Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by
Diplom-Physiker Miguel Preto
born in Lisbon, Portugal
Oral examination: 17 July 2007

N-body studies of galactic nuclei with massive black holes

Referees: Prof. Dr. Rainer Spurzem
Prof. Dr. Ralf Klessen

Abstract

Massive black holes (MBHs) are ubiquituous in galactic nuclei. The collisional dynamics of these dense stellar systems is studied with high resolution, direct N-body simulations and its results are thoroughly compared with those obtained from the Fokker-Planck equation, written in energy space. The predictions from the Fokker-Planck equation are validated by the N-body results and it is concluded that the formation of a Bahcall-Wolf cusp around a MBH is a robust prediction for the old stellar populations of spheroidal systems older than the relaxation time as measured at the hole's influence radius. The mass segregation and the scaling relations predicted for stars in different mass ranges are also confirmed.

N-Body simulations of galaxy mergers show that binary MBHs can form with very high eccentricities after the parent galaxies settle into a newly merged nucleus. These high eccentricities should lead to a coalescence on a very rapid time scale, as the MBHs should emmit strong bursts of gravitational waves at each pericenter passage. The first harmonics of the gravitational wave strain amplitude h are estimated as the massive binary with mass $\sim 10^5 - 10^6 M_\odot$ enters the Laser Interferometer Space Antenna (LISA) frequency band.

We conclude that large-scale N-body modelling of galactic nuclei with one (several) MBH(s) is possible with the currently available software (Sverre Aarseth's NBODY codes and developments therefrom) and hardware (ARI GRAPE cluster).

Zusammenfassung

Massereiche Schwarze Löcher (MSL) sind in Galaktischen Zentren allgegenwärtig. Mit Hilfe hochauflösender, direkter N-Körper-Simulationen wird hier die stoßdominierte Dynamik solcher dichten stellaren Systeme untersucht. Die Ergebnisse werden mit Lösungen der im Energie-Raum formulierten Fokker-Planck Gleichung verglichen. Die Vorhersagen der Fokker-Planck Gleichung werden durch die N-Körper Ergebnisse bestätigt. Es wird gefolgert, dass für die alte stellare Komponente sphäroidaler Systeme, die älter als die Relaxationszeit sind (gemessen am Einflußradius des Schwarzen Lochs), die Entstehung eines

sogenannten *Bahcall-Wolf cusps* um MSL eine robuste Vorhersage ist. Die für Sterne unterschiedlicher Massenbereiche vorhergesagten Skalierungsbeziehungen und Massensegregation werden ebenfalls bestätigt.

N-Körper Simulationen von Galaxienverschmelzungen zeigen, dass, nachdem die beteiligten Galaxien einen neuen gemeinsamen Kern gebildet haben, binäre MSL mit sehr hohen Exzentrizitäten entstehen können. Hohe Exzentrizitäten sollten jedoch zur Verschmelzung der binären MSL auf extrem kurzen Zeitskalen führen, da die MSL bei jedem Durchlauf durch das Perizentrum erhöht Gravitationswellen abstrahlen sollten. Die ersten Harmonischen der Gravitationswellenamplitude h werden für massereiche Binärsysteme im Bereich $sim10^5 - 10^6 M_\odot$ berechnet, die in das Frequenzband der *Laser Interferometer Space Antenna* (LISA) fallen.

Wir folgern, dass die großräumige N-Körper Modellierung galaktischer Zentren, die ein Schwarzes Loch (oder mehrere) enthalten, mit der z.Zt. verfügbaren Software (z.B. Sverre Aarseth's NBODY Codes bzw. den darauf aufbauenden Weiterentwicklungen) und Hardware (z.B. ARI GRAPE cluster) möglich ist.

Dedication

Para a Madalena

Acknowledgments

First and foremost it is a great pleasure to thank my advisor Rainer Spurzem for all the great support and enthusiasm during the last few years. His help and guidance through the technical details of the N-body codes and his always pertinent hints to the relevant literature were, at many times, absolutely essential. In more than one sense, it is true that without his great generosity I would not be here writing these lines. Many thanks for everything, Rainer!

I would also like to thank very much the other members of my PhD thesis committee: Professors Immo Appenzeller, Ralf Klessen and Ulrich Platt.

All my colleagues at the ARI were very nice at all times. I thank all of them for their great simpathy and very humane atittude to life in general. To all of them: Peter Berczik, Ingo Berentzen, Andrea Borch, Jonathan Downing, Andreas Ernst, Jose Fiestas, Ovidiu Furdui, Patrick Glaschke, Chingis Omarov, Kristina Wäcken, many thanks for everything.

In particular, to Jose Fiestas with whom I share the office, I would like to remember here the many enjoyable conversations we had on many subjects ranging from our favorite scientific topics of galactic dynamics, black holes and the Fokker-Planck approximations, to many other things of life.

I would also want to specially thank Ingo Berentzen, with whom I also had extensive discussions on Post-Newtonian dynamics, for his very sound and sober atittude of approaching science and doing research.

I look forward to continue collaborating with you both.

I have also greatly enjoyed the wonderful atmosphere of Heidelberg, quite unique in the world in my opinion. It was also very nice to be able to row with friends from the Heidelberger Ruderklub, and I have enjoyed very much the Neckar valley as seen from the water.

Estou hoje perplexo, como quem pensou e achou e esqueceu.
Estou hoje dividido entre a lealdade que devo
À Tabacaria do outro lado da rua, como coisa real por fora,
E à sensação de que tudo é sonho, como coisa real por dentro.

Álvaro de Campos (Fernando Pessoa), *Tabacaria*, 1928

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Massive black holes in galaxies

Massive black holes (MBH) with masses spanning over a wide range ($10^4 - 10^9 M_\bullet$) are thought to be present in the centers of most galaxies. They reside in the center of galactic nuclei — the densest known stellar systems —, which reach densities well above $10^6 M_\bullet / pc^3$ and extend over radii that ranges from 1 to 10 pc. This is a very small fraction of the volume of a typical galactic bulge which has typically length scales of the order of a few Kpc.

Supermassive black holes (SMBH) are believed to be the engines of quasars and active galactic nuclei (AGNs). Nowadays, they are assumed to have played a critical role in the formation of the basic blocks of structure in the Universe — the galaxies — either by nucleating the proto-galactic overdensities or by promoting bursts of star formation. The quasar 3C273 was the first object to be identified as a "quasi-stellar radio source" (quasar) in the 1960s [Schmidt, 1963]. This object was initially thought to be a star-like object in the Virgo cluster; however, its redshifted lines were the tell-tale signal that it was instead an object at cosmological distances with redshift $z \sim 0.158$. Furthermore, its optical output evidenced a strong variability with a period of about 10 months — this, being the time available for radiation to cross the system's whole extension meant that its size could not be much larger than than ~ 0.25 pc. Then, in the 1970s orbiting X-ray telescopes showed that AGNs are luminous in X-rays, and their X-ray output was shown, in some cases, to vary by a factor of ~ 2 on a time scale of hours further constraining their maximum size to be not larger than ~ 30 A.U. Therefore, from the very beginning, these immensely energetic sources were recognized to be also very compact objects.

Quasars typically emmit more energy than the entire host galaxy but the source of emission occupies regions which are typically smaller than the Solar System. Salpeter (1964) and Zeldovich and Novikov (1967) described the growth of a very massive object, at the center of a galaxy, by accretion with the corresponding dissipation of energy and strong radiative emission. Lynden-Bell (1969) went on to explain the phenomenology of quasars by means of accretion onto a massive black hole. Nowadays, quasars are accepted to result from the accretion of matter into massive black holes in galactic centers; however, in the past, it has been unclear whether these objects really resided in galaxies as they can be much brighter than the host galaxy and that makes the latter's detection rather difficult. Today it is known

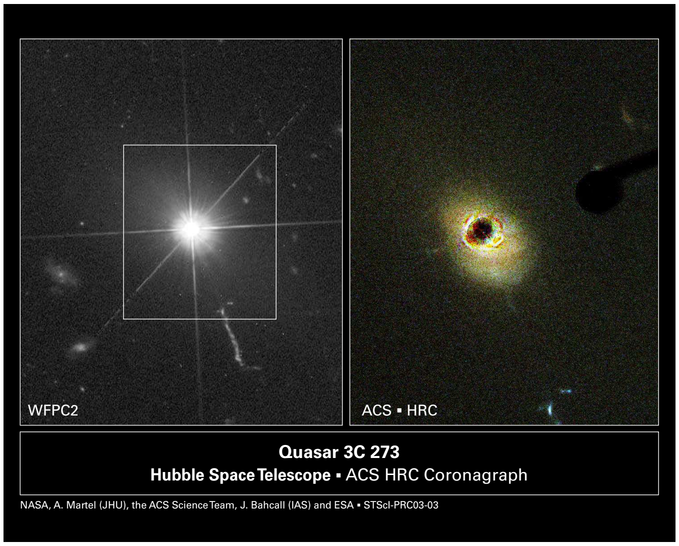

Figure 1.1: Quasar 3C 273 image taken with the NASA Hubble Space Telescope's Advanced Camera for Surveys (ACS). The ACS's coronograph was used to block the light from the quasar in order to reveal the host galaxy.

that the MBHs must have been formed very early in the history of the universe. The Sloan SKy Digital Survey (SDSS) discovered a number of quasars at redshift of $z \sim 6$ [Fan and *et al.* (SDSS Collaboration), 2001] which means that MBHs with masses $\gtrsim 10^9 M_\bullet$ were already formed when the universe was less than 10^9 years old. The SDSS also showed that the number of quasars reached a peak at redshift $z \sim 2$ — approximately the same epoch when the gaalxy mergers peak —, decreasing very sharply since then. The standard interpretation is that quasars are active whenever there is a lot of material to be accreted in the MBH vicinity, and turn off when most of it has been consumed. Galaxy mergers are naturally conducive to the efficient funnelling of gas into the galactic nucleus and therefore are arguably an important driver (albeit not the unique) for quasar activity [Kauffmann and Haehnelt, 2000]. The capture and disruption of stars by MBHs are thought to be insignificant from the energetic point of view in terms of acting as the driving force for quasar activity.

Summarizing: (i) AGNs are extremely compact sources, their flux variability observed on time scales shorter than a day (and flares on time scales of minutes) severely constrain their sizes; (ii) their spectral energy distribution is clearly non-stellar: their power per unit logarithmic frequency interval is roughly constant over seven decades, whereas stars have a much narrower frequency range; (iii) AGNs are very massive since their bolometric luminosity is equal or larger (sometimes by several orders of magnitude) than that of the host galaxy and remain active for time scales longer than 10^7 years. The accepted paradigm is that the source of this nuclear activity is matter accretion onto a central, massive black hole [Rees, 1984, Krolik, 1999]. In some AGNs, indirect evidence also supports the existence of a relativistic regime at the very center from the detection of superluminal motion of radio jets and the broadening of low excitation X-ray emission lines [Ferrarese and Ford, 2005].

More recent searches for MBHs have revealed many of these objects to reside in the centers of nearby, quiescent galaxies. The most compelling observational evidence comes from measurements of the kinematics of stars and gas in the central regions of nearby galaxies. Several of these MBHs in quiescent galaxies have masses $\sim 10^7 - 10^8 M_{\odot}$, which is below the typical quasar mass of $\sim 10^9 M_{\odot}$ (SgrA* in our galactic center is just $\sim 3.4 \times 10^6 M_{\odot}$). This means that several of these neighbor smaller MBHs have never been through a "quasar phase". It is, however, still somewhat of a puzzle why nearby MBHs accrete and radiate so little given the amount of matter (in the form of gas and dust) available at their disposal. Several models exist trying to explain this by means of radiatively-inneficient accretion processes [Blandford and Begelman, 1999, Narayan, 2003].

The searches for MBHs in nearby, quiescent galaxies were motivated by several reasons. On the one hand, weakly accreting MBHs were expected to live in the centers of quiescent galaxies because the cumulative MBH mass density derived from the observed energetics of high redshift quasars is too large, by at least two orders of magnitude, compared to the one inferred from the more local AGN activity [Ferrarese, 2002]. On the other hand, the methods available to detect the tell-tale Keplerian signature in the velocities of neighboring stars and gas can only resolve the central regions for the closest galaxies — even with HST. Finally, there no AGNs among the nearby galaxies.

Black holes are, by definition, dark objects and, if the surrounding regions are not active, their presence can only be inferred by the gravitational effects it imprints on the nearby material. Several techniques exist to detect the presence of a MBH and measure its mass. First, as in NGC4258, Miyoshi et al. (1995) observed a very thin disk of dense molecular material orbiting the center with speeds up to $\sim 10^3$ km/s. Water masers reside in this disk, which is fortuitously oriented with such an inclination that a beam of microwaves is directed along our line-of-sight. From the Doppler-shifted lines at ~ 0.3 pc from the center, the rotation curve is observed to be almost perfectly Keplerian and with the speeds that imply a central mass of about $(3-4) \times 10^7 M_{\odot}$ and central densities of $\sim 2.3 \times 10^9 M_{\odot}/\mathrm{pc}^3$. It is very unlikely that such high mass concentration could result from the presence of a stellar cluster, since it would be very difficult for it to survive these densities for long. In such an environment, the mean inter-stellar distance would be ~ 100 A.U., hence the cluster should rapidly undergo collapse and evaporation due to the very high frequency of stellar collisions and mergers.

The furtuitous arrangement as in NGC4258 does not occur very often, so other methods for the detection and measurement of MBH masses need to be used. For example, if in an AGN, the clouds of gas and dust that surrounds it are irradiated by the central engine, they produce spectral emission lines. The idea, in this case, is to follow the time evolution of the radiative emission of the central engine and, at the same time, that of the emission lines of the surrounding clouds. Then, short-term continuum variability in the output of the former (which is an ubiquitous feature of AGNs) should excite a recognizable feature in the response pattern of the emission lines of the latter. The time lag that the line emmiting regions take to respond to the changes in the continuum emission from the central source provides a measure for the distance between them. This technique is known as reverberation (or echo) mapping [Blandford and Mckee, 1982]. Upon simple assumptions regarding the geometry and the gas motion being predominantly gravitational, a total mass can be derived from the virial theorem $M = fr\sigma^2/G$ where r is the distance between the putative MBH and the surrounding gas measured by the reverberation technique, σ is the velocity dispersion of the clouds measured from the emission line's width; f is a fudge factor depending on the geometry and kinematics of the clouds. The advantages of this method are: (i) it is possible to measure MBH masses in distant and active galactic nuclei; (ii) the regions probed are $\sim 10^3 R_{Sch}$ which is a much smaller scale than that measured locally by kinematic methods; (iii) the implied densities $\gtrsim 10^{10} M_\odot/\mathrm{pc}^3$ mean that alternatives (such as a stellar cluster) to the general relativity singularity are much less likely as an explanation [Ferrarese and Ford, 2005].

In the case of nearby galaxies, the most effective methods for mass determination of the MBH are those based on the kinematic studies of stars orbiting closely around the center. In the Milky Way, SgrA* which is not especially massive or host for energetic phenomena, is nevertheless the closest MBH at only $\sim$ 8.5 Kpc. The Figure shows an infrared image of the Galactic center taken with the 8.2 meter VLT YEPUN telescope at the ESO, Paranal, Chile. The image is very sharp — as a result of the use of adaptive optics to compensate for the turbulent motions of the atmosphere —, and therefore it is possible to identify individual stars orbiting at $\sim 0.005 - 0.01$ pc from the center [Genzel et al., 2003a, Genzel et al., 2003b, Ghez et al., 2005]. The proper motion of these central stars can be monitored over several years. One of the fainter stars in the Figure — named S2 — is located precisely above the inferred position of SgrA*, in this image taken at a time it was at pericenter. This star has been tracked for over 10 years and its orbit traces a Keplerian ellipse with one focus at the inferred SgrA* position and has an orbital period of $\sim$ 15 years . From the motion of all the stars, the measured mass of SgrA* is $\sim 3.4 \times 10^6 M_\odot$ enclosed inside a region $\lesssim 6 \times 10^{-4}$ pc. Therefore SgrA* and the "maser" MBH in NGC4258 are considered to be the most precisely measured MBH masses.

All these MBH detections are still short of a direct observation of a relativistic signature of a black hole — such as the event horizon, frame-dragging, gravitational redshift [Alexander, 2005], or of gravitational waves emmited by the inspiralling and relativistic coalescence of

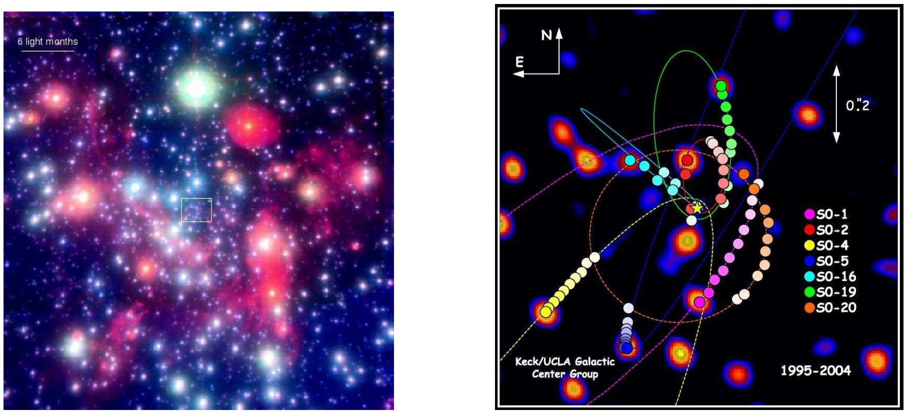

Figure 1.2: Left image: A view of the stars surrounding the MBH at the Galactic center. The MBH is located at the center and the scale is $\sim 0.6 \times 0.6$ pc. Right image: The orbit of stars within the central $\sim 0.04 \times 0.04$ pc of the Galactic center. These orbits constitute the best evidence to date of the presence of a MBH with a mass of $3.4 \times 10^6 M_\odot$ in the center of our galaxy.

compact remnants onto the central MBH [Alexander, 2005, Hopman and Alexander, 2006, Amaro-Seoane et al., 2006]. However, the larger the mass-to-light ratio and density implied by the observations, the more unlikely is that they can be explained without the presence of a MBH. Estimates of a dark cluster (made of brown dwarfs, white dwarfs, neutron stars and stellar BHS) lifetimes against evaporation and collapse are shown in Figure 1.3. The estimated maximum lifetimes attainable by a dark cluster in such an environment is plotted against the inferred density and maximum size allowed by the observational data's resolution scale. According to these estimates, it is still only in the cases of Milky Way's SgrA*, NGC 4258 and Circinus that the dark cluster estimated lifetime is clearly shorter than the estimated age of the galaxy. In all the other cases, this simple argument cannot exclude the dark cluster alternative interpretation for the data. In the case of SgrA*, due to its close proximity, it is possible to constrain its proper motion with remarkable accuracy: measurements, over eight years, with the Very Long Base Array (VLBA) of SgrA*'s position with respect to two extra-galactic radio sources indicates that SgrA* is indeed very close to being stationary (to within ~ 1 Km/s) at the galactic center making the case for a MBH almost inevitable [Reid and Brunthaler, 2004].

In the context of the hierarchical theories of large structure formation in the universe, two outstanding questions are: (i) how far back in the dark halo merger hierarchy do MBHs form; (ii) which type of seeds give rise to the MBH formation — *light* seeds of $\sim 10^2 M_\odot$ remnants of Population III stars at redshifts $z \sim 20$ that form from rare $3.5-\sigma$ peaks of the primordial density fluctuations [Madau and Rees, 2001, Volonteri et al., 2003a]; or *heavy* seeds $\gtrsim 10^4 M_\odot$ resulting from a direct, monolithic collapse [Koushiappas et al., 2004, Begelman et al., 2006].

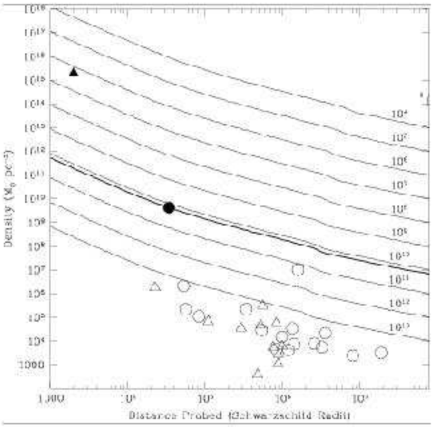

Figure 1.3: The inferred mass density at the innermost resolved radius plotted against the same radius (in units of the Schwarschild radius). The solid curves show the estimated lifetimes of a dark cluster, as explained in the text. The dark triangle denotes the Milky Way measurement; solid circle denotes NGC 4258; open circles correspond to detections made with stellar and open triangles with gas kinematics. The thick solid line corresponds to a lifetime of 1.5×10^9 yrs. From [Maoz, 1998].

These two scenarios have different implications for the detection estimates of gravitational waves by LISA from the inspiralling and merging of binary MBHs [Sesana et al., 2007].

There are also a number of phenomenological scaling relations between the MBH and the large scale galactic bulge where it resides. Kormendy & Richstone (1995), using the eight MBH detections known at the time, realized that there is a correlation between the inferred MBH mass and the blue luminosity of the surrounding bulge — or of the entire galaxy in the case of ellipticals. This could suggest that they are coeval. More recent observations have further substantiated the $M_\bullet - L_B$ relation. In the left panel of the Figure 1.4, the $M_\bullet - L_B$ relation is shown for MBH detections for which the hole's sphere of influence could be resolved by the observations. There is some scatter in the relation, but this was shown to be reduced when the K-band $M_\bullet - L_B$ relation is considered — independently of galaxy type [Marconi and Hunt, 2003]. This is not surprising since near-IR magnitudes are presumably better tracers of mass than B-band magnitudes. On general terms, the scatter in phenomenological relations involving the MBH mass tend to decrease for samples which include only observations that were able to resolve its sphere of influence. This also is not surprising as it is precisely in this innermost region that the MBH potential well dominates the kinematic properties of the stellar system.

The $M_\bullet - \sigma$ relation [Gebhardt et al., 2000, Ferrarese and Merritt, 2000] relates the MBH mass with the large scale bulge's velocity dispersion. This is a very tight correlation — tighter than the Faber-Jackson relation for ellipticals. A consequence of this phenomenological scaling relation is that the MBH mass can be inferred with remarkable accuracy ($\sim 30\%$)

by a single measurement of the velocity dispersion of bulge stars orbiting well outside its influence radius — i.e. very far from the region where the MBH potential well strongly influences the stellar motions at present time. This tight relation may therefore be a hint to a common formation history of the MBH and the spheroidal within which it sits.

The $M_\bullet - \sigma$ relation is now thought to be so fundamental that studies of MBH demographics — both in quiescent [Merritt and Ferrarese, 2001, Yu and Tremaine, 2002, Wyithe and Loeb, 2003, Volonteri et al., 2003b] and in active galaxies — rely strongly on this relation. Models of MBH formation and evolution use the $M_\bullet - \sigma$ relation as a means of calibrating their calculations which must be able to reproduce its slope, normalization and, very importantly, its small scatter while at the same time undergoing the violent merger events shaping galaxies during their evolution.

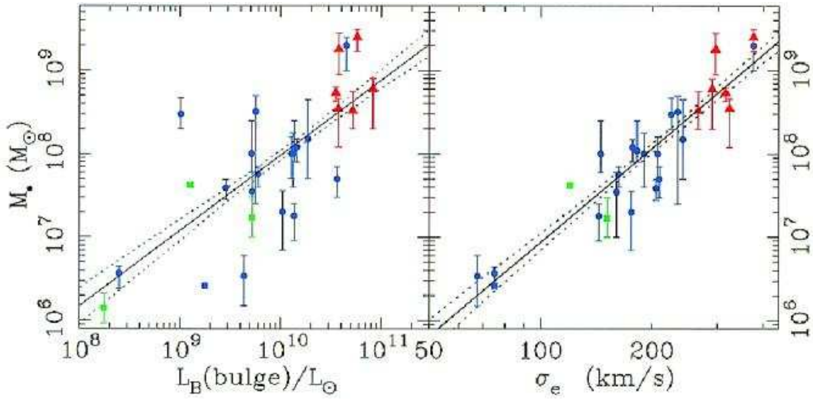

Figure 1.4: Left panel: The $M_\bullet - L_B$ relation. Right panel: The $M_\bullet - \sigma$ relation. (source Gebhardt et al. ApJ 2000)

During the last decade, there has been huge advances in the field with many new observation unveiling the structural relations between the MBH and the surrounding spheroidal stellar systems where they reside. A lot of new phenomenology, with very interesting implications, did percolate to the wider field of galaxy formation and evolution, but this has been only the beginning. Most of these new results were made possible only with the use of the HST which permitted unprecedented resolution power (increased by a factor of ~ 10) to the kinematic study of regions inside the influence radius of the nearby MBHs. HST also propelled the detection of nuclear gas and dust disks which serve as a probe of the MBH potential well. However, it is only for the case of NGC4258 and of our galaxy that an observationally water-tight case can be made for the relativistic nature of the very massive central dark concentration. Moreover, the ultimate direct detection of a relativistic signature — such as the detection of an event horizon or of gravitational waves — is still left for the future. The planned ESA/NASA mission Laser Interferometer Space Antenna (LISA) promises to bring new fundamental contributions to the advancement of this fascinating field of galactic astronomy.

1.2 Detection of gravitational waves

The detection of gravitational waves emmited during the inspiral, plunge and relativistic coalescence of either (i) a binary MBH system; or, (ii) an extreme mass ratio inspiral (EMRI) of a compact remnant (such as a neutron star or a stellar BH) onto a MBH are the most likely gravitational wave sources from which we may expect a direct and unambiguous confirmation about the relativistic singular nature of the central mass concentrations known to be ubiquituous in galactic centers. These sources will emmit gravitational waves (GWs) in the frequency range ($10^{-5} - 1$ Hz) to which the proposed ESA/NASA *Laser Interferometer Space Antenna* (LISA) will be sensitive enough to record. LISA will be probably the most exciting and successful development capable to provide a novel impetus to this field of Galactic Astronomy. In the meantime, other detectors tuned to different frequency ranges and with different sensitivities are being planned, and some are already working in the task of measuring GWs from several type of putative sources.

The properties of gravitational radiation are different from those of electromagnetic radiation — the traditional means with which to perform astrophysical observations. First, electromagnetic waves (EMWs) interact very strongly with ordinary matter (they are absorbed, scattered, and so on) whereas GWs don't. This means that GWs propagate for cosmological distances essentially without any distortion making it possible to probe both very distant and very dense regions which are otherwise inaccessible to us (eg, coalescence and merger of compact remnants and of MBHs, core collapse of stars, etc.). This also means, of course, that detecting the GWs will also be a highly non-trivial task.

EM radiation has typically a wavelength smaller than the size of the emmiting source as is results from microscopic processes (eg. atomic transitions, motion of charges in an astrophysical plasma, etc). On the contrary, the GWs result from the (macroscopic and asymmetric) bulk motions of its very massive source (e.g. the orbital motion of binary members, oscillations of the asymmetric surface of a neutron star, etc), so their wavelenths are typically larger than the size of its source. GWs cannot be used as a means to form an image of the source and to locate it precisely in the sky.

Due to the stochasticity and microscopic nature of the generation of radiation, EMWs are generally phase-incoherent. The GWs, in contrast, are usually phase-coherent similarly to laser light. This crucial property of GWs makes it possible to detect and decode the waveforms through matched filtering techniques so long as reliable models for the GW's functional form for interesting source types are available with the sufficient precision.

Another important consequence of phase coherence is that the observable used in GW detection is its strain amplitude h, which is a quantity that falls off with the inverse of the distance. In most cases, EM radiation observations result from the measurement of a flux and therefore fall off much more rapidly with the inverse square of distance. Measuring GWs is analogous to measuring a coherent $1/r$ EM radiation field. Another nice consequence: an increase of the detector's sensitivity by a factor of two results in an about ten-fold increase in the observable volume (and thus, on average, of the expected number of sources).

GW detectors have a nearly 4π steradian sensitivity over the sky, albeit with the price of having a poor performance in the localization of the sources in the sky.

Compact sources typically emmit GWs in a given frequency band which is set by its size R (stellar radius, binary separation, and so on) and its total mass M; the natural frequency is thus given by

$$f_{GW} = \frac{1}{2\pi}\left(\frac{GM}{R^3}\right)^{1/3} \lesssim \frac{1}{4\sqrt{2\pi}}\frac{c^3}{GM} \approx 10^4\mathrm{Hz}\left(\frac{\mathrm{M}_\odot}{\mathrm{M}}\right). \tag{1.1}$$

The inequality results from the fact that the source cannot be smaller than the Schwarschild radius $2GM/c^2$ of a mass M. Therefore very massive objects such as a MBH, or less massive binaries with very large semimajor axis, will fall on a lower frequency range $10^{-5}\mathrm{Hz} \lesssim \mathrm{f} \lesssim 1$ Hz, whereas compact objects of stellar mass will fall on a higher frequency interval, $f \gtrsim 10^2$ Hz.

In the high frequency band ($1\mathrm{Hz} \lesssim \mathrm{f} \lesssim 10^4$) Hz), there are now several ground-based interferometric detectors in function.

A1. VIRGO is a 3-Km French-Italian ground-based detector with a very sophisticated seismic isolation system to provide very good isolation from low frequency ground oscillation and thus reach very high sensitivity at low frequencies. The low end of its frequency band is set by the mechanical coupling of the detector to ground vibrations, atmospheric motions, human activity and so on; the high end is set by the fact that GWs with $f > 10^4$ Hz are not expected unless there were good reasons to believe in the existence of extremely compact sources with low mass ($\lesssim 1M_\odot$).

A2. LIGO is an American ground-based detector with frequency band and sensitivity curve very similar to those of VIRGO.

A3. GEO600 is a German-English collaboration located near Hannover, Germany.

These ground-based detectors were designed having the following GW sources in mind: (i) coalescing compact remnants; (ii) stellar core collapse (large masses: $\sim 1 - 100M_\odot$ flows in a very compact region at relativistic speeds $v \gtrsim c/5$; however, the degree of asymmetry of the star during collapse is crucial for the wave amplitude, as a perfectly spherical collapse does not generate any GWs, the quadrupole of the mass distribution being the lowest order to contribute to the generation of GWs); periodis sources such as non-axisymmetric rotating neutron stars; stochastic background due to the superposition of very many uncorrelated and unresolved signals (some people speculate about a stochastic wave background generated in the early Universe through a variety of processes such as the amplification of matter fluctuations via inflation, phase transitions as previously unified fundamental interactions separate, vibrating cosmic strings, condensations of branes from higher dimensional spaces, and so on).

In the low-frequency regime, $10^{-5}\mathrm{Hz} \lesssim \mathrm{f} \lesssim 1$ Hz, it is not possible to use ground-based detectors since it is practicably impossible to isolate them from local seismic and ground vibrations of low frequency, and also from the gravitational coupling to fluctuations in the local mass distribution. Also the wavelengths of interest would become too large to detect for any arm length conceivable in a terrestrial detector. Therefore it is necessary to build a detector to operate in the quiet environment of outer space, well removed from the proximity of sources of low frequency noise, in order to reach the required sensitivities for frequencies $f \lesssim 1$ Hz and with an arm length much larger than that of its terrestrial counterparts.

The European Space Agency (ESA) and the North American Space Agency (NASA) are currently designing such a mission: the Laser Interferometer Space Antenna (LISA). The basic concept is the same as that of the ground-based detectors: the distance between widely separated mass probes is constantly monitored for small oscillations due to the passage of GWs. Its arm length will be much larger, $L \sim 5 \times 10^6$ Km, though. Three spacecrafts, where the mass probes will be located, will be placed in orbit such that LISA will form an equilateral triangular constellation, with a 60^o inclination with respect to the plane of the ecliptic, and trailing the Earth's position by 20^o. This orbital motion will also modulate the GW waveforms and will thus make it possible to determine the source's location in the sky.

The coalescence of MBHs binaries will be detectable by LISA out to extrememly large distances, essentially out to redshifts $z \sim 5 - 10$ depending on the mass range considered. This volume is so large that even a low rate of MBH coalescences does not exclude a LISA detection. Another class of priority LISA detections is the extreme mass ratio inspirals (EMRIs) of compact remnants to MBH with typical masses as that of SgrA* $M_\bullet \lesssim \mathrm{few} \times 10^6 \mathrm{M}_\odot$. These EMRIs will be detectable up to a distance of a few Gpc in case of an inspiral of a stellar BH, or few hundreds of Mpc in the case of a neutron star or white dwarf inspiral. LISA will be able to measure the last few years of these EMRIs and will thus be able to probe the spacetime surrounding a MBH in the strong-field regime in order to very if these very massive objects living in galactic centers have indeed the properties associated to Kerr black holes.

1.3 Dynamical modelling of dense stellar systems

From the point of view of a theoretical astrophysicist, it is crucial to model accurately the dynamics of galactic nuclei in order to gain a solid insight about the processes involved in the evolution of the expected gravitational wave sources. The first essential step is to characterize the (quasi-)equilibrium stellar distribution that constitutes the *backbone* large-scale structure of the stellar system. As galactic nuclei are thought to be dominated dynamically by the stellar component (plus the MBH inside its influence radius), the first order model should avoid unnecessary extra-complications associated with gas, star formation, magnetic fields,

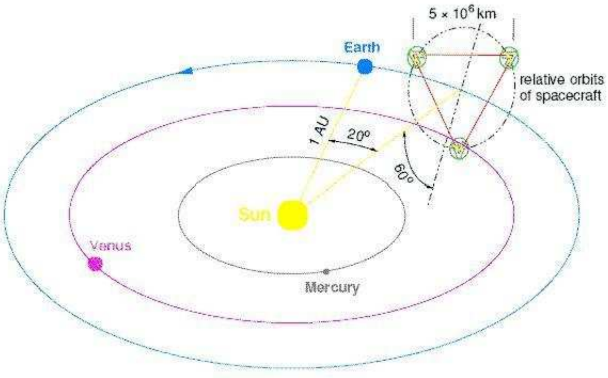

Figure 1.5: Orbital configuration of the Laser Interferometer Space Antenna, LISA.

and so on. This is indeed a very bold working hypothesis, but it is also a powerful and fruitful one in the sense that leads to a very well-defined dynamical problem for which we have, at last, the adequate tools to solve. The galactic nucleus will thus be treated to first order as a gravitational $N-$body problem with very large N ($\sim 10^{7-8}$).

The two-body problem is integrable, was solved long ago by Newton and was for a long time a prototypical example of the models in Physics: a simple, analytic solution of a regular motion. The three-body problem, however, resisted to all attempts at solving it and it was shown by Poncaré in the 19^{th} century that it is not, in general, integrable. On the other limit of large N, self-gravitating systems, when one takes the $N \to \infty$ thermodynamic limit are quite peculiar due to the long-range nature of the gravitational forces, as there are no negative charges to provide for a *mass neutrality* analogous to the charge neutrality on scales larger than the Debye length in plasmas. One consequence, among others, is that self-gravitating systems have a negative heat capacity, which has numerous implications for their evolution — from the nuclear burning in the stellar cores to the core collapse of globular systems. For all these reasons and others, it is well-known that the N-body problem is quite intractable from an analytic point of view, except in a few idealized configurations to a semi-analytic approximation [Lynden-Bell and Wood, 1968, Lynden-Bell and Eggleton, 1980, Goodman, 1984, Goodman, 1987, Weinberg, 1994, Tremaine, 2005] or a purely formal one [Gilbert, 1968], so it is absolutely essential to adopt a numerical approach.

A great deal of effort was devoted over the last four decades to the study of dense stellar systems — in particular, those of globular clusters. For the most part, these studies approached this problem with techniques borrowed from the fields of kinetic theory and plasma physics (it was certainly not a coincidence that Lyman Spitzer was a pioneer in both fields). The formalism of kinetic theory, e.g. the Fokker-Planck equation [Chandrasekar, 1943, Spitzer, 1987, van Kampen, 1992], was applied and adapted to the peculiarities of a system with long-range forces without any shielding effects. Several different methods were developed to compute the solutions of the Fokker-Planck equation: gaseous methods

[Bettwieser and Sugimoto, 1984, Louis and Spurzem, 1991, Amaro-Seoane et al., 2004], direct methods for resolution of the partial differential equation [Cohn, 1979, Cohn et al., 1989, Einsel and Spurzem, 1999, Fiestas et al., 2006], and Monte-Carlo methods [Spitzer and Hart, 1971, Hénon, 1975, Spitzer, 1987, Freitag and Benz, 2002].

All these methods are approximate and incorporate all the assumptions inherent to the Fokker-Planck formalism — neglecting large-angle deflections —, and include the microscopic inter-particle interactions only in an appproximative manner. The only exact way — modulo numerical innacuracies — to compute the solution to the N-body problem is to resort to the direct integration of the Newtonian equations of motion

$$\ddot{\mathbf{r}}_i = -\sum_{j\neq i} G m_j \frac{(\mathbf{r}_j - \mathbf{r}_i)}{|\mathbf{r}_i - \mathbf{r}_j|^3}. \tag{1.2}$$

Naively this seems to be quite straightforward to implement as it is easy to program an integrator for solving these equations of motion. There are, however, many obstacles some of them quite subtle. The galactic nuclei that we want to model have typically a large number of stars, so we need to employ as many particles as necessary to make a one-to-one star simulation. This is simply impossible for the following reasons: (i) to advance a single time step, it is necessary to compute the interactions between all particles which corresponds to a number of operations of the order $\mathcal{O}(N^2)$; (ii) to integrate the system over a relaxation time ($t_{\rm rlx} \propto N$) means that the total number of force computations is of order $\mathcal{O}(N^3)$. This constitutes a heavy computational burden and one is, therefore, forced to make use of special-purpose hardware to accelerate the force computations (the most expensive part of the direct N-body algorithms), parallel-distributed hardware or a combination of both (the latter option was, unfortunately, not available at the beginning of this work). In this work, the decision made was to use the special-purpose hardware GRAPE-6 [Makino and Taiji, 1998, Makino et al., 2003, Fukushige et al., 2005] coupled to a single processor of a standard PC or workstation (host computer) to perform simulations up to $N = 128K$ particles over a relaxation time scale. The GRAPE-6 hardware has specialized pipelines for gravitational force calculations — the most expensive part of a direct, N-body algorithm —, its use speeds up the computation by factors of a few hundreds. In this configuration, the GRAPE-6 hardware computes the pairwise accelerations between all particles in the simulation while the host computer advances them every time using the Hermite integration scheme.

It is necessary to take special care in the integration of compact configurations of stars (or stars plus MBHs) — such as binaries, triples, and so on. These configurations exarcebate the numerical difficulties associated with the singular nature of the Keplerian potential. In our particular problem, there is the extra-factor associated with the large mass ratios between the massive particles employed to model the MBHs and the field particles/stars. A number of sophisticated regularization algorithms (two-body KS regularization, chain regularization)

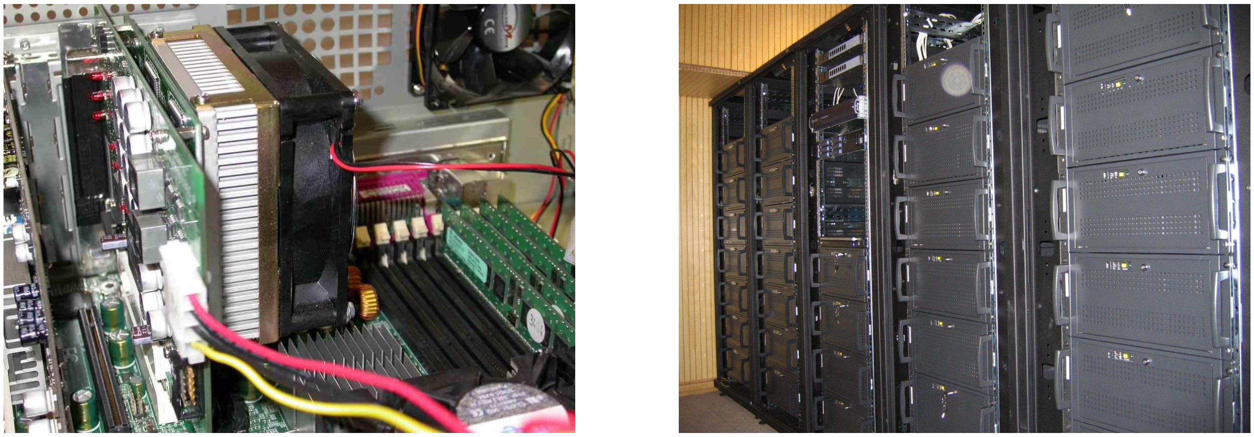

Figure 1.6: Left image: A micro-GRAPE card coupled to a standard PC. Right image: The ARI GRAPE cluster. The ideal computer for the N-body modelling of galactic nuclei and massive black hole(s).

are heavily employed and are seriously challenged by this problem. Fortunately, they were able to deal quite efficiently with these situations and should be able to perform quite well unless the mass ratios rise well above 10^6 [Mikkola, 2006].

A more fundamental and general challenge could be raised concerning the validity of the whole enterprise of numerical simulations of large-scale self-gravitating N-body systems: the exponential orbit instability. This is, in fact, the question of sensitivity to initial conditions in the N-body problem: loosely speaking, if two orbits start with an infinitesimal separation ϵ between them in phase space, how does this phase space separation evolve with time? Will it be kept bounded, will it diverge? It is immediately evident that if they are not kept bounded within a small separation, the reliability of the integration of the *individual* orbits becomes questionable [Miller, 1964]. It can be shown that the e-folding time for the separation of orbits is $\sim 0.7/\ln(1.1\ln(N))$, with a log of a log appearing quite unexpectedly and making an already slow (log-)convergence become excrutiatingly slower [Goodman et al., 1993]. This exponential orbit instability is inescapable as it associated with the intrinsic dynamics, rather than with any systematic integration errors; the truncation error associated with the finite precision of any machine is enough for, even with an exact integration scheme, the loss of accuracy in the integration of *generic, individual* orbits after an integration time no longer than a few dynamical times. Recently this estimate for the e-folding time of the orbit divergence and its *log of a log* nature was confirmed by direct, numerical integration tests [Hemsendorf and Merritt, 2002]. But all is not lost, though: extensive numerical studies have consistently shown that the global quantities of interest for the characterization of the stellar system — the density profile $\rho(r)$, the velocity dispersions $\sigma(r)$, the distribution functions $f(E, J)$, Lagrangian radii and so on — can be systematically reproduced from simulations with initial conditions corresponding to different Monte Carlo realizations from the same cluster model. As an aside, it should be mentioned that the issue of the exponential orbit stability is well known in the field of nonlinear dynamics where people often invoke, for hyperbolic systems, the so-called *shadowing* lemma [Guckenheimer and Holmes, 1983]. This lemma states, loosely speaking, that even if a numerical integration does not, because of the sensitivity to initial conditions, compute the correct orbit corresponding to the chosen initial conditions, it will nevertheless follow closely some other true orbit of the system. Self-gravitating N-body systems are not hyperbolic, alas!, but there has been some work in this direction with simplified models inspired specifically by galactic dynamics [Quinlan and Tremaine, 1992].

1.4 Goals of this work

In this study, our goal is to set the foundations for the N-body modelling of galactic nuclei with massive black holes (MBHs), including the relativistic binary dynamics using the Post-Newtonian equations of motion.

Many studies of the stellar distribution around a central MBH have been done [Bahcall and Wolf, 1976, Bahcall and Wolf, 1977, Lightman and Shapiro, 1977, Cohn and Kulsrud, 1978, Marchant and Shapiro, 1980, Murphy et al., 1991]. In these works, the basic equilibrium properties of a relaxed stellar system around a MBH were extensively studied, but almost all have used the Fokker-Planck formalism with all its inherent assumptions. The computational effort required for the N-body study of the growth of a Bahcall-Wolf cusp around a MBH was too extreme and, moreover, the important question regarding the validation of the approximations made these classical papers remained to be addressed. It was only in recent years, with the advent of the special-purpose computer GRAPE-6, that such a problem could be studied in detail. These N-body studies are not limited by the usual assumptions such as: (i) The stellar system evolves under the driving influence of small-angle scatterings due to uncorrelated, two-body encounters; (ii) The orbit-averaged diffusion coefficients computed under the local approximation and on the assumption of spatial homogeneity. In contrast, using the direct N-body approach one can follow the full dynamics of the stellar system without any assumptions other than the validity of Newton's equations of motion. As a result, in this work it was possible to validate the predictions of the Fokker-Planck equation for stellar clusters around a massive black hole. We have shown that the growth of a Bahcall-Wolf cusp is a robust prediction for the old stellar populations of stellar systems whose relaxation time is less than a Hubble time.

We also study the astrophysically important process of mass segregation and monitor the progressive concentration of the heavier masses in the center of the system. This dynamical process is crucially important for all estimates of close interactions between field stars and the central, MBH. A careful determination of the steady-state radial distribution of the different stellar components is a necessary step in order to later be able to estimate capture rates by the MBH — either by tidal disruptions of main sequence stars, prompt infall of compact remnants, or of slow relativistic inspirals onto the MBH. We have shown that these steady-state distributions of stars can be established with direct $N-$body simulations, given the presently available hardware and software.

In this work, we also have studied the numerical formulation of the Post-Newtonian (PN) equations of motion for the integration of relativistic binary systems. The PN equations in the center of mass frame were implemented with the two-body KS regularization method. This implementation was thoroughly tested and shown to be capable of handling the relativistic inspiral motion until the plunge phase. We have also performed N-body simulations of galaxy mergers, each with a MBH at its center, that are put on an approximately parabolic encounter orbit. We have shown that the MBH binary can be formed, very soon after the parent galaxies have settled into a new larger galaxy, with extremely large eccentricities up to ~ 0.99 or even ~ 0.999. This will lead to a very rapid merger of the MBH binary as the strong emission of gravitational waves, at each pericenter passage, will very rapidly sap the orbital energy of the pair. Other set of N-body calculations also show that very eccentric MBH binaries can be formed, if the MBH components are put initially at symmetric positions about the center and given initial velocities that would lead in isolation to an eccentric Keplerian orbit.

Finally, we estimate the gravitational wave strain amplitudes h that LISA would measure in a three year observation run if these MBH mergers are scaled to represent $\sim 10^5 - 10^6 M_\odot$ MBH binaries.

Chapter 2

Numerical methods

2.1 Hermite Scheme with Block Time Steps

The most central aims of this work is the study of the collisional dynamics of galactic nuclei in the presence of a massive black hole. In order to simulate a set of up to a few hundred thousand particles over a relaxation time scale; observe the collisional growth of a stellar cusp around the central, massive particle it is necessary to follow the evolution of system for several hundreds of dynamical times, as defined at the hole's sphere of influence; very long integrations are also inevitable if we want to follow the inspiral of a massive black hole binary under the gravitational scattering with field stars. The accuracy requirements for such long integrations are very demanding; global quantities such as $\Delta E/E$ should be conserved with a precision of, at least, $\sim 10^{-5}$ per dynamical time. Furthermore, the relaxation process in stellar dynamics results from a series of a very large number of uncorrelated, weak two-body encounters; therefore it is necessary to integrate directly (and accurately) the pairwise interactions between all pairs of particles in the simulation. Hence a direct N-body code is the only possible option available.

The direct $N-$body codes developed by Aarseth over the last fourty years [Aarseth, 1999, Aarseth, 2003] have become the standard in the field of stellar dynamics; its most successful applications have been achieved in the study of globular clusters without any massive black holes, but also in planetary dynamics. In particular, in these stellar clusters's applications the largest mass ratio between particles is typically rather modest ($\lesssim 10-20$); the introduction of a massive particle as a means to model the dynamical effect of the presence of a massive black hole in the stellar cluster is therefore a step into a new dynamical regime that has been almost unexplored before, with only a few notable exceptions [Baumgardt et al., 2004a, Baumgardt et al., 2004b]. Aarseth's codes are also notable in the fact that the exact, singular, form of the Newton's gravitational potential is used without recourse to any softening, and this is done for arbitrarily strong interactions. In this respect, the currently main algorithmic limitations is related to the treatment of the gravitational interaction between particles with very large mass ratios.

The direct N-body integrator for most particles is the standard fourth-order Hermite predictor-corrector scheme with individual and adaptive, block time steps. The time step levels are *quantized*, ie. they are forced to take only values that are powers of two of some basic reference step in order to guarantee the synchronization of all particles at the times when it is necessary to compute the accelerations from all particles in the simulation. See Figure (2.3) for a diagram of the block timestep scheme. This method has also the advantage of being a one-step and self-starting scheme, ie. it is sufficient to specify the positions and velocities of all particles at the initial time to start the integration.

Let us suppose there is a set of N particles with positions $\mathbf{r}_i(t)$ and velocities $\mathbf{v}_i(t)$, where $i = 1, ..., N$, at some time $t = t_0$. The accelarations $\mathbf{a}_i(t)$ and their first time derivative $\dot{\mathbf{a}}_i(t)$ are given by

$$\begin{aligned} \dot{\mathbf{a}}_i(t) &= \sum_{i=1}^{N} Gm_j \frac{\mathbf{r}_{ij}}{r_{ij}^3} \\ \dot{\mathbf{a}}_i(t) &= \sum_{i=1}^{N} Gm_j \left[\frac{\mathbf{v}_{ij}}{v_{ij}^3} + \frac{3(\mathbf{v}_{ij} \cdot \mathbf{r}_{ij})\mathbf{r}_i(t)}{r_{ij}^5} \right], \end{aligned} \tag{2.1}$$

where $\mathbf{r}_{ij} = \mathbf{r_j} - \mathbf{r_i}$, $\mathbf{v}_{ij} = \mathbf{v_j} - \mathbf{v_i}$, $r_{ij} = |\mathbf{r_j} - \mathbf{r_i}|$, and the sums are taken over all particles alike with the exception of those that are selected for special treatment involving any of the regularized schemes. The accelerations and the jerks shown in equations 2.1 are computed at the end of each time step with the help of the special-purpose hardware $GRAPE - 6$ which permits to speed up the computation by factors of several hundred.

The basic idea of the Hermite predictor-corrector integration scheme is very simple. The acceleration and its first derivative are expanded in a Taylor series as follows

$$\begin{aligned} \mathbf{a}_i(t) &= \mathbf{a}_i(t_0) + \mathbf{a}_i^{(1)}(t_0)\Delta t + \frac{1}{2}\mathbf{a}_i^{(2)}(t_0)\Delta t^2 + \frac{1}{6}\mathbf{a}_i^{(3)}(t_0)\Delta t^3 \\ \dot{\mathbf{a}}_i(t) &= \mathbf{a}_i^{(1)}(t_0) + \mathbf{a}_i^{(2)}(t_0)\Delta t + \frac{1}{2}\mathbf{a}_i^{(3)}(t_0)\Delta t^2. \end{aligned} \tag{2.2}$$

Given the values of $\mathbf{a}_i(t)$ and $\dot{\mathbf{a}}_i(t)$ at the beginning and the end of the time step, the second and third derivatives of the acceleration are given by

$$\begin{aligned}
\mathbf{a}_i^{(2)}(t) &= \frac{-3(\mathbf{a}_i(t_0) - \mathbf{a}_i(t)) - (2\mathbf{a}_i^{(1)}(t_0) + \mathbf{a}_i^{(1)}(t))\Delta t}{\Delta t^2} \\
\mathbf{a}_i^{(3)}(t) &= \frac{12(\mathbf{a}_i(t_0) - \mathbf{a}_i(t)) + 6(\mathbf{a}_i^{(1)}(t_0) + \mathbf{a}_i^{(1)}(t))\Delta t}{\Delta t^3}.
\end{aligned} \tag{2.3}$$

This is the Hermite interpolation: it interpolates the orbits of the particles with a 5^{th}-order polynomial for the positions and a 4^{th}-order polynomial for the velocities. This, in turn, requires the knowedge of the 2^{nd} and 3^{rd} derivatives of the accelerations, and hence can be obtained once the value of the acceleration and of its 1^{st} derivative are known at two consecutive instants of time — namely, at the beginning and at the end of each time step.

The integration then proceeds as follows:

1. The particles are sorted according to their new time at the end of next step;

2. Those particles which will have the smaller new time belong to the group to be advanced in the next time step;

3. The predicted positions and velocities for the particles in the next group, obtained directly from the GRAPE-6, are given by

$$\begin{aligned}
\mathbf{x}_{p,i}(t) &= \mathbf{x}_0 + \mathbf{v}_{0,i}(t - t_0) + \frac{1}{2}\mathbf{a}_{0,i}(t - t_0)^2 + \frac{1}{6}\mathbf{a}_{0,i}^{(1)}(t - t_0)^3 \\
\mathbf{v}_{p,i}(t) &= \mathbf{v}_{0,i} + \mathbf{a}_{0,i}(t_0)(t - t_0) + \frac{1}{2}\mathbf{a}_{0,i}^{(1)}(t - t_0)^2.
\end{aligned} \tag{2.4}$$

4. Using the $\mathbf{x}_{p,i}$ and $\mathbf{v}_{p,i}$ for all particles in the current block, the GRAPE-6 returns the accelerations and the jerks at the end of the time step for all of them;

5. The correction is, then, performed for all the particles in the current block according to

$$\begin{aligned}
\mathbf{r}_c &= \mathbf{r}_p(t) + \frac{1}{24}\mathbf{a}_{0,i}^{(2)}\Delta t^4 + \frac{1}{120}\mathbf{a}_{0,i}^{(3)}\Delta t^5 \\
\mathbf{v}_c &= \mathbf{v}_p(t) + \frac{1}{6}\mathbf{a}_{0,i}^{(2)}\Delta t^3 + \frac{1}{24}\mathbf{a}_{0,i}^{(3)}\Delta t^4.
\end{aligned} \tag{2.5}$$

This The time step criteria is determined by the following expression derived empirically by Aarseth (1985)

$$\Delta t = \sqrt{\eta \frac{|\mathbf{a}||\mathbf{a}^{(2)}| + |\mathbf{a}^{(1)}|^2}{|\mathbf{a}^{(1)}||\mathbf{a}^{(3)}| + |\mathbf{a}^{(2)}|^2}}. \tag{2.6}$$

At the beginning of the simulation, in order to determine the first time step, this expression is not useful because the second and third derivatives of the accelerations are not known. For this special case, a simpler starting criteria is employed

$$\Delta t = \sqrt{\eta \left|\frac{\mathbf{a}}{\mathbf{a}^{(1)}}\right|}. \tag{2.7}$$

Note that this simpler time step criteria is, in general, not used since it is much less robust than the former. For instance, in the (artificial) case where for a given particle to be advanced one time step, there is a symmetric configuration of perturbers around it, the resulting acceleration is zero or close to zero, therefore implying a zero (or very small) value for the next time step. This is obviously a very undesirable situation, hence this criteria is not used except for the first time step.

2.2 Canonical KS-Transformation

The gravitational interactions between stars and SMBHs in galactic nuclei with very high stellar density is a very delicate numerical problem. In particular, it is crucial to integrate very accurately the close encounters between stars and BHs (and between BHs if there are more than one). The standard way to deal with two-body close encounters — either bound or unbound — is to use the KS regularization [Stiefel and Scheifele, 1971, Aarseth, 2003]. In KS regularization, the equations for the relative two-body motion are transformed into those of a four dimensional harmonic oscillator — which is very convenient numerically since the singularity has been removed. Then, this is combined with a time transformation $dt = gds$, where $g = R$ is the time step function and s is the ficticious time. As a result, we obtain an Hamiltonian in the extended phase space of a four dimensional harmonic oscillator, $\Gamma = g(H + p_0) = 1/8\mathbf{P}^2 + P_0\mathbf{Q}^2 - G(m_1 + m_2) + Q^2\Phi_{pert}$, where $Q_0 = t$, $P_0 = -E$, $(\mathbf{P}, \mathbf{Q})$ are the four-dimensional KS variables, and Φ_{pert} is an external perturbing potential.

In order to consider this problem in more detail, it is useful to write the Hamiltonian Γ in the extended phase space [Boccaletti and Pucacco, 1996]:

$$\Gamma = H(\mathbf{q}, \mathbf{p}) + p_0, \tag{2.8}$$

where $q_0 = t$ is a phase space coordinate, $p_0 = -H$ is the corresponding conjugate momentum, and the complete extended phase space is $(q_0, \mathbf{q}; p_0, \mathbf{p})$. If we introduce a time step function $dt = g(\mathbf{q})ds$, where s is the ficticious time, and g is an arbitrary function of the coordinates $\mathbf{q}$, we can write down a new Hamiltonian Γ for the extended phase space

$$\Gamma \ = g(\mathbf{q})\left[H(\mathbf{q}, \mathbf{p}) + p_0\right]. \tag{2.9}$$

The Hamilton's equations of motion are then given by

$$\begin{aligned} q_0' &= \frac{\partial \Gamma}{\partial p_0} = g(\mathbf{q}), \qquad \mathbf{q}' = \frac{\partial \Gamma}{\partial \mathbf{p}} = g(\mathbf{q})\frac{\partial H}{\partial \mathbf{p}} \\ p_0' &= -\frac{\partial \Gamma}{\partial q_0} = 0, \qquad \mathbf{p}' = -\frac{\partial \Gamma}{\partial \mathbf{q}}, \end{aligned} \tag{2.10}$$

where the prime denotes differentiation with respect to the ficticious time s. The trajectories on the hypersurface $\Gamma = 0$ in the extended phase space correspond to the solutions of the equations of motion in the original phase space.

In 1921, Levi-Civita introduced a canonical transformation in the 2D plane, taking advantage that the unperturbed two-body motion is planar, through the so-called *Levi-Civita matrix* [Stiefel and Scheifele, 1971]

$$\hat{\mathbf{L}} = \begin{pmatrix} Q_1 & -Q_2 \\ Q_2 & Q_1 \end{pmatrix}. \tag{2.11}$$

According to this transformation, the new and old coordinates are related by

$$\mathbf{q} = \hat{\mathbf{L}}\mathbf{Q}, \qquad \mathbf{p} = \frac{1}{2}\hat{\mathbf{L}}\mathbf{P}/Q^2,$$

hence the time transformed Kepler Hamiltonian

$$\Gamma = R\left[\frac{\mathbf{p}^2}{2} - \frac{G(m_1 + m_2)}{R}\right] \tag{2.12}$$

becomes the regularized Hamiltonian

$$\Gamma_{reg} = \frac{1}{8}\mathbf{P}^2 - E\mathbf{Q}^2 - G(m_1 + m_2) \tag{2.13}$$

which is, in fact, the Hamiltonian of a harmonic oscillator. The constant term $G(m_1 + m_2)$ is, in fact, unconsequential to the dynamics. This represents the complete regularization of the Kepler motion in the plane. However, as soon as we consider an external perturbation, the motion is no longer confined to the plane and a generalization of the Levi-Civita transformation for 3D motion is necessary. Before moving on to the KS transformation for the Kepler perturbed motion, it is important to stress that the KS regularization method is based in two simultaneous, independent transformations: a time transformation and a canonical transformation of the phase space coordinates. This will still be true when we generalize the two-body KS regularization to the N-body chain regularization, and for the perturbed problem as well.

The Levi-Civita transformation is a conformal transformation in the complex plane and thus it is not possible to construct a 3×3 matrix with the same properties as $\hat{L}$. It was only in 1965 that Kustaanheimo & Stiefel, inspired by the theory of spinors, introduced the desired canonical transformation by using a pair of complex numbers, but this results only at the price of working in a 4D space.

The KS transformation can be represented in matricial form as follows

$$\mathbf{q} = \hat{\mathbf{Q}}\mathbf{Q}, \qquad \mathbf{p} = \frac{1}{2}\hat{\mathbf{Q}}\mathbf{P}/Q^2,$$

where the KS matrix $\hat{\mathbf{Q}}$ and $\hat{\mathbf{P}}$ are given by

$$\hat{\mathbf{Q}} = \begin{pmatrix} Q_1 & -Q_2 & -Q_3 & Q_4 \\ Q_2 & Q_1 & -Q_4 & -Q_3 \\ Q_3 & Q_4 & Q_1 & Q_2 \\ Q_4 & -Q_3 & Q_2 & -Q_1 \end{pmatrix}. \tag{2.14}$$

In components, the new and old coordinates are related by

$$\begin{aligned} q_1 &= Q_1^2 - Q_2^2 - Q_3^2 + Q_4^2, \qquad (2.15)\\ q_2 &= 2(Q_1Q_2 - Q_3Q_4), \\ q_3 &= 2(Q_1Q_3 + Q_2Q_4), \\ q_4 &= 0. \end{aligned}$$

We have chosen to show the last expression explicitly in terms of the components in order to highlight the *non-uniqueness* of the KS transformation. In fact, the entries of the 4^{th} line of $\hat{Q}$ could have been identically zero; indeed, these entries may take any values so long as $R_4 = 0$ as there is a whole set of vectors $\mathbf{Q}$ corresponding to a given vector $\mathbf{q}$.

Since $\mathbf{q}_k = (X, Y, Z)^t$ and $R = |\mathbf{q}_k|$, we have

$$Q_1 = \sqrt{(R + |X|)/2} \tag{2.16}$$

$$Q_2 = Y/(2Q_1) \tag{2.17}$$

$$Q_3 = Z/(2Q_1) \tag{2.18}$$

$$Q_4 = 0,$$

and the components of the KS coordinates $\mathbf{Q}_k$ are then given by

$$\mathbf{Q} = \begin{cases} (Q_1, Q_2, Q_3, Q_4)^t, & \text{for } x \geq 0 \\ (Q_2, Q_1, Q_4, Q_3), & \text{for } x < 0 \end{cases} \tag{2.19}$$

Next with the following time transformation

$$\frac{dt}{ds} = R = \mathbf{Q}^2, \tag{2.20}$$

we arrive at the Hamiltonian for a 4-dimensional perturbed harmonic oscillator

$$\Gamma = \frac{1}{8}\mathbf{P}^2 - E\mathbf{Q}^2 + Q^2\Phi_{pert} - G(m_1 + m_2), \tag{2.21}$$

where Φ_{pert} is an external, perturbing potential.

