

Accretion Disks and Star Formation

Suzy Collin, Jean-Paul Zahn

*Observatoire de Paris, Section de Meudon, Place Jansen, 92195
Meudon, France*

Abstract. It is generally admitted that at distances larger than $\sim 10^4$ gravitational radii accretion disks in AGN form a torus made of high velocity gaseous interacting clouds with a small filling factor. We propose here a completely different model for this region, in which unstable fragments give rise to protostars, all becoming massive stars after a rapid stage of accretion. These stars explode as supernovae, which produce strong outflows perpendicular to the disk and induce an outward transfer of angular momentum, as shown by the numerical simulations of Rozycka, Bodenheimer and Lin (1995). So the supernovae themselves can sustain the inflow mass rate required by the AGN. Assuming that the star formation rate is proportional to the growth rate of the gravitational instabilities, one obtains a self-regulated accretion disk made of gas and stars in which the gas is maintained in a state close to gravitational instability. We show that the gaseous disk is able to support a large number of massive stars and supernovae while staying relatively homogeneous. This model could explain the high velocity metal enriched outflows implied by the presence of the broad absorption lines in quasars. It could also account for a pregalactic enrichment of the intergalactic medium, if black holes formed early in the Universe. Finally it could provide a triggering mechanism for starbursts in the central regions of galaxies.

1. Introduction

There is a large consensus that AGN are fueled via accretion disks. The observation of the “UV bump” (Shields 1978, Malkan & Sargent 1982, and many subsequent papers) argues in favor of geometrically thin and optically thick disks, possibly embedded in a hot X-ray emitting corona. Generally these disks are studied using the α prescription for viscosity introduced by Shakura and Sunayev (1973). It is well known that these “ α -disks” have two serious problems at large radii: they are not able to transport rapidly enough the gas from regions located at say one parsec, and they are gravitationally unstable beyond about 0.1 parsec, precisely when the Toomre parameter Q defined as $Q = \Omega c_s / (\pi G \Sigma)$ (Toomre 1964, Goldreich and Lynden-Bell 1965) becomes of order unity (Ω is the Keplerian angular velocity, Σ is the surface density, and c_s is the sound speed).

As an illustration, Fig. 1 taken from Collin & Huré (1998) displays the Toomre radius corresponding to $Q \sim 1$, expressed in Schwarzschild radius $R_S =$

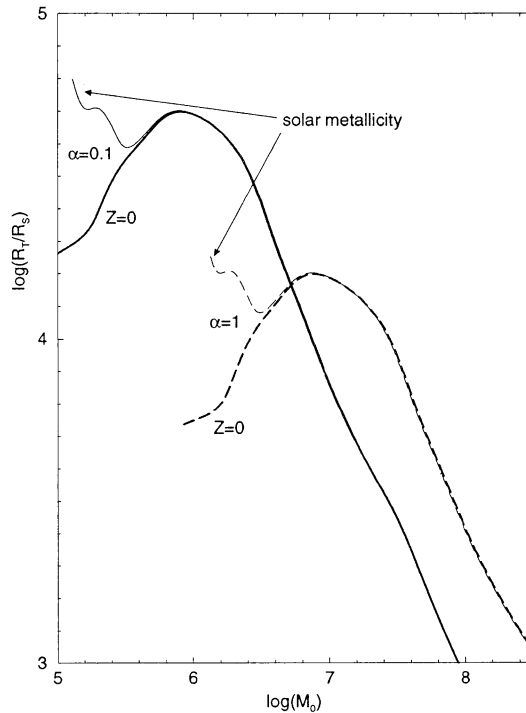


Figure 1. The Toomre radius R_T (normalized to R_S) versus the mass of the black hole M_0 in units of M_\odot . The accretion rate is critical. $\alpha = 1$ is in dashed lines and $\alpha = 0.1$ is in solid lines. The zero metallicity case is plotted in bold lines and the solar metallicity case is in thin lines. After Collin & Huré, 1998.

