THE ASTROPHYSICAL JOURNAL, 288:465–480, 1985 January 15


# SELF-REGULATING GALAXY FORMATION. I. H II DISK AND LYMAN-ALPHA PRESSURE

DONALD P. COX
Department of Physics, University of Wisconsin, Madison, and Space Physics and Astronomy Department, Rice University
*Received 1983 May 3; accepted 1984 August 1*

## ABSTRACT

Assuming a simple but physically based prototype for behavior of interstellar material during formation of a disk galaxy, coupled with the lowest order description of infall, a scenario is developed for self-regulated disk galaxy formation. Radiation pressure, particularly that of Lyman-$\alpha$ (from fluorescence conversion of Lyman continuum), is an essential component, maintaining an inflated disk and stopping infall when only a small fraction of the overall perturbation has joined the disk.

The resulting "galaxies" consist of a two-dimensional family whose typical scales and surface density are expressible in terms of fundamental constants. The model leads naturally to galaxies with a rich circumgalactic environment and flat rotation curves (but is weak in its analysis of the subsequent evolution of halo material).

The galactic family has a natural lower limit on radius around 150 pc, at which point the systems are spheroidal and resemble giant H II regions. The largest radius is roughly 20 kpc, depending somewhat on the epoch of formation. The model seems to apply most naturally to large spirals with radii from 4 to 20 kpc and average column (or surface) densities from 100 to 500 $M_\odot$ pc$^{-2}$.

The formation scenario for a Milky Way–like galaxy would be roughly as follows. The material eventually incorporated into the disk reached a maximum extension of 45 kpc at $t \sim 5 \times 10^8$ years ABB (after big bang). It is sufficiently sheared that the highest velocity material was only marginally bound. During recollapse, angular momentum was azimuthally shared, dictating a final radius smaller by a factor of 4. Disk formation commenced at $t \sim 10^9$ years, achieving an average surface density of 300 $M_\odot$ pc$^{-2}$ only $5 \times 10^7$ years later. At this point, star formation is so rapid in the disk that Lyman-$\alpha$ pressure dominates the disk structure as well as supporting the residual halo against further infall. The Lyman-$\alpha$–supported halo moves toward hydrostatic equilibrium with $M(r) \propto r$. The fragmentation of the halo into bits (many possibly substellar) under the influence of Lyman-$\alpha$ (or supernovae) would leave a galactic system with some 8% of its mass in the disk and the rest in a massive halo, possibly including the Magellanic Clouds.

The formation of such a system would be an extremely dramatic event. The peak luminosity would be of order $10^{46}$ ergs s$^{-1}$, Lyman-$\alpha$ accounting for 10–20%. It would last roughly $4 \times 10^7$ years during which much of the disk material entered the first generation of stars and while *several percent* of the galactic hydrogen burned to helium or beyond by stars with main-sequence lifetimes less than that duration. The scale height of the interstellar medium during formation would be about 150 pc, but the lack of definition of the mean plane would give the first generation of stars an inflated distribution.

*Subject headings:* galaxies: formation — galaxies: structure — stars: formation

## I. INTRODUCTION

Any attempt to explore the formation epoch of spiral galaxies is naturally frustrated on four counts. We need to understand the universe well enough to predict the distribution of infalling mass in angular momentum and time. We need to understand the structure, phases, and mechanisms of the nascent interstellar medium (NISM). We need to be able to determine the star formation rate and zero-age initial mass function (ZAIMF) as functions of themselves and their interactions with the NISM. Finally, we need to understand the feedback effects of the forming system on its own continued infall rate (or that of its neighboring systems). At the present time none of these four elements seem to be amenable to a confidence-inspiring analysis.

Consider in particular the problem of determining the rate of star formation. Its solution cannot be divorced from understanding mass flow through the phases of the interstellar medium (ISM). Mass filters from low density to high. It piles up at stable phases, accumulating until it achieves an adequate transfer rate to the next denser phase. Calculating the rate of mass transfer between phases and from the densest phase into stars involves understanding the heating, cooling, and environmental influences for each phase. In short, one must understand bottlenecks in the mass transfer.

These bottlenecks are at present not well understood even in the solar neighborhood. We know that there are at least two stable phases, diffuse and molecular clouds, and we know something about each. It may be that diffuse clouds are heated by photoelectric ejection off dust grains by starlight. It may be that they are held together by the pressure of a hot, low-density matrix maintained by supernova explosions. But what governs the rate of cloud destabilization that diverts their mass to the denser molecular cloud phase? Do we even know for sure whether the passage is fairly direct or whether the diffuse clouds dissipate and regroup many times before entering the molecular cloud phase?

Molecular clouds are largely opaque because of dust grains, and are self-gravitating. They may be heated in part by cosmic rays and by dissipation of turbulence of as yet unknown origin. They may be supported in part by magnetic fields, the latter kept tied to the material by ionization also due to cosmic rays. Electrons from elements with first ionization potentials below that of hydrogen may at times also be important.

In contrast to this partial understanding, we almost totally lack knowledge about the ISM of a galaxy less than $10^7$ years old, the NISM. Such a system may offer comparable complexity with different forms, or it may offer a remarkable simplicity. It is quite possible that for a while there are no stable diffuse or molecular clouds. The heavy-element abundance is very low, lowering dust content, cooling coefficients, heating rates, and opacities. The galactic magnetic field may not yet have arisen. Cosmic-ray acceleration or trapping may be inefficient. If so, what can we then infer about the NISM state and its star formation rate?

One answer to the above question was proposed by Cox (1983, hereafter Paper I). It was suggested that the densest stable phase available to the NISM was ionized hydrogen at $10^4$ K, with heating and ionization provided by stellar UV. This simple possiblity leads to a trivial equation of state for the medium and a prescription for the formation rate of massive stars. The basic idea is that any gas which is not kept photoionized cools rapidly and forms stars, some of which are hot enough to contribute ionizing photons. The critical star formation rate is that which just manages to keep the residual medium fully ionized.

In order to calculate the critical star formation rate, one needs to know only the efficiency of stars in producing ionizing photons. For time scales longer than the lifetime of the most massive stars, about $3 \times 10^6$ years, an *instantaneous recycling approximation* is appropriate: UV photons are returned immediately from any new generation of stars, making the UV production rate per unit volume $\phi\dot{\rho}_*$, where $\dot{\rho}_*$ is the contemporary density per second going into stars and $\phi$ is the number of ionizing photons returned per gram of star formation.

Summarizing from Paper I, we can write

$$\phi = f_1 f_2 f_3 \frac{7 \text{ Mev}}{\langle h\nu \rangle} \frac{1}{m_p} ,$$

where $m_p$ is the proton mass, 7 Mev is the energy liberated per proton burned, $\langle h\nu \rangle$ is the average ionizing photon energy, $f_1$ is the fraction of the hydrogen which is burned after entering a massive star, $f_2$ is the fraction of the luminosity beyond the Lyman limit, and $f_3$ is the fraction of the total mass in star formation which enters short-lived massive stars. Paper I estimated a consistent set of these parameters to be $f_1 = 0.5, f_2 = 0.3, f_3 = 0.6I$, and $\langle h\nu \rangle = 20$ eV, making $m_p \phi \approx 3000I$ photons per proton. (A slight correction due to the dilution by primordial helium will be introduced shortly.) The uncertain parameter $I$ depends on the initial mass function (IMF). It is approximately 1 for the solar neighborhood IMF and increases with any weighting of the ZAIMF toward more massive stars.

In what follows it will be assumed that the primordial ratio of helium to hydrogen nuclei was 0.1. The gas nuclei number density is $n = n_{\rm H} + n_{\rm He} = 1.1 n_{\rm H}$, whereas the gaseous mass density is $\rho_g = 1.4 m_p n_{\rm H} = (1.4/1.1) m_p n$. It will also be assumed that both hydrogen and helium are singly ionized, making $n_e$ equal to $n$, and that the recombination coefficient (to $n \geq 2$) $\alpha \approx 2 \times 10^{-13}$ cm$^3$ s$^{-1}$ applies to both. (This coefficient is slighly lower than that for a present-day H II region because the temperature is somewhat higher.) Thus the recombination rate per unit volume is $\alpha n^2$, and the critical rate of proton flow into star formation is $\alpha n^2/(m_p \phi)$, providing one ionization per recombination. Then we have (including helium)

$$\dot{\rho}_* = \frac{1.4 \alpha n^2}{\phi} = \left[ 1.4 \left( \frac{1.1}{1.4} \right)^2 \frac{\alpha}{\phi m_p^2} \right] \rho_g^2 .$$

For a volume-average rate, the right-hand side must be divided by the square of the filling factor $f_0$ to get the local density from the average density, and then multiplied by the filling factor to count only the active volume. The net result is that the volume-average star formation rate and gas density are related by

$$\dot{\rho}_* = A \rho_g^2 ,$$

where

$$A = \left( \frac{1.1}{1.4} \right)^2 \frac{1.4}{f_0} \frac{\alpha}{\phi m_p^2} .$$

Two further points are made in Paper I. The first is that this rate probably represents an extreme upper limit to quasi-steady formation of stars. The inclusion of any other bottlenecks should provide a slower rate. The second is that when this rate is compared numerically to the rate which Larson (1969) found was needed to provide acceptable models for elliptical galaxies, the two are equal within uncertainties.

The remainder of this paper takes the first steps toward incorporating this specific NISM and star formation model into a scenario for the formation epoch of spiral galaxies. These steps, numbered by section, include the following:

II. Evaluation of the structure, star formation time scale, and luminosity of a self-gravitating isothermal disk.

III. Recognition of the importance of radiation pressure in disks of interest and a calculation of its magnitude, particularly that due to trapped Lyman-$\alpha$. Discussion of the nature of Lyman-$\alpha$ pressure and its effects on the disk structure.

IV. Discussion of the possibility that Lyman-$\alpha$ might seriously affect the infall rate, and determination of the properties of a halo supportable by Lyman-$\alpha$.

V. Presentation of a first-order infall scenario and the time-dependent properties of the system it would construct.

VI. Evaluation of disk properties at the epoch at which further material is supportable against infall by Lyman-$\alpha$ pressure. This includes determination of the two-dimensional family of disk galaxies which can arise from the three-parameter set of perturbations in the Hubble flow.

VII. Consideration of the fate and subsequent role of a halo supported by Lyman-$\alpha$.

VIII. Presentation of a summary, caveats, promise, and an indication of directions needing further effort.

The four frustrations with which this introduction began are dealt with as follows: the feedback problem is investigated only after the NISM structure problem is eliminated by assumption, the star formation problem by derivation, and the infall problem by naïveté. The feedback result turns out to be significant in that it supports the underlying thesis of this paper, namely: *The eventual properties of spiral galaxies can be controlled by self-regulating physical processes occurring during formation.*

In systems like those about to be explored, the supernova rate would be quite high. These explosions may have considerable impact on the state of the NISM, on the star formation rate, and on the interaction of the disk as a whole with infalling or halo material. In this initial investigation, however, all such effects (but one) are completely omitted. The one effect mentioned from time to time is the possibility that supernovae would maintain a hot matrix (i.e., $T \sim 5 \times 10^5$ K) interspersed through the denser H II phase. This inclusion allows for pressure continuity in a system with gaseous filling factor less than 1, and in addition provides low-opacity channels for photon propagation. All further consideration of supernovae

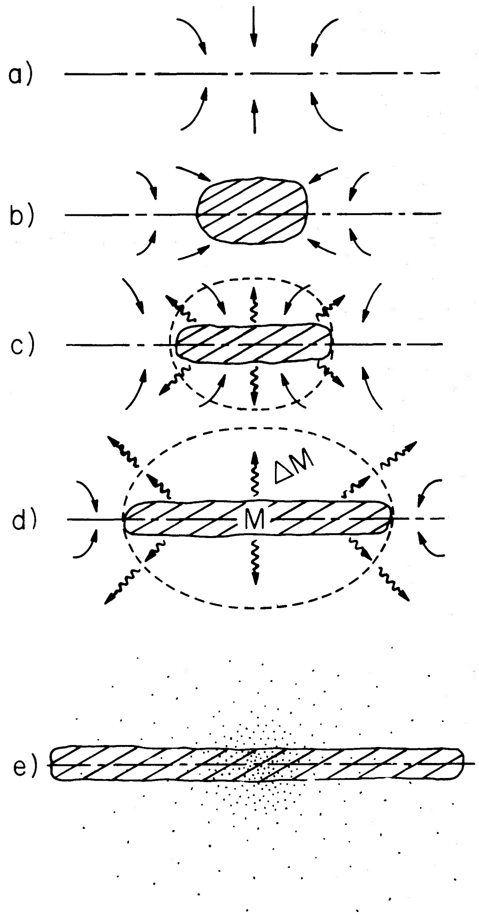


FIG. 1.—The general properties of feedback-controlled models of disk galaxy formation. (*a*) Infalling material converges toward the plane. (*b*) Owing to externally induced shear in the flow field, the convergence is to a disk. The disk material seeks its equation of state, phases, distribution, and star formation rate. (*c*) The disk grows in surface density and radius (higher angular momentum falls in progressively later, although continually accompanied by lower angular momentum material). The scale height shrinks with increased self-gravity. Star formation and evolution contribute to an outflow of energy through the infalling material. At early times there is more material in the "near halo" (*dashed outline*) than in the disk, but the ratio decreases with time. (*d*) At this epoch, the mass in the near halo, $\Delta M$, has dropped to the point where it can be hydrostatically supported by the pressure associated with the outwardly diffusing energy. Infall continues at the outer edge, but for the most part a quasi-static disk-halo system is established. It must, however, evolve dramatically as the disk exhausts its gas supply by star formation. (*e*) A halo which is shattered to stellar (and smaller) fragments at least as rapidly as the much denser disk will lead to the familiar system shown. Such rapid fragmentation of the halo almost certainly requires inducement by energy leaving the disk.

is postponed to a subsequent but parallel investigation of disk galaxy formation with a supernova-dominated NISM.

