BLACK HOLE BINARIES AND X–RAY TRANSIENTS

A. R. King, U. Kolb [1] and L. Burderi[2]

Astronomy Group, University of Leicester, Leicester LE1 7RH, U.K.

ABSTRACT

We consider transient behaviour in low–mass X–ray binaries (LMXBs). We show that if this results from a disc instability, the secondary star must be significantly evolved when mass transfer starts, particularly if the primary is a neutron star. For $P \lesssim 2$d most neutron–star systems will be persistent X–ray sources, whereas the slower orbital evolution of black–hole systems means that most of them are likely to be transient. Both types of transient system must have extreme mass ratios (< 0.1). For orbital periods $P \gtrsim 2$d most LMXBs will be transient regardless of whether the primary is a neutron star or a black hole.

Subject headings: Subject headings: accretion, accretion discs — binaries: close — instabilities — black hole physics

[1] also: Max–Planck–Institut für Astrophysik, Karl–Schwarzschild–Str. 1, D 85748 Garching, Germany

[2] also: Istituto di Fisica, Università di Palermo, Palermo 90100, Italy

1. INTRODUCTION

In recent years several soft X–ray transients (SXTs) have been found to have large mass functions, strongly suggesting that the accreting star is a black hole. The incidence of these black–hole systems among SXTs with known orbital periods (8 out of 14) is much greater than any likely estimate of their incidence among persistent LMXBs (1 out of 29). Moreover, estimates of the mass of the secondary star in the black–hole binaries consistently show that this star must be larger than its main–sequence radius if it is to fill its Roche lobe. In this Letter we shall show that in the context of the disc–instability model for SXTs (e.g. Cannizzo et al. 1983, Lin & Taam 1984) these facts are natural consequences of long–term LMXB evolution.

2. CONDITION FOR A DISC INSTABILITY

The essential ingredient of the disc instability model is the possibility that hydrogen can exist in either neutral or ionized states, depending on circumstances. If a steady state exists in which all of the disc is above the hydrogen ionization temperature $T_H \sim$ 6500 K, the instability will be suppressed and the source will appear as persistent. As is well known, the effective temperature of a steady–state accretion disc powered by local viscous dissipation is given by (e.g. Frank et al, 1992)

$$T_{\rm eff}^4 = \frac{3GM\dot{M}}{8\pi\sigma R^3} \tag{1}$$

at disc radii R much larger than that of the central object (M is the mass of this star, $\dot{M}$ the accretion rate, G the gravitational constant and σ the Stefan–Boltzmann constant). Since $T_{\rm eff}$ decreases as $R^{-3/4}$, the condition $T_{\rm eff} > T_H$ for steady accretion is most stringent at the outer edge R_d of the disc. We define a critical accretion rate $\dot{M}_{\rm crit,\ eff}$ by the equation

$$T_{\rm eff}(R_d) = T_H. \tag{2}$$

Thus the disc instability is suppressed for $\dot{M} > \dot{M}_{\rm crit}$ and we can try to discriminate between steady and outbursting systems by checking this condition.

When applied to CVs this procedure correctly predicts that all disc–accreting systems below the period gap should be dwarf novae (see Section 3). Recently, van Paradijs (1996) has emphasized that LMXB discs are likely to be strongly influenced by irradiation from the central X–ray source, and that this will raise their surface temperatures and tend to stabilize them at lower accretion rates. Specifically, irradiation produces an effective temperature given by

$$T_{\rm irr}^4 = \frac{\eta\dot{M}c^2(1-\beta)}{4\pi\sigma R^2}\left(\frac{{\rm d}H}{{\rm d}R} - \frac{H}{R}\right), \tag{3}$$

where η is the efficiency of rest–mass energy conversion into X–ray heating, β is the X–ray albedo, and $H(R)$ is the local disc scale height. Since the factor in brackets is typically $\propto H/R \sim$ constant (see below), we see that $T_{\rm irr}^4$ falls off only as R^{-2} compared with R^{-3} for $T_{\rm eff}^4$. At the outer disc edge R_d we then expect $T_{\rm irr} > T_{\rm eff}$, and we should define a new (lower) critical accretion rate $\dot M_{\rm crit,irr}$ by the requirement

