A Fokker-Planck Study of Dense Rotating Stellar Clusters

A dissertation presented

by

John Andrew Girash

to
The Department of Physics
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Physics

Harvard University
Cambridge, Massachusetts

November 2009

Dissertation Advisor: **Professor George B. Field** Author: **John Andrew Girash**

A Fokker-Planck Study of Dense Rotating Stellar Clusters

Abstract

The dynamical evolution of dense stellar systems is simulated using a two-dimensional Fokker-Planck method, with the goal of providing a model for the formation of supermassive stars which could serve as seed objects for the supermassive black holes of quasars. This work follows and expands on earlier one-dimensional studies of spherical clusters of main-sequence stars. The two-dimensional approach allows for the study of rotating systems, as would be expected due to cosmological tidal torquing; other physical effects included are collisional mergers of individual stars and a bulk stellar bar perturbation in the system's gravitational potential. The 3 Myr main-sequence lifetime for large stars provides an upper limit on the allowed simulation times. Two general classes of initial systems are studied: Plummer spheres, which represent stellar clusters, and "$\gamma = 0$" spheres, which model galactic spheroids.

At the initial densities of the modeled systems, mass segregation and runaway stellar collisions alone are insufficient to induce core collapse within the main-sequence lifetime limit, if no bar perturbation is included. However, core collapse is not a requirement for the formation of a massive object: the choice of stellar initial mass function (IMF) is found to play a crucial role. When using an IMF similar to that observed for dense stellar clusters (weighted towards high masses but with a high-mass cutoff of $M_{\max} \lesssim 150 M_\odot$) the simulations presented here show, in all cases, that the stellar system forms massive ($250 M_\odot$) objects by collisional mergers of lower-mass stars; in almost all such cases the presence of a stellar bar allows for sufficient additional outward transport of angular momentum that a core-collapse state is reached with corresponding further increase in the rate of formation of massive objects. In contrast, simulations using an IMF similar to that observed for field stars in general (which is weighted more towards lower masses) produce no massive objects, and reach core collapse only for initial models which represent the highest-density galactic spheriods.

Possible extensions of the work presented here include continuing to track stellar populations after they evolve off the main sequence, and allowing for a (possibly changing) nonspherical component to the overall system potential.

Contents

Acknowledgements

The full list of people to whom I am thankful for helping over the course of this project is too long to mention completely, so first off I would like to express my great appreciation to everyone in the Department of Physics and at the Center of Astrophysics while I was there. Of individual thanks, first and foremost I'm grateful for my thesis advisor George Field; ever since we started this project I've been amazed by your knowledge; as it progressed, was also impressed and aided greatly by your wisdom; and now as we approach the end, am very thankful for your patience.

To my thesis committee, Professors Christopher Stubbs, Melissa Franklin and Lars Hernquist, I give my thanks for your unhesitating willingness to join me in this, and for your wise thoughts and feedback.

No less impressive than George Field's patience is that of my wonderful spouse Reebee, who without complaint put up with countless late-night coding sessions, weekends spent debugging, and holidays taken over by writing. I couldn't have done this without your support and encouragement.

Both for their friendship and their specific help in writing and reading this dissertation I am grateful and indebted to Paul Janzen, Warren Brown, J.D. Paul and Ken Rines. My long-term officemates along the way – Dan Koranyi, Andi Mahdavi and Vit Hradecky – all provided welcome, supportive friendship, lessons on the niceties of caffeine, and healthy distractions both intellectual and not. Thanks for not letting me take myself too seriously! When not busy at our respective computers Mike Westover and Dan Green gave me all I could handle on the tennis court, while David Charbonneau, Pauline Barmby, Ian Dell'Antonio, Saurabh Jha, Hannah Jang-Condell, Jan Kleyna, Ted Pyne, Brian Schmidt, Ann Bragg, Dave Latham, Herb Shea and Thomas Dignan each in their own ways helped keep me sane and balanced along the way.

I thank Josh Grindlay for getting me started in astrophysics, and Jean Collins, Peg Herlihy, Beth Rigos and Sheila Ferguson for the umpteen things they did to help keep me going in astronomy and in physics.

Finally to Jim Wilkinson and all my colleagues at the Bok Center goes my heartful gratitude for your support and enthusiasm as I've led somewhat of a dual life at Harvard, providing me the opportunity for an substantive professional career while still completing my graduate work. I look forward to continuing the one while building on the other.

This work was supported by NASA through G. Field's research grants, by funding from the Harvard Deptartment of Physics and the Center for Astrophysics, and by a Natural Sciences and Engineering Resource Council of Canada "1967" Scholarship.

To my parents.

List of Figures

List of Tables

Chapter 1

Introduction

> "The challenge of simulating a dense star cluster with a million stars is formidable, because of the enormous ranges in spatial and temporal scales that have to be modeled simultaneously." [2]

1.1 Astrophysical Motivation: Quasar Massive Black Holes

The observation of high-redshift quasars, *i.e.,* those with redshift $z > 6$, implies a problem for understanding how they formed. The prevailing model for the underlying astronomical object is that of a supermassive black hole accreting material gravitationally [3]. Each high-redshift quasar thus must have had a supermassive black hole in place within the first 10^9 years after the big bang, which turns out to be a rather strong physical constraint. Observations of high-redshift gamma-ray bursts and massive galaxies place similar limits on the formation timescales of the first massive stars [4] and of the first massive galaxies [5].

Of the two possible basic building blocks for a massive object in the early universe – gas, and stars that have already coalesced from the gas – direct formation from the collapse of a region of gas may seem to have certain advantages, in that it avoids the intermediate step of forming stars. But the required gas cooling times are barely consistent with current structure formation models (and even then require reliance on "extremely rare" overdensities) and additionally the centrifugal barrier that arises due to tidally-induced rotation is a "key obstacle" which must be overcome by any model of massive black hole formation [6]. As will be described in more detail below, few previous studies have examined the evolution and possible core collapse of rotating dense stellar clusters, and none have modeled whether such systems can form massive stars which could then serve as seeds for quasar black holes.

The task of forming a massive object from stellar-mass objects faces its own hurdles: even in dense environments stars are less affected by collisions than is gas (although stellar collisions have the benefit of not necessarily causing shock heating) and so energy and angular momentum transfer due to collisional processes are not as effective *a prioi* for stars as they are for gas. However, stellar systems can overcome this handicap through the formation of a large, coherent dynamical subsystem – the most observationally prevalent of which is a "stellar bar" – which can act as a bulk perturbation in the gravitational potential of the system and so can effectively transport angular momentum outwards, thus allowing the system's core to contract further. This mechanism is described further in §1.1.3 and developed in detail in §3.5.

The goal of the current project is to generalize the approach of prior researchers, who developed one-dimensional models which simulated nonrotating stellar systems (*e.g.,* Quinlan and Shapiro [7] and other references given below), to two dimensions in order to address the questions: What about rotation, does it inhibit collapse and can it be overcome? The anticipated answer is that, despite initial rotational support against collapse, the presence of a stellar bar will allow sufficient transport angular momentum outwards so that collisional stellar mergers and mass segregation can produce a massive object (perhaps $\simeq 10^{2-3} M_{\odot}$) in the core of a dense stellar cluster such as could be found at the center of a newly-forming galaxy. Such an object could then undergo growth via accretion to reach supermassive size ($\sim 10^{6-8} M_{\odot}$) within a Hubble time [8] and be observed as a quasar or active galactic nucleus, thus bringing together the formation of early-universe stars, galaxies, and massive collapsed objects.

The remainder of this chapter expands on each of the above ideas, starting with dense stellar systems, and moving on to how massive black holes may form in such stellar clusters. How rotation impacts the system is then discussed. Finally an overview of the motivation for and development of the numerical simulation technique used is given.

1.1.1 Stellar Systems

A primary motivation for this study was the work of Quinlan and Shapiro [1], who developed a Fokker-Planck model of an idealized nonrotating, spherically symmetric, dense cluster of compact stars, and found "rapid buildup of massive black holes in the cluster core resulting from successive binary mergers and mass segregation." Subsequently, Quinlan and Shapiro studied clusters of solar-mass main sequence stars and found that it was"remarkably easy for massive stars to form through multiple stellar mergers in dense galactic nuclei" [7]. After performing a complemetary line of simulations using the specialized GRAPE N-body computing cluster, Portegies Zwart and McMillan [9] speculated that a sufficiently-dense $10^6 M_{\odot}$ stellar cluster which initially formed ~ 30 pc from the Galactic center can spiral inward due to dynamical friction before being disrupted by the Milky Way's tidal field. While a less-dense cluster would be disrupted before reaching the Galactic center and so merely contribute its stars to the Galactic bulge, the more-dense cluster dissolves closer to the center of the Galaxy and leaves behind its central massive black hole which can then grow to supermassive size by accretion.

In early theoretical work on stellar systems Begelman and Rees [10] found that "runaway coalescence" of stars (*i.e.,* stellar mergers combined with mass segregation) could lead to the formation of a central massive object. Again, the system's evolution was a race between coalescence and disruption, the latter in this case being due to gas released by earlier stellar evolution and collisions. Lower-density clusters were susceptible to disruption and dissipation, while more-dense systems retained most of the gas which went to form new stars. Stellar clusters were thus explicitly linked to active galactic nuclei: "Dense star clusters may be responsible for some of the low-level manifestations of activity in galatic nuclei; but they are probably merely precursor stages of the more spectacular quasar-type phenomena, which develop after a massive object has formed."[1] The authors note that runaway stellar merging combined with mass-segregation provides "one of the quickest routes to the formation of a massive object in a dense stellar system".

Star Cluster Observations

Dense stellar clusters are fairly common in the nuclei of large galaxies: 75% or more of late-type (classification Scd-Sm) spiral galaxies have nuclear clusters, as do at least 50% of earlier type (Sa-Sc). These nuclear clusters are as compact as globular clusters found in the Milky Way, with a typical (half-light) radius of $2-5$ pc (Böker [11], and references therein), but they are massive. The dynamical mass of a nuclear cluster is in the range $10^6 - 10^7 M_\odot$, *i.e.,* at the very high end of the globular cluster mass function – and similar to the cluster masses used here. Böker concludes that nuclear clusters are "an intriguing environment for the formation of massive black holes because of their extreme stellar density".

An impediment to building massive black holes in stellar clusters is that the initial mass function (IMF) which describes the initial distribution of stars across the possible range of masses is observed to have a cutoff at around $150M_\odot$ [12], as described in more detail in §2.6.3. Thus there is a gap between the IMF's upper cutoff and the $250M_\odot$ minimum stellar mass required to leave behind a remnant massive black hole after going supernova [13]. If black hole formation is to be feasible this gap must be overcome by dynamical processes in the cluster, some candidates for which are considered in the next section.

1.1.2 Scenarios for Massive Black Hole Formation in Dense Clusters

Observationally, Ebisuzaki *et al.* [14] note the existence of what they refer to as the "missing link" between stars and supermassive black holes: an IMBH seen in or near a compact stellar cluster in the center of galaxy M82. They take this as evidence for supermassive black holes being created through mergers of IMBHs, which themselves form via the merging of massive stars in compact clusters. Coleman Miller and Colbert [15] also note that IMBHs exist in the current universe, and that inner galactic bars can contribute to the growth of a $\lesssim 1000M_\odot$ black hole; they also postulate SMBHs from early IMBH mergers, perhaps as stellar clusters percolate to the galaxy center and merge.

Approaching the problem from the other direction, taking the existence of high-redshift (*e.g.,* $z = 6.41$) quasars as a given and examining timescales by which a supermassive black hole (SMBH) may grow from an earlier IMBH, Tyler *et al.* [16] conclude that a supermassive star (whether formed as such or as a stellar-merger product) was a more likely SMBH progenitor than any of a primordial IMBH, a large Population III star, or the merger product of smaller Population-III black hole remants. Each of latter would be required to form very early and at the extreme upper end of their physically possible mass ranges, followed by continuous Eddington-limit accretion. Again, they did not consider rotation.

The overall difficulty of explaining high-redshift quasar black holes is summarized by Düchting [17]: "The existence of sufficient primordial black holes would unaccountably close the universe" unless the spectrum of masses is fine-tuned. It is "quite difficult to construct viable

---

[1] This citation also includes Rees' prototypical supermassive black hole formation-scenarios flowchart.

physical models.” Remnants of Population III stars, or IMBHs formed from the collapse of early star clusters, also have “serious problems” accounting for $z = 6$ quasars.

In reviewing possible formation processes for massive central black holes in nonrotating dense stellar clusters with a few thousand to a few million stars, Rasio *et al.* [18] found two possible paths. If core collapse of the cluster proceeds quickly enough – occurring before the typical lifetime of a large main-sequence star, $\sim$ 3 Myr [19]– stellar collisions can produce a single supermassive star which then evolves into a intermediate-mass black hole (IMBH) of mass $10^2 - 10^4 M_\odot$. Alternatively, if large individual stars go supernova before core collapse and leave stellar-mass black hole remnants, the supernova mass loss may reverse core collapse somewhat for smaller clusters, while larger systems may still collapse relativistically [1]. In either case the initial cluster evolution is dominated by the most massive stars, *i.e.,* by the high end of the stellar initial mass fuction (IMF). The former case of forming a massive object from collisions of main-sequence stars is the object of this dissertation.

While also plausible, modeling the evolution of a cluster of relativistic stellar remnants is a sufficiently different as to require a separate, later study. Since main-sequence stellar collisions and cluster dynamics “depend crucially on each other” [20] it is unclear what initial conditions to assume for a post-main-sequence cluster. And restricting the simulations to a main-sequence stellar population also avoids the complications of how to model stellar mass loss due to post-main-sequence stellar evolution in which the liberated gas could do any of: remaining as gas in the cluster; evaporate from the system; or coalesce into new stars, all while the initial stars are transforming into stellar-mass remnants [7]. While also interesting, following this stage of the cluster evolution would greatly increase the complexity of an already quite complicated model.

1.1.3 Rotation due to Tidal Torquing

Rotation is ubiquitous in the universe. A protosystem’s expected rotation from being spun up by tidal torques due to cosmological perturbations is “almost independent of the perturbation spectrum” and only weakly correlated with the individual overdensity being considered [21]. An initial angular momentum is an “important piece of realism ... while the probability that the cloud from which a star cluster originates, has zero total angular momentum is very small, practically all models have assumed that.” [22].

Recently, some work has been done on modeling rotating stellar clusters, and the picture that emerges is not always intuitive. Rotating isotropic systems were found to be dynamically stable in the N-body studies performed by Meza [23], but the systems were found to be able to rotate “very rapidly” without becoming oblate. Rotation can accelerate the time to core collapse in single-component globular clusters [22] and was found in 2-component globular clusters to increase the speed of dynamical evolution via the enhanced outward transfer of angular momentum, but only if dynamical friction does not dominate. (However, this also results in rotation being a check on mass segregation: massive stars’ angular momentum J prevents them from sinking as far in as they would otherwise, due to the “gravo-gyro instability” [24].)

In general, rotation in dense stellar systems is important for galactic nuclei, although the vast bulk of previous studies have focused on globular clusters [25]. Most prior work on larger systems has either been to model a specific system and/or to study systems with a pre-existing massive black hole (*e.g.,* the modeling of dwarf elliptical galaxy M32 by Arabadjis [26], and references therein). On galactic scales Barnes and Efstathiou [21] did find that rotation can inhibit collapse in N-body simulations, but suggest caution with analytic predictions of angular momentum evolution in complex stellar systems.

As rotation does not necessarily induce flattening of the system, the simulations performed here assume sphericity for simplicity's sake, and the potential is constructed so that the model is self-consistent with that assumption [27].

Seeding Black Holes from Low Angular Momentum Material

One avenue for overcoming the rotational-support problem is to postulate that only the least-rotating material contributes to the formation of a seed black hole. Koushiappas *et al.* [6] modeled SMBH seeds from the gas with the least angular momentum in "rare-peak" haloes of the early protogalaxies, under the assumption that the distribution of the gas has a significant low-angular-momentum tail, which can be extrapolated to lower masses and smaller radii from the distribution found for dark matter in cosmological simulations. With this assumption they were able to produce seed IMBHs of $10^5 M_\odot$ for any halo larger than $7 \times 10^7 M_\odot$ – but no black hole at all for smaller haloes, including the bulges of most disk galaxies. While giving promising results for large galaxies, they acknowledge that the rotational barrier is a key concept and that their model still requires another undetermined mechanism for outward transport of angular momentum in order to achieve the gas's collapse to a black hole, one which cannot operate for lower-mass haloes and so does not account for any continuation of the galaxy/central-object scaling relation to lower-mass galaxies [11]. Thus another solution, which does not depend on the specific requirement of unusually rare low-rotation initial conditions, is warranted to account for at least some of the observed massive black holes in disk galaxies.

Stellar Bars as Bulk Transporters of Angular Momentum

Although some angular momentum J is expected to be transported outwards in a rotating stellar system via the effects of shear between the faster-rotating inner regions and the slower-rotating outer regions, a bulk perturbation in the potential can be much more effective than individual stars at transporting angular momentum. In general, any nonaxially-symmetric component of the potential in a rotationally supported gravitational system fosters J transport over large distances, due to gravitational torques [28]. The amount of transport can be as much as an order of magnitude greater than that in a purely spherical stellar system [29]. Specifically, a bar-like perturbation in the potential can provide a much more effective angular momentum transfer mechanism compared to what independently-orbiting stars can effect [30], with J being transported from inner to outer regions [31] and eventually to the stellar system's halo region – the presence of which does not stabilize the system against the formation of a bar perturbation in the first place [32]. The bar excites a gravitational wake in the stellar distribution: as the bar passes through a region, other stars experience a shear force along its orbital path. This has a net drag effect specifically on individual stellar orbits in or near resonance with the bar's rotational frequency [33]. This "resonance drag" in finite rotating systems is the analogue of classical dynamical friction in an infinite, homogeneous medium [34].

Bar perturbations were first observed within the spiral structure of galaxies and are still commonly associated with galactic scales. However, inner bars of scale $\lesssim$ 100 pc are common and form before larger-scale bars [35]. While galactic bars may be largely gas, inner bars are mostly stellar and not good at gas transport [36]. There are dynamical arguments for why having both outer and inner stellar bars should be the standard model, with the inner bar decoupling from the outer system [37]. Observationally, for example, the S0 galaxy NCG 4621 has a 60 pc counterrotating (*i.e.,* decoupled) core, the "smallest to date" [38]. Thus, subject to tests confirming that a given system satisfies the dynamical instability criterion for a bar to form (as given in §3.6), a stellar bar is invoked here on the scale of the stellar cluster being modeled in order to foster angular momentum transport outward from the cluster center.

1.2 Motivation for Technique

1.2.1 Comparison of Fokker-Planck and N-body Methods

The two main techniques for modeling collisional stellar systems such as globular clusters, open clusters, or galactic nuclei, are through *N-body* simulations in which particles representing a number of stars are allowed to interact via relevant gravititational and other physical processes, and through integration of the *Fokker-Planck* equation which represents the stellar population statistically as a distribution function f. When compared against each other on simple (spherical, nonrotating) systems the two techniques agree to a large extent ([22], [39]) and so the choice of which method to use depends on the requirements of a given study.

N-body simulations have the advantage of being "direct" in that individual particles are tracked and can be followed through the entire simulated evolution of the system over time, and the physical effects that dominate their behavior can be analyzed. A disadvantage is that it is not always apparent how to incorporate a given physical effect into the N-body construct in the first place. Nor do individual particles in the simulation correspond to individual objects in astronomy (*i.e.,* stars). The typical N in current N-body studies is still low compared to the number of stars in real astronomical systems, and how to extrapolate between "particles" and "stars" is not apparent [22]. Also N-body methods are noisy in the sense that a given simulation cannot be relied on to represent the evolution of a system; instead one must run a suite of simulations in order to get a statistically robust picture, which will be the case as long as N is not up to realistic particle numbers, preventing truly star-by-star modeling [25].

On the other hand, the statistical nature of the Fokker-Planck equation "captures the essence" of continuum mechanics [40] but does not afford direct analysis of how different physical processes contribute to the observed behavior of the system. All one knows is how the stellar distribution function f evolves over time. However, Fokker-Planck does allow for different aspects of the physics to be added or removed at will; thus comparing integrations with and without a given effect included can accomplish the same result. This ability to include new dynamical aspects in an *ad hoc* manner permits us to study a discrete spectrum of stellar masses (*e.g.,* as previously done by Amaro-Seoane [41]), collisional mergers of stars (*e.g.,* Quinlan and Shapiro [7]) and a bulk perturbation in the gravitational potential such as the stellar bar.

1.2.2 Orbit Averaging and the Third Integral

As developed in Chapter 3, the Fokker-Planck approximation[2] is appropriate when small angle, 2-body scattering dominates the global evolution of the system. When that is the case, energy transport can be treated as analogous to heat conduction in a collisional gas – producing a "fairly good description of what happens in N-body simulations" [22]. The full Fokker-Planck equation tracks the time-evolution of the system in all 6 dimensions of position and velocity (or more generally, momentum) which is both computationally prohibitive and usually not necessary for describing the system's evolved properties.

In order to simplify the problem one can attempt to eliminate the positional dependence of the Fokker-Planck equation, leaving only the momentum dependence to describe the dynamics. A typical method used to achieve this is *orbit averaging* in which the Fokker-Planck equation and all dynamical quantities on which it depends are, as the name implies, averaged over a full stellar orbit before being applied. These orbit-averaged Fokker-Planck models are found to "treat very well the diffusion of orbits according to the changes of their constants of motion, taking into account the potential and the orbital structure of the system in a self-consistent way" [41].

One is thus left with a description of a stellar system based on three constants of motion, corresponding to the three dimensions of momentum. This is consistent with the strong Jeans theorem, which requires in the general case that there be three integrals of the motion, even if only two are well determined [39]. Starting with Goodman's 1993 dissertation [43] standard practice has been to use energy E and angular momentum component J_z as the constants of motion, and to simply ignore any third integral. Comparison to N-body results has shown this to be reasonable if flattening of the system is not extreme [22]. For such systems, however, it is in principle possible to use J^2 as a proxy for the third integral, although this would be "extremely difficult numerically and physically" [39]. Strictly speaking, ignoring the third integral is theoretically valid if and only if the system's gravational potential is spherically symmetric, even if its velocity distribution is not [42] – hence the restriction to studying rotating, but gravitationally-spherical systems.

1.2.3 Choice of Canonical Variables

Depending on the choice for constants of the motion, the two-dimensional orbit-averaged Fokker-Planck equation can be used to study different dynamical apsects of a stellar cluster. With the assumption of isotropy of the velocity dispersion, the traditional choice of (E, J_z) can represent an axisymmetric rotating system, *e.g.,* Kim *et al.*'s single-mass [39] and two-mass clusters [24]. Alternatively (E, J^2) can be used to study the somewhat simpler problem of anisotropy in spherically symmetric systems, as first done by Cohn [44] and more recently by Takahashi ([45], [46]) and Fiestas *et al.* [25]. The latter problem is simpler in that the

---

[2] *Approximation* here refers to the idea that the Fokker-Planck equation can be derived by expanding the collisional Boltzmann equation under the assumption of weak encounters, *i.e.,* $\delta v \ll v$, and truncating the resulting series after the second-order terms [42].

distribution can be averaged over angular momentum J before coefficients of the Fokker-Planck equation are calculated, providing a great savings in complexity and calculation time.

However, once incorporation of *ad hoc* terms into the Fokker-Planck equation is desired, *e.g.,* to represent collisional stellar mergers (as originated by Quinlan and Shapiro [7] in a one-dimensional study) energy E is no longer necessarily an integral of the motion. Any change in the potential will cause changes in the stellar distribution function when it is expressed as $f(E, J^2)$ or as $f(E, J_z)$. What is required are the radial action I and the tangential action (*i.e.,* the overall angular momentum) J, which are still adiabatic invariants as the potential evolves, and so also is $f(I, J)$. Binney and Tremaine [42] pointed out this exact advantage of using the canonical actions as coordinates for direct numerical computation of the Fokker-Planck equation, as a method for stellar dynamics simulations. Additional benefits of this approach are that the Fokker-Planck equation takes a particularly clean form and that the procedure of orbit-averaging is then a conceptually simple averaging of quantities over the 2π change in (an) orbital angle variable, as will be seen in Chapter 3.

With this form it is also possible to straightforwardly incorporate terms describing collisional stellar mergers into the Fokker-Planck equation. The main drawback is that a large number of conversions are required between the (E, J, J_z) coordinates in which dynamical quantities are calculated and the (I, J) space in which the system evolution is tracked by the Fokker-Planck equation for $f(I, J)$. Integration of the two-dimensional Fokker-Planck equation – even in its more clean form, as when expressed in term of actions I and J – is sufficiently challenging that this tradeoff remains beneficial. Effectively the additional complexity of the dynamics has been separated out of the Fokker-Planck equation itself and into these coordinate conversions. The result is a three-step simulation process. First the Fokker-Planck equation is used to update the distribution function $f(I, J)$ at a new timestep. Then the effects of any additional physical processes – in this case, collisional stellar mergers – are calculated and $f(I, J)$ updated accordingly. Finally, given the new $f(I, J)$, consistent values for physical quantities such as the stellar density and gravitational potential are found in terms of either E and J or the spatial coordinates.

1.3 Overview of Resulting Model and of Remaining Chapters

Given all the above considerations, the basic form of the model developed for this thesis has the following characeristics:

1. The system to be studied is a fairly dense stellar cluster, using a generalization of the initial conditions from Quinlan and Shapiro [7], with an observationally-motivated initial mass function for the stars which make up the cluster;
2. Unlike most prior work on dense clusters, an overall rotation of the system is included, with a value determined by observation;
3. In order to allow for the inclusion of desired additional physical effects beyond the general gravitional interactions of individual stars in the cluster – *i.e.,* a stellar bar perturbation, and collisional merging of individual stars – the Fokker-Planck equation is employed to track the system's evolution; and

4. To capture the physics required while still keeping the problem numerically tractable the Fokker-Planck model is orbit-averaged, and any third integral neglected, so that the system is represented in two dimensions. The canonical actions (I, J) are employed as integrals of the motion in order for the above-mentioned extra physical effects to be incorporated straightforwardly.

With this setup for the model, the basic question addressed here is under what circumstances – using which physical effects, starting with what initial conditions – can the system produce a star which can serve as seed object for an intermediate-mass black hole and eventually a supermassive black hole. For this to be possible, the star must have sufficient mass, *i.e.,* $250 M_\odot$ or more, and it must be formed before the end of the main-sequence lifetime of large stars, *i.e.,* within the first 3 Myr or so of the system's evolution. The simulations will show that a bar-like perturbation can effect sufficient angular-momentum transfer to allow the creation of a likely seed object in most rotating clusters whose density and velocity dispersion match what is seen in the centers of galaxies, but only if the initial mass function is also set to match that observed in dense clusters – which is more top-heavy that the general field-star IMF.

The remaining chapters are organized as follows:

- Chapter 2 outlines the dynamical and gravitational-potential physics required for the Fokker-Planck model, lists the numerical considerations needed to compute it, and outlines the set of possible initial conditions.
- Chapter 3 derives the diffusion coefficients that incorporate the gravitational effects of the system's individual stars and of any stellar bar perturbation into the Fokker-Planck equation for the evolution of the system, as well as the population effects of collisional mergers. The differencing scheme employed to compute the model is developed, and a brief consideration of the heating and merging effects of binary stars is given.
- Chapter 4 describes the tests performed on the various aspects of the model, both to verify its numerical and physical validity and to establish reasonable values for any numerical input parameters required.
- Chapter 5 details the results of the simulations for the various possible choices of initial conditions and of which physical effects to include.
- Chapter 6 then discusses the simulations' results from an astrophysical standpoint, gives comparisons to what would be expected from previous studies and from general considerations such as timescale arguments, and looks forward to what upcoming observations and future models may be able to provide.
- Tables of frequently-used symbols and subscripts are given in the Appendix.

In the end, the simulations show that dense rotating stellar clusters are a feasible path to the production of seed objects for massive black holes in galaxies. All simulations that were performed with a top-heavy IMF – such as is observed in dense stellar clusters – resulted in at least one sufficiently massive star (and usually several) being formed. The result holds for a range of initial cluster configurations which were chosen to match typical dense nuclear clusters and galaxy centers. Only for models representing giant elliptical galaxy cores was a massive object not seen before the largest stars evolved off the main sequence. For these systems one would need to rely either on post-main-sequence dynamics within the core, or on the presence of a dense nuclear stellar cluster which could evolve independently.

Chapter 2

Dynamical and Gravitational Aspects of the Model

2.1 Overview

The first two sections of this chapter describe the simulation model's dynamical framework, specifically the calculation of orbital parameters *e.g.,* endpoints, frequencies *etc.* These will be needed in Chapter 3 to determine the effect of dynamical friction on the ensemble of stars in the cluster, as described respectively by the diffusion coefficients and the stellar distribution function in the Fokker-Planck equation. The remainder of the chapter then develops the calculations of the cluster's physical properties, such as the mass density profile $\rho(r)$ and the gravitational potential $\Phi(r)$, and ends with a summary of the possible initial conditions used for these quantities.

