THE ASTROPHYSICAL JOURNAL, 292:L41–L44, 1985 May 15

THE COLLAPSE OF DENSE STAR CLUSTERS TO SUPERMASSIVE BLACK HOLES: THE ORIGIN OF QUASARS AND AGNs

STUART L. SHAPIRO[1] AND SAUL A. TEUKOLSKY[1]
Center for Radiophysics and Space Research, Cornell University
Received 1985 January 14; accepted 1985 February 28

ABSTRACT

Fully general relativistic calculations of the gravitational collapse of collisionless equilibrium systems have recently been carried out by Shapiro and Teukolsky. We apply the results of such calculations, together with recent Newtonian Fokker-Planck calculations of the "gravothermal catastrophe," to the dynamical evolution of a dense cluster of compact stars—neutron stars or stellar mass black holes. A plausible scenario for the formation of supermassive black holes via the collapse of such clusters embedded in evolved galactic nuclei is described. The process leads naturally to the birth of supermassive black holes of the "right size" to explain quasars and AGNs: $10^6 \lesssim M/M_\odot \lesssim 10^9$.

Subject headings: black holes — galaxies: nuclei — gravitation — quasars

I. INTRODUCTION

There is strong evidence that quasars, AGNs, and other intense extragalactic radio sources are powered by jets of plasma and magnetic fields and that each jet is produced by a compact supermassive object $M \gtrsim 10^6$–$10^9\ M_\odot$ in the nucleus of a galaxy. A supermassive black hole is the likely candidate for the compact object (see reviews by Begelman, Blandford, and Rees 1984; Rees 1984; and references therein for compelling arguments).

If the black hole hypothesis is indeed correct, an important question remains: *how and under what circumstances did such a supermassive black hole form*? Rees and his collaborators (see especially Rees 1978; Begelman and Rees 1978; and Rees 1984) have sketched possible routes by which a supermassive black hole may form via runaway evolution in active galactic nuclei.

One of the first routes ever proposed by which a dense stellar system might form a supermassive black hole was the path suggested by Zel'dovich and Podurets (1965, hereafter ZP; see also Fackerell, Ipser, and Thorne 1969; Zel'dovich and Novikov 1971; and Misner, Thorne, and Wheeler 1973 for subsequent discussions of their scenario). They argued that the combined effects of *secular* core collapse, i.e., the "gravothermal catastrophe" (which they referred to as "stellar evaporation") and star-star collisions and coalescence would inevitably drive a star cluster core to states of ever-increasing central density and redshift. They further suggested that at sufficiently high central redshift, the cluster would become relativistically unstable, at which point it would undergo catastrophic collapse to a supermassive black hole on a *dynamical* time scale, t_d. The original discussion of ZP focused on an initially Newtonian star cluster composed of stellar mass black holes. They argued that such a cluster would inevitably collapse to form a central supermassive black hole around which any remnant cluster stars would orbit.[2]

In this *Letter* we reassess the scenario of ZP in light of two recent theoretical developments in stellar dynamics. Our reanalysis suggests that, with some modification, this scenario ought to be regarded as a leading contender for the route by which supermassive black holes form in dense galactic nuclei.

First, the gravothermal catastrophe has only recently been elucidated by detailed (1 + 1)– and (2 + 1)–dimensional Fokker-Planck calculations that follow core collapse for many e-foldings (Cohn 1979, 1980; Marchant and Shapiro 1980; Duncan and Shapiro 1982). (The gravothermal catastrophe is the process by which the cumulative effect of distant, two-body gravitational encounters causes the core of a self-gravitating, large-N, Newtonian system to undergo secular collapse on a *relaxation* time scale, $t_r \gg t_d$.) These Newtonian calculations show that advanced core collapse in spherical clusters is *not* described by the simple homologous "stellar evaporation" model, derived by Ambartsumian (1938) and Spitzer (1940) and assumed by ZP. Nevertheless, core collapse does proceed homologously, a result that we shall exploit below.

