Mon. Not. R. astr. Soc. (1976) **175**, 613–632.

A THEORY OF THE INSTABILITY OF DISK ACCRETION ON TO BLACK HOLES AND THE VARIABILITY OF BINARY X-RAY SOURCES, GALACTIC NUCLEI AND QUASARS*

N. I. Shakura

Shternberg Astronomical Institute, Moscow

and R. A. Sunyaev

Institute of Space Reseach Moscow

(Received 1975 October 30)

SUMMARY

Stationary accretion in the inner region of the disk is possible only for one definite value of the viscosity:

$$\eta_c = \frac{4}{9}\frac{m_p}{\sigma_T c}.$$

A linear stability theory is constructed for stationary disk accretion on to black holes. The inner region of the disk, where radiation pressure dominates, is unstable against small perturbations. Growing short wave perturbations take the form of travelling concentric waves. For long wavelengths there are two branches of growing standing waves. The growth rate of one branch decreases rapidly with increasing wavelength; its asymptotic behaviour was found previously by Lightman & Eardley. The growth rate of the other branch increases to a constant limit as wavelength increases; this branch is the thermal instability of disk accretion.

The above instabilities can explain the observed variability of radiation from binary X-ray sources, galactic nuclei and quasars, assuming these objects really do contain accreting black holes.

1. THE GENERAL PICTURE

Accretion of matter with significant angular momentum on to black holes is accompanied by the formation of disks of accreting material. Disks should arise around both black holes of stellar origin in binary systems and supermassive black holes possibly found in galactic nuclei and quasars. The structure and radiation of stationary disks, the theory of which was constructed in the works of Lynden-Bell (1969), Shakura (1973), Shakura & Sunyaev (1973) and Pringle & Rees (1973),† is determined by three parameters: the mass of the black hole M, the accretion rate $\dot{M}$ (or the total disk luminosity $L = \zeta \dot{M} c^2$, where ζ is the efficiency

* Translated from the Russian by R. L. Znajek.

† Relativistic effects were calculated in the survey by Novikov & Thorne (1973), and also in papers by Page & Thorne (1974) and Cunningham (1975), and a thesis by Polnarev (1975).

of energy generation) and the parameter α, characterizing the level of turbulence and/or chaotic small-scale magnetic fields:

$$\alpha = \frac{v_t l_t}{v_s H} + \frac{B^2}{4\pi \rho v_s^2}. \tag{1.1}$$

Here v_t and v_s are respectively the turbulent and thermal velocities of the matter, $B^2/8\pi$ is the energy density of the chaotic magnetic field, $\rho v_s^2/2$ is the thermal energy density of the matter in the disk, l_t is the turbulent mixing length and H is the half thickness of the disk.

The luminosity of an accreting X-ray source is apparently bounded by a quantity of the order of the Eddington luminosity

$$L_c = 1 \cdot 3 \times 10^{38} \left(\frac{M}{M_\odot}\right) \text{erg s}^{-1}$$

obtained by equating gravitational and radiation-pressure forces on plasma near the inner edge of the disk. If matter flows through the outer boundary at a rate substantially above $\dot{M}_c$, the maximum allowed by the Eddington limit, then gas should flow away perpendicularly from the inner region of the disk under the action of the radiation pressure (Shakura & Sunyaev 1973 (referred to below as SS)). In the following we shall be mainly concerned with disk accretion for $\dot{M} < \dot{M}_c = L_c/\zeta c^2$.

At luminosities close to the Eddington limit, $L/L_c \gtrsim \frac{1}{50}(M_\odot/\alpha M)^{1/8}$, the disk consists of three distinct regions. There is the inner zone A, in which radiation pressure is greater than gas pressure and free electron scattering dominates over processes of true radiation absorption. This region emits the dominant part of the energy radiated by the disk, and Compton processes determine the appearance of the observed radiation spectrum. Further away from the hole is zone B, where plasma pressure is greater than radiation pressure, but free electron scattering still dominates radiation transfer. Beyond zone B there may exist a coldest region C where true absorption processes dominate the opacity.

One can show (see Section III) that physical quantities in zone A can be elegantly expressed using the universal constants m_p, c, σ_T and $R_g = 2GM/c^2$. It also follows, from the equations of motion and the condition of energy balance (setting the energy emitted per unit area of the disk Q^- at any radius equal to the rate of energy-generation in the disk Q^+), that stationary disk accretion in zone A can only occur when the viscosity takes one definite value

$$\eta_c = \frac{4}{9} \frac{m_p}{\sigma_T c}. \tag{1.2}$$

This quantity does not depend on the actual form of the viscous forces (turbulent or magnetic), distance from the black hole, luminosity or other parameters of the accretion process. The viscosity given in (1.2) substantially exceeds the maximum possible viscosity for a fully ionized gas and for radiation. Any difference between the quantity (1.2) and the value of the viscosity in the inner zone has to be attributed to a difference between Q^+ and Q^- and the violation of stationarity. The problem of the stability of stationary disk accretion naturally arises.

Rees, Pacholczyk & Pringle (1973), and subsequently Lightman & Eardley (1974), raised the question of the stability of disk accretion. Lightman & Eardley (1974), and also Lightman (1974a, b) investigated the stability of a dynamic equation for disk accretion and came to the conclusion that stationary accretion is unstable in zone A and stable in zones B and C. This fundamentally important and correct result was, however, derived from assumptions that were not completely correct: the dynamic equation relates two most important quantities—the surface density of matter in the disk U and its half-thickness H, and cannot by itself answer the question of the stability of stationary flow. A second equation connecting U and H is required. Lightman & Eardley simplified the problem by using the relation between U and H that occurs in stationary flow. But this relation is a consequence of the equality $Q^+ = Q^-$, which also implies the unlikely equality between the viscosity in the disk and its critical value $\eta_c = \frac{4}{9}(m_p/\sigma_T c)$. In general the viscosity might not be η_c and $Q^+ \neq Q^-$, and also it is necessary to include work done against pressure as the disk expands and contracts, and the mechanical transfer of energy by viscous stresses between regions of the disk at different radii.

Below we shall present the supplementary thermal equation relating U and H when the equalities $Q^+ = Q^-$ and $\eta = \eta_c$ are not satisfied. It is derived from the law of energy conservation and allows a correct analysis of the stability of stationary accretion disks with respect to radial perturbations to be carried out.

Using the non-stationary equations of disk accretion (see below) the axially-symmetric problem was considered and it was assumed that the characteristic radial length scale of the perturbations, Λ, satisfied the condition $H_0 < \Lambda < R$ where $2H_0$ is the thickness of the disk at any given radius (quantities with the suffix zero refer to the unperturbed stationary disk), and that terms of order $(H_0/R)^2$ and $H_0^2/R\Lambda$ were negligible when compared with terms of order $(H_0/\Lambda)^2$. The dynamic and thermal equations derived in this manner form a closed system of equations; analysis of their solutions demonstrated the existence of two branches of perturbations that grow with time (see Fig. 1).

The time-dependent equations of disk accretion and their analysis were presented by us in a short note (Sunyaev & Shakura 1975). Recently we received a preprint by Shibasaki & Hochi (1975), which also demonstrates the possibility of thermal instability connected with the inequality $Q^+ \neq Q^-$.

Perturbation types and their growth rates

In the limit of dominant radiation pressure ($\beta_0 \equiv p_r/(p_r+p_g) \simeq 1$) perturbations with length scales satisfying $2H_0 < \Lambda < 4H_0$ take the form of concentric rings moving across the surface of the disk. Their growth-rate Ω varies from $\Omega = 0$ at $\Lambda = 2H_0$ to $\Omega \simeq \alpha\omega/10$ at $\Lambda = 4H_0$ (ω is the angular velocity of the disk at a given radius).