The fundamental units employed in the model are km/s, $M_\odot$, and pc. With these choices the unit of time is $1\text{pc}/(\text{km/s}) = 0.9778 \times 10^6\text{yr} \simeq 1\text{Myr}$, and Newton's constant of gravitation is $G= 4.299 \times 10^{-3}\,(\text{km/s})^2\,\text{pc}/M_\odot$.

2.2 Dynamics

The most convenient way to parameterize stellar orbits is to use to the canonical radial and angular actions, which together will be referred to as the action vector I_j, with components I_1 (the radial action) and I_2 (the angular action). I_2 is equal to the traditional angular momentum J, while I_1 can be defined from the radial momentum:

$$I_1 = \frac{1}{2\pi}\oint dr\,p_r = \frac{1}{\pi}\int_{r_p}^{r_a} dr[2(E-\Phi) - J^2/r^2]^{1/2} \tag{2.1}$$

in which orbital endpoints r_p, r_a are the star's periapse and apoapse distances from the cluster center, E is the star's total orbital energy and $\Phi(r,\theta)$ is the gravitational potential at distance r and polar angle θ from the center.[1] Azimuthal symmetry of the cluster's main

---

[1] Shlosman [28] points out that the Jacobi energy E of (2.1) isn't a true integral of motion, even for individual stars, when there are nested bar perturbations: the underlying potential is then not a constant.

stellar distribution (and so also of the potential) is assumed.

The advantage of using the canonical coordinates is that the stellar distribution function, when expressed in terms of these actions, is invariant to adiabatic changes in the potential; hence we do not have to solve for a new distribution function just because the potential evolves over time [42]. As seen above however it is often necessary to deal with the traditional variables $r, \theta, \phi, E, J, J_z$ when calculating orbital parameters or any physical quantity that depends on the value of the potential. Alternatively, when the actions (I_1, I_2) are known, (2.1) implicitly defines the stellar orbital energy E.

Following Tremaine and Weinberg [30], we can move between action space and position space via the relations that follow. The radial and angular orbital frequencies Ω_j are found similarly to I_1 :

$$\frac{1}{\Omega_1} = \frac{1}{\pi}\int_{r_p}^{r_a} \frac{dr}{[2(E-\Phi) - J^2/r^2]^{1/2}} \tag{2.2}$$

$$\frac{\Omega_2}{\Omega_1} = \frac{J}{\pi}\int_{r_p}^{r_a} \frac{dr}{r^2[2(E-\Phi) - J^2/r^2]^{1/2}} \tag{2.3}$$

From these we can relate the dynamical quantities in the two coordinate systems:

$$\left(\frac{\partial I}{\partial E}\right)_J = \frac{1}{\Omega_1}, \quad \left(\frac{\partial I}{\partial J}\right)_E = -\frac{\Omega_2}{\Omega_1} \tag{2.4}$$

$$\left(\frac{\partial E}{\partial I_j}\right) = \Omega_j \tag{2.5}$$

where the last relation generalizes to the case including a third integral of motion if we define $I_3 = J_z$ and note that $\Omega_3 \equiv 0$.

Although orbit-averaging will remove any direct dependence of the diffusion coefficients on the canonical angles w_1 and w_2, the angles are needed for intermediate steps in the coefficients' calculation. (The angles are also needed for some of the tests of the coefficients' validity.) They are given by

$$w_1 = \Omega_1 \int_C \frac{|dr|}{[2(E-\Phi) - J^2/r^2]^{1/2}} \tag{2.6}$$

and

$$w_2 - \psi = \int_C \frac{(\Omega_2 - J/r^2)|dr|}{[2(E-\Phi) - J^2/r^2]^{1/2}}. \tag{2.7}$$

The azimuthal angle w_3 is not of interest except in that it contains a random initial phase ϕ.

As shown in Fig. 2.1, ψ is the angle swept out by the star's position vector as measured from the point where it crossed the $\theta = \pi/2$ plane:

$$\sin\psi \sin\beta = \cos\theta, \tag{2.8}$$

This is expected to be the case for clusters such as those modeled here when near the center of a galaxy [36]; however, the inner bars decouple dynamically from the larger-scale galactic bars [37] which will vary on much longer timescales, and so this nicety is expected to produce only a secular change in the underlying potential that averages out over long timescales.

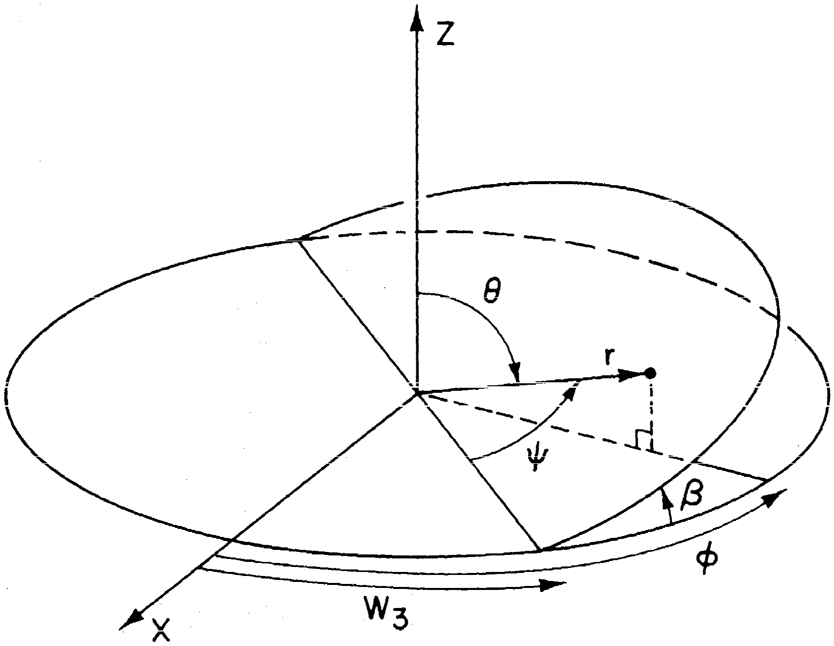

Image source: Tremaine and Weinberg [30].

Figure 2.1: Euler angles for the orbit of a given perturbing potential.

in which β is the orbital inclination, *i.e.,* $J_z = J\cos\beta$. Integration path C is along the orbit, *i.e.,* from r_p to r_a and back to r_p when traversing an entire orbit. For our purposes the quantity $(w_2 - \psi)$ is what will be needed most often. The above relations are intuitive given that $[2(E-\Phi) - J^2/r^2]^{1/2}$ is the radial component of the velocity and the canonical angles obey $d^2w_j/dt^2 = 0$.

2.2.1 Numerically Solving for Orbital Endpoints

The general method for finding the orbital endpoints r_p, r_a is to solve numerically for the roots of $v_r^2 = [2(E-\Phi) - J^2/r^2] = 0$. For small values of either action component I_1 or J this may not be numerically robust and other techniques must be used. If $(E-\Phi)/|\Phi|$ is miniscule (*i.e.,* $\lesssim 10^{-8}$), the star is considered to be "fixed" at the origin. For $(E-\Phi)/|\Phi| \lesssim 10^{-3}$ a quadratic approximation to the potential is employed and the endpoints solved for analytically. The endpoint-solver takes as its natural inputs E and J, but it also requires at least a reasonable guess of I_1 so it can distinguish between more-radial and more-circular orbits.

For very-nearly-radial orbits (taken to be $|J| < z_{\text{tol}}\, I_1$), r_p is set to a value of 10^{-8} in order to avoid numerical problems at exactly $r = 0$. [2] The input parameter $z_{\text{tol}} = 10^{-5}$

---

[2] Alternatively, for nearly-radial orbits $I_1/|J| \lesssim 10^{-3}$, an epicyclic approximation for finding r_p and r_a as well as Ω_j was tried. This was not found to be an improvement and results with the epicycle approximation are not reported here. However, for numerically tiny values of the components of the action (typically $I_j \lesssim 5 \times 10^{-4}$) *i.e.,* for stellar orbits restricted to the cluster center, an epicylcic approximation [42] is employed: $(\Omega_2)^2 \simeq \partial^2\Phi/\partial r^2|_{r=0}$ and $\Omega_1 = 2\Omega_2$, although in practice no values of I_j on the grids used trigger this case.

typically, although other values were tested to make sure the results were robust against such a change.

For moderately-circular orbits (those with $I_1/|J| \lesssim 0.17$), first a value of r for which $v_r > 0$ is found; then r_p and r_a are determined by searching out in both directions until v_r changes sign. For more radial orbits, the entire allowed range of r is searched for subranges over which v_r changes sign. In both cases, the bracketing of the ranges over which v_r changes sign is then made increasingly finer until convergence is reached.

2.2.2 Calculating Orbital Frequencies and other Dynamic Quantities

Once orbital endpoints r_p and r_a are known then (2.3) and (2.2) can be used to find orbital frequencies Ω_j directly (and similarly (2.1) gives the action components I_j if they are not yet determined). In practice doing integrals is very expensive in processor time; thus in general Ω_j is calculated from (2.5). When values for $\partial E/\partial I_j$ are not yet known for a given simulation timestep, however, (2.3) and (2.2) are employed.

Just as in the orbital endpoint calculation, boundary cases require special treatment. When the orbit is highly radial (taken to be when $J/I_1 \lesssim 5 \times 10^{-4}$), (2.3) becomes unstable and so the limiting case of $\Omega_2/\Omega = \frac{1}{2}$ is used. Finally, for orbits that are very nearly circular ($I_1/J < 5 \times 10^{-4}$), even if the solution for r_p and r_a is obtainable, doing further calculus such as (2.3) and (2.2) is numerically difficult and so in this case the orbit is taken to be exactly circular with $r_p = r_a = r_c$, the radius at which $\partial v_r/\partial r = 0$. The above I_1/J requirement is sufficiently strict that this will not affect any dependent calculations.

2.3 Calculation of the Density

2.3.1 Using the Distribution Function in (E, J^2) Space

Traditionally, the (smoothed) stellar mass density ρ is given by the "full" 6-dimensional integral of distribution function F over all velocity space – which gives the total number density of stars – multiplied by the per-star mass m:

$$\rho(\mathbf{r}) = m \int d^3v \, F(\mathbf{r}, \mathbf{v}) \tag{2.9}$$

which results in a total mass M for the cluster of:

$$M = \int \rho(\mathbf{r}) \, dV = m \int d^3r \, d^3v \, F(\mathbf{r}, \mathbf{v}). \tag{2.10}$$

Previous studies assumed a stellar cluster with spherical symmetry and used (E, J^2) as the coordinates of choice; using the transformation of Cohn [44], in (E, J^2) space the above expression becomes

$$M = 8\pi^3 m \int d(J^2) \int dE \, \Omega_1^{-1} F(E, J). \tag{2.11}$$

With azimuthal symmetry, 2π of the above comes from integrating over the velocity-space azimuthal angle ϕ_v. Expanding $d(J^2)$ and applying (for fixed J) $dE = \Omega_1 dI$ from (2.3), the total cluster mass when integrated over the actions is thus

$$M = 16\pi^3 m \int J dJ \int dI\, F(E(I,J),J). \tag{2.12}$$

2.3.2 Using the (E, J^2)-space Distribution Function in Action Space

In order to take advantage of the invariance of the actions, we define a distribution function analogous to F but expressed in action space:

$$M = m \int dI\, dJ\, f(I,J). \tag{2.13}$$

Note that this is by construction an "orbit-averaged" distribution function ([42], [44]) and so all angle-dependence has already been averaged out. Comparing with (2.12) one obtains

$$f = 16\pi^3 JF \tag{2.14}$$

as the conversion between the action-space f and the earlier F.

To calculate the spatial density ρ, one can either perform the integral of (2.9) directly, or first convert it to a 2-dimensional orbit-averaged form that uses f, similar to (2.13). The former requires use of the velocity volume element (see *e.g.,* Cohn and Kulsrud [47])

$$\int_{\phi_v=0}^{2\pi} d^3v = 4\pi \frac{JdJdE}{r^2 v_r} \tag{2.15}$$

and so, if using the two-dimensional f of (2.13) and assuming spherical symmetry,

$$\begin{aligned}
\rho(r) &= 4\pi m \int_{\Phi}^{0} dE \int_0^{J_{\max}} \frac{JdJ}{r^2 v_r} F(E,J) \qquad (2.16)\\
&= 4\pi m \int_0^{J_{\max}} JdJ \int dI \frac{\Omega_1}{r^2 v_r} F(E(I,J),J)\\
&= \frac{m}{4\pi^2} \int_0^{J_{\max}} dJ \int dI \frac{\Omega_1}{r^2 v_r} f(I,J)
\end{aligned}$$

where the endpoints of the I integral are those energetically allowed. This has the advantage of being a direct method of calculation, but is otherwise undesirable for this study for several reasons: the $1/r^2$ dependence means that the ρ calculation is less robust close to the cluster center; the singularities at the orbit endpoints due to the $1/v_r$ dependence are particularly difficult to deal with in the $dIdJ$ form of (2.16), and spherical symmetry is not an intended assumption for this model.

2.3.3 Calculating the Density in Action Space

In order to deal with the above issues it is necessary to develop a differential in terms of the action coordinates themselves. Expanding d^3v in spherical coordinates we have

$$d^3v = v^2 dv\, d\cos\theta_v\, d\phi_v. \tag{2.17}$$

Defining the "tangential" velocity $v_\theta^2 + v_\phi^2 \equiv v_T^2 = (J/r)^2$, then $dJ = r dv_T = rv\, d\cos\theta_v$ and so

$$d^3v = 2vdv\, d\phi_v \frac{dJ}{r} \tag{2.18}$$

and

$$\int_{\phi_v=0}^{2\pi} d^3v = 4\pi\, v\, dv\, \frac{dJ}{r}. \tag{2.19}$$

Note that 2π in the above again comes from integrating over azimuthal angle ϕ_v, and an extra factor of 2 has been inserted in order to allow use of the usual convention that $J \geq 0$ only. Finally, with $\Omega_1 dI = dE = v\, dv$,

$$\int_{\phi_v=0}^{2\pi} d^3v = 4\pi\, \Omega_1\, dI\, \frac{dJ}{r}. \tag{2.20}$$

Inserting this in (2.9) and using the conversion from F to f,

$$\rho(r) = \frac{m}{4\pi^2 r} \int_r \frac{dI\, dJ}{J}\ \Omega_1 f(I, J) \tag{2.21}$$

where as always, the action integral endpoints are determined by which orbits are energetically allowed at r; this is indicated by subscript "r" on the integral sign. This form for the calculation of ρ is more suited for the overall action-space-based model of this study than is (2.16), and is what will be used throughout; the cost of this choice is that the orbital endpoints are now more complicated to calculate.

In what follows, it will often be convenient to use the notation (I_1, I_2) in place of (I, J). Generalizing the expression for ρ to the case in which there are multiple values for the stellar mass,

$$\rho(r) = \frac{1}{4\pi^2 r} \int_r \frac{dI_1 dI_2}{I_2}\ \Omega_1 \sum_q m_q f_q(I_1, I_2). \tag{2.22}$$

where subscript q refers to a particular stellar species (*e.g.,* if the lowest-mass stars in the simulation are of solar mass, then $m_1 = 1M_\odot$ and those stars' distribution function is f_1). The ranges of integration of both of the I_j are the ranges of energetically allowed values, and so depend on $E(I_1, I_2)$ and on $\Phi(r)$.

2.4 Rotational Velocity and Orbital Inclination

It is possible to deduce an expression for the rotational velocity $\varpi(r)$ of a given stellar orbit using only the velocity dispersion and the assumption of velocity isotropy. However, such a relation would depend more heavily on the assumption of isotropy than does anything else in this study, and in the end it does not yield any insight into the distribution of orbits over inclination angle β (where $\cos\beta = J_z/J$), which will also be needed later. Instead, we implement a procedure for first finding the distribution over β and use that to determine ϖ.

2.4.1 Orbital Inclination

Even in a spherically-symmetric gravitational potential, with an overall rotational component to the stellar velocity distribution some quantities will depend on the resulting distribution of orbital inclination. It is convenient to parameterize the inclination dependence of f as having factors that are even and odd in $\alpha \equiv \frac{\pi}{2} - \beta$:

$$h(\alpha) = g(|\alpha|) + \Theta(\alpha) \tag{2.23}$$

where $\int_{-\pi/2}^{\pi/2} d\alpha\, h(\alpha) = \int_{-\pi/2}^{\pi/2} d\alpha\, g(|\alpha|) = 1$ and $\Theta(-\alpha) = -\Theta(\alpha)$. Thus defined, $h(\alpha(\beta))$ can then be applied as a weighting function in averaging any β-dependent quantity, which is essentially a way of defining an effective distribution function:

$$f_\beta(I, J; \beta) \equiv h(\alpha) f(I, J) = [g(|\alpha|) + \Theta(\alpha)] f(I, J). \tag{2.24}$$

Without any prior knowledge of the rotational state of a stellar system other than the net amount of rotation, DeJonghe ([27], [48]) has shown that the statistically most probable form for $\Theta(\alpha)$ is

$$\Theta(\alpha) = \tanh\left(\frac{bJ_z(\alpha)}{2}\right) g(|\alpha|) \tag{2.25}$$

where b is a parameter determined by the total rotational angular momentum. Other than being an even function of its argument, so far $g(|\alpha|)$ is unconstrained. A particularly useful choice is to assume that for a given J, all possible values $0 \leq |J_z| \leq J$ are equally likely [30], consistent with the assumption of isotropy. This is equivalent to setting $g(|\alpha|) = \frac{1}{2}\cos\alpha$:

$$h(\alpha) = \frac{1}{2}\left[1 + \tanh\left(\frac{bJ\sin\alpha}{2}\right)\right]\cos\alpha \tag{2.26}$$

or, rewritten in terms of inclination angle β,

$$h(\beta) = \frac{1}{2}\left[1 + \tanh\left(\frac{bJ\cos\beta}{2}\right)\right]\sin\beta. \tag{2.27}$$

It remains to determine the value of b. For notational convenience we define mass-weighted distribution functions, *i.e.,* $\tilde{f} \equiv \sum_q m_q f_q$ and similarly $\tilde{F} \equiv \sum_q m_q F_q$. Then we define the "population" of a given orbital energy as

$$P(E) \equiv \int_{J_{\min}}^{J_{\max}} dJ\, \tilde{F}(E, J) \tag{2.28}$$

and the total angular momentum for a given energy as

$$J_{\text{tot}}(E) \equiv \int_{J_{\min}}^{J_{\max}} dJ\, J\, \tilde{F}(E, J). \tag{2.29}$$

For a distribution that has no J-dependence (and hence, no rotational streaming), the expected total angular momentum is simply

$$\overline{J}(E) = \frac{1}{2}(J_{\max} + J_{\min}) P(E) \tag{2.30}$$

and so one can deduce the excess of angular momentum at a given E:

$$\Delta J(E) \equiv J_{\rm tot}(E) - \overline{J}(E) \tag{2.31}$$

so that $\Delta J(E) = 0$ for no net rotation. This allows for the following scheme for determining b: for a given orbital energy E, choose $b(E)$ such that the resulting amount of net rotation is consistent with $\Delta J(E)$:

$$\begin{aligned} \Delta J(E) &= \int dJ\, \tilde{F}(E,J) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\alpha\, J_z\, h(\alpha; E) \\ &= 2\int dJ\, J\, \tilde{F}(E,J) \int_{0}^{\frac{\pi}{2}} d\alpha \sin\alpha\, \Theta(\alpha; E) \\ &= \int dJ\, J\, \tilde{F}(E,J) \int_{0}^{1} d(\sin\alpha) \sin\alpha\, \tanh\left(\frac{bJ\sin\alpha}{2}\right) \\ &= \int dJ\, J\, \tilde{F}(E,J) \left(\frac{2}{bJ}\right)^2 \int_0^{bJ/2} dx\, x\, \tanh x \end{aligned} \tag{2.32}$$

which can be expanded to read

$$\int dJ\, \tilde{F}(E,J) \left[J - \frac{1}{2}(J_{\rm max} + J_{\rm min}) - J\left(\left(\frac{2}{bJ}\right)^2 \int_0^{bJ/2} dx\, x \tanh x\right)\right] = 0. \tag{2.33}$$

Root-finding over the above integral equation is a computationally expensive process; before resorting to that, we first attempt an iterative solution by expanding $x\tanh x$ for small $x < \frac{\pi}{2}$ and solving the resultant equation for the leading term in b, using the previous iteration's b value in the higher-order terms. If this does not converge we do rootfind on the exact expression above. For larger values of $bJ/2$, *i.e.,* $bJ/2 > \frac{\pi}{2}$ the integral over x itself asymptotically approaches $\frac{1}{2}$, simplifying the process. In all cases the resulting $b(E)$ is then verified to satisfy (2.32); with this b, (2.27) then describes the orbital inclination distribution of orbits with energy E.

2.4.2 The Mean Rotational Velocity

Further defining the population p at a given r as

$$p(r) \equiv \int_r dI dJ\, \tilde{f}(I,J) \tag{2.34}$$

finding the mean rotational velocity ϖ at a given r is then straightforward:

$$\varpi(r) = \frac{1}{p(r)} \int_r dI dJ\, \tilde{f}(I,J) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\alpha\, \frac{J_z}{r}\, h(\alpha(E(I,J))) \tag{2.35}$$

$$= \frac{1}{r\, p(r)} \int_r dI dJ\, J\, \tilde{f}(I,J) \left(\frac{2}{bJ}\right)^2 \int_0^{bJ/2} dx\, x\, \tanh x \tag{2.36}$$

in which $b = b(E(I,J), J)$ as found in §2.4.1. Descriptively, $\varpi(r)$ is the mean speed at which stars stream in the azimuthal direction at a given distance from the system center, and is found by averaging J_z/r over all inclination angles β as weighted by the relative density $h(\beta)$.

2.4.3 The Coordinate Grid

The self-consistent gravitational potential $\Phi(r)$ needs to be updated after each time step in the simulation, through the iterative solving of Poisson's equation. Up to this point, the dynamics – including the calculation of the orbital inclination distribution $h(\beta)$ – have assumed a spherical gravitational potential $\Phi(r)$ and (smoothed) density $\rho(r)$, *i.e.,* that the cluster's overall ellipticity $e = 0$. This is done simply to allow the construction of a self-consistent model using the two-dimensional stellar distribution function $f(I, J)$, as adding a third integral would be computationally prohibitive.

With the present definitions of dynamical quantities and construction of $h(\beta)$, setting $e = 0$ is mathematically proper [27] and yields a self-consistent model. It is possible that $e \neq 0$ may be more physically motivated, although N-body studies do show that "spherical stellar systems can rotate very rapidly without becoming oblate" [23]. In order to see how much effect the assumption (by construction) of sphericity has, we can add in allowance for ellipsoidal isodensity surfaces in an *ad hoc* manner, using the presumption that the first-order effects of $e(r) \gtrsim 0$ on bulk properties (*e.g.,* anything that depends directly on ρ or on Φ) will outweigh the (small, for small e) errors it introduces into the dynamics proper.[3] Thus unless otherwise specified, ellipticity is pre-defined as $e \equiv 0$ for all results – although provision for the future consideration of the case with a preset $e(r) \neq 0$, which may be allowed to vary along with the coordinate grid values, is built into the model and is the basis for the development presented in §2.5 of the gravitational potential and in Appendix A.2 of numerical aspects of the model.

Thus in the general case there will be a ellipticiy $e(a)$ associated with a given isodensity surface at ellipsoidal radius a; this must be taken into account when calculating quantities that depend on the local position.[4] Equating the exterior value of the potential $d\Phi_e(a)$ of an ellipsoidal shell of mass dM with the analogous value $d\Phi_0(r)$ for a spherical shell of identical mass one obtains

$$-\frac{G\,dM \sin^{-1} e}{ae} = -\frac{G\,dM}{r} \tag{2.37}$$

$$a = r\left(\frac{\sin^{-1} e}{e}\right) \simeq r(1 + \frac{1}{6}e^2) \tag{2.38}$$

where the approximation is used throughout for very small ($< 2 \times 10^{-3}$) values of e. From here on, r can be taken as a "dummy" variable, used for convenience to label gridpoints; the shell's semimajor axis length a from (2.38) is the physically-meaningful radial coordinate.

In cylindrical coordinates, $a^2 = R^2 + \frac{z^2}{1+e(a)}$. The a^2 grid is allowed to dynamically update with each time step, meaning that the values of a^2 chosen to be grid points can change at each timestep so that regions of larger $d\Phi/da^2$ and $d\rho/da^2$ are given a greater density of grid points [49]. When $e = 0$ then $a \equiv r$; this is the case for the majority of simulations

---

[3] The only *inherently* non-spherical aspect of the model's potential or density, namely a "bar perturbation" in the stellar distribution's potential, will be considered in §3.6.3.

[4] Note that, despite allowing for a net overall rotation in the stellar population, the procedure for calculating the non-constant distribution of orbital inclinations $h(\beta)$ does not alter the overall energy balance.

studied here, and for those the coordinate grid is referred to as the "r^2 grid" and not as the "a^2 grid". The numerical consequences of having a finite number of coordinate gridpoints are discussed in §4.7.2.

2.4.4 The Velocity Dispersion

The one-dimensional root-mean-squared (*"rms"*) stellar velocity dispersion σ_o at a given radius a or r is found similarly to ϖ:

$$\sigma_o^2 = (1-\delta)\sigma_1^2 = \frac{(1-\delta)}{p(r)} \int_r dIdJ\, \tilde{f}(I,J)[2(E - \langle\Phi\rangle_a) - \langle v_T^2\rangle_a] \tag{2.39}$$

where δ is an anisotropy parameter that allows for a difference between the radial velocity dispersion σ_1, and σ_o. Subscript "a" on bracket $\langle\cdot\rangle_a$ denotes the possible averaging[5] of the contained quantity over the surface of the isodensity ellipsoid with semimajor axis a. To wit, note that (2.39) assumes when $\delta = 0$, σ_o is identical to $\sigma_1 \equiv \langle v_r^2\rangle^{1/2}$. This is necessary given that the restriction of working with a two-dimensional $f(I_1, I_2)$ means there is not full information on the three-dimensional velocity distribution, and so we must rely on an assumption regarding the degree of velocity anisotropy in order to calculate σ_o^2. It is also why it is not advisable to depend on this calculation of σ_o^2 when determining either the distribution of inclinations $h(\beta)$ or the amount of stellar rotation; those quantities are of central importance to the overall model whereas σ_o is less so.

The expression (2.39) gives the velocity dispersion for all stellar mass species considered together, but it is just as easy to perform the same calculation for a particular stellar mass value by restricting the sums in $p(r)$ and $\tilde{f}$ to that single species.

2.5 Gravitational Potential

As the simulation progresses, new values for the stellar density distribution and the gravitational potential must be determined at each timestep in turn. The new potential at each point on the radial-coordinate grid at a given timestep can be calculated in a manner analogous to that used by Cohn [44]. We write the (inverted) Poisson equation

$$\Phi^{\text{new}} = \mathcal{L}^{-1}\rho[f^{\text{new}}; \Phi^{\text{new}}] \tag{2.40}$$

where $\mathcal{L}^{-1}$ is the inverted Laplacian, derived below. The "new" superscript denotes that, starting with a newly-updated stellar distribution function f, (2.22) and then (2.40) are used to find self-consistent values for ρ and Φ. Note that the allowed ranges of I_1 and I_2 in (2.22) depend on $\Phi(R,z)$, hence the implicit dependence of ρ on Φ in (2.40). For the first step in the interation $\Phi^{\text{new}} = \Phi^{\text{old}}$ is typically used in the right hand side of (2.40); if this fails to converge a guess for Φ^{new} based on a comparison of previous timesteps' potential functions is also attempted.

As done by Cohn [44], we employ the "Aitken δ^2" process [50] to accelerate convergence. This procedure is iterated until convergence is achieved, typically to an accuracy of $\sim 2\%$

---

[5] For $\langle v_T\rangle_a$, simple geometry yields $\langle\frac{1}{r^2}\rangle_a = \frac{2}{3a^2} + \frac{1}{3a^2(1-e^2)} = \frac{a^2}{2}[1 + \frac{e^2}{3(1-e^2)}]$. The average value of the gravitational potential $\langle\Phi\rangle_a$ over the ellipsoid surface must be calculated numerically.

for every point in the Aitken-processed a^2 grid. A final check of conservation of overall energy is also made.