$$T_{\rm irr}(R_d) = T_H \tag{4}$$

rather than (2). To evaluate (2, 4) we assume that R_d is about 70% of the primary's Roche lobe radius R_L, for which in the regime of interest ($M_1 > M_2$, with M_1, M_2 the primary and secondary masses) the approximation $R_L \simeq [0.38 - 0.2\log(M_2/M_1)]a$ (Paczyński, 1971) holds. Here a is the binary separation. For typical mass ratios $M_2/M_1 \sim 0.1$ this gives $R_d \simeq 1.2 \times 10^{11} m_1^{1/3} P_d^{2/3}$ cm, where m_1 is the primary mass measured in $\rm M_\odot$ and P_d the binary period in days. For $T_{\rm irr}$ we need an estimate of $H(R)$: simple disc theory (cf Vrtilek et al., 1990) shows that for an irradiated disc H lies between the two laws $H \propto R^{9/7}, H \propto R^{9/8}$, while $H/R \simeq 0.2$ and $\beta \simeq 0.9$ (de Jong et al, 1996, and references therein). Using these results and $\eta = 0.11$ in (2, 4) gives

$$\dot M_{\rm crit,eff} \simeq 2.9 \times 10^{-9}\, P_3^2\, {\rm M_\odot\ yr^{-1}}, \tag{5}$$

and

$$\dot M_{\rm crit,irr} = \dot M_{\rm irr}\, m_1^{2/3}\, P_3^{4/3}, \tag{6}$$

with $\dot M_{\rm irr} \simeq 5 \times 10^{-11}\,{\rm M_\odot\ yr^{-1}}$ and $P_3 = P/(3\ {\rm hr})$.

3. LMXB EVOLUTION

From the reasoning above we can decide whether a given LMXB evolutionary sequence is likely to produce transient systems simply by comparing its predicted mass transfer rates with $\dot M_{\rm crit,irr}$. It is generally accepted that the LMXB phase follows a common–envelope episode; this leaves a detached system with a compact star (neutron star or black hole) and a low–mass main–sequence secondary star in an orbit with a typical period $\sim 0.5 - 10$ d. The system reaches contact and becomes an LMXB under the influence of two processes: (i) nuclear expansion of the secondary (eventually off the main sequence), on a timescale $t_{\rm MS}$; and (ii) shrinkage of the orbit on the angular–momentum loss timescale $t_{\rm AML}$. We get three cases, depending on the relative size of the two timescales and the initial masses and binary separation (cf Pylyser & Savonije 1988a, b; Singer et al. 1993; Ritter 1994). (We consider only stable mass transfer in this paper, generally assured for $M_2 \lesssim M_1$.)

(a) $t_{\rm MS} \ll t_{\rm AML}$. The secondary has time to evolve off the main sequence before angular momentum losses shrink the Roche lobe sufficiently to cause mass transfer. The secondary evolves as a subgiant, and transfers mass on its nuclear timescale $t_{\rm nuc}$, the binary evolving to longer periods (up to hundreds of days).

(b) $t_{\rm MS} \gg t_{\rm AML}$. Angular momentum losses shrink the binary orbit too rapidly for significant nuclear evolution of the secondary before mass transfer starts. The binary evolves to short periods (hours) with the secondary essentially unevolved at all times.

(c) $t_{\rm MS} \sim t_{\rm AML}$. The secondary is significantly evolved before mass transfer starts, but angular momentum losses shrink the binary orbit more rapidly than it expands. Once contact is established further nuclear evolution of the secondary is frozen. The binary evolves to short periods as in case (b), but this time with a secondary which is significantly larger than its main–sequence radius.