It should be clear at the outset that the model presented is a prototype rather than a final product. It is the simplest of a class envisioned to incorporate feedback, via physical mechanisms of interstellar material, into the description of galaxy formation. The general features of the class are summarized in Figure 1.

## II. THE SELF-GRAVITATING ISOTHERMAL DISK

Assuming that infall is taking place gradually compared with the vertical dynamical time scale of the disk, we can straightforwardly evaluate the disk structure. For systems of interest, mass will remain primarily gaseous during the period of model applicability, but a factor $f_g$, the gaseous fraction, is introduced at the outset to maintain a generality needed much later. It is assumed, however, that the stars when they do form have a scale height similar to that of the gas (although a detailed model would as usual predict a smaller scale height for the stars).

Conditions in an ionized, isothermal, self-gravitating disk with total column density $\sigma$, gaseous column density $\sigma_g = f_g \sigma$, volume filling factor $f_0$ for the H II phase, and average nuclear mass $m = (1.4/1.1)m_p$ are given by

$$n = f_g \bar{\rho}/f_0 m\ , \quad p = 2nkT\ ,$$

$$p/p_0 = \rho/\rho_0 = \mathrm{sech}^2\,(z/H)\ , \quad p_0 = \pi G \sigma \sigma_g/2 = \pi G \sigma^2 f_g/2\ ,$$

$$n_0 = \pi G^2 f_g/4kT\ , \quad H = 2kT/\pi G m \sigma f_0\ ,$$

$$\sigma = 2H\bar{\rho}_0 = 2Hmf_0\, n_0/f_g\ , \quad g_{\max} = 2\pi G \sigma\ ,$$

where $H$ is the scale height, $g_{\max}$ is the gravitational acceleration for newly accreting material and $\bar{\rho}$ is the horizontally averaged total mass density at height $z$. Values at midplane are given the subscript zero. (The hot matrix phase lends continuity to the pressure with no contribution to the mass density.)

The star formation time scale at midplane will be designated $t_m$ and is given by

$$t_m = \frac{\rho_g}{\dot{\rho}_*} = \frac{m_p \phi n_{\mathrm{H}}}{\alpha n^2} = \frac{1}{1.1}\,\frac{m_p \phi}{\alpha n_0}$$

$$\approx \frac{4.3 \times 10^8 I}{n_0(\mathrm{cm}^{-3})}\ \text{years}$$

$$= \frac{1}{1.1}\,\frac{m\phi 4kT}{\alpha G \sigma^2 f_g}\ .$$

It is clear that the instantaneous recycling approximation is not threatened for $n_0 < 100\ \mathrm{cm}^{-3}$, since $t_m$ is then longer than $3 \times 10^6$ years, but note the inverse quadratic dependence on $\sigma$.

Let us now explicitly relate the recombination rate to the total emissivity $\eta$ of the disk in starlight. This emissivity arises primarily in the same massive stars that supply ionizing UV. We assumed in § I (following Paper I) that 30% of the luminosity of massive stars emerged beyond the Lyman limit. The average energy of these photons is $\langle h\nu \rangle \sim 20$ eV, and each such photon balances one recombination. Thus $\eta$ is related to the recombination rate by $0.3\eta/20\ \mathrm{eV} = \alpha n^2$, or

$$\eta \approx (67\ \mathrm{eV})\alpha n^2\ ,$$

i.e., the starlight production rate is 67 eV per recombination. One might wish to round up to 70 eV or 100 eV as time goes on, to include the accumulation of lower mass stars.

The direct Lyman-$\alpha$ emissivity of the gas due to recombination of hydrogen to levels with $n \geq 2$ is $(10.2\ \mathrm{eV})\alpha n n_{\mathrm{H}}$, or $(9.3\ \mathrm{eV})\alpha n^2$. This may be augmented quite a bit by collisional excitation, since the remaining energy of ejected electrons (roughly 5 eV on average) is most likely dissipated by excitation of Lyman-$\alpha$ in a gas with low "metal" abundance. (At least 4 eV of the 20 eV available per ionization must, however, emerge in Balmer and higher continua and lines.) The total Lyman-$\alpha$ emission per recombination (of either hydrogen or helium) can thus be written $(\frac{3}{4})I_{\mathrm{H}}\psi$, where the factor $\psi$, the mean number of Lyman-$\alpha$ photons generated per recombination, is such that the product lies between 9.3 and about 15 eV. Then

$$\eta(\mathrm{Ly}\alpha) = (\tfrac{3}{4} I_{\mathrm{H}}\,\psi)\alpha n^2\ .$$

The luminosity per unit area, $l$ (total) or $l_\alpha$ (Ly$\alpha$), follows from integrating these emissivities over $z$. The key ingredient from the disk structure is

$$\int_{-\infty}^{\infty} n^2(z)dz = \tfrac{4}{3}n_0^2 H .$$

Recalling that 67 eV $= \langle h\nu\rangle/f_2$,

$$l = \frac{4}{3} H\alpha n_0^2 f_0 \frac{\langle h\nu\rangle}{f_2} = \frac{\pi G\alpha}{6kTm} f_g^2 \sigma^3 \frac{\langle h\nu\rangle}{f_2} .$$

Similarly,

$$l_\alpha = \frac{\pi G\alpha}{6kTm} f_g^2 \sigma^3 \left(\frac{3}{4} I_{\rm H}\psi\right) .$$

Notice the extreme sensitivity of the surface brightness to $\sigma$. The mass-to-light ratio is proportional to $\sigma^{-2}$. Two further things are noteworthy: The uncertain $f_0$ has canceled out, and the Lyman-$\alpha$ luminosity is a significant fraction, perhaps 20%, of the total.

The Paper I NISM model thus unambiguously specifies the structure, star formation rate, and luminosity per unit area as functions of the disk surface density. We are now prepared to search for consequences of those properties.

## III. RADIATION PRESSURE IN THE DISK

### a) *Initial Survey*

One searching for radiative feedback is naturally interested in properties of the disk when it reaches a column density corresponding to the Eddington limit, that is, when $A_{\rm T} l/2c = mg_{\rm max} = 2\pi G\sigma m$, where $A_{\rm T}$ is the Thomson cross section. The result, however ($\sigma \sim 0.4$ g cm$^{-2}$), is beyond the range of applicability of the model. (The instantaneous recycling approximation fails.) A somewhat greater opacity is offered by neutral hydrogen (absorbing Lyman continuum), but the limiting surface density is still too high to be of interest.

Another interesting possibility is that a disk emitting primarily O star radiation might behave like an O star itself, with a radiatively driven wind. A single isotropic scattering of each emerging photon would provide a net outward force per unit area (a radiation pressure field distributed through the scattering volume like a body force).

$$p_{\rm rad} = \frac{l}{4c} = f_g^2 \frac{\pi G\alpha\sigma^3}{24mc}\frac{\langle h\nu\rangle}{f_2 kT} \sim \frac{10}{3} f_g^2 \frac{\pi G\alpha\sigma^3}{mc} .$$

Such a pressure would become important to disk stability when comparable to $p_0 = \pi G\sigma^2 f_g/2$. This occurs when

$$\sigma \sim \frac{1}{f_g}\frac{3mc}{20\alpha} \sim \frac{0.05 \text{ g cm}^{-2}}{f_g} \sim \frac{225\ M_\odot \text{ pc}^{-2}}{f_g} ,$$

$$H \sim \frac{f_g}{f_0}\frac{40\alpha kT}{3\pi G m^2 c} \sim \frac{f_g}{f_0}(42 \text{ pc}) .$$

Clearly, with $f_0 \sim \frac{1}{2}$ and $f_g \sim 1$, one would arrive at a familiar-looking galactic surface density and scale height at just the epoch that the galaxy becomes highly repulsive.

A detailed mechanism for photon scattering is absent from the above picture. However, since a significant fraction of the total luminosity is converted to highly trapped Lyman-$\alpha$, whose scattering one can hope to model, we now proceed to that specific example.

### b) *The Optical Depth to Lyman-$\alpha$ Scattering*

Assessing the dynamical importance of Lyman-$\alpha$ requires evaluation of its radiation pressure gradient. In a highly scattering system, radiation pressure is essentially isotropic and is equal to one-third of the energy density. Energy density, in turn is approximately the product of emissivity and average photon trapping time. Finally, the trapping time follows from solution of the diffusion problem, whose input must be system properties including scattering optical depth and line width.

In order to estimate the optical depth for Lyman-$\alpha$ scattering, we must be somewhat more specific about the distribution of material and its ionization fraction. A Strömgren radius (at densities of interest) is roughly a factor of 5 smaller than the scale height $H$, so that the optical depth of a half-plane is roughly that of five radial lines of sight through H II regions plus any intervening neutral material. An H II region has a neutral fraction of order $10^{-4}$ and a Lyman continuum optical depth of roughly unity. At $10^4$ K the optical depth of Lyman-$\alpha$ at line center is $10^4$ times the continuum optical depth, i.e., $10^4$ per H II region radius, or $5 \times 10^4$ for the H II region portion of the line of sight in the half-plane. This optical depth through photoionized regions, although large, is utterly negligible compared with the potential contribution of neutral material between H II regions, unless that material comprises less than $10^{-4}$ of the total gaseous mass.

Thus we must attempt to estimate the ratio of interstitial neutral gas to gas within ionized regions. (We have already assumed in § I that the ionized gas contains most of the mass, if not the Lyman-$\alpha$ opacity.) A formal estimate of the amount of neutral material follows from the requirement that it be the location of star formation responsible for keeping the rest ionized. We can write the total star formation rate as $M_g/t_m$, where $M_g$ is the total gaseous mass and $t_m$ is the star formation time scale derived in § II. Equating this required rate with the available rate, written $M_n/t_n$, where $M_n$ is the neutral mass between H II regions and $t_n$ is the mean time for neutral material to convert to stars, the interstitial mass fraction is then

$$\frac{M_n}{M_g} \approx \frac{t_n}{t_m} ,$$

where the NISM model tells us that $t_m \sim 4.3 \times 10^8\ I$ years/$n_0$(cm$^{-3}$). If gravitational free fall were responsible for star formation, then $t_n \sim t_{\rm ff} = 4.3 \times 10^7$ years/$[n_0({\rm cm}^{-3})]^{1/2}$ (e.g., Spitzer 1968) and $M_n/M_g \sim [n_0({\rm cm}^{-3})/100I^2]^{1/2}$. The implied neutral fraction is fairly high at all densities of interest (i.e., 8–30 cm$^{-3}$, as we shall see). More likely, however, neutral material is compacted by H II region expansion, stellar wind ram pressure, and supernova explosions, so that $t_n$ may be roughly equal to the evolution time of nearby massive stars. In that case, $t_n \sim 4 \times 10^6$ years and $M_n/M_g \sim n_0({\rm cm}^{-3})/100I$. For the highest densities of interest ($\sim 30$ cm$^{-3}$ with $I \sim 3$), the neutral mass between H II regions could still be almost 10% of the total. In general the implied Lyman-$\alpha$ optical depth ranges from about $10^7$, when Lyman-$\alpha$ first becomes important (§ IV), to $2 \times 10^8$. Finally, as an extreme lower limit to $t_n$, consider that at a density of order 10 $\mathrm{cm}^{-3}$, the radius of a region containing 100 $M_\odot$ is of order 4 pc. Its crossing time at a velocity of 10 km $\mathrm{s}^{-1}$ is $4 \times 10^5$ years, making this approximately the minimum time to crush such material into a star. With $t_n \sim 4 \times 10^5$ years, the neutral mass fraction is reduced by a factor of 10 from the last estimate above, making the highest optical depth of interest $2 \times 10^7$.

The last estimate, however, brings up an additional complication. Although on average each Strömgren sphere must see the formation of one massive star during its lifetime, with such a short $t_n$, only one in 10 would be driving such formation (on a fraction of its periphery) at any given time. The distribution of neutral material would thus be extremely patchy, and it is possible that the patches would not have the opportunity to interact with a large proportion of the diffusing Lyman-$\alpha$.

The results of this section are admittedly not very satisfying. We are left knowing only that $\tau_0$, the half-plane optical depth at line center for scattering Lyman-$\alpha$, probably lies in the range $10^6$–$10^8$, with a cross section profile corresponding to roughly $10^4$ K. Fortunately this knowledge turns out to be sufficient to first order, and we now turn to the diffusion problem at high optical depth.

### *c) Lyman-α Diffusion*

The diffusion of a Lyman-$\alpha$ photon out of its parent cloud and the disk as a whole will consist of a large number of scatterings with short mean free path (involving little time and little progress) followed by a small number of scatterings in the wing of the line, with larger mean free paths (much more time between scatterings and much more progress) (e.g., Osterbrock 1962; Walmsley and Mathews 1969; Adams 1972, 1975). The total trapping is approximately linear in the scale, and fortunately the ratio of trapping time to scale is only weakly dependent on the line-center optical depth $\tau_0$. Adams (1975) calculated trapping times versus $\tau_0$ for the case of a homogeneous plane with photons originating at midplane. The results (shown in his Fig. 1) imply typical photon paths of about $15H$ (trapping times of $15H/c$) for $\tau_0 \lesssim 10^6$ and paths of $13H(\tau_0/10^6)^{1/3}$ for $\tau_0 \gtrsim 10^6$. The inclusion of extreme inhomogeneity in both source (recombination) and scatterer distributions, velocity structure, and the dispersed distribution of the original photon source throughout the disk can be expected to lower the average trapping time somewhat for our galaxy disk model. The effects are small, however, the bulk of progress being made far in the line wing where detail is insignificant. Considering Adam's results as an upper limit and recognizing the extreme insensitivity of the result to $\tau_0$, the trapping time written as

$$t = (15H/c)\delta$$

gives an adequate representation of the situation. The fudge factor $\delta$ is expected never to deviate by more than a factor of 3 from unity.