Second, we can report that, for the first time, it is now possible to numerically integrate the full Einstein equations for the dynamical evolution of an arbitrary spherical, collisionless configuration in general relativity (Shapiro and Teukolsky 1985*a*, *b*). These (Vlasov) integrations enable one to follow on the computer the evolution on dynamical time scales of relativistic star clusters, even during epochs characterized by total gravitational collapse leading to the formation of a supermassive black hole.

[1]Departments of Astronomy and Physics, Cornell University.

[2]Following their semiquantitative analysis, ZP urged that "it is very desirable to perform calculations for such a model." With no progress reported by 1971, Zel'dovich and Novikov (1971) again exhorted: "A numerical investigation of such collapse is badly needed, but no one up to now has had the fortitude to attempt it."

These fully relativistic numerical integrations provide dramatic confirmation of the original speculation of ZP that star clusters become relativistically unstable at sufficiently high central redshift, $z_c \gtrsim 0.5$ (see Fig. 1 [Pl. L1]). While this instability had already been demonstrated rigorously in perturbation theory (Ipser and Thorne 1968; Ipser 1969*a*, *b*; Fackerell 1970), the numerical integrations can follow its nonlinear growth and determine the ultimate fate of unstable clusters. In particular, the integrations show quite generally that clusters of sufficiently high z_c *do* undergo catastrophic collapse to a black hole on a dynamical time scale. They further reveal that *even in the case of extremely centrally condensed configurations with extensive Newtonian halos, an appreciable fraction of the total mass ultimately collapses to a central black hole in a few crossing times*, provided the central core is sufficiently relativistic. This result is due to the cascading nature of collapse in collisionless configurations, a collective phenomenon predicted by ZP and called "avalanche-type contraction" by them.

The recent Fokker-Planck and general relativistic dynamical calculations provide fresh impetus for reconsidering the ZP supermassive black hole formation scenario.

II. THE BASIC SCENARIO

The starting point of our calculations is a very dense, Newtonian star cluster embedded in an evolved galactic nucleus. The cluster is assumed to consist entirely of neutron stars or stellar mass black holes. Various authors have performed evolutionary calculations of dense ($\gtrsim 10^8$ stars pc^{-3}), but otherwise normal, galactic nuclei initially composed of main-sequence stars (Gold, Axford, and Ray 1965; Spitzer and Saslaw 1966; Spitzer and Stone 1967; Colgate 1967; Sanders 1970; Spitzer 1971). They find that the evolution is invariably dominated by an epoch of stellar collisions and coalescence and the buildup of massive stars. These stars undergo supernova explosions that can leave behind compact stellar mass remnants. Consequently, *the endpoint of this "coalescence phase" is very likely a cluster of neutron stars or stellar mass black holes* (Begelman and Rees 1978). We infer from the calculations of Colgate (1967) and Sanders (1970) that such a cluster might originally consist of 10^7–(2×10^8) compact objects, each of 1–10 $M_\odot$, confined to a region 0.01–0.1 pc in radius and moving with a velocity dispersion 800–2000 km s^{-1}. We shall ignore any residual gas or angular momentum in the cluster and assume spherical symmetry.

The subsequent evolution of a Newtonian cluster of compact stars is characterized by two distinct epochs: (1) an initial long, low-redshift ($z_c < z_{\rm coll} \approx 10^{-2}$) "point-mass" epoch during which the core undergoes secular collapse via the gravothermal catastrophe, and (2) a later short, high-redshift ($z_c \gtrsim z_{\rm coll}$) "finite-radius" epoch during which compact star collisions and coalescence dominate the core evolution. During the brief collision-coalescence period, a cluster core of neutron stars undergoes an internal "phase change" to a core of stellar mass black holes: when such collisions become dynamically important, the velocity dispersion in the core v_c is less than the escape speed from the neutron star surface. Consequently, 1.4 $M_\odot$ neutron stars collide at free-fall velocity, which results in coalescence and little mass loss (Seidl and Cameron 1972; Gilden and Shapiro 1984). Upon cooling by rapid neutrino emission (with a time scale $\approx$ minutes), the coalesced neutron star collapses to form an approximately 2.8 $M_\odot$ black hole, which will later collide with another compact star, and so forth.