For wavelengths $\Lambda \geqslant 4H_0$ there are two branches of growing perturbations, which take the form of standing waves. For the lower branch (see Fig. 1) the growth-rate drops with increasing wavelengths, and in the limit $\Lambda \gg H_0$ the amplitude of the viscosity perturbations η_1/η_0 is only a small fraction of the amplitude of the surface density perturbations U_1/U_0 and height perturbations H_1/H_0. Thus the equality $Q^+ = Q^-$ is only valid on this branch for long wavelengths, and the solution turns out to be the instability discovered by Lightman & Eardley (1974). But the growth-rate of this branch is small: $\Omega \approx \alpha\omega(H_0/\Lambda)^2$. One is more interested in the other branch of standing waves, where growth-rate increases with increasing wavelength, and in the limit $\Lambda \gg H_0$ tends to the value $\Omega \simeq 0.2\alpha\omega$. On this branch when $\Lambda \gg H_0$ the amplitude of the surface-density perturbations U_1/U_0 is small in comparison with perturbations in the disk's thickness H_1/H_0, viscosity and other

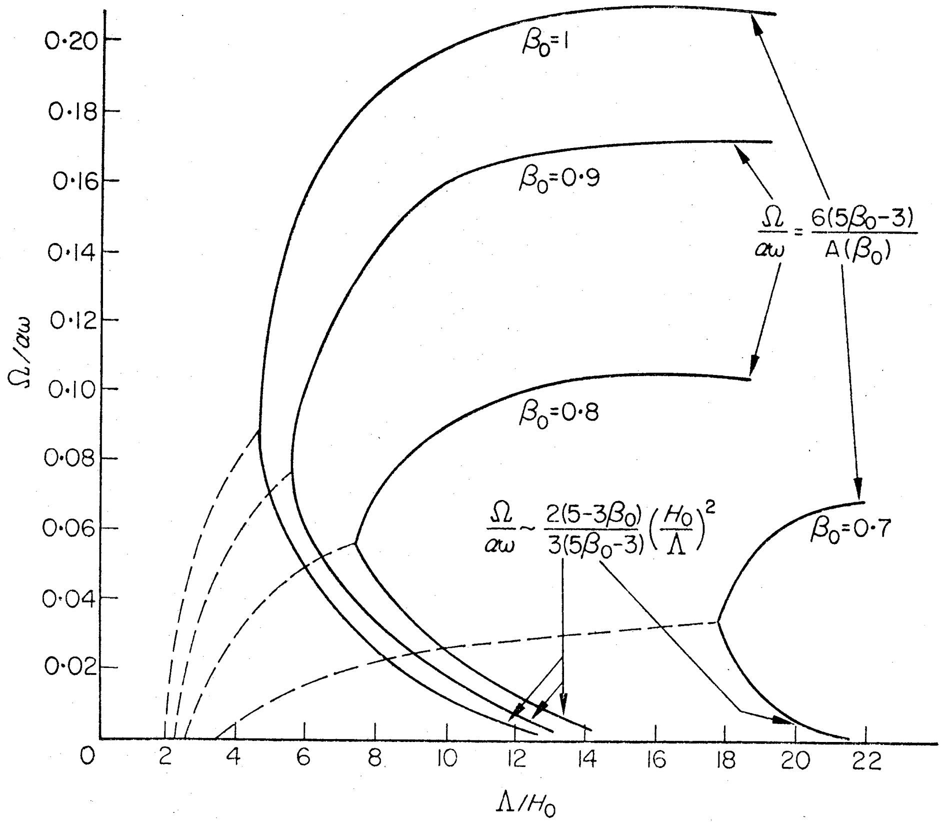

FIG. 1. *Dependence of the instability growth rate on wavelength when radiation pressure dominates, $\beta_0 > \frac{3}{5}$. For $\beta_0 < \frac{3}{5}$ disk accretion is stable. Broken lines denote travelling waves, continuous lines denote standing waves. Upper branches correspond to thermal instabilities and lower branches to dynamic instabilities. The long wavelength limit of the lower branches is the instability discovered by Lightman & Eardley.*

physical quantities. The growth of these perturbations is due to a thermal instability; the equality $Q^+ = Q^-$ does not hold on this branch.

If the importance of radiation pressure is reduced, i.e. with smaller β_0, the maximum value of the growth-rate of the thermal instability falls, the wavelength at which growing perturbations are just possible increases, and in the limit $\beta_0 \to \frac{3}{5}$, we have $\Omega \to 0$ and $\Lambda_{\min} \to \infty$. (Recall that for $\beta_0 = 1$ the wavelength of the smallest non-diminishing perturbation is equal to the thickness of the disk ($\Lambda_{\min} = 2H_0$).) Both branches of the solution then become the line $\Omega = 0$.

For $\beta_0 < \frac{3}{5}$ we only have branches with negative Ω, i.e. small perturbations will decay in a region where plasma pressure dominates. Thus the boundary between the stable and unstable zones of the disk is given by $\beta_0 = \frac{3}{5}$.

The linear stability theory presented here does not take account of boundary conditions. We shall accept its conclusions when a large number of wavelengths can fit into a region whose size is characteristic of changes in the physical quantities β_0, ω, etc., of the unperturbed solution (i.e. when its size is of order R). Perturbations with $\Lambda \sim R$ are described by more complicated equations. Their growth necessitates a continual increase in mass-flow $\dot{M}$ across the boundary of the stable and unstable zones. However, because of the stability of the solution in the outer zone from where the matter is flowing, $\dot{M}$ remains fixed and equal to $\dot{M}_0$, thus favouring stability. The constancy of $\dot{M}_0$ prevents the growth of perturbations on very large length scales, which would completely wreck the disk structure of the accretion. On the other hand small perturbations of length scale $\Lambda \ll R$ do not destroy the disk, but give rise to quasi-periodic temporal fluctuations in density, thickness, energy-generation and consequently in the emergent radiation flux. The spectral composition of the radiation will also vary.

Non-linear regime

An individual mode of growing standing perturbations consists of alternating nodes and antinodes. At a node the importance of radiation pressure diminishes, and when $\beta = \frac{3}{5}$ is reached, the instability stabilizes there. At an antinode, where $\beta \to 1$, the growth of H relative to H_0 enables additional energy to be radiated through the sides of the ring; this increases Q^- and of course stabilizes the situation. But the main factor hindering the growth of the perturbation at antinodes appears to be the cessation of mass flow $\dot{M}$ into these regions because of the stabilization of the accretion rate in the nodes. On the other hand an antinode with wavelength Λ might break up into smaller perturbations, and the latter in their turn into perturbations on even smaller scales, right up to $\Lambda \simeq 2H$. In a non-linear regime the antinodes determine the accretion rate into regions of smaller radius, i.e. the accretion rate $\dot{M}(R, t)$ and luminosity in the inner region change with time.

Evidently, near the boundary separating the stable and unstable zones, where $\beta_0 = \frac{3}{5}$, long-wave perturbations grow slowly, and determine the rate at which matter flows into the inner region of the disk. The flow proceeds in 'batches', and the instability time scale at the outer zone, i.e. the formation time of these regions $t \sim 1/\Omega(R_A) \sim (R_A/H_0)^2(1/\alpha\omega(R_A))$ is the longest time-scale for perturbations of disk accretion. Because the time required for growth to non-linearity at the outer boundary is approximately constant, the formation time of separate blocks is more or less constant as well. This could be the cause of the quasiperiodicity of the observed flux for long time scales. At smaller radii the quantities ω and β_0 are bigger, and so the growth of perturbations is also bigger. The smallest fluctuation time $t \sim 1/\Omega(R_0) \sim 1/\alpha\omega(R_0)$ is associated with the region near the inner edge of the disk. Thus the examined instability should cause fluctuations in the luminosity of the disk, the amplitude of the rapid fluctuations being modulated by the slower ones, and a fluctuation spectrum should show variations on all time scales from a minimum to a maximum. As $\dot{M}_0$ is fixed at the outer edge of the disk, averaging over times $t > t_{max}$ should always give the same value for the luminosity of the disk in all wave bands.

Azimuthal perturbations

The thermal instability should undoubtedly have azimuthal as well as radial modes, with roughly equal growth rates. For perturbations of long wavelength, the azimuthal modes should lead to the appearance of a spiral structure. For small length scales the joint action of azimuthal and radial modes should cause hills and hollows to appear on the surface of the disk. Energy generation would be maximum at the hills, i.e. they should be observed as hot spots on the disk surface. As a result of the Doppler effect the rotation of hot spots around the black hole should lead to quasi-periodic variations in luminosity with the orbital period $T = 2\pi/\omega$ (Sunyaev 1973).