When the gravitational potential Φ is taken to be spherically symmetric – as is the case for most of the simulations performed – the procedure is conceptually straightforward, if numerically complex. The mass density $\rho[f^{\text{new}}; \Phi^{\text{old}}]$ is updated from the new distribution function using (2.22) and then the solution of the Poisson equation is simply

$$\Phi^{\text{new}}(r) = -4\pi G \int_r^\infty \rho \, \frac{dr}{r} \tag{2.41}$$

after which $\rho[f^{\text{new}}; \Phi^{\text{new}}]$ and Φ^{new} are iteratively solved for until convergence is achieved.

For the more general, but rarely needed case in which there is an overall and variable ellipticity $e(a)$ to the potential, a somewhat more complex procedure than given above is required. One still starts with the inverted Poisson equation as given in (2.40) but now the derivation of the inverted Laplacian $\mathcal{L}^{-1}$ is more involved. Details are given in the Appendix.

2.6 Initial Conditions

It is standard in stellar-dynamical work to first consider a Plummer sphere distribution [44] before moving on to other possible potential-density pairs. The Plummer sphere is the reference potential for this study: it fits the light curves of globular clusters [51]. An alternative choice of initial potential which fits the surface brightness of galactic spheroids is the "$\gamma = 0$" model [52]. These two choices have the shared advantage of providing well-defined values for the initial distribution function $f(I, J)$ in addition to the potential $\Phi(r)$ and density $\rho(r)$.

To introduce rotation into either of these density-potential pairs requires an *ad hoc* alteration of the distribution function as described below in §2.6.2, which is followed by descriptions of the stellar bar model and the possible choices for the stellar initial mass fuction (IMF).

2.6.1 Potential-Density Pairs

Base Model: The Plummer Sphere

Previous studies of dense-cluster dynamics have often started with a Plummer sphere ([7], [23], [53], [41]), going as far as refering to nonrotating single-mass Plummer spheres as a "standard testbed" [54]. Following Quinlan and Shapiro [1], the gravitational potential and stellar mass density for a Plummer model of total mass M are

$$\Phi(r) = -\frac{GM}{(r^2 + r_{\text{core}}^2)^{1/2}} \tag{2.42}$$

and

$$\rho(r) = \frac{3M}{4\pi r_{\text{core}}^3}(1 + r^2/r_{\text{core}}^2)^{-\frac{5}{2}}. \tag{2.43}$$

The corresponding energy-space distribution function for mass species m_q is

$$\bar{F}_q(E) = \frac{24\sqrt{2}N_q}{7\pi^3 G^5 M^5} r_{\rm core}^2 E^{7/2}. \tag{2.44}$$

where N_q denotes the total number of stars of individual mass m_q. Note that a general non-rotating distribution function is not necessarily independent of J, but the Plummer model's is.

In particular the specific models first employed by Quinlan and Shapiro [1], such as those listed in Table 2.1, have also been commonly used (*e.g.,* as done by Rasio *et al.* [18]) for later studies of nonrotating systems, and so are modified for use here as described below in §2.6.2. The models in Table 2.1 have two desirable properties: the initial relaxation time $t_r < 1 \times 10^3$ Myr so the total collapse time is expected to be $\lesssim 1 \times 10^4$ Myr; and the escape velocity satisfies $\sigma_1 < v_{\rm esc} \simeq 600$km/s and so stellar collisions result in coalescence of the bulk of the combined stellar material, not disruption and dissipation of it [7]. The naming convention is that models with the same middle digit in their name share a common value for initial velocity dispersion $\sigma_o(0)$, while the final letter indicates the initial central relaxation time $t_r(0)$.

Relating the models to astronomical systems, model E1B best describes globular clusters and is not studied extensively here. E2A and E2B could be very-dense globular clusters [55], or the nuclei of bulgeless spiral and dwarf elliptical galaxies (as shown in Freitag *et al.* [56], [57]). The densities of models E4A and E4B match those of massive young stellar clusters [56]. The cores of giant elliptical galaxies are observed to have line-of-sight velocity dispersions $\sigma_{\rm LOS}$ near 350 km/s [58]; this is not quite a match for the $\sigma_o = 400$ km/s of models E4A and E4B. However, near the observed centers of galaxies $\sigma_{\rm LOS}$ is more directly comparable to σ_1 (as given by 2.39) and it will be shown in §5.2.2 that for model E4B, $\sigma_1(0) \simeq 340$ km/s, a good match with the observations of giant elliptical galaxy cores. Previous simulations of nonrotating systems have found that a very massive star – identified as a possible precursor to an intermediate-mass black hole – can form in a system with a velocity dispersion of "many hundreds of km/s" [19], so the high value for the velocity dispersion should not be an impediment.

Alternative Model: the "$\gamma = 0$" Sphere

The potential-density pair for the "$\gamma = 0$" sphere is

$$\Phi_\gamma(r) = -\frac{GM}{2r_{\rm core}}\left[1 - \frac{r^2}{(r + r_{\rm core})^2}\right] \tag{2.45}$$

$$\rho_\gamma(r) = \frac{3M}{4\pi}\frac{r_{\rm core}}{(r + r_{\rm core})^4} \tag{2.46}$$

with distribution function

$$\bar{F}_q(E) = \frac{3MN_q}{2\pi^3(GMr_{\rm core})^{3/2}}\left((2\epsilon)^{\frac{1}{2}}\frac{3-4\epsilon}{1-2\epsilon} - 3\sinh^{-1}\sqrt{\frac{2\epsilon}{1-2\epsilon}}\right), \quad \epsilon \equiv -\frac{Er_{\rm core}}{GM}. \tag{2.47}$$

Of other possible models, note that the potential of Jaffe's [59] model, and the density of Hernquist's [60] model, diverge as $r \to 0$ and so are inappropriate – or at least inconvenient – for use here.

Carrying over values of M and r_{core} from Table 2.1 to the case of the $\gamma = 0$ sphere, only one resulting model has central density $\rho(0)$ and velocity dispersion σ_1 in the required ranges: "G2A", *i.e.,* a $\gamma = 0$ sphere with the same values for M and r_{core} as model E2A. If instead of equating r_{core} one uses common values of the half-mass radius $r_{1/2}$, a potential model "G3C" results with core radius $r_{\text{core}} = 0.20$ pc. Model G3C has the same total mass $M = 1.1 \times 10^7 M_\odot$ and the same half-mass radius $r_{1/2}$ as model E2B, albeit with a very large initial central density of $\rho(0) = 1.4 \times 10^9 M_\odot/\text{pc}^3$; this model is likely of most use as a demonstration case rather than as a representation of a realistic astronomical system. However an intermediate model with $r_{\text{core}} = 0.29\text{pc}$ (*i.e.,* half that of model E2B's $r_{\text{core}} = 0.58\text{pc}$) features a less extreme initial density $\rho(0) = 5.0 \times 10^8 M_\odot/\text{pc}^3$. This model has an initial relaxation time similar to that of model G2A but a somewhat larger initial central velocity dispersion and is thus labelled "G3A".

Numerically calculated values for velocity dispersion σ_1, density ρ and relaxation time t_r for these $\gamma = 0$ sphere models are given with the overall simulation results in §5.3.

2.6.2 Introducing Rotation

As described in §1.1.3, rotation due to cosmological tidal torques is ubiquitous in the universe and is only weakly dependent on how a particular system was formed. Barnes and Efstathiou [21] considered various models for formation of objects in the early universe, and in terms of Peebles' dimensionless spin parameter

$$\lambda \equiv \frac{J_{\text{rot}} |W_{\text{grav}}|^{\frac{1}{2}}}{G M^{2.5}} \tag{2.48}$$

they determined that the typical value due to cosmological tidal torques is $\lambda \simeq 0.05$, almost independent of the perturbation spectrum. At the stellar cluster scale required here this is consistent with the measurement by Cervantes-Sodi *et al.* [61], who found for a sample of SDSS galaxies that $\lambda = .04 \pm .005$ with a weak trend of increasing λ with decreasing mass.

In order to introduce rotation into the simulations, we alter the distribution of $F_q(E)$ so that the total angular momentum of the cluster J_{rot} produces the desired value for λ, while conserving N_q. This results in a "tilted" $F(E, J)$ in which the resulting excess total amount of J is attributed to J_{rot}:

$$\frac{1}{(J_{\text{hi}} - J_{\text{lo}})} \int_{J_{\text{lo}}}^{J_{\text{hi}}} dJ \, J F(E, J) = \left(1 + \frac{J_{\text{rot}}}{J_{\text{tot}}(E)}\right) \bar{F}_q(E) \tag{2.49}$$

with the tilted distribution function

$$F_q(E, J) = (1 + Q_{\text{rot}}) \bar{F}_q(E) \tag{2.50}$$

Model	$\rho(0)$ $[M_\odot/\mathrm{pc}^3]$	$\sigma_o(0)$ [km/s]	M $[M_\odot]$	r_{core} [pc]	$t_r(0)$ [yr]	Corresponding Astronomical object
E4A	3.0×10^8	400	1.8×10^7	0.24	4.6×10^7	core of a
E4B	1.0×10^8	400	3.1×10^7	0.42	1.4×10^8	giant elliptical galaxy
E2A	4.0×10^7	200	6.2×10^6	0.33	4.6×10^7	nucleus of dwarf elliptical
E2B	1.3×10^7	200	1.1×10^7	0.58	1.4×10^8	or bulgeless spiral galaxy
E1B	1.8×10^6	100	3.6×10^6	0.79	1.4×10^8	globular cluster

Table 2.1: The set of initial condition models. For purposes of the numerical simulation, the quantities taken as fundamental parameters are total mass M and core radius r_{core}. The central density $\rho(0)$, velocity dispersion $\sigma_o(0)$ and relaxation time $t_r(0)$ listed here are the resulting analytic values for a Plummer sphere distribution. Model names are taken from Quinlan and Shapiro [1].

in which the amount of tilt $Q_{\mathrm{rot}}(E, J)$ is given by

$$Q_{\mathrm{rot}}(E, J) = \frac{(\zeta + 1) J_{\mathrm{rot}}}{\langle J \rangle_E \, J_{\mathrm{tot}}} \left(\frac{2}{\Delta J} \right)^{\zeta - 1} (J - \langle J \rangle_E)^\zeta \, . \tag{2.51}$$

In short, this procedure simply replaces $\bar{F}_q(E)$ with a new $F_q(E, J)$ that conserves number and gives the same total angular momentum, but that has a non-flat J-dependence of the form $(J - \langle J \rangle_E)^\zeta$ which results in a net total rotational angular momentum J_{rot}.

Because the (I, J) grid (on which the Fokker-Planck coefficients are calculated) is square but the corresponding (E, J) grid (representing the space on which dynamical quantities must be calculated) is not, we must take care in how ΔJ is determined in the rotating case. In the above, $\langle J \rangle_E = \frac{1}{2}(J_{\mathrm{hi}}(E) + J_{\mathrm{lo}}(E))$ is the average value of J on the grid for a given E. Similarly, ΔJ is the span of J values over which $\bar{F}_q(E)$ is being tilted; to strictly satisfy (2.49) one would take it to equal $J_{\mathrm{hi}}(E) + J_{\mathrm{lo}}(E)$ but in practice this results in values of E which do not span a large range of J being "overloaded" with more than their share of rotation. Thus instead[6] we simply set $\Delta J = (J_{\mathrm{max}} - J_{\mathrm{min}})$. Polynomial power $\zeta \geq 1$ is a free parameter that sets the shape of the tilt. We usually apply a linear tilt $\zeta = 1$.

The new rotating distribution $F(E, J)$ is no longer a solution of the Poisson equation that matches the Plummer potential and density profiles given by (2.42) and (2.43). This "problem" is easily overcome by using the mechanisms of the main simulation itself, as described earlier in this chapter, to find the $\Phi(r)$ and $\rho(r)$ that do correspond to the new $F(E, J)$, before starting the simulation proper based on these initial conditions.

2.6.3 Initial Mass Function

The spectrum of masses which is input to the model is expected to play a large role in its dynamical evolution [54]. A simple power-law $dN/dm \propto m^{-\alpha}$ represents perhaps the

---

[6] The cost of this change is that the resulting value for λ of the newly-determined distribution $F(E, J)$ is slightly lower than intended. This would happen in any case however, as the new distribution also has a lower value for $|W_{\mathrm{grav}}|$ than the original $\bar{F}_q(E)$ did, due to the additional overall rotation. All values for λ quoted in Chapter 5 are "true" values in that these two effects have been taken into account by recalculating (2.48).

simplest form for the initial mass function (IMF). The traditional Salpeter IMF has $\alpha = 2.35$. Other observationally-determined IMFs include the Miller-Scalo [54]:

$$\frac{dN}{dm} \propto \begin{cases} m^{-1.4}, & 0.1 < m/M_\odot < 1.0 \\ m^{-2.5}, & 1.0 < m/M_\odot < 10 \\ m^{-3.3}, & 10 < m/M_\odot \end{cases} \tag{2.52}$$

and the Kroupa IMF [62]:

$$\frac{dN}{dm} \propto \begin{cases} m^{-1.3}, & 0.08 < m/M_\odot < 0.5 \\ m^{-2.2}, & 0.5 < m/M_\odot < 1.0 \\ m^{-2.7}, & 1.0 < m/M_\odot \end{cases} \tag{2.53}$$

A problem with the above IMFs is that none was determined for the specific case of dense clusters near centers of galaxies. However conditions in the centers of galaxies are sufficiently different from those elsewhere that all but the highest-density molecular clouds (number density $n > 10^4/\text{cm}^3$) will be shredded by tidal forces [12]. Recent observational work on such IMFs – most notably for the Arches cluster located 30pc from the center of the Milky Way, as well as the Quintuplet and Central clusters, and R136 in the Large Magellanic Cloud, has led to the following picture of the IMF in the centers of galaxies (Figer [12], [63], Oey and Clarke [64], and references therein, but see Elmegreen [65] for a dissenting opinion):

- galaxy-center IMFs are top-heavy, with $\alpha \simeq 1.7 - 1.8$ for $M \gtrsim (6-10)M_\odot$, or $\simeq 1.8 - 1.9$ when differential extinction is taken into account;
- stellar masses below $\sim 2M_\odot$ do not add appreciably to the total mass of the stellar cluster;
- there is a fundamental, initial upper mass cutoff of $\sim 150M_\odot$;
- central densities can reach $\simeq 10^6 M_\odot/\text{pc}^3$.

Notable exceptions to the high-mass cutoff include the Pistol Star and its "twin" FMM362 in the Quintuplet cluster; despite a cluster age of at least 4 Myr these stars have masses in the $150 - 200M_\odot$ range with an expected stellar lifetime of 2.5 – 3 Myr. One explanation is that these stars are the product of stellar mergers of stars in the $100M_\odot$ range [12].

Thus an IMF for the Arches cluster, representative of dense clusters near the centers of galaxies, can be expressed as

$$\frac{dN}{dm} \propto \begin{cases} m^{-1}, & 2 \lesssim m/M_\odot < 8 \pm 2 \\ m^{-1.8\pm0.1}, & 8 \pm 2 < m/M_\odot \lesssim 150 \end{cases} \tag{2.54}$$

and a simplified version using somewhat overly-conservative choices of parameters is

$$\frac{dN}{dm} \propto m^{-1.9}, \quad 2 < m/M_\odot < 150. \tag{2.55}$$

It is the above IMF that is used for the "Arches" simulations in this study. Some other cases here use the Salpeter IMF of $\alpha = 2.35$ but most start with the Kroupa IMF which for a lower-mass bound of $1M_\odot$ corresponds simply to $\alpha = 2.7$; both of these were also used by Amaro-Seoane [41] for an analogous study of dense but nonrotating clusters which had already formed massive central objects, *i.e.,* black holes or supermassive stars.

With all relevent dynamical quantities now calculable, the Fokker-Planck equation's coefficients for the system can now be determined; the validity of the overall simulation technique and its various components tested; and the actual simulations performed using the initial conditions given above. These steps comprise the next three chapters respectively.

Chapter 3

Derivation of the Fokker-Planck Diffusion Coefficients

3.1 The Orbit-Averaged Fokker-Planck Equation

As described in Chapter 2, it is not advantageous to use energy E and angular momentum J (or its vertical compenent J_z) for the integrals of the motion here as prior two-dimensional, spherically-symmetric studies have done [44]; any change in the gravitational potential of the system will cause changes in the distribution function when it is expressed as $F(E, J)$ or as $F(E, J_z)$. What is required are the radial action I_1 and tangential action $I_2 \equiv J$, which are adiabatic invariants as the potential evolves – and so also is $f(I_1, I_2)$ given the assumption of weak encounters. As given by Binney and Tremaine [42], orbit-averaging is then a simple averaging of quantities over the 2π change in (an) angle variable. It is customary to average over the radial angle. The orbit-averaged Fokker-Planck equation then takes a particularly simple form:

$$\frac{\partial}{\partial t} f(I_i, t) = -\frac{\partial}{\partial I_j}[fC_j] + \frac{1}{2}\frac{\partial}{\partial I_i \partial I_j}[fD_{ij}] \tag{3.1}$$

where $f(I_i, t) = f(I_1, I_2, t)$ is the distribution function and C_j, D_{ij} are the drift and diffusion coefficients. Summation over repeated indices i and j is implied.

Physically, the Fokker-Planck equation can be understood as a collisionless Boltzmann equation $df/dt = 0$ with a collisional term added[1] to account for particles (here, stars) scattering in and out of a given volume of phase space. When the collisional term is expanded in a Taylor series and truncated after the second order term, the remaining two terms are the drift and diffusion coefficients; as stated by Binney & Tremaine [42], they describe the expected rate of change in I_i or $I_i I_j$ for a given test star with actions (I_i, I_j). When the coefficients satisfy the relations

$$D_{ij} = D_{ji} \tag{3.2}$$

---

[1] "Collisional" here refers to the relatively distant gravitational encounters that dominate the overall dynamical scattering, and is not to be confused with the close-range "collisional mergers" of stars which are treated in §3.8.

and

$$C_j = \frac{1}{2}\frac{\partial}{\partial I_i} D_{ij} \tag{3.3}$$

then the Fokker-Planck equation can be recast in a flux-conservative form

$$\frac{\partial}{\partial t} f(I_i, t) = \frac{1}{2}\frac{\partial}{\partial I_j}[D_{ij}\frac{\partial}{\partial I_i} f]. \tag{3.4}$$

It will be seen below that this flux-conservative form has a distinct advantage when it comes to finite differencing.

The Fokker-Planck approximation remains valid as long as number $N = \int dI_1 dI_2 f \gg 1$, and requires that the time step Δt satisfy $t_{dyn} \ll \Delta t \ll t_r$. We have a dynamical time (as defined by Tremaine and Weinberg [30]) of $t_{dyn} \equiv \sqrt{3\pi/16G\rho} \lesssim 10^4$ yr, and a relaxation time of $t_r > 10^6$ yr or more, and so in practice finding an appropriate size of timestep Δt is not difficult.[2]

Generalizing to the case of multiple distribution functions f_q and adding *ad hoc* terms L_q, G_q to account for losses and gains due to stellar mergers, as well as analogous parameters B_q and R_q for stellar births and deaths ("remnants") due to stellar evolution, the most-general Fokker-Planck equation in this study is

$$\frac{\partial}{\partial t} f_q(I_i, t) = \frac{1}{2}\frac{\partial}{\partial I_j}[D_{ij}\frac{\partial}{\partial I_i} f_q] - L_q + G_q - B_q + R_q \tag{3.5}$$

in which subscript q refers to a given subpopulation of stars labeled "q", usually distinguished by being of mass m_q.

The remainder of this chapter develops the physics of the various terms on the right hand side of (3.5). First the general perturbation potential is converted into action space for use in the diffusion coefficients D_{ij}. The specific case of field-star perturbers is developed, which is then used to derive expressions for D_{ij}. The other relevent perturbation potential, that of a stellar bar, is then constructed and its diffusion coefficients also developed. How to finite-difference the Fokker-Planck equation using these diffusion coefficients follows, and expressions for the stellar merger loss and gain terms L_q and G_q are then derived. The chapter closes with a discussion of the possible effects of binary heating on the model and the reasons for not including it in the calculation; despite its omission from the Fokker-Planck model, the amount of possible binary heating that could have occured is still tracked througout the simulations in order to confirm that the reasons for omitting it remain valid.

3.2 The Perturbing Potential

3.2.1 General form of the expansion

The drift and diffusion coefficients describe the rate of transfer of action (momentum) to a star due to interactions with other individual stars or with a bulk stellar bar perturbation.

---

[2] For a given test star, the relaxation time is given by $t_r = \max_{ij}(\frac{I_i I_j}{D_{ij}})$. In Chapter 5 the median relaxation time [7] of $t_{rq} = \sigma_1^3/[4\pi\sqrt{3/2}G^2 m_q^2 \ln(0.4N)]$ is used instead, in order to indicate the relaxation time of stars of a particular stellar mass in position space, instead of in action-space as considered here.

To do so, first it is necessary to express the perturbing potential Φ_* in action space. This has been done by Tremaine and Weinberg [30]; the following summarizes their results. Starting in traditional spherical coordinates Φ_*, when expanded in spherical harmonics is

$$\Phi_* = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \Phi_{lm}(r) Y_{lm}(\theta, 0)\, e^{im(\phi - \Omega_* t)}. \tag{3.6}$$

(Label "$*$" can refer to any perturbing mass, *i.e.,* either to a stellar bar as a whole or to a single field star.) Expressed in action space, (3.6) becomes

$$\Phi_* = \sum_{m=0}^{\infty} \sum_{k,n=-\infty}^{\infty} \Psi_{knm}(I_1, I_2)\, e^{i(kw_1 + nw_2 + mw_3 - \Omega_* t)} \tag{3.7}$$

where

$$\Psi_{knm} = \sum_{l=0}^{\infty} \left(\frac{2}{1 + \delta_{m0}} \right) V_{lnm}(\beta) W_{klnm}(I_1, I_2) \tag{3.8}$$

and in which β is the perturber's orbital inclination (thus $\beta = 0$ for a bar in the rotational plane of the system). Quantity $V_{lnm}(\beta)$ contains the effect of rotating the frame of reference so that it is aligned with the orbital plane, and $W_{klnm}(I_1, I_2)$ is the strength of the interaction of the (k, n, m) resonance for multipole expansion term l, as will be seen below[3]. In particular, W_{klnm} has the form

$$W_{klnm}(I_1, I_2) = \frac{1}{\pi} \int_0^{\pi} dw_1\, \cos[kw_1 - n(\psi - w_2)] \Phi_{lm}(r) \tag{3.9}$$

in which one may note that angles w_1 and $(\psi - w_2)$ are functions of r as given by (2.6) and (2.7). By inverting (2.6) one obtains $r(w_1)$ and by extension $(\psi - w_2)(w_1)$ as required for use in (3.9). Various symmetries of spherical harmonics and in the expansion are of relevance:

- Φ_* is real, and all perturbers studied have real Φ_{lm} for all (l, m), so $\Phi_{l-m} = (-)^m \Phi_{lm}$;
- the system is symmetric about the $z = 0$ plane, so $(l + m)$ must be even;
- $V_{lnm}(\beta) = 0$ unless $(l + n)$ is even;
- the above two items result in V_{lnm} being real for all (lnm), and so Ψ_{knm} is real for all values of (knm);
- $V_{l-n-m} = (-)^m V_{lnm}$ and $W_{-kl-n-m} = (-)^m W_{klnm}$, and so $\Psi_{-k-n-m} = \Psi_{knm}$.

Note that the choice of radial basis functions $\Phi_{lm}(r)$ is not unique [23]; the current method is most convenient for dealing with inclined orbits while still allowing development of the diffusion coefficients in terms of orbital resonances, below.

---

[3] For completeness' sake, we note here that $V_{lnm}(\beta) = (-)^{(m-n)/2} r_{lnm}(\beta) Y_{ln}(\frac{\pi}{2}, 0)$, where r_{lnm} is the Slater rotation matrix $r_{lnm} = \sum_t (-)^t \frac{\sqrt{(l+n)!(l-n)!(l+m)!(l-m)!}}{(l-m-t)!(l+n-t)!t!(t+m-n)!} \tan^{2t+m-n}(\frac{\beta}{2}) \cos^{2l}(\frac{\beta}{2})$ [30]. The sum is taken over all t such that the factorials' arguments are all positive semidefinite, and similarly $r_{lnm} \equiv 0$ if any of $(l \pm n)$ or $(l \pm m)$ is negative.

3.2.2 Effect of Orbital Inclination

Many of the physical quantities that determine the diffusion coefficients depend either on orbital inclination β and/or the polar angle of the curent position θ. For the former dependence we invoke the orbital inclination distribution $h(\beta)$ developed in §2.4.1. Quantities such as $V_{lnm}(\beta)$ are then averaged over $(0, \beta_{\max})$ with weighting function $h(\beta)$; notationally this will be denoted by angle brackets with subscript $h(\beta)$. Thus for any quantity y

$$\langle y \rangle_{h(\beta)} \equiv \frac{\pi}{\beta_{\max}} \int_0^{\beta_{\max}} d\beta\, h(\beta)\, y(\beta). \tag{3.10}$$

Normally $\beta_{\max} = \pi$, which accounts for both prograde and retrograde orbits. As the β dependence of $V_{lnm}(\beta)$ is strictly trigonometric, if the form of $h(\beta)$ allows it and if no part of Φ_{lm} as contained in W_{klnm} depends on β, the averaging of (3.8) can be done analytically; else it must be done numerically. This will be of help later, when weak approximations will be all that are needed to satisfy the requirements for doing analytical averaging.

A further example is given by the square of (3.8):

$$\overline{\Psi^2}_{knm} \equiv \left(\frac{2}{1+\delta_{m0}}\right)^2 \left\langle \left[\sum_{l=0}^{\infty} V_{lnm}(\beta) W_{klnm}(I_1, I_2)\right]^2 \right\rangle_{h(\beta)}. \tag{3.11}$$

The above $\overline{\Psi^2}_{kmn}$ is what will be used in the diffusion coefficients below. To calculate the average, the squared sum is expanded and evaluated term by term; this is more efficient than averaging over the entire squared sum at once. For convenience, the triple (knm) will be written as $(\ell_1 \ell_2 \ell_3)$ in most sums.

The direct calculation of (3.11) takes an unacceptably long time, largely due to the complexity of the calculation of $\Phi_{lm}(r)$ as will be seen below; it is sped up by evaluating the integrands at preset values of θ and w_1, and interpolating from those values when performing the integrations. Similarly, solving (2.6) for $r(w_1)$ and then integrating (2.7) for $w_2 - \psi$ each time they are needed in (3.9) would be computationally prohibitive, and so they are also calculated on a grid and interpolated for.

The ability to transform away the θ dependence of $Y_{lm}(\theta, 0)$ in going from (3.6) to (3.7) notwithstanding, other quantities that depend not on β but instead on θ directly require a slightly more subtle treatment. The approach here is to average any such quantity over its allowed range of $\frac{\pi}{2} - \beta \leq \theta \leq \frac{\pi}{2}$ before doing the averaging over β itself. Allowing there to also be a purely β-dependent factor y:

$$\langle y; x \rangle_{\theta} = \frac{\pi}{\beta_{\max}} \int_0^{\beta_{\max}} d\beta\, h(\beta) y(\beta)\, \frac{2}{\pi} \int_0^{\frac{\pi}{2}} d\psi\, x(\theta(\psi, \beta)). \tag{3.12}$$

The inner averaging is actually performed over ψ instead of θ because, averaged over all stellar orbits of a given (I, J), $d\psi/dt$ is a constant whereas $d\theta/dt$ is not. (See (2.7) and note that the phases of w_1 and w_2 are independent; Figure 2.1 gives a visual depiction.) To convert between θ and ψ, (2.8) is used.

3.3 Field Star Perturbers

For an individual star of mass m_* (not to be confused with spherical harmonic index m) at position $(r_*,\, \theta_*,\, \phi_* = \Omega_* t + \phi_o)$ in the field, the potential at (r, θ, ϕ) is

$$\begin{aligned} \Phi_*(\mathbf{r};\mathbf{r}_*) &= -\frac{Gm_*}{|\mathbf{r}-\mathbf{r}_*|} \qquad (3.13)\\ &= -Gm_* \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} \frac{4\pi}{(2l+1)} \sum_{m=-l}^{l} (-)^m Y_{lm}(\theta,\phi) Y_{lm}(\theta_*,\phi_*)\\ &= -Gm_* \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} \frac{4\pi}{(2l+1)} \sum_{m=-l}^{l} (-)^m Y_{lm}(\theta,0) Y_{lm}(\theta_*,0) e^{im(\phi-\phi_*)} \end{aligned}$$

with the traditional $r_<$ and $r_>$ respectively being the smaller and larger of r and r_* . Initial phase ϕ_o can be absorbed into the initial phase of canonical angle variable w_3 and so has already been dealt with in the previous section. Comparison with (3.6) thus yields

$$\Phi_{lm}(r;\mathbf{r}_*) = -Gm_* \frac{r_<^l}{r_>^{l+1}} \frac{4\pi}{(2l+1)} (-)^m Y_{lm}(\theta_*,0). \qquad (3.14)$$

The field star distribution is defined as the overall cluster distribution, excluding those stars specifically assigned to a stellar bar (which is labeled "B" to denote that it is a bulk perturber, different from individual stars) and so is expressed in terms of the canonical actions: $f_*(I_1', I_2') = \sum_q f_q(I_1', I_2') - f_{\mathrm{B}}(I_1', I_2')$. (Here, primes indicate actions of the perturbing field star, and not of the stellar orbit being perturbed.)