Finally, the regularized equations of motion are then

$$\begin{aligned} \mathbf{Q}'' &= \frac{1}{2}E\mathbf{Q} + \frac{1}{2}R\hat{\mathbf{Q}}^t\left(\mathbf{F}_{2,p} - \mathbf{F}_{1,p}\right), && (2.22) \\ E' &= 2\mathbf{Q}' \cdot \hat{\mathbf{Q}}^t\left(\mathbf{F}_{2,p} - \mathbf{F}_{1,p}\right), && (2.23) \\ t' &= \mathbf{Q} \cdot \mathbf{Q} = R, \end{aligned}$$

where $R = |\mathbf{q}| = Q^2$, $\mathbf{F}_p$ are the perturbing forces acting on each of the binary's components. Several points about this last set of equations are worth of some comments: (*i*) these equations are regular at collision ($R = Q = 0$); (*ii*) if $\Phi_{pert} \neq 0$, the energy E is no longer conserved and there is an extra energy equation to be solved; (*iii*) as a consequence of (*ii*), the oscillator frequency is no longer a constant; (*iv*) the number of equations to be integrated is now ten, rather than six, plus the KS transformation of phase space coordinates — however, this increase of computational effort is largely compensated by the fact that the singularity has been removed; (*v*) the external force (tidal) perturbations, on the relative two-body motion from particles at very large distances, can be neglected since they scale as $F_{pert} \sim 1/R^3$, only the force on the center of mass needs to be considered in this case.

In the perturbed case, the center of mass motion of the binary needs to be considered explicitly. In the N-body code, a ficticious particle is created to represent the binary's c.m. and the usual equation of motion

$$\ddot{\mathbf{r}}_{c.m.} = \frac{m_1\mathbf{F}_{1,pert} + m_2\mathbf{F}_{2,pert}}{m_1 + m_2}, \tag{2.24}$$

is used; it is then advanced by the Hermite integrator as an ordinary particle.

Note that it was only with the introduction of the KS regularization scheme that [Szebehely and Peters, 1967] were able in 1967 to solve numerically, for the first time, the famous

Pythagorean problem of celestial mechanics ([Szebehely and Peters, 1967]; see also [Diacu and Holmes, 1996] for historical context) — this work led to important insights on the problem of binary formation which is so crucial for the modelling of stellar clusters.

2.3 Chain Regularization

In the vicinity of a central BH, we will need to deal with very compact configurations of stars and the two-body KS regularization may at times not be sufficient. This occurs in several important circumstances for us: (i) The motion of several particles orbiting the massive particle with very small semimajor axis and/or very high eccentricities; (ii) The interaction of a field star with a massive black hole binary (slingshot effect); (iii) The formation and dynamical evolution of triple massive black hole configurations after a galaxy merger, and so on.

The chain algorithm (Mikkola & Aarseth, 1990, 1993) provides an elegant solution for this problem. Given a number ($N \geq 2$) particles, this integration algorithm orders them in a linear chain according to the relative distances (we have found more convenient, for our purposes, due to the high mass ratio between the black hole particle and the field stars, to sort the particles by forces instead, as we discuss below): it starts by selecting the closest two particles and connects them, then it searches for the particle closest to either one of the ends of the chain and adds it to the closest end, and so on, until all particles (selected according to an empirical criteria based either on the distance or interaction strength) are included in the chain. After rewritting the equations of motion for the relative vector separations of consecutive particles along the chain, a two-body KS transformation is then applied to each of them. This is a generalization of the two-body KS transformation to the N-body chain regularization. For the chain regularization, a slighlty different time transformation, $ds = gdt$ with $g = 1/L$, where L is the Lagrangean, is more appropriate than the simplest generalization of the two-body case $g = R_1 R_2 ... R_N$ — see discussion in [Mikkola and Aarseth, 1990, Mikkola and Aarseth, 1993]. However, there is nothing fundamental about the choice of time step function and its choice is based entirely on the observed performance from numerical experiments.

It should be noted that, in the case $N \geq 3$, singular terms are present in the regularized Hamiltonian — they are precisely those that correspond to the interaction terms between non-consecutive members of the chain. These terms do not cause any numerical problems so long as the $N - 1$ chain vectors are the smallest separations/stronger interactions between the chained particles (as they should by construction). As a consequence, after each integration step, it is necessary to check whether the chain particles are still ordered according to the present distances (or forces) — otherwise switching of the positions of the particles in the chain must be done [Mikkola and Aarseth, 1990, Mikkola and Aarseth, 1993]. The main limitation of the chain regularization is that it is not appropriate to integrate systems with extremely large mass ratios since the total energy, which appears explicitly in the Hamiltonian, then becomes dominated by the heavy particles. It can, however, handle well up to mass ratios of the order $\sim 10^6$; but for 10^9, the total energy error is already of the same order of that of the small masses [Mikkola and Aarseth, 2002]. The highest mass ratio in our calculations never exceed 10^4, so we are quite safe to use the chain regularization scheme. As a corollary, the motion of test particles cannot be integrated by the chain regularization since the masses appear in the denominator of some terms in the Hamiltonian.

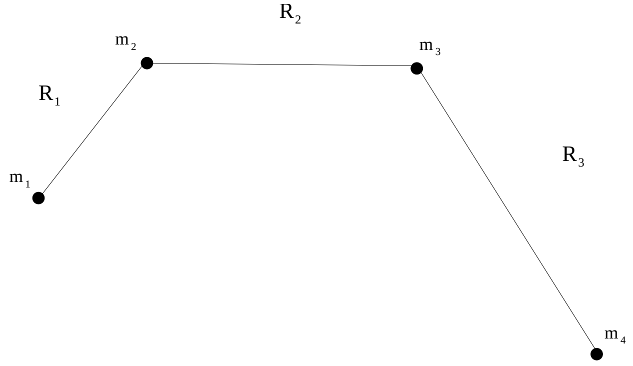

Figure 2.1: Chain diagram for $N_{ch} = 4$.

Let us now suppose there are N_{ch} particles to be included in the chain during a N-body calculation. The Hamiltonian for these chain particles can be written as

$$H = \frac{\mathbf{p}_{cm}^2}{2M_{ch}} + \sum_{i=1}^{N_{ch}} \frac{\mathbf{p}_i^2}{2m_i} - \sum_{1 \le i \le j \le N_{ch}} \frac{m_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|} - \Phi_{pert}, \tag{2.25}$$

where the subscript c.m. denotes the center of mass and Φ_{pert} is the perturbing potential due to the perturbing field stars not included in the chain.

After the chain vectors are defined, and the particles have been relabelled from 1 to N_{ch} along the chain, and going to the chain's center of mass frame, the following generating function [Boccaletti and Pucacco, 1996]

$$S = \sum_{i=1}^{N_{ch}-1} \mathbf{W}_i \cdot (\mathbf{q}_{i+1} - \mathbf{q}_i), \tag{2.26}$$

defines the canonical transformation $S : (\mathbf{q}, \mathbf{p}) \to (\mathbf{R}, \mathbf{W})$ that yields the chain's phase space coordinates. For a generating function of the type $S = S(\mathbf{q}, \mathbf{W})$:

$$\begin{aligned} \mathbf{p}_k &= \frac{\partial S}{\partial \mathbf{q}_k} = \mathbf{W}_{k-1} - \mathbf{W}_k \quad (k = 2, 3, ..., N_{ch-1}) \qquad (2.27) \\ \mathbf{p}_1 &= -\mathbf{W}_1, \quad \mathbf{p}_N = \mathbf{W}_{N_{ch}-1}, \end{aligned}$$

while

$$\mathbf{R}_k = \frac{\partial S}{\partial \mathbf{W}_k} = \mathbf{q}_{k+1} - \mathbf{q}_k. \tag{2.28}$$

Let us, for illustrative purposes, consider the situation where four particles were included in the chain. As a result of these transformations, the (singular) chain's Hamiltonian, in the chain's c.m. frame, and ommiting the external potential energy, is given by

$$\begin{aligned} H = \frac{1}{2}\left(\mu_{12}\mathbf{W}_1^2 + \mu_{23}\mathbf{W}_2^2 + \mu_{34}\mathbf{W}_3^2\right) - \left(\frac{m_1 m_2}{R_{12}} + \frac{m_2 m_3}{R_{23}} + \frac{m_2 m_4}{R_{24}}\right) \\ -\frac{1}{m_2}(\mathbf{W}_1 \cdot \mathbf{W}_2) - \frac{1}{m_3}(\mathbf{W}_2 \cdot \mathbf{W}_3) - \Phi_{n.c.}, \end{aligned} \tag{2.29}$$

where $\mu_{ij} = (m_i + m_j)/m_i m_j$ and $\Phi_{n.c.}$ represent the interaction potential between non-consecutive chain particles

$$\Phi_{n.c.} = \frac{m_1 m_2}{|\mathbf{R}_1 + \mathbf{R}_2|} + \frac{m_2 m_4}{|\mathbf{R}_2 + \mathbf{R}_4|} + \frac{m_1 m_4}{|\mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3|}. \tag{2.30}$$

Applying the KS transformations, $\mathbf{R}_k = \hat{\mathbf{Q}}_k \mathbf{Q}_k$ and $\mathbf{W}_k = \hat{\mathbf{Q}}_k \mathbf{P}_k/(2Q_k^2)$, and the time transformation $dt = g ds$, we find the regularized Hamiltonian $\Gamma_{reg} = g\left[H(\mathbf{P}, \mathbf{Q}) + p_0\right]$. We write down this Hamiltonian, for the case $N_{ch} = 4$ and for a time transformation $g = R_1 R_2 R_3$:

$$\begin{aligned} \Gamma_{reg} = \frac{1}{8}\left(\mu_{12} R_2 R_3 \mathbf{P}_1^2 + \mu_{23} R_1 R_3 \mathbf{P}_2^2 + \mu_{34} R_1 R_2 \mathbf{P}_3^2\right) - (m_1 m_2 R_2 R_3 + m_2 m_3 R_1 R_3 \\ + m_3 m_4 R_1 R_2) - \frac{1}{4}\left(\frac{R_3}{m_2}\mathbf{P}_1^t \hat{\mathbf{Q}}_1^t \hat{\mathbf{Q}}_2 \mathbf{P}_2 + \frac{R_1}{m_3}\mathbf{P}_2^t \hat{\mathbf{Q}}_2^t \hat{\mathbf{Q}}_3 \mathbf{P}_3\right) - R_1 R_2 R_3 \left(\Phi_{n.c.} + E\right), \end{aligned} \tag{2.31}$$

where E is the chain's total energy at the initial time t_0.

Having written down this Hamiltonian help us to recognize the following properties: (*i*) the equations of motion are regular for the chain vectors $R_1, R_2, R_3 \to 0$: hence consecutive particles along the chain may have collisions or close encounters without numerical difficulties; (*ii*) it reduces to the old triple regularization scheme [Aarseth and Zare, 1974] for $N_{ch} = 3$, and to the two-body KS regularization when $N_{ch} = 2$; (*iii*) although the non-chained distances are still singular, collisions or close encounters for which $R_{13}, R_{14}, R_{24} \to 0$ can still be dealt with provided we switch the particles within the chain in order to regularize the relevant interaction; (*iv*) it is possible to integrate close triple, quadrupule, etc. encounters without numerical problems so long as the singular distances are always larger than all the regularized ones (and this can always be achieved by switching). It should also be mentioned that switching is not too frequent for most interactions.

We need now to consider the effect of the external forces on both the chain's internal degrees of freedom and on its center of mass. Denote the perturbing acceleration acting on the chain member m_k by $\mathbf{F}_k$, and assume that the forces are dependent on time. The perturbing potential Φ_{pert} due to the field particles is given by

$$\Phi_{pert} = \sum_i m_i \mathbf{r}_i \cdot \mathbf{F}_i(t), \tag{2.32}$$

and therefore the induced accelerations on the chain are

$$\begin{aligned} \dot{\mathbf{W}}_1 &= -\frac{\partial \Phi_{pert}}{\partial \mathbf{R}_1} = -m_1 \left(\mathbf{F}_1 - \mathbf{F}_{cm}\right), & (2.33) \\ \dot{\mathbf{W}}_k &= -\frac{\partial \Phi_{pert}}{\partial \mathbf{R}_k} = -m_1 \left(\dot{\mathbf{W}}_{k-1} - \dot{\mathbf{p}}_k\right), & (2.34) \\ \dot{\mathbf{W}}_{N-1} &= \dot{\mathbf{p}}_N, \end{aligned}$$

where $\dot{\mathbf{p}}_k = m_k \left(\mathbf{F}_k - \mathbf{F}_{cm}\right)$, for $k = 1, 2...., N_{ch}$. Using the properties of the KS transformations, the effect of the external potential on the KS momenta is given by

$$\mathbf{P'}_k = 2g\hat{\mathbf{Q}}^t \dot{\mathbf{W}}_k, \tag{2.35}$$

and the rate of change of the chain's internal energy is

$$E' = 2\sum_{k=1}^{N_{ch}} \hat{\mathbf{Q}}_k^t \dot{\mathbf{W}}_k. \tag{2.36}$$

Note that the prime denotes here again differentiation with respect to the ficticious time. It is important to stress that the external forces are not restricted to be conservative ones; in fact, several types of dissipative forces can be included (*eg.* dissipative terms due to relativistcs corrections etc.)

2.4 Two-Body KS and Chain Interfaces

In a N-body code, the KS/chain particles will not be isolated so one must take care of the time evolution of the KS/chain's center of mass due to the stars outside the chain. The chain's stars are also affected by the gravitational potential from the rest of the galaxy and thus external forces must be taken into account for the time evolution of the internal degrees of freedom. Finally, the chain membership will evolve over time, so it is thus crucial to develop appropriate criteria for absorption of new stars into the chain, as well as to release (emmit) stars from the chain.

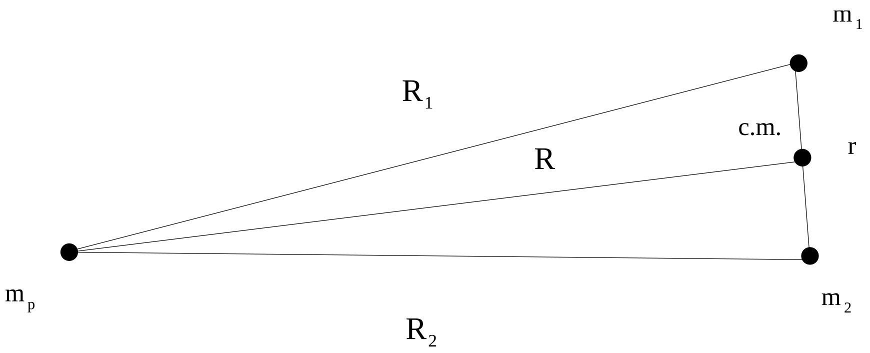

Figure 2.2: Tidal estimate for $N_{ch} = 2$.

Let us suppose there are two particles of masses m_1 and m_2 in the chain; then, the differential force from each field particle of mass m_p on the chain stars is given by

$$\Delta\mathbf{F} = Gm_p\left[\frac{\mathbf{R}_2}{R_2^2} - \frac{\mathbf{R}_1}{R_1^2}\right] = Gm_p\left[\frac{\mathbf{R}-\mu_1\mathbf{r}}{|\mathbf{R}-\mu_1\mathbf{r}|^3} - \frac{\mathbf{R}+\mu_2\mathbf{r}}{|\mathbf{R}+\mu_2\mathbf{r}|^3}\right], \qquad (2.37)$$

where $\mathbf{R_i}$ is the vectorial separation between the chain particles and the perturbing star m_p, $\mu_i = m_i/(m_1+m_2)$, $\mathbf{R}$ is the separation between the center of mass and m_p, and $\mathbf{r} = \mathbf{r_2} - \mathbf{r_1}$ $(R, R_i \gg r)$. We can perform a multipole expansion and, keeping the lowest order term only, obtain $|\Delta\mathbf{F}| \sim (Gm_p/R^3)r$. The dimensionless parameter γ is defined to be the ratio between the differential (tidal) force $\mathbf{\Delta F}$ due to m_p on the chain particles and the force $\mathbf{F}$ between them:

$$\gamma = \left|\frac{\Delta\mathbf{F}}{\mathbf{F}}\right| = \frac{m_p}{(m_1+m_2)}\left(\frac{r}{R}\right)^3. \qquad (2.38)$$

If γ is small enough, it is a fair approximation to neglect the differential (tidal) effect due to m_p, and consider only its effect on the chain's center of mass. Therefore, there is a critical radius R_{crit} inside which field particles are treated as perturbers of the chain; this critical radius is

$$R_{crit} = \left[\frac{m_p}{(m_1+m_2)\gamma_{min}}\right]^{1/3} r. \qquad (2.39)$$

All particles inside the critical radius R_{crit} will be considered as *perturbers* of the chain — ie. each chain particle feels the gravitational force from the perturbers. On the other hand, the particles outside R_{crit} affect the chain only through the force they exert on its center of mass. We find $\gamma_{min} = 10^{-6}$ to provide an adequate criteria for the identification of perturbers.

The *center of mass approximation* is used for computing the force on the field particle m_p, due to the chain's particles, whenever they are far enough so that it is not necessary to resolve the chain members. In order to be more quantitative, we will define a new dimensionless parameter λ_p which is the ratio between the differential force from the chain particles m_1, m_2 and the force from their center of mass on the field stars:

$$\lambda_p = \frac{|\Delta \mathbf{F}|}{|\mathbf{F}|} \sim \frac{G(m_1+m_2)r}{R^3} \times \frac{R^2}{GM_{cm}} = \frac{r}{R}. \tag{2.40}$$

Then there is a critical radius $\mathcal{R}_{crit}$ above which the force from the chain is computed by adopting the c.m. approximation

$$\mathcal{R}_{crit} = \lambda_p r = \gamma_{min}^{-1/3} r. \tag{2.41}$$

For our choice $\gamma_{min} = 10^{-6}$, $\lambda_p = 100$ obtains. We have motivated these expressions by focusing on the particular case $N_{ch} = 2$ — a few modifications are required to deal with the general case. The separation r between the binary members was previously the obvious measure for the chain size; a natural generalization for $N_{ch} > 2$ is to take the largest distance between any of the chain stars and the central massive particle. Note that we are specializing for the single black black hole case. In the case of a massive black hole binary, it will be the largest distance between any chain star measured with respect to the binary's center of mass.

The perturbers will affect the center of mass through the familiar expression

$$\ddot{\mathbf{r}}_{cm} = \frac{m_1 \mathbf{F}_1 + m_2 \mathbf{F}_2}{m_1 + m_2}, \tag{2.42}$$

where $\mathbf{F_1}$ and $\mathbf{F_2}$ are the total perturbing forces (per unit mass), on each binary member, due to the perturbers.

To advance the chain's center of mass, we need to create a new particle in the N-body code that will be advanced by the Hermite integrator in the same manner as an ordinary particle. Since we need to know the coordinates of the perturbers for the integration of the chain's internal degrees of freedom, we require that the chain's center of mass stepsize is always equal to the smallest step size among all particles in the simulation at a given instant of time.

In a N-body calculation, the chain membership will have to evolve over time; hence criteria for absorption and emission of particles from the chain are necessary. The chain is ideally suited for treating compact configurations, including a single binary (in which case it reduces to the two-body KS regularization) — so we should choose stringent criteria for chain membership. In fact, if we allow too large number of particles to become members of the chain, the integration can be slowed down severely. However, in some situations, *eg.* massive BH binary embedded in a steep cusp of field stars, it can handle efficiently very large number of members. There are two basic parameters to decide whether to accept a field particle into the chain: *DTMIN*, the field particle should have a step size smaller than this value; *RMIN*, the field particle should be absorbed only after it has entered the region within this radius measured from the single BH (in the case of a BH binary, we would set $RMIN = \beta a$, where a is the binary's semimajor axis and β is a constant of order unity); in the case of a single massive particle a is a certain critical radius for absorption into the chain. The emission of stars from the chain is done whenever a chain star crosses (in the outward sense) the radius $REM = \alpha RMIN$, where α is a constant slightly larger than one to avoid continuing absorption and emission of the particle every orbital period. In practice, REM is chosen to be $\sim (1.0 - 1.5) RMIN$.

To minimize the number of switchings, and thus possible loss of accuracy, we have chosen to use the force between chain particles, rather than distances, as the sorting quantity for the choice of chain vectors, which is a better measure of the pairwise interaction strength. We have verified that this means that the BH(s) will remain most of the time on the bottom of the chain, decreasing very much the number of switchings involving the BH(s).

Te chain's equations of motion are integrated with the Bulirsch-Stoer extrapolation method [Deuflhard, 1985,J.Stoer and Bulirsch, 1993] using the midpoint method as the basic integrator. A given basic step of the chain is completed via gradually decreasing substepsize evaluations; the results are then extrapolated to zero stepsize using a rational function. This integrator is very accurate as it provides automatic error estimate and chooses the stepsize based on the estimate for the relative error, as measured by $\Gamma = (H - E)/L < \epsilon$. We choose $\epsilon = 10^{-10}$ or $\epsilon = 10^{-12}$. This method provides very accurate solutions for non-stiff equations — there is, of course, some perturber configurations that may conceivably cause difficulties if the stepsize needs to be decreased by too large a fraction in order to meet the relative error criterium.

The implementation of the KS/chain regualrization consists mainly of the following steps (not necessarily by this order):

- check whether there are particles due for forming a KS pair, or absorption or emmission to the chain;
- evaluate and sort mutual distances (or forces);
- relabel the particle's indices and form chain vectors;
- transform to KS coordinates and momenta;
- apply time transformation g;
- determine the perturbers;
- determine the field particles for which the KS/chain must be resolved;
- advance the KS pair/chain forward in time with Hermite scheme/Bulirsch-Stoer method;
- advance the KS pair/chain's c.m. forward in time as a N-body particle.

Note that all decisions concerning emission/absorption, perturbers, chain resolution, switching, etc are done at every step. Since the Bulirsch-Stoer method involves the evaluation of the vector field (perturbers included) at different subintervals of a given step, we perform a low order prediction for the perturbers at each substep.

2.5 The Chang-Cooper differencing scheme for the $1D$ Fokker Planck equation

The Fokker-Planck equation in energy space which is used for the modelling of spherically symmetric stellar clusters with isotropic distribution of velocities is of the following general form

$$\frac{\partial u}{\partial t} = \frac{1}{A(x)} \frac{\partial}{\partial x} \left[B(x,t)u + C(x,t) \frac{\partial u}{\partial x} \right], \tag{2.43}$$

it is defined in the domain $0 \leq t < t_f$, and $0 \leq x < \infty$. The coefficients A, B and C are all positive functions of their arguments and have familiar physical interpretation: $4\pi^2 A(x)dx$ is the spherically averaged volume element; B is the dynamical friction coefficient; and C is the heating coefficient. The quantity x represents, in our case, the stars's binding energy E and $u(x,t)$ represents the single particle distribution function $f(E)$.

The goal is to find the long term solution $u(x,t)$ to the above partial differential equation given some initial and boundary conditions. A large number of standard finite difference schemes exist to obtain stable, convergent solutions to (2.43), provided that a fine enough x-mesh Δx and time step Δt are employed. Chang & Cooper 1972, motivated by questions arising in the kinetic study of plasma physics, introduced an implicit differencing scheme that preserves some of the intrinsic properties of the exact solutions to the Fokker-Planck equation. In particular: (i) The positiveness of the solution $u(x,t) \geq 0$ for all times $t > t_0$, provided an initial condition $u(x,t_0) \geq 0$; (ii) The strict conservation of particle number, $N = \int_0^{+\infty} dx \; u(x,t) \; A(x)$, so long as no external sources or sinks are present; as if, for instance, the boundary conditions are of the flux-null type; (iii) An accurate resolution of the steady-state solution (more about this below). Furthermore, their new scheme also allows the adoption of larger time steps Δt and wider mesh Δx while keeping a very high accuracy.

To construct a numerical solution of equation (2.43), $u(x,t)$ is represented by a discrete set $u_j^n = u(j\Delta x, n\Delta t)$. We will usually adopt a constant time step and a logarithmic mesh in energy space for all our calculations. For definiteveness, let us employ the forward difference for the time detivative and the centered difference for the derivative with respect to x

$$\begin{aligned} \left.\frac{\partial u}{\partial t}\right|_j &= \frac{u_j^{n+1} - u_j^n}{\Delta t} \\ \left.\frac{\partial^2 u}{\partial x^2}\right|_j &= \frac{u_{j+1} - 2u_j + u_{j-1}}{\Delta x^2}. \end{aligned} \tag{2.44}$$

In order to write down the difference scheme to integrate (2.43), it is costumary to rewrite it in the following form

$$\frac{\partial u}{\partial t} = \frac{1}{A(x)}\left[\alpha u + \beta\frac{\partial u}{\partial x} + \gamma\frac{\partial^2 u}{\partial x^2}\right], \tag{2.45}$$

where $\alpha = \partial B(x,t)/\partial x$, $\beta = B(x,t) + \partial C(x,t)/\partial x$, and $\gamma = C(x,t)$.

One of the simplest differencing implicit schemes is given by

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{1}{A_j}\left[\alpha_j^n \tilde{u}_j + \beta_j^n \frac{\tilde{u}_{j+1} - \tilde{u}_j}{2\Delta x} + \gamma_j^n \frac{\tilde{u}_{j+1} - 2\tilde{u}_j + \tilde{u}_{j-1}}{\Delta x^2}\right], \tag{2.46}$$

where $\tilde{u}$ is a linear combination of u_j^n and u_j^{n+1} depending on the specific differencing scheme; $j \in [0, J]$, $J\Delta x$ would be the maximum energy chosen to be large enough so that $u_{j>J} = 0$ (in general, we will adopt some boundary condition at large, but finite, x).

Let us now look at some of the properties of this differencing scheme; with that goal in mind, take the total number of particles to be given approximately by*

$$N^n = \sum_{j=0}^{J} A_j u_j^n \Delta x. \tag{2.47}$$

The exact conservation of particle number is, by definition, given by the condition: $N^{n+1} = n^n$. Hence, summing over all the mesh points $j = 0, ..., J$

$$\begin{aligned} \frac{N^{n+1} - n^n}{\Delta t \Delta x} \quad = \quad & \sum_{j=1}^{J-1}\left[\frac{\gamma_{j+1}^n - 2\gamma_j^n + \gamma_{j-1}^n}{\Delta x^2} + \frac{\beta_{j-1}^n - \beta_{j+1}^n}{2\Delta x} + \alpha_j^n\right]\tilde{u}_j + \\ & \left(\frac{\alpha_J^n}{\Delta x^2} + \frac{\beta_J^n}{2\Delta x}\right)\tilde{u}_{J+1} + \left(\frac{\gamma_{J-1}^n - \gamma_J^n}{\Delta x^2} + \frac{\beta_{J-1}^n}{2\Delta x} + \alpha_J^n\right)\tilde{u}_J + \\ & \left(\frac{\gamma_1^n - \gamma_0^n}{\Delta x^2} + \frac{\beta_1^n}{2\Delta x} + \alpha_0^n\right)\tilde{u}_0 + \left(\frac{\gamma_0^n}{\Delta x^2} - \frac{\beta_0^n}{2\Delta x}\right)\tilde{u}_{-1}. \end{aligned} \tag{2.48}$$

The terms under the addition sign in (**??**) are called the *interior* terms ($j = 1, ..., J-1$), while the others refer to the boundaries ($j = -1, 0, J, J+1$). This is the discretized version for the exact rate of change of the total number of particles

$$\frac{\partial N}{\partial t} = \left[B(x,t) + C(x,t)\frac{\partial u}{\partial x}\right]_0^{+\infty}. \tag{2.49}$$

---

[*] A more exact representation would not affect the results.

Note that the latter equation — which is the exact conservation law — has only boundary terms; consequently, even if appropriate zero-flux boundary conditions were imposed, the present differencing scheme would not conserve particle number due to the existence of the generally non-vanishing interior terms (which will effectively act as spurious sources and sinks of particles). It can be shown, by Taylor expanding the coefficients α, β, and γ, that the error in particle number conservation is $\propto \Delta x^2$.

Chang & Cooper start by applying the differencing scheme directly to the Fokker-Planck equation in its canonical form (2.43). In order to proceed, they had to define a generalized flux $F = Bu + C\partial u/\partial x$

$$\frac{u_{j+1}^n - u_j^n}{\Delta t} = \frac{1}{A_j \Delta x}\left(\tilde{F}_{j+1/2} - \tilde{F}_{j-1/2}\right). \tag{2.50}$$

Summing over all mesh points

$$\sum_{j=0}^{J} A_j \frac{\Delta x}{\Delta t}\left(u_j^{n+1} - u_j^n\right) = \sum_{j=0}^{J}\left(F_{j+1/2} - F_{j-1/2}\right) = F_{J+1/2} - F_{-1/2}, \tag{2.51}$$

and the discrete version of the exact conservation law obtains. The numerical rate of change of the particle now depends on the boundary terms only, and thus conservation is exactly achieved if zero-flux boundary conditions are adopted, regardless of the mesh or of the step size. Note that to define $\tilde{F}$ it is necessary to specify $\tilde{u}_{j\pm 1/2}$, the usual choice being $u_{j\pm 1/2}^{n+1} = \frac{1}{2}(u_j^{n+1} + u_{j+1}^{n+1})$, but this can be sown to result in negative u_j if the mesh is not fine enough. In order to avoid this problem, Chang & Cooper adopted the following solution:

$$u_{j+1/2}^{n+1} = (1 - \delta_j)u_{j+1}^{n+1} + \delta_j u_j^{n+1}, \tag{2.52}$$

for $0 \leq \delta_j \leq 1/2$. This means that the $'j + 1/2'$ value will be somewhere between the j and the $j + 1$ values. By comparing the theoretical and numerical scheme's properties of the equilibration process and of the steady-state, they arrived at the following prescription

$$\delta_j = \frac{1}{w_j} - \frac{1}{e^{w_j} - 1}, \tag{2.53}$$

which will range from 1/2 to 0, while w goes from to 0 to $+\infty$, and $w_j = \Delta x B_{j+1/2}^n / C_{j+1/2}^n$. The resulting Chang & Cooper scheme is therefore given by

$$\begin{aligned} \frac{u_j^{n+1} - u_j^n}{\Delta t} &= \frac{1}{A_j \Delta x}\left[C_{j+1/2}^n W_j e^{w_j} u_{j+1}^{n+1} - \right. \\ &\left. \left(C_{j+1/2}^n W_j + C_{j-1/2}^n W_{j-1} e^{w_{j-1}}\right) u_j^{n+1} + C_{j-1/2}^n W_{j-1} u_{j-1}^{n+1}\right], \end{aligned} \tag{2.54}$$

where $W_j = w_j/(e^{w_j} - 1)$.

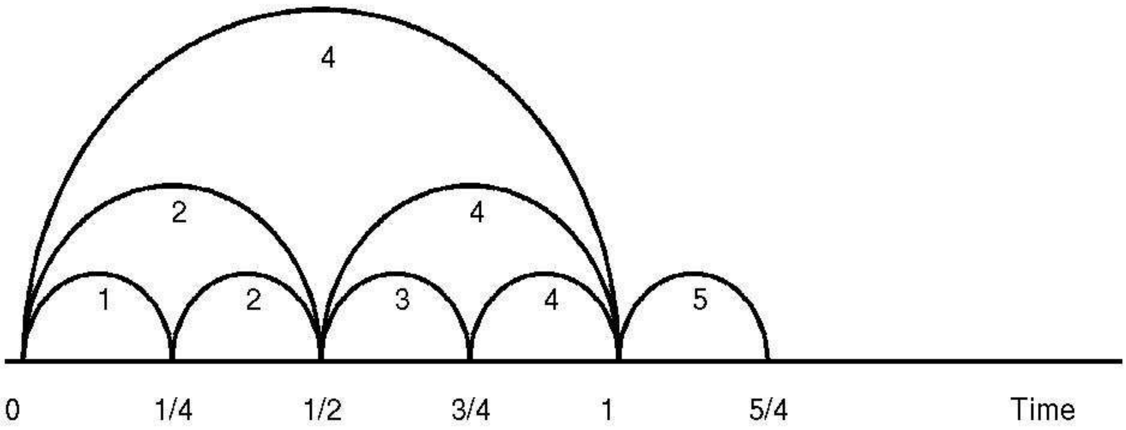

Figure 2.3: Diagram of the block time step scheme. Suppose there are three time step Δt levels: $\Delta t = 1/4, 1/2$ and 1. At time $t = 0$, the smallest new time is $t = 1/4$, so the first block of particles is integrated from $t = 0$ to $t = 1/4$. Then, the smallest new time is $t = 1/2$, and the second block of particles is integrated: some from $t = 0$ to $t = 1/2$ and some from $t = 1/4$ to $t = 1/2$. This is repeated at all time steps. It is only at the instants of time which are integer multiples of the largest time step level (in this case, $\Delta t = 1$) that all the particles in the simulations are synchronized. At all other times, if output is needed or for some other reason, it is necessary to make some prediction.

Chapter 3

Stellar cusps around a massive black hole in galactic nuclei: Theory

3.1 Introduction

The question of how will the stellar distribution around an (intermediate-)massive black hole (I)MBH evolve has been originally addressed in the 1970's following the detection of X-ray sources in globular clusters. These observations inspired the hypothesis that IMBHs, of masses ranging from $10^2 M_\odot$ to $10^4 M_\odot$, could reside in the center of globular clusters [Peebles, 1972, Bahcall and Wolf, 1976, Lightman and Shapiro, 1977, Cohn and Kulsrud, 1978].

From the beginning, people have searched for collisional equilibria since globular cluster are significantly older than the estimated relaxation times of their cores, $t_{\text{rlx}} \sim 10^8$ yrs, so these systems should have had enough time to reach a steady state independently of what would have been their initial conditions at the time of formation. Peebles 1972 has derived, using a simple scaling argument, a stationary density law $\rho(r) \propto r^{-9/4}$; this result was subsequently criticized by Bahcall & Wolf 1976 which showed that it implied a violation of any physically reasonable boundary conditions — the black hole would behave as a matter source, generating a large net stellar flux outwards, rather than as a sink for matter as it should. Using the Fokker-Planck formalism, Bahcall & Wolf 1976 predicted that the solution was rather $\rho(r) \propto r^{-7/4}$ (or $f(E) \propto E^{1/4}$), corresponding to an asymptotic steady-state with a net quasi-zero stellar flux. As this was a one-dimensional calculation (in E-space, neglecting any dependence on J), it took no account of any loss cone effects. In their following paper on the subject, Bahcall & Wolf 1977 estimated, through an approximate, but still one-dimensional calculation, that the inclusion of a loss cone sink term would not lead to significant changes to the '7/4-cusp' solution over most of the relevant range ($E_h \lesssim E \ll E_t$ in energy, or $r_t \ll r \lesssim r_h$ in radius). Shapiro & Lightman 1977 and Cohn & Kulsrud 1978, in a more detailed two-dimensional, in (E, J) space calculation, did a boundary layer analysis of the loss cone effects and were able to confirm the preliminary results from the previous authors.

After this series of seminal papers, several authors have studied this problem further with a higher level of detail and complexity, by the inclusion of other physical effects — such as stellar formation and evolution, stellar collisions, and so on —, but mainly through Monte-Carlo calculations that made the same simplifying assumptions as the Fokker-Planck formalism on which, in fact, they are based [Marchant and Shapiro, 1980, Murphy et al., 1991, Freitag and Benz, 2002]. It has been therefore an open problem since then how to assess the validity of the Fokker-Planck approximations in this context, and how their eventual breakdown would affect the solution properties. For instance, Lin & Tremaine 1980 have suggested that strong encounters (which are not taken into account in any of the Fokker-Planck based calculations) could lead to a significant ejection rate of stars from the stellar Bahcall-Wolf cusp around the central black hole. They have also estimated that this ejection rate could be, at least, an order of magnitude higher than the rate at which stars are tidally disrupted, or consumed, by the black hole. Moreover, it was still remained to be shown convincingly that the Bahcall-Wolf 7/4 cusp could be recovered from a more exact calculation such as a direct N-body calculation which did not rely in any of the simplifying assumptios inherent to the Fokker-Planck formalism. One of the goals of this work is to show that the Bahcall-Wolf cusp is recovered consistently with direct N-body techniques and it is therefore a robust theoretical prediction for stellar systems which are old enough with respect to their relaxation time.

3.2 Basic quantities

The Schwarschild radius of a non-rotating black hole of mass $M_\bullet$ is given by

$$r_S = \frac{2GM_\bullet}{c^2} \approx 9.5 \times 10^{-8}\text{pc}\left(\frac{M_\bullet}{10^{-6}M_\odot}\right). \tag{3.1}$$

The gravitational potential of the MBH is dominant inside of its so-called influence radius r_h

$$r_h = \frac{GM_\bullet}{\sigma^2} \approx 0.43\text{pc}\left(\frac{M_\bullet}{10^{-6}M_\odot}\right)\left(\frac{100\text{Km/s}}{\sigma}\right)^2, \tag{3.2}$$

where σ is the $1D$-velocity dispersion. In the case of a singular isothermal sphere (SIS), $\rho(r) = \sigma^2/2\pi G r^2$, the velocity dispersion $\sigma =$ cte. and the stellar mass within r_h is $2M(< r_h)$. For this reason, and since the definition in equation (3.2) is not practical as σ^2 depends on r, it is costumary to use an alternative definition of r_h: it is the radius which encloses a stellar mass equal to $2M_\bullet$,

$$M_*(< r_h) = 2M_\bullet. \tag{3.3}$$

This is equivalent to the definition (3.2) in the case of a SIS. The $1D$-velocity dispersion inside the radius of influence scales as: $\sigma^2 \sim GM_\bullet/r$.

If stars approach very close to the MBH, they may be broken apart by the very strong tidal forces it generates. Although the precise radius at which this destruction happens depends on the orbital shape and on the star's internal structure, the tidal disruption radius r_t is given approximately by

$$r_t \sim \left(\frac{M_\bullet}{m_*}\right) r_* \approx 2.15 \times 10^{-6}\text{pc} \left(\frac{M_\bullet}{10^6 M_\odot}\right), \tag{3.4}$$

where the numerical values were estimated for a star of Solar mass and radius. The ratio to the hole's event horizon is

$$\frac{r_t}{r_S} \approx 22.7 \left(\frac{M_\bullet}{10^6 M_\odot}\right)^{-2/3}, \tag{3.5}$$

meaning that a MBH like that residing in the Milky Way's center can effectively tidal disrupt stars of the Solar type if they approach sufficiently close. On the other hand, black holes with $M_\bullet \gtrsim 10^8 M_\odot$ will swallow Solar type stars whole without disrupting them.

The dynamical time scale is the time a star takes to complete one orbit

$$t_{dyn} \sim 2\pi\frac{r}{v} \sim 2\pi\sqrt{\frac{r^3}{GM(< r)}} \sim 9.44 \times 10^4\text{yr} \left(\frac{r}{1\text{pc}}\right)^{3/2} \left(\frac{10^6 M_\odot}{M}\right)^{1/2}, \tag{3.6}$$

where r is the size of the orbit and $M(< r)$ is the total mass enclosed within r.

The relaxation time is the average time it takes for a typical star to change its initial velocity by $\Delta v \approx v$ or, equivalently, the average time for the succession of weak, uncorrelated gravitational encounters between stars to establish locally a Maxwellian velocity distribution. For the case of a spatially homogeneous distribution of equal mass stars, the relaxation time will be of order

$$\begin{aligned} t_{\text{rlx}} \;&\sim\; 0.34\frac{\sigma^3}{G^2 \rho m_* \ln\Lambda} \approx \\ & 5.4 \times 10^9\text{yrs.} \left(\frac{\sigma}{100Km/s}\right)^3 \left(\frac{10^6 M_\odot/pc^3}{\rho}\right) \left(\frac{M_\odot}{m_*}\right) \left(\frac{10}{\ln\Lambda}\right). \end{aligned} \tag{3.7}$$

Note that roughly $t_{\rm rlx} \sim \sigma^3/\rho \sim r^{\gamma-3/2}$ in a stellar cusp with a density profile such as $\rho \sim r^{-\gamma}$ around the MBH. The Milky Way, as well as several other galaxies in the Local Group are known to have central stellar cusps with slopes $\gamma \approx 3/2$ for $r \lesssim r_h$; therefore, $t_{\rm rlx}$ in these spheroids should be approximately constant with radius inside their hole's sphere of influence.

The stellar clusters that reside at the centers of galactic nuclei can be classified according to their age $t_{\rm age}$ and their relaxation time $t_{\rm rlx}$. In the case where $t_{\rm age} \ll t_{\rm rlx}$, the cluster still retains some memory of its formation process; we say it is a *collisionless* cluster. This is the case for most of the confirmed MBH detections, for which $t_{\rm rlx} > t_{\rm Hubble}$. On the other limit, $t_{\rm age} \geq t_{\rm rlx}$, the system has forgotten its *initial conditions*, the typical stars have had enough time to diffuse appreciably in phase space, and for its macroscopic properties to have evolved significantly; we call it a *collisional* cluster. This is the case of smaller MBHs - e.g. in the Milky Way, M32, and many expected small, compact galactic nuclei predicted according to the hierarchical models for structure formation ([Volonteri et al., 2003b, Sesana et al., 2007]). The study of these *collisional* stellar clusters around MBHs is central in the planned efforts for the detection and analysis of gravitational waves from *extreme mass ratio inspirals* (EMRIs) and from MBH binary's inspirals and mergers [Hopman and Alexander, 2006, Amaro-Seoane et al., 2006].

In an idealized smooth and spherical potential, the energy E and angular momentum J are integrals of motion for each star. However, any finite realization of a smooth potential by a discrete set of particles (stars) will suffer from energy (and angular momentum) relaxation — as the stars *see each other*, thereby exchanging gradually energy and angular momentum, in a dynamical process that leads to the secular evolution of the macroscopic properties of the stellar cluster. The relaxation time $t_{\rm rlx}$ is a measure of the average time it takes for a typical star to change its velocity by $\Delta v/v \approx 1$, through a series of a very large number of *uncorrelated* weak encounters (each of which obeys $\Delta v/v \ll 1$). The relaxation process is reminiscent of the behavior of a gas of molecules. However, the details are significantly different: in the case of a molecular gas, the intermolecular forces are strongly repulsive at very short-range and negligible otherwise; while the gravitational forces between stars are very long-range. Therefore, while the gas molecules travel freely with constant velocity between sudden brief and strong accelerations when they collide with another molecule, the accelerations felt by a typical star do not vary rapidly and its dynamics is determined to 0^{th} *order* by the force field generated by the stellar cluster as a whole. This means that to 0^{th} *order*, we can model the gravitational field acting on a star as if it originated from a smooth density distribution rather than from a set of point masses. The question of knowing how accurate is it to model a stellar cluster as being composed by a smooth mass distribution and a collective force field is solved by the classification of the stellar systems as being collisionless.

On Figure (3.1) it is shown a clear trend of $t_{\rm rlx}$ with the galactic bulge's luminosity. The relaxation times are typically larger than the Hubble time for galactic spheroids with absolute magnitude higher than $M_V \approx -18$ — approximately the Milky Way's bulge luminosity. This reflects the well-known observational fact that the nuclei of brighter galaxies have stellar cusps with shallower slopes, and thus with larger relaxation times. The bulges with heavier black holes — precisely those bulges that are more massive and more luminous, according to the observational correlation between $M_\bullet$ and M_{bulge} — are more likely to have cusps that are less steep and hence to be in the collisionless regime, still retaining the memory of their formation history.

The formation of massive black hole binaries may be very common when two galaxies merge, each with a massive black hole lurking at its center. Let us define here some important quantities associated to massive black hole binaries which will be studied in Chapter 5 of this work. Denote by $M_{\bullet 1}$ and $M_{\bullet 2}$ the masses of the MBH binary components, $\mu = G(M_{\bullet 1}+M_{\bullet 2})$, and the reduced mass $\mu_{red} = M_{\bullet 1}M_{\bullet 2}/(M_{\bullet 1}+M_{\bullet 2})$. The binary's specific (per unit reduced mass) binding energy is given by

$$E_b = \frac{G(M_{\bullet 1}+M_{\bullet 2})}{r} - \frac{1}{2}v^2 = \frac{GM_{\bullet 1}M_{\bullet 2}}{2a} = \frac{\mu}{2a}, \tag{3.8}$$

where a is the binary's semimajor axis. The orbital period P of the massive binary is

$$P = 2\pi\left(\frac{a^3}{\mu}\right)^{1/2} \approx 9.37\times 10^4\text{yrs.}\left(\frac{10^6 M_\odot}{M_{12}}\right)^{1/2}\left(\frac{a}{1\text{pc}}\right)^{3/2}, \tag{3.9}$$

where $M_{12} = M_{\bullet 1} + M_{\bullet 2}$. The relative (circular) velocity is

$$v_c = \sqrt{\frac{\mu}{a}} = 65.7\text{km/s}\left(\frac{M_{12}}{10^6 M_\odot}\right)^{1/2}\left(\frac{1\text{pc}}{a}\right)^{1/2}. \tag{3.10}$$

The massive black hole binary becomes a *hard binary* once its specific energy E_b exceeds the $1D$- velocity dispersion σ_{1D}^2 of the stellar field in the surrounding nuclei. This can be recasted in terms of a critical semimajor axis a_h below which the binary is *hard*:

$$a \lesssim a_h \equiv \frac{\mu}{8\sigma_{1D}^2} \approx 0.054\text{pc}\left(\frac{M_{12}}{10^6 M_\odot}\right)\left(\frac{100\text{km/s}}{\sigma}\right)^2. \tag{3.11}$$

The radius of influence r_h of the binary can be expressed in terms of the mass of its two components as follows:

$$r_h \equiv \frac{2\mu}{\sigma(r_h)^2} \approx 0.86\text{pc}\left(\frac{M_{12}}{10^6 M\odot}\right)\left(\frac{100\text{km/s}}{\sigma}\right)^2. \tag{3.12}$$

Or, alternatively, a more practical definition for r_h is as the radius at which the enclosed stellar mass is twice the MBH binary total mass:

$$M(< r_h) \equiv (1-2)M_{12}. \tag{3.13}$$

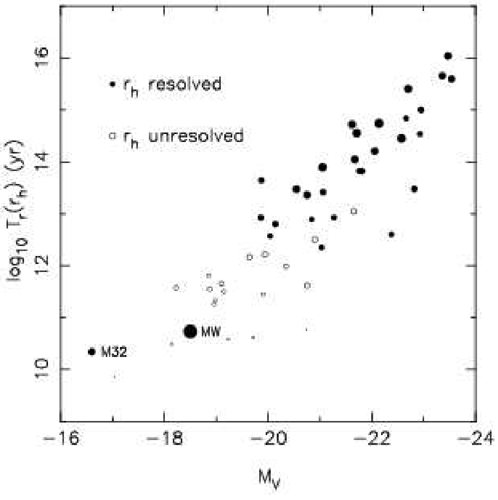

Figure 3.1: Estimates of the relaxation time $t_{\rm rlx}$ at the influence radius of the vblack hole as a function of the absolute visual magnitude of the galaxy (the stellar bulge, in the case of the Milky Way). It is only for the spheroids fainter than absolute magnitude $M_V \sim 18$ that the $t_{\rm rlx}$ drops below the Hubble time. (figure from [Merritt and Szell, 2006]).

A number of approximations and simplifications are made in this work, some in the Fokker-Planck, others in the N-body calculations, and some on both. For instance, the loss cone effects are neglected: in terms of the Fokker-Planck equation they would require to make the integration in (E, J) instead of E only; in the N-body runs, it would be necessary to define a capture radius and absorb the particles that would cross it. This is relevant in case one wants to estimate the capture rate of objects by the MBH, and therefore is out of the scope of this work. We have adopted spherically symmetric galactic models which is the adequate first step, but the inclusion of axisymmetric, rotating or triaxial models will be relevant in the future (although, in the case of triaxiality, the Fokker-Planck equation is unable to cope with it). A lot of complicated physics is left out: star formation and evolution, star collisions, gas dynamics, massive clumps such as giant molecular clouds or globular clusters, and so on. A realistic modelling of all these effects and components would be not only out of the current computational and algorithmic capabilities, but its results would also be extremely difficult to interpret. We believe that simplicity is a main virtue in the art of modelling and thus prefer to restrict ourselves to well defined problems until some fundamental breakthrough — be it either conceptual, algorithmic or in hardware — opens the way to a higher level of complexity.

3.3 Heuristic derivation of the Fokker-Planck equation and its assumptions

Under the influence of the smooth potencial $\Phi(\mathbf{x})$, the DF $f(\mathbf{x}, \mathbf{v})$ evolves according to the collisionless Boltzmann equation (*CBE*)

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} - \nabla\Phi \cdot \frac{\partial f}{\partial \mathbf{v}} = 0, \tag{3.14}$$

which physically means that the probability density in phase space around each star is conserved over time. The *CBE* can be derived from the Liouville equation in Γ-space [Binney and Tremaine, 1987] provided that the N-body DF is *separable* — or, equivalently, that the n-point correlation function vanishes, for all n —, and physically this means that the star's locations in phase space are completeley uncorrelated so that the probability for finding one star at position $\mathbf{x}$ and velocity $\mathbf{v}$ is completely independent from the location of all the other stars. This constitutes the *0th-order approximation* to the stellar dynamics.

When gravitational encounters between stars are taken into account, the phase space density around each star evolves at a rate determined by the collision operator $\Gamma(f) \equiv \left(\frac{\partial f}{\partial t}\right)_{enc}$. * The general expression that governs the DF f($\mathbf{x}$,$\mathbf{v}$) is given schematically by

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} - \nabla\Phi \cdot \frac{\partial f}{\partial \mathbf{v}} = \begin{cases} 0, & t_{\rm age} \ll t_{\rm rlx} \\ \left(\frac{\partial f}{\partial t}\right)_{\rm enc}, & t_{\rm age} \geq t_{\rm rlx} \end{cases} \tag{3.15}$$

The DF provides a *macroscopic* description of the stellar system; in contrast, the N-body calculations provides the full information about the *microscopic* dynamics on the basis of its evolution. The time evolution of the DF caused by the gravitational encounters is governed by the *Master equation*

$$f(\mathbf{v}, t) = \int f(\mathbf{v} - \Delta\mathbf{v}, t) R(\mathbf{v} - \Delta\mathbf{v}, \Delta\mathbf{v}, t)\ d(\Delta\mathbf{v}), \tag{3.16}$$

where $R(\mathbf{v} - \Delta\mathbf{v}, \Delta\mathbf{v}, t) d(\Delta\mathbf{v})$ represents the transition probability, per unit time, that a star with velocity $\mathbf{v}$ will experience a velocity change $\Delta\mathbf{v}$. The Master equation is an integro-differential equation and it is, in general, very difficult to solve - even numerically. In order to derive the more tractable Fokker-Planck equation, we start by expanding both sides of

---

*But not including the accelerations from the smooth component of the potential — they are already included in the left hand side of the Boltzmann equation.

the Master equation in Taylor series :

$$
\begin{aligned}
f(\mathbf{x},\mathbf{v}) + \frac{\partial f(\mathbf{v},t)}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2 f(\mathbf{v},t)}{\partial t^2}(\Delta t)^2 + \mathcal{O}(3) = \\
\int d(\Delta \mathbf{v})\left(fR - \frac{\partial}{\partial v_i}(fR)\Delta v_i + \frac{1}{2}\frac{\partial^2 (fR)}{\partial v_i v_j}\Delta v_i \Delta v_j + \mathcal{O}(3)\right).
\end{aligned}
\tag{3.17}
$$

We then divide by Δt and take the limit when $\Delta t \to 0$ to obtain an expression for the Fokker-Planck collision operator

$$
\left(\frac{\partial f}{\partial t}\right)_{enc} = -\frac{\partial}{\partial v_i}\left(f\langle \Delta v_i\rangle_t\right) + \frac{1}{2}\frac{\partial^2}{\partial v_i \partial v_j}\left(f\langle \Delta v_i \Delta v_j\rangle_t\right) + \mathcal{O}(3),
\tag{3.18}
$$

where the diffusion coefficients are given by

$$
\begin{aligned}
\langle \Delta v_i\rangle_t &= \lim_{\Delta t\to 0}\frac{\langle \Delta v_i\rangle}{\Delta t} \\
\langle \Delta v_i \Delta v_j\rangle_t &= \lim_{\Delta t\to 0}\frac{\langle \Delta v_i\rangle\langle \Delta v_i\rangle}{\Delta t},
\end{aligned}
\tag{3.19}
$$

and

$$
\begin{aligned}
\langle \Delta v_i\rangle &= \int R(\mathbf{v},\Delta\mathbf{v})\Delta v_i d(\Delta\mathbf{v}) \\
\langle \Delta v_i \Delta v_j\rangle &= \int R(\mathbf{v},\Delta\mathbf{v})\Delta v_i \Delta v_j d(\Delta\mathbf{v}).
\end{aligned}
\tag{3.20}
$$

The terms of third order $\langle \Delta v_i \Delta v_j \Delta v_k\rangle$ and higher order were neglected, since they are smaller by order of the Coulomb logarithm $\ln\Lambda$. † It is in this truncation of the full series development that that enters the assumption regarding the small deflections.

The diffusion coefficients measure the rate at which the stars diffuse in phase space. Note that, somewhat unintuitively, the first two lower-order terms are usually of the same order of magnitude [Binney and Tremaine, 1987, Spitzer, 1987]. The second-order term $\langle \Delta v_i \Delta v_j\rangle$ determines the rate at which stars undergo a *random walk* in phase space; while $\langle \Delta v_i\rangle$ represents the rate at which stars drift in phase space — which is usually called dynamical friction.

---

† $\ln\Lambda$ is usually not very large, $\sim 7-15$ for $N \sim 10^5 - 10^8$; [Goodman, 1983] has shown that they do not change appreciably the results.

The precise expressions for the diffusion coefficients can be derived through straightforward, but lengthy, calculations (see appendix of Binney & Tremaine 1987). The result is:

$$D(\Delta v_i) = 4\pi^2 G m_f (m_f + m_t) \ln\Lambda \, \frac{\partial}{\partial v_{t,i}} h(\mathbf{v}_t) \tag{3.21}$$

$$D(\Delta v_i \Delta v_j) = 4\pi^2 G m_f^2 \ln\Lambda \, \frac{\partial^2}{\partial v_{t,i}\partial v_{t,j}} g(\mathbf{v}_t), \tag{3.22}$$

where the functions f and g are the well-known Rosenbluth potentials [Rosenbluth et al., 1957]

$$h(\mathbf{v}_t) = \int d\mathbf{v}_t \, \frac{f(\mathbf{v}_t)}{|\mathbf{v}_t - \mathbf{v}_f|} \tag{3.23}$$

$$g(\mathbf{v}_t) = \int d\mathbf{v}_t \, f(\mathbf{v}_t) |\mathbf{v}_t - \mathbf{v}_f|. \tag{3.24}$$

In the case of an isotropic distribution of velocities, the stellar DF depends on $v_f \equiv |\mathbf{v}_f|$ only, and the diffusion coefficients can be expressed in a simpler form. Taking the direction parallel to the test star's velocity $\mathbf{v}_t$:

$$D(\Delta v_\|) = 4\pi^2 G^2 m_f (m_f + m_t) \ln\Lambda \frac{\partial h(v_t)}{\partial v_t}, \tag{3.25}$$

where the Rosenbluth potential h for an insotropic velocity distribution is

$$h(v_t) = 4\pi \left[\frac{1}{v_t} \int_0^{v_t} dv_f \; f(v_f) v_f^2 + \int_{v_t}^{+\infty} dv_f \; f(v_f) v_f \right]. \tag{3.26}$$

Obviously, by symmetry considerations, the first order diffusion coefficient vanishes in the directions perpendicular to the test star's velocity

$$D(\Delta v_\perp) = 0. \tag{3.27}$$

The other diffusion coefficient is given by

$$D(\Delta v_i \Delta v_j) = 4\pi^2 G m_f^2 \ln\Lambda \left[\frac{\delta_{ij}}{v_t} \frac{\partial g(v_t)}{dv_t} + \frac{v_{t,i} v_{t,j}}{v_t^2} \left(\frac{\partial^2 g(v_t)}{dv_t^2} - \frac{1}{v_t} \frac{\partial g(v_t)}{dv_t} \right) \right]. \tag{3.28}$$

Taking $i = j$ to be the direction parallel to $\mathbf{v}_t$, then

$$D(\Delta v_{\|}^2) = \langle (\Delta v_{\|})^2 \rangle = 4\pi^2 G m_f^2 \ln\Lambda \frac{\partial^2 g(v_t)}{dv_t^2}, \tag{3.29}$$

but if $i = j$ is taken in the plane perpendicular to the test star's velocity, then

$$D(\Delta v_{\perp}^2) = \langle (\Delta v_{\perp})^2 \rangle = 2 \times 4\pi^2 G m_f^2 \ln\Lambda \frac{\partial g(v_t)}{dv_t}, \tag{3.30}$$

where the extra factor 2 results from the fact that $D((\Delta v_x)^2) = D((\Delta v_y)^2) = 1/2 D((\Delta v_{\perp})^2)$. The Rosenbluth potential g for the case of an isotropic velocity distribution reads

$$g(v_t) = \frac{4\pi v_t}{3}\left[\int_0^{v_t} dv_f \left(3v_f^2 + \frac{v_f^4}{v_t^2}\right) f(v_f) + \int_{v_t}^{+\infty} dv_f \left(\frac{3v_f^3}{v_t} + v_t v_f\right) f(v_f)\right]. \tag{3.31}$$

In order to derive the Fokker-Planck equation, a number of simplifying assumptions were made:

A1. The scattering is dominated by a succession of weak ($\Delta v/v \ll 1$), uncorrelated gravitational encounters and the effects from strong encounters ($\Delta v/v \gtrsim 1$) are neglected;

A2. The *local approximation* assumes most of the scattering is due to encounters with impact parameter b much smaller than the system's size ($b \ll R_{\text{size}}$). This implies that the transition probability that governs the diffusion rate obeys the condition $R(\mathbf{v} - \Delta\mathbf{v}, \Delta\mathbf{v}, \Delta\mathbf{x} \neq 0; t) = 0$, i.e. each encounter produces a small kick in the star's velocity, but leaves its position unchanged since the encounters occurs during the time $\sim b/v \ll t_{\text{dyn}}$; this fact, in turn, implies that the diffusion coefficients $D(\Delta x_i) = D(\Delta x_i \Delta x_j) = D(\Delta x_i \Delta v_j) = 0$ vanish identically. Therefore, when the diffusion coefficients are computed explicitly they can be treated as if, during each gravitational encounter, the stars moved along Kepler hyperbolae unaffected by the potential from the remaining stars.

A3. The DF changes very little during a dynamical time, its evolution is slow and occurs over a relaxation time scale. Therefore, for computational efficiency, we want to *average out* the fast changes felt by a star along each orbit in the smooth potential from the slow changes (accumulating over many orbits) due to the two-body encounters. So, we compute and employ the *orbital-averaged* diffusion coefficients, otherwise these would be dependent on the star's position along its orbit and would increase the dimensionality of the phase space necessary for the analysis.

A4. Triple and higher-order encounters, being of lower probability, are neglected.

A5. The Master equation and, consequently, the Fokker-Planck equation are not time-reversible unlike the microscopic system of stars, evolving under the Newton's laws of motion. [‡] In fact, this situation is completely analogous to the derivation of the Boltzmann equation for the evolution of the DF of a gas of molecules. The irreversibility is implicitly introduced by the assumption of uncorrelated encounters — the molecular chaos assumption — which is equivalent to assuming that the transition probability R and f are statistically independent, so the scattering rate $\Gamma(f)$ can be obtained as the product of two (independent) probabilities as fR.

3.4 Fokker-Planck equation in energy space

According to Jeans theorem [Binney and Tremaine, 1987,Spitzer, 1987], the DF of a spherical galaxy with an isotropic velocity distribution must depend on the star's energy $E = -\frac{1}{2}v^2 - \Phi(\mathbf{x})$ only. Since $f(\mathbf{x}, \mathbf{v})$ denotes the number of stars with position $\mathbf{x}$ and velocity $\mathbf{v}$, we can compute the number of stars with energy E, $N(E)$, by

$$\begin{aligned} N(E) &= \int d\mathbf{x}\; d\mathbf{v}\; f(E)\; \delta(E - \frac{1}{2}v^2 - \Phi(\mathbf{x})) = \\ &= 16\pi^2 \int dr\; r^2 \int dv^2 f(E)\; \delta(\frac{1}{2}v^2 + \Phi - E) = \\ &= 16\pi^2 f(E) \int_0^{r_{max}(E)} dr\; r^2 \sqrt{2(E - \Phi(r))} = 16\pi^2 p(E) f(E), \end{aligned} \tag{3.32}$$

where the radial range available for a star of energy E goes from zero to $r_{max}(E)$ — the latter being a root of the equation $E = \Phi(r_{max})$, for which the star's radial velocity vanishes. The quantity $p(E)$ denotes the surface enclosing the region of phase space with volume $q(E)$ available for a star of energy E:

$$\begin{aligned} p(E) &= \int_0^{r_{max}(E)} dr\; v\, r^2 = \int_0^{r_{max}(E)} dr\, r^2 \sqrt{2(E - \Phi)} \\ q(E) &= \frac{1}{3} \int_0^{r_{max}(E)} dr\; v^3\, r^2 = \frac{1}{3} \int_0^{r_{max}(E)} dr\, r^2\, (2(E - \Phi))^{3/2}\,. \end{aligned} \tag{3.33}$$

Note that $p = \partial q / \partial E$.