Spectral changes

It has already been remarked in SS that when the total luminosity of the disk is close to the Eddington limit and $\alpha \sim 1$, Compton processes in the region of greatest energy generation produce hard X-rays. The derived spectrum closely resembles the observed spectrum of Cyg X-1. If, however, the luminosity is several times smaller than the Eddington limit, hard X-rays are absent during stationary accretion. Instability in the radiation pressure zone leads to the appearance of a hard X-ray tail at substantially smaller luminosities, but not smaller than $L/L_c = \frac{1}{50}(M_\odot/\alpha M)^{1/8}$. The point about the latter limit is that there the unstable region where $\beta_0 > \frac{3}{5}$ disappears. The instability of the accretion produces regions (in the antinodes) where $\dot{M}$ grows relative to its mean value, and can reach the critical value. This implies an increase in the local rate of energy generation, and a rise in the electron temperature. Compton processes harden the spectrum, giving it a characteristic hard tail. Meanwhile in the regions with high plasma density and relatively low temperature (the nodes), bremmstrahlung processes generate a large number of low-frequency photons. These diffuse into the regions of high energy generation (the antinodes) where it is difficult for new photons to be created; there they are Comptonized and help in removing energy from the high-temperature regions.

It should be said that the typical spectrum of an accretion disk, derived in SS, can be approximated by a sum of several power-law spectra. Thus if observers say that in this or that narrow part of the spectrum the observations are consistent with a power law, one cannot, on that basis alone, deduce that the radiation from any particular object is non-thermal. As hard radiation comes from the most perturbed region, it should vary strongly with time. Softer radiation, emitted by the outer regions where the instability grows slowly or is completely suppressed, should fluctuate less.

Black holes in binary systems

Up to the present time all the properties of Cyg X-1 indicate that it is the most plausible black hole candidate. The variability of its luminosity is naturally explained by the model proposed here. In the spring of 1971 observations by UHURU revealed a sharp transition in the distribution of energy radiated by the source. In the 1–6 keV range the energy flux fell by nearly a factor of 10, and there appeared simultaneously a high energy tail in the X-ray range $h\nu > 10$ keV.

The overall X-ray luminosity did not change, or perhaps it even increased slightly. Analysis of the observations has shown that the X-ray flux fluctuates on time scales from about 10 s down to a few milliseconds. The amplitude of the millisecond variations is greater at higher energies. On averaging over more than 10 s, it was found that the fluctuations were smoothed out in all parts of the energy spectrum. The source was observed in this condition for several years. In the spring of 1975 observations from the X-ray satellites *Ariel* and *ANS* and the space station *Salyut*-4 showed a reversal: the 1–6 keV X-ray flux increased by about a factor of 10, while the flux at around 8 keV stayed practically unchanged. Unfortunately there is no data on the hard part of the spectrum. It is natural to associate these transitions with the appearance and disappearance of the unstable zone in an accretion disk around a black hole in Cyg X-1. The X-ray luminosity of Cyg X-1, $L_X \simeq (0{\cdot}6\text{–}1{\cdot}0) \times 10^{38}$ erg s^{-1}, is close to the value above which the instability appears, assuming the mass of the hole to be $\sim 10\,M_\odot$. When $L/L_c < 1/50$, a large part of the energy is radiated in the 1–6 keV region, and hard quanta are absent. But if the luminosity is increased to slightly above the value $L \sim L_c/50$, the disk becomes unstable in the region of maximum energy generation. $\dot{M}$ becomes variable, and where $\dot{M}$ is close to the critical value, hard X-rays are generated, while the 1–6 keV flux falls because of Compton processes. Thus when the luminosity is close to $L_c/50$ a small change in the accretion rate $\dot{M}$ can cause a substantial redistribution of the energy radiated by the disk. Another mechanism capable of influencing the spectral composition of the radiation is any variation of the parameter α in the accreting matter. It would seem natural for turbulence to be more easily generated in the unstable region of the disk, and for the parameter α to grow there.

Supermassive black holes

The first model of disk accretion on to supermassive black holes was proposed by Lynden-Bell (1969) with the aim of explaining the observational appearances of galactic nuclei. It was then successfully developed by Lynden-Bell & Rees (1971), and others.

Like the theory of stationary disk accretion, the theory of its instability can be applied to black holes with very different masses. However, when we change over to supermassive objects, the observational appearance of disk accretion is completely different from when the black holes are of stellar origin.

(1) When the mass of the black hole is increased the Eddington limit grows linearly: $L_c \simeq 10^{38} (M/M_\odot)$ erg s^{-1}, and black holes with $M \simeq 10^8$ or $10^9 M_\odot$ can through accretion radiate at 10^{46} or 10^{47} erg s^{-1}.

(2) The gravitational radius R_g and the radius of the last stable circular orbit ($3R_g$ for the Schwarzschild metric and $R_g/2$ for the extreme Kerr metric), which corresponds to the inner edge of the disk, are both proportional to M. Thus the area of the region which radiates most of the energy is effectively proportional to M^2, and the flux of energy radiated per unit area $Q \sim L/R^2 \sim M^{-1}$. If the disk surface radiated like a black body, then the radiation temperature, given by the Stefan–Boltzmann law $Q = bT_{\text{eff}}^4$, would vary like $M^{-1/4}$. The transition from an X-ray source with $M \simeq 10 M_\odot$ to $M \simeq 10^9 M_\odot$ with $L/L_c \simeq 0{\cdot}1$–1 remaining unchanged is accompanied by the temperature dropping by a factor of 100 to $\sim 10^5$ K and energy being mostly radiated in the ultraviolet.

(3) *Temporal variation*. The instability of disk accretion discovered here gives a minimum time for radiation fluctuations of $t_{\min} \approx 2\pi/\omega_{\max} \approx 2\pi\sqrt{R_g^3/GM} \sim M$, proportional to the mass of the black hole. If in the source Cyg X-1 ($L \simeq 10^{38}$ erg s^{-1}, $M \simeq 10 M_\odot$) variations are observed over times ranging from several milliseconds to tens of seconds, then for consistency with the model of disk accretion on to supermassive black holes ($L \simeq 10^{46}$ erg s^{-1}, $M \simeq 10^9 M_\odot$) one expects to observe variability in quasars over times ranging from a few days to many years.

Observations by Lyutiy and others (see Lyutiy & Pronik 1975) show the presence of quasi-regular and chaotic variability in the nuclei of the Seyfert galaxies NGC 4151, 1275 and the quasar 3C 273. Evidently these objects possess quasiperiodic phenomena of the same type as those observed in the X-ray source Cyg X-1. They only differ in the mass of the black holes, and consequently in the types of radiation and the variability time scales. If our model is correct, then by using a similarity transformation many observed properties of disk accretion on to black holes in binary systems can be transferred to quasars and galactic nuclei.

2. DYNAMICS OF DISK ACCRETION

We shall use cylindrical polar coordinates (R, ϕ, z) such that the z-axis is perpendicular to the plane of the disk, the latter being the plane $z = 0$.

To a first approximation matter in the disk flows along circular Keplerian orbits with velocity

$$v_\phi = \omega R = \sqrt{\frac{GM}{R}}. \qquad (2.1)$$

The Keplerian rotation law (2.1) is not that of a rigid body as $\partial\omega/\partial R \neq 0$, and so frictional forces between neighbouring layers, which are assumed to be much smaller than the centrifugal force $v_\phi{}^2/R$ and the gravitational force $-GM/R^2$, should redistribute angular momentum. Consequently a radial velocity component $v_\mathrm{r} \ll v_\phi$ will appear in the disk.

A useful variable for describing the disk is the surface density

$$U(R) = 2\int_0^H \rho(R, z)\, dz. \qquad (2.2)$$

Here $H(R)$ is the half-thickness of the disk at any given radius, while $\rho(R, z)$ is the volume density.

We shall consider a ring in the disk of width δR, between radii R and $R+\delta R$. The mass of this ring is $\delta M = 2\pi UR\,\delta R$ and the angular momentum is

$$\delta K = 2\pi U v_\phi R^2\,\delta R.$$

(a) *Equation of continuity*

The temporal variation in the mass δM of a ring of width δR is given by the difference $2\pi\,(\partial/\partial R)(Uv_\mathrm{r}R)\,\delta R$ between the rates at which matter flows in and out of the ring, because of the equation of continuity

$$\frac{\partial U}{\partial t} = -\frac{1}{R}\frac{\partial}{\partial R}(Uv_\mathrm{r}R) = \frac{1}{2\pi R}\frac{\partial \dot{M}}{\partial R}. \qquad (2.3)$$

Here and below $\dot{M}(R, t)$ is the mass crossing a cylindrical surface in unit time (the local accretion rate).