During both epochs the number of core stars decreases while the central velocity dispersion and redshift increase. Ultimately the core becomes relativistic ($z_c = z_{\rm crit} \approx 0.5$), at which point it undergoes catastrophic gravitational collapse to a black hole. The hole rapidly grows outward beyond the core, eventually swallowing a significant fraction of the total mass of the compact star cluster in a few mean dynamical time scales via the "avalanche" instability.

The subsequent dynamical evolution of a dense galactic nucleus with a central supermassive black hole has been extensively studied (Hills 1975; Rees 1978; Young, Shields, and Wheeler 1977; Frank 1978; McMillan, Lightman, and Cohn 1981; Duncan and Shapiro 1983). Gas liberated during the collisions of residual main-sequence stars and tidal disruptions by the hole is consumed by the hole, causing it to grow significantly. This gas accretion can efficiently generate AGN or quasar luminosities at the Eddington limit, provided the hole acquires a mass $\gtrsim 10^6$–10^9 $M_\odot$ (see Shapiro 1985 for a brief summary and references).

III. BASIC EQUATIONS

The core of a Newtonian cluster of compact stars evolves approximately according to the homological differential equations ($c = G = 1$)

$$\dot{E}_c/E_c = (-\alpha_1/t_r) + (1/t_{\rm coll}), \tag{1}$$

$$\dot{N}_c/N_c = (-\alpha_2/t_r) - (1/t_{\rm coll}), \tag{2}$$

$$\dot{m}_c/m_c = 1/t_{\rm coll}, \tag{3}$$

(cf. Lightman and Shapiro 1978). Here $N_c(t)$ is the number of core stars, $m_c(t)$ is the mean stellar mass, and $E_c(t)$ is the characteristic core binding energy. These quantities are related to the characteristic core density n_c, rms velocity v_c, and radius r_c according to the virial theorem

$$E_c \approx \tfrac{1}{2} N_c m_c v_c^2 \approx \tfrac{1}{4}(N_c m_c)^2/r_c, \tag{4}$$

and the relations

$$n_c \approx N_c/(4\pi r_c^3/3) \approx 2 \times 10^9 \ {\rm pc}^{-3}(N_{c,8}^{-2} v_{c,3}^6 m_{c,*}^{-3}), \tag{5a}$$

$$r_c \approx 0.2 \ {\rm pc}\,(N_{c,8} v_{c,3}^{-2} m_{c,*}), \tag{5b}$$

where $N_{c,8} \equiv N_c/10^8$, $v_{c,3} \equiv v_c/10^3$ km s^{-1}, and $m_{c,*} \equiv m_c/M_\odot$. The central relaxation time scale is given by Spitzer and Hart (1971):

$$t_r \approx v_c^3 \bigg/ \left[\left(\frac{3}{2}\right)^{1/2} 4\pi m_c^2 n_c \ln(0.4 N_c)\right]$$

$$\approx 0.8 \times 10^8 \ {\rm yr}\,(N_{c,8}^2 v_{c,3}^{-3} m_{c,*} \Lambda_8^{-1}), \tag{6}$$

where $\Lambda_8 \equiv \ln(0.4 N_c)/\ln(0.4 \times 10^8)$. The corresponding dynamical time scale $t_d \ll t_r$ is

$$t_d \approx r_c/v_c \approx 200 \text{ yr} \left(N_{c,8} v_{c,3}^{-3} m_{c,*} \right). \tag{7}$$