In case (a) the entire structure of the secondary is controlled by the mass of its helium core, quite independently of the secondary's total mass M_2: a semianalytic description is available (Webbink, Rappaport & Savonije, 1983) which can be used to show that (King, 1988)

$$-\dot{M}_2 \simeq 4.0 \times 10^{-10} P_d^{0.93} m_2^{1.47}\ \mathrm{M_\odot yr^{-1}}. \tag{7}$$

where m_2, P_d are M_2, P measured in $\mathrm{M_\odot}$ and days respectively. Comparison with (6) shows that, except possibly for rather massive secondary stars, the systems are likely to be transient. Since this evolution requires binary periods $\gtrsim 2$ d, and cases (b, c) have shorter periods, we can already predict that most LMXBs with periods $\gtrsim 2$ d will be transient. We note that this conclusion is independent of whether the primary is a neutron star or a black hole. The observational data are sparse, but not in conflict with this conclusion, particularly when we take account of the obvious selection effect against finding transients (in a subsequent paper we shall show that the recurrence time of long–period systems are probably so long that most will remain undiscovered).

Case (a) cannot supply the low values ($\lesssim 2$ d) which characterize the majority of measured SXT periods. For these we must turn to the other two cases. Case (b) is familiar from CV evolution: the mass transfer rate is a function of the orbital period. Fig. 1 shows evolutionary sequences calculated using a bipolytrope evolution code (Kolb & Ritter 1992) for various primary masses. For $M_1 = 1.4\mathrm{M_\odot}$ the predicted transfer rate is well below $\dot{M}_{\rm crit,eff}$ for systems below the period gap $P < 2$ hr), showing that disc instabilities can occur at these periods. This agrees with the observation that all non–magnetic CVs at these periods are dwarf novae. As is well known, the situation is much less clear above the period gap, since $\dot{M}_{\rm crit,eff}$ is very close to the predicted transfer rate; this is slightly higher than for the LMXB case shown as the typical primary mass in CVs is smaller (see e.g. Kolb 1996).
If irradiation is allowed, as appropriate in LMXBs, we see immediately that all these systems are likely to be steady. We thus reach another conclusion: short–period LMXBs containing neutron stars and main–sequence secondaries are likely to be persistent rather than transient.

In contrast to the nuclear–evolution case (a), the mass transfer rate in case (b) is quite sensitive to the primary mass M_1. First, larger M_1 raises the orbital angular momentum of the binary as $J \propto M_1^{2/3}$ for $M_1 \gg M_2$: if the angular momentum loss rate $\dot{J}$ is independent of M_1 (as is the case for e.g. magnetic stellar wind braking) the transfer rate is reduced as $-\dot{M}_2 \propto M_1^{-2/3}$, since

$$-\frac{\dot{M}_2}{M_2} \sim -\frac{\dot{J}}{J} \tag{8}$$

(e.g. King, 1988). On the other hand we know that angular momentum loss via gravitational radiation actually increases, with $-\dot{M}_2 \propto M_1^{2/3}$ (cf eq. 9 below). Since magnetic braking dominates for periods above the gap, where most LMXBs are found, the net effect is to reduce the predicted transfer rates but to flatten the curve of $-\dot{M}_2$ vs P at shorter orbital periods, where gravitational radiation becomes more important. This can be seen in Fig. 1 (middle panel), with $M_1 = 10\mathrm{M}_\odot$. The smaller contrast between the angular momentum losses above and below the period gap means that the latter is narrower, and indeed the minimum period is longer. (For sufficiently high M_1 the gap actually disappears altogether, cf Fig. 1, bottom panel, where M_1 is an unrealistically high $50\mathrm{M}_\odot$.) Mukai (1994) proposed that the reduced transfer rates for $M_1 = 10\mathrm{M}_\odot$ were responsible for the prevalence of black–hole systems among SXTs: however, while $-\dot{M}_2$ is indeed reduced below $\dot{M}_{\rm crit,eff}$, disc instabilities are still forbidden by irradiation, i.e. $-\dot{M}_2 > \dot{M}_{\rm crit,irr}$. We have to extend our conclusion that LMXBs with main–sequence secondaries are likely to be persistent rather than transient to the black–hole case also.