To deal with the undesired perturber radial coordinate r_*, we perform an additional orbit-averaging of the field star's contribution to the overall perturbing potential as given by (3.14), this time averaging over the perturbing field star's orbital range. It is more straightforward to directly average over r_* weighted by $1/|v_{r*}|$ than to convert to the corresponding canonical angle w_1' first, which would accomplish the identical task but require an additional step. Both $Y_{lm}(\theta_*, 0)$ and v_{r*} depend on θ_* (the latter via $\Phi(\mathbf{r}_*)$) which can at the same time be averaged over using the technique of (3.12):

$$\begin{aligned} \left\langle 1; \langle \Phi_{lm}(r;\mathbf{r}_*)\rangle_{w_1'} \right\rangle_{\theta_*} &= \left\langle 1; \int_{r_{p*}}^{r_{a*}} dr_* \frac{\Phi_{lm}(r)}{|v_{r*}|} \Bigg/ \int_{r_{p*}}^{r_{a*}} \frac{dr_*}{|v_{r*}|} \right\rangle_{\theta_*} \qquad (3.15)\\ = \; -\frac{4\pi G m_*}{(2l+1)} (-)^m & \left\langle 1; Y_{lm}(\theta_*,0) \int_{r_{p*}}^{r_{a*}} \frac{dr_*\; r_<^l / r_>^{l+1}}{[2(E_* - \Phi(\mathbf{r}_*)) - {I_2'}^2/r_*^2]^{\frac{1}{2}}} \Bigg/ \int_{r_{p*}}^{r_{a*}} \frac{dr_*}{|v_{r*}|} \right\rangle_{\theta_*} . \end{aligned}$$

Of course for orbits that are very nearly circular ($r_{p*} \simeq r_{a*}$), r_* is well-defined for any given value of I_1' (noting $I_2' \gtrsim 0$) and so the above averaging over r_* is unnecessary; one can then simply take $\Phi_{lm}(r) = \Phi_{lm}(r; r_*)$. Otherwise the full form of (3.15) is required, and is what is used for the $\Phi_{lm}(r)$ factor in (3.9).

3.4 The Diffusion Coefficients

With knowledge of the perturbation coefficients $\overline{\Psi}_{\ell_1\ell_2\ell_3}$ we can now proceed to compute the drift and diffusion coefficients C_j and D_{ij}. The former were derived by Van Vleck [66] in a quantum mechanical context, although a more related calculation is that of Tremaine and Weinberg [30] who derived the dynamical friction exerted on an external satellite by a stellar system. Our situation is analogous, but requires the generalization of considering both components of the action instead of only I_2. Thus the derivation of D_{ij} here follows a similar form.

Denoting first-order approximations by Δ_1, Hamilton's equations for the system are

$$\Delta_1 I_j = \frac{\partial \chi_1}{\partial w_j}, \quad \Delta_1 w_j = -\frac{\partial \chi_1}{\partial I_j} \tag{3.16}$$

with generating function

$$\chi_1 = -\mathrm{Re}\left[\sum_{\ell_3=0}^{\infty} \sum_{\ell_1,\ell_2=-\infty}^{\infty} \Psi_{\ell_1\ell_2\ell_3} \frac{e^{i(\ell_p w_p - \omega t)}}{i(\ell_p \Omega_p - \omega)}\right] \tag{3.17}$$

in which $\Omega_p = \frac{dw_p}{dt}$, and $\omega = \ell_3\Omega_* + i\eta$ is the frequency of the perturbation potential including a slow "turning on" of the perturber in the distant past effected by $\eta > 0$. Throughout this section repeated index p will indicate summation over all possible values 1,2,3 of the index, and the same character p when used in different factors indicates distinct sums, *e.g.,* $(\ell_p\Omega_p)(\ell'_p\Omega_p) = (\ell_1\Omega_1 + \ell_2\Omega_2 + \ell_3\Omega_3)(\ell'_1\Omega_1 + \ell'_2\Omega_2 + \ell'_3\Omega_3)$. So, to second order

$$\begin{aligned}
\Delta_1 I_i \, \Delta_1 I_j &= \frac{\partial \chi_1}{\partial w_i} \, \frac{\partial \chi_1}{\partial w_j} \\
&= \mathrm{Re}\left[\sum_{\ell_3=0}^{\infty} \sum_{\ell_1,\ell_2=-\infty}^{\infty} \ell_i \overline{\Psi}_{\ell_1\ell_2\ell_3} \frac{e^{i(\ell_p w_p - \omega t)}}{i(\ell_p\Omega_p - \omega)}\right] \mathrm{Re}\left[\sum_{\ell'_3=0}^{\infty} \sum_{\ell'_1,\ell'_2=-\infty}^{\infty} \ell'_j \overline{\Psi}'_{\ell_1\ell_2\ell_3} \frac{e^{i(\ell'_p w_p - \omega' t)}}{i(\ell'_p\Omega_p - \omega')}\right] \\
&= e^{2\eta t} \sum \sum \ell_i \overline{\Psi}_{\ell_1\ell_2\ell_3} \left[\frac{\eta \sin(\ell_p w_p - \ell_3\Omega_* t) - (\ell_p\Omega_p - \ell_3\Omega_*) \cos(\ell_p w_p - \ell_3\Omega_* t)}{(\ell_p\Omega_p - \ell_3\Omega_*)^2 + \eta^2}\right] \\
&\quad \times \sum{}' \sum{}' \ell'_j \overline{\Psi}'_{\ell_1\ell_2\ell_3} \left[\frac{\eta \sin(\ell'_p w_p - \ell'_3\Omega_* t) - (\ell'_p\Omega_p - \ell'_3\Omega_*) \cos(\ell'_p w_p - \ell'_3\Omega_* t)}{(\ell'_p\Omega_p - \ell'_3\Omega_*)^2 + \eta^2}\right]
\end{aligned} \tag{3.18}$$

where for notational convenience, a prime on a quantity indicates that any indices it takes are primed, *i.e.,* $\omega' \equiv (\ell_3'\Omega_* + i\eta)$ *etc.*. To second order, the rate of change of $(\Delta I_i\, \Delta I_j)$ is

$$\begin{aligned}
\frac{d}{dt}[\Delta_1 I_i\, \Delta_1 I_j] &= 2\eta\, [\Delta_1 I_i\, \Delta_1 I_j] + e^{2\eta t} \qquad (3.19)\\
\times &\quad \left[\sum\sum \ell_i \overline{\Psi}(\ell_p\Omega_p - \ell_3\Omega_*) \frac{\eta\cos(\ell_p w_p - \ell_3\Omega_* t) + (\ell_p\Omega_p - \ell_3\Omega_*)\sin(\ell_p w_p - \ell_3\Omega_* t)}{(\ell_p\Omega_p - \ell_3\Omega_*)^2 + \eta^2}\right.\\
&\quad \left.\times \sum{}' \sum{}' \ell_j' \overline{\Psi}' \frac{\eta\sin(\ell_p' w_p - \ell_3'\Omega_* t) - (\ell_p'\Omega_p - \ell_3'\Omega_*)\cos(\ell_p' w_p - \ell_3'\Omega_* t)}{(\ell_p'\Omega_p - \ell_3'\Omega_*)^2 + \eta^2}\right]\\
+ &\quad e^{2\eta t}\left[\sum\sum \ell_i \overline{\Psi} \frac{\eta\sin(\ell_p w_p - \ell_3\Omega_* t) - (\ell_p\Omega_p - \ell_3\Omega_*)\cos(\ell_p w_p - \ell_3\Omega_* t)}{(\ell_p\Omega_p - \ell_3\Omega_*)^2 + \eta^2}\right.\\
\times &\quad \left.\sum{}' \sum{}' \ell_j' \overline{\Psi}' (\ell_p'\Omega_p - \ell_3'\Omega_*) \frac{\eta\cos(\ell_p' w_p - \ell_3'\Omega_* t) + (\ell_p'\Omega_p - \ell_3'\Omega_*)\sin(\ell_p' w_p - \ell_3'\Omega_* t)}{(\ell_p'\Omega_p - \ell_3'\Omega_*)^2 + \eta^2}\right]
\end{aligned}$$

in which for brevity $\overline{\Psi} \equiv \overline{\Psi}_{\ell_1\ell_2\ell_3}$ and $\overline{\Psi}' \equiv \overline{\Psi}'_{\ell_1\ell_2\ell_3}$.

The diffusion coefficient for an individual perturber is $\frac{d}{dt}(\Delta_1 I_i\, \Delta_1 I_j)$, averaged over all initial phases of the w_p; as these are assumed to be distributed randomly, most of the terms in (3.19) vanish or cancel out. Considering each of the three "grand sums" (*i.e.,* each quantity contained in outer sets of brackets in 3.19) separately, this happens by the following reasons:

1. all terms without either $\ell_p = \ell_p'$ for all p, or else $\ell_p = -\ell_p'$ for all p, vanish upon the averaging;
2. all cross terms of the form $[\pm\sin(\ell_p w_p - \ell_3\Omega_* t)\cos(\ell_p w_p - \ell_3\Omega_* t)]$ also vanish upon averaging over (any) initial phase angle;
3. $\ell_p = -\ell_p'$ requires that $\ell_3 = 0$; thus in each grand sum, for each term with a given $\ell_1 = -\ell_1'$ and $\ell_2 = -\ell_2'$ there is an equal-magntitude, opposite signed term with $\ell_1 = \ell_1'$ and $\ell_2 = \ell_2'$, and so all terms with $\ell_3 = 0$ cancel out;
4. in the second and third grand sums, the remaining terms can all be grouped into compound terms having angle dependence of the form $[\cos^2(\ell_p w_p - \ell_3\Omega_* t) - \sin^2(\ell_p w_p - \ell_3\Omega_* t)]$, which also vanishes upon averaging over initial phase.

With the above, the only remaining terms are those from the first grand sum having $\ell_p = \ell_p'$ and $\ell_3 > 0$, and excluding cross-terms. Using angle brackets to denote averaging over initial phases, the reduced form of the averaged perturbation is thus

$$\begin{aligned}
\left\langle \frac{d}{dt}[\Delta_1 I_i\, \Delta_1 I_j] \right\rangle_w &= 2\eta\, e^{2\eta t} \times \qquad (3.20)\\
&\sum_{\ell_3=1}^{\infty} \sum_{\ell_1,\ell_2=-\infty}^{\infty} \ell_i \ell_j \overline{\Psi}^2_{\ell_1\ell_2\ell_3} \left\langle \left[\frac{\eta\sin(\ell_p w_p - \ell_3\Omega_* t) - (\ell_p\Omega_p - \ell_3\Omega_*)\cos(\ell_p w_p - \ell_3\Omega_* t)}{(\ell_p\Omega_p - \ell_3\Omega_*)^2 + \eta^2}\right]^2 \right\rangle_w\\
&= \eta\, e^{2\eta t} \sum\sum \ell_i \ell_j \overline{\Psi}^2_{\ell_1\ell_2\ell_3} \left[\frac{1}{(\ell_p\Omega_p - \ell_3\Omega_*)^2 + \eta^2}\right]\\
&= \eta\, e^{2\eta t} \sum\sum \ell_i \ell_j \frac{\overline{\Psi}^2_{\ell_1\ell_2\ell_3}}{|(\ell_p\Omega_p - \ell_3\Omega_*) - i\eta|^2}
\end{aligned}$$

in which $\langle \sin(\ell x+b)\cos(\ell x+b)\rangle_x = 0$ and $\langle \sin^2(\ell x+b)\rangle_x = \langle \cos^2(\ell x+b)\rangle_x = \frac{1}{2}$ have been used to eliminate the trigonomic factors.

The analogous derivation of the drift coefficients performed by Tremaine and Weinberg [30] found that the effect of the first-order perturbation vanishes, and that to lowest order

$$\left\langle \frac{d}{dt}[\Delta_2 I_j] \right\rangle_w = \frac{1}{2}\eta\, e^{2\eta t} \sum\sum \ell_i \ell_j \frac{\partial}{\partial I_i} \frac{\overline{\Psi}^2_{\ell_1\ell_2\ell_3}}{|(\ell_p\Omega_p - \ell_3\Omega_*) - i\eta|^2}. \tag{3.21}$$

From (3.20) and (3.21) it can be seen that (3.2) and (3.3) do indeed hold; from here on the flux-conservative form of the Fokker-Planck equation (3.4) will be assumed and no further explicit consideration of the drift coefficients is required.

In order to be used in the Fokker-Planck equation, the limit $\eta \to 0$ must be taken, which allows us to use the relation $\delta(x) = \frac{1}{\pi}\lim_{\eta\to 0} \eta|x - i\eta|^{-2}$. Thus, one obtains

$$\left\langle \frac{d}{dt}[\Delta_1 I_i\, \Delta_1 I_j] \right\rangle_w = \pi \sum_{\ell_3=1}^{\infty} \sum_{\ell_1,\ell_2=-\infty}^{\infty} \ell_i \ell_j \overline{\Psi^2}_{\ell_1\ell_2\ell_3} \delta(\ell_p\Omega_p - \ell_3\Omega_*). \tag{3.22}$$

To obtain the full diffusion coefficient one must integrate over the full phase space of the perturber. With $f_*(I'_1, I'_2)$ denoting the distribution function of the perturbers under consideration and invoking the β-averaged expansion coefficients of the perturbing potential from (3.11) one obtains [4]

$$\begin{aligned} D_{ij} &= \int d^2I'\, f_*(I'_1, I'_2) \left\langle \frac{d}{dt}[\Delta_1 I_i\, \Delta_1 I_j] \right\rangle_w \\ &= \pi \int d^2I'\, f_*(I'_1, I'_2) \sum_{\ell_3=1}^{\infty} \sum_{\ell_1,\ell_2=-\infty}^{\infty} \ell_i \ell_j \overline{\Psi^2}_{\ell_1\ell_2\ell_3} \delta(\ell_p\Omega_p - \ell_3\Omega_*). \end{aligned} \tag{3.23}$$

Note that in this formalism the effect of the dynamical friction which underlies the interactions is expressed in terms of resonances between the orbital frequencies of any star and the secular frequency of the perturbing potential (which is either that of the bar, or the sum of all potentials of field stars of a given orbital frequency Ω_*). In practice, for a field star perturber, D_{ij} is evaluated by using the delta-function that describes the resonance to eliminate one of the action integrals and then integrating over the other. For the perturbation due to a stellar bar (for which Ω_* is taken to be a common value for all stars comprising the bar), the resonance delta-function does not integrate out in (3.23) and will have to be handled separately as will be seen below.

3.4.1 Mass Segregation

Because $\overline{\Psi}_{\ell_1\ell_2\ell_3}$ scales linearly with m_*, and f_* is the number-weighted distribution function, one can see that D_{ij} also scales with m_*; this results in the same mass segregation effect as described by Quinlan and Shapiro [7] in which the tendency towards equipartion of energy

---

[4] The expression here differs from the one derived by Tremaine and Weinberg [30] who additionally had an integral over dJ_z, but in our case that's already been taken care of through the averaging over β. Also, f_* here is already "orbit averaged" and no summing over the ranges of w_1 and w_2 is needed.

causes lighter stars to receive kinetic energy from heavier stars, which then preferentially sink towards the center.

Mass segregation is observed in globular clusters, and can be qualitatively described using energy equipartition. The segregation can be unstable in that a population of more-massive stars decouples dynamically from the lower-mass population if no equipartition is possible; whether this happens depends on m'_*/m_* and the ratio of total masses of the two populations [67]. Mass segregation is one example of the complex dynamics that can arise when there is a spectrum of masses in the system.

As a more specific example, McMillan *et al.* [68] calculate that if $m'_* \leq 20M_\odot$, then a subpopulation forms in a globular cluster core in $\simeq 0.2$ of the half-mass relaxation time t_{rh}, and the central density of the larger-mass stars increases by 2-3 orders of magnitude for "typical IMFs"; in denser clusters it is possible for this to occur before many stars evolve to supernovae. Even a modest range of masses can produce enough mass segregation to greatly increases the rate of core collapse [54]). The minimum central density determined by McMillan *et al.* as leading to mass-segregation-driven core collapse in globular clusters is similar to that used in some of the cases studied here.

Is Mass Segregation Primordial, or Dynamic?

Whether mass segregation in stellar clusters is primordial or is the result of dynamical evolution is a matter of some debate: N-body simulations of globular cluster dissolution performed by Baumgarft *et al.* [69] on the specialized GRAPE6 computing engine matched observed properties of globular clusters when an initial mass segregation was assumed and did not not match when no initial segretation was included; they also deduced a "near-universal" mass function for low-metalicity star formation environments. However, using the same "NBODY4" code McMillan and Portegies Zwart [55] found no firm evidence for *a priori* mass segregation in young dense clusters. In Fokker-Planck (nonrotating) models of individual clusters, Amaro-Seoane [41] found no evidence of initial mass segregation while acknowledging that other models for young clusters had predicted it. While none of the above studies looked at exact analogues of the early-universe stellar clusters studied here, the conservative approach is still not to assume any initial mass segregation, so that any mass segregation observed in the simulations is the result of the dynamical friction. Collisional mergers of large stars can also be a source of effective mass segregation, in that mergers should preferentially occur in the more-dense central region of the system.

3.5 The Stellar Bar: General Considerations

As stated in the Chapter 1, although some angular momentum is expected to be transported outwards via the effects of shear between the higher-Ω_2 inner regions and the lower-Ω_2 outer regions, a bulk perturbation in the potential is predicted to be much more effective than individual stars at transporting angular momentum, as indicated by the appearance of the square of the perturbation coefficient $\overline{\Psi^2}_{\ell_1\ell_2\ell_3}$ in the overall Fokker-Planck diffusion coefficient (3.23).

Most work on stellar bars, both observed and modeled, has been on the scale of entire galaxies. But "inner bars" are common in observed samples of galaxies even though the formation and dynamics of nested bars is still poorly understood [28]. Even for galactic-scale bars few studies address the issue of "strength" of the bar, in terms of the overall mass associated with the bar perturbation. Typical values for the strength are in the range of 5-15% of the overall mass available in the relevant volume: for their N-body simulations Weinberg and Katz [33] assumed the bar comprised 15% of the disk and halo mass within the bar's radius (which they took to be half the corotation radius, although their results were not sensitive to this choice); and Athanassoula [32] found in N-body simulations that a stellar bar quickly grew to $\simeq 5\%$ of the mass within the radius of their stellar disk, regardless of whether the potential within that radius was dominated by the disk or by the halo.

In terms of orbital composition, bars are elongated along constituent orbits' long axes [70], and chaotic orbits contribute to the bar's overall makeup [32]. Chaotic orbits are not considered here, but this is in line with stellar dynamics as practiced since the 1940s in which the chaotic effects of N-body interactions are ignored in general [71].[5] Preliminary 3-dimensional modeling of stellar bars by Athanassoula [32] shows orbits similar to, but more general than, those of two-dimensional models. He finds that a halo component in the galaxy's overall gravitational potential does not stabilize the system against forming a stellar bar (*cf.* also [72]), and infers that (as yet undetermined) bar models with non-isotropic distribution functions would be expected to lead to even stronger bar growth.

Dynamically, a "bar perturbation" has typically been taken as being synonymous with the presence of an additional quadrupole ($m = 2$) component in the potential (*e.g.,* Binney and Tremaine [42]), and it will be taken as so here. It is worth noting that Athanassoula [32] has recently found some bars to have relatively strong $m = 4$, and in some cases even $m = 6$ or $m = 8$, components. These modes represent a possible avenue for further study but to lowest order would be expected to simply enhance the rate of transport of angular momentum within the cluster.

3.6 The Stellar Bar: Implementation

Early in the study of bar dynamics, a high central density $\rho(0)$ was argued to preclude bar formation due to the disruptive presence of an inner Linblad resonance (ILR), and the "Toomre" $T_{\mathrm{rot}}/|W| \gtrsim 0.14$ instability criterion [73] was formulated as a test for whether or not such a non-axisymmetric perturbation could form in a system. However, high-$\rho(0)$ barred galaxies with ILRs are observed, and in general bars are observed to be much more ubiquitous than the Toomre criterion allows. Thus the true trigger of bar formation remains unknown [74]. Of particular interest is the finding of Gadotti and de Souza [75] that bars can form in Plummer sphere potentials which are embedded in nonspherical halos but which possess no disk component. Many studies (*e.g.,* Athanassoula [32]) note that bars are very common features.

That being said, all simulations performed here possess $T_{\mathrm{rot}}/|W| \gtrsim 0.17$ and thus satisfy not only the Toomre criterion but also the $T_{\mathrm{rot}}/|W|$ value of 0.171 calculated by Christodoulou *et al.* [76] (which was for stellar Maclaurin spheroids, but was the strictest bar-forming criterion found in the literature). This is not by design, but is a function of using an

---

[5] Gurzadyan [71] goes on to state that it would be expected from ergodic theory that N-body interactions and the ensuing chaotic effects would dominate over 2-body interactions – however, it is emphasized that this applies for the case containing a central massive object, unlike the situation here.

observationally-motivated value for the overall rotation parameter of $\lambda \simeq 0.05$ as discussed in §2.6.2. Thus the bar is simply assumed to exist.

N-body studies show that bars form over 1–2 rotational periods but last for many more ($\gtrsim 10$) [77], and that the stars trapped in the bar do not undergo much individual dynamical evolution [78]. Some N-body simulations show the bar losing strength over time, but this effect has been found not to be due to angular-momentum transfer weakening the bar directly but is instead due to a vertical-buckling instability of the bar in the presence of background gas [79], which does not pertain here. This has led us to build the nonaxisymmetric potential from a fraction of the lowest-mass component of the stellar population, and to force those stars to orbit in "lock step", sharing a common orbital frequency and phase (*i.e.,* it is assumed that they all orbit at a frequency Ω_B whenever the diffusion coefficient calculation requires knowledge of the orbital frequency). We assume there is no transfer of stars from bar to field or vice versa, consistent with studies showing the same stars remain trapped within the bar ([80], [78]). N-body studies show that the bar frequency is a compromise of the orbital frequencies of its component stars [70], so our Ω_B is set by either conserving the total angular momentum of those stars, or by simply averaging over their individual orbital frequencies; each possibility is dicussed below.

Informed by the considerations of §3.5, the most straightforward way to incorporate a stellar bar into our model is to assume that a fraction of the stars are trapped in the bar from the outset, that the resulting nonaxial perturbation in the potential is dominated by the quadrupole term, and that the structure rotates in bulk at a common angular speed. With these assumptions the stars comprising the bar can be considered as simply a subset of the overall stellar population for the purposes of calculating updated values of the distribution function f or stellar density ρ, while for determining the bar's contribution to the diffusion coefficients D_{ij} the bar is treated as a semi-solid object with a single overall rotation speed Ω_B. Allowing the bar distribution to evolve like the field stars avoids any problem of inserting the bar binding energy by hand.

The mass fraction assigned to the bar here is typically 1% of the total system mass, for simplicity taken from the lowest stellar mass in the IMF being used (although the choice of stellar mass in the bar was found to have little or no effect on the simulation's results). Using the definition of bar strength suggested by Weinberg and Katz [33] as the ratio of the bar's mass to the total mass within half of the corotation radius, this gives a bar strength of 5-10% depending on the initial cluster model being considered, well within the 5-15% range of previous studies.

One possible concern is that the general formalism of resonant interactions represented by the diffusion coefficients (3.23) is only valid when the bar's rotation speed Ω_B does not change too slowly [30]; presumably this is because if $d\Omega_\mathrm{B}/dt$ is too small, the back-reactions of the resonant interaction on the bar itself will not be smoothly distributed over frequency space and may disrupt the stellar orbits comprising the bar. In this study a typical value for the timescale of the rate of change of the bar frequency $\Omega_\mathrm{B}/(d\Omega_\mathrm{B}/dt)$ is $\gtrsim 10 - 200$ bar rotation times; this timescale is consistent with the range of Weinberg's [34] calculation for a general bar interacting with a halo and with Athanassoula's [72] N-body bar simulations.

The back-reaction issue may explain why, when a bar mass-fraction of $\gtrsim 2\%$ (corresponding to a bar strength much greater than 10%) was tested, the simulations were unstable to numerical divergences in which the results were not consistent for different choices of numerical stepsize Δt. Backing off to a bar fraction of 1% of the total system mass allowed for numerical stability, at the cost of being a slightly conservative choice for the strength of the bar perturbation.

3.6.1 Bar Speed Determined by Angular Momentum Conservation

Being made up of point masses, the bar's effective moment of inertia $\mathcal{I}_{\rm eff}$ about the polar axis can be expressed as simply

$$\mathcal{I}_{\rm eff} = \mu_{\rm B} \int \rho_{\rm B}\, dV\, r^2 \sin^2\theta \tag{3.24}$$

in which $\rho_{\rm B}$ represents the smoothed mass distribution of stars in the bar. (Throughout this study, all quantities ρ are smoothed stellar densities.) The factor $\mu_{\rm B}$ allows for *ad hoc* tuning of the moment of intertia, and is intended to account for the fact that the bar is not a solid object but is a collection of individual stars streaming through the bar's overall structure; $\mu_{\rm B}$ can even in principle be negative but in general is of order $\mu_{\rm B} \simeq \rho_{\rm B}/\rho$ [81], and so normally we will set it as an input parameter. (Note that Athanassoula [32] and others treated the bar as a rigid object, but they did not "build" their bar models out of the consituent stars of the system.)

Quantities depending on the radial coordinate, *e.g.,* $\rho_{\rm B}$ and e, are stored in terms of ellipsoidal coordinate a^2 and so it is convenient to peform the integration over a instead of r. The mass element is then $\rho_{\rm B} dV = 4\pi\rho_{\rm B} a^2 (1-e^2)^{1/2}\, da$, and one achieves

$$\begin{aligned}
\mathcal{I}_{\rm eff} &= 2\pi\mu_{\rm B} \int_{-1}^{1} d\cos\theta \int da\, \rho_{\rm B}\, a^2 (1-e^2)^{1/2} a^2 \left(\sin^2\theta + \frac{\cos^2\theta}{1-e^2}\right) \sin^2\theta \\
&= 4\pi\mu_{\rm B} \int da\, a^4 \rho_{\rm B}\, (1-e^2)^{3/2} \int_0^1 d\cos\theta\, \frac{1-\cos^2\theta}{1-e^2(1-\cos^2\theta)} \\
&= 2\pi\mu_{\rm B} \int da^2\, a^3\, \rho_{\rm B} \left[\frac{1-e^2}{e^2}\left(\frac{\sin^{-1} e}{e} - \sqrt{1-e^2}\right)\right].
\end{aligned} \tag{3.25}$$

The above uses standard tables and trigonometric identities to do the angular integration[6] [82]. The quantity enclosed in square brackets in (3.25) approaches a value of $\frac{2}{3}$ in the limit of $e \rightarrow 0$ as is expected when finding the moment of inertia of a spherical shell; we set it to be exactly $\frac{2}{3}$ for tiny values of e ($\lesssim 0.001$, which gives an accuracy in $\mathcal{I}_{\rm eff}$ of better than one part in 10^6) in order to avoid any numerical problems of dividing by a vanishlingly small denominator.

The bar frequency is then

$$\Omega_{\rm B} = \frac{J_{\rm B}}{\mathcal{I}_{\rm eff}} \tag{3.26}$$

where $J_{\rm B}$ is the total angular momentum of the stars in the bar, $J_{\rm B} = \int dI\, dJ\, f_{\rm B}(I,J) J$. Note that the integral for $J_{\rm B}$ should properly be of J_z and not of the full J. We do not invoke an isotropy argument to approximate J_z/J because it is more straightforward to simply modify $\mu_{\rm B}$ accordingly, increasing it by a factor of $\simeq 2$ or possibly $\simeq \sqrt{2\pi}$ over the

---

[6] Substitution $z = 1 - e^2(1-\cos^2\theta)$ aids in the evaluation of the term with $\cos^2\theta$ in the numerator of the second line in (3.25).

value $\mu_{\mathrm{B}} \simeq \rho_{\mathrm{B}}/\rho$ given above.