The central two-body collision time scale is

$$t_{\text{coll}} \approx \frac{1}{n_c \sigma_{\text{coll}} v_\infty} \approx 6 \times 10^{13} \text{ yr} \left(N_{c,8}^2 v_{c,3}^{-5} m_{c,*} \right), \tag{8}$$

where $v_\infty \approx 2^{1/2} v_c$ is the typical asymptotic relative velocity, and σ_{coll} is the collision cross section. For any compact object with radius $R < 4m$, the critical specific angular momentum for test particle capture is $\tilde{l} = 4m$, implying

$$\sigma_{\text{coll}} \approx \pi \tilde{l}^2 / v_\infty^2 = 16 \pi m^2 / v_\infty^2. \tag{9}$$

We employ equation (9) as an approximation to σ_{coll} for both neutron stars and black holes. (Note that a 1.4 $M_\odot$ neutron star constructed from, e.g., a moderately stiff TNI equation of state has $R \approx 5.1m$ and $\tilde{l} \approx 4.1m$.)

The constants α_1 and α_2 appearing in equations (1) and (2) may be fixed by matching to the detailed Fokker-Planck calculations for advanced core collapse in the absence of collisions ("point-mass" epoch). These calculations give

$$\begin{aligned} d \ln n_c / d \ln r_c &= -\beta, & \beta &= 2.23, \\ \dot{n}_c / n_c &= \xi / t_r, & \xi &= 3.6 \times 10^{-3}, \\ d \ln v_c^2 / d \ln n_c &= \gamma, & \gamma &= 0.10, \end{aligned} \tag{10}$$

(Cohn 1980). Solving equations (1)–(3) with $t_{\text{coll}} = \infty$ yields $\gamma = 1 - (2/\beta)$, $\alpha_1 = (5 - 2\beta)\xi/\beta = 8.72 \times 10^{-4}$, $\alpha_2 = (3 - \beta)\xi/\beta = 1.24 \times 10^{-3}$ (hence $\alpha \equiv \alpha_1/\alpha_2 = 0.701$) and

$$\tau = \left\{ \left[1 - (3\gamma/2) \right] \xi \right\}^{-1} t_r = 330 t_r, \tag{11}$$

where τ is the time remaining for the gravothermal catastrophe to drive secular core collapse to completion ($N_c \to 0$). We note that "stellar evaporation" theory gives $\alpha = \alpha_1 = 0$ and $\alpha_2 = 7.4 \times 10^{-3}$, which results in quite different evolution.

It proves convenient to replace equation (1) by an analogous equation for z_c. Cohn (1980) finds

$$z_c \approx \phi_c \approx 14 v_c^2/3 \approx 5 \times 10^{-5} v_{c,3}^2 \tag{12}$$

at late times, where ϕ_c is the central potential. Equation (1) then becomes

$$\dot{z}_c / z_c = \left[(\alpha_2 - \alpha_1)/t_r \right] + 1/t_{\text{coll}}. \tag{1'}$$

We solve equations (1′), (2), and (3) subject to two constraints: (i) $\tau \leq H^{-1} = 2 \times 10^{10}$ yr; and (ii) $N_c(z_{\text{crit}}) \geq 1$. Constraint ($i$) guarantees that secular core collapse to a relativistic state occurs in less than a Hubble time. Constraint (ii) ensures that this collapse will not stall before reaching $z_c = z_{\text{crit}}$ because of the diffusion of all the core stars out into the halo.