Thus, to explain the occurrence of SXTs at periods $\lesssim 2$ d we are forced to consider case (c). The main difference between this and case (b) immediately above is that the secondary has a lower mass for a given orbital period, i.e. $M_2 < M_2(\mathrm{MS})$, where $M_2(\mathrm{MS}) \simeq 0.33P_3\mathrm{M}_\odot$ is the mass of a main–sequence star that would just fill the Roche lobe at the current period. Simple scaling suggests that this effect significantly lowers the expected transfer rates. Using the form of Verbunt & Zwaan (1981) for the magnetic braking rate, with the radius of gyration set to $(0.2)^{1/2}$ and the calibration parameter to unity, and the standard form (e.g. Landau & Lifschitz 1958) for the gravitational radiation losses, the mass transfer rate is

$$-\dot{M}_2 = \dot{M}_{\rm MB}\, m_1^{-2/3}\hat{m}_2^{7/3}\, P_3^{5/3} \;+\; \dot{M}_{\rm GR}\, m_1^{2/3}\hat{m}_2^2\, P_3^{-2/3}, \qquad (9)$$

with $\hat{m}_2 = M_2/M_2(\mathrm{MS})$, $\dot{M}_{\rm MB} \simeq 2\times10^{-9}\,\mathrm{M}_\odot\ \mathrm{yr}^{-1}$ and $\dot{M}_{\rm GR} \simeq 7.6\times10^{-11}\,\mathrm{M}_\odot\mathrm{yr}^{-1}$, and we have assumed $M_1 \gg M_2$. For neutron–star primaries ($m_1 \simeq 1.4$) even quite evolved secondary stars still give rather high transfer rates, driven essentially by the magnetic braking term above; for example, $\hat{m}_2 \simeq 0.6$ implies $-\dot{M}_2 = 10^{-8}\,\mathrm{M}_\odot\ \mathrm{yr}^{-1}$ for the period $P = 18$ hr of Sco X–1. However for black–hole primaries ($m_1 \sim 10$), it appears possible to arrange $-\dot{M}_2 < \dot{M}_{\rm crit,irr}$, as required for SXT behaviour, if M_2 is significantly below the main–sequence mass. Fig. 2 shows the result of a computation using a full stellar evolution code (Mazzitelli 1989, adapted to treat binary evolution by Kolb & Ritter 1992), which confirms this impression. We can get a simple analytic representation of this by combining (6) and (9), obtaining an expression for the minimum primary mass required for SXT behaviour:

$$M_1(\mathrm{min,\ SXT}) = \frac{M_{\rm min}\hat{m}_2^{7/4}P_3^{1/4}}{(1 - h\hat{m}_2^2 P_3^{-2})^{3/4}}, \qquad (10)$$

where $M_{\rm min} = (\dot{M}_{\rm MB}/\dot{M}_{\rm irr})^{3/4}\mathrm{M}_\odot \simeq 16\mathrm{M}_\odot$ and $h = \dot{M}_{\rm GR}/\dot{M}_{\rm irr} \simeq 1.52$. Clearly SXT behaviour is impossible for any primary mass unless the denominator in (10) is positive, i.e.

$$\hat{m}_2 \lesssim k\, P_3 \;\; (\mathrm{GR\ alone}), \qquad (11)$$

where $k = (\dot{M}_{\rm irr}/\dot{M}_{\rm GR})^{1/2} \simeq 0.81$, which is simply the requirement that the mass transfer rate driven by GR alone must certainly be less than the critical rate. If the GR rate is well below this value, the second term in the denominator of (10) is negligible, and magnetic braking gives the requirement

$$M_1 > M_1({\rm min,\ SXT}) \simeq 16{\rm M}_\odot\, \hat{m}_2^{7/4}\, P_3^{1/4} \quad ({\rm MB\ alone}). \tag{12}$$

Clearly *both* the requirements (11) and (12) must hold for SXT behaviour, and indeed the true condition (10) is stronger than either. For example, the SXT UY Vol (=0748-676) has a neutron star primary, since it shows X–ray bursts. With $M_1 = 1.4{\rm M}_\odot$ and $P = 3.82$ hr, we find $M_2/M_2({\rm MS}) < 0.235$ from (10).