3.6.2 Bar Speed Determined by Angular Frenquency Conservation

Alternatively, one may assume that the bar pattern speed is simply the average over all stars comprising the bar of those star's orbital frequencies, as has been found to be the case in N-body simulations of barred systems [70]. This case is very straightforward:

$$\Omega_{\mathrm{B}} = \frac{1}{M_{\mathrm{B}}} \int d^2 I \, m_{\mathrm{B}} f_{\mathrm{B}} \Omega_2 \tag{3.27}$$

As stated above, for simplicity in this study the bar is assumed to consist only of stars of a single mass, *i.e.,* $m_{\mathrm{B}} = m_1$. The total mass in the bar is simply $M_{\mathrm{B}} = m_{\mathrm{B}} \int d^2 I \, f_{\mathrm{B}}$. This method of determining Ω_{B} has the advantage of not requiring knowledge of the bar's moment of inertia and so involves no assumptions about μ_{B}. In general we employ (3.27) by default; when (3.26) is used it will be specifically noted.

3.6.3 Bar Perturbation

As the physical effect of the bar is largely to add an additional quadrupole term to the overall potential, the bar can be modeled numerically by the following three terms in the expansion (3.6), with all other terms zero:

- the $\Phi_{2\pm2}$ term that describes the quadrupole interaction;
- the Φ_{20} term that incorporates the ellipsoidal nature of the bar [83]; and
- the Φ_{00} term that sets the zero point of the bar's potential.

Only the $m = 0$ terms in (3.6) contribute to the value of Φ_{B} averaged over phase angle ϕ; because of this they are useful in determining how large the $m = 2$ term is, despite the fact that they do not directly contribute to any transfer of angular momentum themselves. The "bar potential" Φ_{B} is simply the fraction of the overall potential attributable to the subpopulation of stars that comprise the bar (as found by using ρ_{B} in place of ρ in 2.41). The requirement that the expansion sum, averaged over ϕ, matches the known value of the bar potential $\Phi_{\mathrm{B}}(a, \theta)$ implies that at a given semimajor radius value a,

$$\Phi_{00}(a) Y_{00} + \Phi_{20}(a) Y_{20}(\theta, 0) = \Phi_{\mathrm{B}}(a, \theta). \tag{3.28}$$

Evaluating (3.28) at the pole and equator ($\theta = 0$ and $\theta = \frac{\pi}{2}$ respectively), and solving for Φ_{20} and then Φ_{00} gives

$$\Phi_{20}(a) = \frac{2}{3}\sqrt{\frac{4\pi}{5}} \left[\Phi_{\mathrm{B}}\left(a, 0\right) - \Phi_{\mathrm{B}}\left(a, \frac{\pi}{2}\right)\right] \tag{3.29}$$

and

$$\Phi_{00}(a) = \sqrt{4\pi} \left[\frac{1}{3}\Phi_{\mathrm{B}}\left(a, 0\right) + \frac{2}{3}\Phi_{\mathrm{B}}\left(a, \frac{\pi}{2}\right)\right]. \tag{3.30}$$

For the strength of the quadrupole, Φ_{22} is chosen so that the more elliptical the orbits of the stars in the bar are, the more the Φ_{22} term contributes. By the symmetry of the bar, the $m = \pm 2$ terms in (3.6) are equal, leading to

$$\Phi_{2\pm2}(a)Y_{2\pm2}\left(\frac{\pi}{2}, 0\right) = \frac{1}{2}\left(1 - \frac{\langle r_p \rangle}{\langle r_a \rangle}\right)\Phi_{\mathrm{B}}\left(a, \frac{\pi}{2}\right) \tag{3.31}$$

i.e., circular orbits contribute nothing to the bar quadrupole, and purely radial orbits make for a maximally strong bar. Y_{22} is at its largest magnitude at $\theta = \frac{\pi}{2}$, so this guarantees that the bar potential as expressed by this expansion remains negative at all ϕ while its average over ϕ vanishes. The averages in (3.31) are over all bar stars whose orbits cross radial coordinate a. When a is greater than the outmost point $a_{\max}$ contained in the numeric radial grid, each Φ_{lm} is assumed to drop off as $1/a^l$ (or $1/r^l$ when the underlying potential is spherical symmetric) from the value at $a_{\max}$: $\Phi_{lm}(a > a_{\max}) = \left(\frac{a_{\max}}{a}\right)^l \Phi_{lm}(a)$.

Now that the strength of the bar perturbation has been characterized, the remaining difficulty is in evaluting the integrals of (3.23). Note that, by incorporating Φ_{B} as defined for the entire bar as a whole in the calculation of the bar's Φ_{lm} (and hence Ψ_{knm}) coefficients, the integration over the bar's distribution function f_{B} has effectively already been performed; this contrasts with the field-star case in which the Φ_{lm} (and Ψ_{knm}) are defined for individual stars and only afterwards is the distribution function integrated over. Thus (3.22), with its "bare" δ-function, is actually the proper expression to use for the bar's contribution to the diffusion coefficients.

The problem then is that, for the bar, the δ-function is then no longer a function of an integration variable. This can be thought of as being due to that, by construction, the stars comprising the bar's gravitational potential orbit in "lock step", sharing a common bulk orbital frequency, and so the δ-function no longer has a distribution of frequencies on which to act. Fortunately, the fact the calculation is being done on a numeric grid imposes a finite scale over which any bulk resonance effect must be "smeared" anyway – namely the difference $\Delta\Omega(I_1, I_2)$ between the nearest points on the action-space grid – and so the resonance effect must be widened from δ-function-width to at least this size. Thus to take the place of the δ function in the bar's diffusion coefficient, we insert a Gauusian window function $\mathcal{F}_{\mathrm{B}}$:

$$\mathcal{F}_{\mathrm{B}}(\Omega) \equiv \mathcal{A}e^{-(\Omega-\ell_3\Omega_{\mathrm{B}})^2/2(\varepsilon\ell_3\Omega_{\mathrm{B}})^2} \tag{3.32}$$

so that for the (ℓ_1, ℓ_2, ℓ_3) resonance, the closer the frequency $\Omega = \ell_1\Omega_1(I_1, I_2) + \ell_2\Omega_2(I_1, I_2)$ for any given gridpoint (I_1, I_2) is to $\ell_3\Omega_{\mathrm{B}}$, the more effect of the resonance it experiences. For the width of the resonance we compare the bar pattern speed to the timescale of the bar's slowing down due to dynamical friction:

$$\varepsilon = \dot{\Omega}_{\mathrm{B}}/\Omega_{\mathrm{B}}^2. \tag{3.33}$$

According to Tremaine and Weinberg [30], ε as defined above is of order the fractional mass density of the bar and so it is set accordingly as an input parameter. The normalization factor $\mathcal{A} = 1/(\sqrt{2\pi}\varepsilon\ell_3\Omega_{\mathrm{B}})$. With this window function we can define an effective bar diffusion coefficient

$$\mathcal{D}_{ij} = \pi \sum_{\ell_3=1}^{\infty} \sum_{\ell_1,\ell_2=-\infty}^{\infty} \ell_i\ell_j\overline{\Psi^2}_{\ell_1\ell_2\ell_3}\mathcal{F}_{\mathrm{B}}(\Omega). \tag{3.34}$$

The above $\mathcal{D}_{ij}$ is used in place of D_{ij} in (3.22); this scheme allows the bar to act as a single, bulk perturber of frequency Ω_{B} whose effects are felt only on the "cells" in the action space grid over which the line of resonances $\delta(\ell_1\Omega_1 + \ell_2\Omega_2 - \ell_3\Omega_{\mathrm{B}})$ falls. The model's computer code issues a warning if the smearing width $\varepsilon\ell_3\Omega_{\mathrm{B}}$ is smaller than the inter-gridpoint spacing of $\Delta\Omega(I_1, I_2)$ values for any action-space grid cell in which the resonance does fall.

3.7 Finite Differencing Scheme

3.7.1 Numerical Stability

Any finite differencing must be numerically stable to be useful, and so here we employ the *Crank-Nicholson* method, which averages implicit and explicit schemes and is guaranteed to be numerically stable for any stepsize – although the diffusion coefficients, which are very weakly nonlinear and are very expensive to calculate, are treated explicitly. The resulting sparse matrix is solveable using standard methods as shown below.

The Fokker-Planck equation (3.4) expressed in Crank-Nicholson form is

$$\begin{aligned}\frac{1}{\Delta t}(f_{xy}^{(\tau+1)} - f_{xy}^{(\tau)}) = \frac{1}{4}\sum_{(\tau)}^{(\tau+1)} \Bigg[& \frac{1}{\Delta I_{x+x-}}\left(D_{11x+y}\frac{f_{x+1y}-f_{xy}}{\Delta I_{x+1\,x}} - D_{11x-y}\frac{f_{xy}-f_{x-1y}}{\Delta I_{x\,x-1}}\right) \\ & + \frac{1}{\Delta J_{y+y-}}\left(D_{22xy+}\frac{f_{xy+1}-f_{xy}}{\Delta J_{y+1\,y}} - D_{22xy-}\frac{f_{xy}-f_{xy-1}}{\Delta J_{y\,y-1}}\right) \\ & + \frac{1}{\Delta I_{x+x-}}\left(D_{12x+y}\frac{f_{x+y+}-f_{x+y-}}{\Delta J_{y+y-}} - D_{12x-y}\frac{f_{x-y+}-f_{x-y-}}{\Delta J_{y+y-}}\right) \\ & + \frac{1}{\Delta J_{y+y-}}\left(D_{21xy+}\frac{f_{x+y+}-f_{x-y+}}{\Delta I_{x+x-}} - D_{21xy-}\frac{f_{x+y-}-f_{x-y-}}{\Delta I_{x+x-}}\right)\Bigg]\end{aligned} \tag{3.35}$$

The notation in (3.35) requires explanation: superscripts denote the current already-solved-for timestep (τ) and the to-be-solved-for timestep $(\tau{+}1)$; they are implicit for all f values on the right hand side. The averaging over (τ) and $(\tau{+}1)$, *i.e.,* over the explicit and implicit differencings of (3.4), constitutes the Crank-Nicholson scheme. Subscripts xy denote points on the two-dimensional grid in action space, along the radial (first subscript) and tangential (second) directions. A trailing "±" indicates the position partway between the labeled gridpoint and the one ±1 gridpoints away (*e.g.,* $D_{22xy-} = D_{22x(y-1)+}$ if "partway" is taken to be "halfway", as will be discussed in §3.7.3). Thus (3.35) is an example of differencing the flux-conservative diffusion equation "as it stands" [84]. Finally, $\Delta I_{zz'} \equiv (I_z - I_{z'})$ and similarly for $\Delta J_{zz'}$.

Note that the diffusion coefficients are treated purely explicitly; as are Φ, ρ *etc.*, they are only dependent on the overall ensemble f and not upon any particular f_{xy}, and so are similarly considered to be part of the "snapshot" of the system at timestep τ over which the evolution of f is calculated.

3.7.2 Time Splitting

Operator splitting [84] (also called "time splitting" when referring to a time-evolution operator such as used here) is used to evaluate each of the lines in (3.35) individually. Using fractional increases in τ to conceptually denote this splitting, this can be written as

$$\frac{f_{xy}^{(\tau+\frac{1}{3})} - f_{xy}^{(\tau)}}{\Delta t} = \frac{1}{4}\sum_{(\tau)}^{(\tau+\frac{1}{3})} \frac{1}{\Delta I_{x_+x_-}} \left(D_{11x_+y}\frac{f_{x+1y} - f_{xy}}{\Delta I_{x+1\,x}} - D_{11x_-y}\frac{f_{xy} - f_{x-1y}}{\Delta I_{x\,x-1}} \right) \tag{3.36a}$$

$$\frac{f_{xy}^{(\tau+\frac{2}{3})} - f_{xy}^{(\tau+\frac{1}{3})}}{\Delta t} = \frac{1}{4}\sum_{(\tau+\frac{1}{3})}^{(\tau+\frac{2}{3})} \frac{1}{\Delta J_{y_+y_-}} \left(D_{22xy_+}\frac{f_{xy+1} - f_{xy}}{\Delta J_{y+1\,y}} - D_{22xy_-}\frac{f_{xy} - f_{xy-1}}{\Delta J_{y\,y-1}} \right) \tag{3.36b}$$

$$\begin{aligned} \frac{f_{xy}^{(\tau+1)} - f_{xy}^{(\tau+\frac{2}{3})}}{\Delta t} = \frac{1}{4}\sum_{(\tau+\frac{2}{3})}^{(\tau+1)} \frac{1}{\Delta_{\pm}^2} \Big(& f_{x_+y_+}(D_{12x_+y} + D_{21xy_+}) - f_{x_+y_-}(D_{12x_+y} + D_{21xy_-}) \\ & -f_{x_-y_+}(D_{12x_-y} + D_{21xy_+}) + f_{x_-y_-}(D_{12x_-y} + D_{21xy_-}) \Big) \end{aligned} \tag{3.36c}$$

in which for convenience $\Delta_{\pm}^2 \equiv (\Delta I_{x_+x_-}\Delta J_{y_+y_-})$ has been used and terms rearranged in (3.36c). Collecting terms for $(\tau + \frac{1}{3})$ on the left and for (τ) on the right, (3.36a) becomes

$$\begin{aligned} & -\frac{1}{4}\frac{D_{11x_-y}}{\Delta I_{x\,x-1}} f_{x-1y}^{(\tau+\frac{1}{3})} + \left(\frac{\Delta I_{x_+x_-}}{\Delta t} + \frac{1}{4}\frac{D_{11x_+y}}{\Delta I_{x+1\,x}} + \frac{1}{4}\frac{D_{11x_-y}}{\Delta I_{x\,x-1}} \right) f_{xy}^{(\tau+\frac{1}{3})} - \frac{1}{4}\frac{D_{11x_+y}}{\Delta I_{x+1\,x}} f_{x+1y}^{(\tau+\frac{1}{3})} \\ & \quad = \frac{1}{4}\frac{D_{11x_-y}}{\Delta I_{x\,x-1}} f_{x-1y}^{(\tau)} + \left(\frac{\Delta I_{x_+x_-}}{\Delta t} - \frac{1}{4}\frac{D_{11x_+y}}{\Delta I_{x+1\,x}} - \frac{1}{4}\frac{D_{11x_-y}}{\Delta I_{x\,x-1}} \right) f_{xy}^{(\tau)} + \frac{1}{4}\frac{D_{11x_+y}}{\Delta I_{x+1\,x}} f_{x+1y}^{(\tau)} \end{aligned} \tag{3.37}$$

which can be seen to form a tridiagonal linear set of equations in x that can be solved for the unknowns $f_{xy}^{(\tau+\frac{1}{3})}$, for a given stepsize Δt and (fixed) y. Routine *TRIDAG* from Press *et al.* [84] is used to solve (3.37) for each y value in turn. Given knowledge of the full $f_{xy}^{(\tau+\frac{1}{3})}$, the solution of (3.36b) for $f_{xy}^{(\tau+\frac{2}{3})}$ is entirely analogous, with the radial and tangential coordinates swapping roles.

Time splitting is not mandatory for setting up differencing, but it greatly simplifies the scheme both conceptually and mathematically. It will be shown in §4.6 that time splitting does result in a more accurate and robust differencing.

3.7.3 Ensuring a Positive-Definite Distribution

In addition to numerical stability, when tracking the evolution of a distribution function of real objects it is greatly advantageous to apply a differencing scheme which guarantees a positive-definite solution for every gridpoint and at each timestep. For the one-dimensional Fokker-Planck equation there exists the *Chang-Cooper* spatial differencing scheme which (when combined with the Crank-Nicholson time-differencing described above) has both these qualities [85]. The Chang-Cooper scheme consists of a method to formulate a working prescription for the "half-grid" points such as f_{x_+y}: for example, if one defines a parameter

$\delta_{x(y)}$ using $f_{x+y} \equiv (1-\delta_{x(y)})f_{x+1y} + \delta_{x(y)}f_{xy}$ then Chang-Cooper provides a calculation for $\delta_{x(y)}$ which guarantees f will remain positive-definite.

A one-dimensional differencing scheme does not necessarily generalize to two or more dimensions, however: Chang-Cooper does not give a general form for calculating both $\delta_{x(y)}$ and $\delta_{(x)y}$. Fortunately when the Fokker-Planck equation is specifically cast in its flux-conservative form (3.4), the Chang-Cooper method reduces trivially to the case of "centrally weighting" all derivatives that involve the distribution function: $\delta_{x(y)} = \frac{1}{2}$ and $\delta_{(x)y} = \frac{1}{2}$ for all x and y. Thus $f_{x+y} = \frac{1}{2}(f_{x+1y} + f_{xy})$ *etc.* for the half-grid values of f; substituting these into (3.36c) and recollecting for the various values of f on the gridpoints (*i.e.*, f_{xy} and its eight immediate neighbors) yields a difference equation for $f_{xy}^{(\tau+1)}$ analogous to that of (3.37) for $f_{xy}^{(\tau+\frac{1}{3})}$. The ensemble of such equations for all values of x and y forms a sparse matrix system which is solved using the routine *LINBCG* from Press *et al.* [84].

3.7.4 Numerical Boundary Conditions

At the edges of the (I, J) grid the differencing schemes as described above require knowledge of f_{xy} values beyond the grid proper. Denoting the highest values of x and y on the grid as X and Y, for $x = 1$ or $x = X$ and for $y = 1$ or $y = Y$ numeric boundary conditions must replace (3.37) and its analogues for $(\tau+\frac{2}{3})$. Using the $x = X$ case as an illustration, possibilities include:

$$f_{Xy}^{(\tau+\frac{1}{3})} = f_{X-1y}^{(\tau+\frac{1}{3})} \tag{3.38a}$$

$$f_{Xy}^{(\tau+\frac{1}{3})} = \left(1 + \frac{\Delta I_{XX-1}}{\Delta I_{X-1X-2}}\right) f_{X-1y}^{(\tau+\frac{1}{3})} - \left(\frac{\Delta I_{XX-1}}{\Delta I_{X-1X-2}}\right) f_{X-2y}^{(\tau+\frac{1}{3})} \tag{3.38b}$$

$$f_{Xy}^{(\tau+\frac{1}{3})} = f_{X-1y}^{(\tau)}. \tag{3.38c}$$

These are all straightforward extrapolations; cases (a) and (c) are taken directly from Strikwerda [86], while case (b) has been generalized for a non-uniform grid.

An option which does not require any extrapolation consists of employing a one-sided (towards the interior) differencing on the boundary instead:

$$\frac{\Delta f(I_X)}{\Delta I} = \frac{f_{Xy} - f_{X-1y}}{\Delta I_{XX-1}} \tag{3.38d}$$

which can then be used in place of Chang-Cooper's centered differencing to derive analogues of equations (3.35) and (3.36) for use on the boundaries. This scheme effectively causes $\Delta^2 f/\Delta I^2$ and $\Delta^2 f/\Delta J^2$ to vanish on the boundaries, and so is referred to as the *natural-spline* method (although the full differencing of the Fokker-Planck equation only vanishes if *e.g.*, $D_{11X+y} = D_{11X-y}$ as well). It is not a numerical boundary condition *per se*, but rather a new differencing scheme for the boundary points that avoids the need for any numerical boundary condition.

Experimentation with all four possible boundary condition schemes showed that only the natural-spline treatment of (3.38d) accurately tracked an analytically-solvable test case, as shown in Figure 4.12 and described in §4.6.

3.8 Stellar Mergers

3.8.1 Rates of Loss due to Mergers

Modeling stellar mergers due to collisions requires the transfer of stars from one f_q to another, *e.g.,* if a star of mass $m_{\mathcal{Q}}$ with actions $(\mathcal{I}, \mathcal{J})$ and one of mass $m_{q'}$ with actions (I', J') combine into a new star of mass $m_q = m_{\mathcal{Q}} + m_{q'}$, then $f_{\mathcal{Q}}(\mathcal{I}, \mathcal{J})$ and $f_{q'}(I', J')$ must be decreased, and f_q increased at a dynamically appropriate value of the action (I, J).[7]

This alteration of f_q, $f_{\mathcal{Q}}$ and $f_{q'}$ is accomplished via the loss (L_q) and gain (G_q) terms added to the Fokker-Planck equation, as in (3.5). Note that these terms are not part of the differencing scheme, as they are simple rates of change and so can be calculated directly. The terms L_q and G_q are similar to those developed by Quinlan and Shapiro [1], who used a one-dimensional distribution function $F(E)$ only. We adapt them to the two-dimensional $f(I, J)$, calculating the probability of collision given cross section σ. (Context should prevent any notational confusion between the cross section and the components of the velocity dispersion, *e.g.,* σ_ϕ^2.) Expressing the cross section as a sum of powers of the collisional speed $|\Delta\vec{v}|$ with mass-dependent coefficients, *i.e.,*

$$\sigma \equiv \sum_\alpha \sigma_\alpha(q, q')|\Delta\vec{v}|^\alpha \tag{3.39}$$

then the rate of collisions of a given "target" star of mass m_q and actions (I, J) at radius r with all stars of mass $m_{q'}$ is

$$\dot{\mathcal{N}}_{qq'}(r, I, J) = \frac{\rho_{q'}}{m_{q'}} \sigma(q, q')|\Delta\vec{v}| = \sum_\alpha \frac{\sigma_\alpha(q, q')}{4\pi^2 r} \int_r \frac{dI'\, dJ'}{J'} \Omega_1' f_{q'}(I', J')|\Delta\vec{v}|^{\alpha+1} \tag{3.40}$$

where primes are used to indicate quantities dependent on the actions being integrated over, and the only dependence upon the target star's actions (I, J) is in $\Delta\vec{v} = |\vec{v} - \vec{v}'|$. The above makes use of the density (2.21) as an operator, but in doing so introduces an ambiguity due to the lack of full information about r, $\vec{v}$ and $\vec{v}'$ in the orbit-averaged distribution function $f_{q'}$. This ambiguity will be dealt with below.

Given values for the actions, the magnitudes of the radial $(|v_r(r)|, |v_r'(r)|)$ and tangential $(|\vec{v}_T| = J/r, |\vec{v}_T'| = J'/r)$ components of the stellar velocities are well-defined, but there is ambiguity in the latter's subcomponents due to lack of information regarding J_z/J. A conservative approach is to assume that the relative orientation of the tangential parts of the velocity vectors is random, *i.e.,* if γ is the angle between $\vec{v}_T$ and $\vec{v}_T'$, so that

$$|\Delta\vec{v}|_\pm^2 \equiv (|v_r| \mp |v_r'|)^2 + (|\vec{v}_T| - |\vec{v}_T'|\cos\gamma)^2 + |\vec{v}_T'|^2 \sin^2\gamma \tag{3.41}$$

then γ is randomly distributed. This is conservative because it assumes no "collimation" of $\vec{v}_T$ and $\vec{v}_T'$ due to the effect of overall cluster rotation. Taking the average over random angle γ and over the possible relative signs of the radial components one can define for any

---

[7] In practice there isn't usually a mass bin q such that $m_q = m_{\mathcal{Q}} + m_{q'}$ exactly, and so interpolation over the two bins nearest $(m_{\mathcal{Q}} + m_{q'})$ is required; this is discussed in §3.8.3.

given function $g(|\Delta\vec{v}|)$,

$$\langle g(|\Delta\vec{v}|)\rangle_\gamma \equiv \frac{1}{2\pi}\int_0^\pi d\gamma \sum_\pm g\left(|\Delta\vec{v}|_\pm\right). \tag{3.42}$$

The r dependence is easily eliminated by integrating over the orbit of the target star. Thus the orbit-averaged rate of collisions for a single star is

$$\dot{N}_{qq'}(I,J) = \int_0^\pi \frac{dw_1}{\pi}\left\langle \dot{\mathcal{N}}_{qq'}(r(w_1),I,J)\right\rangle_\gamma = \int_{r_p}^{r_a} \frac{\Omega_1 dr}{\pi[2(E-\Phi)-J^2/r^2]^{1/2}}\left\langle \dot{\mathcal{N}}_{qq'}(r,I,J)\right\rangle_\gamma \tag{3.43}$$

where $\dot{\mathcal{N}}_{qq'}$ is given by (3.40), in which the endpoints of the action integrals therein are those energetically allowed at r. In practice the integral over dr above is brought inside the one over $dIdJ$, in which case one integrates over all (I,J) and restricts the integration over r to the range, if any, over which the orbits overlap. A possible simplifying assumption for the integral in (3.43) is due to the fact that the target star spends the bulk of its time at or near its orbital endpoints $r_p = r(0)$ and $r_a = r(\pi)$.)

With this in hand the total loss rate due to stellar collision of all stars with given values of action (I,J) and mass m_q is simply

$$L_q(I,J) = f_q(I,J)\sum_{q'} \dot{N}_{qq'}(I,J). \tag{3.44}$$

Of course, in reality there is a continuum of masses and so the sum in (3.44) should properly be an integral, but in this study stellar masses are defined on a finite grid of values and so the sum is over those values.

3.8.2 Rates of Gain

The calculation of the gain rate $G_q(I,J)$ is conceptually similar but slightly more cumbersome, owing to the difficulty in determining the new velocity vector of a star which is the product of a merger of two smaller stars. In G_q, I and J now denote the actions of the star of mass m_q produced by the collision, and so we introduce the notation $(\mathcal{I},\mathcal{J})$ for the target star's actions. Thus at any given r, conservation of momentum requires

$$m_q^2 J^2 = [(m_q - m_{q'})\mathcal{J} + m_{q'}J'\cos\gamma]^2 + m_{q'}^2 J'^2 \sin^2\gamma \tag{3.45}$$

and

$$m_q v_r(r) = (m_q - m_{q'})\vartheta_\pm \pm m_{q'}v_r' \tag{3.46}$$

in which $\vartheta_\pm$ has been used for the target star's radial velocity and the choice in relative sign accounts for the possibilities that the stars have like or opposite radial velocity directions at the time of collision. Note that this is the *same* "±" as in (3.41), although of opposite sign, and does not need to be independently averaged over. After solving (3.45) for $\mathcal{J} = r|\vec{\vartheta}_T|$ and (3.46) for $\vartheta_\pm$, the target star's orbital energy is simply $\mathcal{E}_\pm = \frac{1}{2}(\vartheta_\pm^2 + |\vec{\vartheta}_T|^2) + \Phi(r)$, and then $\mathcal{I}_\pm$ can be found directly from (2.1).

The analogue of (3.43) in the merger-gain case is the rate of producing stars of mass m_q and actions (I,J) from those of masses $m_\mathcal{Q} \equiv m_q - m_{q'}$ and $m_{q'}$ and with respective actions

$(\mathcal{I},\mathcal{J})$ and (I',J'). Because we are "back-constructing" what $\mathcal{I}$ and $\mathcal{J}$ are from knowledge (I',J') and (I,J), the orbit-averaging is only over that portion of the orbit of a star with actions $(\mathcal{I},\mathcal{J})$ that overlaps with the product star orbit as determined by its actions (I,J). (Consideration of the overlap of the orbit having actions (I',J') with that of (I,J) is already implicit in the use of $\rho_{q'}(r)$ in (3.40)). Labeling the target star's radial frequency as $\Omega_{\mathcal{W}}$, its associated canonical angle variable as $\mathcal{W}_1$, and its orbital endpoints $\mathcal{R}_p$ and $\mathcal{R}_a$,

$$\begin{aligned} \dot{N}'_{\mathcal{Q}q'}(I,J) &= \left\langle \int_{\mathcal{W}_{\min}}^{\mathcal{W}_{\max}} d\mathcal{W}_1\, \dot{\mathcal{N}}_{\mathcal{Q}q'}(r(w_1),\mathcal{I},\mathcal{J}) \right\rangle_{\gamma} & (3.47a) \\ &= \left\langle \int_{\max(\mathcal{R}_p,r'_p)}^{\min(\mathcal{R}_a,r'_a)} \frac{\Omega_{\mathcal{W}}dr}{|\vartheta_{\pm}|} \dot{\mathcal{N}}_{\mathcal{Q}q'}(r,\mathcal{I},\mathcal{J}) \right\rangle_{\gamma} . & (3.47b) \end{aligned}$$

Note that in this case we integrate over $d\mathcal{W}_1$ so that contributions to $\dot{N}'_{\mathcal{Q}q'}(I,J)$ over the entire orbit of the merging stars are considered, whereas in the merger-loss case we orbit-averaged over dw_1 in order to eliminate the r dependence.

To determine $\mathcal{W}_1(r)$ (or $\mathcal{R}_p(\mathcal{I},\mathcal{J})$ and $\mathcal{R}_a(\mathcal{I},\mathcal{J})$) one must solve (3.46) and (3.45) – which is only possible within the integral over $d\mathcal{W}_1$ (or over dr), and which in turn requires knowledge of $\mathcal{W}_1(r)$ in order to determine the integral's endpoints. To avoid the impasse, the form

$$\dot{N}'_{\mathcal{Q}q'}(I,J) = \left\langle \int_{\max(r_p,r'_p)}^{\min(r_a,r'_a)} \frac{\Omega_{\mathcal{W}}dr}{|\vartheta_{\pm}|} \dot{\mathcal{N}}_{\mathcal{Q}q'}(r,\mathcal{I},\mathcal{J}) \right\rangle_{\gamma} \tag{3.48}$$

is used instead, with the additional requirement that only dynamically-allowed solutions of (3.45) are allowed: *i.e.,* for a given value of γ we define $\dot{\mathcal{N}}_{\mathcal{Q}q'}(r,\mathcal{I},\mathcal{J}) \equiv 0$ if $\mathcal{J} < 0$ or is complex; this is equivalent to only considering orbital ranges that do indeed overlap.