We can derive, by analogy with the last section, the Fokker-Planck equation in energy space; being one-dimensional the energy space is the most appropriate to study the dynamics of spherical clusters with isotropic velocity distributions. Following Spitzer (1987), we can

---

[‡] The reversibility will be lost also at the microscopic level once we introduce relativistic effects to the dynamics - see Chapter 5.

compute $(\partial N(E)/\partial t)_{enc}$ by defining $\Psi(E,\Delta E)d(\Delta E)$ as the probability, per unit time, that a star of energy E undergoes an energy change ΔE within the range $d(\Delta E)$. Perform again a series expansion of the Master equation this time written in energy space — rather than velocity space, as before —, to obtain an expression for the Fokker-Planck collision operator in energy space:

$$\left(\frac{\partial N}{\partial t}\right)_{enc} = -\frac{\partial}{\partial E}\left(N\langle\Delta E\rangle_V\right) + \frac{1}{2}\frac{\partial^2}{\partial E^2}\left(N\langle(\Delta E)^2\rangle_V\right). \tag{3.34}$$

The orbit-averaged diffusion coefficients are averages over all stars of energy E within the available phase space volume V, and are given by the following expressions

$$\begin{aligned}\langle\Delta E\rangle_V &= \frac{\int_0^{r_{max}} dr\langle\Delta E\rangle v r^2}{\int_0^{r_{max}} drvr^2}\\ \langle(\Delta E)^2\rangle_V &= \frac{\int_0^{r_{max}} dr\langle(\Delta E)^2\rangle v r^2}{\int_0^{r_{max}} drvr^2}.\end{aligned} \tag{3.35}$$

In energy space, the Fokker-Planck equation takes the form of the conservation of particle number in energy space

$$\begin{aligned}4\pi^2 p(E)\frac{\partial f}{\partial t} &= -\frac{\partial R(E)}{\partial E}\\ R(E,t) &= -D_E f - D_{EE}\frac{\partial f}{\partial E},\end{aligned} \tag{3.36}$$

and the orbit-averaged diffusion coefficients are given by

$$\begin{aligned}D_E(E) &= -64\pi^2 G^2 m_*^2 \ln\Lambda \int_E^{+\infty} dE'\, p(E')f(E')\\ D_{EE}(E) &= 64\pi^2 G^2 m_*^2 \ln\Lambda\left[q(E)\int_{-\infty}^{E} dE'\, f(E') + \int_E^{+\infty} dE'\, q(E')f(E')\right].\end{aligned} \tag{3.37}$$

In order to relate the diffusion coefficients expressed in terms of the binding energy E to those expressed in terms of the velocity, it is useful to note that under the local approximation two-body encounters last much less than a dynamical time, $t_{enc} \ll t_{\rm dyn}$,

$$\begin{aligned}\Delta r &= 0, \quad \text{which means } \Phi_{before} = \Phi_{after}\\ &\Delta v \neq 0.\end{aligned}$$

If the test's star velocity before the encounter is

$$\mathbf{v}_t = v_t \hat{e}_1 = v_\parallel \hat{e}_1, \tag{3.38}$$

then

$$\begin{aligned} E_0 &= \frac{1}{2} v_t^2 = \frac{1}{2} v_\parallel^2 + \Phi(r) \\ E_0 + \Delta E &= \frac{1}{2} \left(v_\parallel + \Delta v_\parallel \right)^2 + \frac{1}{2} \Delta v_\perp^2 + \Phi(r). \end{aligned} \tag{3.39}$$

Hence

$$\Delta E = \frac{1}{2} (\Delta v_\parallel)^2 + \frac{1}{2} (\Delta v_\perp)^2 + v_\parallel \Delta v_\parallel, \tag{3.40}$$

and therefore

$$\begin{aligned} \langle \Delta E \rangle &= \frac{1}{2} \langle (\Delta v_\parallel)^2 \rangle + \frac{1}{2} \langle (\Delta v_\perp)^2 \rangle + v_\parallel \langle \Delta v_\parallel \rangle \\ \langle (\Delta E)^2 \rangle &= \frac{1}{4} \langle (\Delta v_\parallel)^4 \rangle + \frac{1}{4} \langle (\Delta v_\perp)^4 \rangle + 2 \langle (\Delta v_\perp)^2 \rangle \langle (\Delta v_\parallel)^2 \rangle + \\ &+ v_\parallel^2 \langle (\Delta v_\parallel)^2 \rangle + \langle (\Delta v_\perp)^2 \rangle \langle \Delta v_\parallel \rangle = \\ &= v_\parallel^2 \langle (\Delta v_\parallel)^2 \rangle + \mathcal{O}(3). \end{aligned}$$

Finally to derive the diffusion coefficients in energy space: start by substituting (3.21), (3.29) and (3.30) into these equations; then, orbit-average the resulting expressions according to (3.35), to obtain the orbit-averaged diffusion coefficients in energy space as defined by equation (3.37).

3.5 The Bahcall-Wolf cusp

We want to determine the steady-state solution for the stellar distribution around a massive black hole. We start by making an *ansatz* for the solution well inside the blach hole's sphere of influence, $r \ll r_h$:

$$f(E) \propto E^p, \qquad \rho(r) \propto r^{-(p+3/2)}. \tag{3.41}$$

Furthermore, in order to solve the Fokker-Planck equation, being a partial differential equation, it is necessary to specify the boundary conditions. We choose flux-free boundary conditions: the inner boundary condition is $f(E_t) = 0$ where $E_t \sim GM_\bullet r_t$ is the tidal energy that corresponds to the tidal radius at which stars are either disrupted by the tidal force or consumed by the black hole. We will also consider the limit case when $E \to +\infty$. The outer boundary condition corresponds to the value of the DF at some low value of the binding energy to the stellar cluster and the black hole, $f(E_0) = f_0$ — the precise value will depend on the adopted model, usually a Dehnen model.

At steady-state, $\partial f/\partial t = 0$, meaning the stellar flux $R(E,t)$ from (3.36) is a constant independent of time and of energy. If we substitute our *ansatz* in the expression for $R(E,t)$, we arrive at an algebraic equation and solve it for the exponent p. Thus:

$$R(E) \propto \tag{3.42}$$

$$-\frac{E}{16}\int_E^{+\infty} dE'(E')^{p-5/2} + \frac{p}{24}\left[E^{-3/2}\int_{-\infty}^{E} dE' E'^p + \int_E^{+\infty} dE'(E')^{p-3/2}\right] = \text{const},$$

hence

$$R(E) \propto -\frac{E}{16}\frac{E^{p-3/2}}{p-3/2}\bigg|_E^{+\infty} + \frac{p}{24}\left(E^{-3/2}\frac{E^{p+1}}{p+1}\bigg|_{-\infty}^{E} + \frac{E^{p-1/2}}{p-1/2}\bigg|_E^{+\infty}\right) = \text{const}. \tag{3.43}$$

We can immediately see that convergence in the limit $E \to +\infty$ requires the exponent to be $p < 1/2$. It is, of course, also clear from the physical point of view that $E_t < +\infty$; however, violating this condition on the value of the exponent p would lead nevertheless to an unphysically large negative stellar flux at the highest energies contradicting the nature of the black hole as a sink for matter. On the other hand, if we set the (constant) stellar flux to zero $R(E) = 0$, in the limit where $E \to +\infty$, we get a precise value of $p = 1/4$ for the exponent and therefore $f(E) \propto E^{1/4}$ is called the *zero-flux* solution. In the more realistic case where the tidal energy $E_t < +\infty$ is very high, but finite, we can see from the Figure 3.3 that the numerical solution of the Fokker-Planck equation adapts itself as close as possible to the zero-flux solution in a way that agrees with the inner boundary condition — which turns to be an excellent approximation within the energy range $E_h \ll E \ll E_t$ and we call this the *nearly zero-flux* solution.

An isothermal distribution , $f \propto e^{E/\sigma^2}$, would also translate into a zero-flux solution but, for the same reasons as before, doesn't satisfy the inner boundary condition. The nearly-zero flux solution has a small steady-state stellar flux given by

$$R(E) \sim \frac{N(<E_t)}{t_{\text{rlx}}(E_t)} \propto r_t, \tag{3.44}$$

since the flux is fixed by the rate at which stars diffuse at the tidal energy E_t (tidal radius r_t). At steady-state, this nearly-zero net flux results, of course, from the near-cancellation of much larger fluxes with both signs at all energies, the exception being at the tidal energy where, by definition, there is no outward flux. It is thus the inner boundary that breaks the symmetry between inner and outward fluxes; as a result when $E_t \to +\infty$, the assymetry becomes smaller and smaller, until in the limit reduces to zero and we recover the exact zero-flow solution. For completeness, we should mention that the steady-state condition can also be met with the exponent $p = 3/4$ — the solution found by Peebles (1972) —, but this physically unnacceptable because it entails a huge negative stellar flux away from the black hole turning it into a matter source. During the transient process, while scattering proceeds but steady-state has not yet been reached, there is a larger net inward stellar flux, so long as the initial cusp is less steep than the $p = 1/4$ and $\gamma = 7/4$ solution, and vice-versa if it is steeper. The formation of the Bahcall & Wolf cusp occurs over a time period of roughly a relaxation time as measured at the (initial) black hole's radius of influence r_h, on longer time scales the stellar cusp has effectively forgotten its initial conditions.

Since the gravitational potential changes as the stellar distribution evolves, equation (3.36) should also contain a term describing the adjustement in f due to changes in E as the gravitational potential evolves (e.g. Spitzer 1987, eq. 2-86). However changes in $\rho(r)$ only take place well within r_h in our simulations, and the potential is dominated by the fixed (no mass accretion to the black hole) mass of the black hole in this region. Significant changes to the gravitational potential happen on a time scale much longer than that of our integration (of roughly a single relaxation time at the hole's sphere of influence), and note that even for the smallest and densest galactic nuclei this time is of the order of the Hubble time. Hence, assuming a fixed potential in the Fokker-Planck equation is an excellent approximation for our purposes.

Figures 3.2 and 3.3 show the growth of a Bahcall-Wolf cusp for two initial Dehnen models with $\gamma = 1/2$, differing only in the mass of the central black hole particle: $0.01M_{cl}$ and $0.05M_{cl}$ respectively. The 7/4 cusp grows over approximately half of a relaxation time, as it is measured at the hole's influence radius r_h. The main difference between the two calculations is indeed the value of r_h (≈ 0.26 and ≈ 0.66); this difference reflects in the maximum radius out to the inner cusp extends to. In both cases, the inner cusp extends out to $r \lesssim 0.1 r_h$. The solution $f(E)$ tends asymptotically to the expected $+1/4$ slope, although there is a slight increase to a higher value for energies very close to E_t. This is a result of the boundary condition $f(E_t) = 0$. The total number of stars per unit energy $N(E) \sim p(E)f(E) \sim E^{-5/2}E^{1/4} \sim E^{-9/4}$ for $E_h \ll E \ll E_t$. This means that the inner cusp, although it reaches very high densities, it is nevertheless built from a very small number of stars occupying a very small volume.

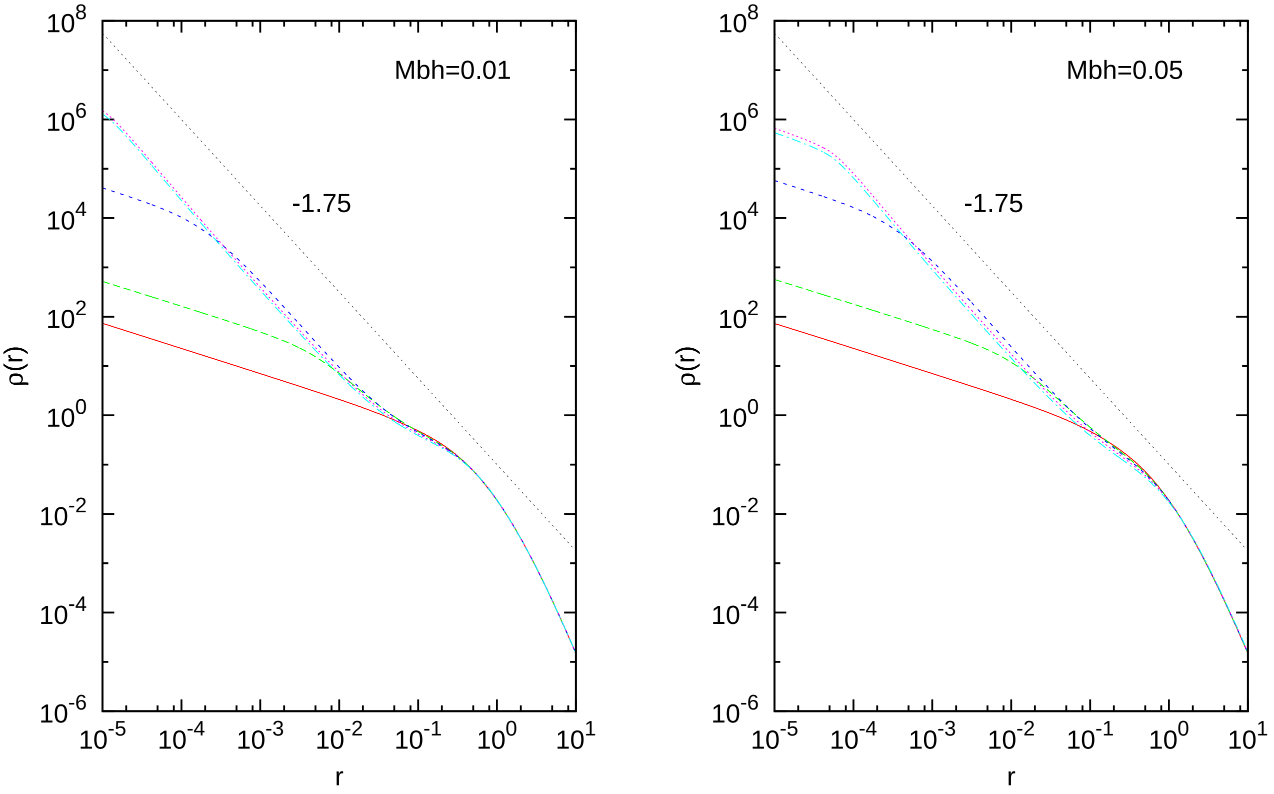

Figure 3.2: Evolution of the density profile $\rho(r)$ according to the Fokker-Planck equation.The initial condition is a Dehnen model with $\gamma = 1/2$ (lower curves)a black hole with mass $M_\bullet = 0.01$ (left panel) and $M_\bullet = 0.05$ (right panel). The asymptotic $\rho(r) \propto r^{-7/4}$ solution is established after roughly $(0.5 - 0.6)t_{\rm rlx}(r_h)$. This solution is valid out to $r \sim 0.1r_h$ in both cases. The right arrow signals points to the influence radius r_h, the left arrow points to $0.1r_h$.

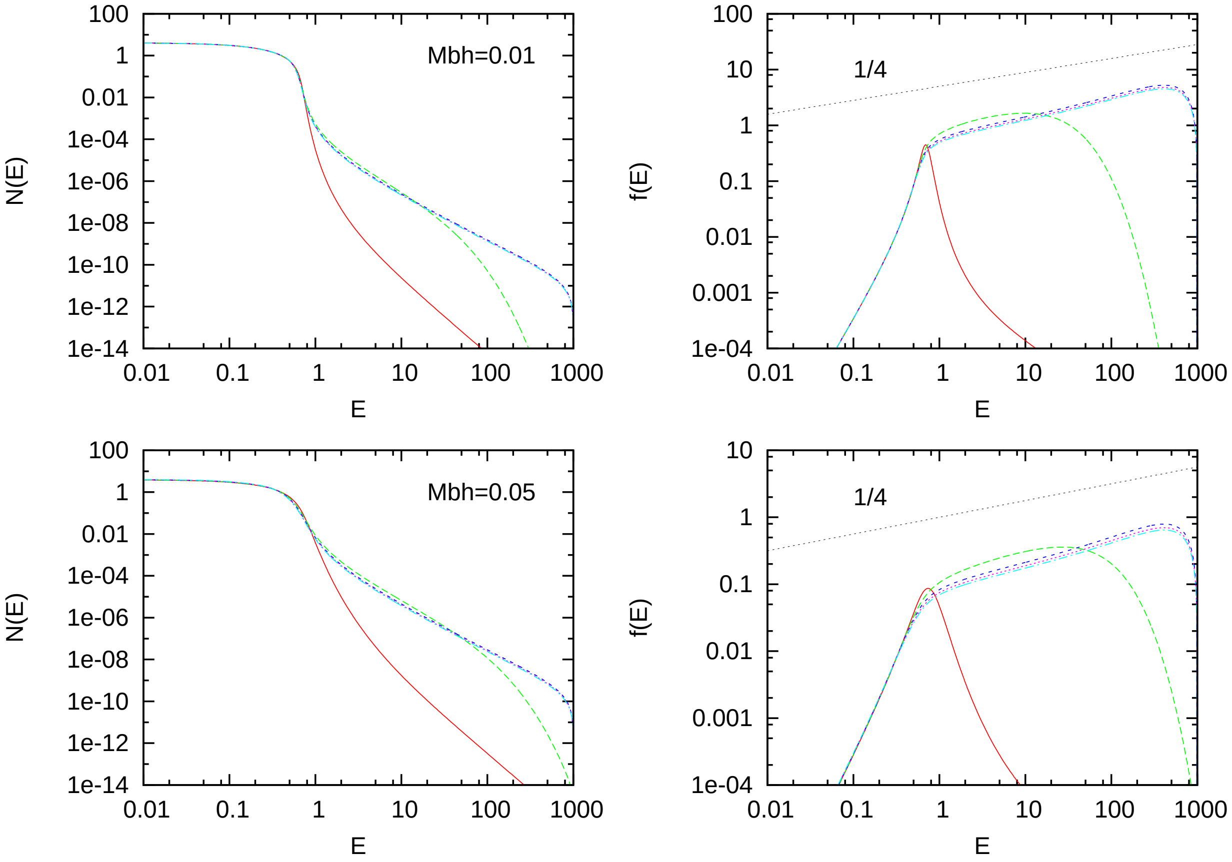

Figure 3.3: Evolution of the phase space density for the models in Figure 3.2. $N(E)$ is the number of stars per unit energy, $f(E)$ the number of stars per unit volume (phase space $d\mathbf{x}d\mathbf{v}$). The asymptotic solution is attained after the same time as ρ in Figure 3.2.

3.6 Multi-mass steady-state solution

The Fokker-Planck equation for the multi-mass case is given by [Spitzer, 1987, Chernoff and Weinberg, 1990]:

$$\begin{aligned}\frac{\partial f_i}{\partial t} = \frac{16\pi^2 G^2 m_*^2 \ln\Lambda}{p(E)} \frac{\partial}{\partial E} \sum_j \mu_j \Bigg[&- \left(\int_E^{+\infty} dE' p(E') f_j(E') \right) \mu_i f_i(E) + \\ &\mu_j \left(q(E) \int_{-\infty}^{E} dE' f_j(E') + \int_E^{\infty} dE' q(E') f_j(E') \right) \frac{\partial f_i}{\partial E} \Bigg] ,\end{aligned} \tag{3.45}$$

where $\mu_j \equiv m_j/m_*$ is the ratio of the mass of the j^{th} stellar component, m_* is some fiducial mass scale, and f_i is the DF for the i^{th} component. In the case of a single mass population, one would have $m = m_*$ and $\mu = 1$. The Fokker-Planck equation can be rewritten in a more compact form, as follows:

$$\frac{\partial f_i}{\partial t} = \frac{16\pi^2 G^2 \ln\Lambda}{p(E)} \frac{\partial}{\partial E} \sum_j m_j \left[-m_i D_E^{(j)}(E) f_i(E) + m_j D_{EE}^{(j)} \frac{\partial f_i}{\partial E} \right] , \tag{3.46}$$

where $D_E^{(j)}$ and $D_{EE}^{(j)}$ are the first and second order diffusion coefficients of the i^{th} component, respectively. These coefficients add linearly the contribution from the scattering of stars with the other stars in the same mass bin $j = i$, to those contributions that arise from the scattering with stars that belong to the other mass bins $j \neq i$.

Bahcall & Wolf (1977) investigated the steady-state solutions for the more realistic case of a stellar system with more than one mass component, again under the assumptions inherent to the Fokker-Planck formalism. They have found a scaling relation between the exponents p_j of the power law solutions $f(E_j) \propto E_{p_j}$ in terms of the stellar masses of the different components. It is easy to sketch a simple derivation of such scaling relations. For simplicity, we start to consider the case with two components only; the stellar fluxes for both are then given by

$$\begin{aligned} R_1(E) = R_{11}(E) + R_{12}(E) = m_1 & \left[D_E^{(1)} m_1 f_1(E) + m_1 D_{EE}^{(1)} \frac{\partial f_1}{\partial E} \right] + \\ & m_2 \left[D_E^{(2)} m_1 f_1(E) + m_2 D_{EE}^{(2)} \frac{\partial f_1}{\partial E} \right] \\ R_2(E) = R_{21}(E) + R_{22}(E) = m_1 & \left[D_E^{(1)} m_2 f_2(E) + m_1 D_{EE}^{(2)} \frac{\partial f_2}{\partial E} \right] + \\ & m_2 \left[D_E^{(2)} m_2 f_2(E) + m_2 D_{EE}^{(2)} \frac{\partial f_2}{\partial E} \right], \end{aligned} \tag{3.47}$$

where R_{11} is the stellar flux of the component 1 due to gravitational scattering with itself; R_{12} is the stellar flux of component 1 due to scattering with component 2, and so on. Now, suppose that the stellar system evolves towards a steady-state in which both components settle to a zero flux solution, $R_1(E) = R_2(E) = 0$ and their DFs are

$$f_1(E) \propto E^{p_1}, \qquad f_2(E) \propto E^{p_2}. \tag{3.48}$$

Then, we can write

$$p_i(E) = \frac{E}{f_i(E)} \frac{\partial f_i}{\partial E} = \frac{\partial \ln f_i}{\partial \ln E}. \tag{3.49}$$

Substituting this *ansatz* in (3.47), we find that the zero-flux steady-state solution reads

$$\begin{aligned} R_1(E) &= m_1 \left[-D_E^{(1)} m_1 f_1(E) + m_1 D_{EE}^{(1)} \frac{p_1 f_1(E)}{E} \right] + \\ & \quad m_2 \left[-D_E^{(2)} m_1 f_1(E) + m_2 D_{EE}^{(2)} \frac{p_1 f_1(E)}{E} \right] = 0 \\ R_2(E) &= m_1 \left[-D_E^{(1)} m_2 f_2(E) + m_1 D_{EE}^{(1)} \frac{p_2 f_2(E)}{E} \right] + \\ & \quad m_2 \left[-D_E^{(2)} m_2 f_2(E) + m_2 D_{EE}^{(2)} \frac{p_2 f_2(E)}{E} \right] = 0. \end{aligned} \tag{3.50}$$

Since, by assumption, the stellar fluxes of both components vanish identically for all energies, we can subtract them as follows

$$R_1 - \frac{m_1}{m_2} R_2 = m_1^2 \frac{D_{EE}^{(1)}}{E} \left(p_1 - \frac{m_1}{m_2} p_2 \right) + m_2^2 \frac{D_{EE}^{(2)}}{E} \left(p_1 - \frac{m_1}{m_2} p_2 \right) = 0. \tag{3.51}$$

We know that the diffusion coefficients $D_{EE}^{(1)}$ and $D_{EE}^{(2)}$ are always positive for all positive energies E (as we consider bound stars only, with $E > 0$, and $f(E) = 0$ for $E < 0$); hence in order for (3.51) be valid for all E,

$$p_1 = \frac{m_1}{m_2} p_2 \tag{3.52}$$

must be true. Note that this result can be readily extended for an arbitrary number of m components assuming they all have reached a steady-state zero-flux solution; in this more general case, we obtain:

$$p_j = \frac{m_j}{m_i} p_i. \tag{3.53}$$

Bahcall & Wolf (1977) have derived this result for the special case where the gravitational potential is that of the black hole only, but this simple derivation doesn't suffer from that limitation and is valid when the stellar potential in taken into account. The scaling law (3.52) shows that heavier components of the population establish a steeper density profile than the light ones. In the extreme case of a population of test stars, $p = 0$, and this is the shallowest profile allowed $\gamma = 3/2$. It should be stressed that this does not constitute a proof that all solutions obey this scaling relation independently of the initial conditions and of the boundary conditions — the power law solution $f(E) \propto E^{p_i}$ is an assumption. All numerical solutions developed this scaling relation so long as the boundaries are chosen in a way there is an energy range $E_h \ll E \ll E_t$ for the asymptotic solution.

3.7 Mass segregation and equipartition of kinetic energy

In the presence of a range of stellar masses, the exchange of energy between the different components tends to drive the system locally towards the *equipartition* of kinetic energy — eg. $m_1 < v_1^2 >= m_2 < v_2^2 >$, if there are only two components. During this process, the heavier stars, on average, loose kinetic energy and fall towards the bottom of the stellar plus hole potential well; at the same time, the lighter stars, on average, gain kinetic energy, and move out of the potential well. This tendency towards equipartition of kinetic energy is therefore the microscopic mechanism underlying the mass and spatial segragation of the stellar populations.

In fact the diffusion coefficients are the moments of the velocity changes due to stellar scattering. There are only three independent diffusion coefficients: $D(\Delta v_\parallel)$, $D((\Delta v_\parallel)^2)$, and

$D((\Delta v_\perp)^2)$. These can be combined to find the rate of change of kinetic energy

$$\begin{aligned}
D(m_* \Delta E) &= m_* \sum_{i=1}^{3} \left(v_i D(\Delta v_i) + \frac{1}{2} D(\Delta v_i \Delta v_j) \right) = \\
&= m_* \left(v_\parallel D(\Delta v_\parallel) + \frac{1}{2} D((\Delta v_\parallel)^2) \frac{1}{2} D((\Delta v_\perp)^2) \right) = \\
&= 16\pi^2 G^2 m_* m_f \ln\Lambda \left[m_f \int_v^{+\infty} dv_f v_f f_f(v_f) - \frac{m_*}{v} \int_0^v dv_f v_f^2 f_f(v_f) \right]. \qquad (3.54)
\end{aligned}$$

We have made use of the local approximation to write the second equality. The first integral acts as a *heating* term, only the field stars with $v_f > v_t$ contribute to the heating; the second integral acts as a *cooling* term, only the field stars with $v_f < v_t$ contribute to the friction. Note that when the test star moves with low velocity, the heating term dominates and the star gains kinetic energy; in the other limit, when the test star acquires high velocity, the opposite happens, the friction term is dominant and the test star loses kinetic energy. The tendency towards equipartition is, effectively, the result of the concurring contributions from these two terms.

If both stellar populations have Maxwellian velocity distributions, the mean rate of change of the kinetic energy for, say m_2 viewed as a population of test stars evolving in a sea of stars with mass $m_1 < m_2$, averaged over the velocities of all test stars of mass m_2 is given by [Spitzer, 1987]

$$\begin{aligned}
\frac{d\bar{E}_2}{dt} &= \frac{4}{\sqrt{\pi}} \alpha^3 \left(\frac{3}{2\sigma^2} \right)^{3/2} \int_0^{+\infty} dv_2 v_2^2 e^{-\frac{3v_2^2}{2v_1^2}} D(m_* \Delta E) = \\
&= 8\sqrt{6\pi} G^2 m_2 \rho_1 \ln\Lambda \frac{\bar{E}_1 - \bar{E}_2}{(\sigma_1^2 + \sigma_2^2)^{3/2}}, \qquad (3.55)
\end{aligned}$$

where $E_i = 1/2 m_i \sigma_i^2$ and $\sigma^2 = \sigma_1^2 + \sigma_2^2$. We can obtain a symmetric for $d\bar{E}_1/dt$ simply by interchanging the subscripts 1 and 2, and it follows that $n_1 d\bar{E}_1/dt + n_2 d\bar{E}_2/dt = 0$ as it should for consistency. From (3.55), the scattering between the two populations drives them to a steady-state in equipartition of kinetic energy: if the heavy component has larger kinetic energy, as in our calculations, encounters with the light stars leads to a net loss in kinetic energy, while the symmetric equation governs how the light component undergoes the corresponding net gain in kinetic energy. Suppose now that $m_2 \gg m_1$, the characteristic time scale for establishing equipartition, $t_{eq} \sim \left| \bar{E}_2 - \bar{E}_1 \right| / \left| \dot{\bar{E}}_2 - \dot{\bar{E}}_1 \right|$, is given in order of magnitude by:

$$t_{eq} \sim \frac{0.0814\sigma^3}{G^2 m_1 m_2 \ln\Lambda (n_1 + n_2)} t_{rlx,1} \sim 0.239 \frac{m_1}{m_2} \left(1 + \frac{\sigma_2^2}{\sigma_1^2} \right)^{3/2} \frac{n_1}{n_1 + n_2} t_{rlx,1}, \qquad (3.56)$$

where $\sigma^2 \approx \sigma_1^2 \approx \sigma_2^2$ which is indeed true given our initial conditions (see Chapter 4), n_i is the number density of component i and $t_{rlx,1}$ is the relaxation time of the light component if it were to evolve in isolation. Note, however, this derivation was made under two implicit assumptions: (*i*) The stellar potential is static (this cannot be strictly true, as it will certainly evolve as equipartition and mass segregation proceed); (*ii*) And, more importantly in our case, it was assumed there is no external potential present, e.g. no black in the center of the cluster.

In fact, in the absence of a black hole, Inagaki & Wiyanto(1984) and Inagaki & Saslaw(1985) have shown, by Fokker-Planck calculations of self-consistent star clusters, and more recently Khalisi et al. (2006) with N-body simulations, that the different populations of stars can either move towards or away from equipartition, depending on the number fraction and on the total mass residing in each of the components.

If a massive black hole is present at the center of the cluster, the system can *never* reach equipartition. This can be easily verified by computing the expected value of the velocity dispersion for each component that obeys the steady-state Bahcall & Wolf cusp solution. Either from the Jeans's equation for a spherical system with an isotropic velocity tensor, or directly from the definition of the second moment of the velocity distribution, we arrive at the following result for the velocity dispersion of component i

$$\langle v_i^2 \rangle = 3\langle v_{r,i}^2 \rangle = \frac{6GM_\bullet}{(5+2p_i)r}. \tag{3.57}$$

The explicit dependence of the velocity dispersion for each component i on the exponent p_i of its $f_i(E)$ means that the equipartition parameter ξ must always be different from one

$$\xi \equiv \frac{m_2\langle v_2^2 \rangle}{m_1\langle v_1^2 \rangle} = \frac{m_2}{m_1}\frac{5+2p_1}{5+2p_2} \neq 1. \tag{3.58}$$

This result warrants a simple physical interpretation. Inside the cusp where the black potential is dominant: the heavy stars, as they on average lose kinetic energy to the lighter stars, fall deeper into the potential well of the black hole where they will settle with a larger kinetic energy; whilst the lighter ones, gaining on average kinetic energy, move outwards to radii where they will also settle but with smaller kinetic energy. As a result it is not possible, inside the cusp, for the heavy stars to undergo a net loss of kinetic energy, while the lighter stars receive a net gain of kinetic energy, as it would be required for reaching equipartition. This is an even more drastic situation than the case of a stellar system without central point masses where it is possible, under restricted circumstances given approximately by the Spitzer criterium, to reach equipartition in the center.

The velocity dispersion of each component still scales as $r^{-1/2}$, as expected on general grounds, but the ratio of the amplitudes do not converge to the relation $\sigma_i^2 \propto m_i^{-1}$ expected from equipartition arguments. It should be stressed that the predicted velocity dispersion in the cusp is only very weakly dependent on the stellar mass (via the exponent p, it changes by less than 10% over an entire stellar range). This is a very general prediction for the velocity distributions of old stars around a massive black hole and it is consistent with observations of the Galactic center, as well as for the isotropy and non-rotation [Alexander, 2005].

Note that this is not, however, a stric proof since it relies on an unproven, although very reasonable and compelling, assumption about the distribution function of all components $f_i(E) \propto E^{p_i}$. All numerical experiments (either of N-body or of Fokker-Planck type) have confirmed that this is indeed the asymptotic solution independently of initial conditions.

Figures 3.4 and 3.5 show the growth of a Bahcall-Wolf cusp over slightly less than half a relaxation time as measured at the hole's sphere of influence. In both cases, the initial conditions was a Dehnen model with an inner slope $\gamma = 1/2 + \delta$, with two mass components of mass ratio equal to 5 and 3 respectively, and equal number fraction for each component $f_1 = f_2 = 1/2$. The models have a central black hole mass $M_\bullet = 0.01$ and $M_\bullet = 0.05$, respectively. The main difference between these two runs is the radius of influence and, consequently, the largest radius out to which the inner cusp forms. This is very clearly seen from the density runs with radius for the component of lower mass. The density cusp of the heavy stars extends out to $\sim 0.1 r_h$ in both cases. In the $M_\bullet = 0.01$ case, the cusp of the light component assumes the asymptotic $\gamma = 1.5$ slope only for $r \lesssim (6-8) \times 10^{-3}$ well inside $0.1 r_h$, whilst in the $M_\bullet = 0.05$ case the same slope extends out to $r \approx 0.02$ again inside $0.1 r_h$. Therefore, in the N-body calculations, due to the finite number of particles, it will be considerably more difficult to resolve the Bahcall-Wolf cusp not only for the cases with lower black hole mass, but also for the light component. It can be seen from the plots the light component shows a variable (negative) slope over a larger radial range than the heavy component — the latter reaches its asymptotic slope more rapidly.

Both figures show the final cusp for the heavy component with the form $f_h(E) \sim E^{p,h}$, with $p_h \sim 0.3 - 0.4$ which is a slightly larger value than that predicted for the single-mass case. The same result was seen in the Bahcall & Wolf 1977 approximate calculations. There is no contradiction with the derivation of the $p = 1/4$ cusp since the latter derivation was valid for the single mass case only. The scaling relations for the exponents was obtained from an independent derivation that did not constrain the value for the slope of the heaviest component. We see that, when more than one mass component is present, the heavier one can develop a density profile that it is not very far from that of the isothermal sphere. On the other hand, the slope of the light component is very close to that predicted from the scaling relation. $p_l \approx m_l/m_h p_h \approx 1/5 p_h \approx 0.07$ and ≈ 0.12 for each case.

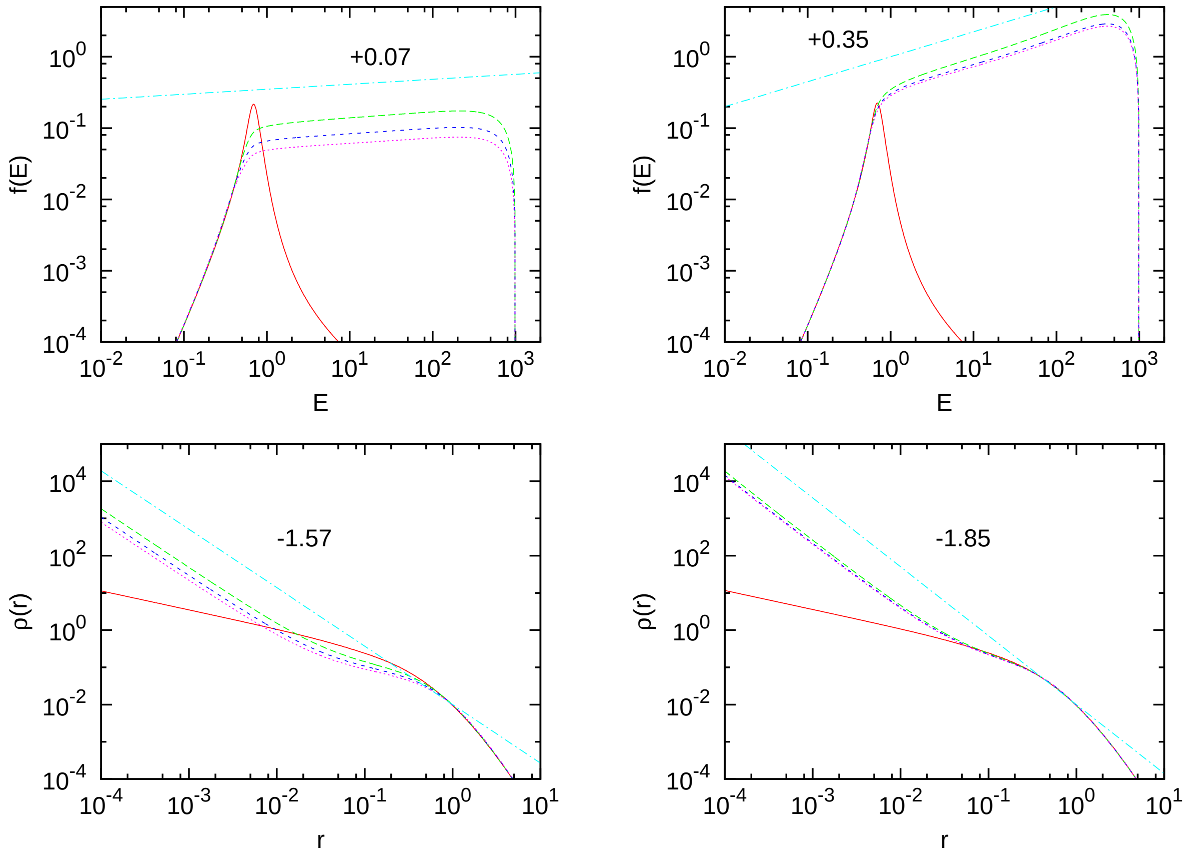

Figure 3.4: Evolution of the mass density $\rho(r)$ and phase space density $f(E)$ for Dehnen models with two-mass components, $\gamma = 1/2$ and black hole mass $M_\bullet = 0.01 M_{cl}$. The mass ratio between the stars in each component is $R = 5$, there is the same number of stars of each mass, $f = 1/2$. Upper left panel: The light population evolves to the asymptotic phase space solution $f_l(E) \propto E^{0.07}$, with $p_l = m_l/m_h p_h$. Upper right panel: The heavy population phase space density reaches the asymptotic value $\gamma_h \approx 0.35$ above the 1/4 value of the single mass case. Lower left panel: The spatial density of the light component extends out to less than $0.1 r_h$. Lower right panel: The spatial density of the heavy component has an inner slope that extends out to $\sim 0.1 r_h$. The time for the growth of the cusp is $\approx 0.3 t_{rlx}$, which is evidence of the accelerating effect that the energy exchange between stars with different masses has on the dynamical evolution.

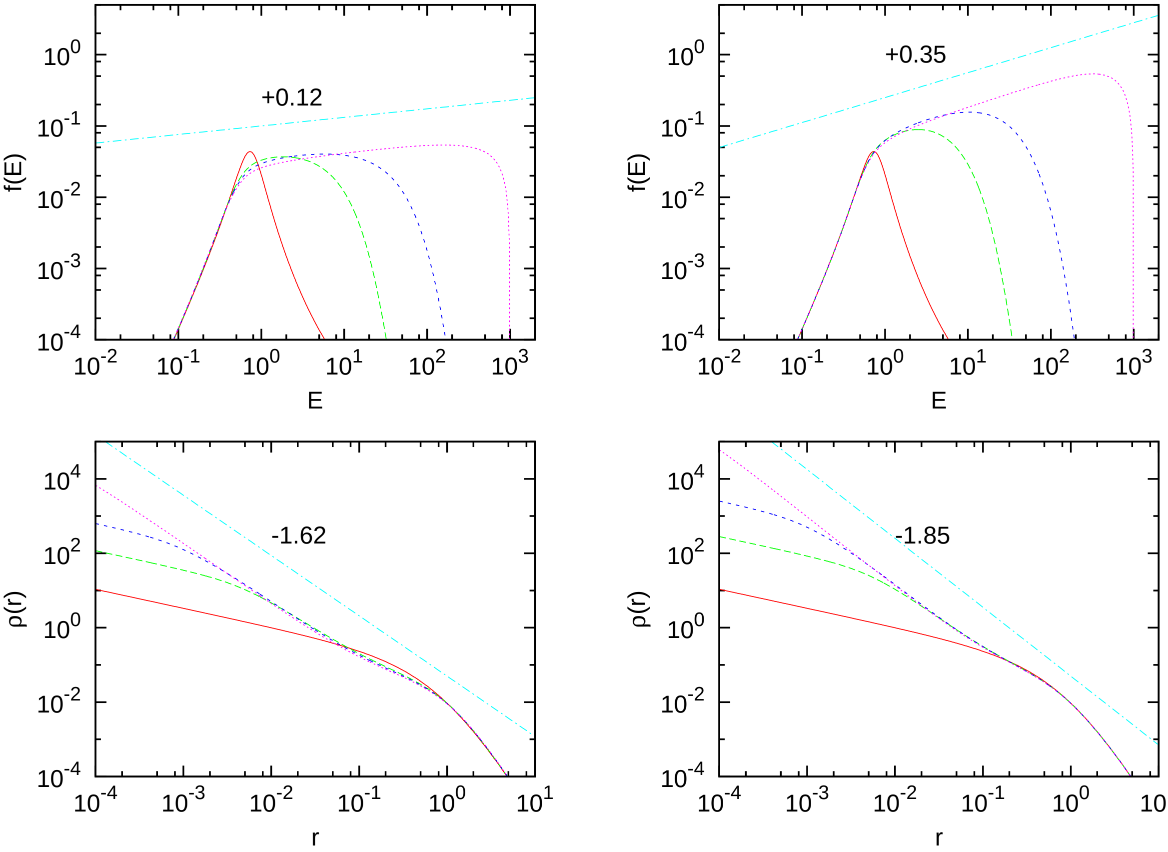

Figure 3.5: Evolution of the mass density $\rho(r)$ and phase space density $f(E)$ for Dehnen models with two-mass components, $\gamma = 1/2$ and black hole mass $M_\bullet = 0.05 M_{cl}$. The mass ratio between the stars in each component is $R = 3$, there is the same number of stars of each mass, $f = 1/2$.

Chapter 4

Numerical results: N-body and Fokker-Planck comparison

4.1 Introduction

The distribution of stars around a massive black hole is a classic problem in galactic dynamics. Many such distributions are possible, depending on the initial state, the mode of black hole formation, and the time since its formation. However, a number of plausible scenarios predict a steeply-rising stellar density within the black hole's sphere of influence, $r \lesssim r_h, \sigma^2 = GM_\bullet/\sigma^2$, with $M_\bullet$ the black hole mass and σ^2 the $1D$ stellar velocity dispersion outside of r_h. Such models receive support from the observed run of stellar density with radius near the centers of the nearest galaxies known to contain black holes: the Milky Way, M31, and M32. The nucleus of each galaxy has a density cusp with $\rho \sim r^{-\gamma}$, $\gamma \approx 1.5$ within the black hole's sphere of influence [Lauer et al., 1998, Genzel et al., 2003a, Schödel et al., 2007]. This is consistent with the slopes predicted by the so-called adiabatic growth models, in which a black hole with a small initial mass is embedded in a star cluster and its mass increased to some final value, on a time scale long compared with orbital periods. The final density in the adiabatic growth model follows a power law at $r \lesssim r_h$, with index γ_f that depends on the initial stellar distribution. If the initial model is a non-singular isothermal sphere, then $\gamma_f = 3/2$ [Peebles, 1972, Young, 1980]. However, a non-singular isothermal sphere seems a rather ad-hoc guess for the initial state, and initial models without flat cores give steeper final slopes, $1.5 \lesssim \gamma_f \lesssim 2.5$ [Lee and Goodman, 1989, Quinlan et al., 1995, Merritt, 2006].

Uncertainties in the initial state are less consequential if the stellar cluster is old compared with the relaxation time $t_{\rm rlx}$ for gravitational scattering to redistribute energy between stars. In this case, one expects the collisional transport of mass and energy to set up a steady-state distribution whose functional form is independent of the initial phase space density. Peebles (1972) first addressed this problem and derived a power-law index $\gamma = 9/4$ for the stellar density within r_h. Peebles obtained this solution via a scaling argument and by setting the flux of stars in energy space to a constant non-zero value. Shapiro & Lightman (1976) and Bahcall & Wolf (1976) criticized the Peebles derivation on the grounds that the implied flux is unphysically large, if fact divergent if the solution is extended all the way to the black hole and it violates the boundary condition by turning the black hole to a source. A physically more reasonable solution would have a nearly zero flux of stars into the black hole and treating it as a sink. Bahcall & Wolf (1976) repeated Peebles's derivation, setting the phase space density $f(E)$ to zero at the black hole's tidal disruption radius. They found that f evolves, in a time only slightly longer than $0.5t_{\rm rlx}$, to a steady-state in which the flux is close to zero at all energies. The zero-flux solution has $f \propto E^{1/4}$ within the black hole's sphere of influence, with $E \geq 0$ the binding energy; the corresponding stellar density is $\rho \propto r^{-7/4}$.

The Bahcall-Wolf solution has been verified in a number of subsequent studies, almost all of which were based on the Fokker-Planck formalism [Cohn and Kulsrud, 1978, Marchant and Shapiro, 1980, Freitag and Benz, 2002, Amaro-Seoane et al., 2004]. The $f \propto E^{1/4}$ and $\rho \propto r^{-7/4}$ character of the solution has been found to be robust, at least at radii where the capture or destruction of stars occurs in a time long compared with orbital periods. But verifying the Bahcall-Wolf solution in an N-body integration is also clearly desirable, since an N-body calculation is free of many of the simplifying assumptions inherent to the Fokker-Planck equation, including the restriction to small-angle scattering, and the neglect of spatial inhomogeneities in the derivation of the (orbit-averaged) diffusion coefficients. But the N-body approach is challenging: large particle numbers are required to resolve the cusp, and accurate integration schemes are needed to follow accurately the motion of stars near the black hole. In this chapter, we combine sophisticated N-body codes with the special-purpose GRAPE hardware and show that the formation of a Bahcall-Wolf cusp can be convincingly reproduced without any of the approximations that go into the Fokker-Planck formalism. Our results provide a clear demonstration of the applicability of the direct force N-body techniques to the collisional evolution of dense star clusters, and highlight the usefulness of the direct N-body techniques for the undestanding of the dynamical evolution of galactic nuclei containing massive black holes. The results presented in this Chapter constitute the first milestone towards a full, detailed direct N-body modelling of galactic nuclei harboring a massive black hole at its center.

4.2 Initial conditions

A crucial element of our method is the use of initial conditions that represent a precise steady state of the collisionless Boltzman equation. Embedding or growing a massive particle in a pre-existing stellar system can easily result in the formation of a density cusp with slope similar to that predicted by Bahcall & Wolf (1976), but for reasons having nothing to do with collisional relaxation. For instance, as discussed in the Introduction, adiabatic growth of a black hole produces a density profile within the hole's sphere of influence of $\rho \sim r^{-\gamma}$,

$1.5 \lesssim \gamma \lesssim 2.5$; hence a cusp that forms collisionlessly can easily mimic a $\rho \propto r^{-7/4}$ collisional cusp. To avoid any possibility of non-collisional cusp formation in our simulations, which could seriously compromise the interpretation, we generated initial coordinates and velocities from the steady-state phase space density $f(E)$ that reproduces the Dehnen density law in the gravitational potential including both the stars and the black hole [Dehnen, 1993, Tremaine et al., 1994].

The Dehnen model's density law is defined by [Dehnen, 1993]

$$\rho(r) = \frac{(3-\gamma)M_{cl}}{4\pi} \frac{a}{r^{\gamma}(r+a)^{4-\gamma}}, \qquad 1/2 < \gamma \leq 3, \tag{4.1}$$

where M_{cl} represents the total mass of the stellar cluster (not including the black hole's mass), a is a break radius that roughly separates the inner region of the cusp with slope $-\gamma$ from the outer parts whose density decays much faster with slope $\gamma - 4$. The gravitational potential $\Phi_*(r)$ from stars that corresponds to (4.1) is given, through Poisson's equation, by

$$\Phi_*(r) = \frac{GM_{cl}}{a} \times \begin{cases} -\frac{1}{2-\gamma}\left[1 - \left(\frac{r}{r+a}\right)^{2-\gamma}\right], & \gamma \neq 2 \\ \\ \ln\left(\frac{r}{r+a}\right), & \gamma = 2 \end{cases} \tag{4.2}$$

This model was generalized by Tremaine et al. (1994) for the case where there is a massive point particle at the center of the cluster. In this event, the density law remains unaltered, but the gravitational potential generated by the massive particle needs to be added to the stellar contribution in order to obtain the overall potential

$$\Phi(r) = \Phi_*(r) - \frac{GM_{\bullet}}{r}. \tag{4.3}$$

The Dehnen model has an isotropic velocity distribution, so its DF depends on the energy E only, $f = f(E)$. The density $\rho(r)$ and the $f(E)$ are related, in case of an isotropic velocity distribution, through the following equation [Binney and Tremaine, 1987]

$$\rho(r) = 4\pi\ M_{cl} \int_0^{\Psi(r)} dv\ v^2 f(E)\sqrt{2(\Psi(r) - E)}, \tag{4.4}$$

where $E = -\frac{1}{2}v^2 - \Phi(r) = \Psi(r) - \frac{1}{2}v^2$. The bound stars have $E > 0$; unbound stars are not included so $f(E) = 0$ for $E < 0$. Following Binney & Tremaine (1987), since $\Psi(r)$ is a monotonic function of r in any spherical potential, it follows that we can effectively regard ρ as a function of Ψ rather than of r, and write

$$\rho(\Psi) = 4\pi\ M_{cl} \int_0^{\Psi} dE\ f(E)\sqrt{2(\Psi - E)}. \tag{4.5}$$

Next, differentiate both sides with respect to Ψ to obtain

$$\frac{1}{\sqrt{8}\pi M_{cl}}\frac{d\rho}{d\Psi} = \int_0^{\Psi} dE\ \frac{f(E)}{\sqrt{\Psi - E}}. \tag{4.6}$$

This is an Abel integral equation whose solution is well-known and is readily computed by the Eddington's formula

$$f(E) = \frac{1}{\sqrt{8}\pi^2 M_{cl}}\frac{d}{dE}\int_0^{E} \frac{d\Psi}{E-\Psi}\frac{d\rho}{d\Psi}. \tag{4.7}$$

In this way, it is clear that, given the Dehnen's density law, one can easily determine an isotropic DF that supports it, including the case in which there is a massive particle at the center of the cluster model.

It is instructive to dedicate some space to highlight a few of the asymptotic properties of the Dehnen's model. In the limit of large energy E (or, equivalently, for small radius $r \ll r_h$ near the black hole), the Dehnen density is asymptotically of the form $\rho(r) = \rho_0(r_h/r)^{\gamma}$; the velocity dispersion for $r \lesssim r_h$ is given by $\sigma^2(r) \sim GM_\bullet/r$. Using Eddington's equation we find in this limit

$$f(E) = \frac{\rho_0}{(2\pi\sigma_h^2)^{3/2} m_*}\ \frac{\Gamma(\gamma+1)}{\Gamma(\gamma-1/2)}\ \left(\frac{E}{\sigma_h^2}\right)^{\gamma-3/2}, \qquad 1/2 < \gamma < 3, \tag{4.8}$$

for the DF deep inside the stellar cusp, *ie.* for $E \gtrsim E_h$ or $r \lesssim r_h$. We can see from (4.8) that, in the presence of a massive particle at the center, the $f(E)$ is not well-defined for cusp slopes shallower than $\gamma \leq 1/2$. The model with $\gamma = 1/2 + \delta$, with δ as small as one wishes, does indeed yield a cusp with the shallowest slope allowed for a solution of the CBE in the case of isotropic velocities. It can be shown, however, that an equilibrium solution of the CBE can be achieved together with slopes shallower than $\gamma = 1/2$ so long as some tangential-biased anisotropic tensor is allowed [Merritt, 2006]. We shall not pursue this in this work.

Another property worth mentioning is the following: although the $f(E)$, for $\gamma > 3/2$, rises steeply as $E \to +\infty$, the number of stars $N(E)$ deep inside the cusp is relatively small; it fact it decreases very rapidly with increasing E according to

$$N(E) = 16\pi^2 p(E) f(E) = \frac{32\pi^2}{5}\frac{\rho_0}{(2\pi\sigma_h^2)^{3/2}}\frac{(GM_\bullet)^3}{m_*}\frac{E^{\gamma-4}}{\sigma_h^{2\gamma-3}}, \tag{4.9}$$

since $p(E) \sim \frac{\sqrt{2}\pi}{16}(GM_\bullet)^3 E^{-5/2}$ for $r \ll r_h$. For $\gamma < 3/2$, the asymptotic behavior of $f(E)$ is a monotonic decrease at high energies, and that of $N(E)$ as well.

For $1 < \gamma < 2$, the Dehnen models without a central black hole are cold at the center, as their velocity dispersion decreases towards zero at the origin; for $0 < \gamma < 1$, they tend to a constant value at the center; only for $2 < \gamma < 3$ does the velocity dispersion diverge as $r^{-\gamma}$ at the center. However, and as expected on general grounds, the addition of the black hole to the cluster's center changes the $f(E)$ and generates the usual velocity dispersion profile $\sigma(r) \propto r^{-1/2}$, for $r < r_h$ and for all $\gamma > 1/2$. This is, of course, one of the classical tell-tale signs of the presence of a massive black hole at the center of a stellar cluster. One remark on the outer slope of the Dehnen model: its density decreases outwards much more steeply than what one expects realistically for a galactic nucleus. For instance, in the case of the Milky Way, the power law dependence of the density cusp extends to well outside the hole's radius of influence [Alexander, 2005, Schödel et al., 2007]. We can interpret this faster decline of the density of our models, at larger radii outside of the cusp, as an effective cut-off that serves as a trick for not waisting too much time computing the dynamics of regions that are basically stationary, and non-evolving, during the time scales we are interested in.

In summary, the Dehnen model density satisfies $\rho(r) \propto r^{-\gamma}$ at small radii, and the $f(E)$ that reproduces the Dehnen's $\rho(r)$ in the presence of a central black hole is non-negative for all $\gamma > 1/2$; hence $\gamma = 1/2 + \delta$ is the shallowest central profile that can be adopted if the initial conditions are to be represent a precise steady-state. As shown in Table 1, most of our runs used initial conditions with this minimum value of γ. Henceforth we adopt units such that the gravitational constant G, the total stellar mass M_{cl}, and the Dehnen scale length a are all equal to one.

Table 1 also gives the other important parameters of the N-body integrations. The influence radius of the black hole r_h was defined as the radius at which the enclosed stellar mass at $t = 0$ was equal to twice the black hole mass. This is, in fact, equivalent to the more standard definition $r_h = GM_\bullet/\sigma^2$ when $\rho \propto r^{-2}$. The relaxation time $t_{\rm rlx}$ was computed at $t = 0$, $r = r_h$ from the standard expression (eq. 2-62 of Spitzer 1987), setting $\ln\Lambda = \ln(r_h\sigma^2/2Gm_*)$. Both r_h and $\sigma(r_h)$ for the initial models are computed from analytical expressions from Tremaine et al. (1994). This definition of Λ is equivalent to equating $b_{\rm max}$, the maximum impact parameter for encounters in Chandrasekar's theory, with r_h. This choice is motivated by the expectation that $b_{\rm max}$ for stars near the center of strongly inhomogeneous stellar systems should be of order the radius at which the density faloff begins to abate, e.g. the core radius if there is a core [Maoz, 1993, Merritt, 2001]. In our simulations, this radius is of order the Dehnen scale length $a \approx$ a few$\times r_h$ at $t = 0$, decreasing to a fraction of r_h after the formation of the collisional cusp; hence a choice of $b_{\rm max} \approx r_h$ seems appropriate. Furthemore, as pointed out by Spitzer (1987), there are several other factors that contribute to constrain the value $b_{\rm max}$. For instance, if $\rho(r)$ decreases with a characteristic distance scale a, then the auto-correlation function of the fluctuating force that acts on the test star decreases very strongly for time scales longer than a/σ; therefore, fluctuations on spatial scales larger than a should not contribute very much for the diffusion coefficients of stars within the cusp. Concurrently, if there is a great discrepancy between the orbital periods of test and perturbing stars, the perturbations become adiabatic and produce no net effect on the secular evolution of the stellar system. We stress that Chandrasekar's $\ln\Lambda$ is a poorly-defined quantity in strongly inhomogeneous and evolving systems, and our choice is at best approximate. Nevertheless, we will see that the time scaling determined by this choice of $\ln\Lambda$ results in a very good correspondence between the evolution rates seen in N-body and Fokker-Planck models. When scaling the Fokker-Planck results to the N-body results, the only free parameter is $\ln\Lambda$. We used the values given in Table 1.

In order to specify completely the initial conditions for our models, there are essentially three independent parameters to take into account: (*i*) the inner density slope, γ; (*ii*) the total number N of stars; (*iii*) and the ratio of the black hole mass to the total cluster mass in stars $M_\bullet/M_{cl}$. All the other quantities in Table 1 can be derived from these.

A systematic study of this parameter space is too expensive in terms of CPU time; in particular, we are severely restricted by the maximum number of particles that a micro-GRAPE board coupled to a single processor can handle in its memory, $N_{max} = 1.28 \times 10^5$. Runs 1 and 5 are exceptions, as they were run with the help of a large GRAPE-6 board which can store up to $N = 2.56 \times 10^5$ particles in the memory.

The total number of particles is a crucial parameter: as we decrease N, the mass ratio between the black hole and the field stars decreases as well. This is a source of problems for the N-body modelling of galactic nuclei, since this small value for N results in an increased amplitude for the black hole's wandering around the galactic center. Given the large discrepancy between the currently attainable values of N and the typical values in nuclei [*], this can be a source of confusion in the interpretation of the N-body results if we don't keep in mind that the latter may produce spurious overestimates for the hole's position and velocity fluctuations, if the adopted N is not large enough.

The discrete encounters between the black hole and the field stars induce a small, irregular (Brownian) motion of the massive black hole. This effect was recognized and estimated early by Bahcall & Wolf (1976). During the last few years, this Brownian motion has been further studied by several authors [Chatterjee et al., 2002, Merritt et al., 2007]. These studies concluded that the *rms* fluctuations of the hole's positions and velocities result from the approximate equipartition of kinetic energy with the field stars; these fluctuations scale as

$$< x^2 >\sim \frac{m_*}{M_\bullet} r_{core}^2, \qquad < \dot{x}^2 >\sim \frac{m_*}{M_\bullet}\sigma^2, \tag{4.10}$$

where r_{core} is the core radius of the stellar distribution and σ^2 is the (constant) velocity dispersion in the core. In galactic nuclei, the core radius is an ill-defined quantity as the

---

[*] The typical mass ratio in our galactic center ranges from $\sim 10^{-7}$ for very low mass stars to $\sim$ few$\times 10^{-5}$ for stellar mass black holes; while the simulations listed in Table 1 span a range from 500 to 6000. These latter values are, in fact, more representative of the *intermediate mass black holes* that may reside in the centers of globular clusters; however, see Sesana et al. 2005 for a cosmological motivation and connection with GW searches.

density scales as $\rho \sim r^{-\gamma}$. However, this result can be in some sense reinterpreted if we substitute $M_\bullet$ in the previous equation by $M_\bullet + M(< r_h)$ and interpret the result in terms of Brownian wandering of the combined system of black hole plus stellar cusp in a background core with constant density instead.

On what concerns the question of cusp formation around a black hole, part of the problem associated with the brownian motion is mitigated by the fact — seen in our simulations and from other authors — that the hole is capable of carrying the cusp with it as it wanders slowly around the center; and that, in fact, the dynamical evolution is confined to the hole's sphere of influence (see plots of the cumulative mass distribution, $M(< r)$).

4.3 Single mass case

As the cusp develops, the density of stars at $r \lesssim r_h$ increases. This is illustrated in Figure (4.1), which shows the mass in stars within a radial distance $0.1 r_h$ from the black hole as a function of time for the Runs 1 to 5 in Table. The distances were defined with respect to the instantaneous position of the black hole particle; the latter wanders like a Brownian particle, but the density peak tends to remain centered on the black hole as it moves — the black hole carries the cusp with itself. For comparison, we also show in Figure (4.1) the same quantity as computed from the Fokker-Planck equation, which can be obtained from the instantaneous density profile according to

$$M(r < 0.1 r_h; t) = 4\pi \int_0^{0.1 r_h} dr\ r^2 \rho(r; t). \tag{4.11}$$

The time scaling of the Fokker-Planck equation was set using the value of $\ln \Lambda$ given in Table — no adjustements were made optimize the fit. (We note that integrations like these could in principle be used to *evaluate* $\ln \Lambda$.) While there are hints of systematic differences in some of the runs, overall the correspondence is very good: clearly, the N-body evolution is quite close to what is predicted from the Fokker-Planck equation.

The N-body runs in Table 1 exhibit a range in N from 0.25×10^6 (Run 1) to 0.1×10^6 (Runs $2 - 4$), showing that the correspondence between N-body and Fokker-Planck results remains good over at least a modest range in particle number. Figure (4.1) also suggests that each of the integrations has reached an approximate steady-state with regard to changes in the density at the final time step. After reaching this steady state in the density profile, the system still undergoes a slow homologous expansion as predicted by the Fokker-Planck equation.

Having demonstrated the reliability of the time scaling of the Fokker-Planck equation, we can make a more detailed comparison with the N-body models. Figure 4.2 shows the evolution of the stellar density $\rho(r)$ in Run 1 compared with the Fokker-Planck prediction. The N-body density, in this case, was computed from snapshots of the particle positions (no time averaging) using MAPEL, a maximum penalized likelihood algorithm [Merritt and Tremblay, 1994]. The radial coordinate was defined again as the distance from the black hole. The correspondence between N-body and Fokker-Planck results is again quite good; the only systematic difference appears at very small radii ($r \lesssim 0.01 r_h$) where the particle numbers are too small for reliable estimates of ρ. The final cusp is well represented by $\rho \propto r^{-7/4}$ at $r \lesssim 0.1 r_h$, as in the other integrations in Table 1. Figure 4.4 shows the evolution of $\rho(r)$ for another four runs from Table 1. In this case, the density was estimated with a kernel-based algorithm [Silverman, 1986]. A superposition of snapshots of the particle's positions in real space — typically ten snapshots centered around the times indicated — was used in order to reduce the noise from the estimates of the density. In all cases, we see a robust cusp growth, for $r \lesssim (0.1 - 0.2) r_h$, over roughly half a relaxation timescale $\approx (0.5 - 0.6) t_{\rm rlx}$, and a steady profile thereafter. We conclude therefore that the cusp predicted by Bahcall & Wolf (1976) can be reproduced with direct N-body methods for a range of black hole masses that span a decade in the mass from $0.005 \times M_{cl}$ to $0.05 \times M_{cl}$. As we have discussed before, it is not feasible to decrease the black hole mass much further without increasing N at the same time. If we decrease the hole's mass further below the minimum value $0.005 M_{cl}$ indicated in the Table, within the particle number limit imposed by the micro-GRAPE's memory, an unrealistically large Brownian motion of the hole around the center would ensue compromising the mechanism for cusp formation. On the other hand, increasing the black hole mass above the maximum value indicated in the Table 1 would lead to a more realistic $M_\bullet / m_*$ ratio, but it would (i) increase $M_\bullet / M_{cl}$ to extremely unrealistic values; (ii) it would also become more expensive in terms of CPU time since r_h increases with the hole's mass and therefore the $t_{\rm rlx} \propto r^{\gamma - 3/2}$ inside the cusp increases as well (at the radius of influence, the contribution from the stars has to be taken into account but this not affect the conclusion).