$2GM/c^2 = 3 \cdot 10^5 M/M_\odot$ cm as a function of the black hole mass in M_\odot for the critical accretion rate $\dot{m} = \dot{M}/\dot{M}_{\text{crit}} = 1$ (we set $\dot{M}_{\text{crit}} = L_{\text{Edd}}/(0.1c^2)$), for two values of the viscosity parameter $\alpha = 0.1$ and $\alpha = 1$, and for two metallicities, a solar and a zero metallicity. The Toomre radius expressed in R_S globally decreases with increasing black hole mass. It is roughly of the order of 10^3 to a few $10^4 R_S$, i.e. comparable to the size of the Broad Line Region (hardly a coincidence!). These results are obtained using real opacities, but they are in good agreement with analytical expressions derived assuming constant opacity coefficient and mean mass per particle μ . For small masses ($M \leq 2 \times 10^7 \alpha^{2/5} M_\odot$) the Toomre radius scales as $\dot{m}^{-8/27} M^{-26/27}$, while for higher central masses, it scales as $\dot{m}^{4/9} M^{-2/9}$ (cf. Huré 1998).

Since the α -prescription is not valid in the gravitationally unstable region, one must find a mechanism accounting for the high rate of mass and angular momentum transport in AGN. It can be achieved by magnetic torques if large scale magnetic fields are anchored in the disk. We assume here that neither turbulent nor large scale magnetic fields are important in accretion disks. Farther from the center the supply of gas can be provided by gravitational torques or

by global non axisymmetric gravitational instabilities, but this is not true in the intermediate region where the disk is locally but not globally self-gravitating.

The second problem lies in the gravitational instability. Lin and Pringle (1987) proposed a prescription allowing for strongly unstable disks. They argued that in a self-gravitating disk, the largest size of the turbulent eddies stabilized by rotation, is Q times the disk thickness. This prescription corresponds to a large (supersonic) viscosity. On the contrary Paczyński (1978) proposed that a self-gravitating disk is maintained in a state of marginal instability (i.e. $Q =$ a few units), in which the energy dissipated by collisions between clumps prevent their collapse.

There are actually two extreme possibilities for a gravitationally unstable disk: either the unstable fragments collapse to compact bodies, and as a consequence the disk evolves rapidly towards a stellar or a non interacting system, or alternatively their collapse is stopped and the disk gives rise to small gaseous possibly interacting entities.

The suggestion of star formation in the self-gravitating region of such an accretion disk has been first made by Kolykhalov and Sunayev (1980). Begelman, Frank, and Shlosman (1989) and Shlosman and Begelman (1989) discussed in more detail the conditions for star formation at about 1-10 pc from the black hole, and its consequences on the disk. They concluded that, unless the disk can be maintained in a hot or highly turbulent state, it should transform rapidly into a flat stellar system which will be unable to build a new gaseous accretion disk and to fuel a quasar. They favor the picture of a disk made of marginally unstable randomly moving clouds with a small filling factor, where the “viscosity” of the disk is provided by cloud collisions. Here we adopt the opposite view that if an unstable fragment begins to collapse, the collapse will continue until a protostar is formed, so the disk will be made both of stars and of gas.

2. The overall star-gas disk

In Collin & Zahn (1998) we describe in detail a model of gas-star disk. We shall only recall here the salient points of this study and summarize the main results.

The model is based on the assumption that, though stars are present in the disk, it stays quasi homogeneous. It means that the supernovae shells do not overlap, a condition which is checked a posteriori for each solution. Another assumption (also checked a posteriori) is that the bulk of the gas is close to marginal instability (which implies that the density is a function only of the radius and of the central mass, as $Q = 1$ is equivalent to $\pi G \rho / \Omega^2 \sim 1$, where ρ is the midplane density of the disk). If the gas would be highly unstable (i.e. $Q \ll 1$), star formation would be too efficient and would immediately transform the gaseous disk into a stellar disk (cf. next subsection).

An important issue concerning such a marginal homogeneous disk is that when massive stars evolve, HII regions and winds create a cavity around them. One can show that the size of the cavity is limited to about one scale height of the disk, H . Therefore a new star cannot form inside a radius H of the star. One can also show that no gap is opened in the disk around the stars by induced density waves unless their mass exceeds a mass m_{gap} which is extremely large. So finally the gaseous disk is able to support a large number of massive stars

while staying relatively homogeneous, since H is very small compared to the radius of the disk (typically $10^{-3}R$).