(b) *Equation of momentum transfer*

The variation in angular momentum δK is given by differences in the transfer of momentum due to radial motions $(\partial/\partial R)(Uv_\phi R^2 v_\mathrm{r})\,\delta R$ and frictional forces $(\partial/\partial R)(W_{\mathrm{r}\phi}R^2)\,\delta R$ and is described by the equation

$$\frac{\partial}{\partial t}(Uv_\phi R^2) = -\frac{\partial}{\partial R}(Uv_\phi R^2 v_\mathrm{r} + W_{\mathrm{r}\phi}R^2) \qquad (2.4)$$

where

$$W_{\mathrm{r}\phi} = 2\int_0^H w_{\mathrm{r}\phi}\, dz$$

is the viscous stress acting across an element $2HR\,d\phi$ of a cylindrical surface of radius R. $2\pi Wr\phi R^2$ is the torque due to viscous forces acting between one layer of matter and another.

Using the continuity equation (2.3) and ignoring deviations from the Keplerian rotation law $(\partial v_\phi/\partial t = R\,(\partial\omega/\partial t) = 0$ (see Appendix)) we obtain from (2.4)

$$\frac{\dot{M}\omega R}{2} = 2\pi\frac{\partial}{\partial R}(W_{\mathrm{r}\phi}R^2). \qquad (2.5)$$

It is clear from this equation that the direction of overall mass-flow (the sign of $\dot{M}$) is given by the dependence of $W_{r\phi}R^2$ on radius.

(c) *Energy dissipation in the disk*

The variation of the total kinetic and potential energies

$$\frac{\partial}{\partial t} 2\pi UR\left(\frac{\omega^2R^2}{2}-\frac{GM}{R}\right)\delta R$$

in the ring is determined by the differences between the rates of flow of these energies into and out of the ring

$$\delta R\,\frac{\partial}{\partial R}\,2\pi Uv_rR\left(\frac{\omega^2R^2}{2}-\frac{GM}{R}\right),$$

and also by the difference in work being done by viscous stresses at the sides of the ring $\delta R\,(\partial/\partial R)\,2\pi W_{r\phi}\omega R^2$ and irreversible energy losses into heat due to friction and differential rotation $\delta R 2\pi W_{r\phi}R^2\;(\partial\omega/\partial R)$. Putting everything together, we obtain the equation

$$\frac{\partial}{\partial t}\,UR\left(\frac{\omega^2R^2}{2}-\frac{GM}{R}\right) = -\frac{\partial}{\partial R}\left(Uv_rR\left(\frac{\omega^2R^2}{2}-\frac{GM}{R}\right)\right) -\frac{\partial}{\partial R}(W_{r\phi}\omega R^2)+W_{r\phi}R^2\frac{\partial\omega}{\partial R}. \quad (2.6)$$

This equation does not include work done by pressure forces nor the kinetic energy of the radial motion, which in this case are smaller by a factor of order $\alpha(H/R)^2$ than the rate at which energy is dissipated into heat. The full energy equation can be found in Lynden-Bell & Pringle (1974). Using the continuity equation (2.3), expression (2.6) can be simplified. The energy dissipation into heat in a 'unit column' is twice as great as the rate at which energy is radiated away per unit surface area (the disk has two sides):

$$Q^+ \equiv -\tfrac{1}{2}W_{r\phi}R\frac{\partial\omega}{\partial R} = \frac{\dot{M}\omega^2}{8\pi}-\frac{1}{2R}\frac{\partial}{\partial R}(W_{r\phi}\omega R^2). \quad (2.7)$$

During radial motion half the liberated potential energy goes into increasing the kinetic energy, and the other half (corresponding to the first term on the right-hand side of (2.7)) should go into heat and mechanical energy. As well as transferring angular momentum, viscous stresses transfer mechanical energy. This process explains the second term on the right of (2.7). We are assuming that energy dissipated into heat is radiated on the spot, and not transported radially.

(2.7) can be derived from the momentum equation (2.5) by multiplying it by ω. Similarly, multiplying (2.4) by ω gives (2.6).

(d) *Hydrostatic equilibrium along the z-direction*

If motions in the disk along the z-direction are subsonic, then the disk is in hydrostatic equilibrium (see Appendix). In the direction perpendicular to the plane of the disk the gas and radiation pressure gradient is balanced by the component of the gravitational attraction of the central body normal to the disk (the selfgravitation of the disk being negligibly small):

$$\frac{\partial p}{\partial z} = -\rho \frac{GM}{R^3} z = -\rho\omega^2 z. \tag{2.8}$$

Here and elsewhere the disk is assumed to be thin, $H \ll R$.

The disk has a definite structure along the z-direction, but from now on we shall only be using physical quantities averaged over z. While averaging we shall assume that the volume density ρ at any given radius does not vary with z, i.e. the disk is homogeneous. (In SS it was shown that the disk was homogeneous when radiation dominated the pressure.) Integrating (2.8), we find

$$p(z) = p_c \left[1 - \left(\frac{z}{H}\right)^2\right] \tag{2.9}$$

where $p_c = \frac{1}{2}\rho\omega^2 H^2$ is the central pressure of the disk. The average pressure p is then

$$p = \frac{1}{H}\int_0^H p(z)\, dz = \tfrac{2}{3}p_c = \tfrac{1}{3}\rho\omega^2 H^2 = \frac{U\omega^2 H}{6}. \tag{2.10}$$

We introduce an averaged sound speed

$$v_s = \frac{\omega H}{\sqrt{3}} \tag{2.11}$$

using the relation $p = \rho v_s^2$. For a thin disk $H/R \ll 1$ this can only be true if $v_s \ll \omega R = v_\phi(R)$, i.e. pressure forces in the disk $(1/\rho)\,(\partial p/\partial R) \sim (GM/R^2)\,(H/R)^2$ are smaller by a factor of about $(H/R)^2$ than the gravitational and centrifugal forces.

(e) *Viscous stresses*

In general the viscous stress $W_{r\phi}$ is determined by the dynamical viscosity η:

$$W_{r\phi} = -2\eta HR \frac{\partial \omega}{\partial R}, \tag{2.12}$$

which for the Keplerian law $\omega = \sqrt{GM/R^3}$ gives

$$W_{r\phi} = 3\eta H\omega.$$

In SS it was shown that neither molecular nor radiant viscosity can play an important role in accretion disks; momentum is transferred by turbulence and by small-scale chaotic magnetic fields. In this case

$$\eta = \tfrac{1}{3}\rho v_t l_t \tag{2.13}$$

where v_t is the turbulent or Alfvèn velocity and l_t is the turbulent mixing length or the length scale of the magnetic field (for more detail see Lynden-Bell (1969) and SS). It is normally assumed that $v_t \lesssim v_s$ and $l_t \lesssim H$.

It is convenient to introduce the parameter

$$\alpha = \frac{v_t l_t}{v_s H} + \frac{B^2}{4\pi\rho v_s^2} \tag{2.14}$$

which describes the excitation level of the turbulence and the importance of the chaotic magnetic fields. Then

$$\eta = \tfrac{1}{3}\alpha\rho v_s H = \frac{\alpha}{3\sqrt{3}}\rho\omega H^2 = \frac{\alpha}{6\sqrt{3}}U\omega H \tag{2.15}$$

and for viscous stresses with a Keplerian rotation law we have

$$W_{r\phi} = 3\eta H\omega = \frac{\sqrt{3}}{6}\alpha U\omega^2 H^2 = \sqrt{3}\alpha p H. \tag{2.16}$$

We remark that in SS and in later work by others viscous stresses were given by the formula $W_{r\phi} = 2\alpha p H = \alpha U v_s^2$. These equations are identical apart from the numerical factor $\sqrt{3/2}$, which for the accuracy we require can be assumed to be unity. From now on we shall use $W_{r\phi}=2\alpha p H$. (Comparison with Sunyaev and Shakura (1975), where we used $W_{r\phi} = \sqrt{3}\alpha p H$, shows that the results obtained are practically unchanged.)

3. PROPERTIES OF STATIONARY ACCRETION

For stationary accretion $\partial/\partial t = 0$ and the dynamic equations integrate easily, resulting in a system of algebraic equations. From the continuity equation (2.3) it follows that the accretion rate $\dot{M} = -2\pi U v_r R =$ const. does not depend on radius, i.e. it is one of the more important quantities characterizing the accretion.

