10^5\,\mathrm{cm}^{-3}$ are satisfied simultaneously, are shown in $n-T$ plane by I, I′, II, and II′, for various mass of perturbation $\mathfrak{M}=\rho\lambda^3$ $(\lambda=2\pi/k)$. Here $\rho=m_{\mathrm{H}}n$ represents the density of gas. The loci of the growth of perturbation in I and I′ are as follows (see Appendix),

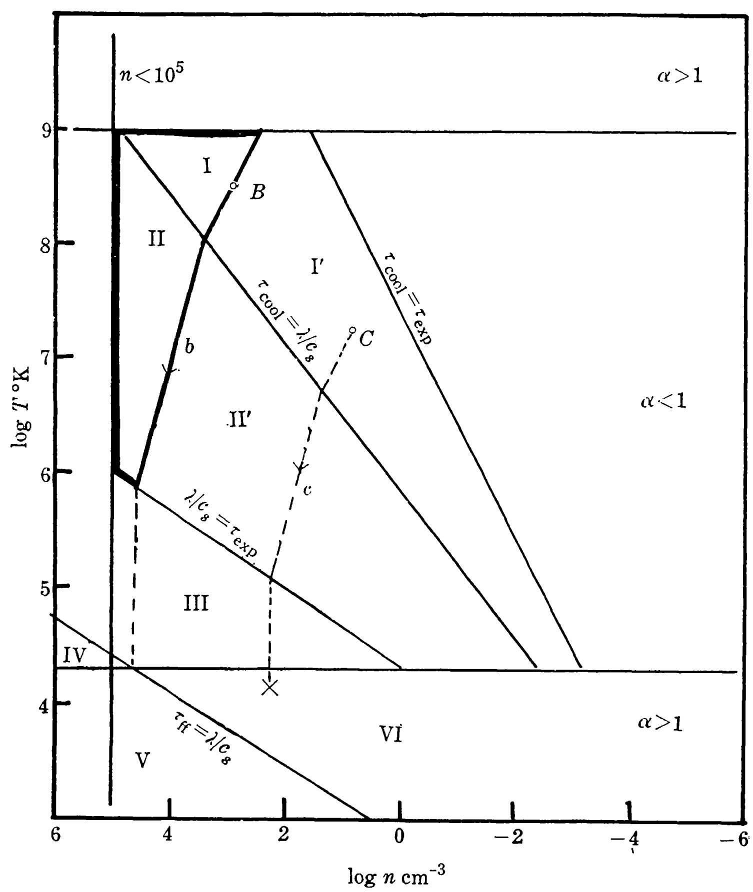

FIG. 2. $n-T$ plane for $\mathfrak{M}=10^{39}$ grams $(=0.5\times10^6\,M_\odot)$.

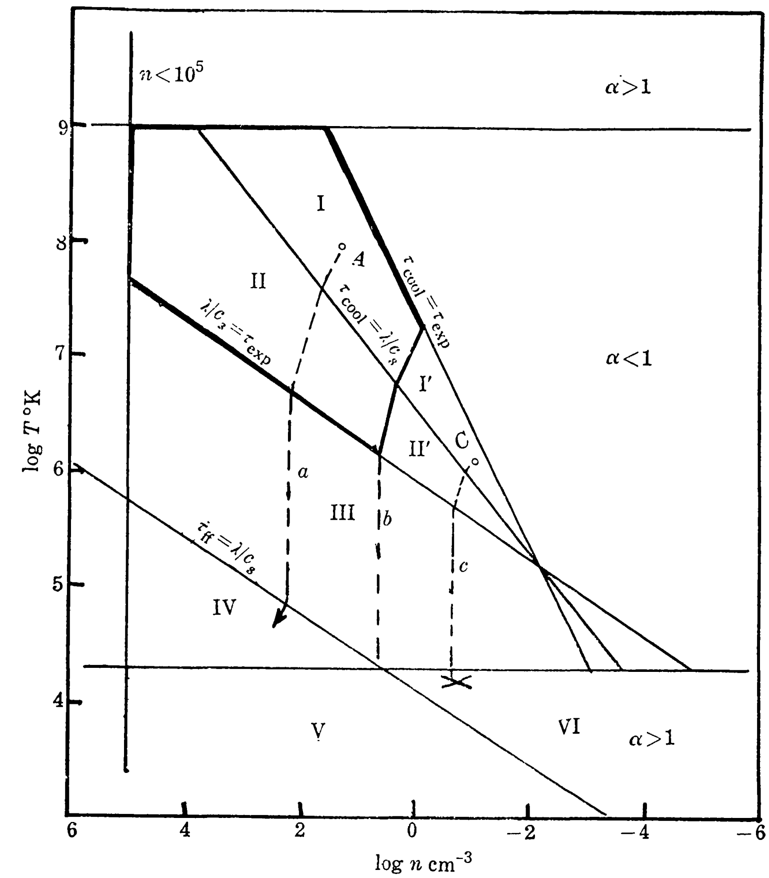

FIG. 3. $n-T$ plane for $\mathfrak{M}=10^{41}$ grams ($=0.5\times10^8\,M_\odot$).