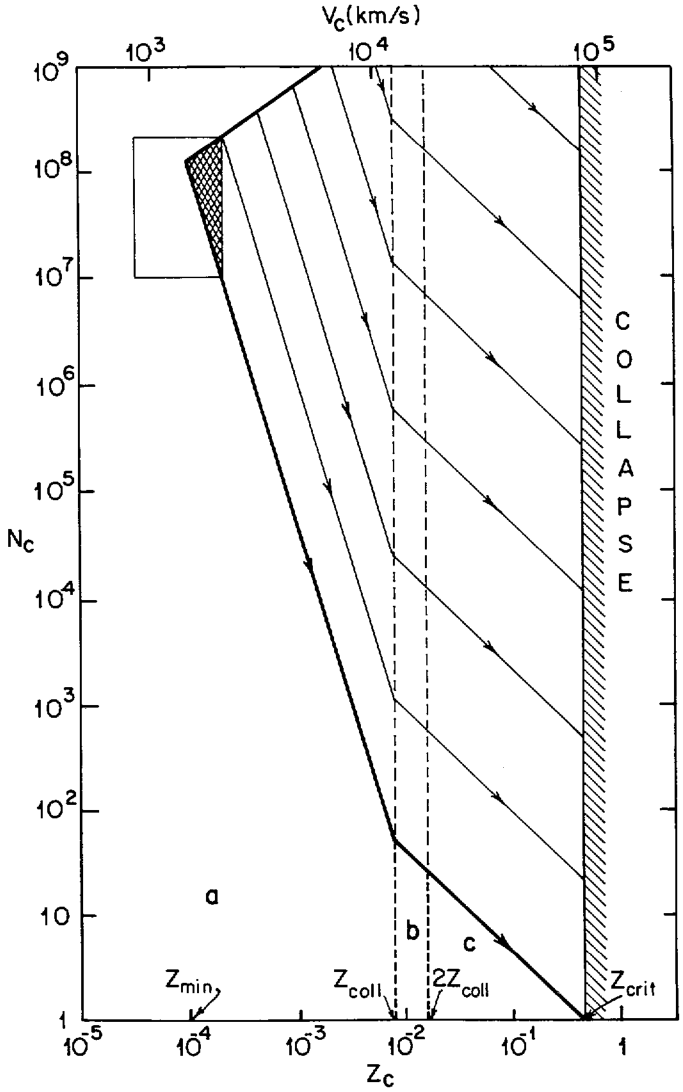

FIG. 2.—Dynamical evolution tracks in the N_c-z_c plane of dense cluster cores consisting of 1.4 $M_\odot$ neutron stars. During phase (a), the gravothermal catastrophe, the core evolves by two-body scattering to central redshift $z_{\text{coll}} \approx 8 \times 10^{-3}$. During phase (b) the neutron stars collide, coalesce, and collapse to form $\sim 2.8\ M_\odot$ black holes. During phase (c) the remnant black holes collide and coalesce, driving the core to central redshift $z_{\text{crit}} \approx 0.5$. At this point the entire cluster becomes relativistically unstable and undergoes catastrophic collapse to a supermassive black hole in a few mean orbital periods. The range of plausible core parameters for dense galactic nuclei lies in the box shown in the upper left of the figure. The hatched region shows the fraction of these capable of evolving to a supermassive black hole in a Hubble time.

IV. APPROXIMATE SOLUTION

Equations (1′), (2), and (3) may be combined to give the identity

$$N_c \propto \left(m_c^\alpha / z_c \right)^{1/(1-\alpha)} = m_c^{2.35} z_c^{-3.35}. \tag{13}$$

This relation may be evaluated separately in the "point-mass" and "finite-radius" epochs to yield an approximate solution for $N_c(z_c)$ and $m_c(z_c)$. During the "point-mass" phase $z_c <$

$z_{\rm coll}$, collisions can be ignored, whereby

$$m_c = \text{constant}, \qquad N_c \propto z_c^{-3.35}, \qquad (z_c < z_{\rm coll}). \quad (14)$$

During the "finite-radius" epoch $z_c > z_{\rm coll}$ collisions dominate, m_c increases, but the total core mass $N_c m_c$ remains constant. Thus

$$m_c \propto z_c, \qquad N_c \propto z_c^{-1}, \qquad (z_c > z_{\rm coll}). \quad (15)$$