We will soon be considering a situation for which Lyman-$\alpha$ pressure is comparable to or even greater than thermal pressure. Krolik (1979, 1983) has suggested that Lyman-$\alpha$ pressure in this regime will drive an instability tending to fragment the material down to a scale such that the line-center optical depth is only moderate per fragment, releasing photons more readily into optically thin crevices. A hot matrix generated by supernovae may also provide macroscopic channels for photon escape. At the smallest scale, the crevices increase the photon mean free path very little if the fragment filling factor remains large. Only if channels at each scale lead sequentially to ones of larger scale can this microcracking lead to an increased escape rate. A rough approximation to such a system is one-dimensional diffusion along a line for which the mean free path increases exponentially with progress achieved. Taking the smallest and largest mean free paths to be $H/\tau_0$ and $H/2$, respectively, I estimate the trapping time to be roughly $q^2$ ln $(\tau_0/2)H/c$, where $q$ is the number of sequential steps in the same direction between doublings of the mean free path. For $\tau_0 \gtrsim 10^6$ and $q \gtrsim 1$, $t > 15H/c$. Thus crack-augmented diffusion seems not to invalidate the previous conclusion regarding the trapping time. On the other hand, in enhancing escape of radiation close to line center, it could significantly alter the frequency profile of emerging flux (backfilling the very dark line center found by Adams), and hence its ability to interact with low-velocity material at high $z$.

### *d) Conditions in the Disk when Lyman-α Pressure Becomes Significant*

The Lyman-$\alpha$ pressure at midplane, one-third of the energy density, is (for a trapping time $15H\delta/c$)

$$p(\mathrm{Ly}\alpha) = \frac{1}{3} f_0 \eta_0(\mathrm{Ly}\alpha) \frac{15H\delta}{c} = \frac{15\delta}{4c} l_\alpha \sim 3\delta \frac{l}{4c} .$$

Since the result is roughly 3 times the previously estimated total single-scattering radiation pressure, we can expect it to become important at an interesting (i.e., galaxy-like) column density. Other sources of opacity are not likely to alter the total radiation pressure significantly.

The radiation pressure becomes important to the disk structure when $p(\mathrm{Ly}\alpha)$ reaches $p_0$, and the parameters of the disk when this occurs are referred to as "critical" in subsequent sections of this paper. Characteristic values are given the subscript 1. At the critical epoch, the disk parameters are

$$\sigma f_g = \sigma_1 \equiv \frac{4kT}{(3/4)I_\mathrm{H}\psi} \frac{mc}{5\delta\alpha} \approx \frac{0.0146 \text{ g cm}^{-2}}{\delta} = \frac{70\ M_\odot \text{ pc}^{-2}}{\delta} ,$$

$$\frac{f_0 H}{f_g} = H_1 \equiv \frac{5\delta\alpha[(3/4)I_\mathrm{H}\psi]}{2\pi G m^2 c} \approx 138\delta \text{ pc} ,$$

$$f_g n_0 = n_1 \equiv \frac{4\pi G m}{25\delta^2\alpha^2} \frac{kTmc^2}{[(3/4)I_\mathrm{H}\psi]^2} \approx \frac{8.1 \text{ cm}^{-3}}{\delta^2} ,$$

$$\frac{t_m}{f_g} = t_{m1} \equiv \frac{1}{1.1} m_p \phi \frac{[(3/4)I_\mathrm{H}\psi]^2}{4kTmc^2} \frac{25\alpha\delta^2}{\pi G m}$$

$$\approx 5.33 l \delta^2 \times 10^7 \text{ years} ,$$

$$p_0 f_g = 2n_1 kT = p_1 \approx \frac{2.4 \times 10^{-11}}{\delta^2} \text{ dyn cm}^{-2} .$$

Expressing the recombination coefficient $\alpha$ in fundamental constants, the critical surface density can also be written (Rybicki 1983)

$$\sigma_1 \approx \frac{0.28}{\delta} \frac{m}{A_\mathrm{T}} \left(\frac{kT}{I_\mathrm{H}}\right)^{3/2} \sim 0.006 \frac{m_p}{A_\mathrm{T}} ,$$

where $A_T$ is once again the Thomson cross section $(8/3)\pi r_0^2 = 8\pi e^4/3(m_e c^2)^2$. Electron-scattering optical depths of order $10^{-2}$ are implied. Similarly,

$$\begin{aligned} H_1 &\approx \frac{8}{3(0.28)}\left(\frac{I_H}{kT}\right)^{1/2} \frac{e^2}{Gm^2} \frac{e^2}{m_e c^2}\left(\frac{e^2}{\hbar c}\right)^2 \delta \\ &\sim 23.3 \frac{e^2}{Gm_p^2} r_0 \left(\frac{e^2}{\hbar c}\right)^2 \delta \\ &\sim 23.3 \frac{m_e}{m_p} \alpha_{fs}^2 N r_0 \delta , \end{aligned}$$

where $\alpha_{fs}$ is the fine structure constant, $r_0$ the classical radius of the electron, and $N$ the "large number" $e^2/Gm_p m_e \sim 2 \times 10^{39}$.

The critical epoch at which Lyman-$\alpha$ pressure reaches $p_0$ is thus characterized by a column density whose magnitude depends only on $m_e$, $m_p$, $e$, and $c$. It knows nothing of perturbations in the universe, galaxies, or even gravity. The scale height of that material depends also on $\hbar$ and $G$ in such a way that it is of order 100 pc, also with no prior knowledge of the nature of galaxies. Both results are sufficiently attractive that it is tempting to suppose that Lyman-$\alpha$ pressure simply halts all further infall when this state is reached. The situation is more complex than that, however, and its discussion is postponed to the next two sections.

### e) *Three Details*

One may wonder about the simple equation of $p(\text{Ly}\alpha)$ and $p_0$ for determining the critical epoch. In examining Adam's (1972) graph of the source function versus optical depth, it seems reasonably convincing that the radiation pressure gradient is of order $p(\text{Ly}\alpha)/H$. In the self-gravitating plane, comparing this gradient with $\rho g$ is equivalent to comparing $p(\text{Ly}\alpha)$ with $p_0$.

The presence of a small amount of dust in this nominally pristine system could seriously deplete Lyman-$\alpha$ pressure and invalidate the results. This difficulty is present over only a narrow range in dust abundance, however, since making the disk optically thick to continuum radiation subjects it to the full single-scattering radiation pressure $l/4c$, effectively reducing $\delta$ only to $\frac{1}{3}$.

The Lyman-$\alpha$ absorption cross section on dust, at solar neighborhood abundances, is roughly $10^{-21}$ cm$^2$ per hydrogen atom (Savage 1983). Since the Lyman-$\alpha$ path is $15H$, its traversed column density is roughly $15 f_g \sigma/2 = 7.5\sigma_1$, approximately $5 \times 10^{22}$ atoms cm$^{-2}$ when Lyman-$\alpha$ becomes significant. Dust absorption is therefore neglible at this epoch only if its abundance is less than about $10^{-2}$ of that in the solar neighborhood. For a dust abundance greater than $10^{-1}$ solar, the continuum opacity in the optical is important at $\sigma \sim 0.05$ g cm$^{-2}$, the critical column density found for single scattering.

The amount of heavy-element contamination present in a galactic disk when it reaches $\sigma_1$ depends on the length of time infall requires to provide that column density and the elemental yields of stars with lives shorter than that time. Even with a modest elemental enrichment, however, it is unclear whether UV absorbing dust would have time to form. Suppose, for example, that it forms in the atmospheres of red giant carbon stars. Throughout the remainder of this paper, it will be assumed that infall was rapid (compared, say, to $t_{m1} \sim 5I\delta^2 \times 10^7$ years) and dust insignificant. The converse situation may well merit further investigation.

Notice finally that as $\sigma$ approaches $30\sigma_1$, Lyman-$\alpha$ would have a high probability of escape via incoherent electron scattering into the far wings of the line (Adams 1983). With these caveats in mind, let us now turn to the effects of the Lyman-$\alpha$ trapping on the structure of the disk.

### f) *Disks with Higher Column Density*

For column densities above the critical value, the calculated Lyman-$\alpha$ pressure in the initial model exceeds the thermal pressure. Two competing possibilities emerge. Either Lyman-$\alpha$ pressure is able to rearrange material such that escape is enhanced, leading back naturally to equipartition, or it finds no rearrangement that will substantially reduce the escape rate, in which case Lyman-$\alpha$ assumes responsibility for support of the disk, thermal pressure becoming insignificant. In either case, one expects a Rayleigh-Taylor–like instability but of a significantly modified character (e.g., Krolik 1979). The two chief differences from the classical situation are that the lighter photon fluid is continously resupplied inside the densest portions of the denser fluid, and there is considerable interpenetration of the two fluids. The cracking described in the last section is expected, but the weak gradient in pressure does not lead to stable condensations. Lumps continually form, dissipate from their own Lyman-$\alpha$ emission, and reform, the system behaving as if opalescent. If the earlier calculation of crack-augmented diffusion is correct, the net escape rate is altered very little, and a disk supported by Lyman-$\alpha$ is the inevitable consequence.

Rather than do the radiative transfer problem in a radiation-supported disk, I shall simply continue to equate $p(\text{Ly}\alpha)$ to the midplane pressure, still necessarily $p_0 = \pi G \sigma^2 f_g/2$. Other surviving relationships are $\sigma = 2H\bar{\rho}_0$ and $n = f_g \bar{\rho}/f_0 m$. In calculating the total luminosity, I shall also assume that

$$\int_{-\infty}^{\infty} n^2 dz \approx \tfrac{4}{3} n_0^2 H$$

continues to be a reasonable approximation. By continuing the assumption that the Lyman-$\alpha$ trapping time is $15\delta H/c$, while recognizing that a slightly lower value of $\delta$ may be appropriate for a radiation-dominated plane, the properties of the Lyman-$\alpha$ supported disk are then

$$H = \frac{f_g H_1}{f_0} \approx \frac{f_g \delta}{f_0} (138 \text{ pc}) ,$$

$$x \equiv \frac{\sigma}{\sigma_1} \approx \frac{\sigma\delta}{0.0146 \text{ g cm}^{-2}} = \frac{\sigma\delta}{70\ M_\odot \text{ pc}^{-2}} ,$$

$$n = n_1 x \approx \frac{8.1 \text{ cm}^{-3}}{\delta^2} x ,$$

$$p = f_g p_1 x^2 \approx 2.4 \times 10^{-11} \frac{f_g x^2}{\delta^2} \text{ dyn cm}^{-2} ,$$

$$t_m = \frac{t_{m1}}{x} \approx 5.33 \times 10^7 \frac{I\delta^2}{x} \text{ years} ,$$

$$t_d = \left(\frac{2H}{g_{\max}}\right)^{1/2} = \left(\frac{1}{2\pi G\bar{\rho}}\right)^{1/2} = 1.2 \times 10^7 \left(\frac{f_g}{f_0}\frac{\delta^2}{x}\right)^{1/2} \text{ years} ,$$

$$\frac{l_\alpha}{c} = \frac{4p_0}{15\delta} = \frac{2\pi G\sigma^2 f_g}{15\delta} = 6.4 \times 10^{-12} \frac{f_g x^2}{\delta^3} \text{ dyn cm}^{-2} .$$

The applicability criterion is $x > 1/f_g$. *The net result of Lyman-$\alpha$ pressure is to stabilize the scale height so that there-*

*after the density is simply proportional to* $\sigma$. The rate of increase of $l_\alpha$ and decrease of $t_m$ with $\sigma$ are also lessened. The above equations assume that thermal pressure is negligible. The actual ratio of thermal to total pressure is approximately $x^{-1}$, a nicety which will continue to be neglected, even though we shall find that the validity of the model is severely strained for values of $x$ greater than about 10.

### IV. RADIATION PRESSURE IN THE HALO

The Lyman-$\alpha$ luminosity per unit area of the disk, in equilibrium, is $l_\alpha = (4c/15\delta)p_0$. If the scale height is perturbed from its equilibrium value, $l_\alpha$ is inversely proportional to $H$, while $p_0$ is unchanged. Thus Lyman-$\alpha$ pressure is able to raise the bulk of material in the disk only to a height $H_1$. As more material is added, nothing about this conclusion changes. Increased radiation pressure will not blow away an excess surface density (excess compared with $\sigma_1$).