The large number of particles in our N-body experiments permits us to go one step deeper in comparing N-body with Fokker-Planck results. It is possible to extract estimates of $f(E)$ from the N-body data sets. We did this as follows. Snapshots of the particle positions and velocities were stored at each N-body unit; such data sets are essentially uncorrelated at radii near r_h. Roughly between 20 to 70 (depending on the runs) of these snapshots were then combined into a single data file, giving an effective N typically of order a few $\times 10^6 - 10^7$. From the the combined data set we computed an estimate of the gravitational potential using standard expressions, assuming spherical symmetry, then computed the phase space volume element $p(E)$ defined above. The particle energies were also computed, and a histogram constructed of $N(E)$. Finally, we used the relation $f(E) = N(E)/16\pi^2 p(E)$ to compute an estimate of the phase space density, assuming an isotropic distribution of particle velocities (see Appendix for a more detailed description). We will see later that the velocity distribution remains essentially isotropic throughout the calculations.

Figure 4.3 shows the results for Run 5 ($N = 1.5 \times 10^5, \gamma = 1/2$). While $f(E)$ is clearly harder to estimate than $\rho(r)$ — it is effectively a 3/2 derivative of ρ and hence noisy — we can see from Figure 4.3 that the $f(E)$ extracted from the N-body runs evolves in a very similar way to the $f(E)$ computed via the Fokker-Planck equation, and that its steady state form is consistent (modulo the noise) with the Bahcall & Wolf solution $f(E) \propto E^{1/4}$ at late times. The same estimate is shown in Figure 4.4 from four other Runs listed in Table 1. We can also see here a very good agreement with the predictions from the Fokker-Planck equation.

The total mass in the stellar cusp ($r \lesssim r_h$) is typically of the order of $\alpha \times M_\bullet$, with $\alpha \lesssim 0.1 - 0.2$, as it can be seen from the plots that represent the cumulative mass distribution. This is a confirmation of the Fokker-Planck prediction that the mass in stars contained within the cusp is small compared with $M_\bullet$, and that the black hole effectively dominates the gravitational potential therein.

Note that that, in our runs, we have not included any dissipative effects resulting from the interaction between stars making very close pericenter passages to the black hole (eg. tidal disruption, prompt infall, or relativistic inspiral). It has been discussed in Chapter 2 that the characteristic spatial scales for these physical phenomena are very small and do not affect in a noticeable way the global structure of the stellar cusp at the radii that our particle numbers are able to resolve. This effect was quantified by the approximate Fokker-Planck calculations of Chapter 2 (see also [Bahcall and Wolf, 1977, Hopman and Alexander, 2006]). A few N-body calculations done with a mass spectrum and implementing the capture of stars by the black hole fully confirm these qualitative assertions.

Figure 4.4 shows the formation of Bahcall-Wolf cusps around black hole of masses 0.01, 0.025 and 0.05 of the total stellar cluster mass. All of them show a robust growth of the asymptotic cusp — both in spatial and phase space densities — on a time scale $\approx (0.5 - 0.6) t_{\rm rlx}(r_h)$. We conclude, therefore, from our N-body experiments, that the predictions of the Fokker-Planck equation are fully validated for stellar clusters made of stars with the same mass around a massive black hole.

G	γ	N	$M_\bullet/m_x$	$M_\bullet/M_{cl}$	r_h	$\ln\Lambda$	t_{rlx}	$T_{max}(T_{max}/t_{\mathrm{rlx}})$
$G1$	1/2	$250K$	2500	0.01	0.26	8.0	1690	1500(0.887)
$G2$	1/2	$100K$	500	0.005	0.19	6.6	454	500(1.101)
$G3$	1/2	$100K$	1000	0.01	0.26	7.1	764	800(1.047)
$G4$	1	$100K$	1000	0.01	0.17	6.6	227	200(0.881)
$G5$	1/2	$150K$	1500	0.01	0.26	7.5	1034	860(0.832)
$G6$	1/2	$120K$	1200	0.025	0.43	7.8	1844	1636(0.887)
$G7$	1/2	$120K$	6000	0.05	0.66	8.3	3912	2472(0.632)

Table 4.1: Single mass runs. 1^{st} column: slope of the Dehnen's model inner cusp at $t = 0$; 2^{nd}: number of particles in the simulation; 3^{rd}: ratio of BH mass to field star's mass; 4^{th}: ratio of BH mass to total cluster mass in stars; 5^{th}: influence radius; 6^{th}: relaxation time measured at the BH's influence radius; 8^{th}: maxumun integration time.

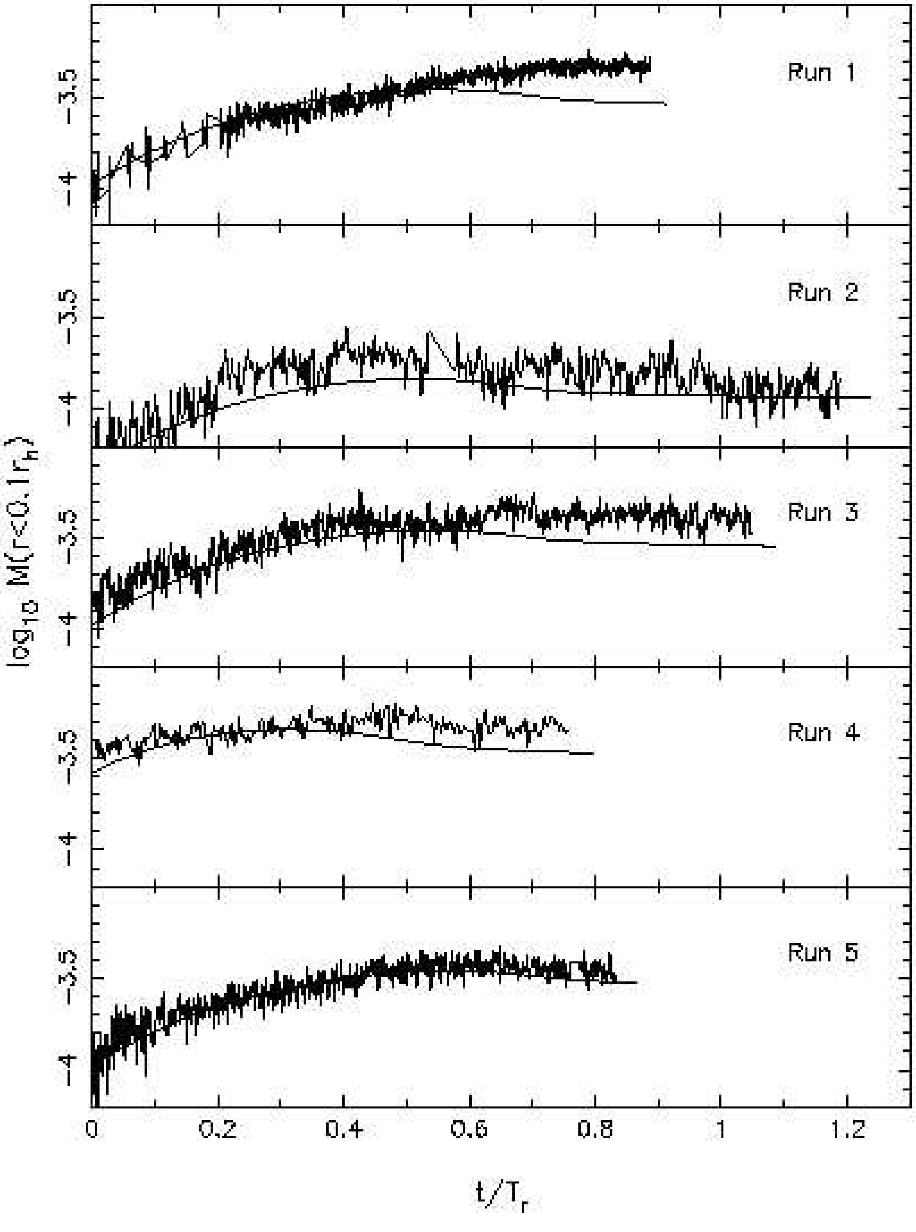

Figure 4.1: Evolution of the mass in stars within a distance $0.1r_h$ from the black hole, where r_h is the influence radius of the black hole measured at the time zero. Noisy curves are from the N-body runs; smooth curves are solutions to the Fokker-Planck equation [Preto et al., 2004].

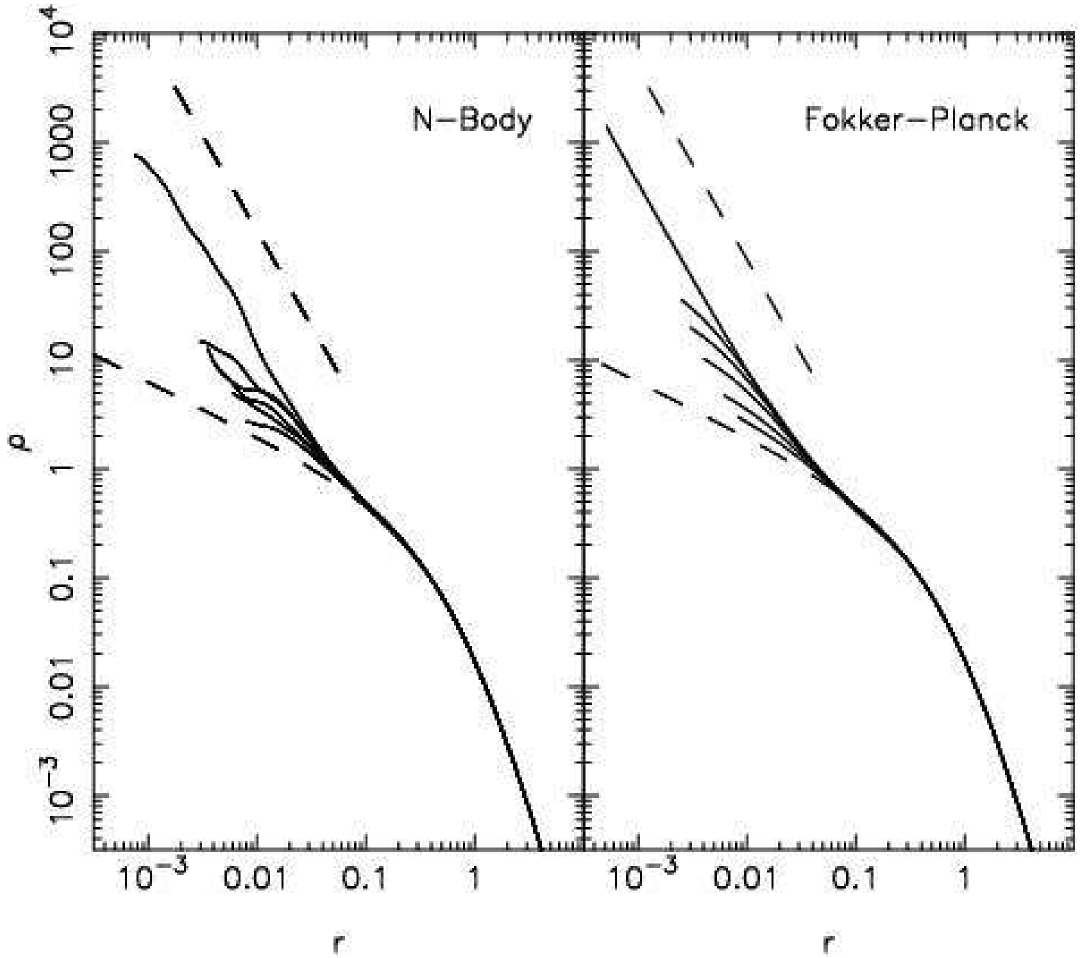

Figure 4.2: Evolution of the mass density profile around the black hole. Left panel: N-body $\rho(r)$ was estimated from the particle positions at times $t = 100, 200, 300, 400, 500, 1500$ via a maximum penalized likelihood algorithm. Right panel: Densities predicted from the Fokker-Planck equation at the same times; scaling of the time unit used the value of $\ln \Lambda$ given in 4.1. Lower dashed curves show the density at $t = 0$; upper dashed curves show $\rho \propto r^{-7/4}$, the asymptotic solution to the Fokker-Planck equation [Preto et al., 2004].

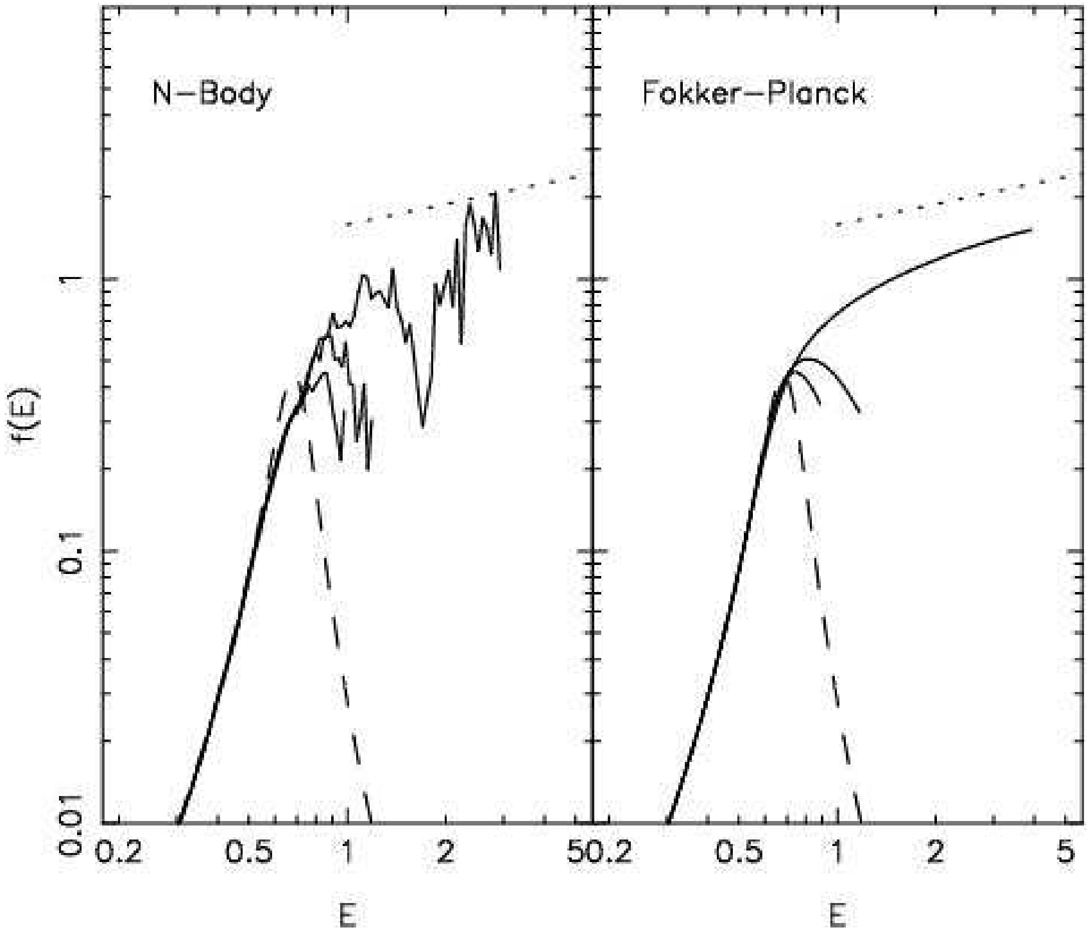

Figure 4.3: Evolution of the phase space density of stars around the black hole in Run 5. Left panel: N-body; $f(E)$ was estimated from the particle energies at times $t = 180, 300, 600$ as described in the text. Right panel: Densities predicted from the Fokker-Planck equation at the same times; scaling of the time unit used the value of $\ln \Lambda$ given in Table 1. Lower dashed curves show $f(E)$ at $t = 0$; upper dashed curves show $f \propto E^{1/4}$, the asymptotic solution to the Fokker-Planck equation [Preto et al., 2004].

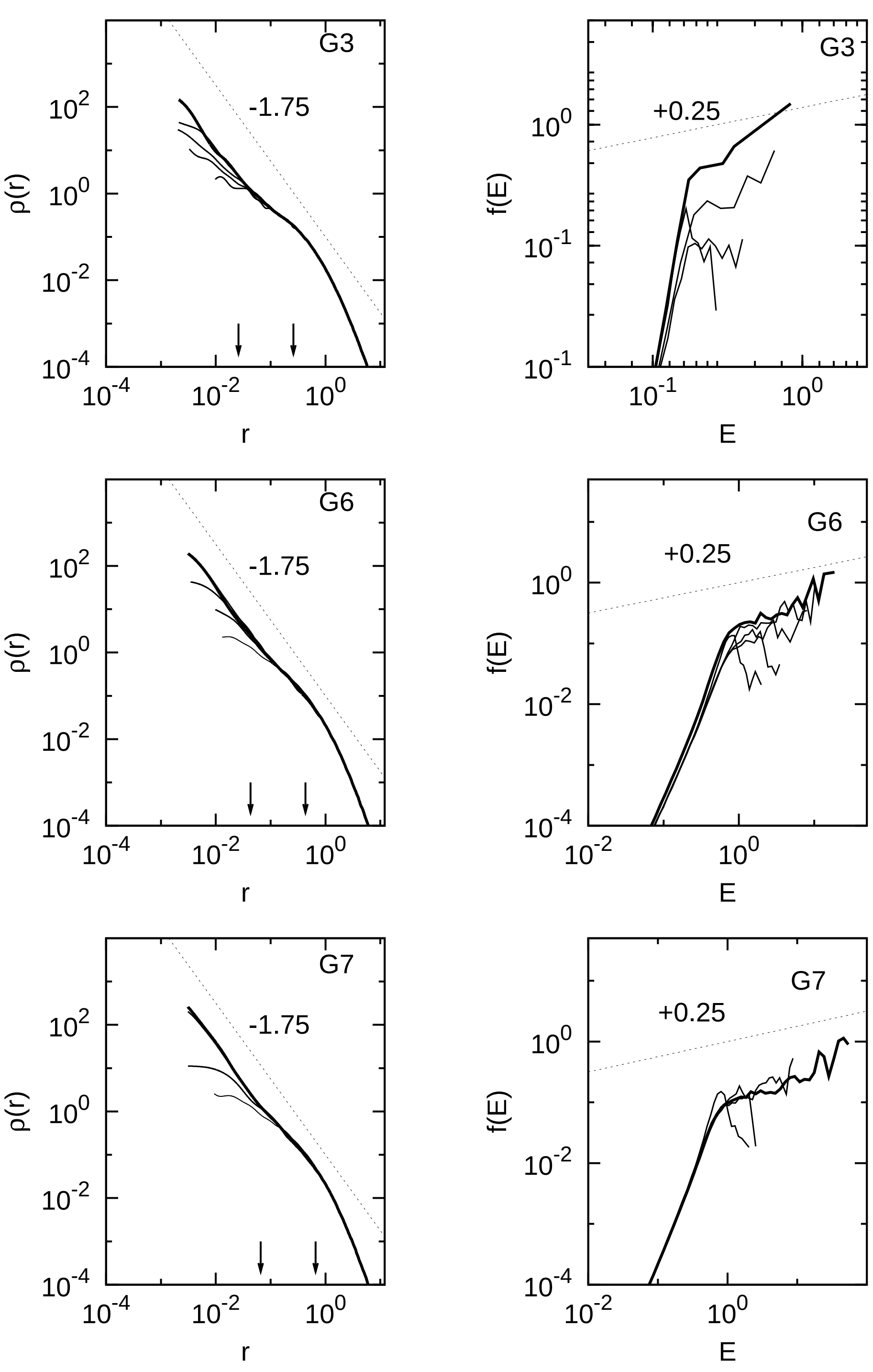

Figure 4.4: Growth of a Bahcall-Wolf cusp around the black hole in Runs 3, 6 and 7. Left panels: N-body evolution of the mass density profile of stars around the black hole for three black hole masses, $M_\bullet = (0.01, 0.025, 0.05)M_{cl}$. The density $\rho(r)$ was estimated from the particle's positions via a kernel density estimation method. The asymptotic solution $\rho \propto r^{-7/4}$ was reached in $\approx (0.5 - 0.6)t_{rlx}(r_h)$. Right panels: N-body evolution of the phase space density of stars around the black hole computed from the particle's energies. The evolution of $f(E)$ occurs on the same time as that of $\rho(r)$.

4.4 Double mass case

The initial conditions for the two-component runs require the specification of another two parameters in addition to those already present in the single mass case. The new parameters are the mass ratio of the individual stars in each component, $R = m_{*,2}/m_{*,1}$; and the number fraction of, say, the heavy component $f = N_2/N$, with $f \in [0, 1]$ and $N = N_1 + N_2$ where N_1 and N_2 are the number of stars in each of the two components. We need, in total, to specify five independent parameters. Henceforth, component 2 will denote the heavy stars and component 1 the light stars: $m_2 > m_1$. Given the values for R and f, together with the total number N of stars, we obtain the stellar mass for each component via

$$\begin{aligned} m_{*,1} &= \frac{1}{(1-f)+fR}\frac{1}{N} \\ m_{*,2} &= \frac{R}{(1-f)+fR}\frac{1}{N}, \end{aligned} \tag{4.12}$$

and the total mass of stars in each population by

$$\begin{aligned} M_1 = N_1 m_{*,1} &= \frac{1}{(1-f)+fR}\frac{N_1}{N} \\ M_2 = N_2 m_{*,2} &= \frac{R}{(1-f)+fR}\frac{N_2}{N}. \end{aligned} \tag{4.13}$$

The number of stars N_1 in the light component is $N_1 = (1-f)N$. In the case of a cluster with more than one stellar mass, the total mass density is given by

$$\rho(r) = 4\pi \int_{m_l}^{m_h} dm\ m \int_0^{\Psi} dE\ f(E,m)\sqrt{2(\Psi - E)}, \tag{4.14}$$

where m_l and m_h are the lightest and heaviest masses, respectively. In the particular case of a two-component cluster

$$f(E,m) = f_l(E)\delta(m - m_l) + f_h(E)\delta(m - m_h). \tag{4.15}$$

We have chosen to perform a series of runs where essentially the parameters R and f are varied; otherwise, the black hole mass at $M_\bullet = 0.05$, the initial cusp slope $\gamma = 1/2$, and the total number of stars $N = 120K$ are kept fixed. The runs $G8$, $G9$ and $G10$ are the exception in that they were made with a lower black hole mass 0.01. The total mass of the stellar cluster remains equal to one in all cases. The set of two-component runs and corresponding parameters is listed in Table 2.

The procedure to start a two-component run is the following: start to generate a Monte Carlo realization of a Dehnen model as for a single mass run; choose the parameters R and f, and assign randomly a mass $m_{*,1}$ or $m_{*,2}$ to every star according to a probability N_1/N and N_2/N, respectively. The resulting Monte Carlo generated model has a density profile that is practically indistinguishable from those of the equal mass models. The model is in an almost dynamical equilibrium; deviations from virial equilibrium are always smaller than $\sim 1-2\%$ — as compared with $\sim 1/\sqrt{N}$, or roughly $0.2-0.3\%$ for $N = 120K$, in the equal mass case. On a dynamical time scale after the beginning of the simulation, phase mixing occurs and the system also effectively reestablishes the virial equilibrium as close as possible to unity. Following this prescription, the two-component models start without any mass segregation and are not in equipartition since both components are endowed with the same velocity dispersion. An initial state with no mass segregation may be motivated astrophysically as follows: current theories of galaxy formation predict the typical galaxy is built up through a large series of minor mergers and a few larger mergers. A merger between the nuclei of the galaxies probably induces violent relaxation and therefore will erase — partially or totally — the memory from the mass segregation state of the parent nuclei.

The Coulomn logarithm $\ln\Lambda$ is, for the multi-mass case, an even more poorly defined quantity than before because, not only the model is inhomogeneous, but now there is an extra complication brought by the interaction between stars with different masses. In this situation, it is costumary to substitute the single mass m_* of the old expression for $\ln\Lambda$ by the average stellar mass $< m_* >$ of the model; and the total velocity dispersion, taking into account both components, is used in the expression for $\ln\Lambda$. The time scale for diffusion in phase space of a given test star does not depend on its own mass (see equation (3.54), since $t_{\rm rlx} \sim v^2/d < (\Delta v)^2 > /dt$ corresponds to the first integral in 3.54). On the other hand, the characteristic time scale for the test star to decellerate by dynamical friction does depend upon its own mass (second integral from 3.54): $t_{df} \sim v/d < \Delta v > /dt \approx (m_1/m_2)t_{rlx,2}$ as it was argued there. Therefore, the energy exchange between the heavy stars of mass m_2 and light stars with mass m_1 happens at a faster rate than that between equal mass stars. As a result, the heavy stars can loose kinetic energy fast so long as there is a sufficiently large amount of mass made up of the lighter component. This is the case in real galaxies where the lighter component is usually dominant both in mass and number. Hence this imples that there is a significant amount of mass segregation before any significant amount of relaxation has taken place and we expect to observe interest dynamical evolution on a time scale significantly shorter than $t_{\rm rlx}$. The equipartition time scale turns out, however, to be of little usefulness since, as discussed on Chapter 3, the system is kept at all times far from equipartition.

Figures 4.8, 4.9 and 4.10 show the evolution in time of each component's mass inside the radii $0.1r_h$, $0.333r_h$ and $0.666r_h$. There is a very nice agreement in all runs between the curves obtained from the Fokker-Planck equation and the N-body simulations, although the N-body curves for the heavy component shows a tendency, in some runs, to rise above the corresponding Fokker-Planck curves reflecting the mass overdensity of heavy stars at the innermost resolved radii in those runs. The curves for the light component, on the other hand, agree almost perfectly between the two methods in all runs.

Figures 4.11 and 4.12 show the results of two mass components simulations with mass ratios of 3 and 5, respectively. The initial conditions were Dehnen models with an inner cusp slope of $\gamma = 1/2 + \delta$, black hole masses of $M_\bullet = 0.01$ in both cases. G9 can be directly compared to the Fokker-Planck calculations displayed in Figure 3.4, as the parameters are exactly the same in both cases. Focusing first on G9, it can be seen that the asymptotic inner slope $\gamma \approx 1.75$ of the heavy component extends outwards up to a radius $\approx 0.1 r_h$ as in the Fokker-Planck result; furthermore, the N-body run shows a hint of a slight increase in the cusp slope above the 1.75 value which is entirely consistent with the Fokker-Planck result. However, in order for this latter feature to be conclusive, it would be necessary to increase considerably the particle number to resolve the radial region inside 10^{-3}. This is a general conclusion from this work: it will be extremely important to make further simulations of the cusp formation with higher particle numbers in order to increase the resolution at the center. The N-body density profile for the light component settles into a 1.5 cusp slope for $r \lesssim 0.05 r_h$, also in agreement with the Fokker-Planck equation. The results from the run G8 are qualitatively similar to those of G9, if there is any difference is in that they display an even more clean realization of Bahcall-Wolf cusps and corresponding scaling relations.

In Figure 4.12 the cumulative mass distribution is shown. These were constructed in the following manner: first, the particles were separated according to which mass component they belong to; second, they were sorted, within each component, with respect to the radial distance relative to the instantaneous position of the black hole particle; third, the total stellar mass enclosed inside the radius at each star's position was recorded and plotted as a function of the above mentioned radial distance. Formally, these plots correspond to the volume integral of the spherically symmetric mass density distribution: $M(<r) = 4\pi \int_0^r dr'\ r'^2\ \rho(r',t)$ which is $\propto r^{3-\gamma}$ in the case $\rho \propto r^{-\gamma}$. This is a more direct and crude measure than the kernel density estimation method used to recover $\rho(r)$, and it is used here also to serve as a check on the reliability of the latter. In general, we find that the results are consistent between these two methods.

A note of caution should be made here concerning the interpretation of the plots in the pages ahead. In some cases, especially when the number fraction of stars in the heavy component is rather low, a superposition of snapshots was made to reduce the noise in estimating $\rho(r)$ and increase the number of particles when plotting $M(<r)$. In the latter case, for example, in order to keep constant the total mass plotted, the mass of each particle was divided by the total number of snapshots employed. As a result of this procedure, in some cases the curve of the heavy component seems to extend to smaller value in mass than that of the lighter component. This is, obviously, not a real feature and results only from having chosen a different normalization for each case.

The more interesting parameters to vary in the two-mass component runs are the mass ratio R of the stars and the number fraction f_1, f_2 in each component. The equal number fraction case studied in the runs G8 and G9, corresponding to typical cases studied originally by Bahcall & Wolf 1977, is very artificial since in all realistic mass distributions there is inevitably a trend for the decrease of the number of stars with increasing stellar mass. For this reason, a set of simulations was performed where the number fraction in heavy stars is varied and takes the following values: 3.75×10^{-3}, 7.5×10^{-3}, 0.1 and 0.2. In all cases, there are very clear realizations of the Bahcall-Wolf cusp for both stellar components, with cusp slopes that are consistent with the Fokker-Planck predictions.

The data sorted out, for each component, in order to produce the $M(< r)$ plots was also used to estimate the numerical values of the cusp's inner slopes via a least-squares fit method. In Tables 3 and 4, there are columns for the values of γ_l and γ_h measured from all the particle's positions inside r_h, or $0.1r_h$ (the values corresponding to the latter are shown between parenthesis). The asymptotic values expected for Bahcall & Wolf cusps are $\gamma_l = 1.5$ and $\gamma_h = 1.25$. From the inspection of the tables, it is easy to recognize that, for most cases, the asymptotic slopes obtained from the N-body runs for $r \lesssim r 0.1 r_h$ are consistent with the Fokker-Planck solutions; in contrast, for radii $\gtrsim r_h$ the contribution from the star's self-gravity is already strong enough to perturb the density profile away from those values. That this is indeed the case can be confirmed by the fact that the total mass in stars is $M(< r_h) \approx (1.5 - 2)M_\bullet$, as indicated in the Tables. Note that $M(< 0.1r_h) \approx (0.1 - 0.2)M_\bullet$ showing that the cusp's gravitational potential is indeed dominated by the hole's potential. The brackets in the Table around N_{cups} and M_{cusp} refer to the average values taken from the superposition of snapshots ($\sim 10 - 20$) used to produce the $\rho(r)$ and $M(< r)$ plots, and to make the least-squares fit to the data as well.

The Fokker-Planck calculations for a two-component cluster already predicts a cusp slope greater than 7/4 for the heavy component. This is also obtained from the N-body runs which, furthermore, show an interesting and intriguing extra-feature: the heavy stars develop an overdensity at the innermost resolved radii. This tendency is more pronounced in the cases with higher mass ratio, eg. runs G32, G35 and G49. At the same time, the light stars seem to undergo a slight expansion, with the consequent decrease in density, and settle into a quasi steady-state with a cusp slope ~ 1.5 broadly consistent with the Bahcall & Wolf predictions. However, we can also observe in these same runs some oscillations in the value of γ_l which, at times, decrease to values such as $1.2 - 1.4$.

This earlier expansion in the light component for two-component runs with larger mass ratios can also be seen from the time evolution of $M(< r)$ in figures 4.9 and 4.10 — both from the Fokker-Planck and the N-body curves. It did not escape our attention that these overdensities in the heavy stars at the smallest radii — if present in real galaxies — might translate into an overdensity in the form of a central sub-cluster of stellar mass black holes. This has already been suggested on theoretical grounds [Miralda-Escudé and Gould, 2000]. If this were true, it could increase the estimates of EMRI event rates and therefore the chances for LISA to detect gravitational waves emmited by these objects as they slowly inspiral towards the central, massive black hole. In fact, what is required to produce more EMRIs are physical mechanisms which speed up the diffusion of stars to higher energies, or small pericenters, but slows down once these small pericenters are reached (those at which the timescale for energy loss due to gravitational wave emission is shorter than the local relaxation time). A sudden strong effect such as a close encounter with another star may, for some encounter parameters, be precisely such a mechanism. This is, however, outside the scope of this work and it is left for (near-)future work.

While we have no certain explanation yet for this effect, it could indicate a consistent deviation from the Fokker-Planck predictions if it is not simply noise due to the small particle numbers at the inner regions. It is therefore important to perform simulations with better resolution at the center. In case this result persists then, it is quite possible it could arise from a non-negligible ejection rate of the light stars from the cusp, due to close encounters with the heavy stars. The heavy star would, by energy conservation, become more strongly bound to the black hole. The importance of strong encounters is entirely neglected under the approximations made in the derivation of the Fokker-Planck equation. The usual justification for neglecting strong encounters between stars is that the relaxation time is shorter than the mean time between encounters that lead to $\gtrsim 90^o$ deflection angle by a factor $1/8 \ln \Lambda$, where $\ln \Lambda \approx 7 - 8$ in our simulations and $\ln \Lambda \approx 15$ for a galactic nuclei with $N = 10^7$ stars. Now, this certainly suggests that strong encounters are rare events and thus should not affect the mean values of the global properties of the stellar cluster. This is indeed confirmed by the N-body simulations reported in this thesis. However, none of this implies a zero rate of close encounters nor does it imply that they do not happen at all.

If this hypothesis is correct, the dependence on the mass ratio R is clearly understandable from the physical point of view: stars can escape from the cusp as a result of a single strong encounter with another star and, by energy conservation, its encounter pair will fall deeper in the hole's potential well. Moreover, the rate of escape of stars due to strong encounters with another single star is very sensitive to the stellar mass distribution and can be increased by an order of magnitude, or more, if an equal mass population is substituted by a spectrum of masses [Hénon, 1969]. In fact, Hénon made an estimate of the rate of escape of stars from a Plummer model and concluded that it increases approximately as $(m_2/m_1)^{5/2}$. We need to be aware, nonetheless, that his calculation was made for a system with a homogeneous core density and constant velocity dispersion, rather than for the more complex situation we face of an inhomogeneous stellar cusp around a massive black hole.

The appearance of an overdensity at small radii, for the heavy component, whenever the mass ratio R is sufficiently large is a very interesting feature in the results of these simulations. It is very suggesting of the possibility of the formation, or existence, of a central sub-cluster — presumably, built up of heavy compact remnants — at very small distance from the central massive black hole. This question is, by itself, worth of a serious independent study and for now it is left for a future work.

G	R	f_h	$M_\bullet/m_h$	M_h/M_{cl}	$t_{\rm rlx}$	$T_{max}(T_{max}/t_{\rm rlx})$
$G8$	3	1/2	800	0.75	871	527 (0.605)
$G9$	5	1/2	720	0.833	871	389 (0.447)
$G30$	5	0.1	1680	0.357	3912	1000 (0.256)
$G32$	20	0.1	870	0.690	3912	696 (0.178)
$G49$	50	0.1	708	0.847	3912	1781 (0.0.455)
$G51$	2	0.2	3600	0.333	3912	1395 (0.357)
$G50$	3	0.2	2800	0.428	3912	1511 (0.386)
$G35$	50	0.2	1296	0.923	3912	1059 (0.271)
$G70$	50	0.0075	164	0.274	3912	1302 (0.333)
$G71$	20	0.0075	342	0.131	3912	1340 (0.342)
$G72$	20	0.00375	342	0.131	3912	1321 (0.338)

Table 4.2: Double mass runs. 1st column: heavy-to-light component mass ratio ; 2nd: fraction of particles in the heavy component; 3rd: ratio of BH mass to heavy star's mass; 4th: ratio of BH mass to total cluster mass in stars; 5th: single-mass cluster's relaxation time measured at the BH's influence radius ; 6th: maxumun integration time.

Gal	$M_\bullet$	R	f_h	r_{cusp}	$< N_{cusp,l} >$	$< M_{cusp,l} > /M_\bullet$	γ_l
$G8$	0.01	3	0.5	0.26 (0.06)	860 (34)	0.358 (0.014)	2.010 (1.575)
$G9$	0.01	5	0.5	0.26 (0.06)	852 (93)	0.237 (0.026)	2.120 (1.545)
$G30$	0.05	5	0.1	0.66 (0.06)	8428 (129)	0.957 (0.015)	1.758 (1.445)
$G32$	0.05	20	0.1	0.66 (0.06)	5757 (59)	0.314 (0.0032)	1.933 (1.484)
$G49$	0.05	50	0.1	0.66 (0.06)	2879 (20)	0.077 (5.36×10^{-4})	1.356 (1.430)
$G51$	0.05	2	0.2	0.66 (0.06)	3311 (81)	0.442 (0.011)	1.283 (1.349)
$G50$	0.05	3	0.2	0.66 (0.06)	2731 (112)	0.309 (0.013)	1.721 (1.456)
$G35$	0.05	50	0.2	0.66 (0.06)	9066 (45)	0.134 (6.93×10^{-4})	1.973 (1.550)
$G70$	0.05	50	0.0075	0.66 (0.06)	4884 (37)	0.572 (0.004)	2.048 (1.442)
$G71$	0.05	20	0.0075	0.66(0.06)	9744 (121)	1.358 (0.017)	1.831 (1.347)
$G72$	0.05	20	0.00375	0.66 (0.06)	9660 (116)	1.435 (0.017)	1.818 (1.414)

Table 4.3: Double mass runs: Light component. 1st column: BH mass ; 2nd column: heavy-to-light component mass ratio ; 3nd: fraction of particles in the heavy component; 4th: radius for measuring cusp slope; 5th: average number of light stars inside the cusp; 6th: average mass in light stars inside the cusp (in units of the BH mass); 7th: estimate cusp slope for the light component.

Gal	$M_\bullet$	R	f_h	r_{cusp}	$< N_{cusp,h} >$	$< M_{cusp,h} > /M_\bullet$	γ_h
$G8$	0.01	3	0.5	0.26 (0.06)	1250 (97)	1.562 (0.121)	1.575 (1.256)
$G9$	0.01	5	0.5	0.26 (0.06)	1276 (86)	1.772 (0.119)	1.632 (1.261)
$G30$	0.05	5	0.1	0.66 (0.06)	1465 (62)	0.832 (0.035)	1.318 (1.460)
$G32$	0.05	20	0.1	0.66 (0.06)	1233 (45)	1.344 (0.049)	1.351 (0.857)
$G49$	0.05	50	0.1	0.66 (0.06)	987 (33)	1.322 (0.049)	1.225 (0.873)
$G51$	0.05	2	0.2	0.66 (0.06)	1187 (55)	0.314 (0.012)	1.224 (1.320)
$G50$	0.05	3	0.2	0.66 (0.06)	3846 (104)	1.308 (0.035)	1.328 (1.04)
$G35$	0.05	50	0.2	0.66 (0.06)	2347 (312)	1.727 (0.230)	1.397 (0.937)
$G70$	0.05	50	0.0075	0.66 (0.06)	121 (10)	0.709 (0.059)	1.030 (0.476)
$G71$	0.05	20	0.0075	0.66(0.06)	291 (30)	0.809 (0.083)	0.996 (1.052)
$G72$	0.05	20	0.00375	0.66 (0.06)	168 (18)	0.500 (0.054)	0.982 (0.983)

Table 4.4: Double mass runs: Heavy component. 1st column: BH mass ; 2nd column: heavy-to-light component mass ratio ; 3nd: fraction of particles in the heavy component; 4th: radius for measuring cusp slope; 5th: average number of heavy stars inside the cusp; 6th: average mass in heavy stars inside the cusp (in units of the BH mass); 7th: estimate cusp slope for the heavy component.

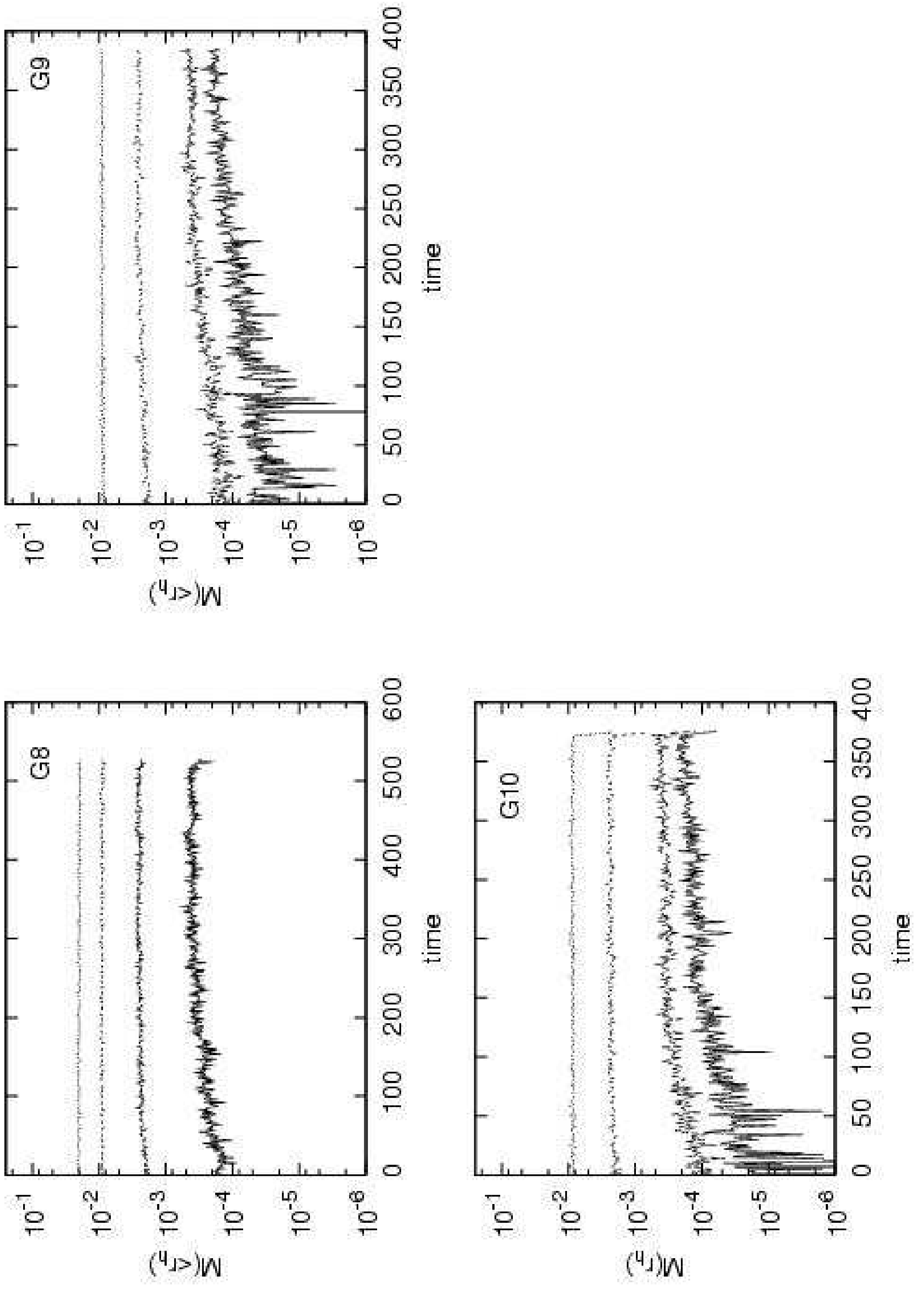

Figure 4.5: Time evolution of the total mass (light and heavy components) $M(< r, t)$ enclosed inside radius r. The radii in the figure are $G8 : 0.1r_h$, $0.333r_h$, $0.666r_h$, r_h and $G9 : 0.05r_h$, $0.1r_h$, $0.333r_h$, $0.666r_h$, where r_h is the radius of influence of the black hole measured at the initial time.

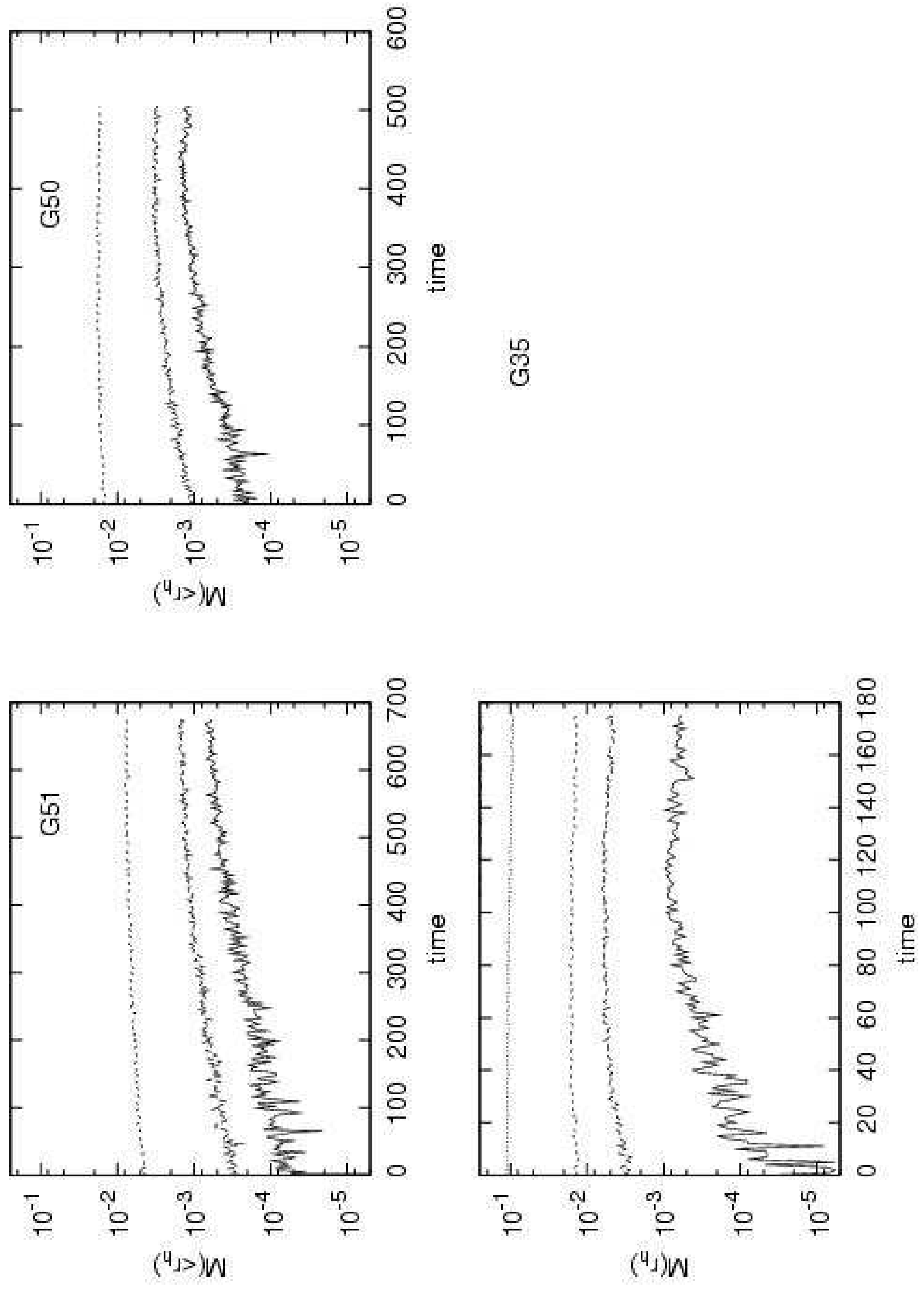

Figure 4.6: Time evolution of the total mass (light and heavy components) $M(< r, t)$ enclosed inside radius r. The radii in the figure are $G51 : 0.05r_h, 0.1r_h, 0.333r_h$, $G50 : 0.05r_h, 0.1r_h, 0.333r_h$ and $G35 : 0.05r_h, 0.1r_h, 0.333r_h, 0.666r_h$,where r_h is the radius of influence of the black hole measured at the initial time.

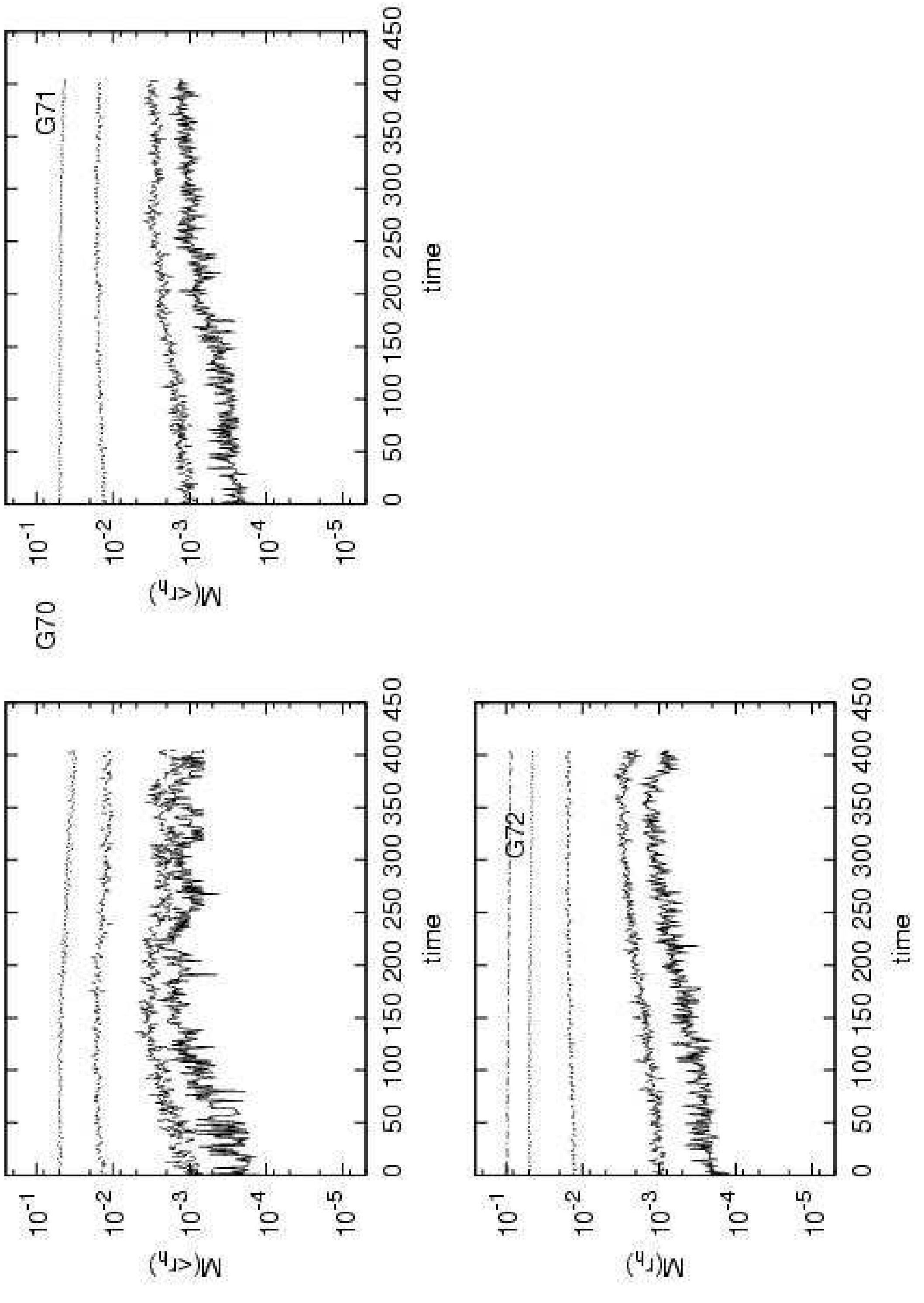

Figure 4.7: Time evolution of the total mass (light and heavy components) $M(< r.t)$ enclosed inside radius r. The radii in the figure are $G70$: $0.05r_h$, $0.1r_h$, $0.333r_h$, $0.666r_h$, $G71$: $0.05r_h$, $0.1r_h$, $0.333r_h$, $0.666r_h$ and $G72$: $0.05r_h$, $0.1r_h$, $0.333r_h$, $0.666r_h$, r_h, where r_h is the radius of influence of the black hole measured at the initial time.

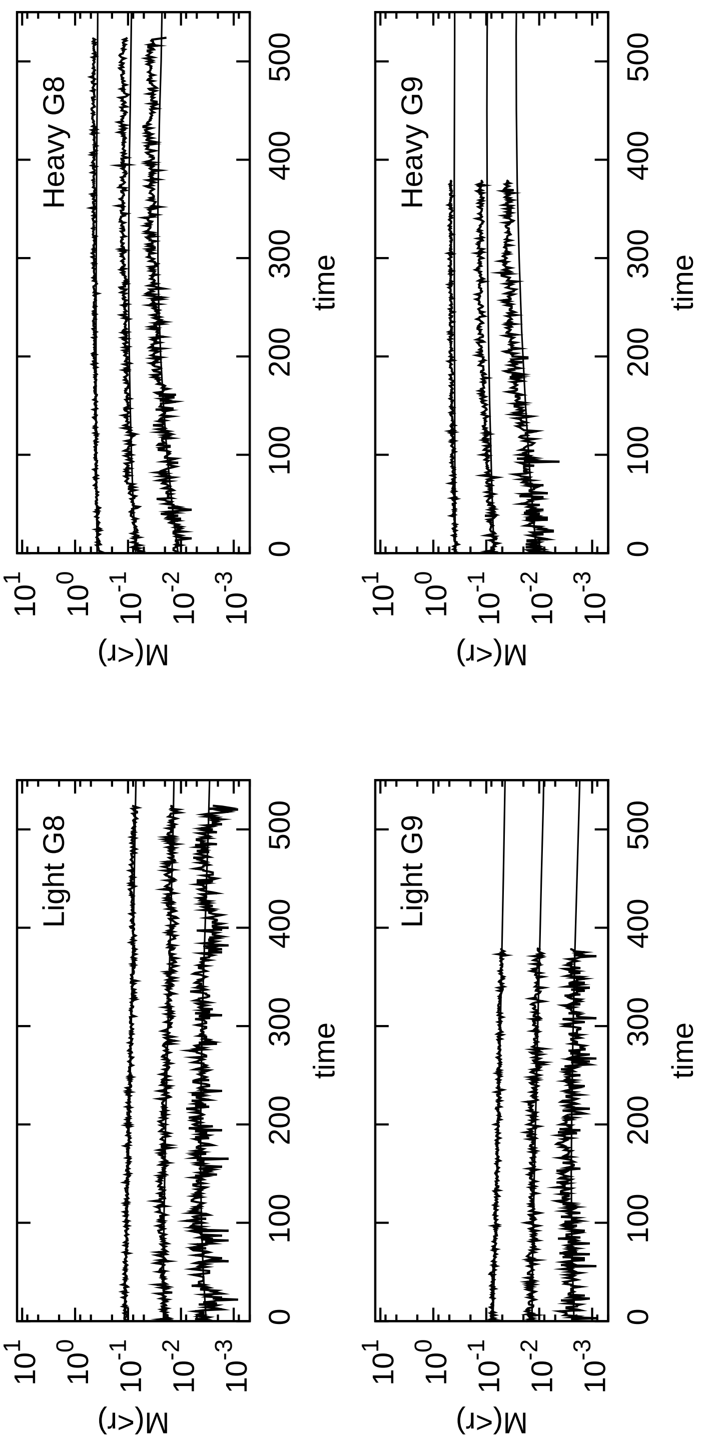

Figure 4.8: Time evolution of the mass enclosed inside the radius r, $M_l(< r, t)$ and $M_h(< r, t)$, for both components; these were computed from the Fokker-Planck equation (smooth curves) and from the N-body runs (noisy curves). Left panel: Light component. Right panel: Heavy component. The radii are defined by the distances $0.1r_h$, $0.23r_h$ and $0.5r_h$ from the black hole, where r_h is the radius of influence of the black hole measured at the initial time.

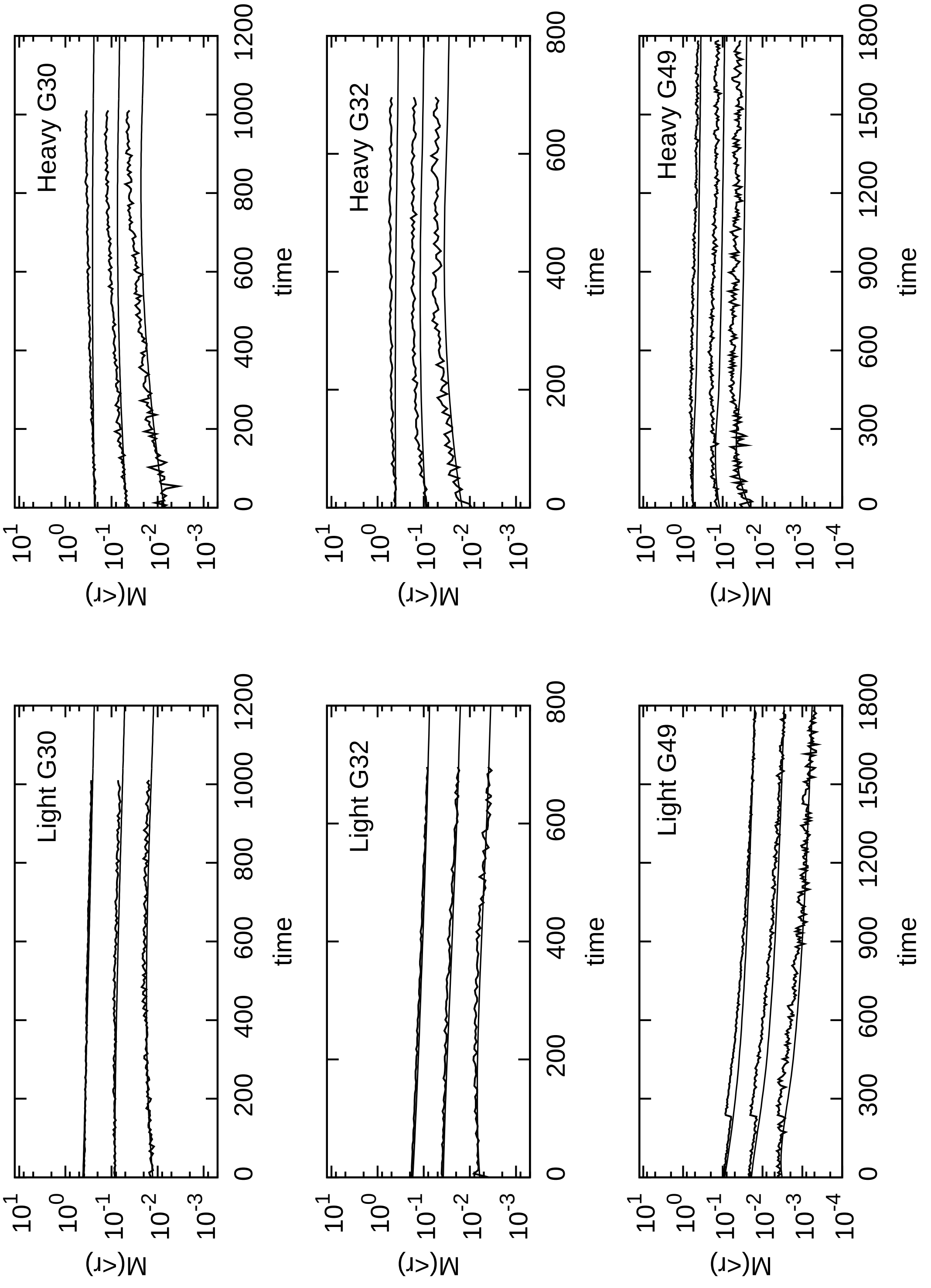

Figure 4.9: Time evolution of the mass enclosed inside the radius r, $M_l(< r, t)$ and $M_h(< r, t)$, for both components; these were computed from the Fokker-Planck equation (smooth curves) and from the N-body runs (noisy curves). Left panel: Light component. Right panel: Heavy component. The radii are defined by the distances $0.1r_h$, $0.23r_h$ and $0.5r_h$ from the black hole, where r_h is the radius of influence of the black hole measured at the initial time.