$$\frac{d\ln T}{d\ln\rho}=-1+O(k^{-2}c_s^{-2}\tau_{\mathrm{cool}}^{-2})\approx-1\,,\qquad \text{if}\quad \frac{k}{kc_s}\ll\tau_{\mathrm{cool}}\,. \tag{14}$$

In II and II′ a perturbation grows along a locus

$$\frac{d\ln T}{d\ln\rho}=-\left(1+\frac{1-\alpha}{\alpha}\right)+O(k^2c_s^2\tau_2^{\mathrm{cool}})\approx-2\,,$$

$$\text{if}\quad \frac{1}{kc_s}\gg\tau_{\mathrm{cool}}\quad\text{and}\quad \alpha=\frac{1}{2}\,. \tag{15}$$

When the system goes into region III, of Figures 2 and 3, where the characteristic time scale of thermal contraction, $1/kc_s$, is larger than that of cosmological evolution but the time scale of radiative cooling is small enough, it may be approximated that the contraction does not proceed but the temperature decreases by radiative cooling.

7. *Gravitational Contraction*

The condition (C) that the gas cloud generated initially by thermal instability is followed by gravitational contraction is written by

$$\tau_{ff}<\frac{\lambda}{c_s}\,, \tag{16}$$

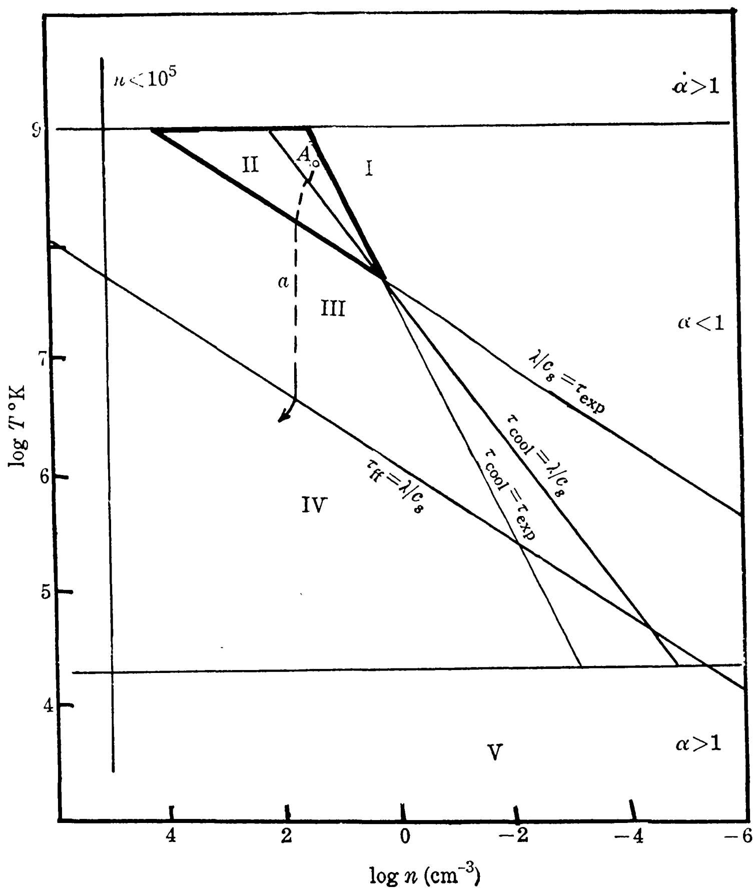

FIG. 4. $n-T$ plane for $\mathfrak{M}=10^{44}$ grams ($=0.5\times10^{11}\,M_\odot$).

where $\tau_{ff}=(G\rho)^{-1/2}$ and $\lambda=(\mathfrak{M}/\rho)^{1/3}$ represent the time scale of gravitational free-fall and the linear scale of the gas cloud, respectively. The inequality (16) is satisfied in regions IV and V of Figures 2–4.

8. *Formation of Proto-Galaxies*

A system whose evolutionary path in $n-T$ plane can proceed to regions IV or V, as shown, for example, by a path a which starts from I or II, can grow to be a gravitationally bound system. On the other hand, a system whose path goes into region VI, where neither thermal condensation nor gravitational contraction occurs, will be diluted again to be a cosmic gas due to internal thermal pressure. Such a case is shown by a locus c which starts from region I′ or II′. Locus b which seprates regions I and I′ or regions II and II′ is marginal.