A separate question is whether Lyman-$\alpha$ pressure can blow away material that is not yet a part of the disk. Consider, for example, the possibility that Lyman-$\alpha$ is retrapped in a planar layer of scale $z_H > H_1$ in material with very large velocity dispersion. In that case, the already broad profile will quickly be redistributed over the effective Doppler width of this layer and must diffuse to get out. The trapping time in that case will be roughly 2 $(\ln \tau_H)z_H/c$ $(\tau_H < 10^6)$, where $\tau_H$ is the effective line-center optical depth of the layer. The trapped energy per unit area (per side) is $l_\alpha$ $(\ln \tau_H)z_H/c$, and the pressure $p_H = l_\alpha (\ln \tau_H)/(3c) = 4p_0 (\ln \tau_H)/45\delta = f_g \sigma g_{\max} (\ln \tau_H)/45\delta$, where $f_g$ and $\sigma$ refer to the values in the disk. The radiation force per gram is $p_H/\rho_H z_H = p_H/\Delta\sigma$, which exceeds $g_{\max}$ only for $\Delta\sigma < f_g \sigma (\ln \tau_H)/45\delta$, rather less than $\sigma$ itself but independent of $z_H$. Thus a few percent of $\sigma$ can be undergoing active expulsion at any given time. The cumulative effect, however, depends on the time required for expulsion. The maximum amount of expellable material in time $t_m$ is[1]

$$\sigma_e = \frac{p_H t_m}{v_e} = \frac{f_g \sigma t_m g_{\max} \ln \tau_H}{45\delta v_e} ,$$

where $v_e$ is the escape velocity. The product $\sigma t_m = \sigma_1 t_{m1}$ is constant while $g_{\max} = 2\pi G\sigma$. Thus

$$\frac{\sigma_e}{\sigma} \sim \frac{2\pi G\sigma_1 t_{m1} \ln \tau_H}{45\delta v_e} \sim \frac{2.3I \ln \tau_H \text{ km s}^{-1}}{v_e} .$$

Clearly the ratio of expellable mass to total disk mass is small. A similar conclusion applies to the stopping of rapidly infalling material. The fact that rapidly infalling material cannot be stopped anyway means that one need not worry about the ability of Lyman-$\alpha$ to interact with a severely shifted absorption profile.

As we shall see, what the Lyman-$\alpha$ pressure can do well is support (rather than expel) a rather considerable halo of material in the vicinity of the disk. Rapidly infalling material will enter, but the much larger amount of material in the vicinity, which would enter in time were it not for Lyman-$\alpha$ pressure, can be held off for a time of order $t_m$. Lyman-$\alpha$ can also considerably influence the state of that material, drastically altering its eventual fate.

To illustrate the above point with the simplest possible model of Lyman-$\alpha$ trapping in the halo, assume merely that once the Lyman-$\alpha$ has diffused out of the disk (with trapping time $15\delta H/c$, it is further trapped in the halo. There, owing to the spherical geometry, the trapping time is written in terms of $R$ (rather than $H$) and reduced by about a factor of 2, i.e., it is $15\delta_H R/(2c)$ (Adams 1975), where a new fudge factor $\delta_H$ once again allows for later improvement. In particular, I assume that an annulus of the disk, from $R$ to $R + dR$, contributes energy

$$dU = \frac{15R\delta_H}{2c} 2\pi R l_\alpha dR$$

uniformly to the volume interior to $R$, making its contribution to the pressure there

$$dp = \frac{15}{4}\frac{\delta_H}{c} l_\alpha \frac{dR}{R} .$$

Similarly (but not quite consistently), the annulus contributes an energy

$$dU = \frac{15\delta_H r}{2c} 2\pi R l_\alpha dR$$

to the volume within $r > R$, making the pressure contribution at $r$

$$dp = \frac{15}{4}\frac{\delta_H}{c}\left(\frac{R}{r}\right)^2 l_\alpha \frac{dR}{R} .$$

After integrating from $R = 0$ to $r_e$, the disk radius, and differentiating with respect to $r$, the pressure gradient is

$$\begin{aligned}\frac{\partial p}{\partial r} &= -\frac{15}{4}\delta_H \frac{l_\alpha}{c}\frac{1}{r_e}\left(\frac{r_e}{r}\right)^3 , \quad r > r_e \\ &= -\frac{15}{4}\delta_H \frac{l_\alpha}{c}\frac{1}{r} , \qquad r < r_e .\end{aligned}$$

Roughly speaking, the energy density is enhanced over the free-streaming value by the factor $15\delta_H/2$ at all radii.

The supportable mass density in the halo is given by $-\partial p/\partial r = \rho g = \rho GM(r)/r^2$. For $r < r_e$, the implication is that the halo mass within $r$ is directly proportional to the disk mass within $r$. The disk mass is $M_D = \pi r^2\sigma$, the halo mass can be written as $M_H = \beta\pi r^2\sigma$, and the halo density is $\rho = (dM_H/dr)/4\pi r^2 = \beta\sigma/(2r)$. The hydrostatic condition is then

$$\frac{15}{4}\delta_H \frac{l_\alpha}{c}\frac{1}{r} = \frac{\beta\sigma}{2r} G(1 + \beta)\pi\sigma ,$$

or

$$\beta(1 + \beta) = \frac{15}{2\pi}\frac{\delta_H}{G\sigma^2}\frac{l_\alpha}{c} = \frac{\delta_H f_g}{\delta} .$$

Thus, for $\delta_H f_g/\delta \approx 1$, $\beta \approx 0.62$; the supportable mass in the near halo $(r < r_e)$ is some 62% of the disk mass.

Extending the analyis beyond $r = r_e$, we have

$$\frac{15}{4}\delta_H \frac{l_\alpha}{c}\frac{1}{r_e}\left(\frac{r_e}{r}\right)^3 = \frac{1}{4\pi r^4} GM \frac{dM}{dr} ,$$

which yields

$$M^2 = (1 + \beta)^2 M_D^2 + 2M_D^2 \frac{\delta_H}{\delta} f_g\left(\frac{r^2}{r_e^2} - 1\right) ,$$

[1] The mass conversion time scale is the length of time the Lyman-$\alpha$ can remain bright.

where $M_D = \pi r_e^2 \sigma$. For $r^2 \gg r_e^2$, the system mass increases linearly with $r$. The rotation curve is flat. For example, assuming $\delta_H f_g/\delta = 1$, $\beta = 0.62$, the rotation velocity at $r = r_e$ (neglecting the deviation from a spherical potential) is $v_{\rm rot} = 1.27(GM_D/r_e)^{1/2}$, while for $r \gg r_e$ it is $1.19(GM_D/r_e)^{1/2}$.

The details of the above analysis depend on the value of $l_\alpha$ for a radiation-dominated disk. For surface densities below $\sigma_1$, correct results are obtained by using $l_\alpha$ for that regime. At any surface density Lyman-$\alpha$ is capable of supporting some material for a time of order $t_m$. For the radiation-dominated disk, the amount of supportable material within $r > r_e$ is essentially $M_D r/r_e$.

In discovering that very little material is either expellable or stoppable but that a significant amount is supportable, we are led to the supposition that infall will continue until such time as the amount of nearby material in the infall drops to the supportable halo of the then existing disk. Support begins, infall stops, and for the most part the material in the disk turns into stars in a time of order $t_m$, after which the Lyman-$\alpha$ brightness drops precipitously and support of the halo material vanishes.

When halo support fails, the material can be expected to resume its infall. If, however, as will be discussed in § VIII, that material has been forced into dense condensations by Lyman-$\alpha$ pressure during the brief period of support, it will remain as an interpenetrating halo rather than joining the disk. Assuming that this occurs, we can proceed to an infall scenario with a specific mechanism for stopping infall at a particular epoch.

## V. A FIRST-ORDER INFALL SCENARIO

This investigation is concerned with the possibility that individual galaxies suppress their own further growth after reaching total masses much smaller than the mass of the perturbations out of which they are growing. The idea is not that galaxy formation is necessarily inefficient, only that by suppressing further growth of dominant perturbations, slower systems are given time to develop.

A nearly sperically symmetric (and therefore improbable) density enhancement growing in a normal matter-dominated universe can be characterized by three parameters; the minimum density $\rho_{\rm min}$ reached by the central peak before it begins recollapse, the total bound mass $M_P$ of the perturbation, and the external torque or velocity shear to which it is subjected. The minimum-density parameter controls the time of initiation of "galaxy formation," the mass parameter governs the flow rate of material into the "galaxy," and the shear parameter controls the spatial extent of the forming disk.

Concentrating on material very close to the peak of the perturbation (the material the galaxy will eventually admit), conditions are particularly simple. Galaxy formation commences at time $t_c = (3\pi/8G\rho_{\rm min})^{1/2}$ ABB, with core formation. At half that time the mass was at minimum density $\rho_{\rm min}$ and maximum extension. Taking the shear as linear over the central portion of the perturbation, the angular momentum distribution (after azimuthal sharing during the collapse) will be that of a uniformly rotating sphere. Returned material will thus be broadly distributed over the disk with a maximum angular momentum per unit mass $j_{\rm max}$ proportional to the square of the maximum distention $r_{\rm max}$. Hence $j_{\rm max} \propto M^{2/3}$, where $M$ is the returned mass at any epoch and equals $(4/3)\pi r_{\rm max}^3 \rho_{\rm min}$.

By following the perburbation growth at early times, in both velocity and density, and assuming that near the peak the density drops quadratically in radius, the return time for the shell containing mass $M$ can be calculated. The result for short times $\Delta t = t - t_c$ after initiation is $M \approx M_p(2\Delta t/t_c)^{3/2}$. Thus $j_{\rm max} \propto M^{2/3} \propto r_{\rm max}^2(M) \propto \Delta t$.

The disk formed by time $\Delta t$ has mass $M$ and maximum angular momentum per unit mass $j_{\rm max} = v_{\rm rot} r_e \approx (GMr_e)^{1/2} \propto r_{\rm max}^2(M)$, where $v_{\rm rot}$ is the rotation velocity and $r_e$ the disk edge radius. Since $M \propto r_{\rm max}^3$, this relation for $j_{\rm max}$ necessarily implies *that $r_e/r_{\rm max}$ is constant during the infall.* In addition, $r_e/r_{\rm max} \propto (j_{\rm max}/r_{\rm max}^2)^2$, so that the recontraction ratio is proportional to the square of the shear. With this result it is convenient to represent the constant shear parameter by $\epsilon$ defined such that

$$\frac{r_e}{r_{\rm max}} = \epsilon^2 .$$

Since $r_{\rm max}(M) \propto (\Delta t)^{1/2}$ while $M \propto (\Delta t)^{3/2}$, the disk radius $r_e$ and average surface density $\sigma$ are each proportional to $(\Delta t)^{1/2}$ so long as $2\Delta t \ll t_c$ (equivalent to $M \ll M_p$, the case of interest for a self-limiting galaxy).

The characteristic parameters of the unimpededly developing disk as functions of its mass are then

$$r_{\rm max} = (3M/4\pi\rho_{\rm min})^{1/3} = (2MGt_c^2/\pi^2)^{1/3} ,$$

$$r_e = \epsilon^2 r_{\rm max} \propto (\Delta t)^{1/2} ,$$

$$\sigma = \frac{M}{\pi r_e^2} = \frac{1}{\pi\epsilon^4}\left(\frac{4\pi\rho_{\rm min}}{3}\right)^{2/3} M^{1/3} \propto (\Delta t)^{1/2} ,$$

$$v_{\rm rot} = \left(\frac{GM}{r_e}\right)^{1/2} = \frac{1}{\epsilon}\left[GM^{2/3}\left(\frac{4\pi\rho_{\rm min}}{3}\right)^{1/3}\right]^{1/2} \propto (\Delta t)^{1/2} ,$$

$$t_{\rm flow} \equiv \frac{r_e}{v_{\rm rot}} = \left(\frac{r_e^3}{GM}\right)^{1/2} = \frac{\epsilon^3}{[(4/3)\pi G\rho_{\rm min}]^{1/2}} = \frac{2^{1/2}}{\pi}\epsilon^3 t_c ,$$

$$t_{\rm rot} = 2\pi t_{\rm flow} = 2^{3/2}\epsilon^3 t_c ,$$

$$\dot\sigma = \sigma/(2\Delta t) .$$

Note that the flow time scale (or rotation period) is a constant, proportional to $\epsilon^3 t_c$ for the forming galaxy. From the point of view of the galaxy itself, this parameter combination is one of two which govern the eventual structure. The galaxy scale at any value of $\sigma$ follows directly from it:

$$r_e = \pi G\sigma t_{\rm flow}^2 .$$

The other significant parameter is the length of time required to reach some particular value of $\sigma$, and in the next section $\Delta t_1$ is introduced, the time to reach $\sigma_1$ and the Lyman-$\alpha$ domination of the disk. The three parameters of the infall reduce to these two, which govern the nature of the galaxy formed.

It is of some interest to estimate the value of $\epsilon$ that would be appropriate to the Milky Way galaxy, checking on the internal consistency of these results. Assuming a mass and formation time, respectively, of $M \sim 10^{11}\ M_\odot$ and $t_c \sim 10^9$ years implies $r_{\rm max} \sim 45$ kpc. It appears that the mass in our Galaxy never extended farther than the Magellanic Clouds. In addition, for an edge radius now of 10–15 kpc, $\epsilon$ is roughly $\frac{1}{2}$. That translates to a rotation period of $3 \times 10^8$ years or a flow time of $5 \times 10^7$ years. The corresponding value of $\sigma$ is roughly 0.07 g cm$^{-2}$, 300 $M_\odot$ pc$^{-2}$, or $5\sigma_1$.