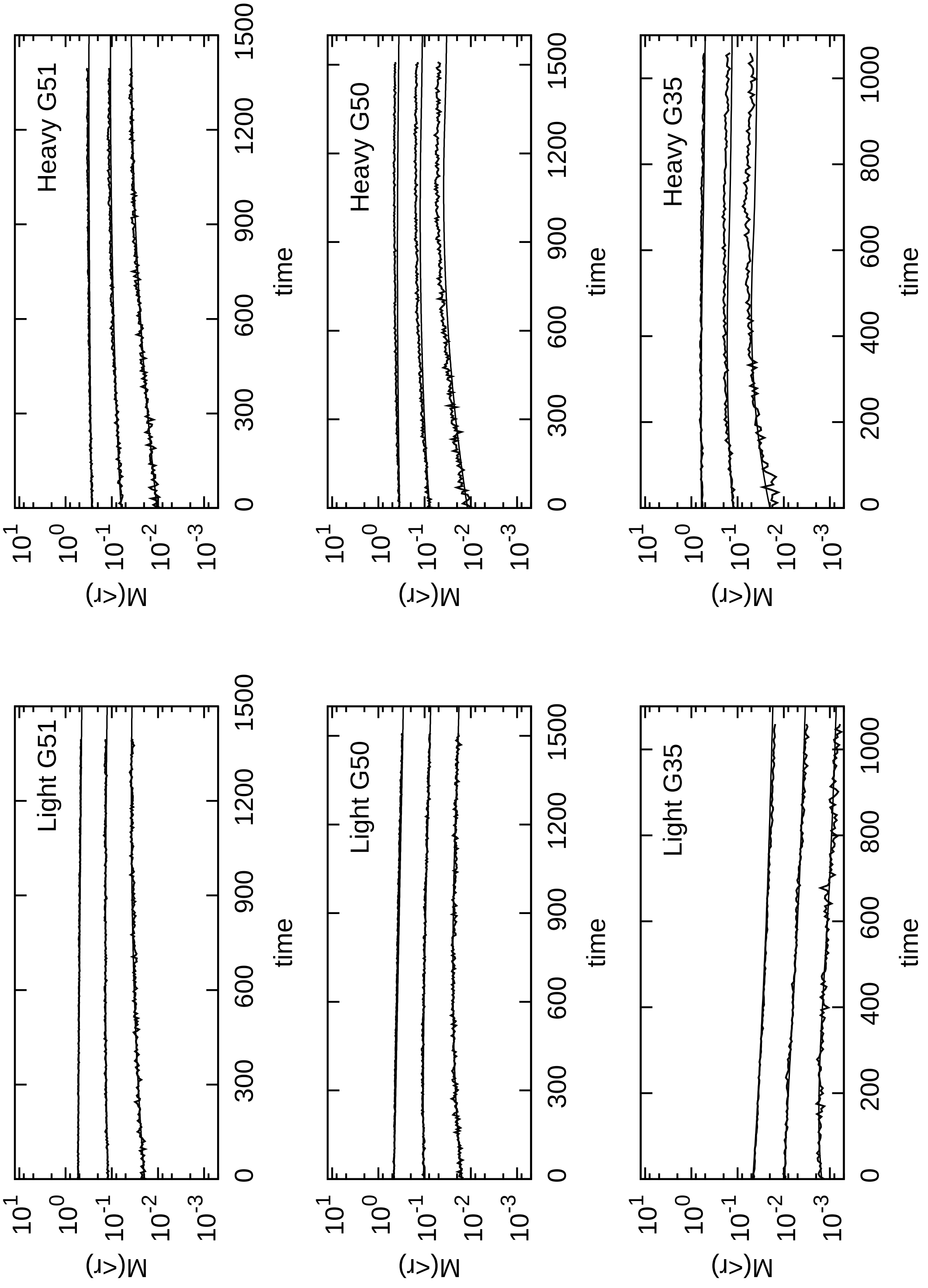

Figure 4.10: Time evolution of the mass enclosed inside the radius r, $M_l(< r, t)$ and $M_h(< r, t)$, for both components; these were computed from the Fokker-Planck equation (smooth curves) and from the N-body runs (noisy curves). Left panel: Light component. Right panel: Heavy component. The radii are defined by the distances $0.1r_h$, $0.23r_h$ and $0.5r_h$ from the black hole, where r_h is the radius of influence of the black hole measured at the initial time.

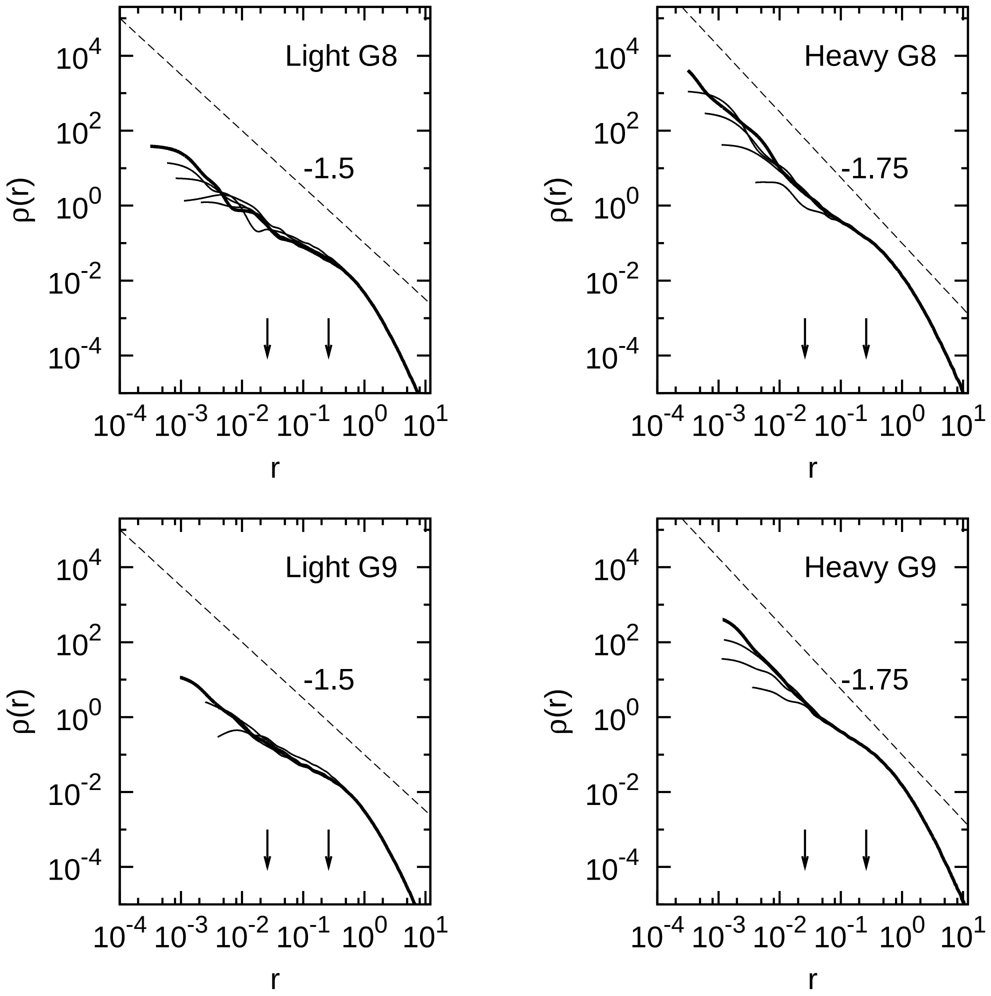

Figure 4.11: Evolution of the mass density profile around the black hole. Left panels: N-body $\rho(r)$ for the light component was estimated from the particle's positions via a kernel density estimation method. Right panels: N-body $\rho(r)$ for the heavy component computed according to the same method. The arrows point to distances $0.1r_h$ and r_h from the black hole. The initial slope was $\gamma_l = \gamma_h = 1/2$; it evolves, after $\approx 0.6t_{rlx}$ for G8 and $\approx 0.45t_{rlx}$ for G9, to a value consistent with the values predicted from the Fokker-Planck equation for radius $r \leq 0.23r_h$ (light component) and $r_h \leq 0.1r_h$ (heavy component).

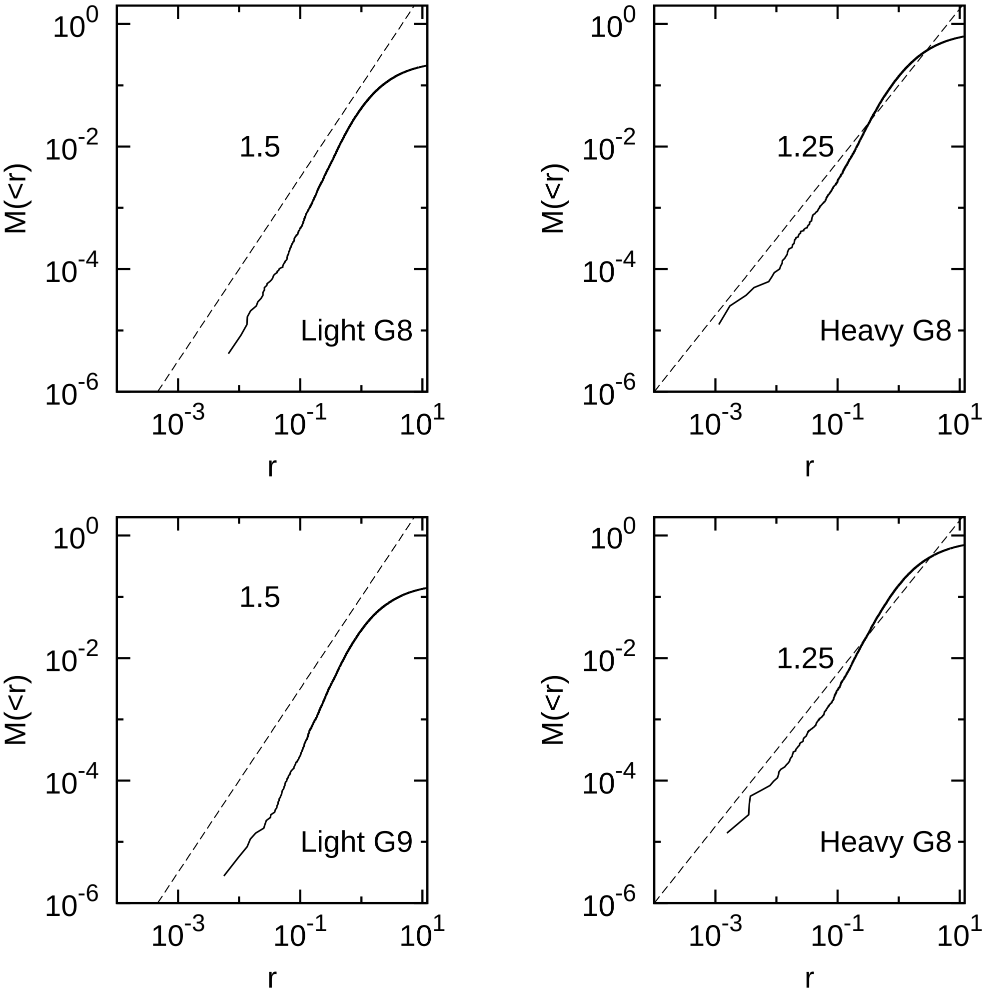

Figure 4.12: The cumulative mass distribution $M(< r)$ at late times was extracted from the N-body runs (see description in the text). The inner slopes $M(< r) \propto r^{3-\gamma}$ inside the stellar cusp, are consistent with those predicted from the Fokker-Planck equation for radii $r \leq 0.23 r_h$ (light component) and $r_h \leq 0.1 r_h$ (heavy component). See Tables 4 and 5 for the numerical values of the inner slopes that were computed from this plot's data.

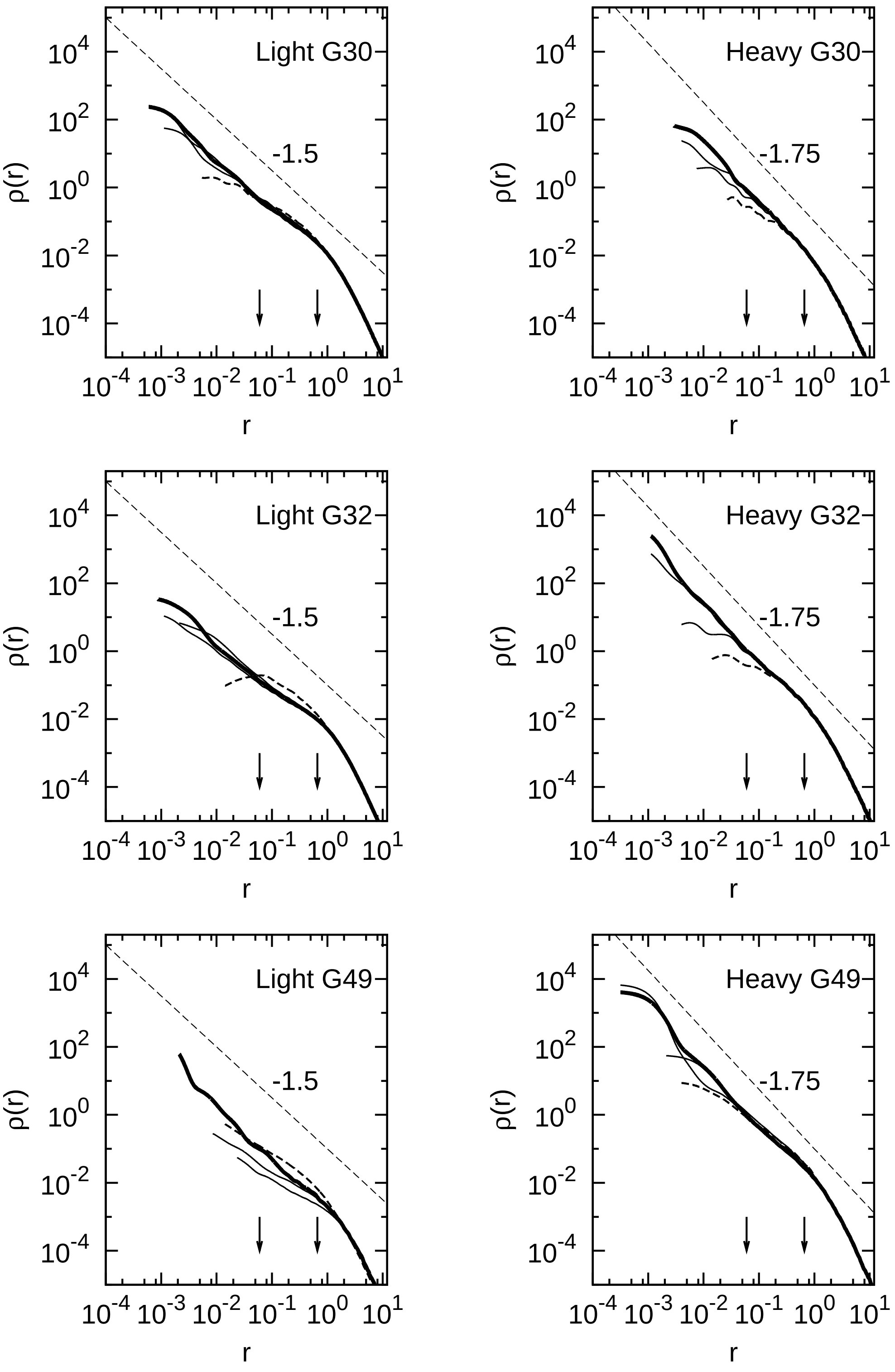

Figure 4.13: Evolution of the mass density profile around the black hole. Left panels: N-body $\rho(r)$ for the light component was estimated from the particle's positions. Right panels: N-body $\rho(r)$ for the heavy component. The arrows point to distances $0.1r_h$ and r_h from the black hole. The initial slope was $\gamma_l = \gamma_h = 1/2$; it evolves, after $\approx 0.t_{rlx}$ to a value consistent with the values predicted from the Fokker-Planck equation for radii $r \leq 0.1r_h$.

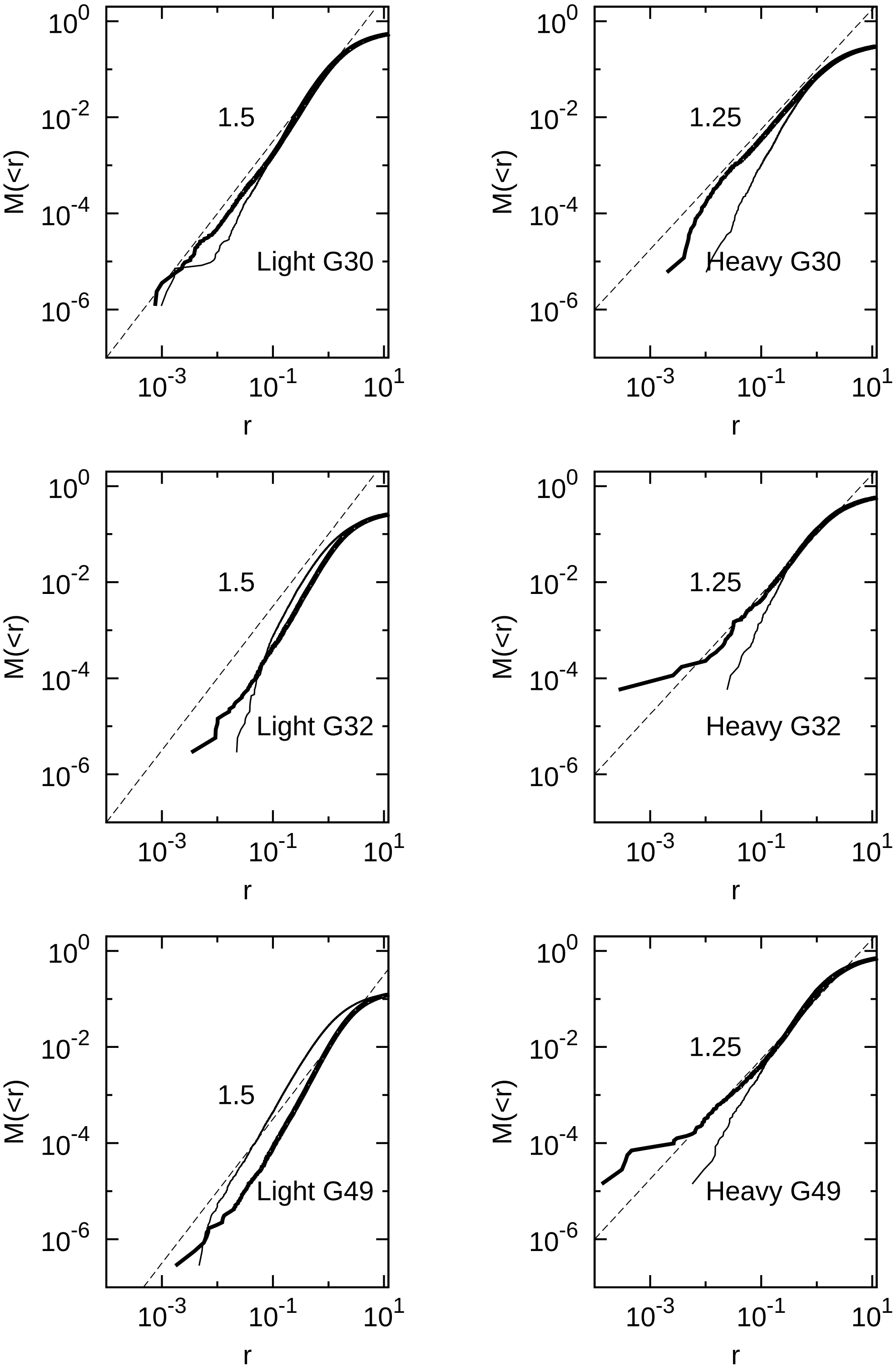

Figure 4.14: The cumulative mass distribution $M(< r)$ at late times was extracted from the N-body runs (see description in the text). The light curve corresponds to the initial time; the heavy curve to the distribution at late times. The inner slopes $M(< r) \propto r^{3-\gamma}$ inside the stellar cusp, are consistent with those predicted from the Fokker-Planck equation for radii $r \leq 0.1 r_h$. See Tables 4 and 5 for the numerical values of the inner slope that were computed from this plot's data.

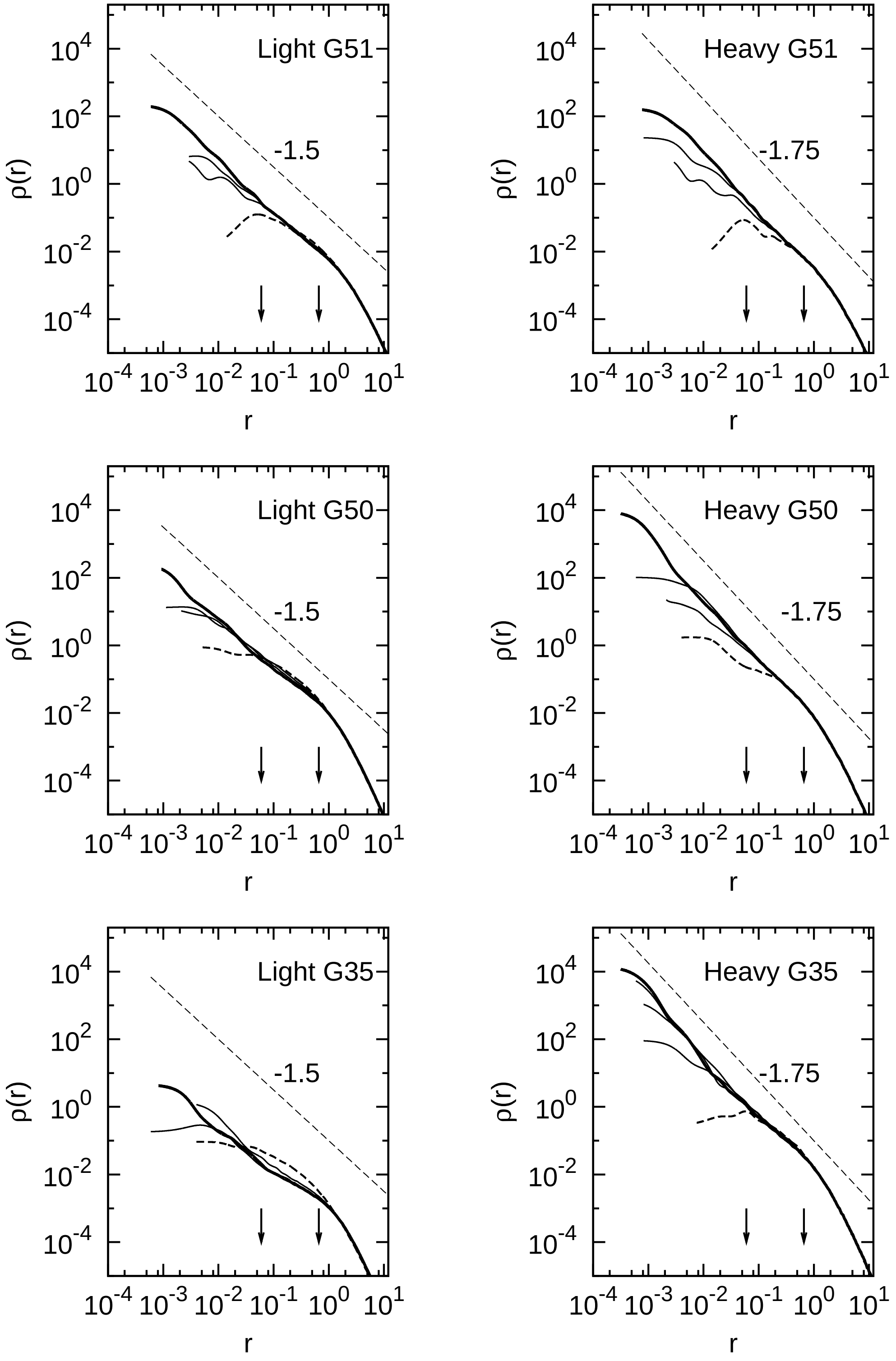

Figure 4.15: Evolution of the mass density profile around the black hole. Left panels: N-body $\rho(r)$ for the light component was estimated from the particle's positions. Right panels: N-body $\rho(r)$ for the heavy component. The arrows point to distances $0.1r_h$ and r_h from the black hole. The initial slope was $\gamma_l = \gamma_h = 1/2$; it evolves, after $\approx 0.t_{rlx}$ to a value consistent with the values predicted from the Fokker-Planck equation for radii $r \leq 0.1r_h$.

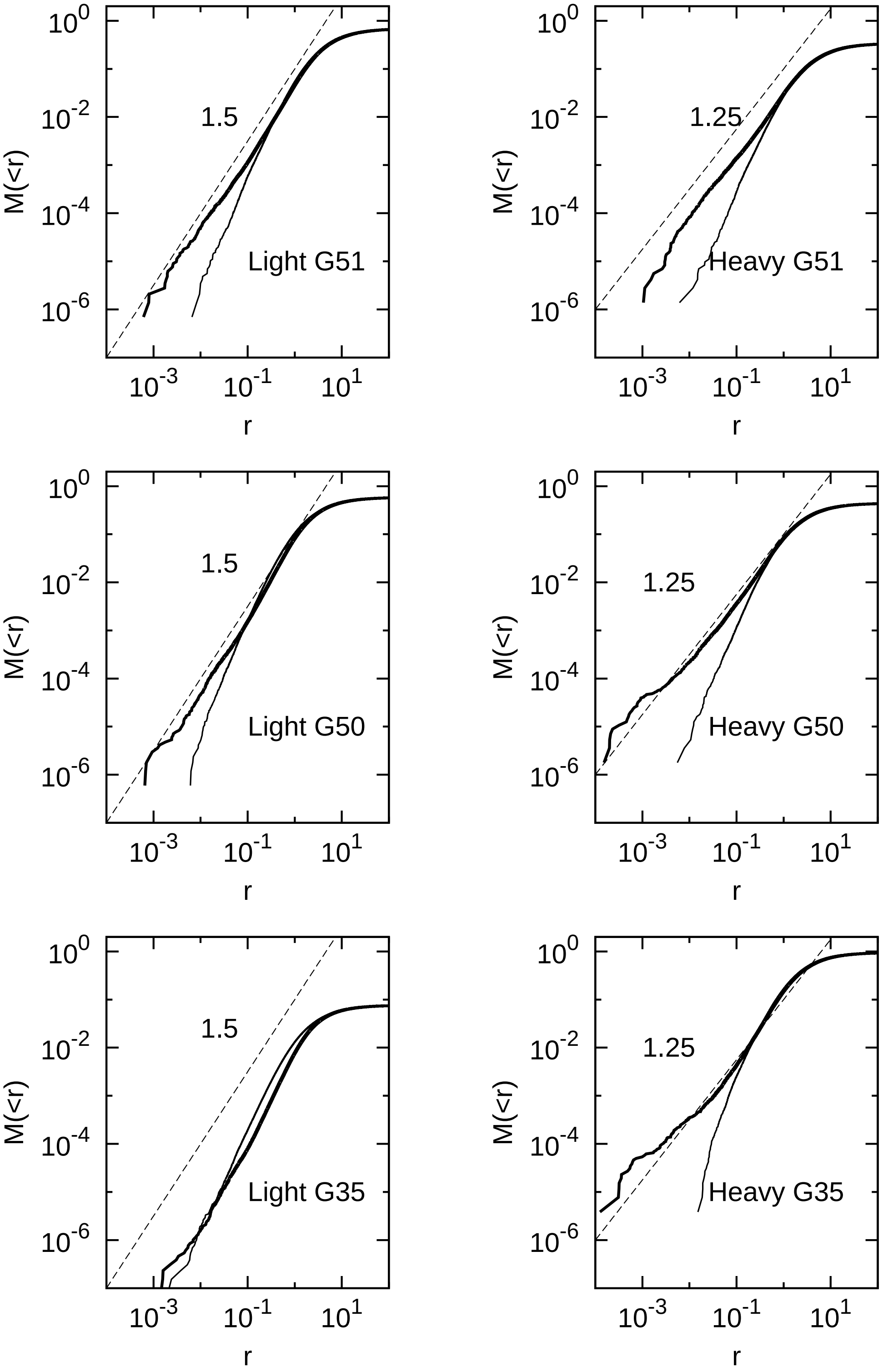

Figure 4.16: The cumulative mass distribution $M(< r)$ at late times was extracted from the N-body runs (see description in the text). The light curve corresponds to the initial time; the heavy curve to the distribution at late times. The inner slopes $M(< r) \propto r^{3-\gamma}$ inside the stellar cusp, are consistent with those predicted from the Fokker-Planck equation for radii $r \leq 0.1 r_h$. See Tables 4 and 5 for the numerical values of the inner slopes that were computed from this plot's data.

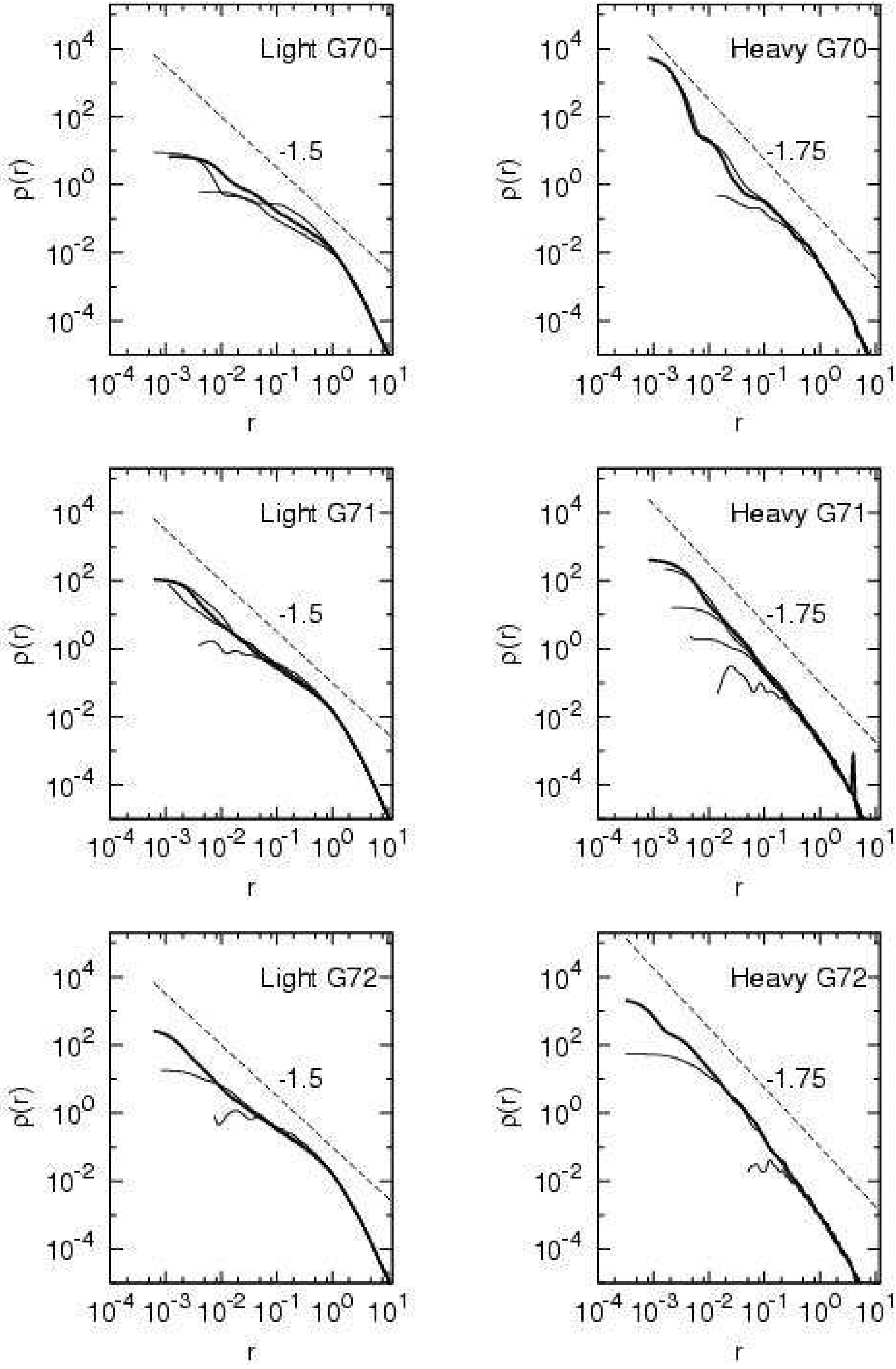

Figure 4.17: Evolution of the mass density profile around the black hole. Left panels: N-body $\rho(r)$ for the light component was estimated from the particle's positions. Right panels: N-body $\rho(r)$ for the heavy component. The initial slope was $\gamma_l = \gamma_h = 1/2$; it evolves, after $\approx 0.t_{rlx}$ to a value consistent with the values predicted from the Fokker-Planck equation for radii $r \leq 0.1 r_h$.

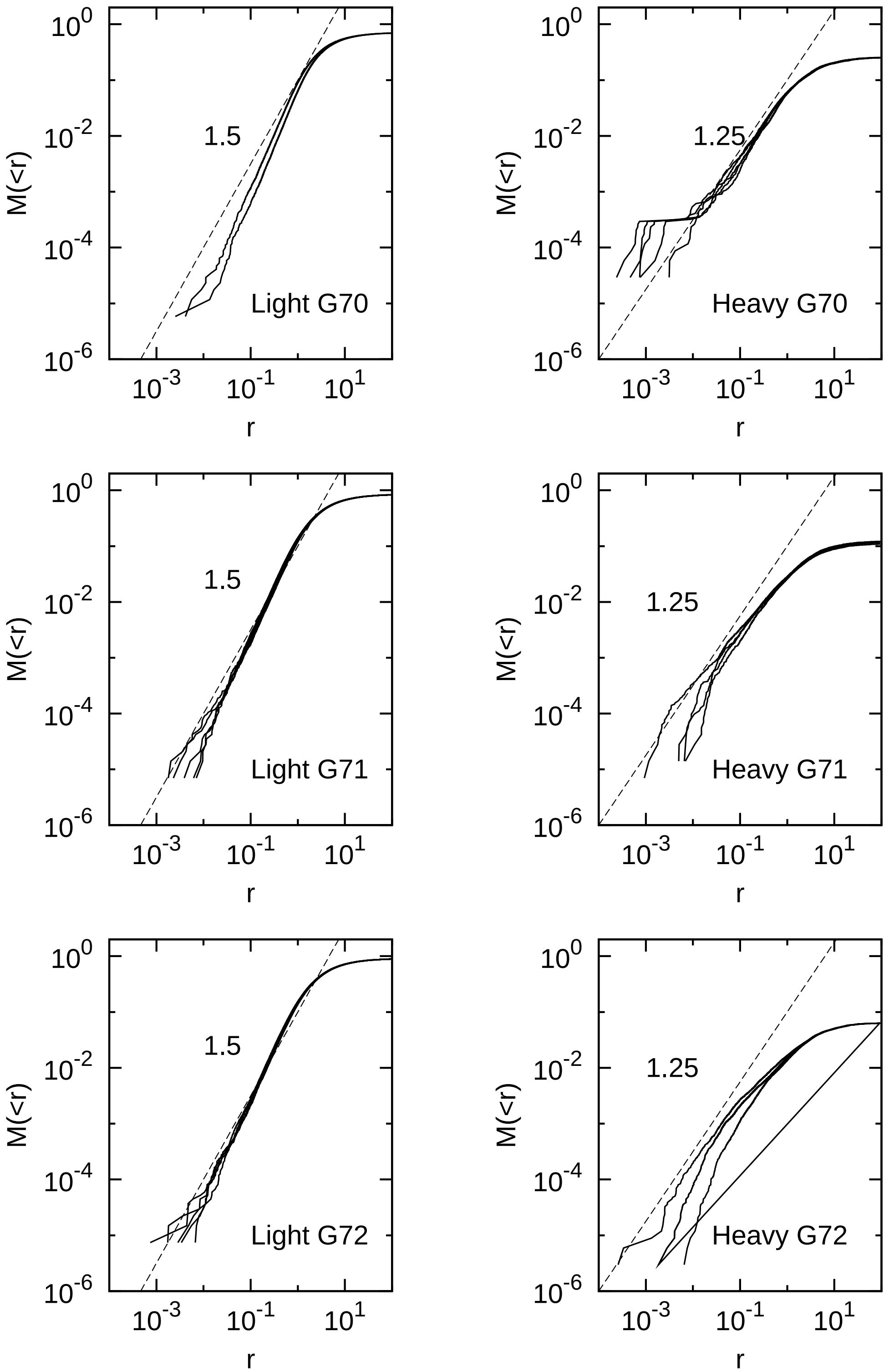

Figure 4.18: The time evolution of the cumulative mass distribution $M(< r)$. The inner slopes $M(< r) \propto r^{3-\gamma}$ inside the stellar cusp, are consistent with those predicted from the Fokker-Planck equation for radii $r \leq 0.1 r_h$.

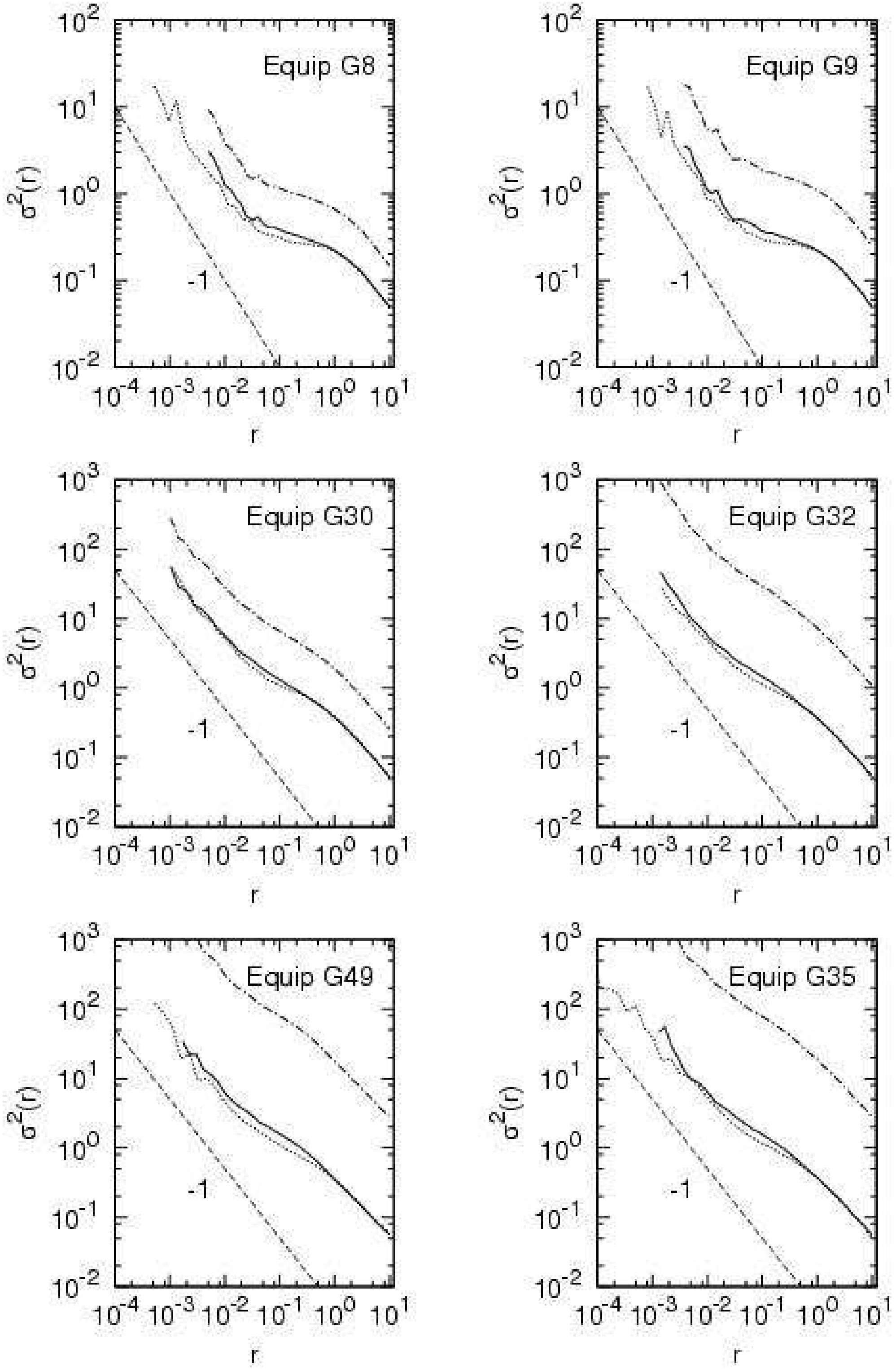

Figure 4.19: The $1D$-velocity dispersions $\sigma_l^2(r)$ and $\sigma_h^2(r)$, at late times, for the two components: full line (light component) and dotted line (heavy component). The velocity dispersions are almost mass-independent in all runs, and are related through $\sigma_l^2 \approx \frac{5+2p_h}{5+2p_l}\sigma_h^2$. The third curve (dashed-dotted line) represents the velocity dispersion the light stars would follow if they were in kinetic energy equipartition with the heavy stars, $\sigma_l^2 = \frac{m_h}{m_l}\sigma_h^2$. Note that the $\sigma^2 \propto 1/r$ is valid inside $\sim 0.1 r_h$, as expected.

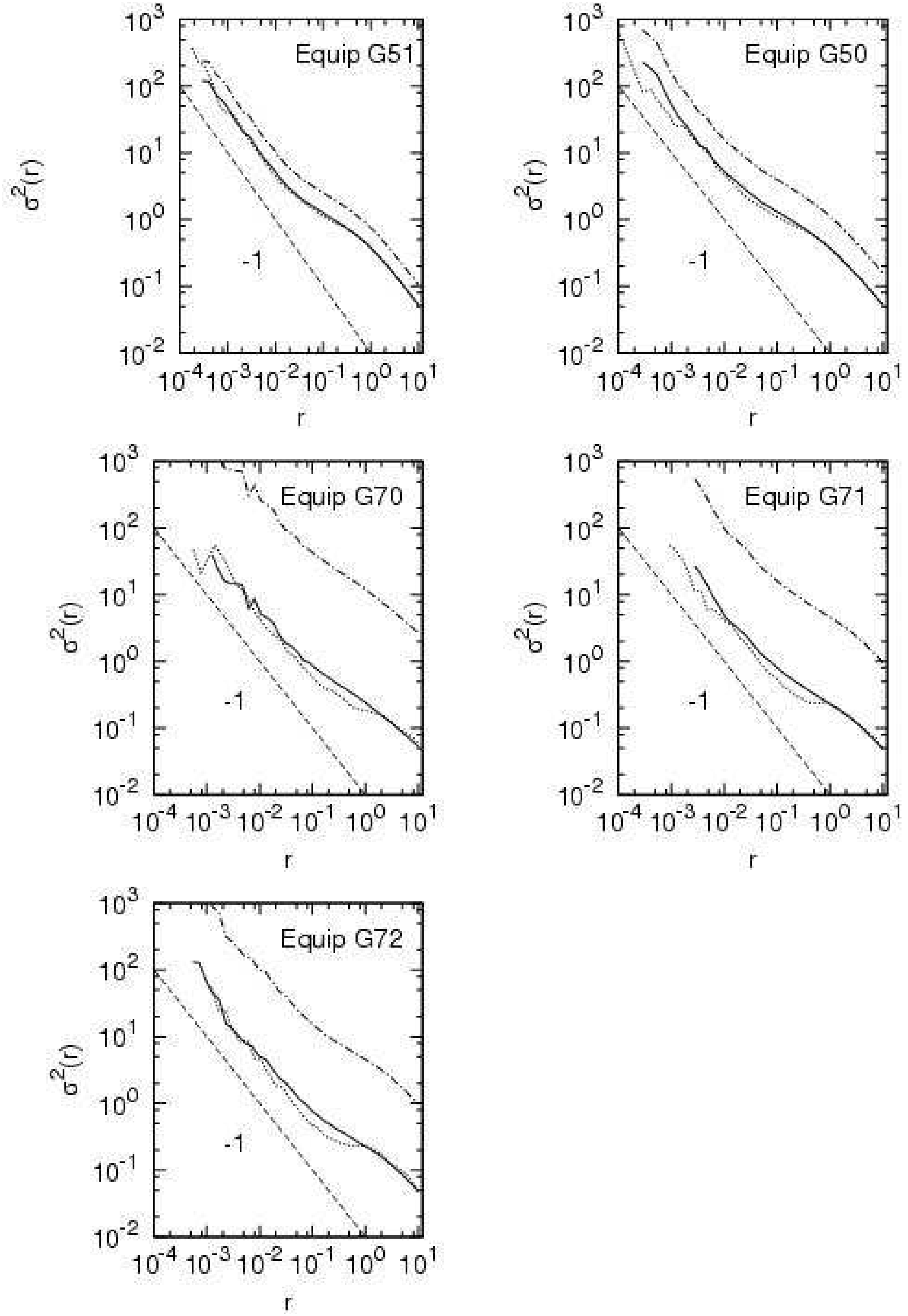

Figure 4.20: The $1D$-velocity dispersions $\sigma_l^2(r)$ and $\sigma_h^2(r)$, at late times, for the two components: full line (light component) and dotted line (heavy component). The velocity dispersions are almost mass-independent in all runs, and are related through $\sigma_l^2 \approx \frac{5+2p_h}{5+2p_l}\sigma_h^2$. The third curve (dashed-dotted line) represents the velocity dispersion the light stars would follow if they were in kinetic energy equipartition with the heavy stars, $\sigma_l^2 = \frac{m_h}{m_l}\sigma_h^2$. Note that the $\sigma^2 \propto 1/r$ is valid inside $\sim 0.1 r_h$, as expected.

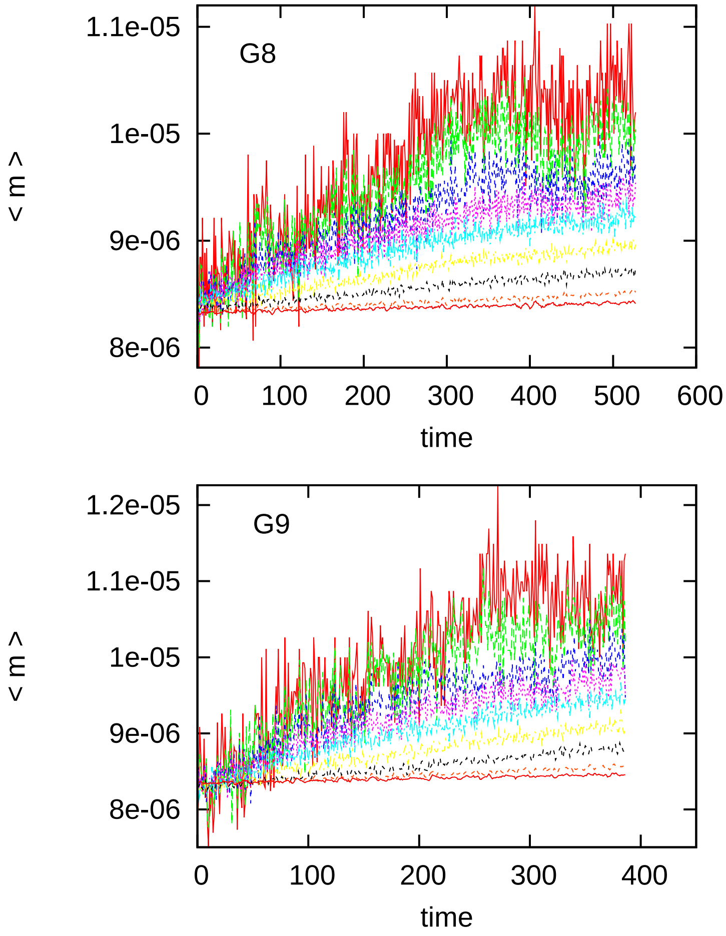

Figure 4.21: Time evolution of the average mass in Lagrangian shells showing the ongoing mass segregation in the stellar cusp around the black hole. The curves correspond to the Lagrangian shells with radial boundaries defined by the Lagrangian radii corresponding to the following fractions 0.001, 0.002, 0.002, 0.005, 0.01, 0.02, 0.06, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75 and 0.9 of the total stellar mass. As mass segregation proceeds, the average mass of the inner shells increase as they become dominated by the heavy stars. The strong, high-frequency oscillations of the average mass reflects the (locally) stochastic nature of the process.

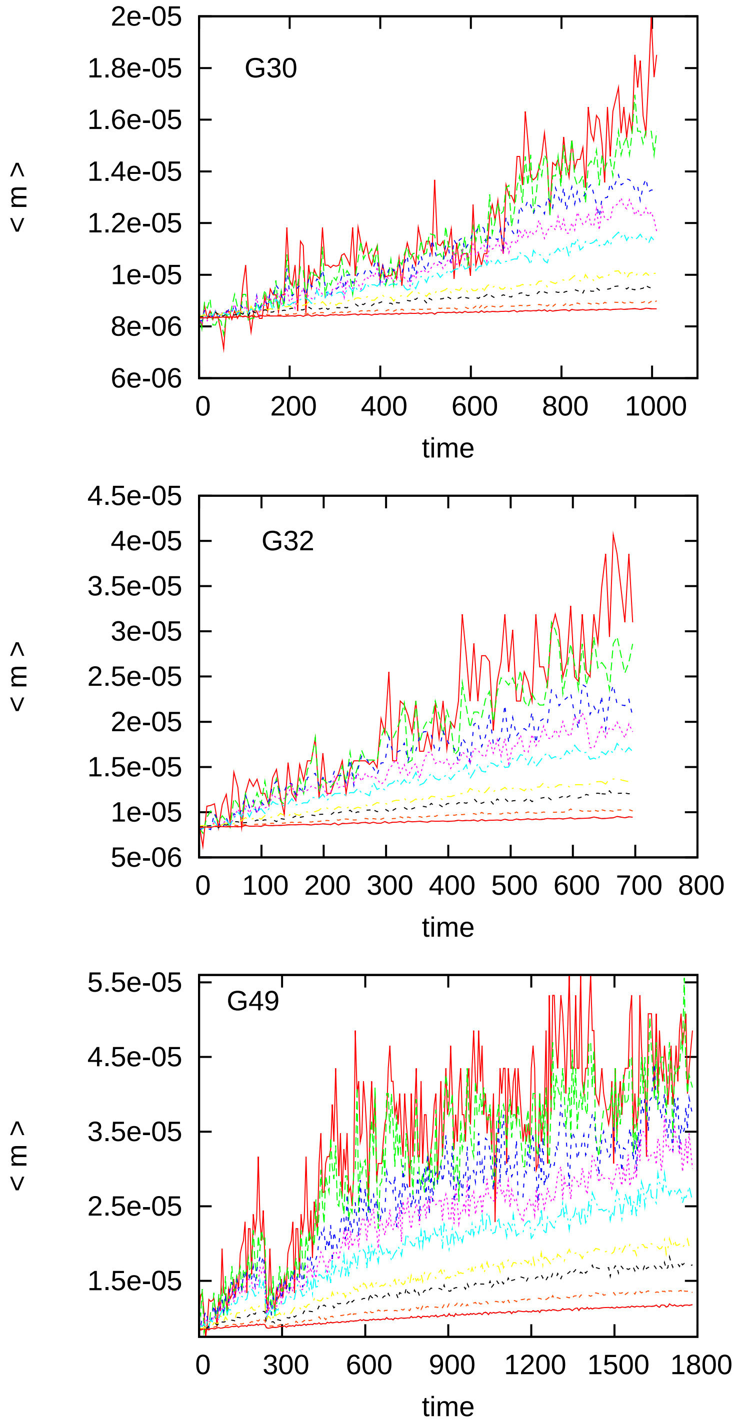

Figure 4.22: Time evolution of the average mass in Lagrangian shells showing the ongoing mass segregation in the stellar cusp around the black hole. The curves correspond to the Lagrangian shells with radial boundaries defined by the Lagrangian radii corresponding to the following fractions 0.001, 0.002, 0.002, 0.005, 0.01, 0.02, 0.06, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75 and 0.9 of the total stellar mass. As mass segregation proceeds, the average mass of the inner shells increase as they become dominated by the heavy stars. The strong, high-frequency oscillations of the average mass reflects the (locally) stochastic nature of the process.

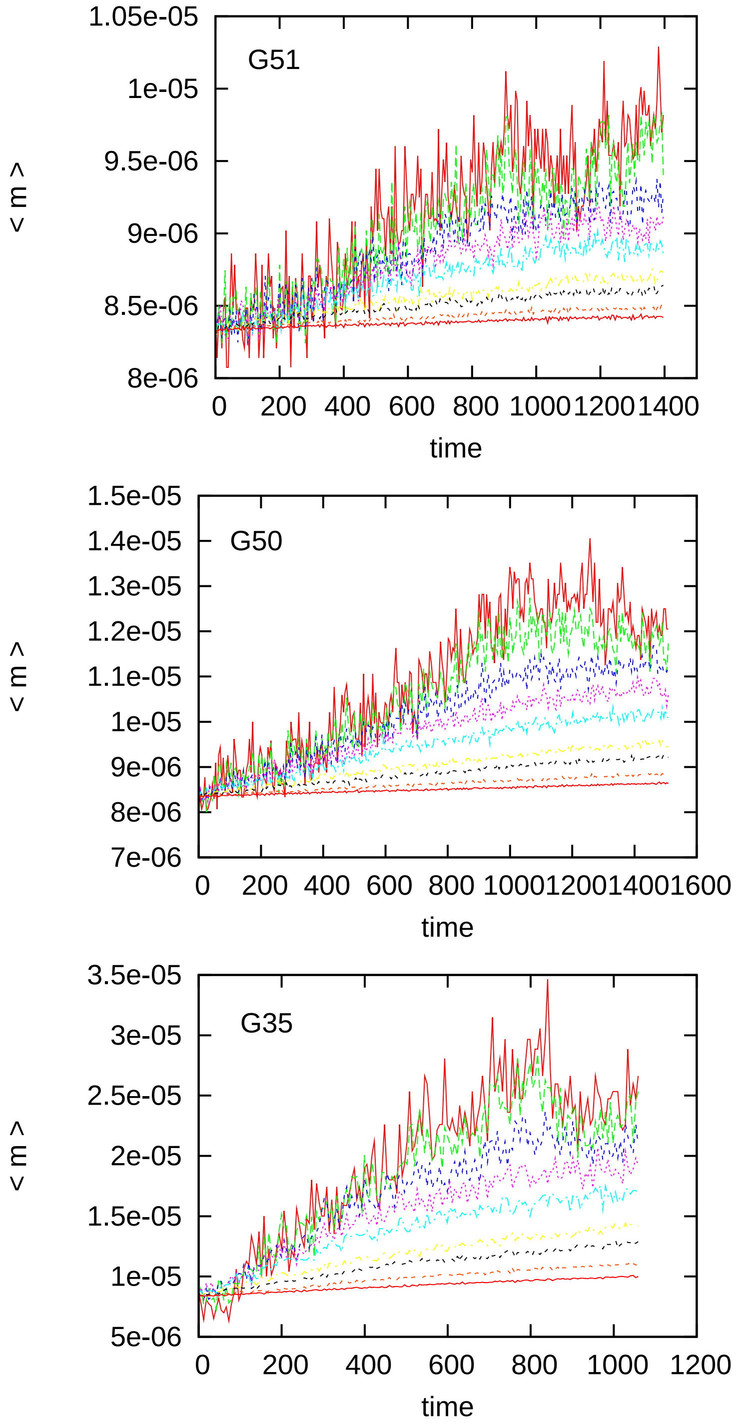

Figure 4.23: Time evolution of the average mass in Lagrangian shells showing the ongoing mass segregation in the stellar cusp around the black hole. The curves correspond to the Lagrangian shells with radial boundaries defined by the Lagrangian radii corresponding to the following fractions 0.001, 0.002, 0.002, 0.005, 0.01, 0.02, 0.06, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75 and 0.9 of the total stellar mass. As mass segregation proceeds, the average mass of the inner shells increase as they become dominated by the heavy stars. The strong, high-frequency oscillations of the average mass reflects the (locally) stochastic nature of the process.

4.5 Multi-mass case

Galaxies are built of stars from a range of masses. The initial mass function (IMF) — the mass distribution with which a stellar population is born in a star formation event — is thought to be reasonably well approximated by a power law, $n(m) \propto m^{-\alpha}$ with $\alpha \approx 2.3$ [Salpeter, 1955, Miller and Scalo, 1979]. This is the so-called Salpeter IMF. Other functional forms such as those adopted by the Scalo and Kroupa IMFs are also prototypical examples that combine different power law slopes for specific stellar mass intervals [Binney and Merrifield, 1998]. However, as stellar evolution proceeds, stars loose mass through winds and sudden, violent events such as supernovae explosions where they loose an appreciable percentage of their original mass and turn into dark, compact remnants.

The mass functions of galactic nuclei are very difficult to constrain observationally. These difficulties are severe even in the case of the closest galactic nucleus at the Milky Way center. There is strong, spatially variable extinction by interstellar dust that blocks most of the radiation in the optical-UV wavelenth range; it is possible, but very difficult, to follow and identify individual stars as they orbit Sgr-A* since the field is heavily crowded [Alexander, 2005].

Surprisingly few $N-$body studies have been done with multi-mass stellar component around a massive black hole with a few notable exceptions. [Baumgardt et al., 2004b, Freitag et al., 2006]. Baumgardt et al. 2004b simulated the evolution of globular clusters with one IMBH in the center, including several additional physical effects such as stellar mass loss and tidal disruptions of stars by the IMBH. Freitag et al. 2006 also did numerical experiments with multi-mass stellar distributions, but mostly employed an Hénon-type Monte Carlo code which shares the assumptions of the Fokker-Planck approach; in this paper, the authors report also one N-body simulation for the case of a two-component cluster.

Baumgardt et al. 2004b did their calculations with initial conditions built from King models, typically with concentration parameter $W_0 = 7.0$, together with a Kroupa IMF having a lower mass limit of $0.1 M_\odot$ and an upper mass limit of $30 M_\odot$, in one series of simulations, and of $100 M_\odot$ in another. They find that the final steady-state density profile has inner slope of $\gamma \approx 1.5$ in all cases, when all masses are included in the density estimation. The stars with masses below $1.2 M_\odot$ are divided into four groups defined by $[0, 0.2]$, $[0.4, 0.6]$, $[0.6, 0.8]$, and $[0.8, 12]$; they find that the inner slopes for these four groups are $\gamma \approx 0.76, 1.24, 1.4, 1.68$ respectively. So, their exponent γ decreases from the heavy masses to the light ones since mass segregation enhances the density of heavy stars in the center. This finding is in qualitative agreement with the theoretical expectations from mass segregation and also with the corresponding results of the present work. However, at a quantitative level, there is a marked discrepancy in that they claim the inner slopes for stars of average mass $< m >$ can be fitted, for their simulations, by the following empirical formula

$$\alpha(m) = 0.75 + \frac{m}{1.1}. \tag{4.16}$$

If this is the case, their results — at least if taken at face value — are in contradiction with the Bahcall-Wolf scaling relations which imply that the minimum asymptotic slope for any component cannot be smaller than $\gamma = 1.5$. For the heavier masses, on the contrary, they find values consistent with the classical $\gamma = 7/4$ cusp. We should admit, at the outset, that our results don't exclude completely any deviations from the Bahcall-Wolf predictions; however, with a simple inspection of our results both for two and multi-mass component runs, we can confidently say that there is not such a pronounced deviation as that hinted by Baumgardt et al 2004b.

In order to try to understand their results, we have made a Fokker-Planck calculation with a choice of parameters as close as possible to theirs. We adopted a Dehnen model as initial condition (the precise form of these should not matter after a relaxation time since by then the system should lost memory from its initial conditions), a black hole mass $M_\bullet = 0.01$, a power law mass function with a Salpeter exponent $\alpha = 2.3$ and computed the evolution of 20 mass bins with a range of mass of $0.1 - 5$.

We can see clearly that the slope of the cusp (in our multi-mass Fokker-Planck calculation) corresponding to the lightest component extends only out to $r \lesssim (0.01 - 0.02) r_h$, while that of the heavier component extends more outwards until $r \lesssim 0.1 r_h$. It is not possible to decide whether the reason for the lower slopes of the light components in the simulations of Baumgardt et al. 2004b was just due to the lack of resolution at sufficiently small radius, but it probably was. It seems that they haven't enough resolution in the center and, therefore, did not reach the radius where the asymptotic Bahcall-Wolf profiles are expected to be seen. In fact, their particle distributions (as inferred from the plots of the density profiles) extend only out to $r \lesssim 0.3 - 0.4 r_h$; the resolution thus attained only allows to measure any slope for a very small radial range (clearly less than a decade). We stress that our N-body runs have a higher resolution at the center since the stellar distributions extend inwards until almost $10^{-3} r_h$ — covering almost a three decade range in radius. On the other hand, in their Figure 7, all components seems to extend their inner slopes out to the same radius in contradiction with our Fokker-Planck and, actually, also with our N-body results.

The two-component simulation reported by Freitag et al. 2006 adopted a Dehnen model with an initial inner slope $\gamma = 1.0$, black hole mass $M_\bullet = 0.1$, a mass ratio $R = 10$, a heavy number fraction $f = 0.05$, and a total number of particles $N = 64K$ (which is only half of our typical value). They find that the light stars only reach the $\gamma = 3/2$ profile for $r \lesssim (3 - 5) \times 10^{-3} r_h$, which is consistent with our own results (compare, for example, with run G71 which is the run with closer parameters to theirs). Note that even in this case our run shows a $\gamma = 3/2$ slope over a wider radial range, presumably because our larger (by a factor of 2) number of particles results in a better resolution at small radii.

We didn't intend to model a specific galactic nucleus, as our main goal is to study the robustness of the Bahcall-Wolf cusp with respect to several plausible mass functions. With this goal in mind, a series of simulations was run, the mass distributions given according to a power law with exponents: $\alpha = 0.3, 1.3, 2.3$. Then, focusing on the latter value $\alpha = 2.3$, the range of the mass spectrum was varied taking the values $20, 50, 50, 100, 600$ in different runs in order to search for any effect on the shape of the Bahcall-Wolf scaling relations. A Kroupa mass function was also tried in a few simulations where a mass spectrum of $0.01 - 1.0$ or $0.2 - 120$ were used in order to probe differently the slopes of this mass distribution. The results we have found in these multi-mass runs confirm and reinforce those already observed in the two-component experiments. In fact, deviations from the Bahcall-Wolf scaling relations — if they happen at all — are small and noticeable only for the cases with a wider interval of mass distribution. In fact, just as it was discussed for the two-component case, it is not clear if the Bahcall-Wolf predictions are violated in any of the cases studied since the slopes of lighter objects seem to converge to -1.5 at the smallest radii. The plots for the cumulative mass distribution indicate very clearly that the light objects are settling down into a $\gamma = 3/2$ cusp at the smallest radii that we are able to resolve with this particle number. This, in contrast with the results from Baumgardt et al. 2004b, seem to be an indication that our simulations were able to resolve deeper into the potential well than theirs, so we can claim with confidence that the Bahcall-Wolf scaling relations are validated by our N-body experiments. It is, in fact, the heavier component — especially for the runs with a larger range of masses — that develops an overdensity at the innermost radii. This is complete agreement with the double-mass runs, and shows that those overdensities were not an artifact resulting from the artificial nature of the mass distribution of the two-component runs. Another discrepance with the Baumgardt et al. is that our density profiles for the total stellar cluster (including all the mass range) have asymptotic inner slopes which are consistent with the 7/4 cusp (and not with 3/2). This is the result of mass segregation: the heavier stars migrate to the center and come to dominate the mass in the inner regions.