4. CONCLUSIONS

We can summarize our conclusions in Table 1.

secondary	primary	
	neutron star	**black hole**
main sequence	all persistent	persistent/some SXTs?
evolved	persistent/SXTs	all SXTs

As can be seen, evolved secondaries are required for SXT behaviour, except possibly for some fairly massive black hole systems. Clearly the combination of a black hole and an evolved secondary is strongly favoured, in agreement with observation. However, we need to explain not only why black holes are common among SXTs, but also why they are far more common (8 out of 14) among transient than persistent LMXBs (1 out of 29), see King et al. 1996 (in prep.) or McClintock (1996). Table 1 shows that this must mean that sufficiently evolved companions are more common in black–hole than neutron–star systems. There is an obvious explanation for this: in the pre–contact phase (cf Section 3) the orbital evolution time $t_{\rm AML}$ of black–hole systems is significantly longer (by a factor $m_1^{2/3} \sim 4$) than for neutron–star systems, for the same reason that the post–contact mass transfer rate is lower for black holes, i.e. that $(\dot{J}/J)_{\rm MB}$ is smaller, (cf the first term of eqn 9). The secondaries in black–hole systems have on average 4 times as long to evolve along or off the main sequence before there is any possibility of their being driven to short periods by angular momentum losses. Yet eqn (10) requires that the secondaries in neutron–star systems should be significantly *more* evolved ($\hat{m}_2 \lesssim 0.25$) than in black–hole systems ($\hat{m}_2 \lesssim 0.75$) for SXT behaviour to occur. Thus for similar initial separations, a considerably smaller fraction of neutron–star than black–hole binaries will become SXTs. Of course a full population synthesis will be required to demonstrate that this conclusion is quantitatively as well as qualitatively valid. A further consequence of the low value of $\hat{m}_2$ in neutron–star SXTs is that the mass ratio must be very small in both types of transient: for $P \lesssim 2$ d we have $M_2/M_1 \simeq 0.33 P_3 \hat{m}_2/m_1$, giving $M_2/M_1 \lesssim 0.059 P_3$ and $M_2/M_2 \lesssim 0.025 P_3$ for neutron–star transients with $M_1 = 1.4{\rm M}_\odot$ and black–hole systems with $M_1 = 10{\rm M}_\odot$ respectively. Such extreme mass ratios are a ubiquitous feature of observed SXTs, lending support to the picture presented here.

We conclude that soft X–ray transient events probably result from accretion disc instabilities. In line with this, King & Ritter (1996, in prep.) find outburst and recurrence times of SXTs in good accord with those expected from simple extrapolations from dwarf nova outbursts, provided that due account is taken of the larger disc size in SXTs. The arguments presented here show that the high incidence of black hole identifications in SXTs given by dynamical mass estimates directly reflects the underlying population rather than observational selection.

We thank Jan van Paradijs for providing us with a copy of his paper in advance of publication, and for very helpful comments on an earlier version of this paper, which led to significant improvements. This work was supported by the UK Particle Physics and Astronomy Research Council through a Senior Fellowship (ARK) and a Rolling Grant to the Leicester Astronomy Group.