It is implicitly assumed that material flowing into a galaxy joins the disk quiescently. This point of view is not altered drastically by a cooling-time estimate for a shock-heated infall material. If a halo containing mass $\Delta M \lesssim M$ within radius $r$ is heated to $T \sim GMm/(5kr)$ and has a cooling time $t_{\rm cool} = 3kT/Ln$, where $L$ is the bremsstrahlung cooling coefficient and $n$ is the average halo density, then it can be shown that

$$\frac{t_{\rm cool}}{t_{\rm flow}} = \frac{M}{\Delta M}\frac{r}{R_{\rm cool}} ,$$

where

$$R_{\rm cool} = 0.95\alpha_{\rm fs}\left(\frac{m_e}{m_p}\right)^{1/2} \frac{e^2}{Gm_p m_e} r_0 \approx 35 \text{ kpc} ,$$

where $\alpha_{\rm fs}$ is the fine structure constant and $r_0$ the classical radius of the electron. Thus, near halos with $\Delta M \sim M$ will cool in less than a flow time by bremsstrahlung alone for galaxies less than 35 kpc in radius. Note that the Lyman-$\alpha$–dominated-disk scale height can be written $H_1 \approx 24\alpha_{\rm fs}(m_e/m_p)^{1/2}R_{\rm cool} = 4 \times 10^{-3}R_{\rm cool}$, that the cooling radius is independent of the mass within it, and that its value is disquietingly similar to the value of $r_{\rm max}$ found for the Milky Way.

One might also be concerned about the amount of star formation which occurs in the infalling material before it reaches the plane. Such star formation will tend to be suppressed by ionizing UV leaking from the plane. But even if not suppressed, it is small. The infall is not as dense as the plane, and the infall time is less than the age of the plane. If they both have the same density dependence for their star formation rates, infalling material will always be much less stellar than the material in the disk.

A bit more detail about the early behavior of the growing perturbation may be of some interest and is included in the Appendix. The next section considers the properties of disks whose orderly infall is interrupted by the onset of Lyman-$\alpha$ support in the halo.

## VI. DISK FORMATION REGULATED BY LYMAN-ALPHA

At the epoch when $\sigma$ reaches $\sigma_1$, taken as fiducial, the system radius is $r_1 = \pi G\sigma_1 t_{\rm flow}^2$, the mass $M_1 = \pi\sigma_1 r_1^2 = \pi^3 G^2\sigma_1^3 t_{\rm flow}^4$, and the incremental time is

$$\Delta t_1 = \frac{t_c}{2}\left(\frac{M}{M_p}\right)^{2/3} = \frac{t_c}{2}\left(\frac{\pi^3G^2\sigma_1^3t_{\rm flow}^4}{M_p}\right)^{2/3} .$$

Expressed in terms of $t_{\rm flow}$ and $\Delta t_1$, conditions at other times are simple functions of

$$x = \sigma/\sigma_1 = (\Delta t/\Delta t_1)^{1/2} .$$

The gaseous fraction, $f_g$, as a function of $\sigma$ must be evaluated before the mass of the supportable halo can be calculated. The UV photon production per unit area is $(4/3)Hf_0\,\alpha n_0^2$, so that the rate of increase of stellar column density (for $\sigma > \sigma_1/f_g$) is

$$\begin{aligned}\dot{\sigma}_* &= \frac{1}{m_p\phi}(1.4m_p)\frac{4}{3}Hf_0\,\alpha n_0^2\\ &= \frac{1.1m}{m_p\phi}\frac{4}{3}f_0 H\alpha n_1^2x^2\\ &= \frac{x^2}{t_{m1}}\frac{2}{3}f_g\left(\frac{2Hmf_0\,n_1}{f_g}\right) = \frac{2}{3}\frac{\sigma_1x^2}{t_{m1}}f_g .\end{aligned}$$

Using the relations $\dot{\sigma} = \sigma/(2\Delta t)$ and $2\Delta t = 2\Delta t_1x^2$, we have $\dot{\sigma} = \sigma_1/(2\Delta t_1 x)$ and

$$\frac{d\sigma_*}{d\sigma} = \frac{\dot{\sigma}_*}{\dot{\sigma}} = \frac{2}{3}\frac{2\Delta t_1}{t_{m1}}x^3f_g = \frac{4}{3}\frac{\Delta t_1}{t_{m1}}x^3\left(\frac{\sigma - \sigma_*}{\sigma}\right) .$$

Defining $x_* = \sigma_*/\sigma_1$,

$$\frac{dx_*}{dx} = \frac{4}{3}\frac{\Delta t_1}{t_{m1}}x^2(x - x_*) .$$

Substituting

$$J = \frac{1}{3}\left(\frac{4}{3}\frac{\Delta t_1}{t_{m1}}\right)x^3 ,$$

the required integration can be written

$$\frac{x_*}{x} = 1 - f_g = \frac{e^{-J(x)}}{[J(x)]^{1/3}}\int_0^{J(x)} J^{1/3}e^J dJ .$$

When integrated numerically and fitted, the result is

$$f_g \approx \frac{4}{4 + 3J + J^2} .$$

This result, together with the definition of $J$, provides the dependence of $f_g$ on $\sigma$ during the unimpeded infall period. (It was assumed, however, that $f_g x > 1$.)

In § IV we found that the ratio of supportable near halo to disk mass, $\beta$, follows from $\beta(1 + \beta) = \delta_H f_g/\delta$. We now need only to examine the infall to discover when the amount of nearby infalling material becomes supportable. This can be done in sophisticated ways, but in the following approach the essential elements are clearer and the dependence on the infall details is weaker.

Nearby infalling material would enter the galactic disk over a time period comparable to $t_{\rm flow} = r_e/v_{\rm rot}$. Hence the amount of material must be of order $\Delta\sigma \approx \dot{\sigma}t_{\rm flow} = \sigma t_{\rm flow}/(2\Delta t)$ and becomes supportable when $\Delta\sigma \approx \beta\sigma$. Thus support commences when $\Delta t$ reaches $t_{\rm flow}/2\beta$. For $\beta \sim 0.6$, unimpeded infall lasts only for a time of order $0.8t_{\rm flow} \sim t_{\rm rot}/8$, roughly $4 \times 10^7$ years for the Milky Way galaxy.

The disk surface density at infall shutoff follows from $\sigma = \sigma_1(\Delta t/\Delta t_1)^{1/2}$,

$$x = \left(\frac{t_{\rm flow}}{2\beta\Delta t_1}\right)^{1/2} .$$

To find the resulting family of possible galactic structures, it is most convenient to calculate the family loci for constant $J$. $J$ leads immediately to $f_g(J)$ and $\beta(f_g)$.

Mapping can proceed, using $x$ as the implicit variable,

$$\begin{aligned}\Delta t_1 &= \tfrac{9}{4}t_{m1}J/x^3 ,\\ \Delta t &= x^2\Delta t_1 = \tfrac{9}{4}t_{m1}J/x ,\\ t_{\rm flow} &= 2\beta\Delta t = \tfrac{9}{2}\beta t_{m1}J/x .\end{aligned}$$

When such mapping is done, the parameter pairs ($t_{\rm flow}$, $\Delta t_1$) for $J > 1.5$ are found to duplicate (at larger values of $x$) the values for $J < 1.5$. This makes $J > 1.5$ ($f_g < 0.37$) inaccessible and over most of the range available to ($t_{\rm flow}$, $\Delta t_1$) the simpler case $f_g \approx 1$ suffices.

Before performing an actual mapping, it is useful to discuss the rotation velocity. We found previously that $r = \pi G\sigma t_{\rm flow}^2$. It is unclear in this simple model whether $\sigma$ should include the supported halo, but I will assume that it does. Thus

$$r \approx \pi G\sigma_1 x(1 + \beta)\left(\frac{9}{2}\beta\frac{t_{m1}J}{x}\right)^2$$

and

$$
\begin{aligned}
v_{\rm rot} &= \frac{r}{t_{\rm flow}} = (\pi G\sigma_1 t_{m1})\,\frac{9}{2}\,\beta(1+\beta)J \\
&= \left(\frac{9}{2}\frac{\delta_H}{\delta}\,f_g J\right)\pi G\sigma_1 t_{m1} \\
&= \left(\frac{9}{2}\frac{\delta_H}{\delta}\,f_g J\right)\frac{5}{1.1}\,\frac{[(3/4)I_{\rm H}\psi]}{mc^2}\,(m_p\,\phi)\delta c \\
&= \frac{45}{2.2}\,\delta_H f_g J\,\frac{[(3/4)I_{\rm H}\psi]}{\langle h\nu\rangle}\,f_1 f_2 f_3\,\frac{7\ {\rm Mev}}{mc^2}\,c \\
&= \frac{45}{2.8}\,(\delta_H f_g J)\,\frac{[(3/4)I_{\rm H}\psi]}{\langle h\nu\rangle}\,(0.007 f_1 f_2 f_3)c \\
&= 7.8\times 10^{-4} I\delta_H f_g J c \\
&= (232\ {\rm km\ s^{-1}})(I\delta_H f_g J)\ .
\end{aligned}
$$

The third-to-last form for $v_{\rm rot}$ is particularly illuminating, in that it reveals the complete set of assumed parameters leading to the final rotation velocity of the galaxy. Lumping all of the uncertainties in $f_1 f_2 f_3$ into $I$ (see § I), the only other parameter of significance (given $J$, which determines $f_g$) is $\delta_H$.

In addition, $f_g J$ is monotonic in $J$ for the applicable range $J \lesssim 1.5$, so that the maximum value of $f_g J$ is 0.558; *the maximum possible rotation velocity is* $129 I\delta_H$ km s$^{-1}$. Clearly our mapping of galaxies in the ($t_{\rm flow}$, $\Delta t_1$)-plane will produce familiar results only if $I\delta_H \gtrsim 2$. A slight weighting of the ZAIMF toward higher masses, compared with the solar neighborhood IMF, is sufficient and perhaps expected. Once again the model produces galaxy-like numbers without prior knowledge. For purposes of illustration, $\delta_H = 1$, $\delta = 1$, $I = 3$ will be assumed in the numerical example below. The largest rotation velocity will be 388 km s$^{-1}$.

Figure 2*a* shows the properties of disk galaxies for which Lyman-$\alpha$ pressure stops infall, as functions of the two composite parameters $t_{\rm flow}$ and $\Delta t_1$. Recall that $t_{\rm flow}$ is $1/2\pi$ of the edge rotation period (constant during the formation of a disk) and $\Delta t_1$ is the time required for $\sigma$ to reach $\sigma_1$. The infall duration $\Delta t$ is $t_{\rm flow}/(2\beta)$ where $\beta$ varies between 0.29 and 0.61 over the diagram. Hence the time scale for accumulation of the "galaxies" shown is $\Delta t \sim t_{\rm flow}$.

Figure 2*b* presents the same family of galaxies as functions of parameters characterizing the perturbation as a whole. Assuming that galaxy formation took place at $t_c \sim 10^9$ years, the figure presents

$$
\epsilon' = \left(\frac{t_c}{10^9\ {\rm years}}\right)^{1/3}\epsilon = \left(\frac{\pi}{\sqrt{2}}\,\frac{t_{\rm flow}}{10^9\ {\rm years}}\right)^{1/3}
$$

and

$$
M'_P = \left(\frac{10^9\ {\rm years}}{t_c}\right)^{3/2} M_P = \left(\frac{10^9\ {\rm years}}{2\Delta t}\right)^{3/2} M_d\ ,
$$

where $M_d = \pi(\pi G\sigma_1 x t_{\rm flow}^2)^2\sigma_1 x$ is the disk mass at the time support begins. For $t_c = 10^9$ years, $M'_P = M_P$ is approximately the total perturbation mass, $\epsilon' = \epsilon$ is the shear parameter, and $\epsilon^2 = r_e/r_{\rm max}$ is the net shrinkage. Values of $t_{\rm flow} \sim \Delta t$ which are small compared to $10^9$ years imply moderate to small $\epsilon$ and galaxy masses small compared to $M_P$. The fraction $M/M_P$ is in fact $(2\Delta t/t_c)^{3/2} \sim (2t_{\rm flow}/t_c)^{3/2} \sim \epsilon^{9/2}$ and is shown along the right-hand vertical axis in Figure 2*b*.

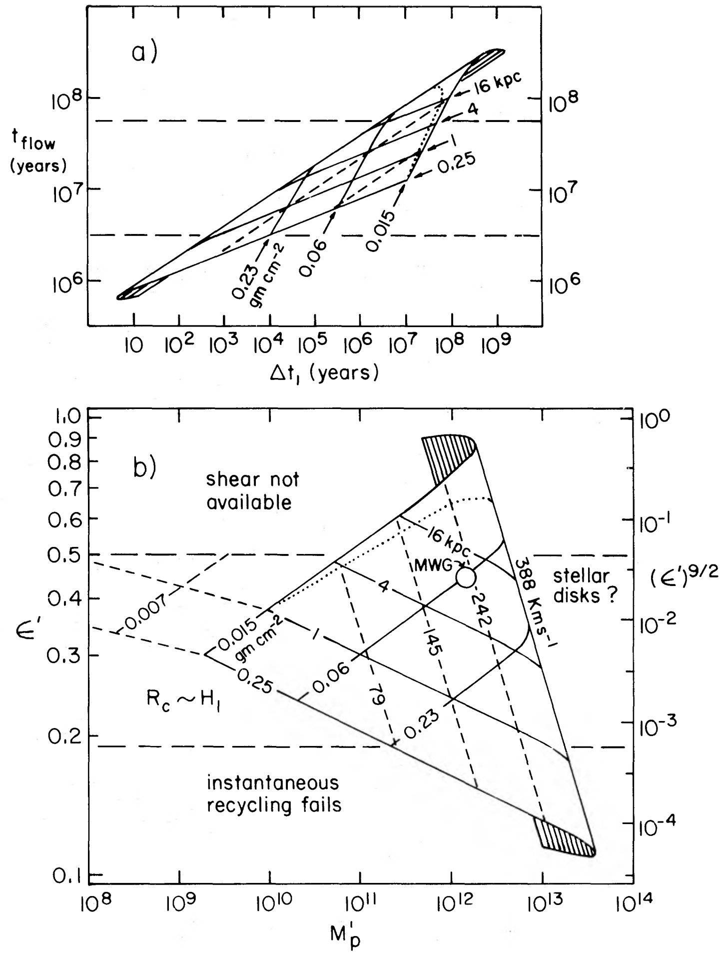


FIG. 2.—The family of galaxies deriving from the H II disk model. (*a*) The average disk column densities and radii are shown as functions of two composite parameters, $t_{\rm flow}$ and $\Delta t_1$. Both are constants for the evolution of a particular system. The former is a measure of the overall evolution rate, while the latter measures the rapidity with which the disk reaches Lyman-$\alpha$ control. (*b*) Normalizing to a formation epoch at $t_c = 10^9$ years, the family is shown as a function of the two remaining perturbation parameters. The perturbation mass is $M_P$, the shear parameter $\epsilon$. Primed values pertain to the normalization described in the text. Loci of constant surface density, radius, and rotation velocity are shown. The right-hand axis indicates the fraction of the initial perturbation which enters the central galaxy by the time Lyman-$\alpha$ stabilizes the halo. Inclusion of the near halo and edge growth raises the incorporation by roughly 3–4. Notice the extreme differences in the breadths of ranges allowed the two parameters $\epsilon$ and $M_P$.