Integrating the momentum equation, we find

$$-\dot{M}\omega R^2 + 2\pi W_{r\phi} R^2 = C_1 \tag{3.1}$$

and using $W_{r\phi} = 2\alpha p H$ we obtain

$$-\dot{M}\omega R^2 + \frac{2\pi}{3}\alpha U\omega^2 H^2 R^2 = C_1.$$

(a) *Boundary conditions*

With decreasing radius the angular velocity of matter in the accretion disk increases, while the angular momentum decreases. Surplus momentum is removed by viscous stresses. Keplerian orbits outside a Schwarzschild black hole exist only for $R > 3R_g$, while for $R < 3R_g$ particles make many revolutions along an unwinding spiral and fall into the black hole. For $R > 3R_g$ radial motion is associated only with the outward loss of angular momentum. For $R < 3R_g$ it could happen purely as a consequence of general relativity, and with constant angular momentum. Outward transfer of angular momentum as a result of viscous stresses would increase the radial velocity.

The equations of motion of test particles with given angular momentum and energy in the field of a Schwarzschild black hole are well known (Zel'dovich & Novikov 1971). Using $\sqrt{3}R_g c$ for the angular momentum and $\sqrt{\tfrac{8}{9}}c^2 = (1-0{\cdot}057)c^2$ for the total energy, which are appropriate for the last stable Keplerian orbit $R = 3R_g$, it is easy to estimate the number of revolutions $n = 9/(\pi\Delta^{1/2})$ a test particle would make in travelling from $R = 3R_g(1-\Delta)$ to R_g, where $\Delta \ll 1$. The smaller the momentum and energy of the particle, the smaller the number of revolutions it makes. Thus the removal of momentum and energy by viscous stresses speeds up the radial motion of the particles. It leads to a rapid break-up of the disk structure. The zone of the disk with $R < 3R_g$ is shown to be narrow. At the boundary of this zone viscous stresses cease to affect noticeably the particle trajectories, and so we can put $W_{r\phi} = 0$. As this zone is narrow, energy transfer from $R < 3R_g$ can be ignored and for simplicity we can put $W_{r\phi} = 0$ at $R = 3R_g$. Substitution of this boundary condition into (3.1) gives us

$$W_{r\phi} = \dot{M}_0 \frac{\omega}{2\pi} \left[1 - \left(\frac{R_0}{R}\right)^{1/2}\right] \quad (3.2)$$

where R_0 is the radius of the inner edge of the disk, equal to the radius of the last stable circular orbit around the black hole, the radius of the star or the effective radius of its magnetosphere. Formal use of the condition $W_{r\phi}(R_0) = 0$ leads to several physical parameters of the disk tending to infinity. The above discussion shows that those infinities are fictional. Thus the inner boundary condition determines the constant in the momentum equation. The outer condition gives the accretion rate $\dot{M}$.

(b) *Thermal balance in the disk*

Substituting expression (3.2) for $W_{r\phi}$ into (2.7), we find the energy emission rate for the disk

$$Q^+ = \frac{3}{8\pi} \dot{M}\omega^2 \left[1 - \left(\frac{R_0}{R}\right)^{1/2}\right]. \quad (3.3)$$

Thus, as Lynden-Bell first remarked, for stationary accretion far away from the inner boundary the radiation flux is three times greater than the difference between the fluxes of gravitational and kinetic energy. This increase results from work done by frictional forces.

Near R_0, however, the dissipation rate falls rapidly and $Q(R_0) = 0$. The total luminosity of the disk

$$L = 4\pi \int_{R_0}^{\infty} QR\, dR$$

is found to be

$$L = \dot{M}_0 (\tfrac{1}{2}) \frac{GM}{R_0} = \zeta \dot{M}_0 c^2,$$

i.e. the total luminosity is equal to the accretion rate $\dot{M}_0$ multiplied by the binding energy for the last stable circular orbit. It is a remarkable property of stationary disk accretion that the total luminosity is independent of the form of the dissipative forces. However, the quantity $W_{r\phi}$ does determine the surface density, internal temperature, optical thickness τ and consequently the spectral distribution of the emitted radiation.

Using $W_{r\phi} = 2\alpha p H$ we obtain yet another useful expression:

$$Q^+ = -\tfrac{1}{2} W_{r\phi} R \frac{\partial \omega}{\partial R} = \eta H R^2 \left(\frac{\partial \omega}{\partial R}\right)^2 = \tfrac{9}{4} \eta H \omega^2 = \frac{\alpha U \omega^3 H^2}{6}. \quad (3.4)$$

From (3.3) and (3.4) it follows in particular that

$$\eta H = \frac{\dot{M}}{6\pi} \left[1 - \left(\frac{R_0}{R}\right)^{1/2}\right]. \quad (3.5)$$

Radiation carries away the thermal energy. If the disk is optically thick, because of electron scattering or true absorption, the radiation flux Q^- is connected with the radiation density ϵ_r at the centre of the disk by the relation

$$Q^- = \frac{4}{3}\frac{\epsilon_r c}{\tau} = \frac{4}{3}\frac{m_p c}{\sigma}\frac{\epsilon_r}{U} \quad (3.6)$$

where σ is the effective cross-section.

For stationary accretion equality of Q^+ and Q^- gives an additional relation for determining the averaged physical quantities in the disk.

(c) *Fundamental equations of stationary disk accretion*

Equations (2.1–2.16, 3.1–3.6) form a system of algebraic equations:

Kepler's law $$\omega = \left(\frac{GM}{R^3}\right)^{1/2}$$

Continuity equation $$\dot{M} = -2\pi U v_r R$$

Momentum equation $$\dot{M}\left[1-\left(\frac{R_0}{R}\right)^{1/2}\right] = 2\pi W_{r\phi}$$

Hydrostatic equation along the z-direction $$p = \frac{U\omega^2 H}{6}$$

Expression for viscous stresses $$W_{r\phi} = 2\alpha p H$$

Energy production by friction $$Q^+ = -\tfrac{1}{2} W_{r\phi} R \frac{\partial \omega}{\partial R}$$

Removal of energy by radiation $$Q^- = \frac{4}{3}\frac{c m_p}{\sigma}\frac{\epsilon_r}{U}$$

Equation of state $$p = \frac{Uk(T_e+T_i)}{2Hm_p} + \frac{2}{9}\epsilon_r$$

(Here we take into account that the mean pressure is less than the central pressure by a factor of one and a half, $p = 2/3 p_c$.)

Opacity law $$\sigma = \sigma_r + \sigma_{ff} = \sigma_r + \frac{AU}{HT_e^{7/2}}$$

(using Thompson scattering and free–free absorption).

Relation between electron temperature and radiation energy density (in the simplest case $\epsilon_r = aT_e^{\gamma}$).

Relation between electron and proton temperature. Here we shall assume $T_e = T_i$. (However, see Eardley, Lightman & Shapiro (1975) where the case $T_e \neq T_i$ is considered.)

Solving this system, it is easy to find the radial dependence of the physical quantities, having set the accretion rate $\dot{M}$, the mass of the object M, the parameter α and the accretion efficiency ζ (or the inner radius of the disk $R_0 = 3R_g$).

The disk consists of three zones:

(A) The inner zone; $p_r \gg p_g$, $\sigma_T \gg \sigma_{ff}$,

(B) The intermediate zone; $p_g \gg p_r$, $\sigma_T \gg \sigma_{ff}$,

(C) The outer zone; $p_g \gg p_r$, $\sigma_{ff} \gg \sigma_T$.

The parameters of the disk in these zones were derived in SS. Zone A is of most interest to us. In particular, it generates most of the energy when accretion is nearly critical. Below we obtain the disk's parameters in this zone.

(d) *Dimensionless variables*

It is natural to use the following dimensionless variables: mass of the object $m = M/M_\odot$, radius $r = R/R_0$ and accretion rate $\dot{m} = \dot{M}/\dot{M}_c$. The luminosity of the accretion disk is related to the accretion rate by $L = \zeta \dot{M} c^2$, where the dimensionless factor ζ is the binding energy at the inner edge of the disk R_0. In Newtonian theory $\zeta = \frac{1}{2}(GM/R_0c^2) = \frac{1}{4}(R_g/R_0)$ where $R_g = 2GM/c^2 \simeq 3m$ km is the gravitational radius of the centre of attraction. For relativistic objects this formula is an approximation. For the Schwarzschild metric the radius of the last stable circular orbit is $R_0 = 3R_g$ and the approximation gives $\zeta = 0{\cdot}083$ instead of the precise value $\zeta = 0{\cdot}057$. For the limiting Kerr metric, when the moments of rotation of the black hole and the disk are aligned, $R_0 = \frac{1}{2}R_g$ and $\zeta = 0{\cdot}5$ instead of the precise value $\zeta = 0{\cdot}42$.