The value of $z_{\rm coll}$ is determined by equating the gravothermal and collision terms in equation (2) and employing equations (6), (8), (12), and (15): $v_{c,\rm coll} \approx 0.04$ or

$$z_{\rm coll} \approx 8 \times 10^{-3}. \quad (16)$$

In the case of a cluster of neutron stars of mass m, the "phase transition" to a core of stellar mass black holes of mass $2m$ during the collision-coalescence epoch is essentially complete by $z_c \approx 2z_{\rm coll}$, according to equation (15). In any case, the subsequent buildup of stellar mass black holes during this epoch for either compact star scenario does not proceed very far ($t_{\rm coll} \gtrsim$ days) before $z_c \approx z_{\rm crit} \approx 0.5$ and catastrophic collapse of the cluster ensues. By this time the cluster has developed a *very* extreme core-halo structure: most of the mass is in halo stars in highly eccentric orbits. Note that "hard" binaries formed either by three-body encounters or two-body tidal dissipation during this phase decay rapidly via gravitational radiation emission (time scale $\approx m v_c^{-8} \lesssim$ days for $v_c \gtrsim v_{c,\rm coll}$).

The evolutionary tracks of neutron star cores are plotted in Figure 2. At any instant the cores are specified by two parameters: N_c and z_c (or v_c). The two distinct segments comprising each track correspond to the "point-mass" gravothermal and "finite-radius" collision-coalescence epochs. Constraints (*i*) and (*ii*) confine the tracks to lie within the heavy lines in the figure. The range of plausible initial core parameters, according to the calculations of Colgate (1967) and Sanders (1970) (see § II), is indicated by the box in the upper left-hand corner of the figure. The corresponding figure for stellar mass black hole cores is virtually identical.

V. DISCUSSION

Several interesting consequences emerge from Figure 2. First, there exists a *minimum* velocity dispersion below which a core of compact stars cannot evolve to a relativistic state in a Hubble time,

$$v_{c,\rm min} \approx 10^3 m_{c,*}^{\delta} \ \text{km s}^{-1},$$
$$\delta = (1-\alpha)/(7-3\alpha) = 0.061. \quad (17)$$

Second, a finite fraction of those compact star clusters likely to form in dense galactic nuclei is capable of evolving to a supermassive black hole in a Hubble time. Moreover, these clusters have roughly $10^7 \lesssim N_c \lesssim 10^9$ compact stars. Hence, following gravitational collapse, supermassive black holes with masses in the range $10^6 \lesssim M/M_\odot \lesssim 10^9$ can ultimately form. This is the mass range frequently cited in black hole models of such systems (cf. § I).

The high-resolution observations of the nucleus of the Seyfert galaxy NGC 4151 and of M31 by the balloon-borne Stratoscope II telescope revealed stellar conditions close to those capable of generating dense clusters of compact stars that can undergo catastrophic collapse. Schwarzschild (1973) estimates that the stars in the central 3.5 pc of the core of NGC 4151 contain fully $4 \times 10^9\ M_\odot$ at a mean density of $2 \times 10^8\ M_\odot$ pc^{-3} and velocity dispersion of $\gtrsim 1.5 \times 10^3$ km s^{-1}. Light, Danielson, and Schwarzschild (1974) calculate that the inner 0.5 pc of the nucleus of M31 may contain 10^8 $M_\odot$.

Planned observations of these and other nuclei with the high-resolution, optical detectors aboard the Space Telescope should, when used in conjunction with Figure 2, *determine whether stellar conditions in these regions are suitable for triggering the black hole formation scenario we have described.*

This work has been supported in part by National Science Foundation grants AST 81-16370 and PHY 83-05288 at Cornell University.