We may conclude that it is only the system whose evolutionary path starts from I or II that can grow to be a galaxy, and that a system whose locus does not start from I or II does not become a gravitationally bound system.

9. *Results*

We find that the mass of the system, whose evolutionary path starts from region I or II and which grows to be a gravitationally bound system, is restricted as

$$10^{39} \lesssim \mathfrak{M} \lesssim 10^{45} \text{ grams}$$

or

$$10^6 M_\odot \lesssim \mathfrak{M} \lesssim 10^{12} M_\odot \, . \tag{17}$$

This is the first important result of this article. Secondly the temperature of medium at the initial phase of formation of proto-galaxies should be as high as

$$10^6 \,°\mathrm{K} \lesssim T \lesssim 10^9 \,°\mathrm{K} \, . \tag{18}$$

This high temperature is interpreted in next section. The final result is that the epoch of the formation of proto-galaxies is comparatively early in the evolution of the universe, that is

$$3 \times 10^5 \text{ years} < t < 3 \times 10^7 \text{ years} \, . \tag{19}$$

The time scale necessary for the system to pass from region I to region V is roughly 10^6—10^7 years.

10. *A Possible Interpretation of Hot Gas Model*

The temperature of gas just before the formation of proto-galaxies is required to be as high as $T \approx 10^7$—10^9 °K at $t \sim 10^7$ years as seen in the preceding section. If the gas in the universe expands adiabatically as the matter temperature descends as $T \propto R^{-2}$, such a high temperature cannot be realized. Therefore, we must suppose some heating source to explain such a high temperature. The only possible energy source in this stage is cosmic background radiation, whose energy density is much higher than thermal energy density of the gas. If only a small fraction of energy of the black-body radiation transforms to thermal energy of the gas, the gas will be heated to sufficiently high temperature.

Let a fraction, Δu_r, of the energy density of the background radiation, u_r, be transformed to thermal energy, nkT, i.e., $nkT \sim \Delta u_r$, then the value of $\Delta u_r / u_r$ required to heat the gas up to 10^7 °K is only of the order of 10^{-3} at $t \sim 10^7$ years.

A possible mechanism to transform a small fraction of radiative energy to thermal energy is as follows: Even a small kinetic fluctuation in radiation field or turbulence in radiation may blow the gas. The blown gas may cause a strong gaseous turbulence. Decay of the turbulence will be enough to heat the gas to high temperature.

The turbulent velocity of gas is of the order of $v_t \sim 300$ km/s if we assume that $\Delta u_r \sim (1/2) v_t^2 \sim nkT$ and that $T = 10^7$ °K at $t = 10^7$ years. This velocity agrees well with the observed velocity of random motion of galaxies. It seems natural to guess that a part of the turbulent motion of cosmic gas in the stage of formation of proto-galaxies remains as the gross motion of galaxies up to the present epoch. Conversely it is possible to see the observed random motion of galaxies as an evidence for the existense of highly turbulent stage, hence for the high temperature stage just before the formation of galaxies.

11. *Conclusions and Discussion*

A density fluctuation in an expanding universe is considered to be initiated to grow by thermal instability due to radiative cooling. Only if the density excess grows large enough to cause a non-linear gravitational contraction, the region can proceed to from a proto-galaxy. Mass of the proto-galaxies produced by such a mechanism is restricted as $10^6 M_\odot \lesssim \mathfrak{M} \lesssim 10^{12} M_\odot$. The upper limit of the mass range well agrees with that of observed galaxies.

The cosmic time when proto-galaxies were formed was $t \sim 10^7$ years. This shows that galaxies were formed in relatively early stage of the universe, but it does not contradict with the fact that the age of observed galaxies is of the order of 10^{10} years. Moreover the existence of interstellar magnetic filed of $B \sim 10^{-6}$ gauss may be agreeable if we consider that a part of cosmic magnetic field has remained in a galaxy when the galaxy was formed. Existence of the uniform metagalactic magnetic field of order of 10^{-9} gauss at the present epoch has been shown recently by SOFUE, FUJIMOTO, and KAWABATA (1968). This cosmic magnetic field would be 10^{-5} gauss at $t \sim 10^7$ years, because the strength of uniform magnetic field is considered to be vary as $B \propto R^{-2} \propto t^{-4/3}$.