The accuracy of the approximate formula is sufficient for our calculations and below we shall sometimes use it. It is convenient to introduce the parameters $l = R_0/3R_g$ and $s = (1 - r^{-1/2})$. Then $\zeta = 1/12l$ and $\omega = c/\sqrt{6}R_0 l^{1/2} r^{3/2}$.

The critical accretion rate

$$\dot{M}_{cr} = \frac{L_c}{\zeta c^2} = \frac{2\pi m_p c R_g}{\zeta \sigma_T} = \frac{8\pi m_p c R_0}{\sigma_T} \simeq 3{\cdot}8 \times 10^{-8}\, m \quad M_\odot\,\mathrm{yr}^{-1}$$

is intimately connected with the limiting Eddington luminosity

$$L_c = \frac{4\pi m_p c G M}{\sigma_T} = \frac{2\pi m_p c^2 R_g c}{\sigma_T} \simeq 1{\cdot}3 \times 10^{38}\, m \quad \mathrm{erg\,s^{-1}}$$

at which radiation pressure on the electrons $f_r = L\sigma_T/4\pi R^2 c$ in a fully ionized medium is equal to the attractive force $f_g = GMm_p/R^2$ on the protons towards the gravitational centre. In dimensionless form many of the expressions become simple.

(e) *Stationary solution in the inner zone*

Solving the system of equations (3, c), we find

	Extremal position	Extremal value
$H/R = 3r^{-1}\dot{m}s$	$r_{max} = 9/4$	$4\dot{m}/9$
$U = 4\sqrt{\frac{2}{3}}\frac{m_p}{\sigma_T}\frac{r^{3/2}l^{1/2}}{\alpha \dot{m} s}$	$r_{min} = 16/9$	$40\frac{l^{1/2}}{\alpha\dot{m}}\,\mathrm{g\,cm^{-2}}$
$\tau_T = \frac{\sigma_T U}{2m_p}$	$r_{min} = 16/9$	$\frac{8l^{1/2}}{(\alpha\dot{m})}$
$n_e = \left(\frac{2}{3}\right)^{3/2}\frac{1}{\sigma_T R_0}\frac{r^{3/2}l^{1/2}}{\alpha\dot{m}^2 s^2}$	$r_{min} = 25/9$	$\frac{15 l^{1/2}}{\sigma_T R_0 \alpha \dot{m}^2} \simeq 2{\cdot}5 \times 10^{19}\frac{l^{1/2}}{\alpha\dot{m}^2 m}$
$\epsilon_r = \sqrt{\frac{3}{2}}\frac{m_p c^2}{\sigma_T R_0}\frac{r^{-3/2}}{\alpha l^{1/2}}$	—	—
$v_r/v_\phi = 3\alpha\dot{m}sr^{-2}$	$r_{max} = 25/16$	$0{\cdot}25\,\alpha\dot{m}$
$Q^+ = \frac{m_p c^3}{2\sigma_T R_0} r^{-3} s\dot{m}$	$r_{max} = 49/36$	—

Note that whenever α, $\dot{m} < 1$ and $r > 1$ the conditions $H \ll R$, $v_r \ll v_\phi$ which were used in deriving the equations are satisfied. The transfer of mechanical energy within the disk strongly influences its structure. We remarked previously (SS) that the greatest energy generation took place not at the inner edge, but at $r = 49/36$. All other fundamental disk parameters have similar extremes (maxima or minima). This is why we have included here the extremal radii and values as well as the exact expressions themselves.

It is interesting that all the fundamental disk parameters are simply expressed in terms of the universal constants m_p, σ_T and c and also R_0 and the dimensionless quantities r, s, α, $\dot{m}$, l. Assuming that local thermodynamic equilibrium holds in the disk at the boundary between zones A and B, $T_e = T_r = (\epsilon_r/a)^{1/4}$ and it is easy to locate the boundary by equating the plasma and radiation pressures. Equating

$$p_r(r_{AB}) = \epsilon_r/3 \quad \text{and} \quad p_g(r_{AB}) = 2nkT$$

we have

$$\frac{r_{AB}}{(1 - r_{AB}^{-1/2})^{16/21}} \simeq 1{\cdot}2 \times 10^2\,(m\alpha)^{2/21}\, l^{-5/21} \dot{m}^{16/21}.$$

From this it follows that a region where radiation pressure dominates exists only if

$$\dot{m} > \dot{m}_{AB} \simeq \tfrac{1}{50}\,(\alpha m)^{-1/8}\, l^{5/16}.$$

We point out the weak dependence of r_{AB} on α.

(f) *Value of the viscosity*

In models of stationary disk accretion it is assumed that the energy production rate $Q^+ = \eta \frac{9}{4} H w^2$ is equal to the rate of energy loss from the disk surface Q^-. Using the formulas for stationary accretion it is easy to express η in each region in terms of the variable r and the parameters $\dot{m}$, m and α. The equality $Q^+ = Q^-$ enables $\eta = \eta(r, m, \alpha, \dot{m})$ to be found.

The most interesting situation is in the inner zone, where $p_r \gg p_g$. In this zone

$$\epsilon_r = 3p_c = \tfrac{9}{2}p = \tfrac{3}{4}U\omega^2 H,$$

$$Q^- = \frac{4}{3}\cdot\frac{cm_p}{\sigma_T}\cdot\frac{\epsilon_r}{U} = \frac{cm_p}{\sigma_T}\,\omega^2 H.$$

Equating with $Q^+ = \frac{9}{4}\eta\omega^2 H$ we find that

$$\eta = \frac{4}{9}\frac{cm_p}{\sigma_T} \simeq 3{\cdot}5 \times 10^{10} \quad \text{erg s cm}^{-3}.$$

Thus the *value of the viscosity* at which stationary accretion is possible in the interior zone of the disk *is independent of all parameters* of the accretion and is expressed as a product of universal constants. We point out that no assumptions concerning the viscosity went into deriving its equality to this constant. The function $\eta = \eta(r, m, \alpha, \dot{m})$ in zone B of the disk is

$$\eta \simeq 10^{13}(\alpha m)^{1/10}(\dot{m}s)^{4/5}\, r^{-21/20} \quad \text{erg s cm}^{-3}.$$

Analogously the value of η in zone C of the disk is

$$\eta \simeq 2 \times 10^{13}(\alpha m)^{1/10}(\dot{m}s)^{17/20}\, r^{-9/8} \quad \text{erg s cm}^{-3}.$$

For comparison the viscosity of a fully-ionized gas in zone A is

$$\eta_i \simeq 4\times 10^5(\alpha m)^{-1/2}(\dot{m}s)\, r^{-9/4} \quad \mathrm{erg\ s\ cm^{-3}},$$

which is smaller by 5 orders of magnitude than the value of the turbulent or magnetic viscosity necessary for stationary accretion.

The radiant viscosity in zone A is also of interest:

$$\eta_r \simeq \frac{4}{15}\frac{\epsilon_r}{\sigma_T n_e c} = \frac{3}{5}\frac{m_p c}{\sigma_T}\frac{\dot{m}^2 s^2}{r^3 l}.$$

The maximum value of the radiant viscosity is $10^{-2}(m_p e/\sigma_T)(\dot{m}^2/l)$ at $r = 16/9$. Equating with the viscosity $\eta = \frac{4}{9}(m_p c/\sigma_T)$ necessary for the maintenance of stationary disk accretion in the inner zone, we see that for all acceptable values of the paramcter $l \geqslant \frac{1}{6}$ the role of radiant viscosity can be ignored. This was shown previously in SS.