The rate of gain due to the mergers is thus

$$G_q(I,J) = \frac{1}{2}\sum_{q'} \dot{N}'_{\mathcal{Q}q'}(I,J) f_{\mathcal{Q}}(\mathcal{I},\mathcal{J}) \tag{3.49}$$

with an extra factor of $\frac{1}{2}$ inserted to counteract the effect of double-counting;[8] it is arbitrary which is called the "target" star, and it takes one of each to create a merger product. The right side of (3.49) contains a notational sleight of hand: since $(\mathcal{I},\mathcal{J})$ depend on I', J', and γ, the integrals from (3.40) and (3.42) now act as operators on $f_{\mathcal{Q}}(\mathcal{I},\mathcal{J})$. For clarity, the full expression for the gain rate is

$$G_q(I,J) = \sum_{\alpha,q'} \frac{\sigma_\alpha(\mathcal{Q},q')}{16\pi^3} \int \frac{dI'\,dJ'}{J'} \Omega'_1 f_{q'}(I',J') \int_0^{\pi} d\gamma \int_{\mathcal{W}_{\min}}^{\mathcal{W}_{\max}} \frac{d\mathcal{W}_1}{r(\mathcal{W}_1)} \sum_{\pm,\mathcal{J}} f_{\mathcal{Q}}(\mathcal{I}_{\pm},\mathcal{J}) |\Delta\vec{v}|_{\pm}^{\alpha+1} \tag{3.50}$$

in which the final sum is over the two possible sign choices in (3.41), and over (one or both) dynamically-allowed values of $\mathcal{J}$ found from solving (3.45). In practice it is required to do

---

[8] Other studies, such as [7], have included a factor of $1/(1+\delta_{\mathcal{Q}q'})$ instead of $1/2$ to prevent double-counting. The distinction is that here we sum over all possible collisional pairs and attribute their products to the appropriate $m_q = m_{\mathcal{Q}} + m'_q$, as opposed to predetermining the product m value being considered and invoking a delta function $\delta(m_q - m_{\mathcal{Q}} - m'_q)$ to restrict the mass values of the colliding stars as needed. The present method is more efficient computationally as all G_q for a given (I,J) are found simultaneously.

the orbit-averaging integral directly over r and not over $\mathcal{W}_1$, as in (3.47b) and (3.48).

An artifact of this approach is that the symmetry between "target" and "object" stars is not obvious in (3.50). However as the numerical calculations for L_q and G_q are performed by independent sections of computer code, a match between overall merger losses and gains shows the validity of the method, as will be discussed in §4.5.

3.8.3 Mass Bookkeeping

The above derivation made use of the manifest fact that any star which is the product of a purely-inelastic stellar collision and merger will have a mass equal to the sum of the colliding stars' masses: $m_q \equiv m_\mathcal{Q} + m_{q'}$. But the model only tracks stars of a finite number mass values, and so it's possible that m_q does not correspond to a stored value. To account for such intermediate mass values, merger products are interpolated across the nearest values below and above m_q on the mass grid [87]. Referring to these values as m_{q-} and m_{q+} respectively, one has

$$G_{q_-} = \frac{m_{q_+} - m_q}{m_{q_+} - m_{q_-}} G_q \tag{3.51}$$

and

$$G_{q_+} = \frac{m_q - m_{q_-}}{m_{q_+} - m_{q_-}} G_q. \tag{3.52}$$

Note that because actions are conserved in collisions, no interpolation in I or J is required. Thus the merger-gain rates from both (3.51) and (3.52) for a given mass value contribute to the full Fokker-Planck equation (3.5).

One may ask why do we not merely round down the merger-product's mass to the next lowest mass bin value, in order to account for mass loss during a not-completely inelastic stellar collision which produced the merger. The main reasons are (1) that it would also inadvertently reduce the total mass of the cluster, which while it would be a reasonable effect in an ordinary globular cluster would not be expected to occur in these more massive systems; and (2) that there is no way to account for diffuse gas in the calculation of the cluster's gravitational potential, nor in associated dynamical quantities. Thus all collisional mass is assumed to go into the newly-merged star, a similar approach to that taken by Quinlan and Shapiro [7] who surveyed results of hydrodynamical simulations of stellar collisions – most of which find a maximum mass loss per collision of $\lesssim 11\%$ – and concluded that the average mass loss will be much less than that and thus also assumed complete coalescence into the produced larger-mass star. It is also consistent with the finding of Freitag *et al.* [19] that assuming coalescence is "fully justified" for velocity dispersions $v_{rms} \lesssim 300$ km/s – a condition largely satified by the simulations here, the vast bulk of which have a calculated initial velocity dispersion of $\sigma_1 \leq 310$ km/s.

3.8.4 The Delta-function Approximation

The full merger-gain coefficient calculation (3.50) is by far the most computationally intensive part of the model, and so calls for a simplifying approximation. The "delta-function approximation", in which all stars at a given of radial distance are assumed to share a common $|\vec{v}'| = v_{rms}(r)$, was found by Quinlan and Shapiro [1] to provide large performance gains with only a marginal loss in accuracy in most cases. The key to the approximation is to not replace the density $\rho_{q'}$ in (3.40) with its distribution-function integral form (2.21) and instead to substitute $|\vec{v}'|$ with $v_{rms}(r)$ throughout, thus avoiding two computationally-expensive integrals over the action components. However the present two-dimensional study requires velocity components $v'_r \equiv |\vec{v}'| \cos\theta_v = v_{rms}\cos\theta_v$ and $v'_T \equiv |\vec{v}'| \sin\theta_v = v_{rms}\sin\theta_v$, and so the approximation comes at the cost of an averaging over $\sin\theta_v$:

$$G_q(I,J) \simeq \sum_{\alpha,q'} \frac{\sigma_\alpha(\mathcal{Q},q')}{4\pi} \int_0^1 d\sin\theta_v \int_0^\pi d\gamma \int_{\mathcal{W}_{\min}}^{\mathcal{W}_{\max}} d\mathcal{W}_1 \frac{\rho_{q'}(r(\mathcal{W}_1))}{m_{q'}} \sum_{\pm,\mathcal{J}} f_{\mathcal{Q}}(\mathcal{I}_\pm,\mathcal{J})|\Delta\vec{v}|_\pm^{\alpha+1}. \tag{3.53}$$

The actions $(\mathcal{I}_\pm,\mathcal{J})$ at a given position $r(\mathcal{W}_1)$ are determined by the new conservation rules

$$m_q^2 v_T^2 = [(m_q - m_{q'})|\vec{\vartheta}_T| + m_{q'} v'_{rms} \sin\theta_v \cos\gamma]^2 + m_{q'}^2 v'^2_{rms} \sin^2\theta_v \sin^2\gamma \tag{3.54}$$

and

$$m_q v_r(r) = (m_q - m_{q'})\vartheta_\pm \pm m_{q'} v'_{rms} \cos\theta_v \tag{3.55}$$

which, as before, may not have a solution for either or both $(\mathcal{I}_+,\mathcal{J})$ or $(\mathcal{I}_-,\mathcal{J})$ (in which case $f_{\mathcal{Q}}(\mathcal{I},\mathcal{J}) \equiv 0$). The notation $v'_{rms}(r)$ indicates the *rms* velocity specifically of the subpopulation of stars of mass q', making this an even less-severe approximation than in the one-dimensional case. Note that, consistent with the conservative assumption of a randomly-varying γ in §3.8.1, $\vec{v}'$ is also taken to be evenly distributed over θ_v. The averaging is restricted to the range $0 \leq \theta_v < \frac{\pi}{2}$ because the "±" in (3.55) already accounts for $\frac{\pi}{2} \leq \theta_v < \pi$.

3.8.5 The Cross Section

Also following Quinlan and Shapiro [7], for the stellar mergers we take a hard-sphere cross section with gravitational focusing correction:

$$\sigma = \pi(r_*^2 + r_*'^2)\left[1 + \frac{2G(m+m')}{(r_* + r'_*)|\Delta\vec{v}|^2}\right] \tag{3.56}$$

in which m and m' are the individual stellar masses, and r_* and r'_* are the physical stellar radii of main-sequence stars of those masses.

In practice, either term in the above may dominate in a given collision, depending on the size of the relative speed of collision $|\Delta\vec{v}|$; Ebisuzaki *et al.* [14] has found observational evidence that the gravitational focusing term can be important in compact stellar clusters, and Freitag *et al.* [19] claim that the gravitational focusing term dominates when the velocity dispersion $\sigma_1 < 300$ km/s, which is the case for most but not all simulations presented here (the exception being the "E4A" models – which, as mentioned above, almost satisfy that criterion).

3.9 Binary Mergers, Binary Heating

It will be seen in Chapter 5 that two-body mergers, while in some cases fairly frequent, do not dominate the dynamics in our simulations. Given that 3-body collisions are much more likely to form binaries than to result in a direct merger, that Quinlan and Shapiro [7] found that few hard 3-body binaries formed during most of their simulations, and that in the absence of a central massive object Freitag *et al.* [19] found that 2-body mergers start before 3-body binaries can form, we ignore the effects of 3-body collisions for stellar mergers as well as for the effect of binary hardening on the overall energy budget of the system. In addition, following the lead of prior studies ([56], [22]) the effect of primordial binaries is also not considered for simplicity's sake. For determining the rate of stellar mergers this is a conservative approach, as when there is no massive central object primordial binaries are likely to foster collisions (even if once a massive black hole does form the binaries then serve to "grind down" stars instead of growing them) [19]. In their N-body simulations Portegies Zwart and McMillan [9] also found that 3-body encounters (binary+star) increase the rate of mergers, in contrast to prior assumptions.

Binaries can also serve to heat the overall distribution of field stars as the binaries harden [7], with the potential to eventually halt and reverse core collapse. However, in practice for dense clusters this effect is not important prior to the late stages of core collapse and so can be ignored in earlier stages [22]. Also, stellar collisions between binaries and other stars lead to a significant reduction in the amount of heating produced [56]. Finally, binary + field star interactions are most likely for very small ($\delta v \simeq 5$ km/s) relative velocities [88], whereas the models employed here feature velocity dispersions measured in the hundreds of km/s.

For all the above reasons, binary heating has not been incorporated into the simulations performed. However, a test of the maximum possible amount of binary heating the system could have achieved was performed for a variety of models, the results of which are given in §4.8.

Chapter 4

Validity Tests and Model Parameter Choices

4.1 Overview

There are five major calculational components to the simulation method:

1. updating the gravitational potential $\Phi(r)$ and stellar density $\rho_q(r)$ given a new distribution function $f_q(I, J)$ at each timestep;
2. calculating the resulting new values for stellar-dynamical quantities such as the orbital frequencies $\Omega_j(I, J)$, as well as the conversion between orbital angle w_1 and radial position $r(w_1)$ for a given actions (I, J) *etc.*);
3. calculating the full set of diffusion coefficients $D_{ij}(I, J)$;
4. calculating the rates of mass loss $(L_q(I, J))$ and gain $(G_q(I, J))$ for each stellar mass value m_q within the distribution functions $f_q(I, J)$; and
5. finite-differencing the Fokker-Planck equation to update each $f_q(I, J)$ using D_{ij}, L_q and G_q.

Also requiring consideration are choices for the purely-numerical parameters of each simulation:

- the size and range of the grid of actions (I, J) on which many quantities are calculated;
- the size and range of the radial-coordinate r grid for $\rho(\vec{r})$, $\Phi(\vec{r})$ *etc.*;
- the choice of timestep Δt;
- the number of terms used in the spherical-harmonic expansion of Φ; and
- how to "bin" the range of stellar masses studied into discrete values m_q.

Tests of each of these aspects of the overall calculation are presented in turn in this chapter, as is verification that the effect of binary heating is sufficiently small that it need not be included in the model.

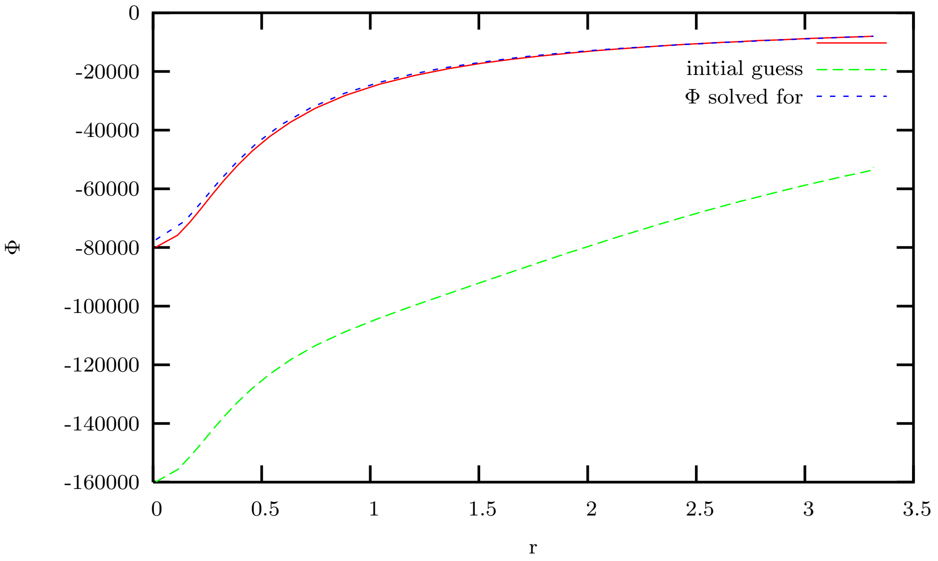

Figure 4.1: Demonstration of the ability of the potential-solver to converge even when given an initial guess for Φ and ρ that are extremely in error. The solid line is the analytical curve for $\Phi(r)$ for a Plummer sphere, the long dashes are the initially input values, and the short dashes are the iteratively solved-for potential. This convergence was obtained after only 4 iterations. The units in this figure are arbitrary for testing purposes.

4.2 Potential-calculating Tests

The first part of the model iteratively calculates the new density ρ and gravitational potential Φ given an updated f; it is most easily tested by giving it an artificially bad (*i.e.,* far from correct) initial guess for $\Phi(r)$ and $\rho(r)$, and verifying that it does converge on the true values. That it does can be seen in Figures 4.1 and 4.2. Of note is that the input functions were close to being "maximally" bad, *i.e.,* any much larger discrepency from the correct values produced unphysical results in intermediate calculations (*e.g.,* $\rho < 0$). Yet in these tests cases the solver converged on the proper solutions in approximately the same number of iterations as it does when used in the actual model calculation.

4.3 Dynamical Tests

4.3.1 Orbital Frequencies

The orbital-dynamic quantities described in §2.2 appear in almost every higher-level calculation in the model. In order to test the accuracy of the calculation of the dynamical quantities' values, two different potential-density pairs for which all quantities are also analytically calculable were used as test cases: the two-dimensional Simple Harmonic Oscillater (SHO), and the isochrone potential [42].

Figure 4.3 shows a sample result of the SHO case; the upper curves are of Ω_1, lower are of Ω_2. The results do not depend on whether the Ω_j were found by directly integrating over

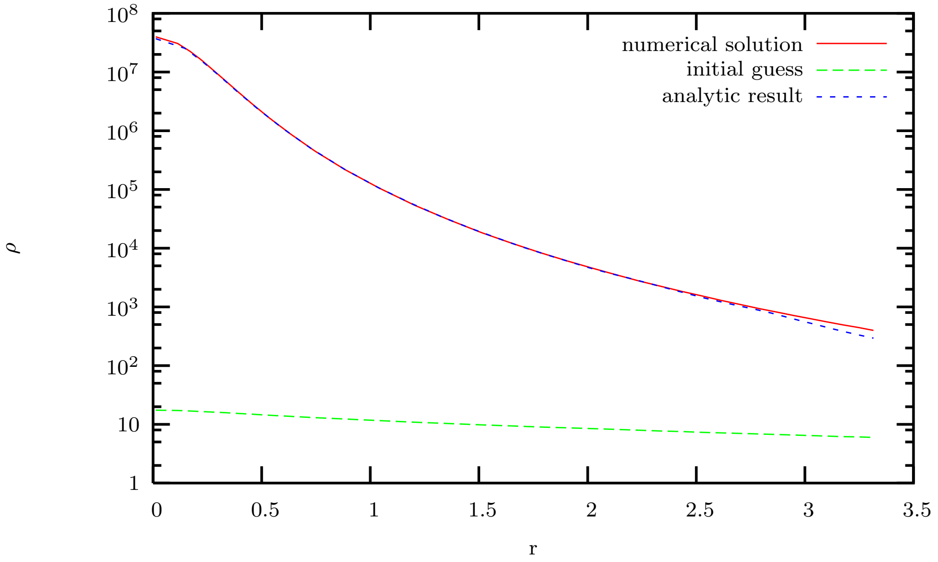

Figure 4.2: The same test of the potential solver as shown in Figure 4.1, but now plotting $\rho(r)$, again in arbitrary units. The initial guess was chosen to be much smaller and flatter than the true solution.

the orbit (*i.e.,* using (2.3) or (2.2)) or by calculating the partial derivative of the energy (using (2.5)); this is consistent with the fact that the numerically-determined values for E and for the orbital endpoints matched the analytic values to within 10^{-3} or better. As can be seen, there is a slight systematic offset in the value found for Ω_1, but otherwise the match to the analytic curves is good; this is a typical result.

When testing the calculation of dynamical quantities, there was a choice in what values to use for the potential $\Phi(r)$: either numerically-determined values (*i.e.,* a "fully-numeric" test which employed the entire calculational machinery), or the analytically-known potential (a "semi-analytic" test which isolated the dynamical calculations only). Figure 4.4 shows results for both, in the case of the isochrone model. Again there is a slight offset between the fully-analytic Ω_1 and the numeric curves. Of note is the close match between the semi-analytic and fully-numeric calculations; this is not surprising given how well the potential-finding algorithm described in Chapter 2 works, as seen in §4.2.

4.3.2 Orbital Angles

As stated in §2.2, the diffusion coefficients do not depend on the canonical angle variables w_j, nor is knowledge of them needed for any of the other calculations in the model; this is largely due to the fact that the sinusoidal dependence of the multipole components of the potential in action space (*i.e.,* as shown in (3.7)) is eliminated upon orbit-averaging.[1]. However, $r(w_1)$ and $[\psi - w_2](r)$ are explicitly required to perform the integral of (3.9) in the diffusion coefficient calculation, and $w_2(r)$ is needed for many of the tests described in this chapter.

---

[1] Although as shown *e.g.,* in (3.47), if desired it is always possible to change variables from r to w_1 by using (2.6).

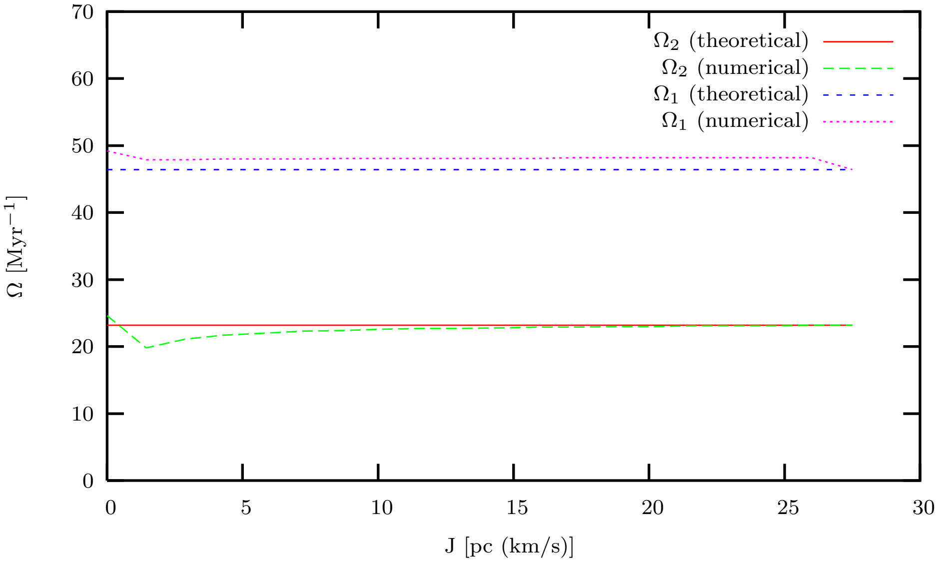

Figure 4.3: Orbital frequencies Ω_1 (upper plot) and Ω_2 vs. J for a Simple Harmonic Oscillater potential, at a constant energy. Straight lines are the analytic values, varying ones are the numerical-found solutions. Numerical values of other dynamical quantities (energy, orbital endpoints) are accurate to within 10^{-3} or better of the analytic values. As explained in the text, numerical curves in this plot were found using the simulation code's calculations of the potential and density; substituting analytically-known potential and density values into the calculation produced identical curves. Similarly, whether the Ω_j were found by integrating over the orbit (using (2.2) or (2.3)) or by taking $\Omega_j = \frac{\partial E}{\partial I_j}$ (2.5) made no difference.

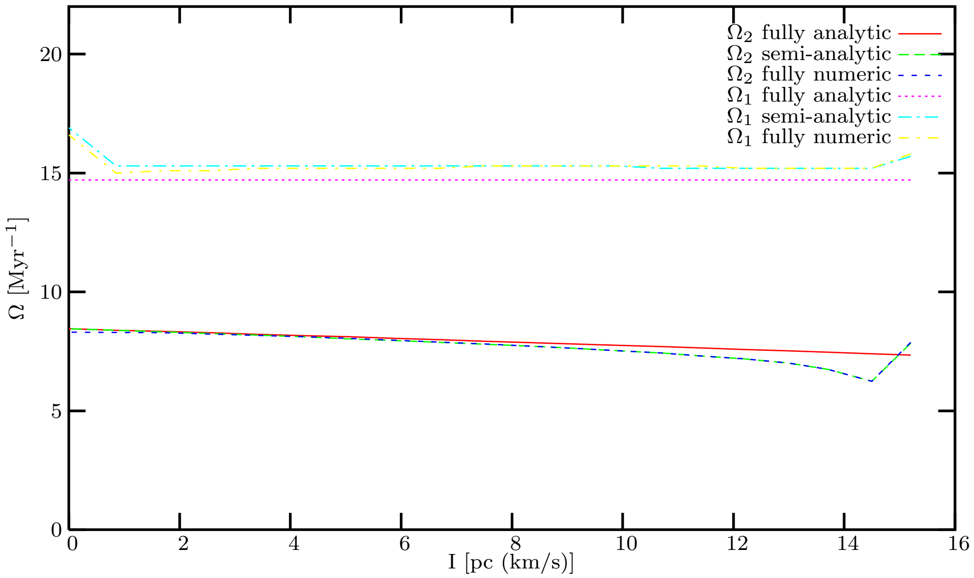

Figure 4.4: Ω_1 (upper plot) and Ω_2 vs. I for an isochrone potential, at a constant energy. Similar to Fig. 4.3 but here the numerically-found Ω curves show both the semi-analytic- and fully-numeric-potential cases.

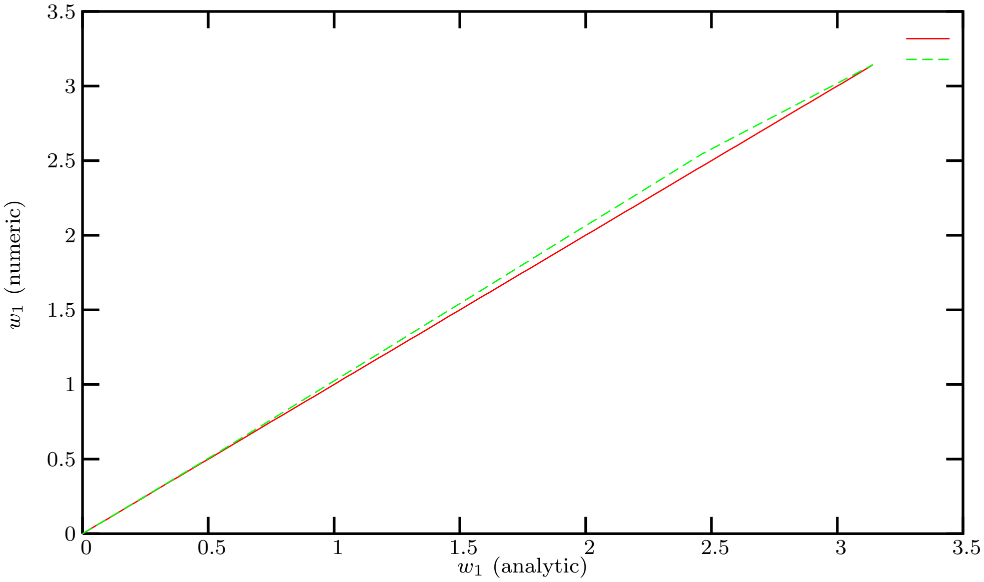

Figure 4.5: Orbital angle $w_1(r)$ interpolated from values calculated using (2.6) v. the analytically-known value of w_1 in an SHO potential is plotted as a dashed line. Results of a semi-analytic test were identical to the fully-numeric results shown here. The solid line shows the diagonal, for comparison.

Figure 4.5 shows the numerically-calculated $w_1(r)$ versus the analytically-known function $w_1(r)$ in an SHO potential. For the calculation of the curve in Fig. 4.5, first $w_1(r)$ was integrated using (2.6) at a finite number of values of r and then the shown curve was interpolated at other points from that grid of values. The technique of calculating a complicated function on a finite grid and then interpolating to determine the function's value at an arbitrary point was employed throughout this study in order to reduce computation time, and so this serves as a test of that procedure as well.

Figure 4.6 is the analogous plot of r, *i.e.,* in which $w_1(r)$ was calculated on a grid and then the inverse funtion $r(w_1)$ found by doing a reverse interpolation. In this case the interpolation does more than merely reduce the required computation time, it is also required to actually invert the function. Figures 4.6 and 4.5 show that the results of the numeric calculation of w_1 and r from Φ and ρ match the true values of those orbital angles.

The other interesting orbital angle term is $(w_2 - \psi)$ as described in §2.2. Figure 4.7 plots $(w_2 - \psi)$ calculated from (2.7) versus the analytically-known values in an SHO potential, for two different values of the actions (I, J) (*i.e.,* across two different orbits). The orbit plotted with "+" in that figure is a typical result, while the one shown with "×" is one of the more extreme cases. It can be seen that although $(w_2 - \psi)$ isn't calculated quite as accurately as w_1 (perhaps because of the delicacy of the $[\Omega_2 - J/r^2]$ factor in (2.7)) neither orbit's values are far from the expected diagonal.

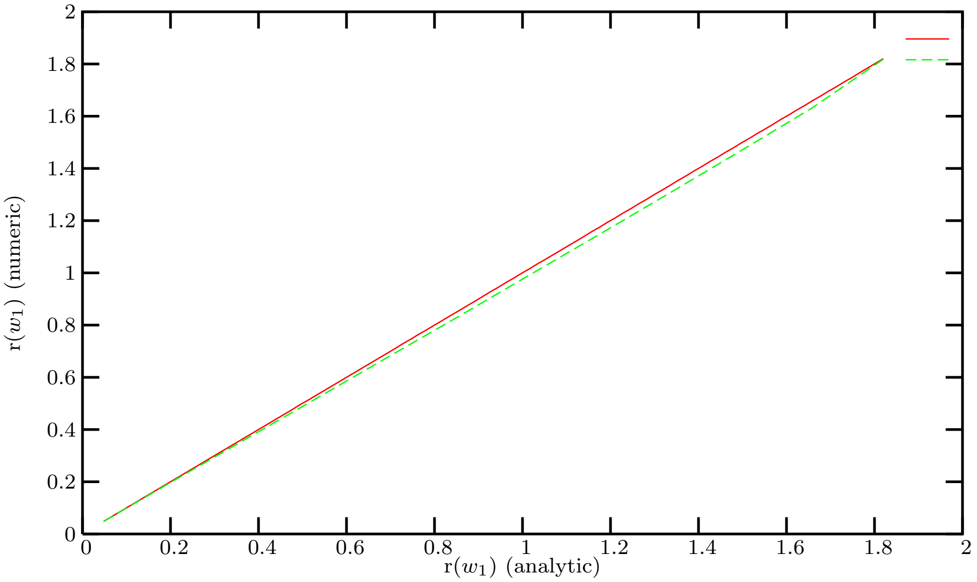

Figure 4.6: As in Fig. 4.5, but here the interpolated inverse function $r(w_1)$ is plotted v. the analytic values of r.