The temperature of gas just before the formation of proto-galaxies is required to be as high as 10^7—10^9 °K. This temperature would be attained by heating through a decay of turbulent motion of gas caused by kinetic fluctuation in cosmic background radiation. The estimated turbulent velocity is of the order of 300 km/s. Observed velocity of random motion of galaxies, which is of the order of 300 km/s, may be an evidence for the existence of such a highly turbulent stage just before the formation of galaxies.

The author acknowledges with hearty thanks his indebtedness to Professor W. Unno, who suggested him the subject, and to Dr. S. Kato of the University of Tokyo for their valuable discussions and advice. He is also grateful to Professor K. Kawabata and to Dr. M. Fujimoto for reading the manuscript with valuable discussions. The original version of this work was performed when the auther was at the University of Tokyo.

Appendix. *Locus of Growth of Thermal Instability in $\rho - T$ Plane*

The basic equations are (FIELD 1965)

$$\frac{d\rho}{dt} + \rho \operatorname{div} \boldsymbol{v} = 0 , \tag{A.1}$$

$$\rho \frac{d\boldsymbol{v}}{dt} + \operatorname{grad} p = 0 , \tag{A.2}$$

$$\frac{1}{\gamma-1} \frac{dp}{dt} - \frac{\gamma}{\gamma-1} \frac{p}{\rho} \frac{d\rho}{dt} + \rho \mathscr{L} - \operatorname{div}(K \operatorname{grad} T) = 0 , \tag{A.3}$$

and

$$p - \frac{\mathscr{R}}{\mu} \rho T = 0 , \tag{A.4}$$

with

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \boldsymbol{v} \cdot \operatorname{grad} ,$$

where $\boldsymbol{v}$ and K denote the velocity of gas and the coefficient of thermal conductivity, respectively. Assuming perturbations of the form

$$a(\boldsymbol{r}, t)=a_1 \exp(\nu t+i\boldsymbol{k}\cdot\boldsymbol{r}) , \tag{A.5}$$

and neglecting the thermal conductivity, we find the linearized equations for the perturbations to be

$$\nu\rho_1+\rho i\boldsymbol{k}\cdot\boldsymbol{v}_1=0 , \tag{A.6}$$

$$\nu\rho\boldsymbol{v}_1+i\boldsymbol{k}p_1=0 , \tag{A.7}$$

$$\frac{\nu}{\gamma-1}p_1-\frac{\nu\gamma p}{(\gamma-1)\rho}\rho_1+\rho\mathscr{L}_\rho\rho_1+\rho\mathscr{L}_T T_1=0 , \tag{A.8}$$

and

$$\frac{p_1}{p}-\frac{\rho_1}{\rho}-\frac{T_1}{T}=0 , \tag{A.9}$$

where $\mathscr{L}_\rho=\partial\mathscr{L}/\partial\rho$ and $\mathscr{L}_T=\partial\mathscr{L}/\partial T$. The subscript 1 in the above expressions represents the perturbed quantities. The characteristic equation for equations (A.6)—(A.9) is given by equation (2).

From equations (A.6) and (A.7)

$$p_1=-\frac{\mu^2}{k^2}\rho_1 . \tag{A.10}$$

By inserting (A.10) into equation (A.9)

$$\left(\frac{\nu^2\rho}{k^2\rho}+1\right)\frac{\rho_1}{\rho}+\frac{T_1}{T}=0 . \tag{A.11}$$

Using equation (A.4), $\gamma RT=c_s^2$, $\nu=1/\tau$ and $T_1/\rho_1=dT/d\rho$ and putting $\mu=1$, we rewrite (A.11) as

$$\frac{d\ln T}{k\ln\rho}=-\left(1+\frac{\gamma}{\tau^2k^2c_s^2}\right) . \tag{A.12}$$

(a) Case 1: $(kc_s)^{-1}\ll\tau_{\rm cool}$. Inserting equation (6) into (A.12) we find equation (14).

(b) Case 2: $(kc_s)^{-1}\gg\tau_{\rm cool}$. Inserting equation (10) into (A.12) we find equation (15).

The thermal instability grows along the locus (14) or (15) in $\rho-T$ plane.

References

FIELD, G. B. 1965, *Astrophys. J.*, **142**, 531.
GOLD, T. and HOYLE, F. 1959, in *Paris Symposium on Radio Astronomy*, ed. R. M. Bracewell (Stanford Univ. Press, Palo Alto), p. 583.
KIHARA, T. 1967, *Publ. Astr. Soc. Japan*, **19**, 121.
KIHARA, T. 1968, *Publ. Astr. Soc. Japan*, **20**, 220.
LIFSHITZ, E. M. 1946, *J. Phys. U.S.S.R.*, **10**, 110.
SOFUE, Y., FUJIMOTO, M., and KAWABATA, K. 1968, *Publ. Astr. Soc. Japan*. **20**, 388.