4. STABILITY THEORY FOR STATIONARY DISK ACCRETION

We shall impose on the solution for stationary disk accretion perturbations of wavelength Λ, satisfying $H \ll \Lambda \ll R$. We shall confine ourselves to axially symmetric $(\partial/\partial\phi = 0)$ perturbations independent of z, $(\partial/\partial z = 0)$. The latter means that at any given radius the disk expands or contracts in the z-direction preserving the uniform distribution of its density. The perturbing motions are assumed to be considerably subsonic. As is shown in the Appendix, ignoring terms of order $(H/R)^2$ and $H^2/R\Lambda$, the disk satisfies the condition of hydrostatic equilibrium:

$$p(R,t) = \tfrac{1}{6}\omega^2(R).U(R,t).H(R,t). \qquad (4.1)$$

(a) *Dynamic equation*

For perturbations of the above form one can show (see Appendix) that the Keplerian rotation law is fulfilled with considerable accuracy:

$$v_\phi^2 = \omega^2 R^2 = \frac{GM}{R}\left[1 + O\left(\frac{H^2}{R\Lambda}\right) + O\left(\frac{H^2}{R^2}\right)\right]. \qquad (4.2)$$

This means that the momentum equation (2.5) holds even in the non-stationary problem. Combining it with the continuity equation and substituting into it expression (2.16) for viscous stress, we get

$$\frac{\partial U}{\partial t} = \frac{\alpha}{2R}\frac{\partial}{\partial R}\frac{1}{\omega R}\frac{\partial}{\partial R}U\omega^2 H^2. \qquad (4.3)$$

We linearize equation (4.3), substituting into it a perturbation solution of the form

$$U = U_0(R)[1+u]$$
$$H = H_0(R)[1+h]$$

where $U_0(R)$ and $H_0(R)$ are the stationary solutions and $u \ll 1$, $h \ll 1$. For perturbations of wavelength $H_0 \ll \Lambda \ll R$ we shall only keep terms of order $(H_0/\Lambda)^2$, ignoring terms of order $(H_0/R)^2$ and $H_0^2/R\Lambda$. After linearization (4.3) becomes

$$\frac{\partial u}{\partial t} = \frac{\alpha\omega H_0^2}{2}\frac{\partial^2}{\partial R^2}(u+2h). \qquad (4.4)$$

Of course the single equation (4.4) relating the functions u and h is insufficient for investigating the stability of stationary disk accretion. A supplementary equation relating u and h is required.

In the well-known papers by Lightman & Eardley (1974) and Lightman (1974), where the question of the stability of disk accretion on to black holes of stellar mass was first posed, use was made of a relation between u and h derived from the condition $Q^+ = Q^-$. In the inner zone of the disk this condition is equivalent to assuming the equality of the viscosity to the fixed value $\eta = \frac{4}{9}(m_p c/\sigma_T)$ which is not generally the case.

(b) *Thermal equation and its linearization*

We will obtain the second equation for u and h from the law of conservation of energy

$$\frac{dE}{dt} = \frac{p}{\rho^2}\frac{d\rho}{dt} + q^+ - q^-. \tag{4.5}$$

Here the first term on the right describes the change in internal energy $E = \epsilon/\rho$ of a gram of matter as a result of work done by pressure forces, while the second and third terms are respectively frictional heating and radiative cooling. Using $d/dt = \partial/\partial t + v_r(\partial/\partial r) + v_z(\partial/\partial z)$ we integrate (4.5) with respect to z (having first multiplied by $\rho\, dz$), assuming uniformity of the expansion or contraction of the disk along the z-axis:

$$\frac{\partial}{\partial t}\epsilon H + p\frac{\partial H}{\partial t} = -\frac{1}{R}\frac{\partial}{\partial R}(\epsilon + p)\, v_r H R + v_r \frac{\partial}{\partial R} pH + Q^+ - Q^-. \tag{4.6}$$

Expressions (3.4) for Q^+ and (3.6) for Q^- are of course correct for non-stationary accretion.

The total thermal energy density is connected with the pressure by the relation $\epsilon = \frac{3}{2}(1+\beta)p$ where $\beta = p_r/(p_r + p_g)$ describes the contribution of radiation to the total pressure. Using (4.1) we get

$$\epsilon = \tfrac{1}{4}(1+\beta)\, UH\omega^2. \tag{4.7}$$

Using the relations derived above we rewrite the thermal equation (4.6) so that the only unknown quantities are U and H:

$$\frac{1}{4}\frac{\partial}{\partial t}(1+\beta)\, UH^2\omega^2 + \frac{1}{6}UH\omega^2\frac{\partial H}{\partial t} = \frac{2\alpha}{3R}\frac{\partial}{\partial R}\frac{5+3\beta}{12}H^2\omega\frac{\partial}{\partial R}UH^2\omega^2 + \frac{v_r}{6}\frac{\partial}{\partial R}UH^2\omega^2 + \frac{\alpha}{4}UH^2\omega^2 - \frac{cm_p}{\sigma_T}\beta H\omega^2. \tag{4.8}$$

We linearize (4.8), keeping only terms of order $(H_0/\Lambda)^2$. Note that in the general case one has to include variations in β. We have

$$p = p_g + p_r = \frac{U}{2H}\frac{k}{m_p}\left(\frac{\epsilon_r}{a}\right)^{1/4} + \frac{\epsilon_r}{3} = \frac{U}{2H}\frac{k}{m_p}\left(\frac{3\beta P}{a}\right)^{1/4} + \beta p. \tag{4.9}$$

Linearizing equation (4.1) gives

$$\frac{p_1}{p_0} = \pi_1 = u + h \tag{4.10}$$

and after linearizing (4.9) using (4.10) we have

$$\frac{\beta_1}{\beta_0} = \frac{1-\beta_0}{1+3\beta_0}(3\pi_1 - 4u + 4h) = \frac{1-\beta_0}{1+3\beta_0}(7h-u) \quad (4.11)$$

where β_0 is the proportion of radiation pressure in the unperturbed flow. Finally, linearizing (4.8), we obtain

$$(8+51\beta_0-3\beta_0^2)\frac{\partial h}{\partial t} + 3(4\beta_0^2+3\beta_0+1)\frac{\partial u}{\partial t} = \tfrac{2}{3}(9\beta_0^2+18\beta_0+5)\,\alpha\omega H_0^2\frac{\partial^2}{\partial R^2}(u+2h)$$
$$+6\alpha\omega[(1+\beta_0)\,u+(5\beta_0-3)\,h]. \quad (4.12)$$

It should be remarked that before linearization the thermal equation was of order $(H_0/R)^2$ relative to the equation of motion (4.4). However, the cancellation of $U_0 w^2 H_0^2$ during linearization equalized the orders of the equations. When linearizing (4.8) we ignored the second term on the right, which is smaller than the other terms by $H_0^2/R\Lambda$.

Thus (4.4) and (4.12) are the two fundamental equations of non-stationary disk accretion. The independent variables are $u(R, t)$ and $h(R, t)$. Through them are expressed all the fundamental properties of non-stationary accretion: in particular $\dot{M}(r, t)$. We shall look for solutions to (4.4) and (4.12) of the form $u = \exp(\Omega t)\,u(R)$, $h = \exp(\Omega t)\,h(R)$. It is convenient to choose as an unknown function the quantity $\psi = u+2h = \exp(\Omega t)\,\psi(R)$, which represents the perturbation of the viscous forces between neighbouring layers. We then obtain for the radial part $\psi(R)$ the equation

$$\Omega\,\frac{A(\beta_0)\,\Omega - 6\alpha\omega(5\beta_0-3)}{B(\beta_0)\,\Omega + 6\alpha\omega(5-3\beta_0)}\,\psi = \frac{2\alpha\omega}{3}H_0^2\frac{\partial^2\psi}{\partial R^2} \quad (4.13)$$

where $A(\beta_0) = 8+51\beta_0-3\beta_0^2$, $B(\beta_0) = 3(4+23\beta_0-3\beta_0^2)$.

(c) *Perturbation types and conditions for their growth*

When investigating the propagation of waves of length $\Lambda \ll R$ we shall ignore radial changes in β_0 and ω. Substituting into (4.13) a solution of the form $\psi = \sin(R/\Lambda)$ we find a dispersion relation which can be conveniently written in the form

$$A(\beta_0)\left(\frac{\Omega}{6\alpha\omega}\right)^2 + \left[B(\beta_0)\left(\frac{H_0}{3\Lambda}\right)^2 + (3-5\beta_0)\right]\frac{\Omega}{6\alpha\omega} + (5-3\beta_0)\left(\frac{H}{3\Lambda}\right)^2 = 0. \quad (4.14)$$

Solution curves for this equation for $\Omega > 0$ are plotted on Fig. 1 for various values of the parameter $\beta_0 > \frac{3}{5}$. In the short-wave region the perturbations take the form of concentric waves running along the disk, as (4.14) then gives complex values for the growth-rate Ω. It is interesting that the smallest possible wavelength for which the real part of the growth-rate is positive is always greater than the disk-thickness $2H_0$. This result is not obtained if one only uses equations (4.4) and the condition $Q^+ = Q^-$. (In the theory of Lightman & Eardley Ω grows without bound as Λ is decreased.) Beginning at some value Λ^* depending on β_0, the equation has two positive roots $\Omega > 0$. These solutions correspond to growing modes of standing waves and have different physical interpretations. For the lower branch in the long wavelength limit the growth-rate diminishes: $\Omega \simeq \frac{2}{3}[(5-3\beta_0)/(5\beta_0-3)]\,(H_0/\Lambda)^2\alpha\omega$ and asymptotically this solution tends to the instability discovered by Lightman & Eardley (1974).