Finally we assume that the regions at the periphery of the disk provide a quasi stationary mass inflow during the life time of quasars or of their progenitors (for instance via global gravitational instabilities induced by merging), equal to the accretion rate on the black hole. In other words we assume that there is neither infall on the inner regions of the disk, nor a strong outflowing mass rate from the disk, except that due to the supernovae.

2.1. Star formation and evolution in the disk

There are several conditions for star formation, besides the fact that the gas should be at least marginally unstable. If the disk is marginally unstable, gravitationally bound fragments can form with a mass M_{frag} of the order of $4\rho H^3$ (Goldreich & Lynden-Bell 1965). They will then form a compact body if the formation time scale, t_{form} , and the cooling time, t_{cool} , are smaller than the characteristic mass transport time in the disk, t_{trans} . t_{form} corresponds to the maximum growth rate of the gravitation instability, which is, according to Wang & Silk (1994), equal to $\Omega^{-1} Q/\sqrt{1-Q^2}$. Unless Q is very close to unity, t_{form} is not much larger than the freefall time $t_{\text{ff}} \sim \Omega^{-1}$.

For an isothermal and spherical collapse the initial fragment gives rise to a dense core of mass m_{frag} in a time t_{collapse} corresponding to the growing rate (Chandrasekhar 1939) $\dot{M}_{\text{collapse}} \sim c_s^3/G$. The corresponding value of t_{collapse} is of the same order as t_{ff} . Actually the collapse should not be spherical, owing to the shear velocity, $\Delta V \sim H\Omega$. In a marginally unstable disk, ΔV is of the order of the sound velocity, so the collapse begins quasi spherically but it would be rapidly dominated by rotational motions unless it gets rid of a large fraction of its angular momentum. This is the case as a large proportion of the angular momentum is given to the protostellar disk, and may also be transformed into orbital motion in binaries or multiple systems. One can also show that under the tidal action of the central mass the small disk is synchronized in about a dynamical time, and this mechanism leads to the suppression of another fraction of the angular momentum. So finally the collapse must proceed as in the spherical case with a characteristic time t_{collapse} .

Once the protostar and the protostellar disk are formed, they undergo a mechanism of accretion and growing proposed by Artymowicz, Lin and Wampler (1993), for stars orbiting around the central black hole and being trapped in the disk (their mechanism is however different from ours). If the accretion would be limited to a region of radius H the accretion rate would be roughly $\dot{M}_{\text{accr}} \sim \Delta V \rho H^2 \sim \Omega m_{\text{frag}}$, and the corresponding accretion time for a star of $10M_{\odot}$ would be equal to $10M_{\odot}/(\Omega m_{\text{frag}})$. The real accretion time is shorter, as one should take into account the mass inflow coming from beyond H and swept up by the Roche lobe, which cannot be estimated simply.

A last important time is t_{migr} , the time for the stars to migrate towards the black hole owing to the mechanism of induced density waves previously mentioned and discussed by Goldreich and Tremaine (1980), Ward (1986, 1988), and Lin and Papaloizou (1986a and 1986b). One can verify that in the solutions found for the disk it is comfortably large, so the stars do not have time to migrate during their evolution. One interesting consequence is that the residual

neutron stars left after the supernova explosions, if they are not ejected from the disk, can undergo other accretion phases, leading to other (presumably powerful) supernova explosions.

All these times are displayed on Fig. 2, and we will see that they satisfy the requirements for star formation.

2.2. Supernovae explosions and angular momentum transport

The expansion of a supernova in the radial direction is stopped at a radius $R_{s,\max}$ when the expansion velocity is equal to the shear velocity at its edge. The shell can still expand in the azimuthal direction and will therefore be stretched and sheared by the differential rotation. It will disappear after about one rotation time. The shell always reaches the disk surface, so the interior is depressurized and the expansion is driven only by momentum conservation. A large fraction of that momentum escapes from the disk, and the momentum P_{disk} supplied to the disk is equal to a fraction of the total momentum P_{tot} carried by the supernova explosion $\sim H/R_{s,\max}$. Since the terminal velocity is of the order of $3P_{\text{disk}}/(4\pi\rho HR_{s,\max}^2)$, one gets $R_{s,\max} \sim 3P_{\text{tot}}/(4\pi\rho\Omega)$.