REFERENCES

Ambartsumian, V. A. 1938, *Ann. Leningrad State Univ.*, No. 22 (Astr. Series, Issue 4).
Begelman, M. C., Blandford, R. D., and Rees, M. J. 1984, *Rev. Mod. Phys.*, **56**, 255.
Begelman, M. C., and Rees, M. J. 1978, *M.N.R.A.S.*, **185**, 847.
Cohn, H. 1979, *Ap. J.*, **234**, 1036.
______. 1980, *Ap. J.*, **242**, 765.
Colgate, S. A. 1967, *Ap. J.*, **150**, 163.
Duncan, M. J., and Shapiro, S. L. 1982, *Ap. J.*, **253**, 921.
______. 1983, *Ap. J.*, **268**, 565.
Fackerell, E. D. 1970, *Ap. J.*, **160**, 859.
Fackerell, E. D., Ipser, J. R., and Thorne, K. S. 1969, *Comments Ap. Space Phys.*, **1**, 134.
Frank, J. 1978, *M.N.R.A.S.*, **184**, 87.
Gilden, D., and Shapiro, S. L. 1984, *Ap. J.*, **287**, 728.
Gold, T., Axford, W. I., and Ray, E. C. 1965, in *Quasi-stellar Sources and Gravitational Collapse*, ed. I. Robinson *et al.* (Chicago: University of Chicago Press), p. 93.
Hills, J. G. 1975, *Nature*, **254**, 295.
Ipser, J. R. 1969*a*, *Ap. J.*, **156**, 509.
______. 1969*b*, *Ap. J.*, **158**, 17.
Ipser, J. R., and Thorne, K. S. 1968, *Ap. J.*, **154**, 251.
Light, E. S., Danielson, R. E., and Schwarzschild, M. 1974, *Ap. J.*, **194**, 257.
Lightman, A. P., and Shapiro, S. L. 1978, *Rev. Mod. Phys.*, **50**, 437.
Marchant, A. B., and Shapiro, S. L. 1980, *Ap. J.*, **239**, 685.
McMillan, S. L. W., Lightman, A. P., and Cohn, H. 1981, *Ap. J.*, **251**, 436.
Misner, C. W., Thorne, K. S., and Wheeler, J. A. 1973, *Gravitation* (San Francisco: Freeman).
Rees, M. J. 1978, *Phys. Scripta*, **17**, 193.
______. 1984, *Ann. Rev. Astr. Ap.*, **22**, 471.
Sanders, R. 1970, *Ap. J.*, **162**, 791.
Schwarzschild, M. 1973, *Ap. J.*, **182**, 357.
Seidl, F. G. P., and Cameron, A. G. W. 1972, *Ap. Space Sci.*, **15**, 44.
Shapiro, S. L. 1985, in *IAU Symposium 113, Dynamics of Star Clusters*, ed. J. Goodman and P. Hut (Dordrecht: Reidel), in press.
Shapiro, S. L., and Teukolsky, S. A. 1985*a*, *Ap. J.*, submitted.
______. 1985*b*, *Ap. J.*, submitted.
Spitzer, L. 1940, *M.N.R.A.S.*, **100**, 396.
______. 1971, in *Galactic Nuclei*, ed. D. O'Connell (Amsterdam: North-Holland), p. 443.
Spitzer, L., and Hart, M. H. 1971, *Ap. J.*, **164**, 399.
Spitzer, L., and Saslaw, W. C. 1966, *Ap. J.*, **143**, 400.
Spitzer, L., and Stone, M. E. 1967, *Ap. J.*, **147**, 519.
Young, P. J., Shields, G. A., and Wheeler, J. C. 1977, *Ap. J.*, **212**, 367.
Zel'dovich, Ya. B., and Novikov, I. D. 1971, *Relativistic Astrophysics*, Vol. **1** (Chicago: University of Chicago Press).
Zel'dovich, Ya. B., and Podurets, M. A. 1965, *Astr. Zh.*, **42**, 963 (English translation in *Soviet Astr.—AJ*, **9**, 742) (ZP).