In both representations Figure 2 locates galactic disks on a bent triangular sheet showing loci of constant $r_e$, $\sigma$, and $v_{\rm rot}$. Along the rolled edge the forward (accessible) section is bounded by a line labeled by the maximum rotation velocity 388 km s$^{-1}$ ($J = 1.5$). Beyond this edge, no Lyman-$\alpha$ stabilized systems exist. The other two edges of the sheet are chosen for convenience. The edge labeled 0.015 g cm$^{-2}$ is the limit of disks which can be supported primarily by Lyman-$\alpha$. Similar models can be constructed beyond this edge using the Lyman-$\alpha$ emissivity of a thermally supported disk. The strong dependence of $l_\alpha$ on $\sigma$ in this regime leads to a more gradual variation of $\sigma$ than that shown within the triangle, but the basic scheme of the model is preserved. (The precise boundary of applicability is shown by the dotted line, along which $f_g\sigma = \sigma_1$.) Dashed lines in Figure 2*b* show extensions of the $r = 0.25$ kpc and $r = 1$ kpc loci beyond this edge, as well as the $x = 0.5$ ($\sigma = 0.007$ g cm$^{-2}$) trajectory.

The third boundary is labeled $r = 0.25$ kpc. Solutions beyond this boundary are also possible in principle, but the calculations involve spherical symmetry, since $r$ and $H$ are comparable. Preliminary calculations indicate that the characteristic scale $R_c$ remains about $H_1$, depending on the filling factor. The clear distinction between disk and halo disappears, but there is a change in $\rho(r)$ dependence from inner to outer regions, at about $R_c$. Apart from the inner few hundred parsecs, such systems have $M \propto r$ but no characteristic mass. They seem to be dwarf spheroids, although I use the name with some caution since they potentially extend to very large radii. During formation the inner region would seem to resemble the huge H II complexes seen *within* other spiral galaxies.

There are two other applicability limits, shown as horizontal dashed lines in both diagrams. The upper line marks $\epsilon' = 0.5$ ($t_{\rm flow} = 5.6 \times 10^7$ years) and reflects the fact that the maximum allowed velocity shear in a gravitationally bound system results in net contraction by a factor of 4 once angular momentum is shared azimuthally. Thus the universe is unable to supply perturbations with $\epsilon > 0.5$. Notice that this upper limit just barely admits the Milky Way.

The lower line at $t_{\rm flow} = 3 \times 10^6$ years ($\epsilon' = 0.19$) marks the failure of the instantaneous recycling approximation to accommodate to changes on time scales of order $t_{\rm flow}$ or $\Delta t$. A related applicability limit is that the star formation time scale must exceed $3 \times 10^6$ years. For $I = 3$ as in this example, $t_m = 1.3 \times 10^9$ years$/n_0$, requiring $n_0 < 400$ cm$^{-3}$, $\sigma/\sigma_1 \lesssim 50$, $\sigma \lesssim 0.7$ g cm$^{-2}$. Recall that incoherent electron scattering also becomes important in this regime. More conservatively, the model's details are in doubt over much of the corner shown with $\sigma > 0.23$ g cm$^{-2}$.

The total range of parameters for which this model (or its cousins at low surface density or small scale) could be applicable is, in Figure 2*b*, to the left of $v_{\rm rot} = 388$ km s$^{-1}$ between $\epsilon = 0.19$ and $\epsilon = 0.5$, excluding the high surface density corner at lower right. It thus contains only systems which are rather severly sheared, but includes a very large range in perturbation mass. It also contains a very broad range of resulting galaxy parameters. Characteristic scales range from roughly 200 pc to 20 kpc. Surface densities range from roughly 0.01 to 0.3 g cm$^{-2}$. Furthermore, the fraction of the perburbation mass which enters the galaxy ranges from a few percent down to 0.0007.

Despite the relatively high specificity of the model thus far, most of the conclusions one might be tempted to draw from these results fall in the category of wild speculation. The problem is that when a system with relatively low $\epsilon$ forms a "galaxy," only a small fraction of the total mass is consumed before infall to that objects stops, or rather pauses, for a time of order $t_m$. The eventual structure of the larger system depends critically on what else is going on, in addition to the effects of the pulse of Lyman-$\alpha$ on the surrounding material.

Consider, however, the high-shear end of the possible range of conditions. For $\epsilon \approx 0.5$, radii range from roughly 4 to 20 kpc, surface densities from 0.015 to 0.08 g cm$^{-2}$, rotation velocities from 80 to almost 400 km s$^{-1}$. All three of those variables alter in direct proportion at constant $\epsilon$. Roughly 3% of the perturbation enters the disk, a similar amount is held in the near halo, and the rest is supported (at least temporarily) in a massive halo reaching roughly 10 disk radii. Were it not for their needing other nearby mass perturbations in order to achieve the high shear, these systems would be good candidates for isolated spiral galaxies with massive halos. Owing to the shear requirements, however, they are more likely to occur near other substantial mass concentrations with contraction times at least as short. Such systems, in general, are more likely among the later forming objects (larger $t_c$), for several reasons. By forming on cluster boundaries where they began as positive density perturbations in comparatively lower density regions, their collapse delay subjects them to maximum shear at maximum extent. In addition, because they form relatively slowly (extended $\Delta t$), they compete badly for dominance in regions containing significant amounts of low-shear material, that is, within clusters.

One other aspect of high-shear systems is that the amount of halo material with which they must cope depends more sensitively on $\epsilon$ than does any other feature. If such material eventually generates the bulge component, the globular cluster host, the dark halo, small satellite galaxies, or the outer H I distribution which is used to measure the rotation curve beyond the optical disk, such manifestations may because of this sensitivity appear rather uncorrelated with the gross properties of the disk component.

## VII. HALO EVOLUTION CONSIDERATIONS

Assuming that halo material is more or less homogeneous when it first stabilizes, its properties are defined by the radiation pressure available to support it. For $r > r_e$, the density of that material is

$$\rho(r) = \frac{5}{2\pi} \frac{\delta_H}{cr} \frac{\mathscr{L}_\alpha}{GM(r)} ,$$

where $\mathscr{L}$ is the total Lyman-$\alpha$ luminosity of the disk. Thus

$$M(r) \approx \left(\frac{10\delta_H \mathscr{L}_\alpha}{Gc}\right)^{1/2} r .$$

The ionization state of this material depends on the amount of UV flux leaking out of the plane. The leakage fraction $f_v$ should be roughly $(1/3)r_s(0)/H$, where $r_s(0)$ is the typical Strömgren radius at midplane. Assuming the typical O star ionizing luminosity to be $1.5 \times 10^{49}$ photons s$^{-1}$ (e.g., Abbott 1982), the leakage fraction for a Lyman-$\alpha$–supported disk is then

$$f_v \sim \frac{1}{20} \frac{f_0^{2/3}\delta^{1/3}}{f_g x^{2/3}} ,$$

and the ionizing flux in the halo at radius $r$ is

$$F = f_v \frac{\mathscr{L}_\alpha}{[(3/4)I_{\rm H}\psi]4\pi r^2} .$$

With these results, the total and neutral densities can be shown to be

$$n(r) = \left(\frac{f_g \delta_H}{3\delta}\right)^{1/2} \frac{H_1 r_e}{r^2} n_0 ,$$

$$n^0(r) = \frac{\delta_H}{\delta} \frac{H_1}{r^2} \frac{1}{A_c f_v} ,$$

where $A_c$ is the average ionization cross section. Since both of these densities are inversely proportional to $r^2$, the ionization fraction is constant. The residual optical depth beyond $r$ to continuum radiation is

$$\tau_c(r) = \int_r^\infty n^0(r) A_c \, dr ,$$

or

$$\tau_c(r) = \frac{\delta_H}{\delta} \frac{H_1}{r} \frac{1}{f_v} .$$

For moderate $f_v$, the homogeneous halo is optically thin and fully ionized. Its residual optical depth to Lyman-$\alpha$, however, remains, significant to at least $r \sim 10^3 H_1 \sim 10^5$ pc.

The halo pressure is $5\delta_H \mathscr{L}_\alpha/(4\pi r^2 c)$, and the characteristic velocity $v_H = (p/\rho)^{1/2}$ is both independent of $r$ and essentially equal to $v_{\rm rot}$. The halo material is thus capable of rearrangement with very high velocity, and the characteristic dynamical time is $r_e v_{\rm rot} = t_{\rm flow}$, at least in the near halo. Consistently, the infall time when support fails is $t_{\rm in} \sim [r^3/GM(r)]^{1/2} \sim (r/r_e) t_{\rm flow}$. One other timescale of interest is the recombination time

$$t_{\rm rec} \sim \frac{1}{\alpha n} \sim \frac{r^2}{H_1 r_e} \frac{1}{\alpha n_0} \sim \frac{1.6 \times 10^5 \text{ years}}{n_0} \frac{r^2}{H_1 r_e} .$$

Typically $t_{\rm rec}$ is small compared with $t_{\rm flow}$ at $r \sim r_e$ but becomes appreciable at large radii. The time to achieve equilibrium with the ionizing UV is $(n^0/n) t_{\rm rec}$, which is always very short, so that the initial ionization state of the infalling material is irrelevant.

The nature of the Rayleigh-Taylor–like instability between radiation and the gaseous fluid differs between disk and halo. In the halo the radiation source is predominantly external, the Lyman-$\alpha$ arising from the disk. In addition, the severity of the overpressure (radiation versus thermal) is much greater, essentially $(v_{\rm rot}/10 \text{ km s}^{-2})^2$ as opposed to only $x$ in the disk. These differences lead to much greater potential for compression of halo material into dense clumps. I have not, however, been able to discern a reliable criterion for judging the crossover point between the opalescent behavior imagined for the disk and the wholesale clump formation which I propose for the halo. It is a question deserving study well beyond the scope of this initial investigation (cf. Krolik 1979).

One can inquire, however, about what sort of clumps might be expected to form in the halo under the effects of the extreme Lyman-$\alpha$ pressure. Such clumps might stabilize when their thermal pressure equals the Lyman-$\alpha$ pressure; but that supposition ignores the penetration of Lyman-$\alpha$. The latter invades clumps easily, and they experience only its relatively mild gradient (which supported them before compression). Lyman-$\alpha$ pressure is not particularly isotropic, however, at the relatively low optical depth of the halo. Thus stable clumps will be ones with a moderate Lyman-$\alpha$ optical depth of their own, experiencing a net compressional force per unit area by differential scattering of the outward flowing Lyman-$\alpha$ component. The outward Lyman-$\alpha$ flux must be $\mathscr{L}_\alpha/(4\pi r^2)$. It carries momentum equal to that flux divided by $c$. The fraction reflected by a clump of optical depth unity at line center will be roughly $1/x_*$, where $x_*$ is the half-width (in Doppler units) of the line, as discussed by Adams (1975). The line width is not particularly sensitive to $\tau_0$, so I will set the fraction at $\frac{1}{8}$, a typical value. We then have the dual conditions

$$\Delta p \sim \frac{1}{8} \frac{1}{c} \frac{\mathscr{L}_\alpha}{4\pi r^2}$$

and, for clumps of scale $R$,

$$A_\alpha n^0 R \sim 1,$$

together with the internal pressure balance $\Delta p \sim 2nkT$ (with $T \sim 10^4$ K) and the ionization balance $F A_c n^0 = \alpha n^2$. $A_\alpha$ is the Lyman-$\alpha$ scattering cross section at line center, approximately $10^4 A_c$. These conditions lead to criteria with $n \propto r^{-2}$, so that the compression factor is independent of $r$. The clump scale increases as $r^2$ and the clump mass as $r^4$.