One can deduce the maximum rate of supernovae supported by the disk without being destroyed. As the cavities created by the blast waves are replenished roughly at the shear velocity, their lifetime is $\sim 1/\Omega$, and this rate is of the order of $\Omega(R^2/R_{s,\max}^2)$. Actually this estimation is rather conservative, as the numerical simulations of Rozyczka et al (1995) show that when the cavity and the shock wave reach their maximum radial extension after a fraction of an orbital time and become strongly elongated, they are also strongly squashed, so the surface of the perturbation seems to be rapidly decreasing.

Supernovae produce a net transfer of angular momentum towards the exterior. The net angular momentum supplied by one supernova is equal to $\Delta J \sim (3/2\pi)P_{\text{tot}}RH/R_{s,\max}$. If all the angular momentum required to sustain the accretion rate is carried by the supernovae, their rate $\mathcal{N}_{\text{SN}} \sim 2\dot{M}R^2\Omega/\Delta J$.

2.3. Solutions

A self-regulation mechanism for the gas density has been proposed for the Galactic disk by Wang and Silk (1994), through the growth time of gravitational instabilities. According to them, the rate of gas transformed into stars is $d\Sigma_g/dt = \Sigma_g \Omega^{-1}\epsilon\sqrt{1-Q^2}/Q$, ϵ being an ‘‘efficiency’’ of star formation which is of the order of 0.1% for massive stars in the Galaxy. Wang and Silk have included the stellar contribution in the Toomre parameter Q , as given by a two fluid approximation. Here it should not be taken into account, as the mean distance between stars exceeds the instability length H .

Since the formation time and the lifetime of the stars are both small compared to the growth time of the black hole or to the active phase of a quasar, a stationary state of the disk can be established if there is a steady mass inflow from outward. One can determine then the star formation rate from the number of supernovae able to sustain the required accretion rate, and solve the radial disk structure in a self-consistent way.

The (vertically averaged) equations for the disk structure are the hydrostatic equilibrium equation, including the stellar contribution, and the energy equation, which should take into account 3 heating mechanisms: 1. heating by the fraction