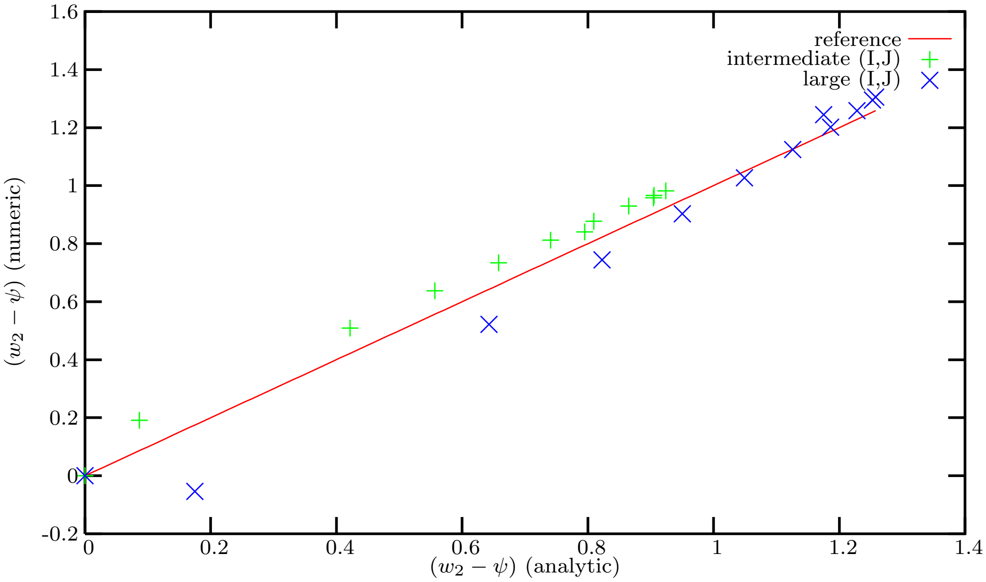

Figure 4.7: Angle term $(w_2 - \psi)(r)$ plotted as numerically-calculated vs. analytically-known values. Points denoted by "+" are for an orbit with intermediate values of actions (I, J) for the SHO potential used, while "×" are for an orbit with (I, J) near the upper end of the range of actions that are still bound in the potential. This figure is done as a scatter plot because, unlike $w_1(r)$, $(w_2 - \psi)(r)$ is not a monotonic function of r. As in Figs. 4.5 and 4.6, $(w_2 - \psi)$ is first integrated from (2.7) for a finite number of r points, and then the values plotted here interpolated are from that grid in r.

4.4 Diffusion Coefficient Tests

4.4.1 Low-level Calculations

The field-star-orbit averaging of the potential's expansion terms as given by (3.15) is one of the more complex steps required in calculating the diffusion coefficients. But for some values of the expansion index l, and over some ranges of r/R_c, the integral in (3.15) is analytically solveable in the test potentials – specifically, for $l = 0$ or 2 in the isochrone potential, and for $0 \leq l \leq 6$ in the SHO potential. In both cases a consistent good match, within 10% or better, was found between the numerically-integrated values and the analytic one, and a match within 2% or better between numeric integrations using analytic or fully-numeric integrands. (Here "fully-numeric" means using only code from the model itself, with the test potential values as inputs.) While encouraging, this test only verifies that the integrals in (3.15) are being performed accurately, and not that they are correct in their form. For that, see below.

4.4.2 Diffusion Coefficients: Reproducing the Potential

As shown in (3.23) the major part of the diffusion coefficient calculation consists of determining the expansion coefficients $\overline{\Psi}_{\ell_1\ell_2\ell_3}$ of the potential in action space, whose relation to the (position-space) potential Φ_* of a given perturber is given by (3.7). Thus a fundamental test of the validity of the diffusion coefficients is to check whether they can be used to reconstruct the overall gravitational potential of a field star or of the bar, effectively reversing the change of coordinates from position-space to action space from which they were calculated. To reconstruct the potential, first we have to add in the $l = 0$ and $l = 1$ terms to (3.7) that don't contribute to the dynamical friction:

$$\Phi(\mathbf{r}) = \sum_{l_3=0}^{\infty} \sum_{l_1,l_2=-\infty}^{\infty} \overline{\Psi}_{\ell_1\ell_2\ell_3}(I,J)\cos(\ell_p w_p - \omega t). \tag{4.1}$$

Not all of the above terms are needed: as stated in §2.2, the angle w_3 contains a random phase ϕ_o and so only terms with $l_3 = 0$ contribute to the sum, due to the cosine factor. Angle w_2 includes ψ, which is measured from the ascending node to the current orbital position and is unconstrained by where in the allowed range of orbital radius r the star currently is; however, this does not imply that terms with $l_2 \neq 0$ do not contribute: even though ψ and w_1 are uncorrelated, ψ and w_2 may be and so w_2 does not act as a random phase as w_3 does.

So upon orbit-averaging, the cosine factor restricts contributions to only those terms with $l_3 = 0$. Note that the cosine factor does not appear in the actual dynamical friction calculation, and so non-zero l_3 terms do contribute there – in fact, their presence is what allows for the resonant interaction between field star and perturber. All values of l_1 and l_2 must be included here however, as w_1, w_2 and the Ψ are all dependent upon r for a given value of the orbital action vector (I, J), and so the only terms that contribute to (4.1) are

$$\Phi(\mathbf{r}) = \sum_{l_1,l_2=-\infty}^{\infty} \overline{\Psi}_{l_1 l_2 0}(I,J)\ \cos(l_1 w_1 + l_2 w_2 - \omega t) \tag{4.2}$$

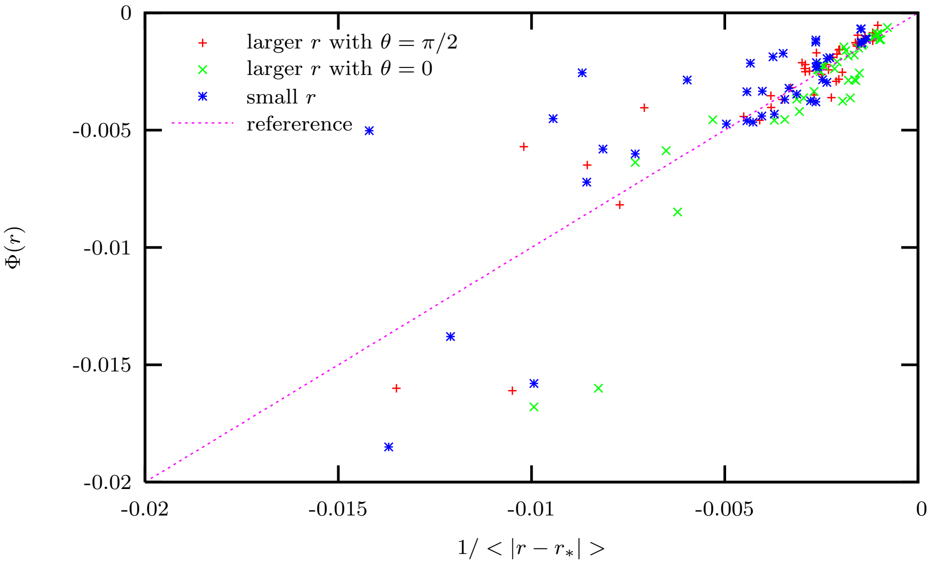

Figure 4.8: Scatter plot showing the results of using (4.2) to reconstruct the potential $\Phi(r) = 1/\langle|r - r_*|\rangle$ of a *single* star within a larger cluster from its diffusion coefficients $\Psi_{l_1 l_2 0}$. The various points plotted represent different choices both of test-point location $\mathbf{r}$ and of object-star action values (I, J). Units are arbitrary but equivalent along both axes.

which can then be averaged over any object-star orbits being considered.

The simplest test case to consider is that of a single object star's potential as given by $\Phi(r) = 1/|r - r_*|$. Figure 4.8 shows the values of $\Phi(r)$ reconstructed using (4.2) from the diffusion coefficients of object stars with various orbital actions (I, J), calculated at several locations $\mathbf{r}$ within the overal cluster. In general the match between the reconstructed $\Phi(r)$ and the expected value $1/\langle|r - r_*|\rangle$ is fairly close, with some exceptions at larger values of $1/\langle|r - r_*|\rangle$, *i.e.,* when the gravitational effect of the object star is greatest. This is not a large cause of concern however: Figure 4.9 shows that the majority of the outliers in Fig. 4.8 are due to object stars on nearly-circular ($I \simeq 0$) or nearly-radial ($J \simeq 0$) orbits; as such orbits are nearly unpopulated in the actual distribution functions $f(I, J)$ used in the simulations they do not contribute strongly to the overall diffusion coefficient values. For the non-extreme orbits that make up the bulk of $f(I, J)$, the diffusion coefficients are seen to be able to reproduce the perturbing potential as expected.

4.4.3 The Bar Perturbation

The calculation of the bar's diffusion coefficients, as shown in (3.34), was constructed so that the bar potential should describe a quadrupole while still averaging to the underlying aggregate potential of its constituent stars. Figure 4.10 shows that this calculation works as designed: near the cluster's equatorial plane the bar potential, as reconstructed from the diffusion coefficients using (4.1), is dominated by the Φ_{22} term, but towards the pole the θ dependence of the spherical harmonic Y_{22} allows Φ_{00} and Φ_{20} to contribute most to the overall bar potential.

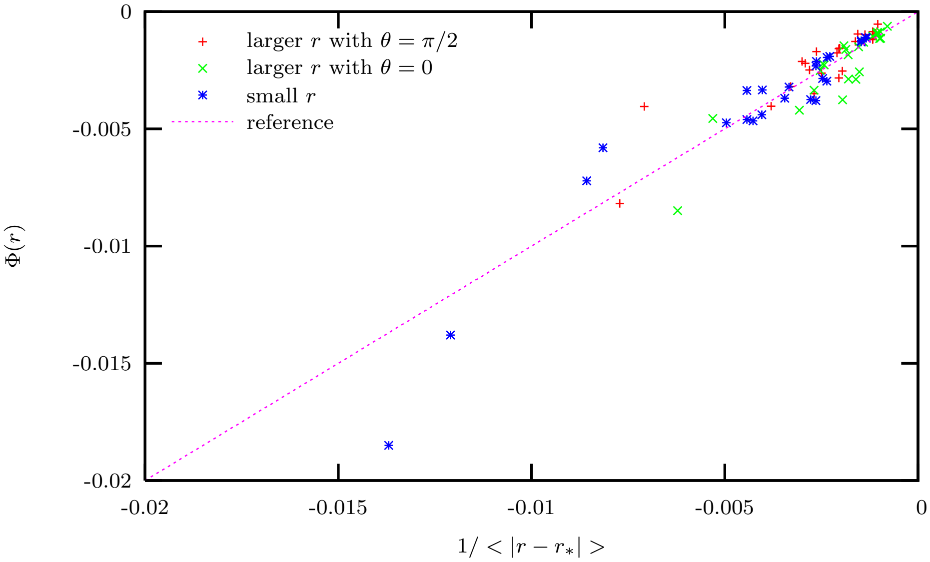

Figure 4.9: Similar to Fig. 4.8, but with circular and purely-radial object-star orbits removed.

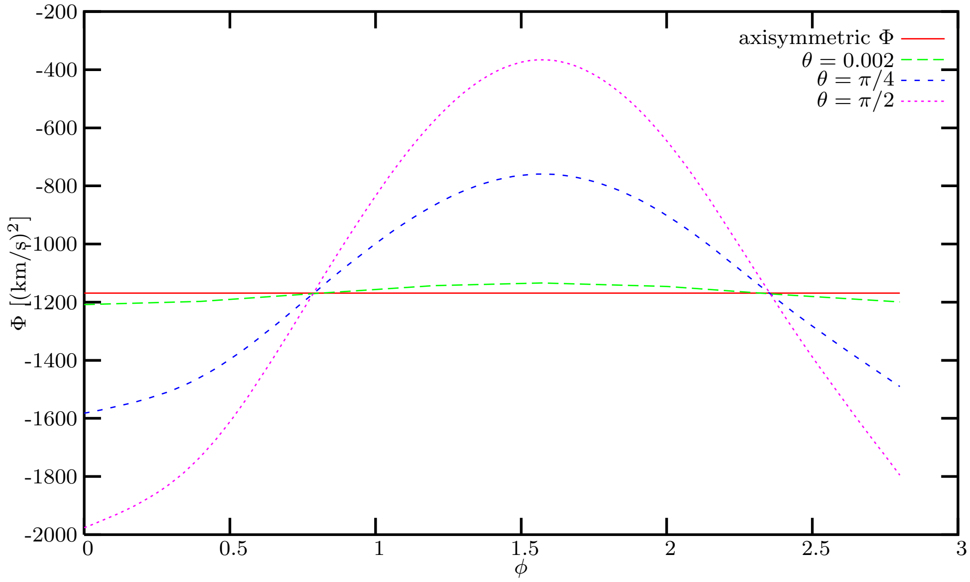

Figure 4.10: The gravitational potential of the bar v. azimuthal angle ϕ, as reproduced by summing (3.7) using the diffusion coefficients of (3.34). This plot is for $r=1.079$pc using the "E1B" model. The flat line is the base axisymmetric potential that the same stars would have if they were not assumed to be in the bar. The curved lines show the bar potential at different values of polar angle θ, from $\theta=0.002$ for the least-curved line through $\theta=\pi/4$ and $\theta=\pi/2$ for the most-curved.

4.5 Merger Losses & Gains

The calculations of the rates of mass loss and gain due to stellar mergers, L_q (as given by (3.44)) and G_q (by (3.50)), are done independently, offering the opportunity to compare them for consistency. Figure 4.11 shows the results for the three models primarily studied: the E4B model with Kroupa- or Arches-style IMF, and the E2B model with Arches-style IMF. The results shown are taken from the same simulations described in detail in Chapter 5 and so have upper mass limits of $125M_\odot$ and either 9 mass bins (for the Arches-style IMF given by (2.55)) or 10 bins (for the Kroupa IMF of (2.53)) between which merger rates are calculated.

Perfect agreement between the loss and gain calculations would result in $\langle L_q \rangle = \langle G_q \rangle$. As can been seen from the figure the fit is not perfect but is close for the most part, with the largest deviation being at late times in the E2A model. Comparing the three E4B runs shown, it is clear that including a minimal stellar bar in the model results in a substantial increase in the merger rates, but using an Arches-style IMF produces much larger merger rates from the start which also increase more over the simulation time; this will be explored more in Chapter 5.

These results, and similar ones for other models, are taken to validate the overall method of calculating L_q and G_q. However as the fit between mass loss and gain rates is not perfect, at each timestep the calculated gain rate $G_q(I, J)$ is multiplied by a correction factor $\langle L_q \rangle / \langle G_q \rangle$, *i.e.,* $G_q(I, J)$ is normalized so that the effective value of the net rate of mass gain due to mergers $\langle G_q \rangle$ agrees with the average rate of mass loss $\langle L_q \rangle$. (G_q is adjusted to agree with L_q instead of the other way around because L_q is a much simpler calculation and in practice shows less fluctuation over simulated time.)

4.6 Testing the Differencing Scheme

In order to test the differencing scheme analytic solutions of the Fokker-Planck equation (3.4) for both constant and varying D_{ij} were employed. Both solutions have the form $f(x, y, t) = Ae^{k_x x + k_y y + \omega t}$. In the constant-coefficient case, the dispersion relation thus produced is $\omega = \frac{1}{2} D_{xx} k_x^2 + D_{xy} k_x k_y + \frac{1}{2} D_{yy} k_y^2$. For an analytic solution with non-constant diffusion coefficients, if one takes $D_{xy} = 0$, $D_{xx}(x) = P_x e^{-k_x x}$ and $D_{yy}(y) = P_y e^{-k_y y}$ with $P_i = const$, then $\omega = 0$ and a static system ($\frac{\partial f}{\partial t} = 0$) results.

Figure 4.12 shows that the numerical boundary condition of (3.38d) tracks the true solution closely, only failing to keep up on the edge at which the per-timstep change is greatest – this is not surprising given that (3.38d) implements a natural spline to deal with the edge effects and will underestimate changes for which the second derivative is far from zero. This is not a large concern in practice, as tests show the discrepency is worst when the "crossterm" diffusion coefficient D_{xy} is of similar magnitude as the smaller of D_{xx} and D_{yy} (as is the case in Fig. 4.12); in the actual model D_{xy} is typically much smaller than either D_{xx} or D_{yy} (although it does become of similar size when all three are close to infinitessimal). Also, in the simulations the D_{ij} don't increase exponentially near the edges, meaning the natural spline is a better approximation than it is in these tests.

Figure 4.13 is similar to Fig. 4.12 except that it plots the constant-f case and not the

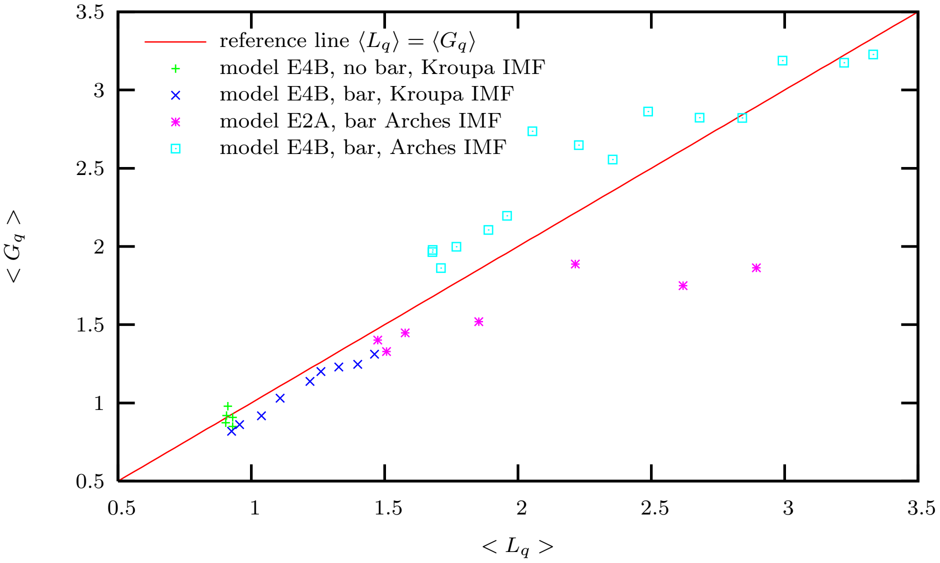

Figure 4.11: Comparison of stellar-merger loss and gain terms L_q and G_q for the primary models studied, averaged over actions (I, J) and stellar-mass type. Units are arbitrary, but consistent within a given model. In general $\langle L_q \rangle$ and $\langle L_q \rangle$ both increase over time, so the progression in this plot is left-to-right for a given model. Full results for these simulations are given in Chapter 5.

constant-coefficient one.

It can been seen that use of the numerical boundary condition of (3.38a) produced a much poorer match to the true solution. Use of any of the other conditions of (3.38) resulted in similar or worse matches. Similarly, not using a time-splitting scheme was unsatisfactory, as shown by the dotted line in the figures. The Courant stability-criterion values [84] for the numeric tests plotted were all $\xi = \frac{1}{4}$, but doubling the timestep and so doubling ξ made no difference in the time-split case, although it did make the non-timesplit case even worse than that shown. Increasing the ξ further (*i.e.,* past $\xi \simeq 1$) and/or taking a much larger number of timesteps did start to produce larger discrepencies towards the edges where f is largest, presumably due to increased inaccuracies caused by the natural-spline numeric boundary condition more severely underestimating the second derivative of f there.

4.7 Model Parameters

4.7.1 Grid size: action space

In order to have confidence in the model's results we must show that the simulations are stable against changes in parameters which are purely numerical: grid sizes, timestep sizes *etc.* Figure 4.14 shows a representative comparison of using different sizes of action-space grids. In this and several other tests, a 40x40 grid was found to give results similar to those of larger grids. (In most cases, a 38x38 grid was also sufficient, but anything smaller would deviate.) For the full simulations discussed in Chapter 5 a 40x40 grid was used unless otherwise noted.

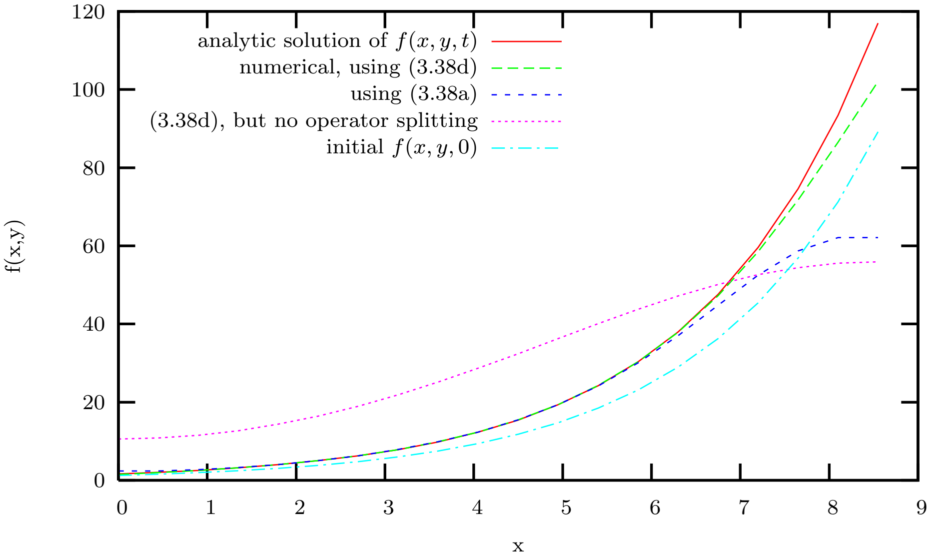

Figure 4.12: Comparison of different finite-differencing and boundary-condition schemes. The solid line is the analytical solution. The long- and medium-dashed lines are numerical solutions using (3.38d) and (3.38a) respectively. The dotted line is a numerical solution also using (3.38d) but in which (3.35) is solved all at once, *i.e.,* without the operator-splitting of (3.36). The dot-dashed line is the $t = 0$ starting solution. The numeric solutions are shown after 30 timesteps. (Doubling Δt and halving the number of timesteps produced curves indistinguishable from those shown for the (3.38d) and (3.38a) cases, and a similar curve for the no-operating-splitting case that was only slightly different.) This figure plots a slice midway through the range of y used.

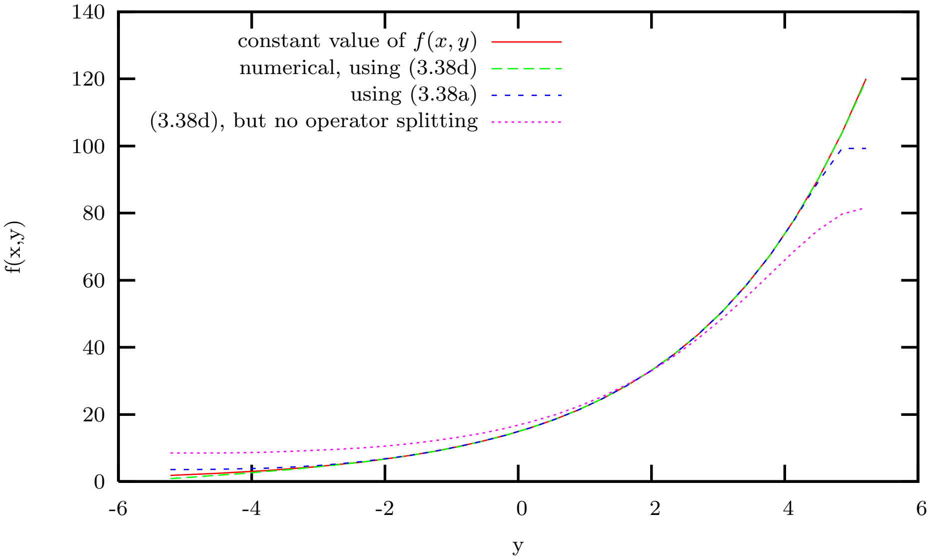

Figure 4.13: As in Fig. 4.12 but for the constant-coefficient, static-solution case (thus the analytic solution is identical to the $t = 0$ curve at all times). Here $f(x,y)$ is plotted for a slice at constant x.

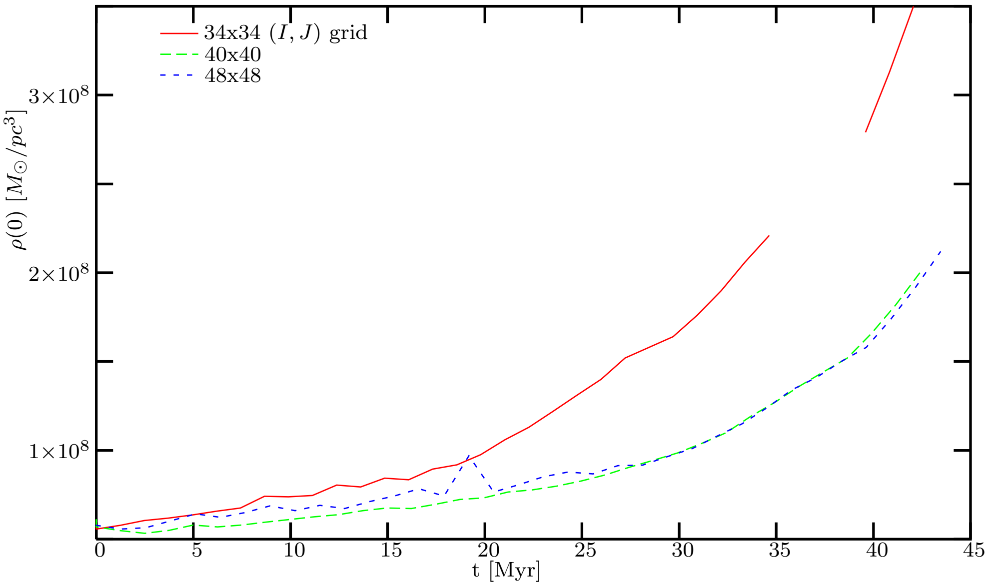

Figure 4.14: Comparison of results of using different-size (I, J) grids. The break in the otherwise smooth 34x34 curve is a spurious artifact of the output-calculating procedure and does not affect the model's further evolution. All other parameters are the same for each run, and each is constrained to having the same r^2 grid as the others at each timestep. (These runs used r^2 grids with 50 points, a maximum l value of 9 and a 3-component Miller-Scalo-style IMF.)

Choosing an even larger action-space grid is possible, but at a considerable expense of computer time: in addition to simply having more gridpoints on which all calculations need to be performed, in practice a larger grid also required smaller timesteps in order to remain stable.

4.7.2 Grid size: radial coordinate

The situation for the grid in the radial coordinate r^2 is similar to that for (I, J) in that a too-course grid produces evolution that is unstable to accumulated errors; however using a too-fine r^2 grid results in noisy output which causes difficulty for the model's mechanism of solving for the new potential $\Phi(r)$. In practice a 50-point r^2 grid produced a good balance between these two effects, although sometimes at the cost of an artificially slower evolution of the model as compared to finer grids.

Figure 4.15 shows a comparison of use of 44-, 50- and 60-point grids. The 60-point case evolves slightly more quickly, as there are more gridpoints near the dense center of the cluster, but it aborts after only a few timesteps. Unlike the (I, J) grid, which is fixed, in the simulations described in Chapter 5 the r^2 grid is allowed to dynamically self-update at each timestep, as described in §2.4.3. In practice the dynamic updating of the r^2 grid helps the finer-grid case adjust better than shown here: for the runs shown in Fig. 4.15 the dynamic updating was disabled in order to allow for a direct comparison between grid choices. Even finer 66- and 74-point grids were also tried and gave results similar to the 60-point case; a 99-point grid was found to require too-small timesteps with no benefit.

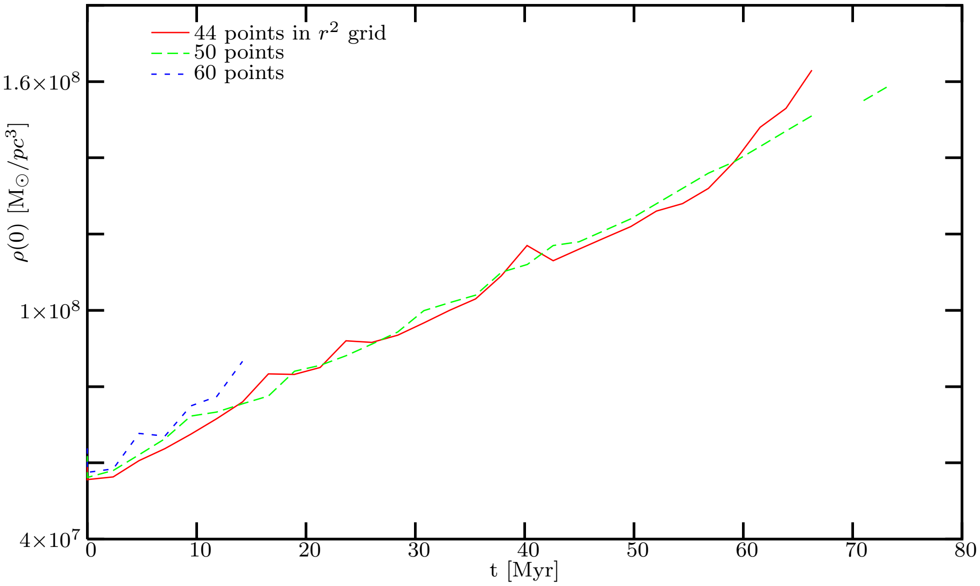

Figure 4.15: Comparison of runs with r^2 grids containing 44, 50 and 60 points; other parameters (timestep, action-space gridsize *etc.*) were identical for each of the 3 cases. In order to do a direct comparsion, each run was prevented from adjusting its r^2 grid dynamically. These runs were for model E4B with $l_{\max} = 4$ and an initial rotation parameter $\lambda = 0.05$. The mass spectrum was 90% $1M_\odot$ and 10% $2M_\odot$.