A few exploratory multi-mass run tests were done for which a tidal disruption radius was defined at $r_t = 10^{-7}$, in our N-body units. The idea was to check whether the cusp slopes for the inner density profile would be lowered due to the capture of stars by the MBH. They were evolved over the same time scale as these runs, but neither qualitative nor quantitive differences were found. The number of stars captured was never found to be above ten.

In the multi-mass runs, the light and heavy stars were defined as follow: the total mass range was divided into four intervals (equally spaced in mass); the light stars correspnd to the first interval, the heavy to the fourth interval.

We conclude that N-body simulations reproduce the scaling relations predicted by the Fokker-Planck calculations. We monitor the growth of Bahcall-Wolf cusp at the same time as the system undergoes very strong mass segregation. The heavy stars come to dominate the innermost regions around the black hole, while the stars with lower mass distribute themselves with a less steep density profile. The heavier stars tend, in many cases, to develop an overdensity at the innermost regions: this could have interesting consequences for the event rates of extreme mass ratio inspirals (EMRIs) which are one of the main observational targets for the upcoming LISA.

Once again, these conclusions should be taken with the proviso that only with a substantial increase of particle number will it be possible to provide a definitive answer — in the form of cusp with an inner slope $\gamma = 3/2$ for the light objects, and $\gamma = 7/4$ for the heavy objects, spanning several decades in radius. Galactic nuclei have a very large number of stars $N \sim 10^7 - 10^8$ so we are still several orders of magnitude below.

We plan to follow up this question in the near future with the help of the ARI GRAPE cluster once the fully parallelized, GRAPE-adapted NBODY6++ code is ready for production runs. In that case, it will be possible to reach $(1 - 3) \times 10^6$ particles — an order of magnitude above the total number employed in this work.

G	γ	$M_\bullet$	R	α	r_{cusp}	$t_{\rm rlx}$	$T_{max}(T_{max}/t_{\rm rlx})$
$G12$	1/2	0.05	20	0.3	0.06	3912	500 (0.128)
$G16$	1/2	0.05	20	1.3	0.06	3912	880 (0.225)
$G20$	1/2	0.05	20	2.3	0.06	3912	820 (0.210)
$G44$	1/2	0.05	100	2.3	0.06	3912	365 (0.093)
$G46$	1/2	0.05	600	2.3	0.06	3912	815 (0.208)
$G52$	1/2	0.05	600	Kroupa	0.06	3912	810 (0.207)

Table 4.5: Multi-mass runs. 1st column: slope of the Dehnen's model inner cusp at $t = 0$; 2nd column: BH mass; 3rd: range of stellar mass; 4th: mass function power-law index α, $n(m) \propto m^{-\alpha}$ (Kroupa mass function for G52); 5th: $0.1 r_h$, where r_h is the influence radius; 6th: relaxation time of the (equivalent) single mass cluster as measured at r_h; 7th: maximum integration time.

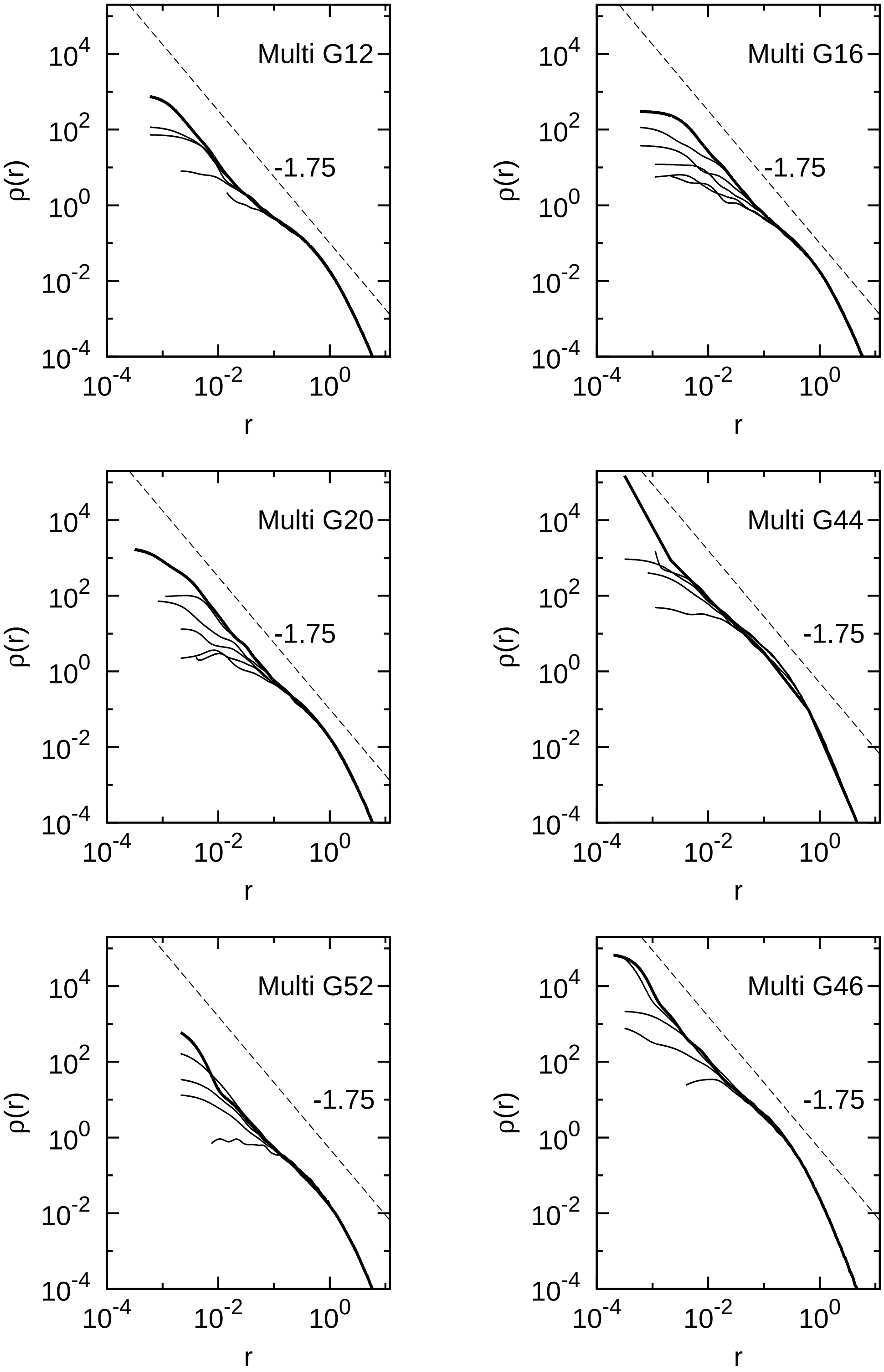

Figure 4.24: Evolution of the total mass density profile, including all stars covering the whole mass interval, around the black hole. An asymptotic, steady-state solution with a cusp slope $\gamma \approx 7/4$ is reached over a fraction of one relaxation time measured at r_h for a single mass cluster.

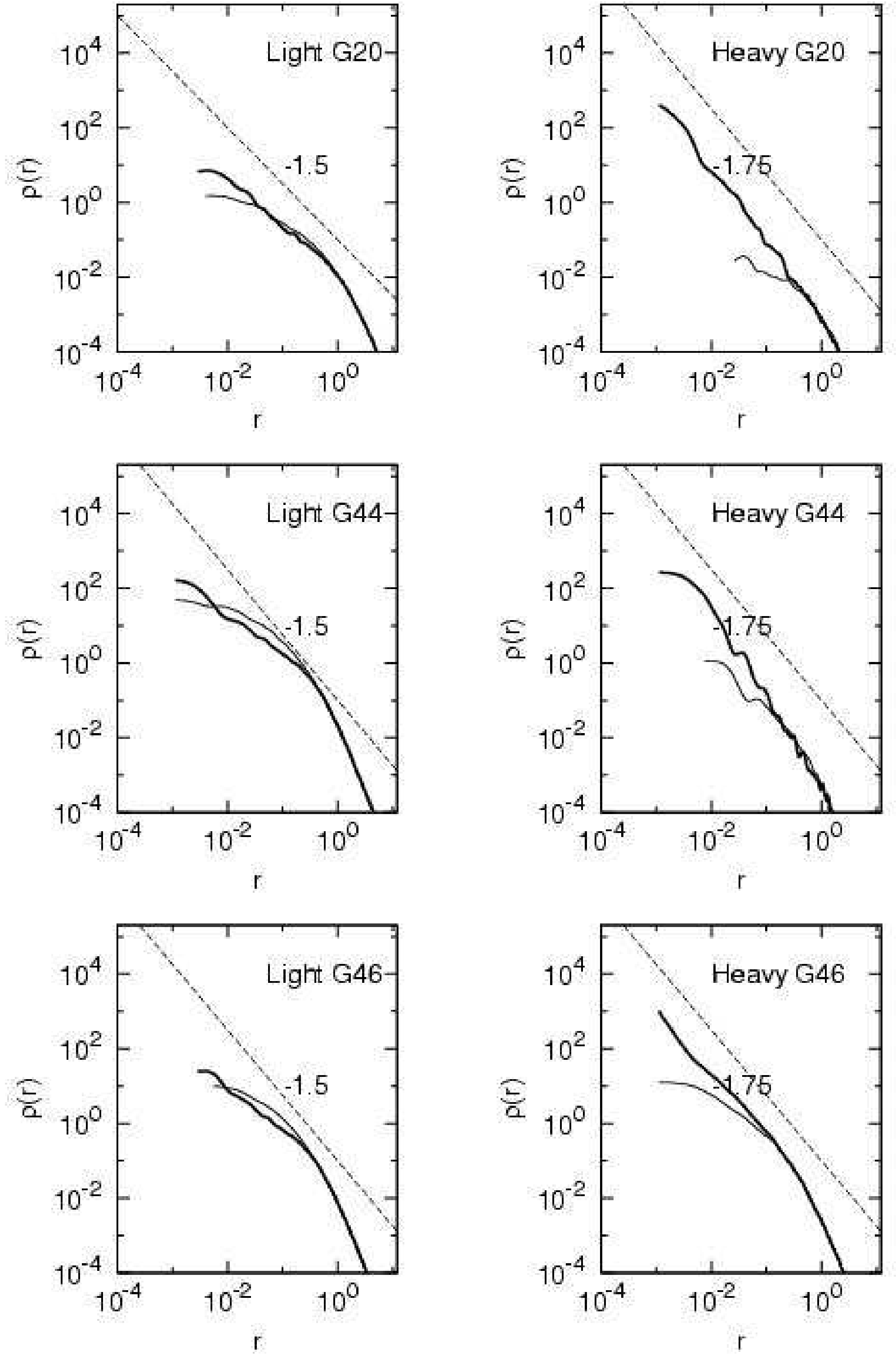

Figure 4.25: Mass density profiles $\rho(r)$. Left: lighter stars in the population. Right: heavier stars in the population (see description in the text). Thin curves for $t = 0$; thick curves for late times.

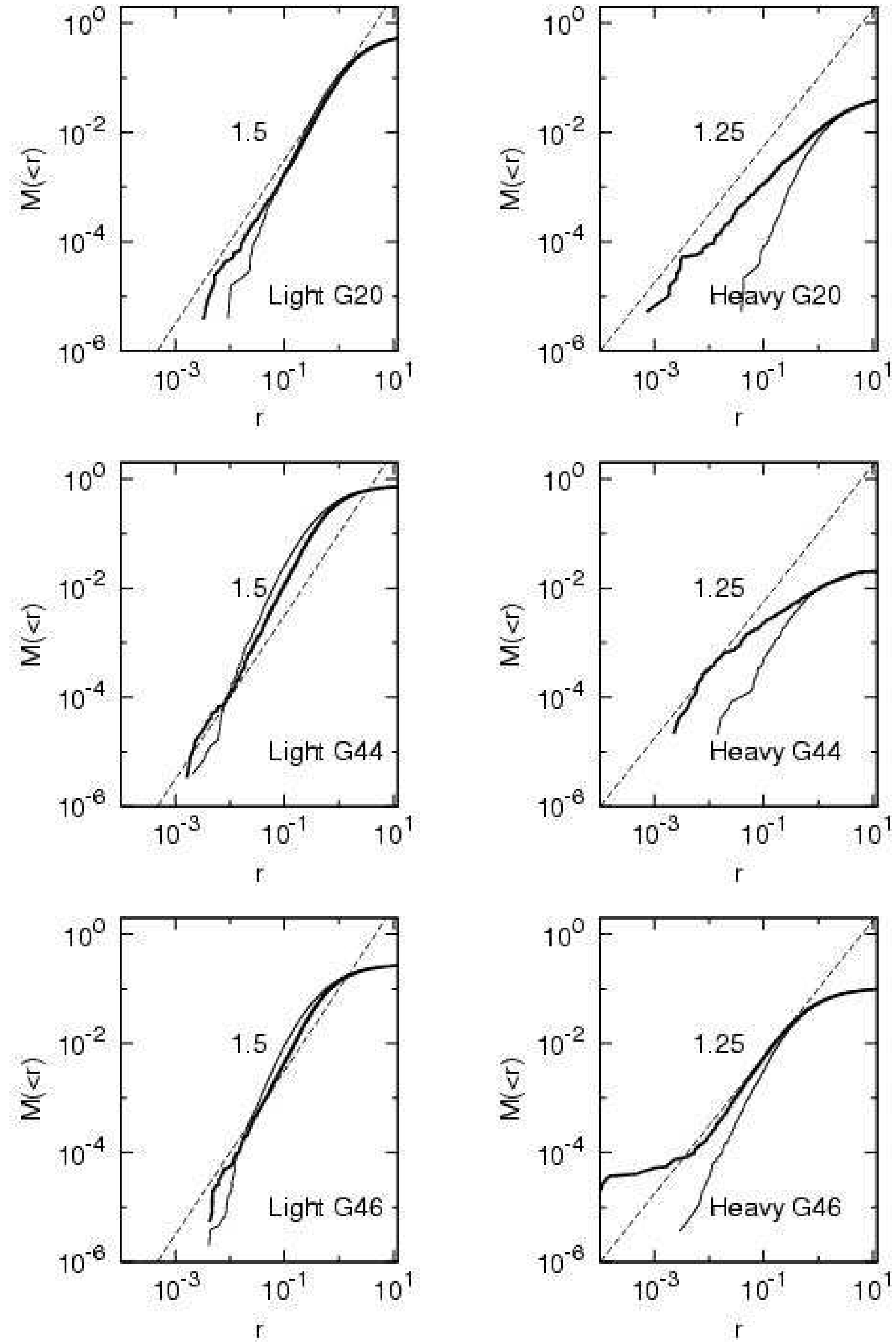

Figure 4.26: Cumulative mass distributions $M(< r)$. Left: lighter stars in the population. Right: heavier stars in the population (see description in the text). Thin curves for $t = 0$; thick curves for late times.

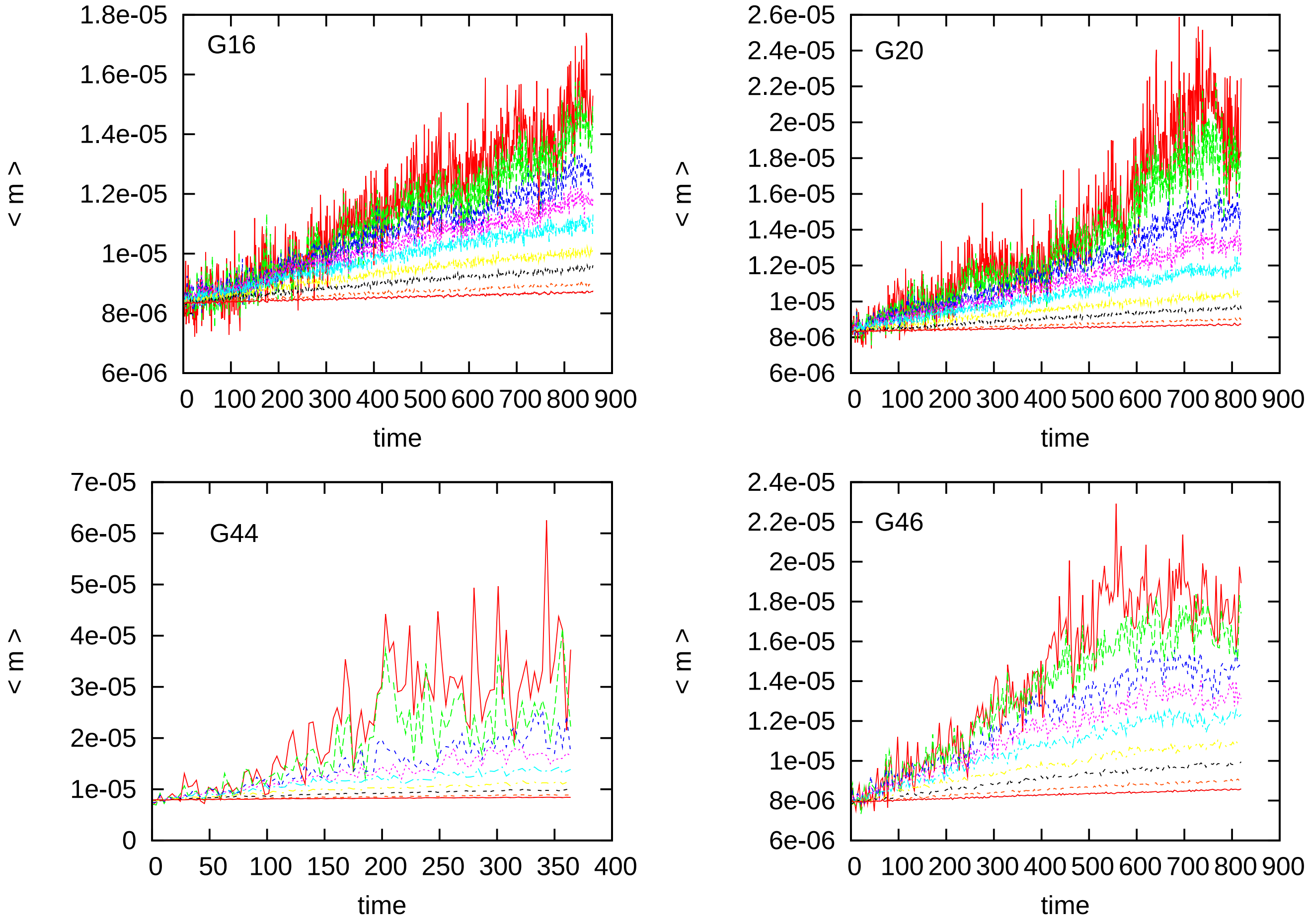

Figure 4.27: Time evolution of the average mass in Lagrangian shells showing the ongoing mass segregation in the stellar cusp around the black hole. The curves correspond to the Lagrangian shells with radial boundaries defined by the Lagrangian radii corresponding to the following fractions 0.001, 0.002, 0.002, 0.005, 0.01, 0.02, 0.06, 0.1, 0.2, 0.3, 0.4, 0.5, 0.75 and 0.9 of the total stellar mass. As mass segregation proceeds, the average mass of the inner shells increase as they become dominated by the heavy stars. The strong, high-frequency oscillations of the average mass reflects the (locally) stochastic nature of the process.

Chapter 5

Gravitational waves from Post-Newtonian binary systems

5.1 Propagation of gravitational waves

In the absence of gravity, the spacetime is flat and Minkowski; in the presence of a weak gravitational field, it will be a nearly Minkowski spacetime. Consider a manifold on which coordinates exist such that the metric can be written everywhere in the form

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \tag{5.1}$$

where $|h_{\alpha\beta}| \ll 1$ and $\eta_{\alpha\beta} = \mathrm{diag}(-1, 1, 1, 1)$ is the Minkowski metric of special relativity. $h_{\alpha\beta}$ represents the metric perturbation; it contains the gravitational waves (GWs) but also non-radiative degrees of freedom. The perturbation metric behaves as a tensor with respect to Lorentz transformations, but not under more general coordinate transformations. The gauge transformations result from very small changes of coordinates: $x^{\alpha'} = x^\alpha + \xi^\alpha(x^\beta), |\xi^\alpha_{,\beta}| \ll 1$. Both these transformations preserve the splitting of the metric in (5.1). Gauge transformations should be viewed as defining equivalence classes among tensors: any two tensors related by such a transformation should produce equivalent physical effects. The perturbation metric $h_{\alpha\beta}$ allows to write down all basic quantities that describe the linearized theory: the Christoffel symbols, the Riemann and Ricci tensors, and so on. These can be found in the general textbooks and review articles [Landau and Lifshitz, 1972, Schutz, 1985, Flannagan and Hughes, 2006], so we see no point in reproducing them here.

It is costumary to choose the harmonic (or Lorentz) gauge in radiation problems; in this case, it is defined by a divergenceless trace-reverse metric perturbation:

$$\bar{h}^{\mu\nu}_{,\nu} = 0, \tag{5.2}$$

where $\bar{h}_{\alpha\beta} = h_{\alpha\beta} - 1/2\eta_{\alpha\beta}h$ and $\bar{h} = \bar{h}^{\alpha}_{\alpha} = -h$. The linearized Einstein tensor $G_{\alpha\beta}$, in the harmonic gauge, is given by

$$G_{\alpha\beta} = -\frac{1}{2}\Box\bar{h}_{\alpha\beta}, \tag{5.3}$$

i.e. it reduces to the $4D$-wave operator on the trace-reversed metric perturbation tensor. The linearized Einstein's field equations are a set of linear wave equations:

$$\Box\bar{h}_{\alpha\beta} = -\frac{16\pi G}{c^4}T_{\alpha\beta}, \tag{5.4}$$

where $T_{\alpha\beta}$ is the stress-energy tensor for matter, and $\Box = -\partial_t^2 + \nabla^2$ is the d'Alembertian operator; in vacuum they become homogeneous wave equations

$$\Box\bar{h}_{\alpha\beta} = 0. \tag{5.5}$$

A class of solutions to the weak field equations is then given by the Fourier superposition of plane wave solutions

$$\begin{aligned} \bar{h}_{\alpha\beta}(t,\mathbf{x}) &= \mathrm{Re}\int d^3k\ A_{\alpha\beta}(\mathbf{k})e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} = & (5.6) \\ &= \mathrm{Re}\int d^3k\ A_{\alpha\beta}(\mathbf{k})e^{ik_\alpha x^\alpha}. & (5.7) \end{aligned}$$

Note that $\bar{h}_{\alpha\beta}$ is constant on a hypersurface on which $k_\alpha x^\alpha =$ cte.; $k^2 = \omega^2 - c^2|\mathbf{k}|^2 = 0$ is the dispersion relation for the waves. From the Lorentz gauge condition, $k^\alpha A_{\alpha\beta} = 0$ obtains and therefore $A_{\alpha\beta}$ is orthogonal to the 4-vector k. These solutions are precisely the (linear) gravitational waves.

It is costumary to assume some further conditions: (i) the spacetime is globally vacuum, for which the matter stress-energy tensor $T_{\alpha\beta} = 0$ everywhere; (ii) specialize to a sub-gauge of the harmonic gauge: the transverse-traceless(TT) gauge in order to make the metric perturbation to be purely spatial and traceless:

$$\begin{aligned} h_{tt} = h_{\alpha t} &= 0 \\ h = h^{\alpha}_{\alpha} &= 0. \end{aligned} \tag{5.8}$$

In the TT-gauge, the metric perturbation becomes very simple

$$\bar{h}^{TT}_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_{xx} & h_{xy} & 0 \\ 0 & h_{xy} & -h_{xx} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.9}$$

Summarizing: this very simplified form for the metric perturbation was obtained after imposing the following restrictions: (i) $A^{\alpha}_{\alpha} = 0$ (one restriction); (ii) $A_{\alpha\beta}U^{\beta} = 0$ where U^{β} is some 4-velocity (3 restrictions for 4 equations, but only three are independent); (iii) $k^{\alpha}A_{\alpha\beta} = 0$, the harmonic gauge condition (4 restrictions). All gauge freedom is now used and the two remaining components of $A_{\alpha\beta}$ are gauge-invariant quantities. By the symmetry of the $G_{\alpha\beta}$ and $T_{\alpha\beta}$, there are a priori ten independent partial differential equations; imposing the harmonic gauge conditions (iii), we are left with six; (ii) reduces to three independent equations; and, finally, the traceless condition (i) leaves us in the end with only two independent equations.

The equivalence principle (EP) states that all bodies accelerate in the same manner under a given gravitational field. As a result, a free-falling body (i.e. moving under the influence no force other than gravity) doesn't depend on anything else than its initial position and velocity — it is independent of its internal structure.

In order for a detector to measure a gravitational wave, its state has to be somehow altered by the passing wave. For instance, were the wave to be perfectly uniform, its effect on the detector would be indistinguishable from a uniform acceleration of all its parts and, according to the EP, undetectable by its own measurements. In order to detect the GW, the apparatus would need to experience the non-uniformities of the gravitational field across its own *cross section*. Such non-uniformities are tidal effects as they stretch the spacetime geometry — loosely speaking, GWs are travelling tidal forces.

Conceptually, the simplest detector of GWs would measure the time evolution of the proper distance between two bodies. As it is a tidal effect, the bigger the separation, the larger the oscillation amplitude imprinted by the passing wave on the proper distance between them (provided, of course, that they are separated by less than a wavelength). This can be seen as follows: bodies in free-fall follow the geodesics of the spacetime metric; the geodesic equation is given by:

$$\frac{d^2x^{\alpha}}{d\tau^2} + \Gamma^{\alpha}_{\beta\gamma}\frac{dx^{\beta}}{d\tau}\frac{dx^{\gamma}}{d\tau} = 0, \tag{5.10}$$

where τ is the proper time measured by the observer travelling along the geodesic. In linearized theory, $v \ll c$ and the geodesic equation reduces to

$$\frac{d^2x^{\alpha}}{d\tau^2} + \Gamma^{\alpha}_{00} = 0. \tag{5.11}$$

However, $\Gamma^{\alpha}_{00} = 1/2(2h^{TT}_{j0,0} - h^{TT}_{00,j}) = 0$, since $h^{TT}_{\alpha 0} = 0$ (in the harmonic gauge, the metric perturbation is purely spatial as we've seen above). As a result, $d^2x^{\alpha}/d\tau^2 = 0$ obtains. In the TT-gauge, the coordinate location of a slowly moving free-falling particle is unnafected by the passage of the GWs — the coordinates oscillate in synchrony with the waves. Here enters the tidal nature of the GWs: the proper separation between two free-falling bodies oscillates under the influence of the GW, whilst their coordinates separation is kept constant.

The proper distance $L(t)$ between the two bodies at $z = 0$, separated along the X-axis by L_0 at the initial time, is affected by the passage of the wave according to

$$\begin{aligned} L(t) &= \int_0^{L_0} dx\, \sqrt{g_{xx}} = \int_0^{L_0} dx\, \sqrt{1 + h^{TT}_{xx}(t, z=0)} \approx \\ &\approx \int_0^{L_0} dx\, \left(1 + \frac{1}{2} h^{TT}_{xx}(t, z=0)\right) = L_0 \left(1 + \frac{1}{2} h^{TT}_{xx}(t, z=0)\right). \end{aligned} \tag{5.12}$$

The fractional length change is thus

$$\frac{\delta L}{L} \approx \frac{1}{2} h^{TT}_{xx}(t, z=0). \tag{5.13}$$

This fraction is purely scalar and therefore it is gauge invariant. The dimensionless magnitude h of a gravitational wave is called the wave *strain*. The phase shift $\delta\phi$ accumulated by a photon travelling back and forth of a laser interferometer arms depends on the time evolution of the proper separation of the mirrors attached to the mass probes, as these oscillate due to the passage of the GWs.

The concept of energy transported by a gravitational wave was for a long time a contentious issue. In GR, energy cannot be localized in a region much smaller than a wavelength since, according to the EP, test particle's (observers) won't feel the passage of a uniform field. However, if the detectors are sufficiently extended to experience the non-uniformity of the gravitational field they should respond in the manner just described. If they were placed inside a medium which offers some resistance to their motion, they will loose energy and heat the surrounding medium. Therefore, the GW would have dissipated some of its energy to the medium in the form of heat so we that GWs must transport energy. *

---

*In fact, more recently full numerical GR simulations have shown that GWs can accumulate into small regions of spacetime (analogously to shock waves in a gas or a fluid) and, due to their self-gravitating nature, if they are strong enough, may collapse locally and form black holes.

5.2 Generation of gravitational waves

GWs are generated by the matter source term on the right hand side of the linearized Einstein's equation

$$\Box \bar{h}_{\alpha\beta} = -\frac{16\pi G}{c^4} T_{\alpha\beta}. \tag{5.14}$$

In order to compute the leading order contribution to the spatial components (the only non-vanishing components in the harmonic gauge) of the metric perturbation, consider a non self-gravitating source whose internal motions are slow compared to the speed of light (*slow motion sources*) †

The non-homogeneous wave equation (5.14) can be solved as usual by means of Green's functions and the solution is

$$\bar{h}_{\alpha\beta}(t, \mathbf{x}) = \frac{4G}{c^4} \int d^3x' \, \frac{T_{\alpha\beta}(t - \frac{|\mathbf{x}-\mathbf{x}'|}{c}, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}, \tag{5.15}$$

where the quantity $t - |\mathbf{x} - \mathbf{x}'|/c$ is the retarded time — it accounts for the time lag associated with the propagation of the GW, at the speed of light, from the source at $\mathbf{x}'$ to the observer at $\mathbf{x}$. When the GW reaches the wave zone, it is safe to write $|\mathbf{x} - \mathbf{x}'| = r + \mathcal{O}(L/r)$. Hence, to leading order, the wave strain is given by

$$\bar{h}_{ij}(t, \mathbf{x}) = \frac{4G}{rc^4} \int d^3x' \; T^{ij}(t - \frac{r}{c}, \mathbf{x}'). \tag{5.16}$$

Note that, in the TT gauge, only the spatial part of the metric contributes to the radiative degrees of freedom. After some algebra, this can be shown to be [Flannagan and Hughes, 2006]

$$\bar{h}_{ij}(t, \mathbf{x}) = \frac{2G}{rc^4} \frac{\partial^2}{\partial t^2} \int d3x' \; \rho x'^i x'^j = \frac{2G}{c^4} \frac{d^2 I_{ij}(t-r)}{dt^2}, \tag{5.17}$$

where $I_{ij} = \int d^3x' \; \rho x'^i x'^j$ is the 2^{nd} moment of the mass distribution. It is still necessary to project out from the metric perturbation the non-TT parts; this is done by projecting the strain amplitudes onto the plane perpendicular to the direction of the wave propagation $\mathbf{n} = \mathbf{x}/r$ $(r = |\mathbf{x}|)$:

$$\bar{h}_{ij}(t, \mathbf{x}) = \frac{2G}{rc^4} \frac{d^2 \mathcal{I}_{ij}(t-r)}{dt^2} \left[P_{ik}(\mathbf{n}) P_{jl}(\mathbf{n}) - \frac{1}{2} P_{kl}(\mathbf{n}) P_{ij}(\mathbf{n}) \right], \tag{5.18}$$

---

† Travelling at those velocities does not permit to cover a distance of the order of the wavelength.

where $P_{ij} = \delta_{ij} - n_i n_j$ is the projection operator [Misner et al., 1972] and $\mathcal{I}_{ij} = I_{ij} - \frac{1}{3}\delta_{ij} I$ is the quadrupole moment tensor for the source's mass distribution.

As it is with other waves, the energy transported by GWs is proportional to the square of the time derivative of its (strain) amplitude. The energy flux falls off with r^{-2}, but when integrated over all directions over a sphere of radius r, the total luminosity $\mathcal{L}$ is independent of distance. The luminosity $\mathcal{L}$ radiated by a matter system is therefore given, to lowest order, by the Einstein's quadrupole formula:

$$\mathcal{L}(t) = \frac{G}{5c^5}\left(\frac{d^3\mathcal{I}_{ij}}{dt^3}\right)^2 . \tag{5.19}$$

For a two-point mass binary system,

$$\mathcal{I}_{ij}(t) = \sum_{\alpha=1}^{2} m_\alpha \left[r^i_\alpha(t) r^j_\alpha(t) - \frac{1}{3}\delta_{ij} r^2_\alpha(t) \right] . \tag{5.20}$$

In the case of a bound binary system, one can to lower order insert the solution of the Keplerian motion into (5.19) and (5.20) to find an explicit expression for the instantaneous luminosity $\mathcal{L}(t)$ value; orbit-averaging $\mathcal{L}(t)$ gives the averaged rate at which the binary system emmits gravitational radiation. The result is given by [Peters and Matthews, 1963]

$$\mathcal{L} = \left\langle \frac{dE}{dt} \right\rangle = \frac{32G^4}{5c^5}\frac{M_1^2 M_2^2 (M_1+M_2)}{a^5(1-e^2)^{7/2}}(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4), \tag{5.21}$$

where a is the semimajor axis and e the eccentricity of the Keplerian ellipse. Assuming a balance between the energy radiated to infinity in the form of GWs and the Keplerian energy $E = 1/2v^2 - GM/r$,

$$\frac{dE}{dt} = -\langle \mathcal{L} \rangle, \tag{5.22}$$

where μ_{red} is the binary's reduced mass.

The first moments of the mass distribution multipole expansion of a spinless source's stress-energy tensor $T^{\alpha\beta}$ are: (i) the source's monopole moment, which is its mass-energy M_0; (ii) its dipole moment M_1; (iii) the quadrupole moment M_2. They are given by

$$\begin{aligned} M_0 &= \int d^3x \ \rho = M \\ M_1 &= \int d^3x \ \rho x_i = M L_i \\ M_2 &= \int d^3x \ \rho x_i x_j = M L_{ij}. \end{aligned} \tag{5.23}$$

A simple physical argument gives some insight on why the quadrupole is the lowest moment able to contribute to the emission of GWs. In order to receive GWs from a source at a distance r, it is necessary that it produces a time-varying strain h; for dimensional reasons, the strain would be given by

$$\begin{aligned} h &\sim \frac{G}{c^2}\frac{M_0}{r} \\ h &\sim \frac{G}{c^3}\frac{d}{dt}\left(\frac{M_1}{r}\right) \\ h &\sim \frac{G}{c^4}\frac{d^2}{dt^2}\left(\frac{M_2}{r}\right). \end{aligned} \tag{5.24}$$

(5.25)

Conservation of mass-energy prevents the monopole term to produce a time-varying strain amplitude h. The time derivative of the dipole moment is equal to the total momentum of the source which is, for an isolated system, a conserved quantity. Again, no time-varying strain h. Finally, there is no conservation law to constrain right hand side of the last equation to be constant, so the quadrupole is indeed the lowest order moment that can give rise to GWs.

5.3 The Post-Newtonian Approximation

The Post-Newtonian (PN) approximation is based on the assumptions that the gravitational field is weak everywhere — inside the source as well: the so-called *weakly stressed sources* —, and of slow motion for both its internal and external degrees of freedom. For PN sources the small parameter ϵ, on which the expansion series is based, is defined according to

$$\epsilon = \mathrm{m}ax\left\{\left|\frac{T^{0i}}{T^{00}}\right|, \left|\frac{T^{ij}}{T^{00}}\right|^{1/2}, \left|\frac{\Phi}{c^2}\right|^{1/2}\right\}, \tag{5.26}$$

and it is always less than unity. The $T^{\alpha\beta}$ is the matter stress-energy tensor. Its domain of validity is the near zone of the source, i.e. a regoin surrounding the source with an extension much smaller than the wavelength of the gravitational waves. $|\Phi/c^2|^{1/2} \ll v/c$ for sources with negligible self-gravity, i.e. whose motion is determined by non-gravitational forces; in our case, we are interested in astrophysical binaries which are (weakly) self-gravitating in the sense that $|\Phi/c^2|^{1/2} \sim \mathcal{O}(v/c)$.

The main application of the PN equations of motion to stellar dynamics occurs in the study of the complete inspiral of compact binary systems (with components such as neutron stars, stellar-mass, intermediate-mass and massive black holes) which reach orbital velocities of order $v \sim (0.3 - 0.5)c$ as they approach the last moments before relativistic coalescence.

This fact raises, of course, concerns about the convergence and accuracy of the PN perturbation series which are, in fact, still far from having been completely solved [Blanchet, 2006]. There are not many results yet from the direct integration of the full Einstein's equations for the binary problem ‡, but preliminary results show that the PN approximation is able to locate ICO of two black holes of comparable mass ($M_{\bullet 1} \approx M_{\bullet 2}$) if the full expansion up to 3PN order is employed (see [Blanchet, 2006], Figure 1 and references therein).

The PN equations of motion for the problem of two compact bodies are currently known up to the 3.5PN order [Blanchet, 2006, Blanchet and Iyer, 2003, Memmesheimer et al., 2004]. Up to this order, there is a clear separation between the conservative part of the dynamics (the even terms $(v/c)^{2n}$ of the PN expansion) from the part associated to the radiation damping (the odd terms of the same expansion, the lower being the 2.5PN term $\propto (v/c)^5$) that result from the emission of gravitational waves by the source.

The direct Post-Newtonian method to derive the conservative part of the equations of motion starts from the 3PN metric of an extended source. In the case of compact bodies, the point particle limit has to be taken, and the matter tensor is therefore composed of delta functions. As a result, it is necessary to employ the Hadamard regularization method to solve the problem of the self-field divergences that result from taking the (singular) limit of a point particle from the original field theory. We do not dwell with these technical details in this work, and take them as given from the theoretical work of the relativity community.

The PN equations of motion — although derived from General Relativity — should be used and interpreted in a Newtonian fashion: the particle's trajectories are to be viewed as if evolving in the absolute space of Newtonian dynamics. Nonetheless, some properties entailed by their relativistic origin, must be fullfilled:

A1. The PN equations of motion are invariant — at least, in harmonic coordinates — under PN-expanded Lorentz transformations;

A2. They must reduce, in the limit when one of the masses goes to zero, to the geodesics of the Schwarschild metric;

A3. They are conservative whenever the odd (2.5 and 3.5 PN) terms due to radiation reaction are turned-off. There is an expression for a conserved *energy* E and *angular momemtum* J up to 3PN order.

Since General Relativity is a causal theory, the gravitational interaction cannot be instantaneously transmitted between two bodies located at some distance from each other. Furthermore, the speed of propagation of physical signals cannot be faster than the speed of light. It was already understood by Laplace, on general grounds, that a finite speed of propagation c for the gravitational interaction would lead to fundamental corrections to the

---

‡Consisting of a large number of of highly nonlinear partial differential equations, Einstein's equations for the two-body problem pose very delicate problems to the numerical analyst as well as a heavy burden to the computational resources.

Newtonian motion of the planets [Damour, 1982]. Laplace predicted that these corrections would be of order v/c (which is wrong), would add a damping term to the equations of motion (which is true, although of order $(v/c)^5$), and woud lead to a shrinkage and a circularization of the orbit, as well as a secular decrease of the orbital period. The secular acceleration of the Moon, already measured in Laplace's time, led him to conclude (wrongly) that the speed of propagation of gravity would have to be at least 7×10^6 faster than the speed of light c. The lowest order damping term to the equations of motion was, in reality, first derived by Eddington in 1924 eight years after the publication of the original Einstein's article on General Relativity.

The concept of a damping force associated with the finite speed of propagation of an interaction was discussed by Lorentz in the beginning of the 20^{th} century in the context of the problem of the motion of electrically charged particles. By direct calculation from first principles of the total force acting on a small extended, charged particle, and taking into account the contribution from its own self-field, led him to the well-known formula

$$\mathbf{F} = \frac{2}{3}\frac{e^2}{c^3}\ddot{\mathbf{v}}, \tag{5.27}$$

where $\mathbf{v}$ is the electron's velocity.

The computation of the radiative reaction in General Relativity is much more complex since the theory is nonlinear and the trajectory of a given particle has to conform with the metric solution of the field equations (*metric induces motion of masses, which perturb the metric, which induces changes to the motion, which further perturb the masses, and so on...*) — in contrast with linear theories as electromagnetism where it possible to specify the motion of a charge and then compute the field that results. The easiest way to understand how the radiation reaction from the emission of gravitational waves affects the motion of the source is to use an *energy balance argument*. An accelerated particle radiates gravitational waves, hence it loses energy, and therefore the appearance of damping terms (and consequent loss of time reversibility) in its equations of motion is inevitable. By equating the negative of the energy radiated in gravitational waves to infinity per unit time (power or luminosity) to the rate of loss of orbital energy, the 2.5 PN acceleration term in the equations of motion can be derived [Peters and Matthews, 1963, Peters, 1964, Blanchet and Schäfer, 1989].

The computation of the energy and angular momentum fluxes at the lowest order, using Einstein's quadrupole formula, requires only the computation of the motion at Keplerian order. Peters & Matthews (1963) and Peters (1964) in two seminal papers assumed a balance equation between the fluxes at infinity (wave zone) and the emission from the source (near zone) to obtain the lowest order radiation damping term. In this way, Peters (1964) was able to derive equations for the secular change of the orbital elements of a two-point mass binary; the orbit-averaged equations read

$$
\begin{aligned}
\left\langle \frac{da}{dt} \right\rangle &= -\frac{64}{5}\frac{G^3 M_1 M_2 (M_1+M_2)}{c^5 a^3 (1-e^2)^{7/2}} \left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right) \\
\left\langle \frac{de}{dt} \right\rangle &= -\frac{304}{15}\frac{G^3 M_1 M_2 (M_1+M_2)}{c^5 a^4 (1-e^2)^{5/2}} \left(1+\frac{121}{304}e^2\right) .
\end{aligned}
\tag{5.28}
$$

A great deal of effort was dedicated to the systematic derivation of the full PN equations of motion to higher orders by many researchers — notably by the Paris group led by Thibault Damour and the Jena group led by Gerhard Schäfer [Damour and Deruelle, 1981, Damour and Deruelle, 1985, Damour and Schäfer, 1988, Schäfer and Wex, 1993]. For a recent comprehensive review see [Blanchet, 2006]. As a result of all this work the full equations of motion up to 3.5PN order for spinless particles are now known. This is a highly non-trivial achievement, completed only several decades after the foundations of General Relativity were laid out by Einstein. In this work, we employ only the development up to 2.5PN order and, for this reason, we write down here the corresponding equations of motion up to that order. The full PN equations of motion up to 2.5PN order, written in the center of mass frame of a binary system, are given by [Blanchet and Iyer, 2003]

$$
\frac{d\mathbf{v}}{dt} = -\frac{GM}{r^2}\left[(1+\mathcal{A})\mathbf{n}+\mathcal{B}\mathbf{v}\right]+\mathcal{O}\left(\frac{1}{c^6}\right) , \tag{5.29}
$$

where $M = M_1 + M_2$ is the binary's total mass, $\mathbf{n} = \mathbf{r}/r$, and the coefficients $\mathcal{A}$ and $\mathcal{B}$ are defined by

$$
\begin{aligned}
\mathcal{A} = \quad & \frac{1}{c^2}\left[-\frac{3\nu}{2}\dot{r}^2+(3+\nu)v^2-\frac{GM}{r}(4+2\nu)\right]+ \\
+ \quad & \frac{1}{c^4}\left\{\frac{15}{8}\nu(1-3\nu)\dot{r}^4+6\nu(\nu-\frac{3}{4})\dot{r}^2v^2+\nu(3-4\nu)v^4+\right. \\
+ \quad & \left.\frac{GM}{r}\left[2\nu(\nu-\frac{13}{4})v^2-2(\nu^2+\frac{25}{2}\nu+1)\dot{r}^2\right]+\frac{G^2M^2}{r^2}(9+\frac{87\nu}{4})\right\}+ \\
+ \quad & \frac{1}{c^5}\left[-\frac{8}{5}\frac{GM}{r}\nu\dot{r}(3v^2+\frac{17}{3}\frac{GM}{r})\right]+\mathcal{O}\left(\frac{1}{c^6}\right)
\end{aligned}
\tag{5.30}
$$

$$\begin{aligned}\mathcal{B} = \quad & \frac{1}{c^2}\left[2(\nu-2)\dot{r}\right] + \\ + \quad & \frac{1}{c^4}\left[3\nu(\nu+\frac{3}{2})\dot{r}^3 - 2\nu(\nu+\frac{15}{4})\dot{r}v^2 + \frac{2GM}{r}\left(1+\frac{41}{4}\nu+2\nu^2\right)\dot{r}\right] + \\ + \quad & \frac{1}{c^5}\left[\frac{8\nu}{5}\frac{GM}{r}\left(v^2+\frac{3GM}{r}\right)\right] + \mathcal{O}\left(\frac{1}{c^6}\right), \end{aligned} \tag{5.31}$$

where $\nu = M_1M_2/M^2$ which → 0 in the test particle limit and → 1/4 in the equal mass case.

Schematically, the PN equations of motion for spinless particles of arbitrary mass ratio are thus of the following form

$$\frac{d\mathbf{v}}{dt} = -\frac{GM}{r^2}\mathbf{n}\left[1+\mathcal{O}(\epsilon)+\mathcal{O}(\epsilon^2)+\mathcal{O}(\epsilon^{2.5})+\mathcal{O}(\epsilon^3)+\mathcal{O}(\epsilon^{3.5})+\cdots...\right], \tag{5.32}$$

where $\epsilon \sim (v/c)^2 \sim GM/rc^2$ is the small expansion parameter. There are no terms of order $\mathcal{O}(\epsilon^{0.5})$ and $\mathcal{O}(\epsilon^{1.5})$ the reason being that the lowest term for emission of gravitational waves is the quadrupole, as discussed before. [§] The odd terms in the PN expansion must be dissipative since a change of sign in the velocity $v \to -v$ and in time $t \to -t$ implies that the corresponding piece of the vector field also changes sign turning an inspiral into an outspiral motion. This split between even-order conservative and odd-order dissipative terms breaks down at $\mathcal{O}(\epsilon^4)$ because of the appearance of tail contributions to the gravitational waves [Blanchet, 2006].

The conservative part of the PN equations of motion in harmonic coordinates can be derived from a generalized Lagrangian that depends not only on the positions and velocities, but also on the accelerations. It is a known theorem that the N-body equations of motion cannot be derived from an ordinary Lagrangian (depending on positions and velocities) beyond the $1PN$ order if the equations of motion are to remain Lorentz invariant (see [Blanchet, 2006] and references therein). They can, nevertheless, be derived from the Lagrangian via the standard Euler-Lagrange equation provided the resulting equations of motion are properly *order-reduced*, i.e. to substitute the accelerations by those of the order just below.

[§] The monopole term $\propto v/c$ is constant by conservation of energy-momentum; the dipole term $\propto (v/c)^3$ is constant due to conservation of total linear momentum.

5.4 Numerical integration of the PN equations of motion

We integrate the PN equations of motion by applying the two-body KS regularization and using the Hermite scheme as integrator. In order to do this, it is necessary to compute not only the accelerations at each time step, but also its time derivatives. The time derivatives of the already long expressions (5.29), (5.30) and (5.31) for the PN accelerations involves a long, but straightforwrad, calculation of a large number of derivatives. This has been already done for the PN equations of motion written in a general inertial frame [Kupi et al., 2006]; it was repeated now for the equations written in the center of mass frame. Extensive comparisons between the codes showed a very good agreement between the numerical results from these two, independently derived, formulations. They constitute a strong numerical check on the long expressions for the time derivatives of the accelerations and give us confidence that they do not contain unnoticed algebraic errors. Other comparisons with the standard Runge-Kutta integrator and symplectic integrators [Preto and Tremaine, 1999] — none of which requires the knowledge of the time derivatives of the accelerations — also produced very good agreement with the numerical integration tests.

The integration of the PN equations of motion with the two-body KS regularization scheme treats the PN terms formally as an external perturbation to the Keplerian motion. Under this scheme, the Kepler motion needs not to be dominant or, in other words, the PN corrections may be large with no detriment to the quality of the integration. This is, therefore, a very useful tool to study the PN motion until the very late stages of inspiral, before the sudden plunge and relativistic coalescence.

A standard check to the accuracy of the integration of a conservative system is the level of energy conservation. Since the even order terms (1PN and 2PN) have an associated conserved energy, it is possible to monitor the fractional energy error $\Delta E/E$, provided that one interprets carefully the results according to the precision to which this quantity is conserved at each given order of the PN expansion. In fact, to (5.32) there is an associated energy E which is schematically of the form

$$E = E_N + \frac{1}{c^2} E_{1PN} + \frac{1}{c^4} E_{2PN} + \mathcal{O}\left(\frac{1}{c^6}\right). \tag{5.33}$$

In Figure 5.1, it is shown the results from two integrations with $c = 3$ and $c = 10$ for an equal mass binary with initial conditions equal to those of a Keplerian circular orbit for the given mass and positions. The integration is done with the radiation damping term turned-off; 1PN and 2PN are turned-on. The green curve represents $\Delta E/E$ for the $c = 3$ case. These high oscillation values do not represent a poor numerical accuracy, but rather the fact that the energy expression (5.33) includes only the first three terms of the series expansion, that we have chosen an unrealistic high value for c and therefore the convergence of the series becomes very slow. It is possible to verify the good accuracy of the integration by taking the time derivative of (5.33)

$$\begin{aligned} \frac{dE}{dt} &= \frac{dE_N}{dt} + \frac{1}{c^2}\frac{dE_{1PN}}{dt} + \frac{1}{c^4}\frac{dE_{2PN}}{dt} + \mathcal{O}\left(\frac{1}{c^6}\right) = \\ &= (\mathbf{v}\cdot\mathbf{a}) - \frac{GM}{r^2}(\mathbf{n}\cdot\mathbf{v}) + \frac{1}{c^2}f_1(\mathbf{x},\mathbf{v},\mathbf{a}) + \frac{1}{c^4}f_2(\mathbf{x},\mathbf{v},\mathbf{a}) + \mathcal{O}\left(\frac{1}{c^6}\right), \end{aligned} \tag{5.34}$$

where the first two terms are the 0^{th} order Keplerian contribution; f_1 corresponds to a complicated expression of 1^{st} order depending on the positions, velocities *and* accelerations; f_2 is the equivalent (longer) expression of 2^{nd} order. After order-reducing this expression, i.e. substituting the accelerations where they appear on the right hand side by the equation of motion, it can be shown (after some algebra) that the terms vanish identically at each order, and only an $\mathcal{O}\left(\frac{1}{c^6}\right)$ oscillation remains. The red curve in Figure (5.1) represents the time derivative dE/dt computed by feeding (5.34) with the values $\mathbf{x}$ and $\mathbf{v}$ from the numerical integration's output, after it has been properly order reduced. It can be readily seen that the accuracy to which the energy is conserved to very high accuracy $\sim 10^{-13}$. The remaining curves in the same figure refer to a similar integration, only now with $c = 10$.

An accurate integration of circular orbits is very important for the determination of the innermost stable circular orbit (ISCO). This is the separation radius r_{ISCO} between the two bodies below which a stable, circular orbit suddenly becomes unstable and plunges — even in the absence of radiation damping. For massless particles orbiting a Schwarschild black hole, the radius of the ISCO is known analytically: $r_{ISCO} = 5GM/c^2$, in harmonic coordinates. This raises the question of whether there exists an analogous ISCO for two particles of comparable mass. In the context of the PN approximation, one can study the linear stability of the circular orbits and obtain an analytic estimate of the radius of the ISCO. This can then be used as another check on the accuracy of the numerical integration.

The first step in this analysis is the determination of the exact frequency for a circular orbit of a PN system. In order to compute this frequency, it is convenient to write the PN equations of motion, in the center of mass frame, in polar coordinates [Blanchet and Iyer, 2003]:

$$\begin{aligned} \dot{r} &= u \\ \dot{u} &= -\frac{GM}{r}\left[1 + \mathcal{A} + \mathcal{B}u\right] + r\omega^2 \\ \dot{\omega} &= -\omega\left[\frac{GM}{r^2}\mathcal{B} + \frac{2u}{r}\right]. \end{aligned} \tag{5.35}$$

The condition for a circular orbit is $\dot{r} = \dot{\omega} = 0$, so $\mathcal{B}_0 = 0$ and $u = 0$ for a circular orbit as well. Therefore, solving for the circular frequency at 1PN order

$$\omega_0^2 = \frac{GM}{r_0^3} \frac{1 - (4 + 2\nu)\frac{GM}{r_0 c^2}}{1 - (1 + 3\nu)\frac{GM}{r_0 c^2}}. \tag{5.36}$$

This can be expanded in the small quantity $GM/r_0 c^2$ to arrive at

$$\omega_0^2 = \frac{GM}{r_0^3} \left[1 + (\nu - 3)\frac{GM}{r_0 c^2}\right] \tag{5.37}$$

which coincides agrees with eqn. (190) in Blanchet (2006) to 1PN order. The same procedure can be followed to 2PN order; in this case, one needs to solve a quadratic equation for the circular frequency ω_0^2; at 3PN order, there will be a cubic equation in ω_0^2 and so on. The different circular frequencies as a function of radius are plotted in Figure (5.4). The corresponding orbits are shown in Figures (5.2) and (5.3). Figure (5.2) shows, for the $c = 3$ case, the orbits resulting from choosing the circular frequency according to the Keplerian, the PN expanded or the exact value given by (5.36). Figure (5.3) shows the corresponding orbits for the $c = 10$ case — it is clear that the convergence of the PN expansion improves with larger value of c and it converges rather rapidly to the circular frequency value. The inclusion of higher order terms, e.g. 3PN, should also improve the convergence. However, it should be noted that the PN expansion parameter is GM/rc^2; for sufficiently small radii, even for realistic values of c, the expressions from the PN expansion will nevertheless show some deviation from the exact value of the circular frequency. In Figure (5.4) the several definitions of circular frequency are plotted as a function of radius. Note that the exact frequency curve for the 1PN case is ill-defined for $r < 1$ (if $c = 3$) and for $r < 0.09$ (if $c = 10$). The convergence trends with increasing c and radius r are immediately captured upon inspection of the plots.

Now it is possible to proceed with the analytic estimate for the location of the ISCO for a binary system with an arbitrary mass ratio. A circular orbit is defined by $r = r_0$, $u_0 = 0$ and $\omega = \omega_0$. Consider then small perturbations to the circular orbit

$$\begin{aligned} r &= r_0 + \epsilon_r \\ u &= \epsilon_u \\ \omega &= \omega_0 + \epsilon_\omega. \end{aligned} \tag{5.38}$$

The equations that govern small perturbations around the circular orbit are then given by

$$\begin{aligned} \dot{\epsilon}_r &= \epsilon_u \\ \dot{\epsilon}_u &= \alpha_0 \epsilon_r + \beta_0 \epsilon_\omega \\ \dot{\epsilon}_\omega &= \gamma_0 \epsilon_u, \end{aligned} \tag{5.39}$$

where the coefficients α_0, β_0 and γ_0 depend only on the parameters of the (unperturbed) circular orbit

$$\begin{aligned} \alpha_0 &= 3\omega_0^2 - \frac{GM}{r_0^2}\left(\frac{\partial \mathcal{A}}{\partial r}\right)_0 \\ \beta_0 &= 2r_0\omega_0 - \frac{GM}{r_0^2}\left(\frac{\partial \mathcal{A}}{\partial \omega}\right)_0 \\ \gamma_0 &= -\omega_0\left[\frac{2}{r_0} + \frac{GM}{r_0^2}\left(\frac{\partial \mathcal{B}}{\partial u}\right)_0\right]. \end{aligned} \tag{5.40}$$

We can rewrite this equation in matricial form

$$\begin{pmatrix} \dot{\epsilon}_r \\ \dot{\epsilon}_u \\ \dot{\epsilon}_\omega \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ \alpha_0 & 0 & \beta_0 \\ 0 & \gamma_0 & 0 \end{pmatrix} \begin{pmatrix} \epsilon_r \\ \epsilon_u \\ \epsilon_\omega \end{pmatrix}, \tag{5.41}$$

and solve for its eigenvalues. The eigenvalues are: $\lambda = 0$ and $\lambda = \pm\sqrt{\alpha_0 + \beta_0\gamma_0}$. The solution is a linear combination of $e^{\lambda t}$ and $e^{-\lambda t}$, so the stability condition is determined by the value of the stability parameter C_0:

$$C_0 = \alpha_0 + \beta_0\gamma_0 < 0. \tag{5.42}$$

The stability parameter C_0 is plotted as a function of r in Figure (5.5) for cases $c = 3$ and $c = 10$. The ISCO is located at radius $r \sim 1.515$ for $c = 3$, and $r \sim 0.136$ for $c = 10$; these are the radii where the curve for C_0 changes sign. The integrator captures very accurately the radius of the ISCO, as can be seen from Figure(5.6): all orbits with $r < r_{ISCO}$ plunge and lead to a relativistic coalescence; for $r > r_{ISCO}$ all circular orbits are stable (numerically, they show no sign of a secular component after 10^3 orbits). The numerical transition is very abrupt as the theory predicts, and happens at the radius estimated by the linear stability analysis.

Once the 2.5PN term is included, there is dissipation and the system ceases to be time-reversible. The 1PN and 2PN terms have no secular component and lead to quasi-periodic behavior — for instance, the orbital precession —, but the 2.5PN radiation damping term introduces a secular effect and leads to a steady loss of energy, inspiralling motion and eventually, at late times, to a sudden plunge and relativistic coalescence. Peters (1964) equations of motion, that completely neglect the quasiperiodic effects associated to the 1PN and 2PN terms, imply a merging time scale given by

$$t_{GW} \sim \left|\frac{E}{\dot{E}}\right| \sim \frac{5c^5}{64G^3}\frac{a^4(1-e^2)^{7/2}}{M_1 M_2(M_1+M_2)}\frac{1}{1+\frac{73}{24}e^2+\frac{37}{96}e^4}. \tag{5.43}$$

This time scale decreases very rapidly with decreasing semimajor axis a and — very importantly for us — with high eccentricities. This is easily understood since gravitating bodies, at small separations, emmit gravitational waves much more strongly and therefore the more so for closer pericenter passages. The physical picture associated with the Peters equations is that of an adiabatic inspiral through a sucessive series of Keplerian orbits until at the late stages (close to the ISCO), the two bodies suddenly plunge towards each other. If the orbit has an appreciable eccentricity it will be circularized with time until only a very small residual eccentricity will remain just before it plunges. Once the relativistic quasiperiodic effects are included with the 1PN and 2PN terms in the equations of motion, how does the merging timescale (5.43) is affected? In Figure (5.7) the red curves describe the motion with the radiation damping term (2.5PN) turned on, but with the other PN terms off. The green curve corresponds to the motion with all PN terms up to 2.5PN included. The timescale for merger depends on the initial conditions: including the conservative PN terms can make it shorter or longer. In the upper plot, the initial velocity is that of a Keplerian circular orbit corresponding to the considered initial separation and masses. As it can be seen from Figure (5.4), at separation r=2 and for $c = 3$, the circular frequency when the PN1 and PN2 terms are included is lower than Keplerian. As a result, if we impose the Keplerian velocities to this system it will have a larger tangential velocity than that required for it keeping circular and will evolve towards a larger radius. The result is that it effectively starts on a rosette-type orbit with a mean radius larger than the initial radius; in other words, with a larger semimajor axis, a non-zero eccentricity and a pericenter equal to the radius of the Keplerian circular orbit. The PN2.5 orbit just inspirals through a succession of nearly circular orbits with decreasing semimajor axis until, after ~ 10 orbits, it approaches the ISCO and plunges. When the exact frequency for the circular orbit of the full 1PN+2PN system is taken, the full PN integration plunges sooner than the 2.5PN one. Figure (5.8) shows the same trend for the $c = 10$ case, but the percentual difference in the merging time scale is smaller; we see that, for circular orbits with realistic values of c, the merging time scales become very similar.

Figure (5.9) shows the evolution in time of the orbital elements of a binary whose initial conditions correspond to those of a Keplerian orbit with eccentricity $e = 0.9$, with $c = 450$ (a typical value for the speed of light in a N-body model of a large galaxy in our units). The difference in merging time is still $\sim 15\%$ despite the fact that c is 45 times larger than that of Figure (5.8). This has a simple physical explanation: when the 1PN and 2PN terms are included, the orbit precesses and, as a result, it makes over time a larger number of pericenter passages than the 2.5PN only (non-precessing) orbit. Therefore, since it is at smaller separations that the bodies dissipate a larger amount of orbital energy, the merging time scale for the full PN orbit is shorter whenever the orbit has a non-negligible eccentricity.

Finally, Figure (5.10) shows the amount of energy lost during the inspiral for the orbits in Figure (5.9). Two curves are plotted on each panel: one, it is the luminosity $\mathcal{L}$ computed from the Einstein's quadrupole formula

$$\mathcal{L} = \left\langle \frac{dE}{dt} \right\rangle = \frac{32G^4}{5c^5} \frac{M_1^2 M_2^2 (M_1 + M_2)}{a^5 (1 - e^2)^{7/2}} (1 + \frac{73}{24} e^2 + \frac{37}{96} e^4). \tag{5.44}$$

The orbital elements $a(t)$ and $e(t)$, obtained from the numerical simulation, are substituted into this formula which is then integrated over the time evolution of the binary. The other curve is obtained directly from the time evolution of the energy E of the system. The two curves fall on top of each other and the agreement is indeed very good.