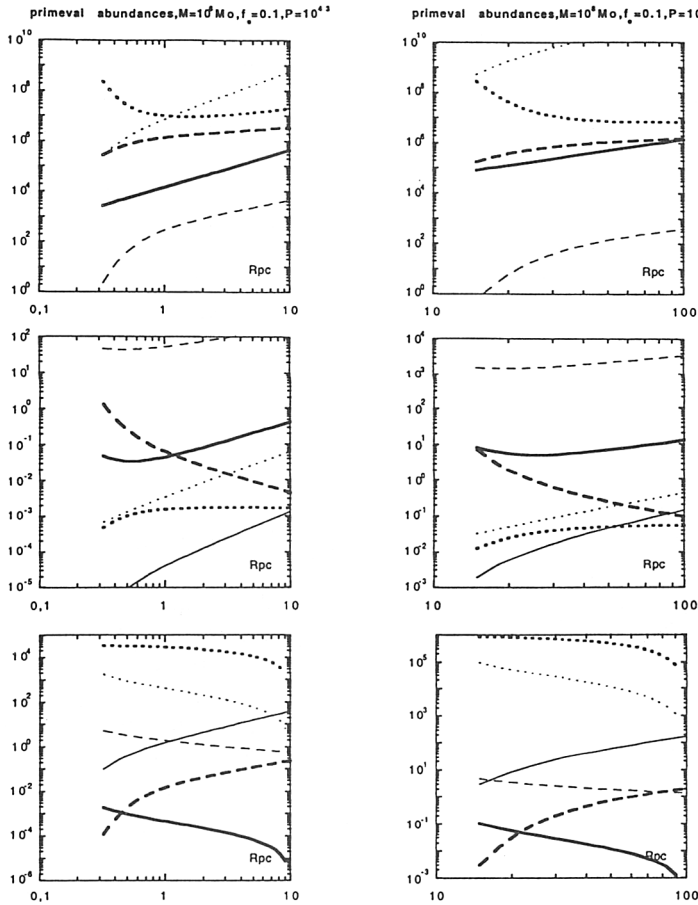


Figure 2. Solutions for the disk made of stars and gas, for primeval abundances. In the top panels are given the characteristic times in year: bold solid lines: orbital time; bold dashed lines: t_{trans} ; bold dot lines: $t_{\text{accr}}(10M_\odot)$; thin dashed lines: t_{cool} ; thin dot lines: $t_{\text{migr}}(10M_\odot)$. In the middle panels, are given several parameters allowing to check that the conditions required by the model are fulfilled: bold solid lines: ratio of the rate of supernovae to the maximum allowed rate; bold dashed lines: ratio of the number of stars to the maximum number allowed by the tidal effect; bold dot lines: ratio of the number of stars to the maximum number allowed by the HII regions; thin solid lines: $\epsilon\sqrt{1-Q_g}$; thin dashed line: m_{gap} in M_\odot ; thin dot lines: $V_{\text{rad}}/V_{\text{Kep}}$. In the bottom panels, several physical parameters of the models: bold solid lines: the rate of supernovae \mathcal{N}_{SN} per year, integrated from the outer edge; bold dashed lines: m_{frag} in M_\odot ; bold dot lines: the mass in gas M_{gas} integrated from the outer edge; thin solid lines: $H/10^{16}\text{cm}$; thin dashed lines: the midplane temperature in 100K $T/100\text{K}$; thin dot lines: the number of stars N_* integrated from the outer edge.

of the stellar luminosity not absorbed in the HII regions, 2. heating by the central source, when the disk flares (as is always the case in this model), 3. heating due to the dissipation of kinetic energy through shock waves induced by supernova shells (it corresponds to the viscous heating in a standard disk). One gets then a complete algebraic system, which can be solved as a function of the central mass and of the Eddington ratio f_E (here equal to \dot{m} for an efficiency factor equal to 10%). As an illustration, Fig. 2 displays the results for two solutions corresponding to primeval abundances, to an Eddington ratio equal to 0.1, and to black hole masses of 10^6 and $10^8 M_\odot$. The luminosity of the stars is set to 10^{38} ergs s^{-1} , their average mass is $30 M_\odot$, and we have assumed $t_{\text{lifetime}} = 3 \cdot 10^6$ yrs and $P = 10^{43}$ g cm s^{-1} . The solutions for solar abundances are not very different.

First are displayed the characteristic times. Except for large black hole masses and accretion rates, the free fall time is about two orders of magnitude less than the mass transport time, which means that $\sqrt{1 - Q_g^2}$ can be as small as 10^{-2} . The assumption that the gas is marginally unstable is therefore self-consistent. The accretion time is at most a few 10^6 yrs, and we have seen that it is probably strongly overestimated. Finally the cooling time is very small and does not set any constraint on the collapse.

We give then several parameters: $\epsilon\sqrt{1 - Q_g}$, m_{gap} , the ratio of the radial to the Keplerian velocity $V_{\text{rad}}/V_{\text{Kep}}$, the ratio of the rate of supernovae to the maximum allowed rate, the ratio of the number of stars to the maximum number allowed by the tidal effect (if the stellar density is too large, star formation is inhibited), and the ratio of the number of stars to the maximum number allowed by the HII regions. These parameters allow to check that the conditions required by the model are fulfilled. One can also check from the product $\epsilon\sqrt{1 - Q_g}$ that a value of ϵ of $\sim 10^{-3}$ is compatible with the condition $\sqrt{1 - Q_g^2} \geq 10^{-2}$. The number of stars is smaller than the number allowed by the HII regions and by the tidal effect. The most constraining condition comes from the maximum allowed number of supernovae. For $M = 10^8 M_\odot$ the number of supernovae is larger than this maximum number. It is comfortably smaller only for small black hole masses or if the momentum provided by a supernova is larger than the standard value. However we have seen that the maximum rate of supernovae is underestimated, so we are confident that the cases where the rate of supernovae is slightly larger than the maximum allowed value are also viable.