REFERENCES

Cannizzo J.K., Wheeler J.C., Ghosh P. 1985, in *Cambridge Workshop on Cataclysmic Variables and Low–Mass X–Ray Binaries*, D.Q. Lamb, J.Patterson (eds.), Reidel, Dordrecht, p. 307

De Jong J.A., van Paradijs J., Augusteijn T. 1996, A&A, in press

Frank J., King A.R., Raine D.J. 1992, *Accretion Power in Astrophysics*, Cambridge University Press

King A.R. 1988, QJRAS 29, 1

Kolb U. 1996, in *Cataclysmic Variables and Related Objects*, IAU Coll. 158, A. Evans, J.H. Wood (eds.), Kluwer Academic publisher, in press

Kolb U., Ritter H. 1992, A&A 254, 213

Landau L., Lifschitz E. 1958, *Classical Theory of Fields*, Pergamon Press

Lin D.N.C., Taam R.E. 1984, in *High Energy Transients in Astrophysics*, S.E. Woosley (ed.), AIP Conf. Proc. 115, p. 83

Mazzitelli I. 1989, ApJ 340, 249

McClintock J.E. 1996, talk at Aspen Workshop on Black–Hole Soft X–ray Transients

Mukai K. 1994, in *The Evolution of X-Ray Binaries*, S.S. Holt, C.S. Day (eds.), AIP Press, New York, p. 79

Paczyński B.E. 1971, ARAA 9, 183

Pylyser E.H.P., Savonije G.J. 1988a, A&A 191, 57

Pylyser E.H.P., Savonije G.J. 1988b, A&A 208, 52

Ritter H. 1994, in *Evolutionary links in the zoo of interactive binaries*, F. D'Antona, V. Caloi, C. Maceroni, F. Giovanelli (eds.), in Memorie della Società Astronomica Italiana 65, p. 173

Singer R., Kolb U., Ritter H. 1993, AG Abstract Series 9, 39

Van Paradijs J. 1996, talk at Aspen Workshop on Black–Hole Soft X–ray Transients, and ApJ, in press

Verbunt F., Zwaan C. 1981, A&A 100, L7

Vrtilek S.D., Raymond J.C., Garcia M.R., Verbunt F., Hasinger G. 1990, A&A 235, 162

Webbink R.F., Rappaport S., Savonije G.J. 1983, ApJ 270, 678

This preprint was prepared with the AAS LaTeX macros v4.0.

FIGURES

Figure 1:
Mass transfer rate $\dot{M}$ versus orbital period P for LMXBs with various primary masses and an unevolved secondary of initial mass $1\mathrm{M}_{\odot}$. Magnetic braking according to (9) is assumed to operate only as long as the secondary has a radiative core. The dotted and dashed lines are the critical mass transfer rates (5) and (6) for standard and irradiated discs respectively. Systems with lower mass transfer rates have thermally unstable accretion discs and are transient.
Top panel: $1.4\mathrm{M}_{\odot}$ neutron star primary. A detached phase occurs when the secondary becomes fully convective, followed by the resumption of mass transfer below the period gap. For CVs with a typical primary mass $\simeq 0.6\mathrm{M}_{\odot}$ the mass transfer rate above the period gap is $\simeq 1.8$ times higher than shown and therefore very close to the classical instability line (cf e.g. Kolb 1996).
Middle panel: $10\mathrm{M}_{\odot}$ BH primary. The mass transfer rate is below the standard stability limit (5) but above the limit for an irradiated disc, i.e. the system is stable at all periods.
Bottom panel: unphysically massive $50\mathrm{M}_{\odot}$ BH primary. Mass transfer via magnetic braking is reduced and that via gravitational radiation increased, so the period gap disappears and the minimum period is longer than for lower primary masses.

Figure 2:
Mass transfer rate $\dot{M}$ versus orbital period P for a $10\mathrm{M}_{\odot}$ BH binary with a slightly evolved secondary. At the onset of mass transfer the $1.2\mathrm{M}_{\odot}$ secondary was close to the end of central hydrogen burning (central H mass fraction 2%). Also shown as a dash–dotted curve is the corresponding track for a CV with $1\mathrm{M}_{\odot}$ white dwarf primary, taken from Singer et al. (1993). For a better comparison the CV track was shifted downwards in $\log \dot{M}[\mathrm{M}_{\odot}\mathrm{yr}^{-1}]$ by 2/3 to compensate for the less massive primary, as (9) suggests. For short orbital periods the BH mass transfer rate is above the rescaled CV transfer rate since gravitational radiation begins to dominate over magnetic braking. The secondary never becomes fully convective, so there is no detached phase. The disk instability lines (5) and (6) are again shown as dotted and dashed lines, see Fig. 1. The system is always transient.