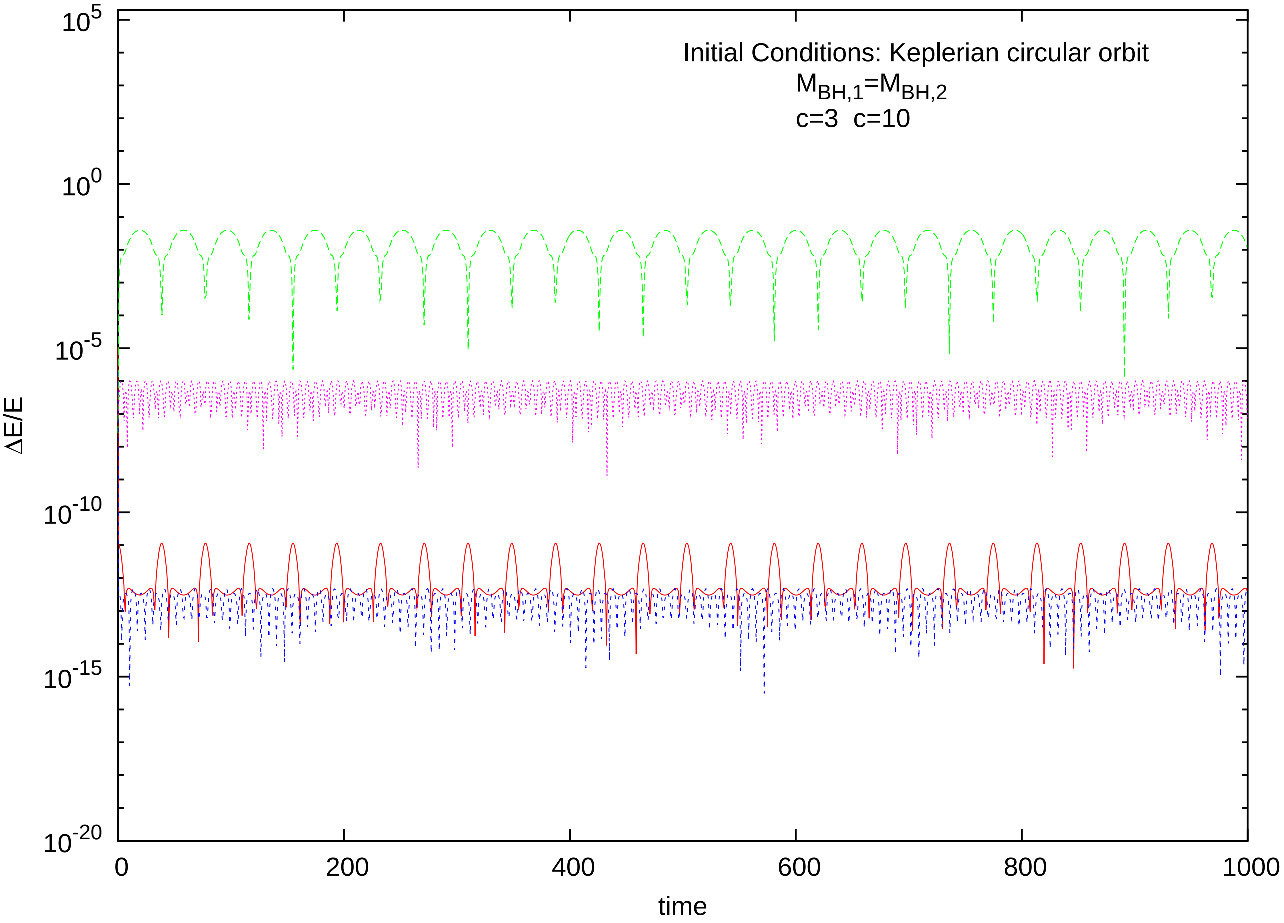

Figure 5.1: Time evolution of the fractional energy error $\Delta E/E$ for a binary system with PN1+PN2 terms turned-on, but with no radiation damping. Green curve: $c = 3$; pink: $c = 10$. Time evolution of dE/dt for the same system. Red: $c = 3$; Blue: $c = 10$. The initial conditions were Keplerian circular orbit: $x_1 = -x_2 = 1.0$ and $v_{y1} = -v_{y2} = 0.5$ for equal masses $M_1 = M_2 = 1.0$.

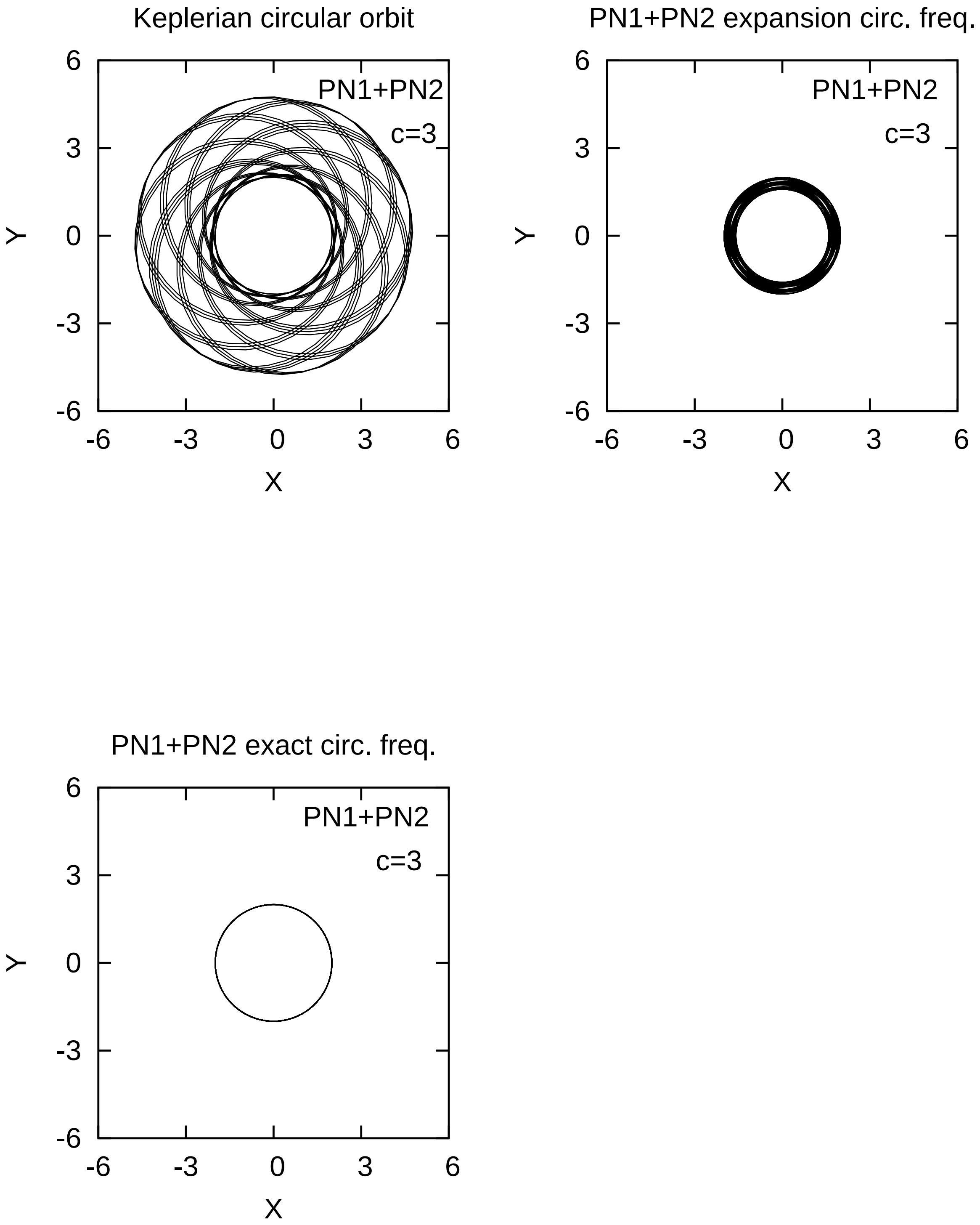

Figure 5.2: Equal mass binary with PN1+PN2 terms, but with no radiation damping. Upper left panel: orbit with initial velocity given for a Keplerian circular orbit; Upper left panel: initial velocities from PN expanded circular frequency; Lower left panel: initial velocities from exact circular frequency. The initial positions are: $x_1 = -x_2 = 1.0$. The masses are: $M_1 = M_2 = 1.0$.

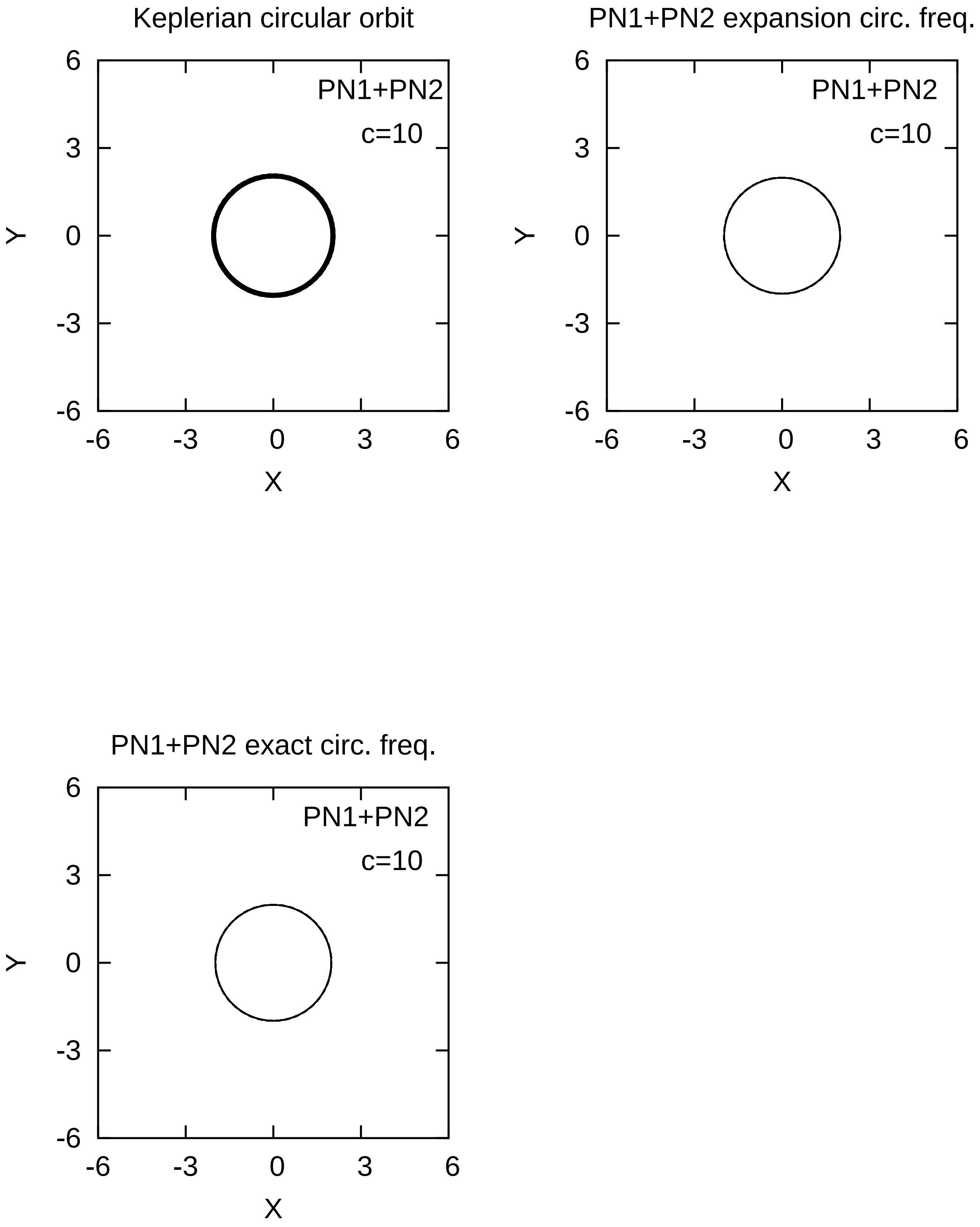

Figure 5.3: The same as Figure 5.2 for the case $c = 10$.

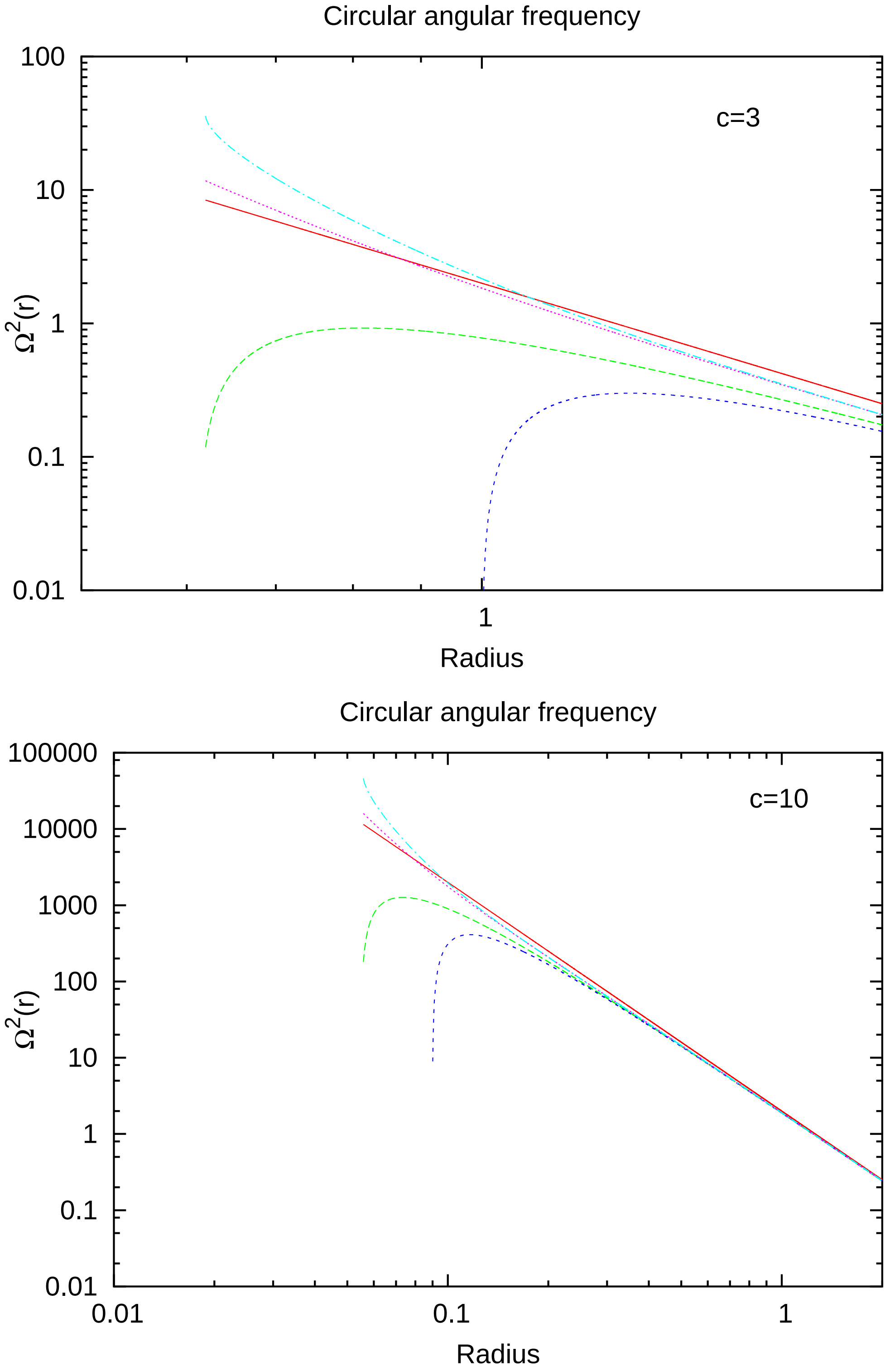

Figure 5.4: Squared circular frequency as a function of radius from different expressions. For both panesls: red curve: Keplerian frequency; green: PN1 expansion circ. freq.; dark blue: PN1 exact freq.; pink: PN1+PN2 expansion freq.; light blue: PN1+PN2 exact freq. The upper panel is for $c = 3$; lower panel for $c = 10$. The equal mass binary $M_1 = M_2 = 1.0$. All curves converge to the same value for large radius; the convergence is faster for larger values of c.

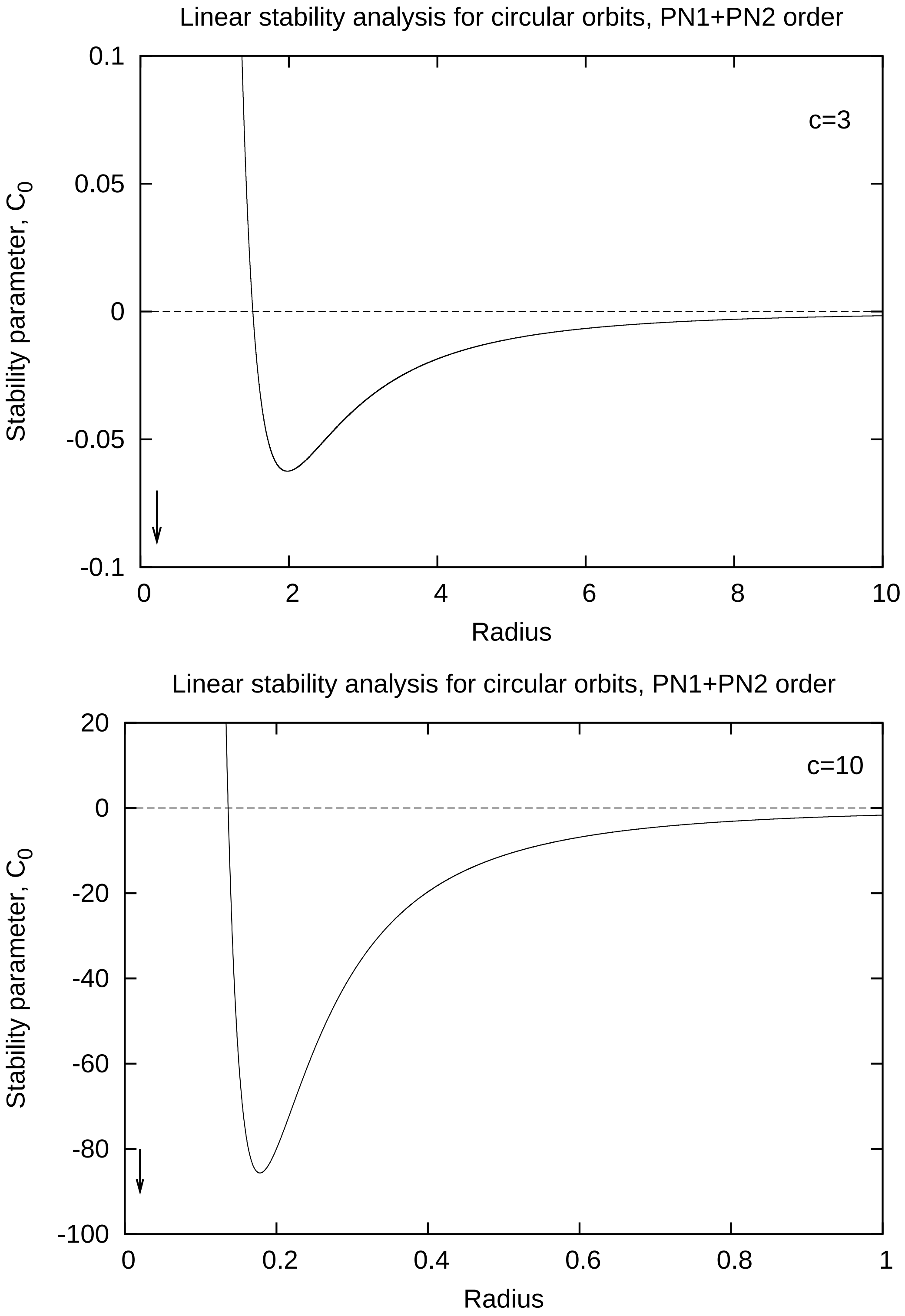

Figure 5.5: Linear stability analysis parameter C_0 for PN1+PN2 circular orbits as a function of radius. C_0 changes sign at the radius corresponding to the (generalized) innermost stable circular orbit (ISCO) for an equal mass binary system ($M_1 = M_2 = 1.0$). For $r > r_{ISCO}$, all circular orbits remain stable after 10^3 orbital periods, without any sign of a secular drift in radius. All circular orbits with $r < r_{ISCO}$ plunge suddenly during the first orbital period. The arrow points to the Schwarschild radius of the black holes. For $c = 3$, r_{ISCO}=1.515; for $c = 10$, $r_{ISCO} = 0.136$.

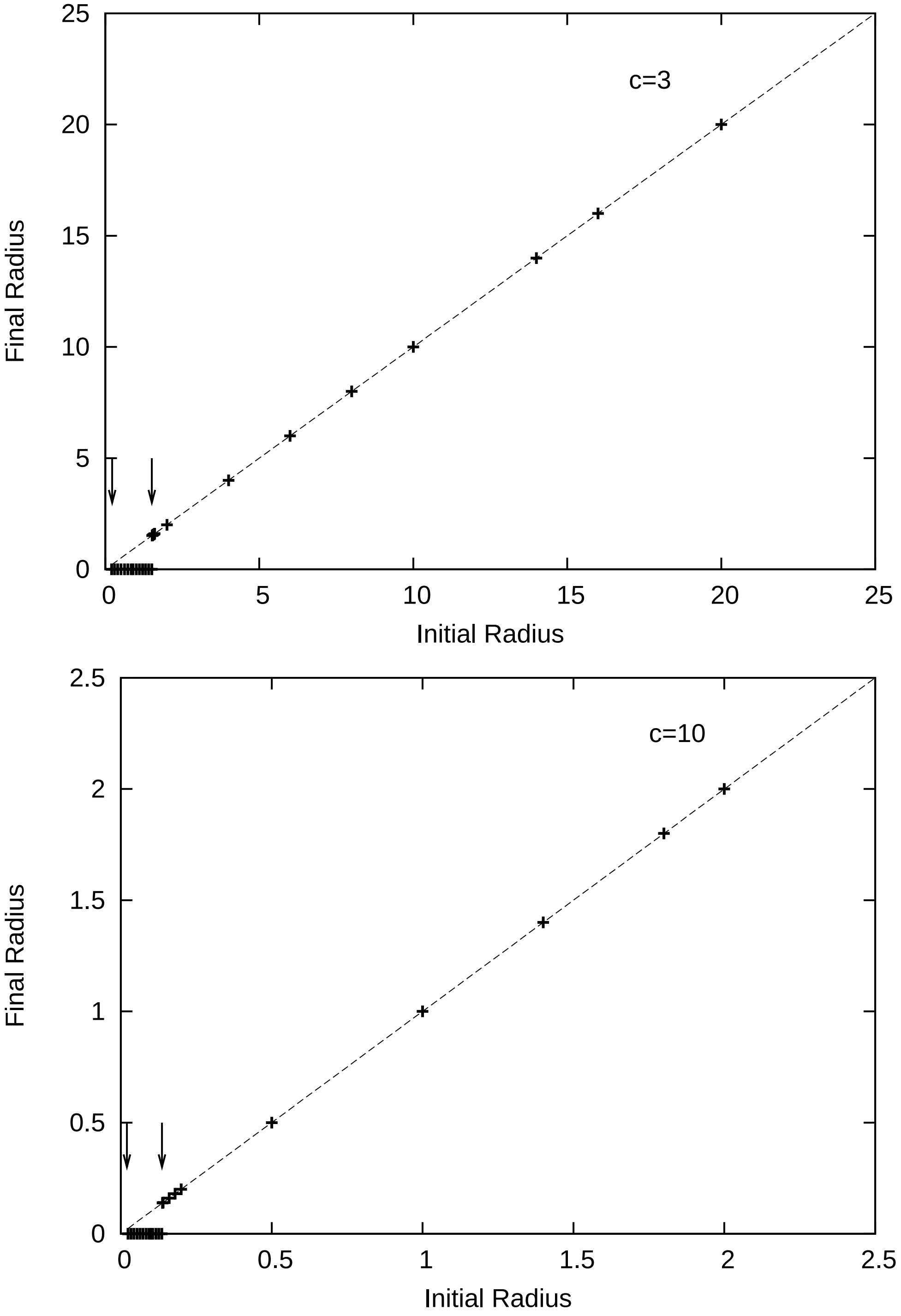

Figure 5.6: The numerical integration captures the ISCO behavior in accordance with the linear stability analysis. The abcissa denotes the radius of the circular orbit at the beginning of the integration; the ordinate denotes the final radius after 10^3 orbital periods. All integrations with $r > r_{ISCO}$ fall right on top of the line $y = x$. For $r > r_{ISCO}$, all circular orbits remain stable after 10^3 orbital periods, without any sign of a secular drift in radius. All circular orbits with $r < r_{ISCO}$ plunge suddenly during the first orbital period. The left arrow points to the Schwarschild radius of the black holes; right arrow: points to r_{ISCO}.

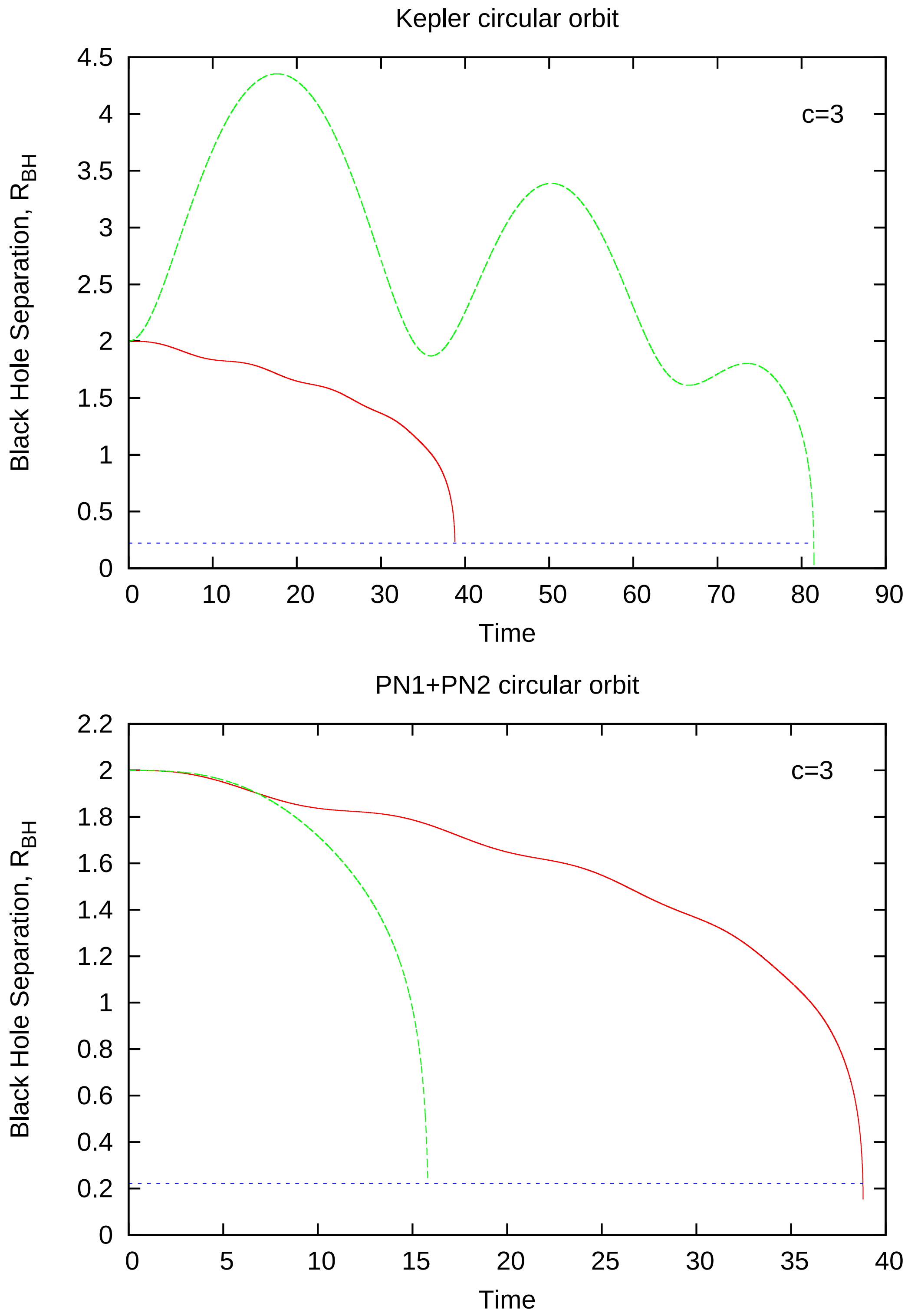

Figure 5.7: Time evolution of an equal-mass binary system with radiation damping. Red curve: PN1 & PN2 turned-off. Green curve: PN1+PN2 turned on. Top panel: initial velocity is that of a Keplerian circular orbit. Lower panel: initial velocity derived from the exact PN1+PN2 circular frequency. $M_1 = M_2 = 1.0$ and $c = 3$.

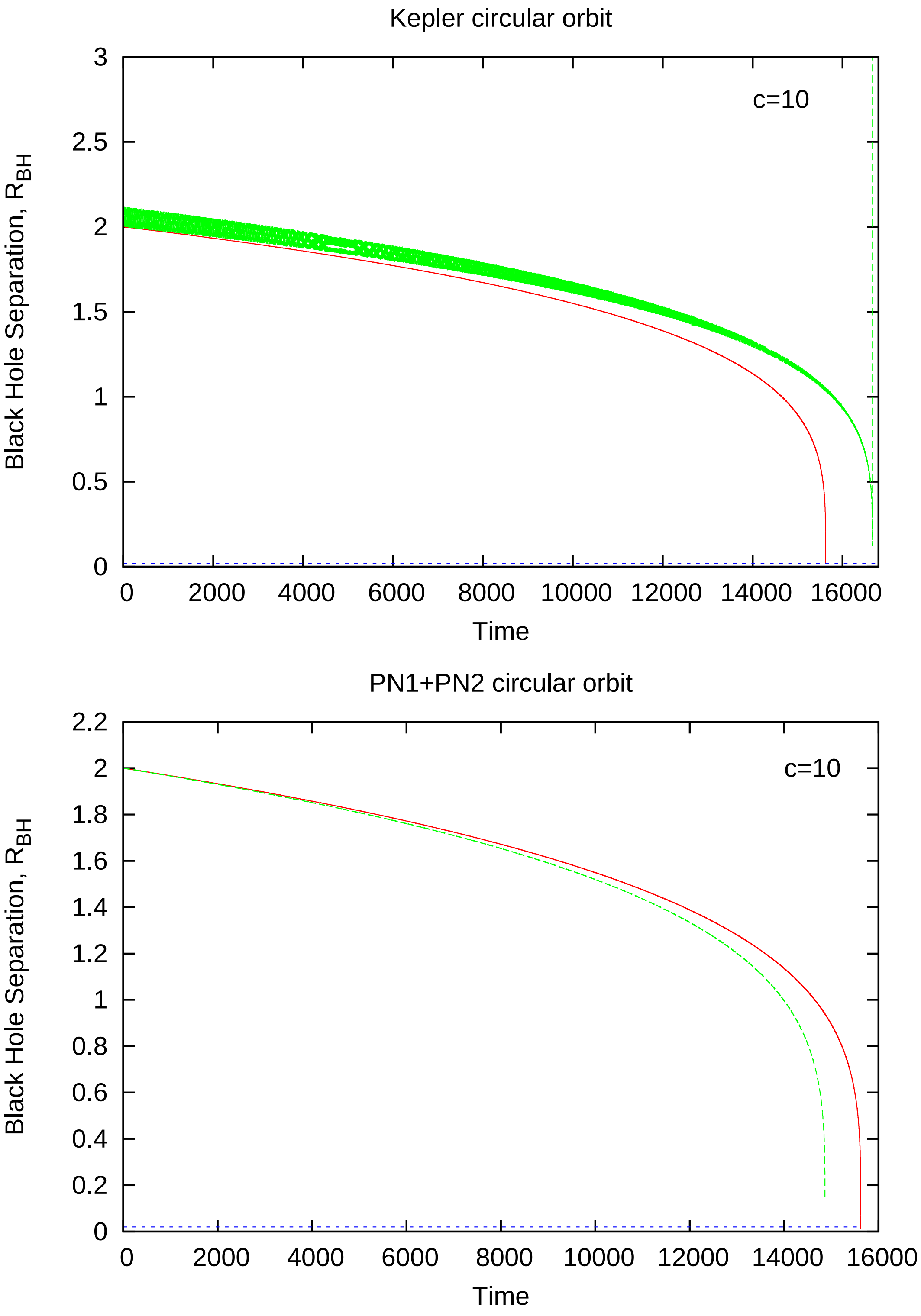

Figure 5.8: The same as Figure 5.7 for the $c = 10$ case.

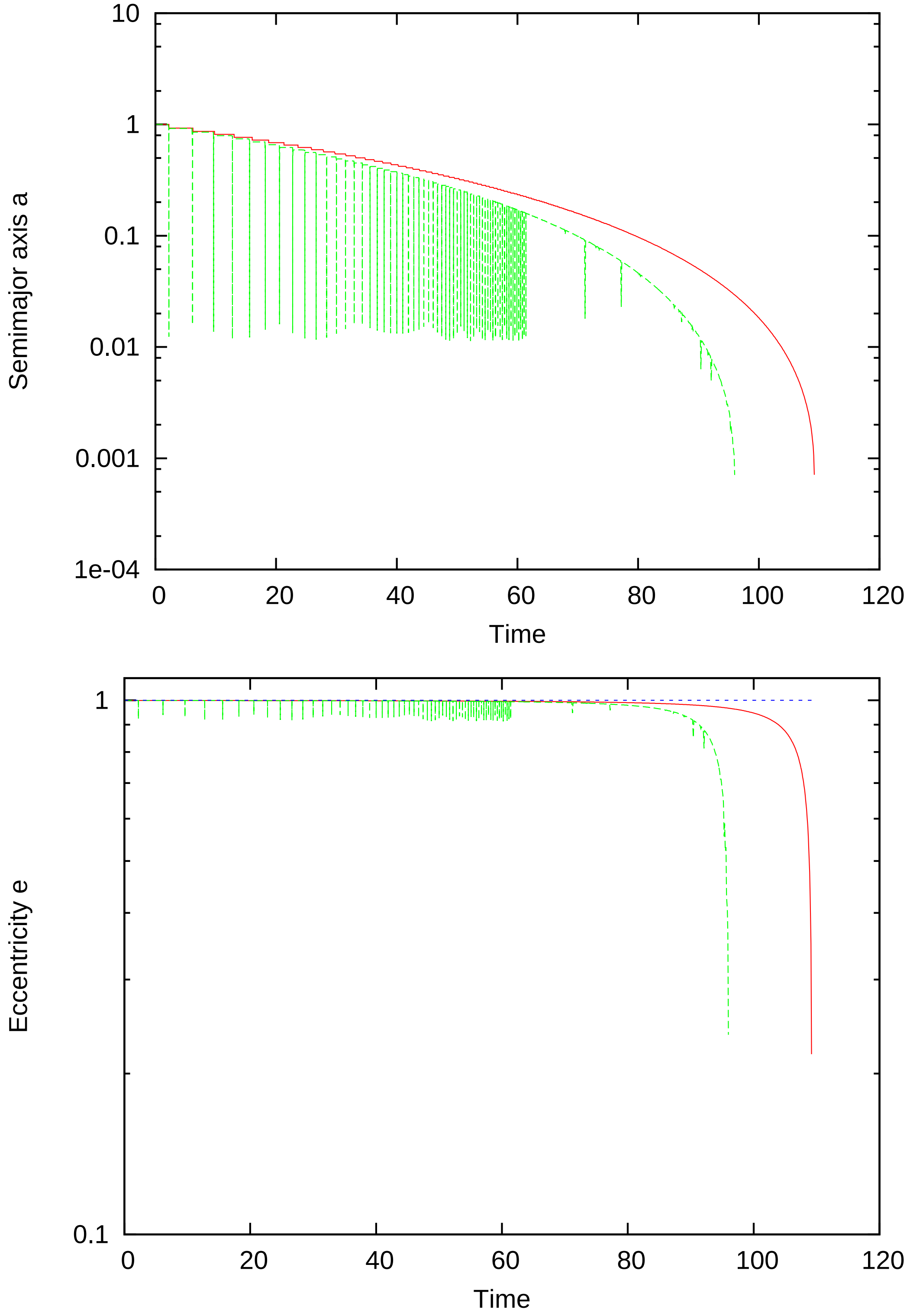

Figure 5.9: Time evolution of the semimajor axis $a(t)$ and eccentricity $e(t)$ for an equal mass binary ($M_1 = M_2 = 1.0$) and $c = 450$. The initial velocities were those of a Keplerian orbit with $e = 0.9$. The binary slowly spirals in for many orbital periods until it suddenly plunges near the ISCO.

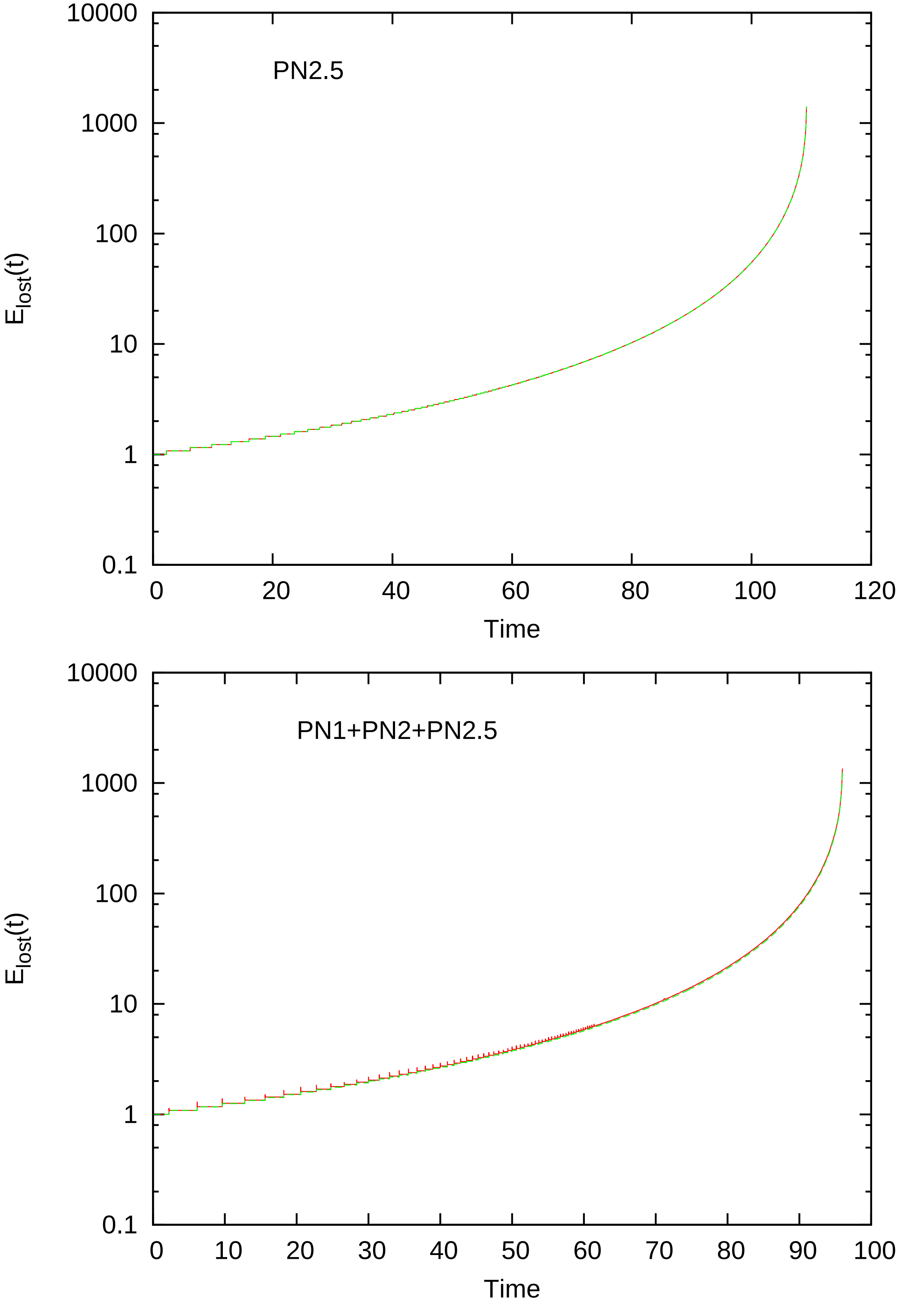

Figure 5.10: Comparison between the energy loss predicted from the Einstein's quadrupole formula with the energy loss computed from the integrated orbit. Top: orbit integrated with 2.5PN terms only; Bottom: orbit integrated with all PN terms on. The agreement is quite good as curves fall on top of each other for both cases. In this test case, the speed of light $c = 450$.

5.5 Gravitational wave signal from massive black hole binaries and LISA detection

Nearly all galaxies contain massive black holes (MBHs) at their centers, and the present day masses are consistent with the idea that they grow by coalescence during hierarchical structure formation [Yu and Tremaine, 2002, Volonteri et al., 2003b, Sesana et al., 2005]. The emission of gravitational waves is a generic prediction of General Relativity, and it is thought that black hole binaries are amongst the most promising sources of GWs — both for ground-based interferometers such as VIRGO and LIGO (for binaries in the $10 - 50 M_\odot$ mass range), as for the planned ESA-NASA satellite mission Laser Interferometer Space Antenna (LISA), for binaries with total mass $\gtrsim 10^4 M_\odot$.

The LISA band of sensitivity will fall in the frequency range $10^{-5} - 1$ Hz, for which GW emission from a cosmological distribution of MBH binaries will be detectable; and also — if their existence is confirmed — from intermediate-mass black hole (IMBH) binaries residing in globular clusters [Miller and Colbert, 2004, Amaro-Seoane and Freitag, 2006]. In addition to individual MBH mergers and inspiralling events, the superposition of unresolved MBH binary signals during the late stages of their evolution creates the so-called residual confusion noise. The question of whether the number of MBH binary events will lead to a stochastic background or not depends very much on the predicted event rates: in models according to which one expects ≈ 10 events per year at redshifts $2 \lesssim z \lesssim 6$, individual detections will be possible [Sesana et al., 2005, Sesana et al., 2007]; in models that predict hundreds or thousands of MBH binaries coalescences per year will produce a strong stochastic background in the LISA data [Berti, 2006]. Most of the published estimates to date were made under the assumption that the MBH binaries will arrive at the LISA frequency band in circular orbits (zero eccentricity). However, all estimates will be seriously affected if the orbital element distribution has a significant component with non-zero eccentricity. Therefore it is of crucial importance to do detailed N-body studies of the MBH binary evolution after the galaxy merger until it reaches the relativistic phase and the LISA frequency band.

Moreover, it is not only the issue of confusion noise amplitudes that is affected by the distribution of orbital elements, but also the mass range for the detectability of individual events resolved by LISA. In short, the physical argument is the following. Circular orbits will only lead to the emmission of GWs at the fundamental frequency which, in the case of GWs, is equal to twice the orbital frequency $f_{GW} = 2 f_{orb}$; however, if the binary reaches the LISA band with some non-zero eccentricity other harmonics of frequency $f_{GW,m} = m f_{orb}$ will also contribute to the signal. Now, since heavier MBHs have larger Schwarschild radius, they coalesce at lower frequencies. This means that, again assuming circular orbits, MBH binaries with $M \gtrsim \text{few} \times 10^7 \mathrm{M}_\odot$ will fall out of the LISA band due to the latter's low-frequency cut-off [Sesana et al., 2005, Sesana et al., 2007]. However, higher harmonics may change the signal-to-noise ratio for binaries currently believed to fall out of the LISA band and there is the possibility that the mass reach of LISA will turn out — for the cases with non-zero eccentricity — to be larger and that it will be able to observe the relativistic inspirals of MBH binaries with total masses of $\sim 10^8 - 10^9 M_\odot$ [Arun et al., 2007].

Galaxy mergers are thought to be common in the universe. If two galaxies, each with a MBH at its center, merge the MBHs will spiral in towards the center of the resultant galaxy under the influence of dynamical friction until they become a bound binary. After this, the binary semimajor axis will decrease because of the dynamical friction from the distant stars. The binary becomes hard when its semimajor axis falls below $a_h = G(M_1 + M_2)/4\sigma^2$ or, equivalently, when its orbital velocity becomes larger than the velocity dispersion of the surrounding stars. At this point, dynamical friction becomes inefficient and the MBH binary continues to harden due to the gravitational slingshot effect. This is the phase that, in most circumstances, is expected to last the longest; it is critical to determine, for a given MBH binary and galaxy environment, whether the time it takes for the binary to shrink up to the point where the relativistic regime becomes important is, or isn't, shorter than a Hubble time [Begelman et al., 1982].

Stars that pass close to the hard binary at distances of a few $(2-3)$ semimajor axis undergo a strong three-body interaction with the MBH binary and are, on average, ejected with high velocity $v_{ej} \sim \sqrt{M_1 M_2/(M_1+M_2)^2} v_{bin}$ — this is the so-called gravitational *slingshot effect* [Merritt and Milosavljevic, 2005]. As the ejected field stars gain kinetic energy, the hard MBH binary loses orbital energy and becomes more strongly bound (it *hardens*); its eccentricity and inclination also wander slowly in a random fashion. If the background density of field stars did not change with time (it does!), the evolution of the binary's orbital elements could be computed via three-body scattering experiments [Quinlan, 1996].

Following Quinlan (1996), the results from these scattering experiments can be described via a set of approximate differential equations. The hardening rate of the MBH binary is given by

$$\frac{d}{dt}\left(\frac{1}{a}\right) = H\frac{G\rho}{\sigma}, \tag{5.45}$$

where ρ is the stellar mass density, σ is the 1D-velocity dispersion, and H is a dimensionless coefficient to be measured from the scattering experiments. The mass ejection rate is given by

$$\frac{dM_{ej}}{d\ln(1/a)} = (M_1 + M_2)J, \tag{5.46}$$

where M_{ej} is the ejected stellar mass and J is another dimensionless coefficient. The evolution of the binary's eccentricity is governed by

$$\frac{de}{d\ln(1/a)} = K. \tag{5.47}$$

However, things are obviously more interesting than this simple picture would imply. First, the mass density that surrounds the MBH binary decreases at it hardens by ejecting stars. Second, at each radius in the galaxy only stars with angular momentum smaller than a given value $J < J_{lc}$ reach close enough to the binary to be ejected: all those stars are ejected after a dynamical time scale. Third, as a result the MBH binary's loss cone drains after just roughly one dynamical time t_{dyn} which means that the supply of stars to the the MBH binary is rapidly exhausted. Fourth, for typical galaxies (virialized, close to spherical symmetry) the only way to replenish the loss cone is via the diffusion of stars in angular momentum space; this is a dynamical process that occurs on a relaxation time scale $t_{\rm rlx} \gg t_{dyn}$. Fifth, by inspecting Figure 3.1 the relaxation time scales range from few $\times\, 10^9$ years for galaxies with smaller MBHs up to much larger than a Hubble for galaxies with MBHs of mass $\sim 10^9 M_\odot$ [Yu, 2002, Milosavljevic and Merritt, 2003, Merritt and Milosavljevic, 2005, Berczik et al., 2005].

The so-called *last parsec problem* consists in that, due to the long relaxation times of giant, luminous galaxies, the binary MBH may stall at a separation of ~ 1 pc [Milosavljevic and Merritt, 2003]. The time scale for the relativistic merger given by equation (5.43), assuming a relativity low eccentricity $e \lesssim 0.3$, is, for large separations, larger than a Hubble time. Therefore, the MBH binary may not merge. However, there are several possible ways out of this problem (without even considering several other processes which may help the binary's shrinkage such as gaseous drag, for example). First, compact less luminous galaxies which are expected to harbour MBHs of mass $M_\bullet \lesssim 10^6 M_\odot$ have shorter relaxation times of the order few $\times\, 10^9$ years. This may be sufficient to provide the necessary flux of stars into the binary's loss cone so that it can reach the relativistic regime in less than a Hubble time. Second, if the MBH is formed with a high eccentricity $e \gtrsim 0.9$, the time scale for relativistic merger may be drastically reduced as it can be seen from the strong dependence on eccentricity of the equation (5.43). Third, deviations from spherical symmetry in the form of a triaxial nucleus, or of an axisymmetric one, and also rotation of the stellar system all tend to keep the binary's loss cone well populated with stars and therefore lead to sustained higher hardening rates. In some cases, all (or some of) these factors may be present simultaneously; this is indeed the case of the galaxy mergers considered in this work.

In order to make a preliminary exploration of these issues, we have made a set of N-body simulations of galaxy mergers with initial conditions that lead naturally, as the outcome of the merger, to the formation of an eccentric MBH binary. The two galaxies were set at time $t = 0$ in an approximately parabolic orbit. In this case, each galaxy was a Dehnen model with inner slope $\gamma = 1$, a black hole particle with mass $M_\bullet/M_{cl} = 0.025$, and with $N = 62$K stars each. The time evolution of the typical galaxy merger can be seen pictorially in the contour plots of Figure 5.11. The lines are the isodensity contours of the stellar distribution of the galaxies and the heavy black dots represent the two MBH particles. The merger happens quite rapidly: it can be seen from the plots that the parent galaxies have become a single one by the time of the fourth snapshot, in less than half of an orbital period of the parent galaxy pair. The encounter is close to frontal, but not quite. It can also be seen that the offspring galaxy is still far from spherical symmetry by the last snapshot in Figure 5.11. When these models are scaled in such a way that the total binary MBH mass is $6 \times 10^5 M_\odot$, the time span from the first to the sixth snapshot is $\sim 10^8$ years.

Figures 5.12 and 5.13 show the time evolution of the MBH binary semimajor axis $a(t)$ and eccentricity $e(t)$, for the galaxy mergers described in Table 5.1. It can be seen that, in five out of six of the cases, the binary forms with $e \gtrsim 0.6$; three of them form with $e \gtrsim 0.8$ and two form with extremely high eccentricity $e \gtrsim 0.95$. The calculations shown here do not include yet the PN terms in the equations of motion of the binary; if they did, GM1 and GM12 would have — according to Peters's timescale equation (5.43) — gone through a relativistic inspiral and merger by $t \sim 150$ and $t \sim 250$, respectively, if we had scaled the units in such a way that each MBH had mass $M_\bullet \sim (1-5) \times 10^5 M_\odot$. For higher masses, the merger time scales would be even shorter. It should be stressed here that the loss cone depletion and, consequently, the *last parsec problem* does not constitute, in this case, an impediment to the transition from the stellar hardening to the relativistic phase. In fact, the binary MBH hardens and reaches the relativistic phase while the galaxy resulting from the merger is still mildly triaxial and has a very well-defined rotation curve — it is thus very far from the idealized¶ non-rotating, spherically symmetric models for which the classical loss cone theory applies [Cohn and Kulsrud, 1978, Milosavljevic and Merritt, 2003]. In the conditions of these galaxy mergers, in contrast with the spherically symmetric models, the loss cone of the resulting galaxy is never emptied; this is a result of the intrinsic dynamics, not of a low N regime [Milosavljevic and Merritt, 2003, Makino et al., 2004]. It should also be stressed that the creation of triaxiality and rotation after the merger, from two initially spherically symmetric progenitors, is a quite generic outcome [Moore et al., 2004]. Therefore, after a galaxy merger there will be always a time interval of the order of a few dynamical times on which the galaxy will have some degree of triaxiality; it will, presumably, evolve over time onto an axisymmetric shape with some non-negligible rotation — all this enhances the loss cone replenishment and, consequently, the binary's hardening rate [Merritt and Poon, 2004, Berczik et al., 2006].

Following Amaro-Seoane & Freitag (2006), we adopt the Quinlan's plus the Peters's equations of motion to extrapolate the time evolution of the binary MBH from the N-body simulation until they reach the LISA band. The Quinlan's plus Peters's equation are

---

¶The non-rotating, spherically symmetric models are unrealistic after the galaxy merger which generically lead to a rotating triaxial outcome. These models are not unrealistic for galaxies closer to equilibrium configurations; for instance, they constitute a good approximation to the galactic nucleus of our galaxy [Alexander, 2005, Schödel et al., 2007] schematically given by

$$\begin{aligned}\frac{da}{dt} &= \left.\frac{da}{dt}\right|_Q + \left.\frac{da}{dt}\right|_P \\ \frac{de}{dt} &= \left.\frac{de}{dt}\right|_Q + \left.\frac{de}{dt}\right|_P\end{aligned} \tag{5.48}$$

Our main goal is to estimate the eccentricity with which the binary reaches the LISA band, and to estimate the first harmonics of the resulting GW strain amplitude h. We do this as follows: (i) choose initial binary's orbital elements from the N-body output at a time when the relativistic terms were still non-dominant ($\lesssim 1\%$); (ii) evolve the PN binary system under the combined Quinlan + Peter's terms; (iii) feed into the expressions for the GW strain the resulting orbital elements $f(t)$ and $e(t)$, where the orbital frequency f was obtained from the semimajor axis $a(t)$ via the 3^{rd} Kepler's law; (iv) the dimensionless coefficients H and K are estimated from the N-body output.

As discussed in sections 5.1 and 5.2, the dimensionless strain h is the natural quantity used to measure the GW strength. As the binary system undergoes the relativistic inspiral still with a semimajor axis much larger than its Schwarschild radius, its secular energy loss is essentially adiabatic. This means that, in this regime, its orbital frequency changes very slowly and for a very small amount during each orbit. In these circumstances, the GW signal it emmits is almost periodic with frequency f, and the total energy carried out by the wave at that frequency is given by $E \propto N(f)h^2$. $N(f)$ is the number of cycles that it spends a frequency interval $\Delta f \sim f$ around the frequency f. Therefore, the *characteristic strain* h_c, in an observation with duration $\tau \sim N(f)f$, is given by $h_c = \sqrt{N}h$. The binary continuously loses orbital energy, its orbital frequency keeps increasing until, eventually, is only close to the ISCO that the orbital frequency changes at a rate comparable to the frequency itself: $N \sim 1$ and $h_c \sim h$.

We want to estimate $N(f)$: assuming that the back-reaction from the GW emission (invoking the energy balance argument) dominates the binary's orbital decay during the relativistic inspiral, it follows:

$$N \sim \frac{f_r}{\dot{f}_r^2} = \frac{5c^5}{96\pi^{8/3}} G^{-5/3}\mathcal{M}^{-5/3} f_r^{-5/3}, \tag{5.49}$$

where $\mathcal{M} = M_1^{3/5}M_2^{3/5}/(M_1+M_2)^{1/5}$ is the binary's chirp mass, and f_r is the orbital frequency is the source's rest frame. The rest frequency shift rate is given by

$$\dot{f}_r = \frac{df_r}{dt} = \frac{df_r}{dt}\frac{da}{dt} = \frac{96\pi^{8/3}}{5c^5}\mathcal{M}^{5/3} f_r^{11/3}, \tag{5.50}$$

where df_r/da is obtained from the Kepler's 3rd law (Peters equations assume Keplerian orbits, which is a very good approximation for the inspiral phase).

From equation (5.18), and plugging the Keplerian orbit parametrized by $r = a(1 - e^2)/(1 + e\cos\phi)$ and $\dot{\phi} = \sqrt{G(M_1 + M_2)a(1 - e^2)}/r^2$ into the quadrupole tensor (5.20) for the binary's mass distribution, one obtains for the orbit-averaged GW strain amplitude

$$h = \frac{8\pi^{2/3}G^{5/3}\mathcal{M}^{5/3}}{\sqrt{10}c^4R(z)}f_r^{2/3}\sqrt{g(n,e)}, \tag{5.51}$$

where $R(z)$ is the comoving distance from the source and $g(n,e)$ is a factor dependent on the binary's eccentricity e and harmonic n; it is given by

$$\begin{aligned} g(n,e) = \frac{n^4}{34}\Bigg\{ &\left[J_{n-2}(ne) - 2eJ_{n-1}(ne) + \frac{2}{n}J_n(ne) + 2eJ_{n+1}(ne) - J_{n+2}(ne)\right]^2 \\ &+(1-e^2)\left[J_{n-2}(ne) - 2J_n(ne) + J_{n+2}(ne)\right]^2 + \frac{4}{3n^2}J_n^2(ne)\Bigg\}, \end{aligned} \tag{5.52}$$

where the J_n are Bessel functions of the first kind [Peters and Matthews, 1963].

Therefore the characteristic strain h_c is given by

$$\begin{aligned} h_c &= \sqrt{N}\ h \sim \frac{1}{\sqrt{3}\pi^{2/3}}\frac{G^{5/6}\mathcal{M}^{5/6}}{c^{3/2}R(z)}\ \sqrt{g(n,e)}\ f_r^{-1/6}, \qquad N \lesssim f\tau \\ h_c &= \sqrt{f\tau}\ h \sim \frac{8\pi^{2/3}}{\sqrt{10}}\frac{G^{5/3}\mathcal{M}^{5/3}}{c^4R(z)}\sqrt{g(n,e)}\ f_r^{7/6}\ \sqrt{\tau}, \qquad N \gtrsim f\tau, \end{aligned} \tag{5.53}$$

where $f \equiv f_{obs} = f_r/(1+z)$. Note that the number of wave cycles observed by the detector in the frequency interval Δf around f is given by $\sim f\tau$.

The knee frequency, in the source's rest frame, that approximately separates these two regimes is given by $f_{r,knee} = N/\tau$. This results in

$$f_{r,knee} = \frac{1}{\pi}\left(\frac{5}{96\pi}\right)^{3/8}\frac{c^{15/8}}{G^{5/8}\mathcal{M}^{5/8}}. \tag{5.54}$$

In the LISA detector, $f_{obs,knee} = f_{r,knee}/(1+z)$. Summarizing: while the inspiral proceeds in the adiabatic phase, the GW's characteristic strain $h_c \propto f^{7/6}$; once the binary chirp rate becomes fast enough (for frequencies above the knee), $h_c \propto f_{-1/6}$.

The results are shown in Figures (5.14) to (5.17). The dimensionless strain is plotted against the frequency as it would be observed by LISA (taking into account the assumed source's redshift); four redshifts were considered: $z = 1, 2, 3,$ and 4. These were chosen on the basis of the cosmological estimates for the source's redshifts for which a larger number of detections per year are expected in the mass range we are considering, $M_\bullet \sim \mathrm{few} \times 10^5 \mathrm{M}_\odot$ [Sesana et al., 2004, Sesana et al., 2007]. The LISA's sensitivity curve was generated with the Online Sensitivity Curve Generator with the default settings, with the exception that we have choosen a three year observation run [Larson et al., 2000, Larson, 2007].

It can be seen from the Figures that the binaries which form with large eccentricity $e \gtrsim 0.8$ all reach the LISA band $f \sim 10^{-4}$ Hz also with high eccentricity (see Figures (5.18) and (5.19)); as a result, several harmonics contribute significantly to the signal and are above LISA's detection threshold. As the binary chirps in the LISA band, it circularizes quite rapidly and one can also see the accompanying monotonic decrease of the higher harmonics's amplitude until the point where the fundamental frequency ($m = 2$) is dominant. It is only in the case of GM7 — whose MBH binary formed with relatively low eccentricity after the galaxy merger — that only the $m = 2$ mode is above the LISA's detection threshold as the binary enters in its frequency range.

We have thus seen that the last parsec problem may be avoided in the cases where the MBH binary forms with very high eccentricity and that it is expected that such binaries will have gravitational wave signals with several high harmonics above LISA's detection threshold.

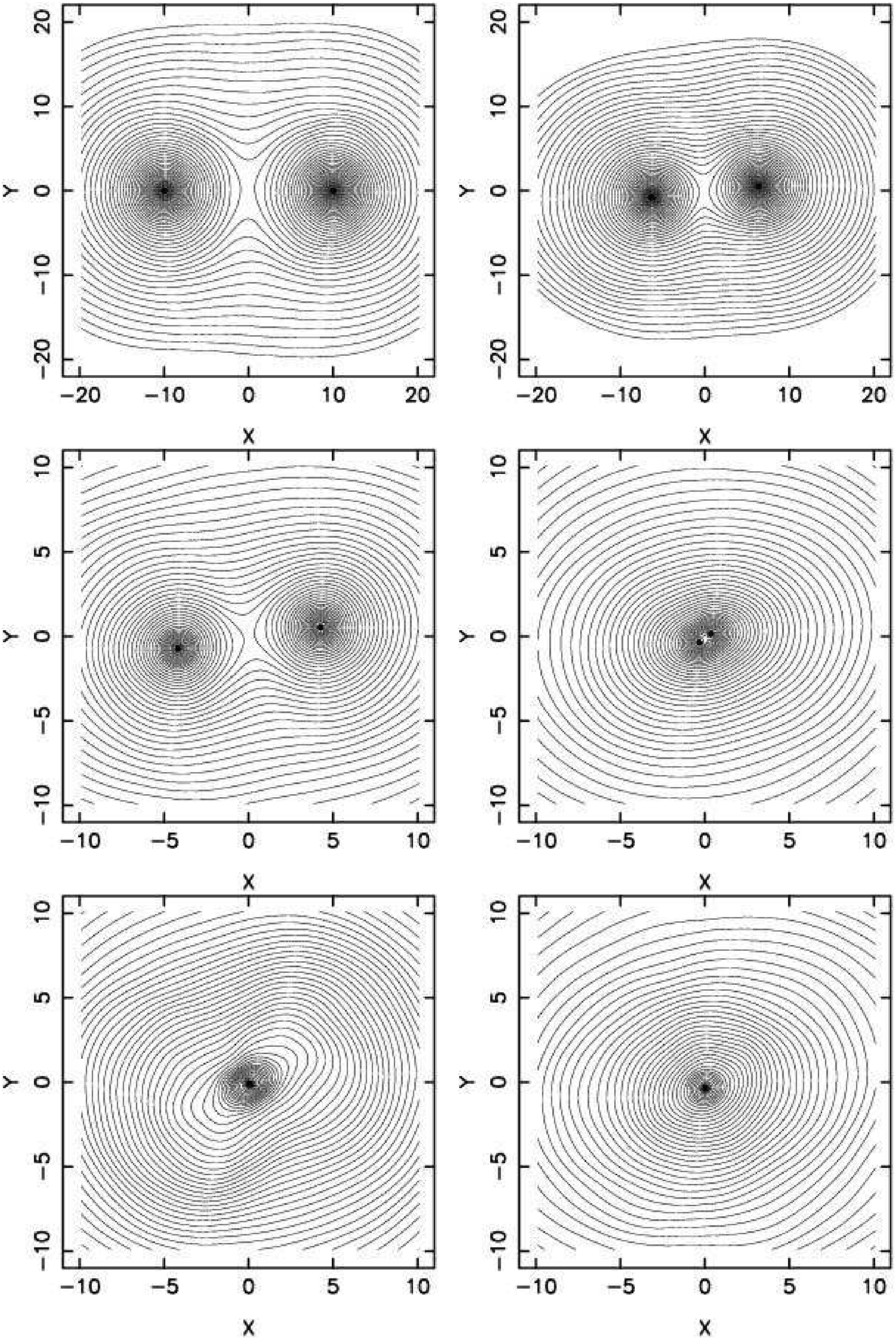

Figure 5.11: Isodensity contours of the stellar mass distribution. The snapshots show the ongoing merger of two initially spherically symmetric Dehnen models, each with a central massive black hole at its center. The galaxies are put on an approximately parabolic orbit and merge quite rapidly in less than half a period of the galaxy pair. After the parent galaxies merge into one, the black holes spiral in to the center where they form a bound pair and harden through three-body slingshot effect. The N-body time units are $t = 0, 50, 60, 70, 80, 200$.