Finally the figures show some interesting physical parameters: the scale height and the midplane temperature of the disk, the mass of the initial fragments, the gaseous mass of the disk, the total number of stars, and the total supernova rate. The last three quantities are integrated from the external edge. One can check that the total mass of the disk (actually of the domain where the model is valid) is much smaller than the black hole mass. The number of stars is small, corresponding to an amount of mass locked in stars of the order of that of the gas. Finally the total supernova rate, which is entirely dominated by the inner regions, varies from 10^{-3} to $1 M_\odot$ per year. We shall come back to this point in the next section.

Although we have considered a stationary model, and shown that it can be accommodated with the picture of a gas-star disk, the most realistic scenario is

that the mass inflow from the periphery of the disk is variable with time, as would be the case if it were achieved by large molecular complexes comparable to those present near the Galactic center. In this case, there will be “low states” and “high states” where the disk would be alternatively “quiescent”, and “active”, i.e. highly perturbed by an intense supernova activity, and star formation will occur in successive “waves” propagating from the outer to the inner regions. An inflow of matter will induce an increased gas density in an outer ring. Before the stars form, accrete and evolve to supernovae, the transfer of mass will not take place, and there will be an accumulation of gas in the ring. After a few 10^6 years, supernovae will induce mass transport towards an inner ring, while the outer ring will be cleared out of its gas until a new mass inflow. Note that during the “active” phases the mass inflow at the periphery could be super Eddington. In this case our description should be modified to take into account the non-stationarity of the process, as one would expect that the corresponding averaged momentum transport (i.e. the accretion rate) be much larger than in the stationary case, as it is not limited by a maximum allowed rate of supernovae.

3. Some implications of the model

In this model the mass accretion rate is of the order of the outflowing mass due to supernovae, although the luminous energy released in supernovae is much smaller than that of the central regions of the disk. Therefore nearby low luminosity AGN should display an increase of flux in the optical range every 10^3 years with the typical light curve of a supernova, and one should also expect to resolve spatially this supernova at a distance of 1-10pc from the central source.

This model has other consequences. Beside the fact that it solves the problem of the mass transport in the intermediate region of the disk, it could give an explanation for the high velocity metal enriched outflows implied by the presence of the Broad Absorption Lines (BALs) in quasars. This problem is discussed in details in Collin (1998), so we recall here only a few points.

There are strong observational evidences for metallicities larger than solar (or at least solar) in the central few parsecs of quasars up to $z \geq 4$ (see the recent systematic study of Hamann 1997). This enriched material is flowing out of the central regions with a high velocity, of the order of $c/30$, as observed in BAL QSOs, which constitute about 10% of the total number of radio quiet QSOs. The phenomenon is generally interpreted as an outflow existing in all quasars, but limited to an opening angle $\sim 4\pi/10$. The outflowing mass rate is quite difficult to estimate, say between 1% to 100% of the accretion rate.

Comparing the observed mass outflow with that given by supernovae, each releasing about $10 M_{\odot}$ of metals out of the nucleus, one sees that the mass outflow rate due to the supernovae accounts easily for the observations. Our computations show indeed that the rate of supernovae is equal to a few 10^{-2} yr^{-1} for a quasar black hole, i.e. an outflow of metals close to the accretion rate. The observed velocities are also easily accounted for by the expanding shells, and finally the location of the phenomenon is in agreement with observations. Relatively larger enrichment in some elements like N and Fe are observed, and could be explained by the oddness of the stars formed in a particular environment. Finally the fact that the opening angle of the BAL region is equal to a

small fraction of 4π is easily explained, the ejection taking place mainly in a cone aligned with the disk axis. Note that in this model the metallicity of the gas fuelling the black hole can be very small, while the observed outflow is always enriched.