Equation (4.12) connects the perturbations u and h:

$$(8+51\beta_0-3\beta_0^2)\left(\frac{\Omega}{6\alpha\omega}\right)h+(4\beta_0^2+3\beta_0+1)\left(\frac{\Omega}{6\alpha\omega}\right)u$$
$$= -(9\beta_0^2+18\beta_0+5)\left(\frac{H_0}{3\Lambda}\right)^2(u+2h)+[(1+\beta_0)\,u+(5\beta_0-3)\,h]. \quad (4.15)$$

Using (4.14) and (4.15) it can be shown that for the lower branch at long wavelengths $u = -h$, i.e. the perturbed viscosity $\eta_1/\eta_0 = u+h \simeq 0$. This means that only in the limiting case of long wavelengths do perturbations grow such that the energy-generation equals the energy-loss: $Q^+ = Q^-$.

There is another instability branch, whose growth rate grows with increasing wavelength, tending to $\Omega \sim 0{\cdot}2\alpha\omega$ when $\beta_0 \simeq 1$. From (4.14) and (4.15) we find that on this branch, in the limit, the perturbation of the surface density becomes small compared to the perturbation of the viscous forces, disk thickness and other quantities: $u \to 0$ as $\Lambda \to W$. This branch is caused by a thermal instability of disk accretion in the zone where radiation pressure dominates.

The perturbations of all physical quantities are uniquely determined by the perturbations u and h. (For example, $Q_1^+/Q_0^+ = u+2h \equiv \psi$, $Q_1^-/Q_0^- = h$, etc.

From the momentum equation (2.5) we have

$$\dot{M} = \frac{4\pi}{\omega R}\frac{\partial}{\partial R}W_{r\phi}R^2 = \frac{4\pi\alpha}{3\omega R}\frac{\partial}{\partial R}UH^2\omega^2R^2 \quad (4.16)$$

and for the relative perturbation of the accretion rate

$$\frac{\dot{M}_1(R,t)}{\dot{M}_0} = (u+2h)+2sR\frac{\partial}{\partial R}(u+2h) \equiv \psi+2sR\frac{\partial\psi}{\partial R}. \quad (4.17)$$

For $H<\Lambda<R$ the second term on the right of (4.17) is greater than the first. This means that the perturbations $\dot{M}_1/\dot{M}_0$ grow at the same rate Ω as ψ, but with an amplitude greater by a factor Rs/Λ:

$$\frac{\dot{M}_1}{\dot{M}_0} \simeq \left[1+\frac{2sR}{\Lambda}\right](u+2h). \quad (4.18)$$

The big change in $\dot{M}_1/\dot{M}_0$ is not associated with increased energy production, but with increased v_r and transfer of mechanical energy along the disk. With decreasing β_0 the value of Ω_{max} falls and that of Ω_{min} grows, so that as $\beta_0 \to \frac{3}{5}$, $\Omega_{max} \to 0$. This means that the external regions, where gas pressure dominates, are stable.

The analysis that has been performed does not use boundary conditions and is valid when a large number of waves can be fitted into distances of order R, i.e. $\Lambda/R \ll 1$. It is qualitatively clear that perturbations with $\Lambda \sim R$ cannot grow because of the stability of the solution for the outer zone, where the matter flows from. The constancy of $\dot{M}_0$ prevents the growth of large-scale perturbations, which would destroy the disk structure of the accretion.

It has to be pointed out that the numerical estimates of the growth rates and characteristic wavelengths of the perturbations depend on the method of averaging along the z-direction of the equations of motion, and in particular of the thermal equation. But as the analysis has shown, the qualitative picture does not change, and the boundary between the stable and unstable zones is always found at $\beta_0 = \frac{3}{5}$.

5. ACKNOWLEDGMENTS

The authors are grateful to Ya. B. Zel'dovich for valuable discussions, to R. L. Znajek for an excellent translation and to M. S. Longair for his assistance.

REFERENCES

Cunningham, C. T., 1975. Preprint.
Eardley, D. M., Lightman, A. P. & Shapiro, S. L., 1975. *Astrophys. J.*, **199**, L155.
Lightman, A. P., 1974. *Astrophys. J.*, **194**, 419 and 429.
Lightman, A. P. & Eardley, D. M., 1974. *Astrophys. J.*, **187**, L1.
Lynden-Bell, D., 1969. *Nature*, **223**, 690.
Lynden-Bell, D. & Rees, M. J., 1971. *Mon. Not. R. astr. Soc.*, **152**, 461.
Lynden-Bell, D. & Pringle, J. E., 1974. *Mon. Not. R. astr. Soc.*, **168**, 603.
Lyutiy, V. M. & Pronik, V. I., 1975. *Variable stars and stellar evolution*, p. 591, eds Sherwood, V. E. & Plaut, L. (IAU Symposium 67), Reidel, Dordrecht.
Novikov, I. D. & Thorne, K. S., 1973. *Black holes*, eds C. DeWitt & B. DeWitt, Gordon and Breach, New York.
Page, D. N. & Thorne, K. S., 1974. *Astrophys. J.*, **191**, 499.
Polnarev, A. G., 1975. Thesis.
Pringle, J. E. & Rees, M. J., 1973. *Astr. Astrophys.*, **21**, 1.
Pringle, J. E., Rees, M. J. & Pacholczyk, A. G., 1973. *Astr. Astrophys.*, **29**, 179.
Shakura, N. I., 1973. *Sov. Astr.*, **16**, 756.
Shakura, N. I. & Sunyaev, R. A., 1973. *Astr. Astrophys.*, **24**, 337.
Shibasaki, N. & Hoshi, R., 1975. Preprint.
Sunyaev, R. A., 1973. *Sov. Astr.*, **16**, 941.
Sunyaev, R. A. & Shakura, N. I., 1975. *Sov. Astr. Letters* (in translation), **1**, No. 8, 2.
Zel'dovich, Y. B. & Novikov, I. D., 1971. *Relativistic astrophysics*, University of Chicago Press.

APPENDIX

We write Euler's equations for axially-symmetric motion in cylindrical polar coordinates:

$$\frac{\partial v_r}{\partial t}+v_r\frac{\partial v_r}{\partial R}-\frac{v_\phi^2}{R} = -\frac{GM}{R^2}-\frac{1}{\rho}\frac{\partial p}{\partial R} \quad (A.1)$$

$$\frac{\partial v_\phi}{\partial t}+v_r\frac{\partial v_\phi}{\partial R}+\frac{v_r v_\phi}{R} = -\frac{1}{\rho R}\frac{\partial}{\partial R}(W_{r\phi}R^2) \quad (A.2)$$

$$\frac{\partial v_z}{\partial t}+v_r\frac{\partial v_z}{\partial R} = -\frac{GM}{R^3}z-\frac{1}{\rho}\frac{\partial p}{\partial z}. \quad (A.3)$$

All components of the viscous stress tensor, except $W_{r\phi}$, are small and are not included in (A.1–3). We shall show that for perturbations with wavelength $H<\Lambda<R$, the perturbation of the toroidal velocity component $v_{\phi 1}/v_{\phi 1}$ is small compared with the perturbations of other quantities. Using for the stationary solution the approximate equations $v_r \simeq \alpha v_{\phi 0}(H/R)^2$ and $v_{s0} \simeq v_{\phi 0}(H/R)$ we obtain from (A.1) two branches with growth rates Ω; $v_{\phi 1}/v_{\phi 0} \simeq \alpha^2(H_0/R)^2\, v_{r1}/v_{r0}$ for one, and $v_{\phi 1}/v_{\phi 0} \simeq \alpha^2(H_0/R)^4\, v_{r1}/v_{r0}$ for the other. The smallness of the changes in angular velocity allows the possibility of ignoring the first term in equation (A.2): $\partial v_\phi/\partial t = 0$.

In the main text of the paper we used only the right-hand side of (A.3)—the condition of hydrostatic equilibrium in the z-direction. Direct substitution of the perturbed solution into (A.3) makes it easy to verify that the terms on the left are smaller by a factor $\alpha^2(H_0/R)^2$ than those on the right.