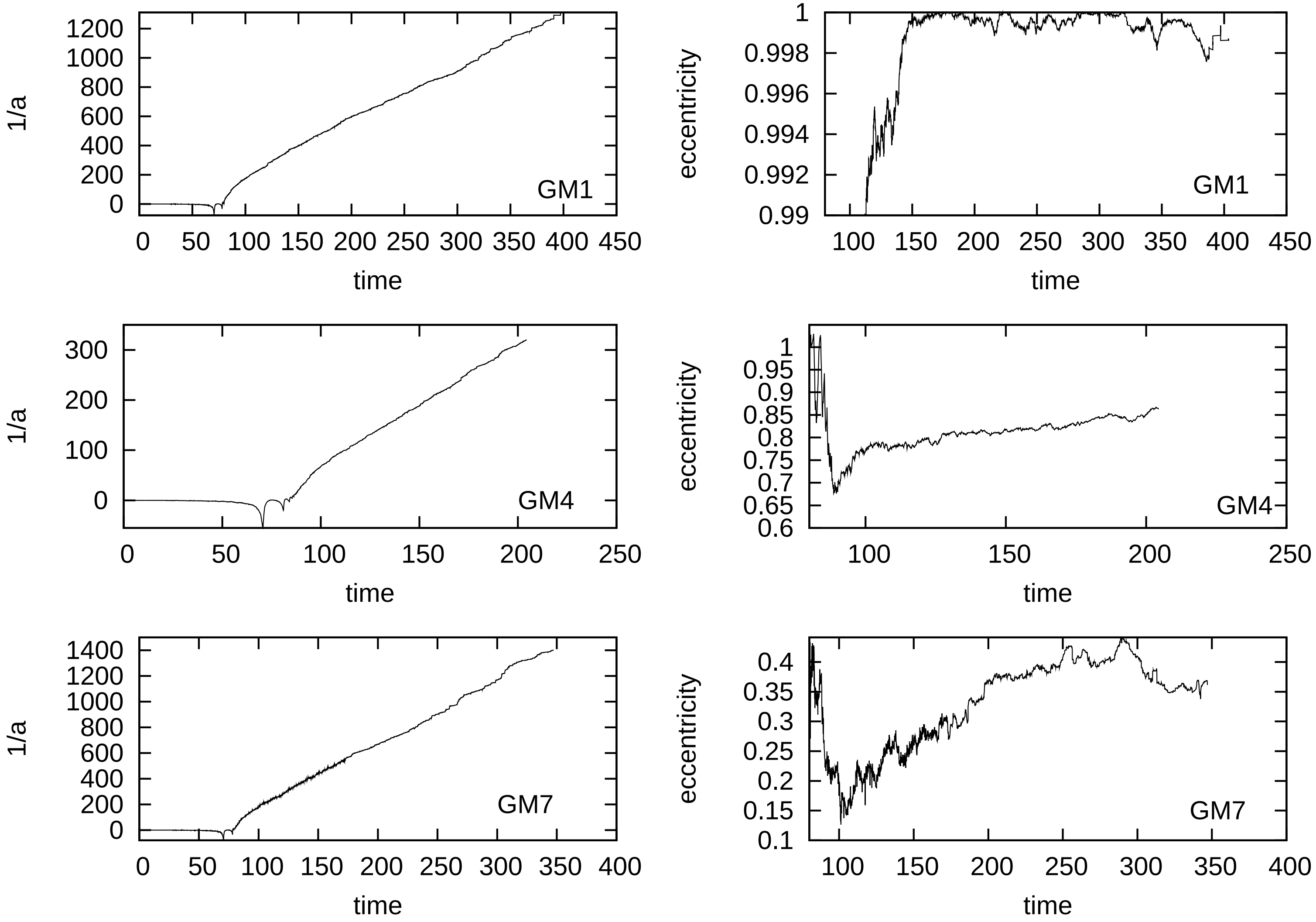

Figure 5.12: Time evolution of the orbital elements $a(t)$ and $e(t)$ of the MBH binary. The binary tends to form with high eccentricity, which tends to slowly grow further over time.

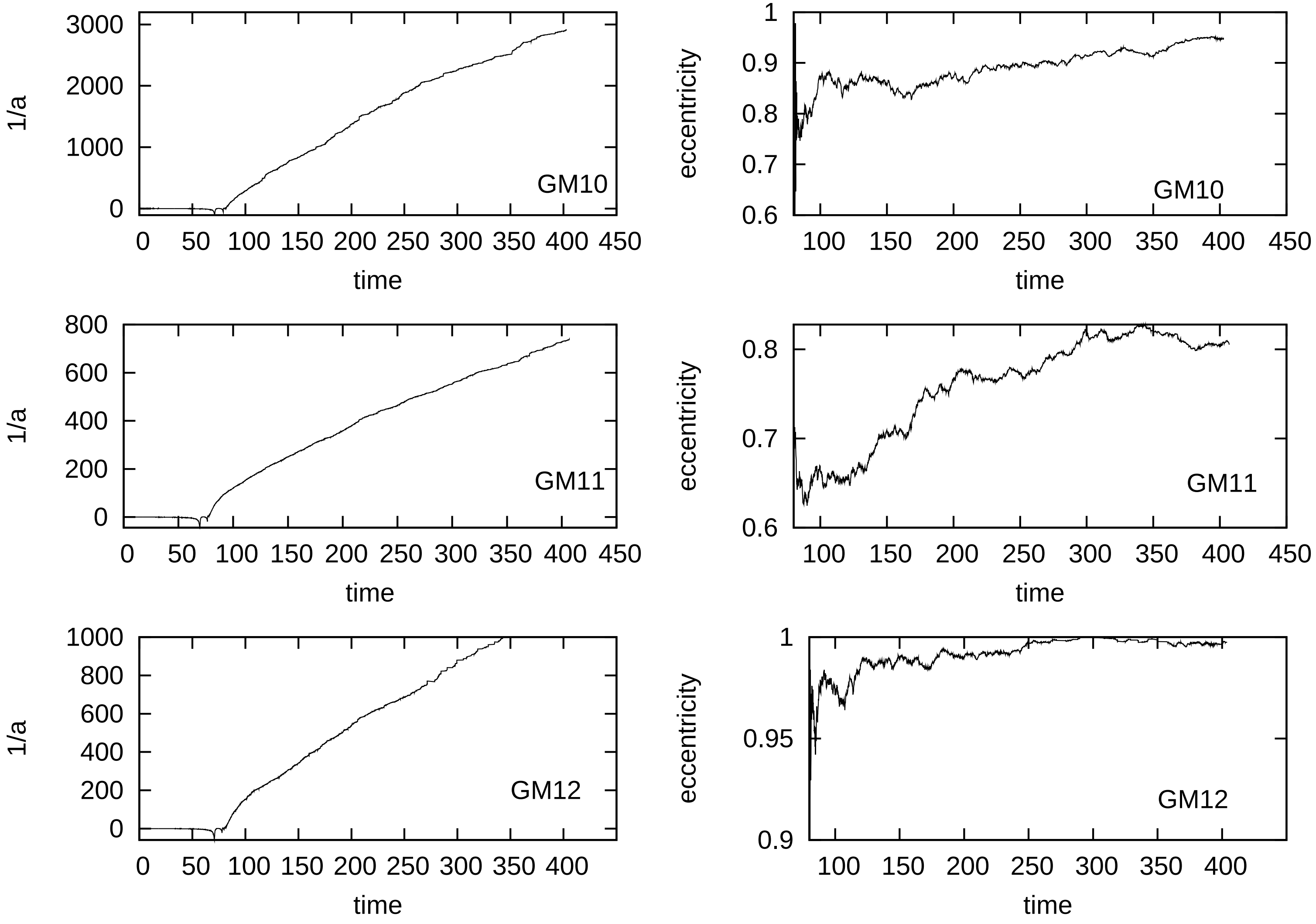

Figure 5.13: Time evolution of the orbital elements $a(t)$ and $e(t)$ of the MBH binary. The binary tends to form with high eccentricity, which tends to slowly grow further over time.

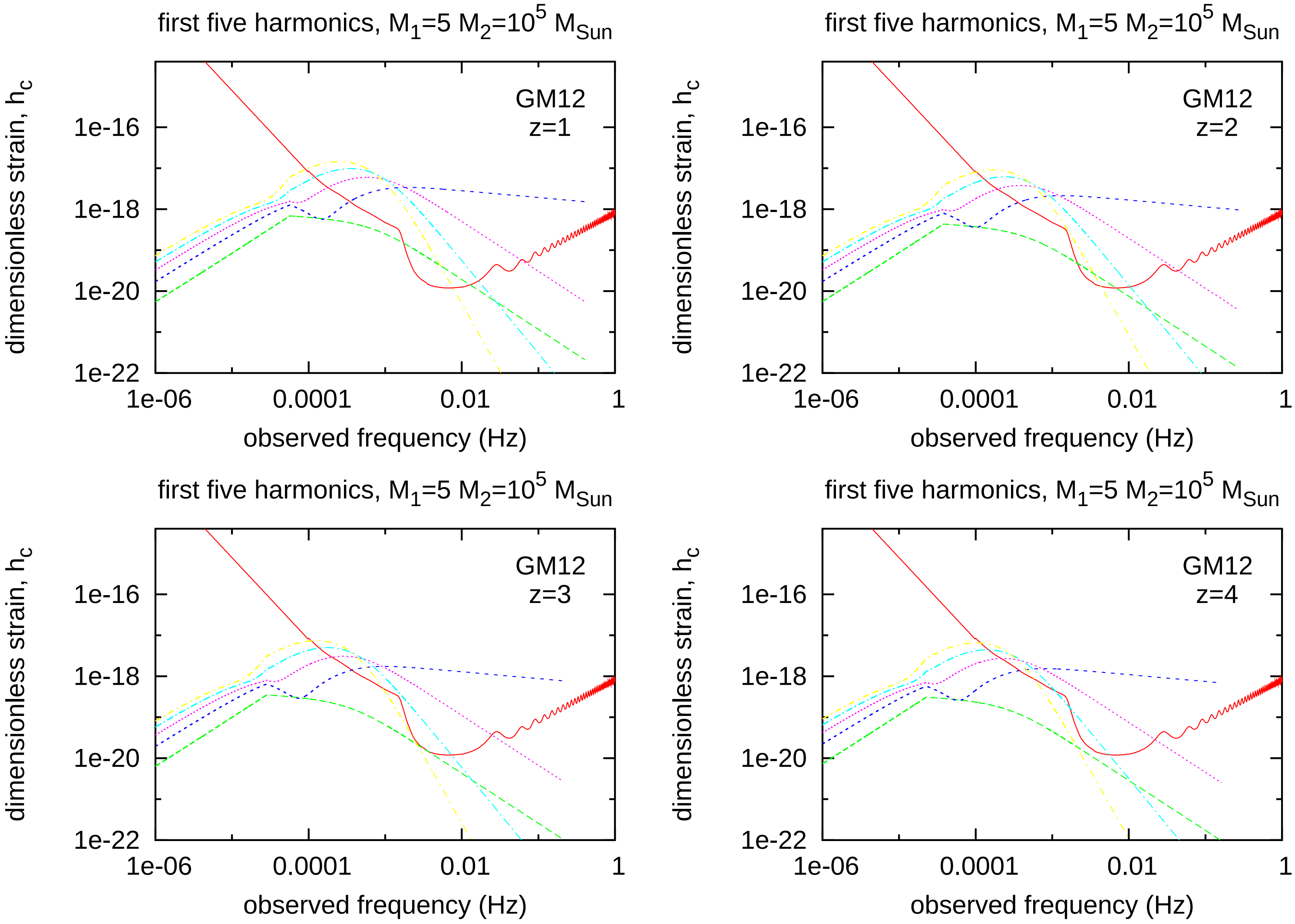

Figure 5.14: First five harmonics of the dimensionless strain amplitude for frequencies in the LISA band. The red curve is the LISA sensitivity curve as obtained from online generated (see text). The green line is the $m = 1$ harmonic, dark blue is $m = 2$, magents is $m = 3$, light blue is $m = 4$ and yellow is $m = 5$. The fundamental frequency corresponding to $m = 2$ becomes dominant as the binary chirps and circularizes its orbit along the LISA band. For all the redshifts considered, several harmonics lie above the detection threshold.

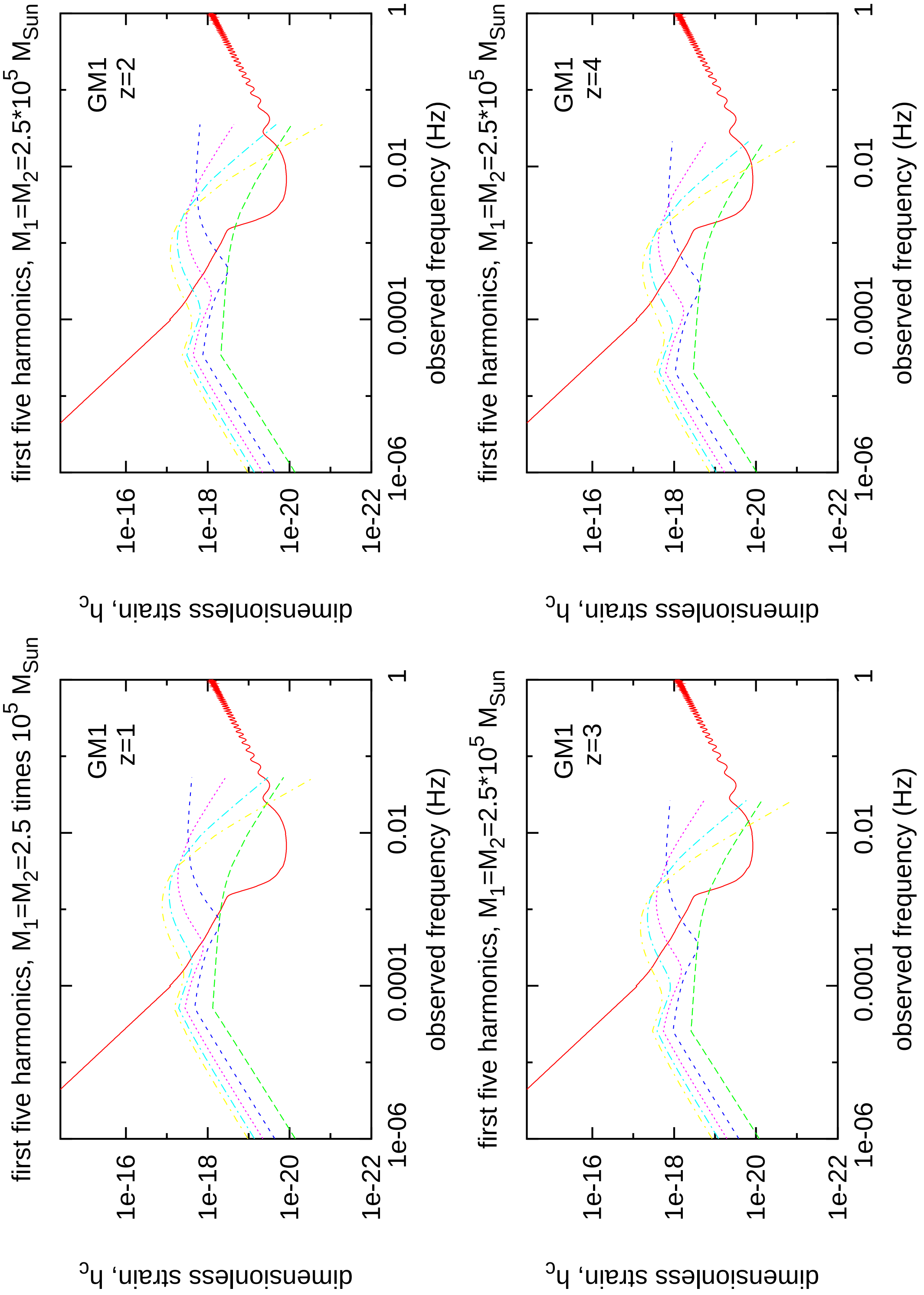

Figure 5.15: First five harmonics of the dimensionless strain amplitude for frequencies in the LISA band. The red curve is the LISA sensitivity curve as obtained from online generated (see text). The green line is the $m = 1$ harmonic, dark blue is $m = 2$, magents is $m = 3$, light blue is $m = 4$ and yellow is $m = 5$. The fundamental frequency corresponding to $m = 2$ becomes dominant as the binary chirps and circularizes its orbit along the LISA band. For all the redshifts considered, several harmonics lie above the detection threshold.

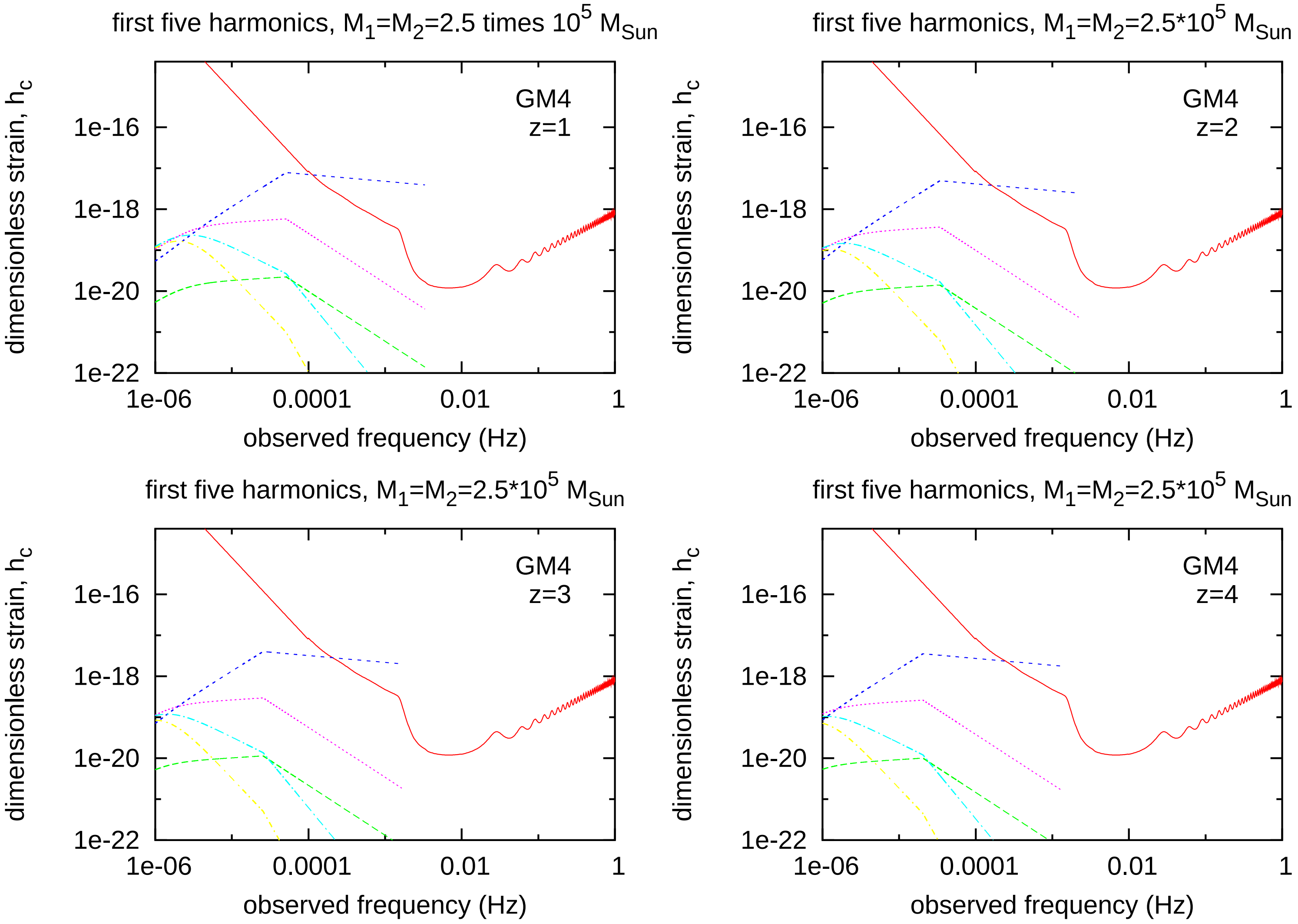

Figure 5.16: First five harmonics of the dimensionless strain amplitude for frequencies in the LISA band. The green line is the $m = 1$ harmonic, dark blue is $m = 2$, magenta is $m = 3$, light blue is $m = 4$ and yellow is $m = 5$. In this case, the binary forms with much lower eccentricity and thus only the fundamental frequency lies above the detection threshold at the lowest redshift considered, $z = 1$.

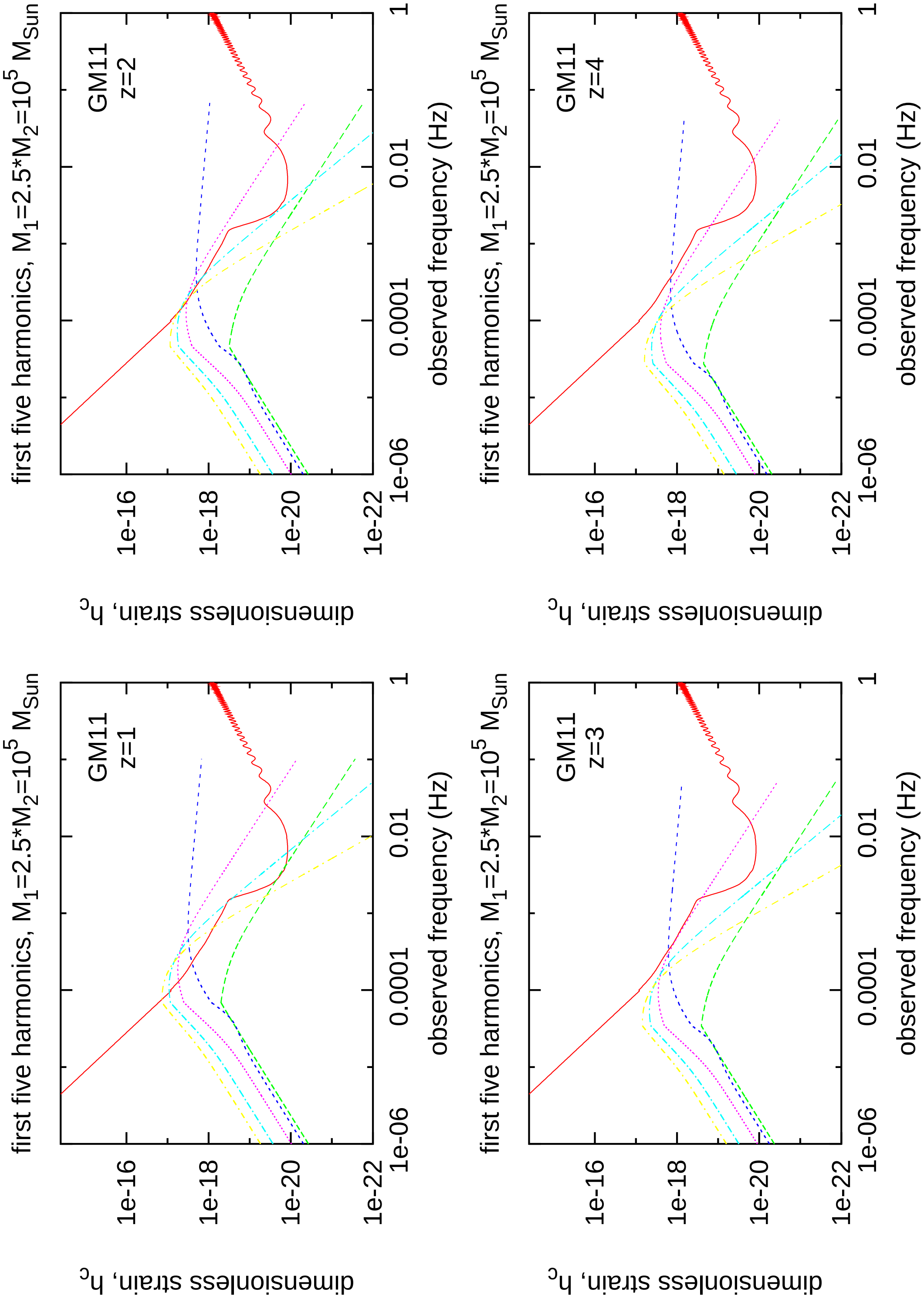

Figure 5.17: First five harmonics of the dimensionless strain amplitude for frequencies in the LISA band. The red curve is the LISA sensitivity curve as obtained from online generated (see text). The green line is the $m = 1$ harmonic, dark blue is $m = 2$, magents is $m = 3$, light blue is $m = 4$ and yellow is $m = 5$. The fundamental frequency corresponding to $m = 2$ becomes dominant as the binary chirps and circularizes its orbit along the LISA band. For all the redshifts considered, several harmonics lie above the detection threshold.

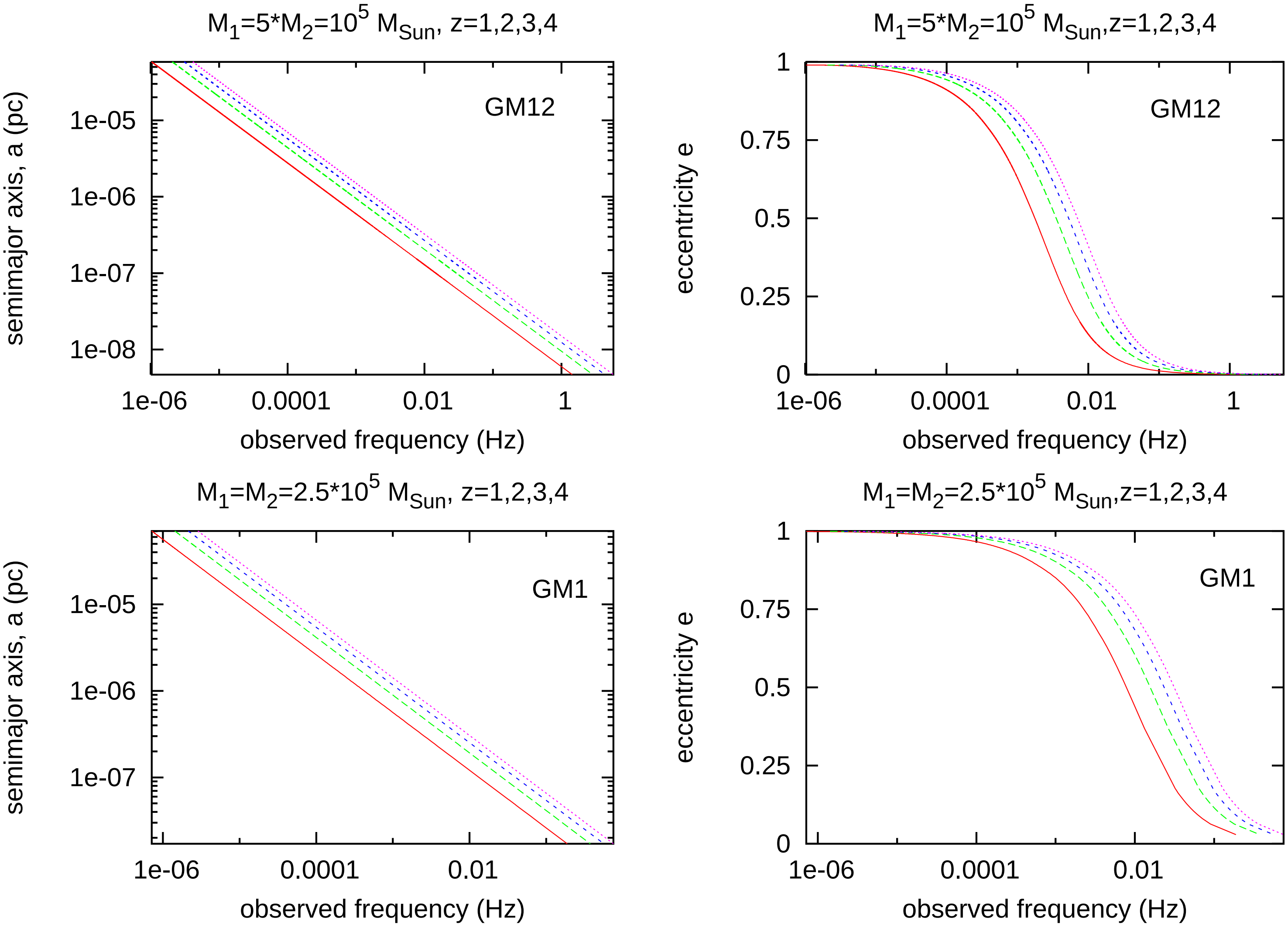

Figure 5.18: Evolution of the binary's orbital elements as it chirps through the LISA frequency band. The orbit is very rapidly circularized and it will have only a residual eccentricity during the last orbits before plunge.

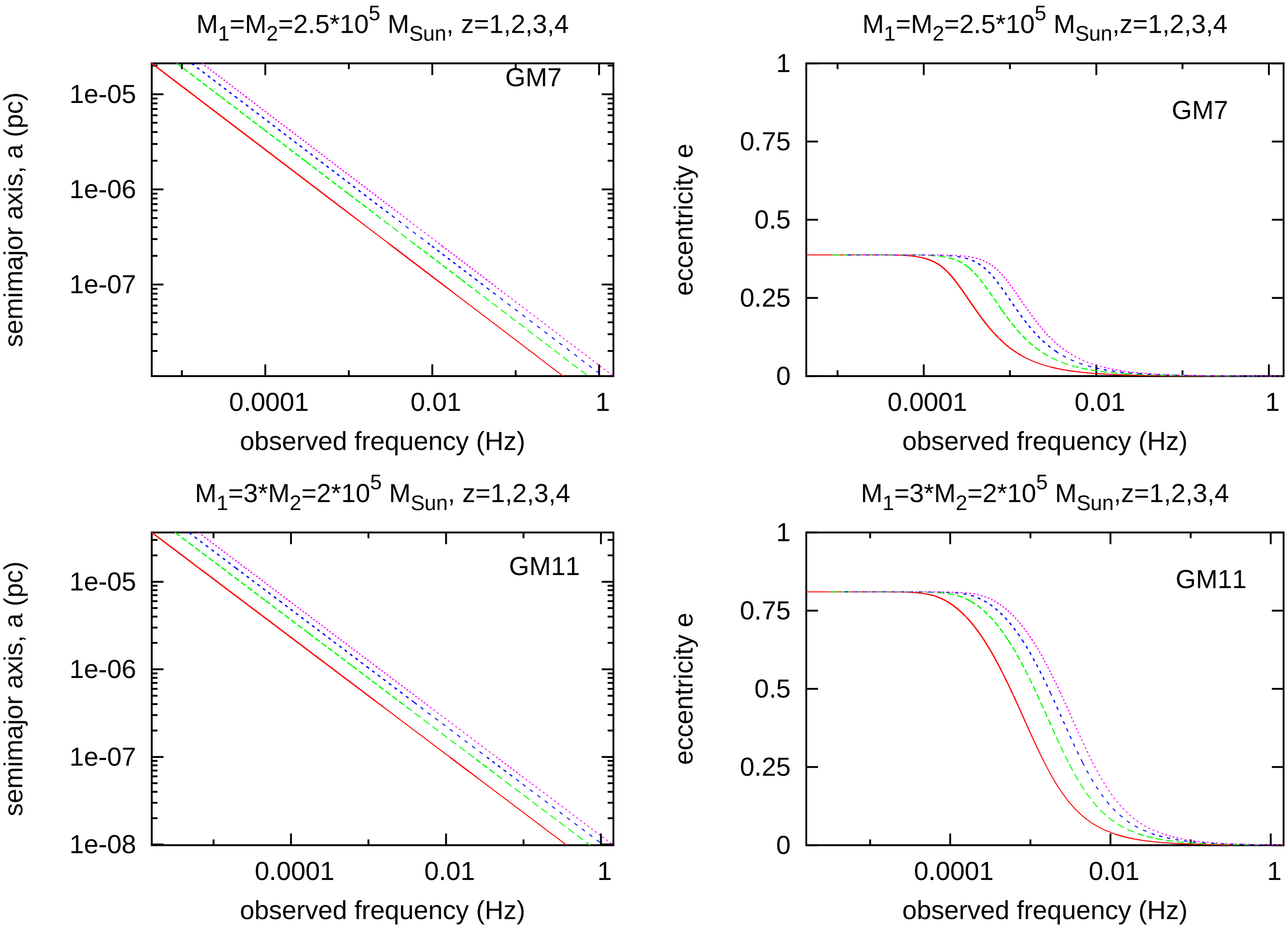

Figure 5.19: Evolution of the binary's orbital elements as it chirps through the LISA frequency band. The orbit is very rapidly circularized and it will have only a residual eccentricity during the last orbits before plunge.

Chapter 6

Discussion

Massive black holes (MBHs) are ubiquituous in galactic nuclei. During the last decade, major breakthroughs were achieved with the detection of a number of MBHs in the nearby galaxies as well as the observational unveiling of a number of structural relations between the MBH and the galactic bulges harboring them — eg. the mass-luminosity $M_\bullet - L$ and the mass-velocity dispersion $M_\bullet - \sigma$ relations. The question of how does the stellar distribution evolves in the presence of an (intermediate) mass black hole (IMBH) was originally addressed in the 1970s following the detection of X-rays sources in globular clusters. The growth of a cusp in the stellar density and velocity dispersion profiles were generic predictions from these now classical studies; the enthusiasm then gradually subsided when it was realized that empirical tests to these predictions would be very hard to come by with the then available telescope resolutions. The more recent observational results have revived the interest in the theoretical study of dense stellar cluster around MBHs.

The recent detections of MBHs — particularly with the Hubble Space Telescope (HST) — in the centers of nearby galaxies with masses $M_\bullet \gtrsim$ few $\times\ 10^6 \mathrm{M}_\odot$ led to a renaissance of interest on the theoretical side of this subject. In fact, the observed stellar density runs with radius near the centers of the nearest galaxies known to contain black holes —- the Milky Way, M31 and M32 — all have a density cusp with $\rho \sim r^{-\gamma}$, $\gamma \approx 1.5$ within the black hole's sphere of influence. This came in marked contrast with the prevailing belief, due to previous observations of lower resolution, that the galactic bulges had a stellar density core with approximately constant density at the center. The presence of a central density cusp around a MBH is consistent with the predictions of more than one model: the so-called adiabatic growth models, the collisional formation of a Bahcall-Wolf cusp. In the former case, a black hole with a small initial mass is initially surrounded by a stellar cluster and its mass is very slowly increased on a time scale longer than the orbital periods but much smaller than the local relaxation time. The final density of such adiabatic growth models is also a power law within the black hole's sphere of influence, but its asymptotic slope may take its value within a range $1.5 \lesssim \gamma \lesssim 2.5$ depending on the model. In contrast, in models of underluminous, compact galactic nuclei for which the local relaxation time is shorter than its age, the old stellar population is expected to relax over a relaxation time scale to a cusp with an asymptotic slope $\gamma = 7/4$, independently of the initial conditions or other details of the model.

For a long time, it has been a serious concern and a source of uncertainty in this field, that all the classic theoretical studies on the collisional cusp dynamics around a MBH have been based on the Fokker-Planck formalism and its inherent simplifying assumptions. These assumptions have never been thoroughly tested in the very specific settings of the very high densities and velocities reached by the stellar distribution around a MBH. Questions have been raised over the years about the validity of the treatment of the collisional dynamics based exclusively on two-body small-angle scattering and the consequent neglect of close encounters. Could this affect the asymptotic steady-state? If it did, it would carry immediately numerous implications for the event rates involving the close interactions between single stars and the MBH at the center (eg. stellar disruptions by the MBH, the slow extreme mass ratio inspirals (EMRIs) of compact remnants onto the MBH which will be one the major sources for LISA, the *paradox of youth* related with the presence of young, massive stars in a region where star formation should be supressed due to the black hole's tidal field, and so on). The interpretation and investigations for all these questions (and others) rest on a theoretical picture for the steady state of the galactic nucleus that needs to be further and completely elucidated. For all these reasons, it is very timely to address the validity of the statistical models of the Fokker-Planck type with a more fundamental approach which assumes nothing more than the validity of the Newton's equations of motion.

It was only recently, with the advent of the special-purpose computer GRAPE-6, that it became possible to do high resolution direct N-body simulations of a galactic nucleus in the presence of a central massive black hole. In this work, we addressed the basic quasi-equilibrium steady state solutions of the stellar distribution around a MBH, and the transition on a relaxation time scale to such steady state with high resolution direct N-body simulations with the help of the GRAPE-6 Titan cluster at the ARI.

We have started by studying the case of a stellar system of single mass stars around a MBH. We compared the results from the direct N-body approach with the predictions of more idealized calculations done using the Fokker-Planck formalism. We have been able to show that the results consistently agree between both approaches and concluded that the growth of a Bahcall-Wolf mass density cusp (both for the mass density profile and in the phase space density) is a robust and generic prediction for the old stellar populations of spheroidal systems older than the relaxation time as measured at the black hole's influence radius. The agreement between both approaches is complete in what concerns the time scales as well. It was the first time that this has been verified with the N-body approach. These results were already published.

In a next step, we have studied in detail the evolution of clusters with two stellar components with a different stellar mass each. The presence of more than one mass component dramatically accelerates the evolution — as compared with the corresponding single mass cluster — due to mass segregation by which the heavy masses drift towards the cluster center whilst the light masses slightly expand outwards. We were able to follow the process of mass segregation in our models, which predict that there should exist a concentration of stellar black holes in the central regions of the most compact, collisional galaxies. Furthermore, we have validated the Bahcall-Wolf prediction for the scaling relations between the asymptotic slopes of the mass density profiles for component with different stellar mass. The agreement is quite impressive and it is the first time that the validity of these scaling relations is unambiguously verified with the N-body approach. We have also verified that, due to the strong mass segregation taking place within the density cusp and also due to a small (but not insignificant) rate of close encounters within the cusp, an overdensity of heavy stars develops at the innermost central regions of the clusters. This is a very exciting prospect as it might constitute in real galaxies a *reservoir* of massive compact remnants at the very center of the density cusp.

In a further step, we have also studied in detail the evolution of clusters with a stellar mass distribution — either in the form of a single power law or a Kroupa mass function. All results observed in the two-component calculations carry on to the multi-mass clusters. In particular, the overdensity at the center of the cluster is also observed in the more realistic multi-mass case and it was not an artifact due to the less realistic nature of the former. We have also checked, via a few test runs with tidal capture of stars by the MBH that all the studied global properties studied in this work are left unchanged and are therefore considered to be robust results.

As a final word concerning this part of the present work, many physical processes present in a galactic nucleus were left out of this study. This is done both out of choice and out of necessity. The models gain much from their simplicity and well-defined nature making it possible a clear interpretation of the results obtained. Moreover, the computational burden posed by these calculation in terms of CPU time is already high enough; furthermore, the algorithmic developments are optimal, at the present time, to deal with the pure stellar dynamics. It will be of uptmost importance to invest a serious effort in order to increase the realism of the models with the inclusion, for instance, of stellar evolution and gas dynamics via SPH in the near future.

The Laser Space Interferometer Antenna (LISA), once it flies and becomes operational, could become one the fundamental drivers of the observational efforts in this field. In fact, the gravitational waves (GWs) emmited by inspiralling MBH binaries or compact remnants inspiralling onto MBH will be the more likely sources of GWs to be detectable by LISA. It is, therefore, very important to understand as much as possible of the interplay of the astrophysical stellar dynamical processes and the dynamical relativistic effects resulting from the the strong interactions that lead to the emmission of GWs.

In this work, we have implemented a numerical formulation of the Post-Newtonian (PN) equations of motion for the integration of relativistic binary systems. The PN equations in the center of mass frame were integrated with the two-body KS regularization method. This implementation was thoroughly tested and shown to be capable of handling the relativistic inspiral motion until the plunge phase. We reach therefore the point where we are capable, from the algorithmic point of view, to deal with the stellar dynamical processes that determine the evolution of the global properties of the galactic nucleus over time scales of the order of the relaxation time while simultaneously following with great accuracy the evolution of the orbital parameters of relativistic binaries — be they massive black holes or stellar mass black holes.

With N-Body simulations of galaxy mergers, we have demonstrated in this work that binary MBHs can be formed with very high eccentricities after the parent galaxies settle into a newly merged object. Since highly eccentric binaries see their relativistic merging time scale strongly reduced, we conclude that it is possible to follow the whole evolution starting before the merger and until the last MBH binary orbits before plunge, even for galaxy models scaled so that the MBH have masses in the range $M_\bullet \sim 10^5 - 10^6 M_\odot$. The last parsec problem — the possibility that the binary MBH shrinkage, due to the three-body slingshot interactions with the field stars, may stall before it reaches a separation small enough that the relativistic effects are capable of driving it to a merger in less than a Hubble time — is, at least under the circumstances thus defined, solved by purely stellar dynamics processes.

It is precisely for this mass range $M_\bullet \sim 10^5 - 10^6 M_\odot$ that LISA's sensitivity will better tuned. We have shown that these eccentric binaries should enter the LISA band with an appreciable eccentricity and so higher harmonics of the wave strain amplitude should be above its detection threshold and contribute to the signal. This leads to very exciting prospects for the modelling of the waveform signals and the extraction of parameters of the sources therefrom.

Bibliography

[Aarseth, 1999] Aarseth, S. (1999). *Publ. Astron. Soc. Japan*, 111:1333.

[Aarseth, 2003] Aarseth, S. (2003). *Gravitational N-Body Simulations.* Cambridge University Press.

[Aarseth and Zare, 1974] Aarseth, S. and Zare, K. (1974). *Cel. Mech. Dyn. Astron.*, 10:185.

[Alexander, 2005] Alexander, T. (2005). *Phys. Rep.*, 419:65.

[Amaro-Seoane and Freitag, 2006] Amaro-Seoane, P. and Freitag, M. (2006). *Astrophys. J. Lett.*, 635:53.

[Amaro-Seoane et al., 2004] Amaro-Seoane, P., Freitag, M., and Spurzem, R. (2004). *Mon. Not. Royal Astron. Soc.*, 352:655.

[Amaro-Seoane et al., 2006] Amaro-Seoane, P., Gair, J., Freitag, M., Miller, M., Mandel, I., Cutler, C., and Babak, S. (2006). Astrophysics, detection and science applications of intermediate- and extreme mass-ratio inspirals. astro-ph/0703495.

[Arun et al., 2007] Arun, K., Iyer, B., Sathyaprakash, B., and Sinha, S. (2007). gr-qc/07041086.

[Bahcall and Wolf, 1976] Bahcall, J. and Wolf, R. (1976). *Astrophys. J.*, 209:214.

[Bahcall and Wolf, 1977] Bahcall, J. and Wolf, R. (1977). *Astrophys. J.*, 216:883.

[Baumgardt et al., 2004a] Baumgardt, H., Makino, J., and Ebisuzaki, T. (2004a). *Astrophys. J.*, 613:1133.

[Baumgardt et al., 2004b] Baumgardt, H., Makino, J., and Ebisuzaki, T. (2004b). *Astrophys. J.*, 613:1143.

[Begelman et al., 1982] Begelman, M., Blandford, R., and Rees, M. (1982). *Nature*, 287:307.

[Begelman et al., 2006] Begelman, M., Volonteri, M., and Rees, M. (2006). *Mon. Not. Royal Astron. Soc.*, 370:289.

[Berczik et al., 2005] Berczik, P., D, M., and Spurzem, R. (2005). *Astrophys. J.*, 633:680.

[Berczik et al., 2006] Berczik, P., D, M., Spurzem, R., and Bischop, H. (2006). *Astrophys. J.*, 642:21.

[Berti, 2006] Berti, E. (2006). *Classical and Quantum Gravity*, 23:785.

[Bettwieser and Sugimoto, 1984] Bettwieser, E. and Sugimoto, D. (1984). *Mon. Not. Royal Astron. Soc.*, 208:493.

[Binney and Merrifield, 1998] Binney, J. and Merrifield, M. (1998). *Galactic Astronomy.* Princeton University Press.

[Binney and Tremaine, 1987] Binney, J. and Tremaine, S. (1987). *Galactic Dynamics.* Princeton University Press.

[Blanchet, 2006] Blanchet, L. (2006). Gravitational radiation from post-newtonian sources and inspiralling compact binaries. *Living Rev. Relativity*, 9(4).

[Blanchet and Iyer, 2003] Blanchet, L. and Iyer, B. R. (2003). *Classical and Quantum Gravity*, 20(4):755.

[Blanchet and Schäfer, 1989] Blanchet, L. and Schäfer, G. (1989). *Mon. Not. Royal Astron. Soc.*, 239:845.

[Blandford and Begelman, 1999] Blandford, R. and Begelman, M. (1999). *Mon. Not. Royal Astron. Soc.*, 303:L1.

[Blandford and Mckee, 1982] Blandford, R. and Mckee, C. (1982). *Astrophys. J.*, 255:419.

[Boccaletti and Pucacco, 1996] Boccaletti, D. and Pucacco, G. (1996). *Theory of Orbits I.* A&A Library, Springer.

[Chandrasekar, 1943] Chandrasekar, S. (1943). Stochastic problems in physics and astronomy. *Rev. Mod. Phys.*, 15(1):1.

[Chatterjee et al., 2002] Chatterjee, P., Hernquist, L., and Loeb, A. (2002). *Phys. Rev. Lett.*, 88:121103.

[Chernoff and Weinberg, 1990] Chernoff, D. and Weinberg, M. (1990). *Astrophys. J.*, 351:121.

[Cohn, 1979] Cohn, H. (1979). *Astrophys. J.*, 234:1036.

[Cohn et al., 1989] Cohn, H., Hut, P., and Wise, M. (1989). *Astrophys. J.*, 342:814.

[Cohn and Kulsrud, 1978] Cohn, H. and Kulsrud, R. (1978). *Astrophys. J.*, 226:1087.

[Damour, 1982] Damour, T. (1982). *Gravitational Radiation*, chapter Gravitational radiation and the motion of compact bodies, page 59. North Holland.

[Damour and Deruelle, 1981] Damour, T. and Deruelle, N. (1981). *Phys. Lett.*, 87A:81.

[Damour and Deruelle, 1985] Damour, T. and Deruelle, N. (1985). *Ann. Inst. Henri Poincaré*, 43:107.

[Damour and Schäfer, 1988] Damour, T. and Schäfer, G. (1988). *Nuovo Cimento*, 101:127.

[Dehnen, 1993] Dehnen, W. (1993). *Mon. Not. Royal Astron. Soc.*, 265:250.

[Deuflhard, 1985] Deuflhard, P. (1985). *SIAM Review*, 27:505.

[Diacu and Holmes, 1996] Diacu, F. and Holmes, P. (1996). *Celestial Encounters.* Princeton Un. Press, Princeton.

[Einsel and Spurzem, 1999] Einsel, C. and Spurzem, R. (1999). *Mon. Not. Royal Astron. Soc.*, 302:81.

[Fan and *et al.* (SDSS Collaboration), 2001] Fan, X. and *et al.* (SDSS Collaboration) (2001). *Astron. J.*, 121:54.

[Ferrarese, 2002] Ferrarese, L. (2002). *Current High-Energy Emission Around Black Holes*, page 3. World Scientific Plublishing, Singapore.

[Ferrarese and Ford, 2005] Ferrarese, L. and Ford, H. (2005). *Space Sc. Rev.*, 116:523.

[Ferrarese and Merritt, 2000] Ferrarese, L. and Merritt, D. (2000). *Astrophys. J. Lett.*, 539:9.

[Fiestas et al., 2006] Fiestas, J., Spurzem, R., and Kim, E. (2006). *Mon. Not. Royal Astron. Soc.*, 373:677.

[Flannagan and Hughes, 2006] Flannagan, E. and Hughes, S. (2006). The basic of gravitational wave theory. gr-qc/0501041.

[Freitag et al., 2006] Freitag, M., Amaro-Seoane, P., and Kalogera, V. (2006). *Astrophys. J.*, 649:91.

[Freitag and Benz, 2002] Freitag, M. and Benz, W. (2002). *Astron. & Astroph.*, 394:345.

[Fukushige et al., 2005] Fukushige, T., Makino, J., and Kawai, A. (2005). *Publ. Astron. Soc. Japan*, 57:1009.

[Gebhardt et al., 2000] Gebhardt, K., Bender, R., Bower, G., Dressler, A., and Faber, S. (2000). *Astrophys. J. Lett.*, 539:13.

[Genzel et al., 2003a] Genzel, R., Schödel, R., Ott, T., Eckart, A., Mouawad, N., and Alexander, T. (2003a). *Astrophys. J.*, 594:812.

[Genzel et al., 2003b] Genzel, R., Schödel, R., Ott, T., and Eisenhauer, F. (2003b). *Nature*, 425:934.

[Ghez et al., 2005] Ghez, A., Salim, S., Hornstein, S., and Tanner, A. (2005). *Astrophys. J.*, 620:744.

[Gilbert, 1968] Gilbert, I. (1968). *Astrophys. J.*, 152:1043.

[Goodman, 1983] Goodman, J. (1983). *Astrophys. J.*, 270:700.

[Goodman, 1984] Goodman, J. (1984). *Astrophys. J.*, 280:298.

[Goodman, 1987] Goodman, J. (1987). *Astrophys. J.*, 313:576.

[Goodman et al., 1993] Goodman, J., Heggie, D., and Hut, P. (1993). *Astrophys. J.*, 415:715.

[Guckenheimer and Holmes, 1983] Guckenheimer, J. and Holmes, P. (1983). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer-Verlag, NY.

[Hemsendorf and Merritt, 2002] Hemsendorf, M. and Merritt, D. (2002). *Astrophys. J.*, 580:606.

[Hénon, 1969] Hénon, M. (1969). *Astron. & Astroph.*, 2:151.

[Hénon, 1975] Hénon, M. (1975). *Dynamics of stellar systems: Proceedings from IAU Symposium no. 69 held in Besancon, France, Septembe r 9-13, 1974*, chapter Two Recent Developments Concerning the Monte Carlo Method. Reidel Pub.

[Hopman and Alexander, 2006] Hopman, C. and Alexander, T. (2006). *Astrophys. J.*, 645:1152.

[J.Stoer and Bulirsch, 1993] J.Stoer and Bulirsch, R. (1993). *Introduction to Numerical Analysis.* Texts in Applied Mathematics. Springer, Berlin.

[Kauffmann and Haehnelt, 2000] Kauffmann, G. and Haehnelt, M. (2000). *Mon. Not. Royal Astron. Soc.*, 311:576.

[Koushiappas et al., 2004] Koushiappas, S., Bullock, J., and Dekel, A. (2004). *Mon. Not. Royal Astron. Soc.*, 354:292.

[Krolik, 1999] Krolik, J. (1999). *Active Galactic Nuclei.* Princeton University Press.

[Kupi et al., 2006] Kupi, G., Amaro-Seoane, P., and Spurzem, R. (2006). *Mon. Not. Royal Astron. Soc.*, 371:L45.

[Landau and Lifshitz, 1972] Landau, L. and Lifshitz, E. (1972). *Teoria do campo.* Editora Mir, Moscovo.

[Larson, 2007] Larson, S. (2007). Lisa: A modern astrophysical laboratory. in LISC, the LISA International Science Community http://www.lisa-science.org/resources/introductory-lisa-material.

[Larson et al., 2000] Larson, S., Hiscock, W., and Hellings, R. W. (2000). *Phys. Rev. D*, 62:062001.

[Lauer et al., 1998] Lauer, T., Faber, S., Ajhar, E., Grillmair, C., and Scowen, P. (1998). *Astron. J.*, 116:2263.

[Lee and Goodman, 1989] Lee, M. and Goodman, J. (1989). *Astrophys. J.*, 343:594.

[Lightman and Shapiro, 1977] Lightman, A. and Shapiro, S. (1977). *Astrophys. J.*, 211:244.

[Louis and Spurzem, 1991] Louis, P. and Spurzem, R. (1991). *Mon. Not. Royal Astron. Soc.*, 251:408.

[Lynden-Bell and Eggleton, 1980] Lynden-Bell, D. and Eggleton, P. (1980). *Mon. Not. Royal Astron. Soc.*, 191:483.

[Lynden-Bell and Wood, 1968] Lynden-Bell, D. and Wood, R. (1968). *Mon. Not. Royal Astron. Soc.*, 138:495.

[Madau and Rees, 2001] Madau, P. and Rees, M. (2001). *Astrophys. J. Lett.*, 551:27.

[Makino et al., 2004] Makino, J., , and Funato, Y. (2004). *Astrophys. J.*, 602:93.

[Makino et al., 2003] Makino, J., Fukushige, T., Koga, M., and Namura, K. (2003). *Publ. Astron. Soc. Japan*, 55:1163.

[Makino and Taiji, 1998] Makino, J. and Taiji, M. (1998). *Scientific Simulations with Special-Purpose Computers — the GRAPE systems.* Wiley.

[Maoz, 1993] Maoz, E. (1993). *Mon. Not. Royal Astron. Soc.*, 263:75.

[Maoz, 1998] Maoz, E. (1998). *Astrophys. J. Lett.*, 494:181.

[Marchant and Shapiro, 1980] Marchant, A. and Shapiro, S. (1980). *Astrophys. J.*, 239:685.

[Marconi and Hunt, 2003] Marconi, A. and Hunt, L. (2003). *Astrophys. J. Lett.*, 589:21.

[Memmesheimer et al., 2004] Memmesheimer, R., Gopakumar, A., and Schäfer, G. (2004). *Phys. Rev. D*, 70:104011.

[Merritt, 2001] Merritt, D. (2001). *Astrophys. J.*, 556:245.

[Merritt, 2006] Merritt, D. (2006). Dynamics of galaxy cores and supermassive black holes. astro-ph/0605070.

[Merritt et al., 2007] Merritt, D., Berczik, P., and Laun, F. (2007). *Astrophys. J.*, 133:553.

[Merritt and Ferrarese, 2001] Merritt, D. and Ferrarese, L. (2001). *Mon. Not. Royal Astron. Soc.*, 320:30.

[Merritt and Milosavljevic, 2005] Merritt, D. and Milosavljevic, M. (2005). Massive black hole binary evolution. *Living Rev. Relativity*, 8:8.

[Merritt and Poon, 2004] Merritt, D. and Poon, M. (2004). *Astrophys. J.*, 606:788.

[Merritt and Szell, 2006] Merritt, D. and Szell, A. (2006). *Astrophys. J.*, 648:890.

[Merritt and Tremblay, 1994] Merritt, D. and Tremblay, B. (1994). *Astron. J.*, 108:514.

[Mikkola, 2006] Mikkola, S. (2006). private communication.

[Mikkola and Aarseth, 1990] Mikkola, S. and Aarseth, S. (1990). *Cel. Mech. Dyn. Astron.*, 47:375.

[Mikkola and Aarseth, 1993] Mikkola, S. and Aarseth, S. (1993). *Cel. Mech. Dyn. Astron.*, 57:439.

[Mikkola and Aarseth, 2002] Mikkola, S. and Aarseth, S. (2002). *Cel. Mech. Dyn. Astron.*, 84:343.

[Miller and Scalo, 1979] Miller, G. and Scalo, J. (1979). *Astrophys. J. Suppl.*, 41:513.

[Miller and Colbert, 2004] Miller, M. and Colbert, E. (2004). *International Journ. Mod. Phys. D*, 13:1.

[Miller, 1964] Miller, R. (1964). *Astrophys. J.*, 140:250.

[Milosavljevic and Merritt, 2003] Milosavljevic, M. and Merritt, D. (2003). *Astrophys. J.*, 596:860.

[Miralda-Escudé and Gould, 2000] Miralda-Escudé, J. and Gould, A. (2000). *Astrophys. J.*, 545:847.

[Misner et al., 1972] Misner, C., Thorne, K., and Wheeler, J. (1972). *Gravitation*. Freeman.

[Moore et al., 2004] Moore, B., Kazantzidis, S., Diemand, J., and Stadel, J. (2004). *Mon. Not. Royal Astron. Soc.*, 354:522.

[Murphy et al., 1991] Murphy, B., Cohn, H., and Durisen, R. (1991). *Astrophys. J.*, 370:60.

[Narayan, 2003] Narayan, R. (2003). *Lighthouses of the Universe*, chapter Why Do AGN Lighthouses Switch Off? Springer-Verlag.

[Peebles, 1972] Peebles, P. (1972). *Astrophys. J.*, 178:371.

[Peters, 1964] Peters, P. (1964). *Phys. Rev.*, 136:1224.

[Peters and Matthews, 1963] Peters, P. and Matthews, J. (1963). *Phys. Rev.*, 131:435.

[Preto et al., 2004] Preto, M., Merritt, D., and Spurzem, R. (2004). *Astrophys. J. Lett.*, 613:L109.

[Preto and Tremaine, 1999] Preto, M. and Tremaine, S. (1999). *Astron. J.*, 118:2532.

[Quinlan, 1996] Quinlan, G. (1996). *New Astron.*, 1:35.

[Quinlan et al., 1995] Quinlan, G., Hernquist, L., and Sigurdsson, S. (1995). *Astrophys. J.*, 440:554.

[Quinlan and Tremaine, 1992] Quinlan, G. and Tremaine, S. (1992). *Mon. Not. Royal Astron. Soc.*, 259:505.

[Rees, 1984] Rees, M. (1984). *Ann. Rev. Astron. Astroph.*, 22:471.

[Reid and Brunthaler, 2004] Reid, M. and Brunthaler, A. (2004). *Astrophys. J.*, 616:872.

[Rosenbluth et al., 1957] Rosenbluth, M., MacDonald, W., and Judd, D. (1957). *Phys. Rev.*, 107:1.

[Salpeter, 1955] Salpeter, E. (1955). *Astrophys. J.*, 121:161.

[Schäfer and Wex, 1993] Schäfer, G. and Wex, N. (1993). *Phys. Lett. A*, 174:196.

[Schmidt, 1963] Schmidt, M. (1963). *Nature*, 197:1040.

[Schödel et al., 2007] Schödel, R., Eckart, A., Alexander, T., Merritt, D., Genzel, R., Sternberg, A., Meyer, L., Kull, F., Moultaka, J., Ott, T., and .Straubmeier, C. (2007). The structure of the nuclear stellar cluster of the milky way.

[Schutz, 1985] Schutz, B. (1985). *A first course in general relativity.* Cambridge University Press.

[Sesana et al., 2004] Sesana, A., Haardt, F., Madau, P., and Volonteri, M. (2004). *Astrophys. J.*, 611:623.

[Sesana et al., 2005] Sesana, A., Haardt, F., Madau, P., and Volonteri, M. (2005). *Astrophys. J.*, 623:23.

[Sesana et al., 2007] Sesana, A., Volonteri, M., and Haardt, F. (2007).

[Silverman, 1986] Silverman, B. (1986). *Density estimation for statistics and data analysis.* Chapman and Hall.

[Spitzer, 1987] Spitzer, L. (1987). *Dynamical evolution of globular clusters.* Princeton University Press.

[Spitzer and Hart, 1971] Spitzer, L. and Hart, M. (1971). *Astrophys. J.*, 164:399.

[Stiefel and Scheifele, 1971] Stiefel, E. and Scheifele, G. (1971). *Linear and Regular Celestial Mechanics.* Springer, Berlin.

[Szebehely and Peters, 1967] Szebehely, V. and Peters, C. (1967). *Astron. J.*, 72:876.

[Tremaine, 2005] Tremaine, S. (2005). *Astrophys. J.*, 625:143.

[Tremaine et al., 1994] Tremaine, S., Richstone, D., Byun, Y., Dressler, A., Faber, S., Grillmair, C., Kormendy, J., and Lauer, T. (1994). *Astron. J.*, 107:634.

[van Kampen, 1992] van Kampen, N. (1992). *Stochastic Processes in Physics and Chemistry.* North-Holland, NY.

[Volonteri et al., 2003a] Volonteri, M., Haardt, F., and Madau, P. (2003a). *Astrophys. J.*, 582:599.

[Volonteri et al., 2003b] Volonteri, M., Madau, P., and Haardt, F. (2003b). *Astrophys. J.*, 593:661.

[Weinberg, 1994] Weinberg, M. (1994). *Astrophys. J.*, 421:481.

[Wyithe and Loeb, 2003] Wyithe, J. and Loeb, A. (2003). *Astrophys. J.*, 590:691.

[Young, 1980] Young, P. (1980). *Astrophys. J.*, 242:1232.

[Yu, 2002] Yu, Q. (2002). *Mon. Not. Royal Astron. Soc.*, 331:935.

[Yu and Tremaine, 2002] Yu, Q. and Tremaine, S. (2002). *Mon. Not. Royal Astron. Soc.*, 335:965.