S. L. Shapiro and S. A. Teukolsky: Space Sciences Building, Cornell University, Ithaca, NY 14853

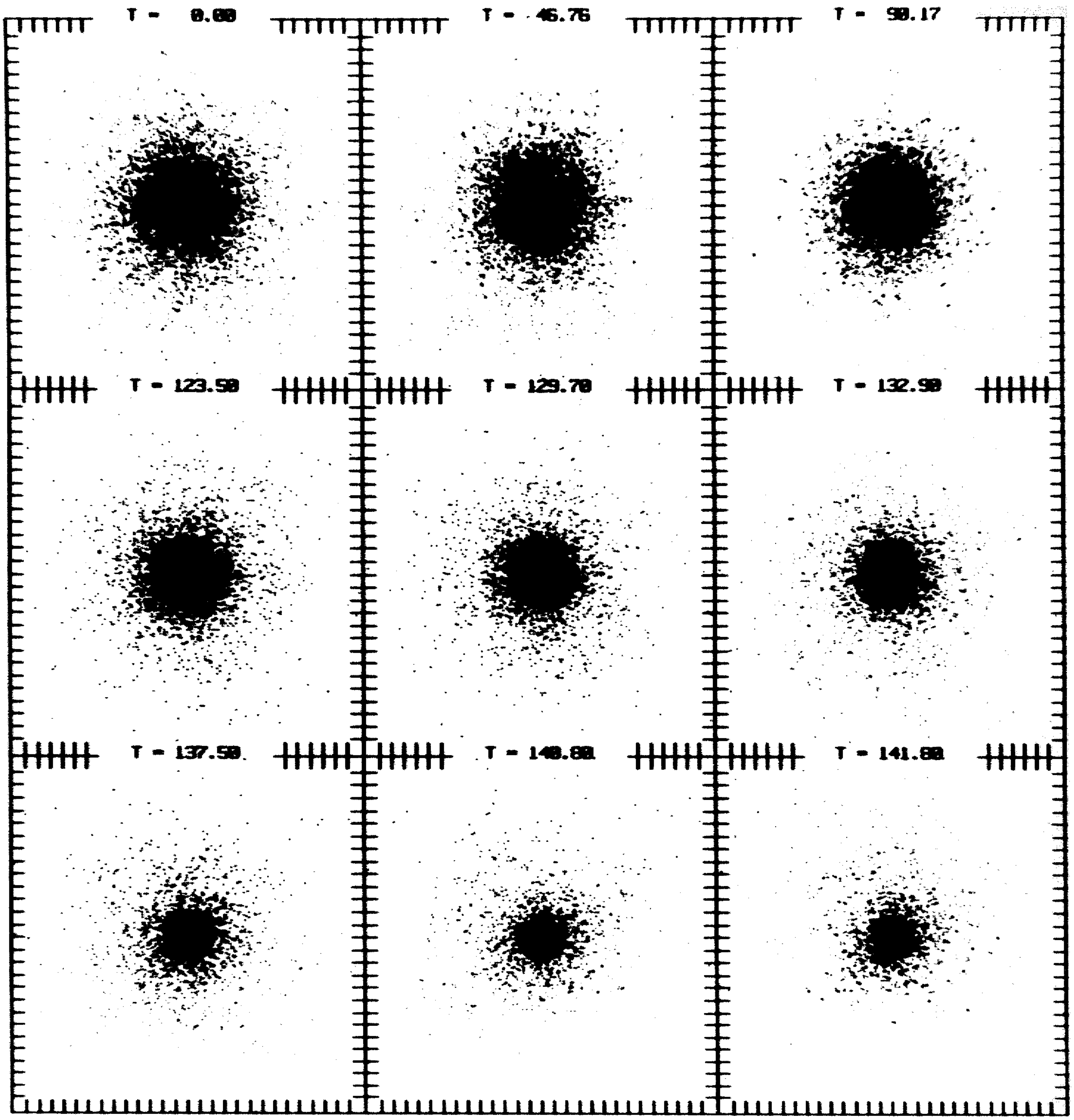

FIG. 1.—Fully relativistic numerical simulation of the gravitational collapse of an unstable relativistic star cluster initially in equilibrium. At $t = 0$ the cluster has radius $R/M = 25$, $z_c = 0.65$, and is very centrally condensed, with just 13% of the mass in the core. When the dynamical integrations were terminated, over 77% of the mass was inside the black hole event horizon. Snapshots of the collapsing configuration are shown at various times. Times and spatial gridpoints are in units of M, the arbitrary total mass of the cluster. For details, see Shapiro and Teukolsky (1985*b*).

SHAPIRO AND TEUKOLSKY (*see* page L42).