A second outcome of the mechanism is to account for a pregalactic enrichment, if massive black holes are created early in the process of galaxy formation, and if galaxy formation takes place through an hierarchical scenario (Silk and Rees, 1998). Massive galaxies will retain their gas, and the supernova shells will compress the interstellar medium, trigger star formation like in the interstellar medium, and induce a starburst. This is an “inside-outside” scenario opposite to the starburst scenario of Hamann and Ferland (1992). A fraction of the enriched gas should however escape from the galaxy, not only due to the nuclear supernovae explosions, but also to the induced starburst. Small galaxies will not retain their gas, which will escape with the enriched gas produced by the supernovae. It will pollute the intergalactic medium (IGM). In particular, if the formation of the black holes precedes the formation of galaxies, it will lead to a pregalactic enrichment of the IGM. According to our computations the mass of metals ejected by the disk is of the order of the mass of the black hole itself. We can therefore estimate the **minimum** enrichment of IGM due to black holes, simply taking the integrated comoving mass density of **observed** quasars, which corresponds to about 10^{-6} of the closure density (Soltan, 1982, and further studies). If these black holes have a typical mass of $10^8 M_{\odot}$, the present mechanism will provide about 10^{-6} of the closure density in metals, i.e. after mixing with the IGM, an average metallicity of a few $10^{-3} \Omega(\text{IGM})_{0.02} Z_{\odot}$, close to the metallicity observed in the $L\alpha$ forest which constitutes the main fraction of the IGM.

Finally, our Galaxy is presently not active and the black hole in the center has a small accretion rate ($< 10^{-4} M_{\odot} \text{ yr}^{-1}$), so it has most probably accreted a large fraction of its mass ($2 \times 10^6 M_{\odot}$) during an early period. The previous estimation leads to an ejection of a few $\sim 10^4 M_{\odot}$ of metals. After mixing with a hydrogen halo of $10^{11} M_{\odot}$, it gives a metallicity of a few 10^{-5} solar, close to that observed in the oldest halo stars.

References

- Artymowicz P., Lin D.N. and Wampler E.J., 1993, *ApJ* 409, 592
 Begelman, M.C., Frank J., Shlosman I., 1989, in “Theory of Accretion disks”, eds. F. Meyer et al., Kluwer Academic Publishers
 Chandrasekhar S., 1939, “An introduction to stellar structure”
 Collin S., Huré J.M., 1998, in press in *A&A*
 Collin S., 1998, to appear in the proceedings of the Conference “From atomic nuclei to galaxies”, held in Haifa, Israel, Physics Reports, Ed. O. Regev
 Collin S., Zahn J.P., 1998, *A&A*, submitted
 Goldreich P., Lynden-Bell D. 1965, *MNRAS* 130, 97
 Goldreich P., Tremaine S., 1980, *ApJ* 241, 425
 Hamann F., 1997, *ApJS* 109, 279

- Hamann F., Ferland G., 1992, *ApJ* 391, 53
Huré J.M., 1998, in press in *A&A*
Kolykhalov P.I., Sunyaev R.A., 1980, *Soviet Astron. Lett.* 6, 357
Lin D.N.C., Papaloizou J.C.B., 1980a, *ApJ* 307, 395
Lin D.N.C., Papaloizou J.C.B., 1980b, *ApJ* 309, 846
Malkan M.A., Sargent W.L.W., 1982, *ApJ* 254, 22
Paczynski B., 1978, *AcA*, 28, 91
Rozyczka, M., Bodenheimer P., Lin D.N.C., 1995, *MNRAS* 276, 597
Shakura N.I., Sunyaev R.A., 1973, *A&A* 24, 337
Shields, G.A., 1978, *Nature* 272, 423
Shlosman I., Begelman M.C., 1989, *ApJ* 341, 685
Shlosman I., Frank J., Begelman M.C., 1989, *Nature* 338, 45
Silk J., Rees M.J., 1998, *A&A* 331, L1
Soltan, A., 1982, *MNRAS* 200, 115
Toomre A., 1964, *ApJ* 139, 1217
Wang B., Silk J., 1994, *ApJ* 427, 759
Ward W., 1986, *Icarus* 67, 164
Ward W., 1988, *Icarus* 73, 330



Eduard Denissyuk and Francesco Bertola