One might hope that these clumps eventually lead to self-gravitating entities and that the near halo produces stars and the galactic bulge, the intermediate halo makes globular clusters, and the far halo makes dwarf galactic companions. The numerical results, however, are not so clear cut. The incremental pressure $\Delta p$ equals only $\delta_H/40$ of the full Lyman-$\alpha$ pressure of the halo, or

$$\Delta p \sim \frac{\delta_H \delta_c}{\delta} \frac{p_0}{120} \left(\frac{r_e}{r}\right)^2 ,$$

where another fudge factor $\delta_c$ is introduced to follow the effects of the uncertainty in clump internal pressure. The clump density is

$$n_c = \frac{\Delta p}{2kT} = \frac{\delta_H \delta_c}{\delta} \frac{f_g n_1 x^2}{240} \left(\frac{r_e}{r}\right)^2 = \frac{\delta_H \delta_c f_g x^2}{30\delta^2} \left(\frac{r_e}{r}\right)^2 \text{ cm}^{-3} ,$$

and the compression factor is

$$\frac{n_c}{n} = \frac{r_e \delta_c x}{140 H_1} \left(\frac{f_g \delta_H}{\delta}\right)^{1/2} .$$

Only relatively large high surface density systems ($r_e x > 140 H_1$) seem vulnerable to any compression at all. The neutral fraction in the clumps is rather small,

$$\frac{n^0}{n_c} = \frac{[(3/4) I_H \psi]}{16kT} \frac{\delta_c \alpha}{A_c c} \frac{1}{f_v} \sim 1.2 \times 10^{-6} \frac{\delta_c}{f_v} ,$$

as are the clump radii,

$$R_c \sim \frac{1}{n^0 A_\alpha} = \frac{1}{n_c} \frac{n_c}{n^0 A_\alpha} \sim (4.5 \text{ pc}) \frac{f_v}{\delta_c n_c} .$$

Finally, the masses of the individual clumps, still for $r > r_e$ only, are

$$M \approx \frac{12 \, M_\odot f_v^3}{\delta_c^3 n_c^2} \sim 10^4 \, M_\odot \frac{f_v^3 \delta^6}{\delta_c^5 \delta_H^2 f_g^2 x^4} \left(\frac{r}{r_e}\right)^4$$

$$\sim 1.3 \, M_\odot \frac{f_0^2 \delta^7}{\delta_c^5 \delta_H^2 f_g^5 x^6} \left(\frac{r}{r_e}\right)^4 .$$

The result for the clump mass is extremely sensitive to a number of factors, making even the order of magnitude suspect. For a galaxy like the Milky Way, however, with $x \approx 5$, clump densities of order 1 cm$^{-3}$ are suggested at $r = r_e$. The other parameters at $r = r_e$ are a flux leakage fraction $f_v \sim 10^{-2}$ (for $f_0 \sim \frac{1}{2}$), a clump radius $\sim 4 \times 10^{-2}$ pc, and a clump mass $\sim 2 \times 10^{-5} \, M_\odot \sim 7 \, M_\oplus$. Masses would be substantially larger (at $r = r_e$) for lower surface density systems. They are also sensitive to the flux leakage fraction, which depends most sensitively on the formation distribution of high-latitude O stars. Overall, the possibilities run from substellar to globular cluster masses at least. But if the analysis contains even a shred of resemblance to reality, there should be strong systematic dependencies of clump mass on system column density and, within a system, on $r/r_e$.

This initial clumping of halo material is not the last word in its evolution. As the inner halo forms clumps, it is possible for it to become optically thick in Lyman-$\alpha$ continuum, shutting off ionization and heating in the outer halo. The outer clumps can then cool rather rapidly and recombine. They may be induced to fragment as their Lyman-$\alpha$ optical depth rises sharply. A similar effect may apply during the decrease in brightness of the plane as star formation lowers the gaseous mass needing to be kept ionized in the disk. Once both ionization and Lyman-$\alpha$ are effectively shut off, and barring any consideration of effects that supernovae would have on the halo, there must be a competition between cooling/recombination on one hand and pressure expansion/infall on the other. If clumps manage to become self-gravitating, they may individually survive. If not, collisions between them can lead to transient high pressures and eventual stability of coalesced groups.

Finally, although one of the ground rules of this investigation has been to omit consideration of supernova effects as much as possible (in part because those effects promise to be as significant as those attributed to Lyman-$\alpha$ and deserve a separate investigation before a merged model is attempted), the relatively low density in the halo makes it a very likely location for supernovae to dominate the outcome. Waves of supernova-driven compression, followed by cooling and condensation, could well shatter the halo.

The above discussion does not leave one confident of any particular outcome of halo evolution. But if we cannot clarify the answer, perhaps we can at least clarify the problem. What one hopes for from halo evolution is that it will generate both the observed stellar halo components as well as a largely unobservable dark halo of substellar objects. But, as we found in the previous section, one hopes for this outcome only in the highly sheared systems leading to halos some thirty or so times more massive than the disk. For less sheared systems, typically with higher surface density and smaller scale, the residual mass in the perturbation ranges up to a thousand times the mass accepted by the galaxy. In that case, one supposes there is some natural cutoff in the amount belonging to the "central object" and that the rest is simply given time to join other perturbations. This may lead to a cluster of galaxies only the central member of which followed the formation scenario herein prescribed.

In regard to this latter possibility, it may be important to realize that there is a fourth parameter of galactic infall, discussed recently by Silk, Szalay, and Zel'dovich (1983). That parameter has to do with the noncentrality of the perburbation flow. It is a defocusing effect that destines different regions in a hierarchical or many-overlapping-wavelength system of perturbations to aim for different centers. Such a parameter is similar to $\epsilon$ in that, in the absence of dissipation, it leads to a system of finite size. If all major perturbations contain similar values of this defocusing parameter, and if in general it leads to only a modest recontraction, then the low $\epsilon$, high surface density systems may never have an opportunity to form. One would then get disk galaxies only when the shear parameter was large enough to dominate over defocusing. All less sheared cases would lead to spheriodal systems dominated by dissipation, somewhat like Larson's (1969) models. In short, it is clearly important to study halo evolution for the systems with sufficiently high $\epsilon$ that the galaxy disks look familiar. It is not so clear whether the lower $\epsilon$ models have any counterpart in reality. If they do, they correspond either to a galactic cluster with an unusually high surface density disk galaxy (at least initially of small radius) at the center, or possibly on occasion to a very massive spheroidal galaxy with a high surface density disk buried deep in its core.

## VIII. SUMMARY

The model presented thus far is largely an impressionistic sketch of how galaxy formation might take place. The individual brush strokes in the previous section may bear little resemblance to detailed behaviors they represent. Having put them to canvas, however, we now step back for three views in broader sweep. The first of these perspectives is a summary of the components of the picture; the second is a portrait of the formation of a particular system, the Milky Way; and the third is a landscape of the possibilities represented in Figure 2.

The most specific aspect of the model is its assumed understanding of the state and self-limited star formation rate of the nascent ISM. By combining this understanding with self-gravitation, many properties of the forming disk can be calculated as simple functions of the column density $\sigma$. These include $p$, $H$, $n$, $\eta$, $\eta_\alpha$, $l$, $l_\alpha$, and $t_m$ or $\dot{\rho}_*$. In order to find the last pair, it is necessary to make an assumption about the high-mass weighting of the ZAIMF for stars, and the parameter $I$ was introduced to express the weighting relative to the Paper I value for the solar neighborhood. (The Paper I weighting, based on Abbott 1982, was already somewhat heavily biased toward high-mass stars as compared with some earlier estimates. See Garmany, Conti, and Chiosi 1982 for a comparison between estimates.)

Sources of radiation trapping in the disk were explored, and it was found that Lyman-$\alpha$ pressure in the disk becomes comparable to thermal pressure as galaxy-like column densities are approached (envisioning $\sigma$ monotonically increasing as a result of infall). The competing possibilities that at higher column densities Lyman-$\alpha$ will either enhance its own escape or take over pressure dominance were investigated, and the latter held sway. The structure of the disk was then recalculated assuming Lyman-$\alpha$ support. The chief result was that the scale height stabilized at $H = 138(\delta f_g/f_0)$ pc, where 138 pc $= 23.3(m_e/m_p)\alpha_{\rm fs}^2(e^2/Gm_p m_e)r_0$. The crossover column density between thermal and Lyman-$\alpha$ support is at $\sigma f_g \delta = 0.0146$ g cm$^{-2} = 70\ M_\odot$ pc$^{-2} = 0.28(kT/I_{\rm H})^{3/2} m/A_{\rm T} \approx 0.006 m_p/A_{\rm T}$. The characteristic midplane density at crossover is 8 cm$^{-3}$, and the mass conversion time scale about $5 \times 10^7$ years.

A simple model was explored for growth of a perturbation in the early universe. In particular, relevant results were presented for the infall of the core of such a perturbation. Initially, three parameters describing an arbitrary but spherically symmetric (except for shear) perturbation were introduced, the perturbation mass $M_P$, the initiation time of disk formation, $t_c$, and degree of shear, $\epsilon$. It is shown that these lead to two interesting formation constants for the subsequent galaxy. The net shrinkage of each mass shell from its maximum extension $r_{\max}$ to subsequent disk $r_e$ is $r_e/r_{\max} = \epsilon^2$. The upper limit to possible values is $\epsilon = \frac{1}{2}$. In addition, $t_{\rm flow} \equiv r_e v_{\rm rot} = (r_e^3/GM)^{1/2} = (2^{1/2}/\pi)\epsilon^3 t_c$. Finally, as infall continues, $r_e$ and $\sigma$ are both proportional to $(\Delta t)^{1/2}$ and $M \approx M_P(2\Delta t/t_c)^{3/2}$ for $2\Delta t \ll t_c$.

To this point the overall model provided a time-dependent picture of a galaxy disk as it accretes matter and assumes an evolving structure consistent with its current surface density. The next task was to learn when the galaxy was capable of rejecting further infall. The mechanism explored was again Lyman-$\alpha$ trapping, this time in the halo. It was found that when the near halo ($r \lesssim r_e$) contained a total mass comparable to what was then in the disk, the Lyman-$\alpha$ pressure gradient could support the halo mass against gravitational attraction. More particularly, when the near halo contained $\Delta\sigma = \beta\sigma$ with $\beta \sim 0.6$, then $\Delta\sigma$ was supportable. This leads immediately to a stabilization time $\Delta t = t_{\rm flow}/2\beta$ (from $\dot{\sigma} = \sigma/2\Delta t \approx \Delta\sigma/t_{\rm flow} = \beta\sigma/t_{\rm flow}$). Thus the duration of unimpeded infall is comparable to $t_{\rm flow} = r_e/v_{\rm rot}$. All properties (essentially $\sigma$, $r_e$, and combinations thereof) follow immediately. Summarized, they are

$$\sigma = \frac{1}{(2\beta)^{1/2}}\left(\frac{2^{1/2}}{\pi}\right)^{1/6}\left(\frac{M_P}{G^2 t_c^4 \epsilon^{15/2}}\right)^{1/3},$$

$$r = \frac{1}{\beta^{1/2}}\left(\frac{2^{1/2}}{\pi}\right)^{7/6}(GM_P t_c^2 \epsilon^{21/2})^{1/3}.$$

(This formula gives the disk radius at $\Delta t$. The results given earlier included an additional factor of $1 + \beta$ expected for disk growth near the edge where the halo mass was not vertically supported.) The disk mass is

$$M = M_P\left(\frac{2\Delta t}{t_c}\right)^{3/2} = M_P\left(\frac{2^{1/2}\epsilon^3}{\pi\beta}\right)^{3/2}.$$

These results are appropriate within the triangular sheets of Figure 2 (or with $\beta \approx x = \sigma/\sigma_1$ they extend beyond the low surface density edge). Conceptually similar results can also be found beyond the low-radius edge, but these systems are spheroidal with constant characteristic radius. In principle the above formulae should be the most significant consequences of this endeavor, but only after the fate of the halo and universe's choice for $\epsilon$ are understood.

The "rolled edge" limit of applicability of the above equations (in Fig. 2) corresponds to systems which collect so much material that they turn to stars before stabilizing. Beyond this edge ($t_m < t_{\rm flow}$, approximately), Lyman-$\alpha$ pressure never stabilizes the infall, and the model as it stands would predict that the entire perturbation mass would gradually fall onto an essentially stellar disk. Such largely stellar systems are very likely candidates for supernova control of disk limitation that will be explored in a subsequent paper. If such a model works, the resulting galaxies should still be systematically larger and higher in surface density and rotation velocity than Lyman-$\alpha$–stabilized disks.

The rolled edge sets a limit to the rotation velocity of Lyman-$\alpha$–stabilized disks equal to

$$v_{\rm rot} \approx 9\delta_H \frac{[(3/4)I_{\rm H}\psi]}{\langle h\nu\rangle}(0.007 f_1 f_2 f_3)c$$
$$\approx 130 I\delta_H \text{ km s}^{-1}.$$

*It is only in setting the location of this limit that the parameter I enters the model results.* If $\delta_H$ is indeed 1, and if ignoring the flattening of the gravitational potential has little effect, the value of this limit argues for $I \sim 2$–3 in order to reach the observed rotation velocities of galaxies.

The mass distribution in the Lyman-$\alpha$–supported halo is such that the rotation curve of the system is flat beyond the edge of the disk. In order for this situation to be maintained after Lyman-$\alpha$ support ceases, the halo mass must clump into long mean free path "bits" before infall resumes. The time to do this is comparable to the time the disk, at much higher density, takes to form itself into stars. Presumably the condensation must be driven by energy percolating out of the disk.

Lyman-$\alpha$ pressure in the halo is high, very much higher than thermal pressure, but the gradient is weak. The estimated clumping factor was not very great and the clump mass calculation was subject to enormous uncertainties. A proper calculation of halo evolution is the most important next step in developing this perspective. It would seem that the potential exists for radical shattering of the halo, perhaps even into very small pieces. But all detail is lacking.