Figure 1

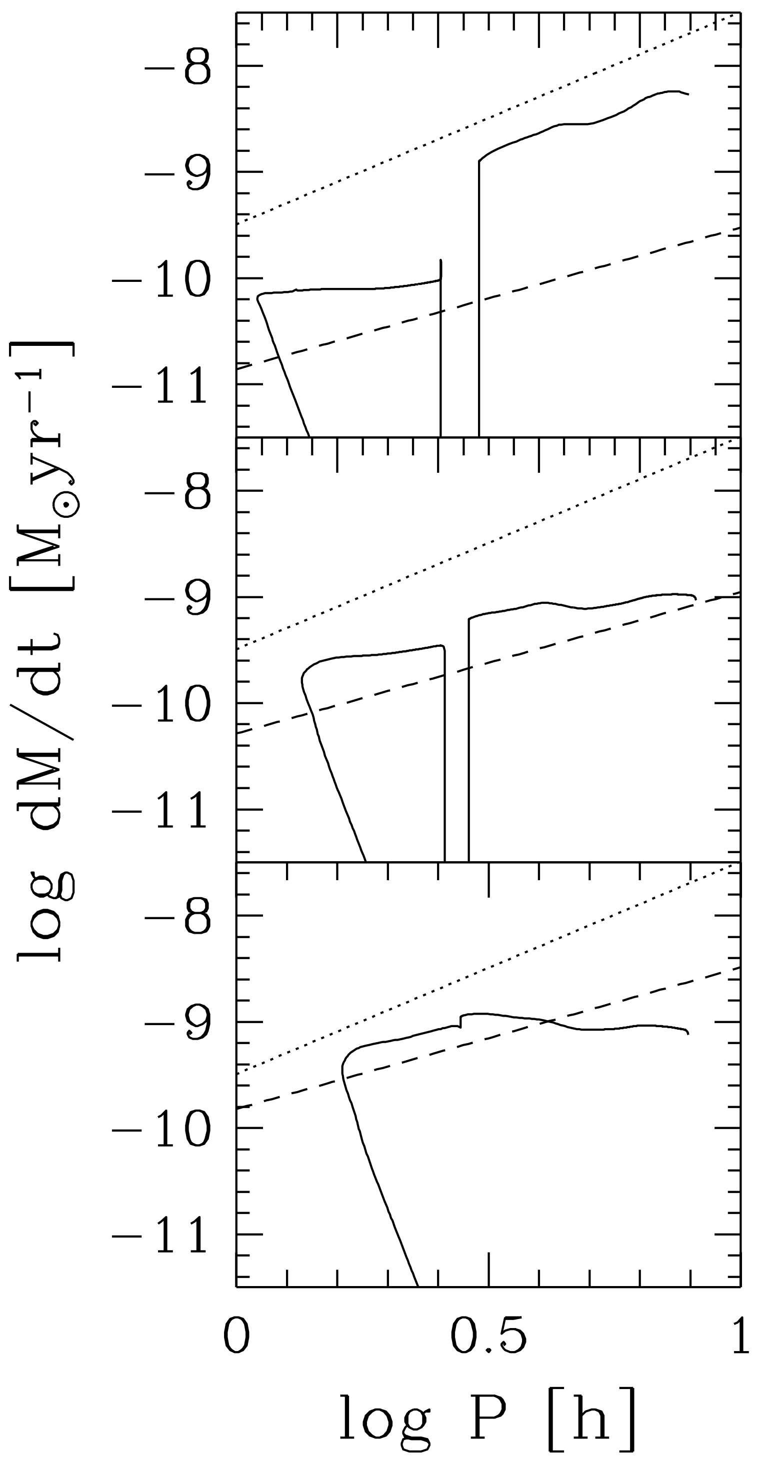

Figure 2