One point not made earlier involves predictability of elemental enrichments in this model. The bulk of disk material forms into stars in a time of roughly $t_m \approx 5 \times 10^7 I/x$ years. But during that time, a fraction $f_3 \approx 0.06I$ went into O stars. Of that, a fraction $f_1 \approx \frac{1}{2}$ was burned to helium or beyond. Given that such stars leave a negligible amount of this processed material in their stellar remnants after explosion, the enrichment level toward the end of this epoch should be $3I$% of the initial hydrogen mass. Quite possibly the IMF evolved rapidly with this enrichment, with an average value of $I$ of the order of 2–3 (from the limit on $v_{\rm rot}$). Still, one would expect the low-mass stars in this first generation to be enriched in the sense that some 3%–9% of their anticipated hydrogen mass will have been converted to helium or heavier elements. The appropriate elemental yields will be those of stars whose lives are shorter than $t_m$. (Recall, however, that when the dust cross section per hydrogen atom reaches $0.01/x$ of the solar neighborhood value, Lyman-$\alpha$ become subject to absorption. Can Lyman-$\alpha$–absorbing dust form in $10^7$ years?)

Among the specific examples calculated in making Figure 2, one which perhaps represents the Milky Way has $M'_P = 1.7 \times 10^{12}\ M_\odot$ and $\epsilon' = 0.46$. Assuming, as there, that $t_c = 10^9$ years and $I = 3$, the Galaxy reaches a surface density $\sigma_1$ at $\Delta t_1 = 2.8 \times 10^6$ years, and its final surface density $4\sigma_1 = 0.06$ g cm$^{-2}$ = 280 $M_\odot$ pc$^{-2}$ at $\Delta t = 4.5 \times 10^7$ years. This point is reached when $f_g = 0.7$, $\beta = 0.47$, $J = 0.5$. Thus the system is still primarily gaseous, with a near halo mass equal to about half the disk mass. The flow time is $4.3 \times 10^7$ years, the edge rotation period $2.7 \times 10^8$ years, and the rotation velocity 242 km s$^{-1}$. Its nominal disk radius is 7.2 kpc, but flow of near halo material to the edge causes growth to 10.6 kpc. The original disk mass is 2.7% of $M'_P$. Adding the near halo raises this by $(1 + \beta)^3$ to 8.5%, or $1.4 \times 10^{11}\ M_\odot$. The remaining material of the perturbation is initially supported by Lyman-$\alpha$ pressure, arranged with $M \propto r$, providing a flat rotation curve to 100 kpc. Support is maintained, however, only for a period of about $t_m = 4 \times 10^7$ years while the disk finishes most of its star formation. Thereafter, halo material at a distance $r$ from the Galaxy takes about $4 \times 10^7 (r/10$ kpc) years to fall in. If it has fragmented into small stable bits before reaching the disk, it can pass through without dissipation, providing a long-term stable halo. The halo will be dark today if the bits are predominantly substellar in size. Globular clusters and a galactic bulge may derive from star formation during the halo fragmentation era.

The disk itself, when infall is first stopped, has a surface density $\sigma = 0.6$ g cm$^{-2}$, a gaseous surface density 0.04 g cm$^{-2}$ (or $N_{\rm H} = 2 \times 10^{22}$ cm$^{-2}$), a central density $n_0 = 32$ cm$^{-3}$, a midplane Lyman-$\alpha$ pressure of $2 \times 10^{-10}$ dyn cm$^{-3}$, thermal pressure of $0.9 \times 10^{-10}$ dyn cm$^{-3}$, a scale height of about 200 pc, a mass conversion time scale of $4 \times 10^7$ years, and $l_\alpha/c = 5 \times 10^{11}$ dyn cm$^{-2}$. Its Lyman-$\alpha$ luminosity is $2 \times 10^{45}$ ergs s$^{-1}$, its total luminosity about $10^{46}$ ergs s$^{-1}$, a truly dramatic object! During the star formation epoch it burns $2 \times 10^9$ of hydrogen, about 4% of the disk mass. Much later in the evolution of this system, when its luminosity has dropped by a factor of $10^2$–$10^3$ and its gas content by a factor of 10–$10^2$, its inhabitants will observe an apparent equipartition between the starlight energy density and the gas pressure. Both are feeble remnants of an earlier time when their comparability was important.

Before leaving the Milky Way, notice that the original star formation epoch in this model lasts roughly one-sixth of a rotation period. The Galaxy could not yet have known about the potential for spiral structure. In addition, the galactic plane would probably be somewhat poorly defined during this time. Locally it is the meeting place of equal and opposite infall momenta, with a scale height of 200 pc. The global average dispersion from the mean plane will likely be rather larger, giving the first generation of stars an inflated final distribution.

The critical test for the basic component of this model will be the observation at high redshift of $10^{11}\ M_\odot$ H II regions emitting $10^{45}$ ergs s$^{-1}$ in Lyman-$\alpha$, powered by $10^{46}$ ergs s$^{-1}$ of O star radiation. They may otherwise resemble modern galaxies only in radial extent. Conversely, the most important area for theoretical extension is in halo evolution.

Turning now to the larger range of possibilities in Figure 2, the predicted values of column density are proportional to $M_P^{1/3}t_c^{-4/3}\epsilon^{-5/2}$, while radii are proportional to $M_P^{1/3}t_c^{2/3}\epsilon^{7/2}$. The ratio $r/\sigma$ increases systematically with $r$ unless there is a strong correlation among the three perturbation parameters. The rotation velocity is proportional to $M_P^{1/3}t_c^{-1/3}\epsilon^{1/2}$. Since $M_P$ is the parameter with by far the largest acceptable range, the rotation velocity essentially measures $M_P$. (Notice in Fig. 2$b$ that the constant $v_{\rm rot}$ lines span only half an order of magnitude in $M_P$.) For perturbation masses in excess of about $10^{12.8}$, the model becomes inappropriate. If S0 galaxies have characteristically high surface densities and rotation velocities, they might have arisen from these more massive perturbations.

Similar models adequately extend this approach to low perturbation masses, providing either low density or very compact systems, depending on $\epsilon$.

The largest problem the model has with low-$\epsilon$ perturbations is in deciding what happens next. Small systems, typically of high surface density, form and evolve rapidly in the core of the perturbation. What, then, happens to the rest of the perturbation, with up to $10^3$ times more mass? The result may depend on a fourth perturbation parameter having to do with the lumpiness of the original perturbation and the defocusing of the infall that lumpiness brings about. These less sheared systems could lead to elliptical galaxies with or without a small embedded disk, or the larger ones could be clusters of smaller ellipticals with a peculiar central object. Then again, they may turn out to lead to nothing ever seen, throwing some small doubt on either the model itself or on the regime of perturbation space available to the early universe. Once again it is halo evolution which is important, this time on a much grander scale.

On a yet grander scale, the UV photons liberated during galaxy formation are easily capable of reionizing residual intergalactic material. Taking a simple example, the generation of $3000I \sim 10^4$ photons per proton entering stars, followed by $f_\nu \sim 10^{-2}$ leakage into the intergalactic medium (IGM), together with assumptions that the universe is predominantly baryonic and flat, that $f_* \sim 0.1$ of the mass enters the galactic stars and $f_{\rm IGM} \sim 0.5$ remains dispersed (the rest in halos), we find roughly 20 photons per ion in the IGM. One photon is required for the initial ionization, one more (approximately) to balance recombination. The remainder are sufficient to hold the neutral fraction to less than $10^{-7}$ and suppress the optical depth of the Lyman-$\alpha$ absoption trough to a value of order $10^{-2}$.

In conclusion, the reader will have noticed that the author has chosen to ignore the current enthusiasm for the idea that the apparently pervasive dark mass consists of matter in an unfamiliar form. The reason, if one is wanted, is that the hard work to determine what sort of halo would evolve in the incredibly hostile environment of a self-regulating galaxy has not yet been done. The dark mass may be in exotic forms, but then again it may not. More generally, one should be aware of the author's innate prejudice against the recent notion that in evolving beyond the radiation-dominated era the universe became not only transparent but dull.

This work was supported in part by NASA grant NGL 50-002-044 to the University of Wisconsin, and by Rice University, where most of it was carried out. I am grateful to Bill Mathews and Ray Talbot for many useful lunchtime conversations in years gone by, and to several people who provided useful help and encouragement. These include George Rybicki, Julian Krolik, Tom Adams, Barry Smith, Larry Auer, and a fellow at the Center for Astrophysics whose name I didn't catch, possibly Stan Owocki, who suggested that the Lyman-$\alpha$ might support the plane, for $\sigma > \sigma_1$. I further appreciate the tolerance of galaxy pundits toward this poorly informed excursion of an ISM modeler into the realm of the nebulae.

## APPENDIX

### GROWTH, AMPLITUDE, SCALE, AND ANGULAR SIZE OF PROTOGALACTIC PERTURBATIONS

A mass shell which reaches the origin at time $t_f$ attains an average minimum interior density $\rho_{\rm min} = 3\pi/8Gt_f^2$. A simple characterization of a perturbation is

$$\rho_{\rm min}(M) = \rho_{\rm min}(0)\left[1 - \left(\frac{M}{M_P}\right)^{2/3}\right],$$

so that the mass shell bounding $M_P$ never quite falls in. The time dependence of infall is then

$$M(t) = M_P\left[1 - \left(\frac{t_c}{t}\right)^2\right]^{3/2} \approx M_P\left[\frac{2(t - t_c)}{t_c}\right]^{3/2},$$

where the last form is that introduced previously for $t - t_c \ll t_c$. For very early times in the development of the perturbation, the average density within a mass shell is

$$\bar{\rho} \approx \frac{1}{6\pi G t^2}\left[1 + \frac{3}{5}\left(\frac{\rho_{\min}(M)}{\rho}\right)^{1/3}\right].$$

For the $\rho_{\min}(M)$ dependence shown, substituting $\rho \approx (6\pi G t^2)^{-1}$ on the right-hand side above yields

$$\bar{\rho} \approx \frac{1}{6\pi G t^2}\left\{1 + \frac{3}{5}\left(\frac{3\pi t}{2t_c}\right)^{2/3}\left[1 - \left(\frac{M}{M_P}\right)^{2/3}\right]^{1/3}\right\}.$$

Then substituting $M \approx (4/3)\pi r^3 \bar{\rho} \approx 4\pi r^3/(18\pi G t^2)$ for $M \ll M_P$, the average density distribution close to the peak of the perturbation is

$$\bar{\rho} \approx \frac{1}{6\pi G t^2}\left[1 + \frac{3}{5}\left(\frac{3\pi}{2}\frac{t}{t_c}\right)^{2/3}\right]\left[1 - \frac{1}{5}\left(\frac{\pi}{3}\frac{1}{M_P t_c t}\right)^{2/3} r^2\right],$$

which has the desired quadratic drop-off with $r$ that is characteristic of a rounded peak, justifying the $\rho_{\min}(M)$ dependence chosen. The local density distribution is similar, but with $\frac{1}{3}$ replacing $\frac{1}{5}$ in the quadratic term.

The relative amplitude of the perturbation is

$$\frac{3}{5}\left(\frac{3\pi}{2}\frac{t}{t_c}\right)^{2/3} = \frac{3}{5}\left(\frac{\rho_{\min}(0)}{\rho}\right)^{1/3}$$

during the matter-dominated phase. This can be written (for $t \ll t_c$) as

$$\frac{\Delta\bar{\rho}_c}{\rho} \approx \frac{3}{5\Omega^{1/3}}\left(\pi\frac{t_H}{t_c}\right)^{2/3}\frac{1}{1+z} \sim \frac{20}{1+z},$$

where $\Omega = 2q_0$, $t_H = H_0^{-1}$, and $z$ is the redshift. The scale of the perturbations would be roughly the radius containing mass $M_P$ at density $1/(6\pi G t^2)$,

$$r \approx \frac{1}{1+z}\left(\frac{2M_P G t_H^2}{\Omega}\right)^{1/3} \sim \frac{3\ \text{Mpc}}{1+z}\left(\frac{M_P}{10^{12}\ M_\odot}\right)^{1/3}.$$

Since these perturbations are very close to our horizon, their presently observed angular diameter would be

$$\theta \approx \frac{r(1+z)H_0\Omega}{c}$$

or roughly $10(M_P/10^{12}\ M_\odot)^{1/3}$ seconds of arc, independent of $z$.

REFERENCES


Abbott, D. C. 1982, *Ap. J.*, **263**, 723.
Adams, T. F. 1972, *Ap. J.*, **174**, 439.
———. 1975, *Ap. J.*, **201**, 350.
———. 1983, private communication.
Cox, D. P. *Ap. J.* (*Letters*), **265**, L61 (Paper I).
Garmany, C. D., Conti, P. S., and Chiosi, C. 1982, *Ap. J.*, **263**, 777.
Krolik, J. H. 1979, *Ap. J.*, **228**, 13.
———. 1983, private communication.
Larson, R. B. 1969, *M.N.R.A.S.*, **145**, 405.
Osterbrock, D. E. 1962, *Ap. J.*, **135**, 195.
Rybicki, G. B. 1983, private communication.
Savage, B. D. 1983, private communication.
Silk, J., Szalay, A. S., and Zel'dovich, Ya. B. 1983, *Sci. Am.* **249**, 72.
Spitzer, L., Jr. 1968, *Diffuse Matter in Space* (New York: Interscience).
Walmsley, C. M., and Mathews, W. G. 1969, *Ap. J.*, **155**, 57.


Donald P. Cox: Department of Space Physics and Astronomy, Rice University, P.O. Box 1892, Houston, TX 77251