Courser grids of 40 or 36 points could not track the center of the cluster with sufficient resolution.

A somewhat more typical set of simulations is shown in Fig. 4.16, in which a broader range of stellar masses is used, a stellar bar is included and the r^2 grid is allowed to update itself dynamically at each timestep. This results in a much smoother evolution in which the 40, 50, 60 and 74-point r^2 grids give almost identical results, although in this case the simulation using a 74-point grid continued longer than the others. Note that the presence of the stellar bar and of a subpopulation of higher-mass stars caused much more rapid rate of change of the central density compared to that of the previous example.

Given these findings a 74-point dynamically-updating r^2 grid was the default choice. In some simulations a 60-point grid was used; those cases are noted.

4.7.3 Timestep size

The choice of timestep size necessarily varied with what cluster parameters were used as an initial condition; a typical choice was a fraction $\lesssim 0.1$ of the cluster's initial central relaxation time $t_{rq}(0)$ as given in §3.1: $\Delta t \lesssim 0.1 t_{rq}(0)$. (Stellar mass m_q here is that of the dominant stellar mass in the cluster.) The goal was to maximize the amount of evolution per timestep in order to avoid inordinately long computation times, while still having confidence in the run's results, as demonstrated by selectively performing similar runs with smaller timesteps and obtaining similar results.

Figure 4.17 shows that the simulation model is stable against different sizes of timestep,

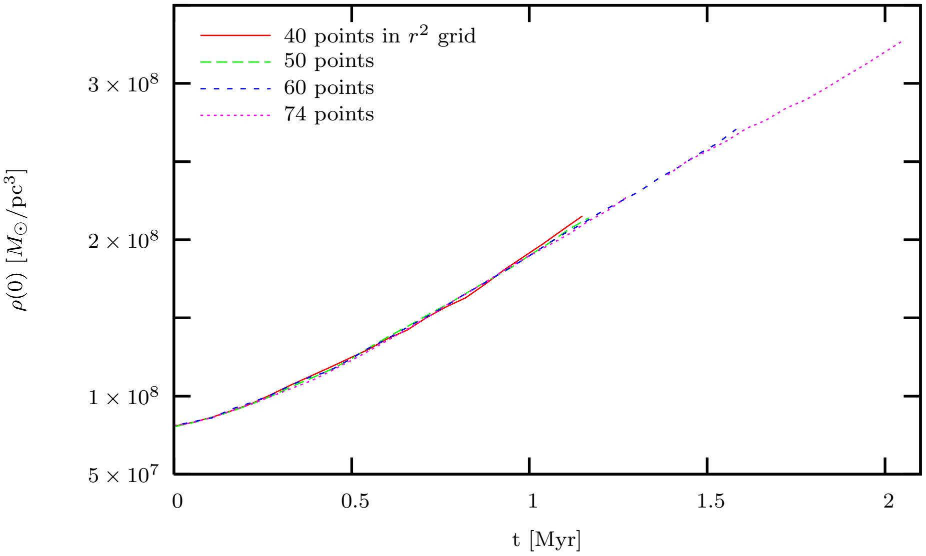

Figure 4.16: Comparison of runs with r^2 grids containing 40, 50, 60 and 74 points using a 40x40 (I, J) grid, $l_{\max} = 3$ and identical timesteps in each case. Dynamic r^2 grids were enabled, but were similar in all 3 cases. These runs were for a 7-mass component cluster (mass range $1-32M_\odot$ in 7 bins) with a moderate amount of rotation ($\lambda_0 = 0.015$), Kroupa-style IMF and with 1% of the cluster mass in a stellar bar.

although for different choices of timestep size the model may abort much earlier than for others (usually due to lack of numerical convergence when iteratively solving for the new Φ or ρ). When this happens, sometimes the run can be stably restarted from a slightly earlier time value but with a smaller timestep, sometimes not; for consistency all simulations presented are single runs starting at $t = 0$.

Figure 4.18 displays a set of runs more representative of the actual simulations presented in Chapter 5: a wider range of masses is employed with a realistic IMF, and the r^2 grid is dynamically updated as the system evolves. Here the largest choice of Δt, while smooth, clearly does not track the progression of the central density as accurately as do the smaller timestep choices. The similarity of the two smaller-timestep runs is interpreted as indicating a reliable outcome, with the extra rise of $\rho(0)$ at $t = 4 - 5$ Myr for the $\Delta t = 0.38$ Myr run taken to be a spurious artifact of the output-calculating procedure as seen earlier. (Note that in any case for the upper end of this range of masses the main-sequence lifetime is only 3 Myr [55].) The full span of the largest-Δt run is shown in Fig. 4.19 for comparison, and shows that the $\Delta t = 0.38$ run becomes possibly unstable around $t \gtrsim 8$ Myr; if reliable results near or after $t \simeq 7$ Myr had been required, it would have been necessary to attempt another smaller-Δt run as confirmation.

4.7.4 Number of expansion terms

The main physical quantity that is approximated by a series expansion is the gravitational potential $\Phi(r)$, which is expressed in term of spherical harmonics, as shown in (3.14). This results in the diffusion coefficients also being representated in the form of a series, as in

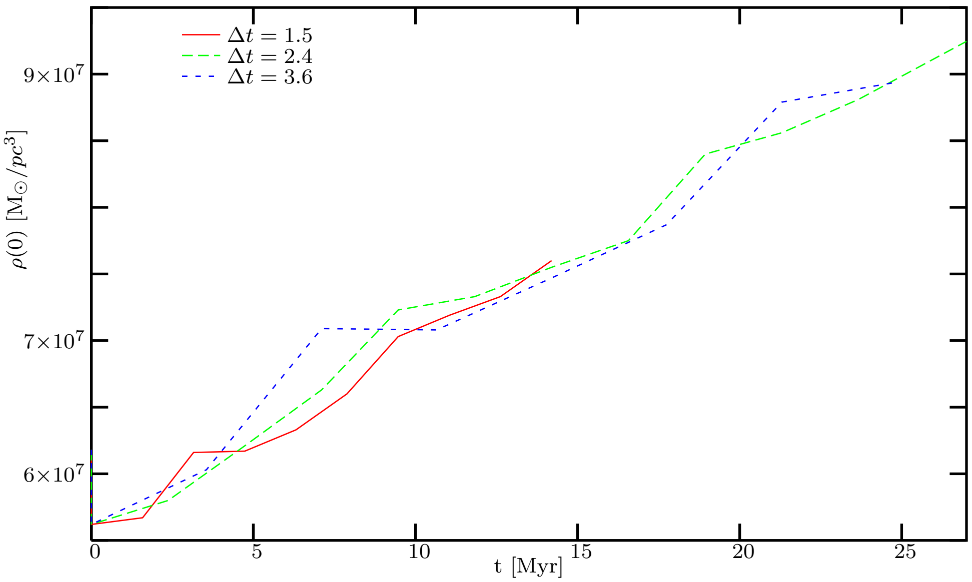

Figure 4.17: Comparison of the use of different-size timesteps. Runs are with r^2-grids of 50 points; the $\Delta t = 2.4$ case continued until $t \simeq 73$, as can be seen in Fig. 4.15. Initial central relaxation time was $t_{rq}(0) = 25.6$ Myr.

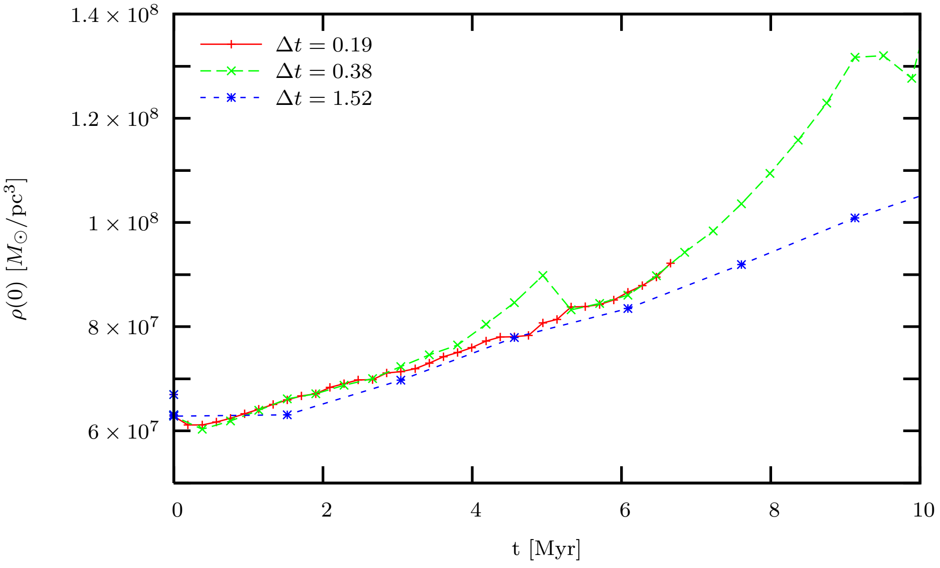

Figure 4.18: Comparison of the use of different-size timesteps. Runs are of model E4B with an r^2-grid of 74 points, a 40x40 (I, J) grid, a maximum l of 2 and an initial rotation given by $\lambda_0 = 0.05$ and a Kroupa IMF with 10 stellar mass bins in the range $1 - 125 M_\odot$. Initial central relaxation time for the $1 M_\odot$ stars that comprised 51% of the cluster was $t_{r1}(0) = 158$ Myr; for comparison the $25 M_\odot$ stars' mass fraction was 2.4% with $t_{r7}(0) = 0.23$ Myr.

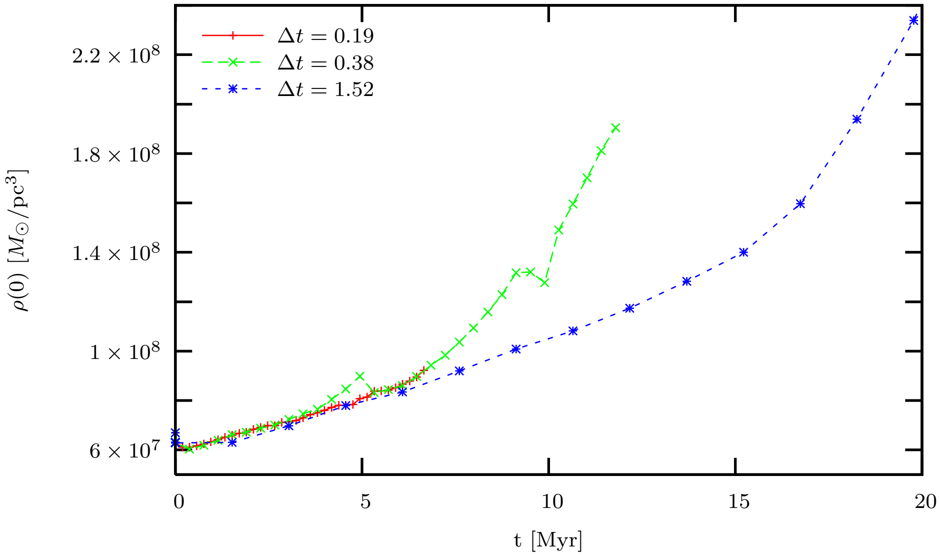

Figure 4.19: Similar to Fig. 4.18 but showing the full range of the $\Delta t = 1.52$ run.

(3.23); physically, cutting off the series at a certain maximum spherical harmonic index $l = l_{\max}$ implies only summing over orbital resonances of indices equal to or lower than $l_{\max}$. Figure 4.20 shows comparisons of different choices of $l_{\max}$ value. The cases with $l_{\max}$ of 3 (*i.e.,* up to the octopole term in the potential) and 4 do not capture the full effect of the dynamical friction, which can be seen as starting to converge around $l_{\max}$ of 5. (The choices $l_{\max}$ =3, 4, 5 and 6 correspond to a total number of expansion terms of 4, 6, 9 and 12 respectively when all allowed values of index m are included.)

When a strong stellar bar perturbation is incorporated in the potential, the situation can change dramatically. As comparing the timescale of Fig. 4.21 with that of Fig. 4.20 shows, the bar potential dominates the dynamical friction. Physically this is because the bar is a bulk perturbation whose total contribution to the potential is squared (*cf.* (3.34)) whereas "field" (*i.e.,* non-bar) stars only enter individually into the Ψ^2 factor in (3.23) before having their contributions summed over. Because the bar is by construction quadrupole-only and its dynamical friction timescale is so much shorter than that for field stars, higher-order resonances become almost irrelevent to the overall cluster evolution. So, unless otherwise stated all simulations which incorporate a bar perturbation have $l_{\max} = 2$ or 3, which results in only two or four terms in the expansion (*e.g.,* the two terms $m = \pm 2$ if $l_{\max} = 2$).

For the "production" simulations presented in Chapter 5 it was found that the choice of $l_{\max}$ needed to be determined on a case-by-case basis. Figure 4.22 shows the behavior of one of the main models used, model E4B with a Kroupa IMF and a stellar bar with mass fraction of 1%. In this case $l_{\max} = 2$ gave a fairly steady evolution; higher values of $l_{\max}$ hinted at a faster increase in central density $\rho(0)$ but were not sufficiently numerically stable to be considered robust; this $l_{\max} = 2$ was used, understanding that it may not quite capture the full progress of the system.

Contrasting this is the case shown in Fig. 4.23, of model E2A with a Kroupa IMF. Here it is clear that $l_{\max} = 5$ and $l_{\max} = 6$ give almost identical results, while $l_{\max} \leq 4$ misses a considerable degree of the increase in central density. However, the $l_{\max} \geq 5$ runs terminate

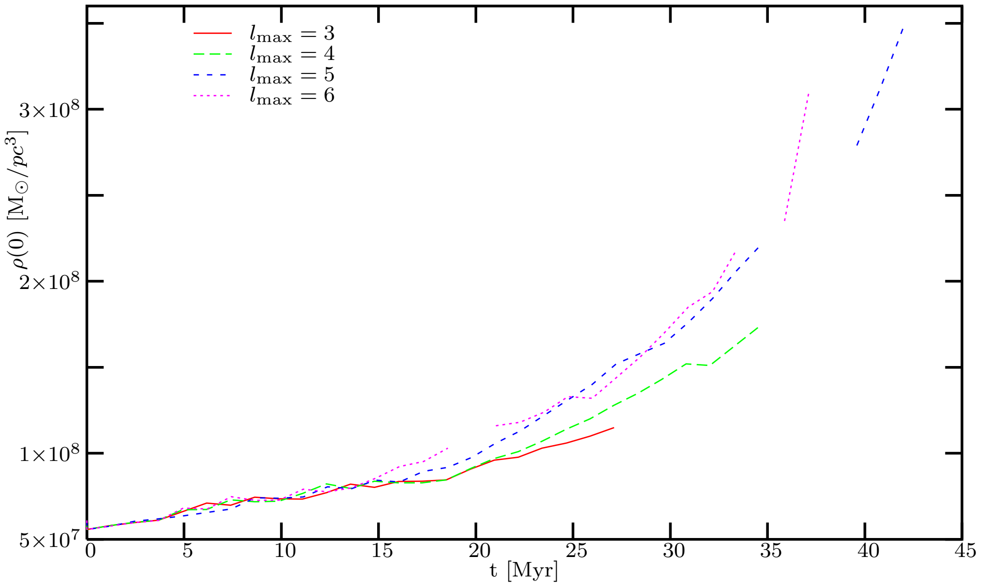

Figure 4.20: Demonstration of the use of an $l_{\max}$ value of 3, 4, 5 and 6 in the diffusion coefficient expansion (3.23) (and by extension in the expansion of the potential (3.14)); this corresponds to a total number of expansion terms of 4, 6, 9 and 12 respectively. This plot is of a rotating 2-component cluster similar to those of Figures 4.15 and 4.17. Each run was constrained to use the same timestep and the same r^2-grid as the others, to allow for direct comparison.

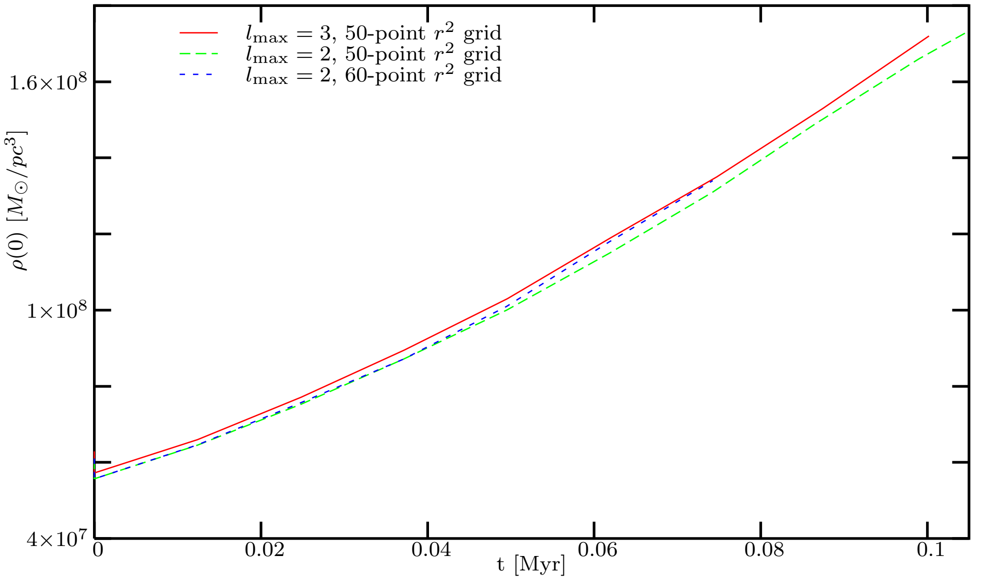

Figure 4.21: Comparison of different numbers of expansion terms and r^2 gridpoints for the case including a bar perturbation. The 50-point, quadrupole-only run continued on a fairly linear path (not shown) until it reached $\rho \simeq 2 \times 10^8$ at $t = 0.2$. This run is for a 2-component cluster with 10% $2M_\odot$ and 90% $1M_\odot$ stars, 4% of which comprised the bar. The rotational parameter was $\lambda = 0.075$. A run with $l_{\max} = 4$ (*i.e.,* 6 expansion terms) was nearly identical to the one shown using $l_{\max} = 3$, although it ended earlier.

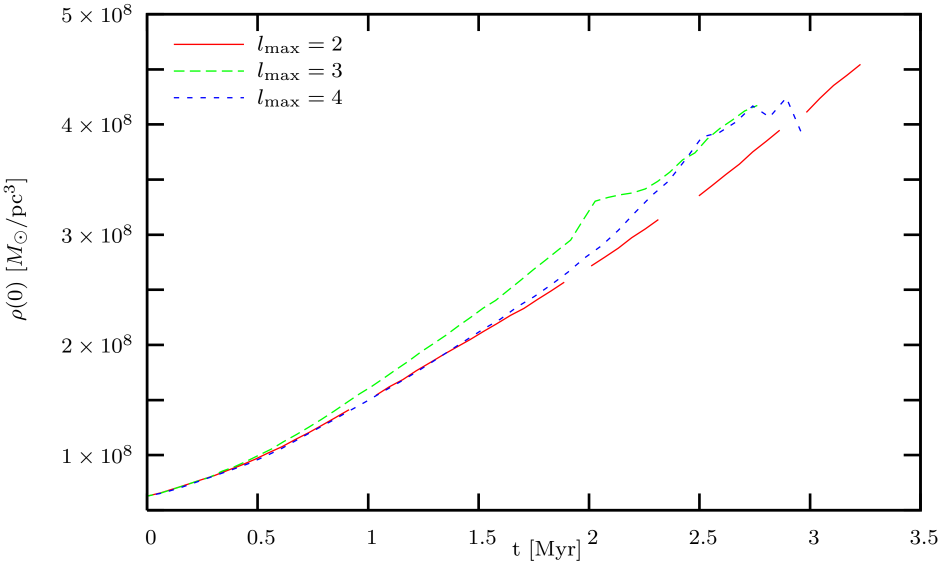

Figure 4.22: Comparison of different numbers of expansion terms for a "production" run: model E4B with a Kroupa IMF, initial rotation parameter $\lambda_0 = 0.05$ and a stellar bar mass fraction of 1%. The stellar mass range is $1 - 125\ M_\odot$.

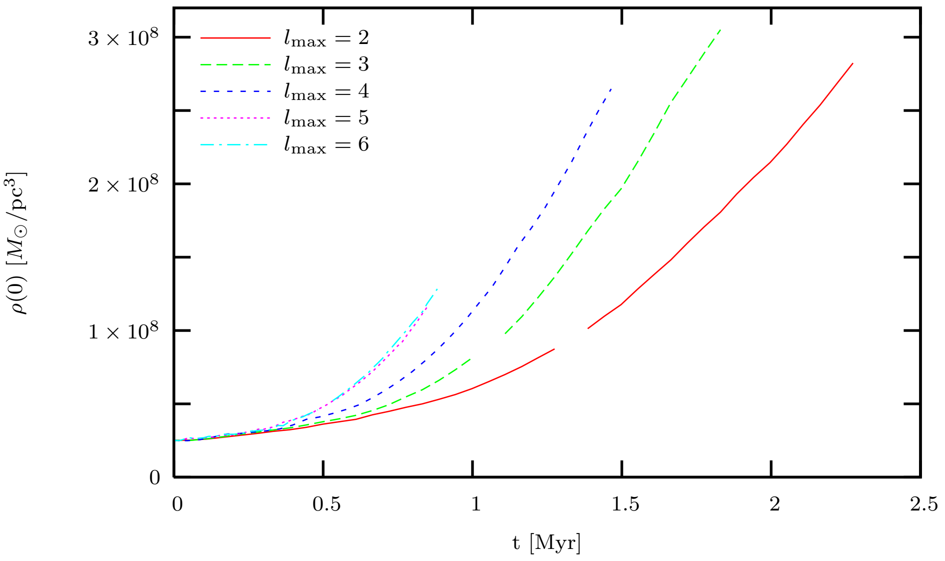

Figure 4.23: Comparison of different numbers of expansion terms for a "production" run: model E2A with a Kroupa IMF, initial rotation parameter $\lambda_0 = 0.05$ and a stellar bar mass fraction of 1%. The stellar mass range is $1 - 125\ M_\odot$.

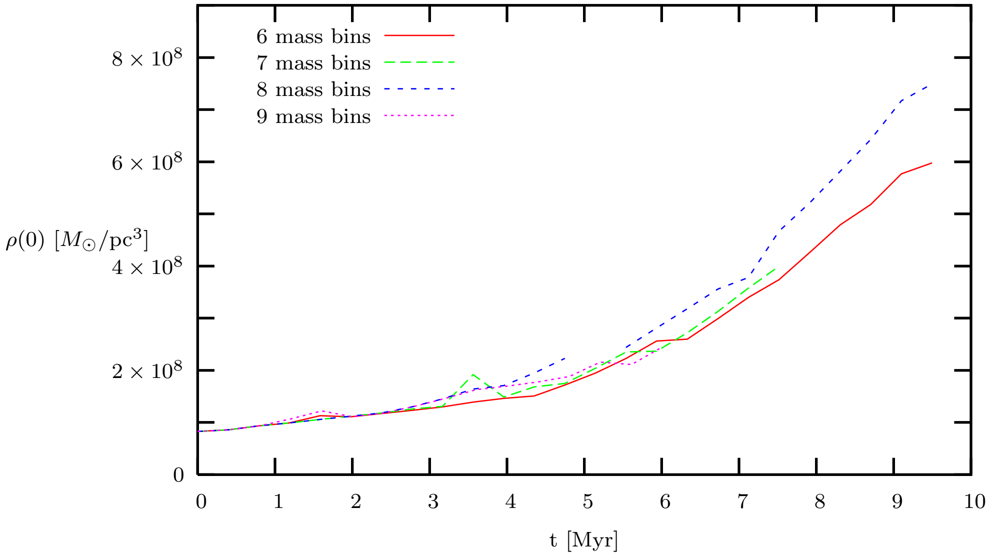

Figure 4.24: Comparison of central density for runs in which the stellar mass spectrum is split into the number of geometrically-spaced mass bins shown. All are model E4B with 50 points in the r^2 grid, a 34x34 (I, J) grid, and an overall mass range of $1 - 32M_\odot$. The IMF is Salpeter-like, but with the uppermost mass bin artifically overpopulated in order to accentuate any differences in the cases.

quite a bit earlier than the lower-$l_{\max}$ simulations, and require considerably more computational time (weeks compared to days) to even get as far as they do. (Taking $l_{\max} = 4$ produces a total of 6 terms in the expansion, while $l_{\max} = 6$ gives 12 terms). So in the case of model E2A/Kroupa an $l_{\max}$ of 4 was used, but only to give an indication of how the system evolves and not to show precise quantitative results.

The above two examples are the extreme cases of those studied; in general the bar perturbation proved dominant enough that $l_{\max} = 2$ or 3 sufficed to describe the system's evolution while still allowing for numerical stability and adequately-short computation times.

4.7.5 Mass Spectrum: Discretizing the Initial Mass Function

The initial mass functions described in §2.6.3 are continuous functions of the individual stellar mass m; in order to separate a given IMF into discrete bins of discrete stellar mass m_q the IMF is simply integrated over the range of masses being considered, with the boundary between bins being the midpoint between them. The lowest mass bin's lower limit is taken to be $\frac{1}{2}m_1$, and the highest-mass m_Q bin is given the same width as it would have if there was an m_{Q+1} bin with the same scaling. As the spectrum of mass bins is geometrically increasing by a factor $m_{q+1} \lesssim 2m_q$, this prescription artifically biases the numerical IMF slightly towards lower masses as compared to the analytical form on which it is based.

Figure 4.24 shows the evolution of the central density for test runs of model E4B with a stellar mass range between 1 and $32M_\odot$ and a Salpeter-like IMF in which the $32M_\odot$ bin has been overpopulated, so that high-mass stars drive the system's evolution more strongly than would occur naturally. Figure 4.25 simlarly shows results for the same runs, but for the

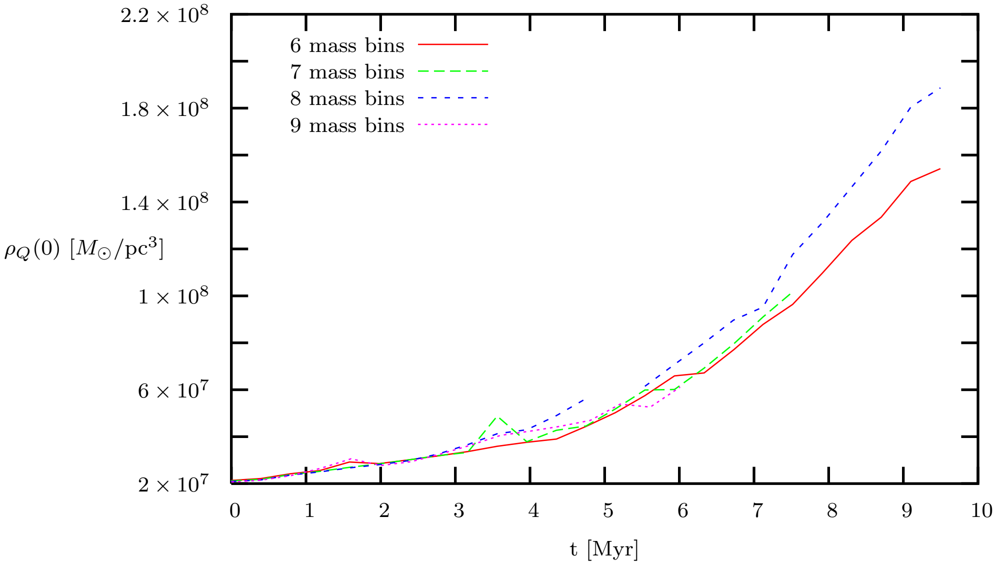

Figure 4.25: Similar to Fig. 4.24 but showing the central density of the uppermost ($m_Q = 32 M_\odot$) stellar mass bin only.

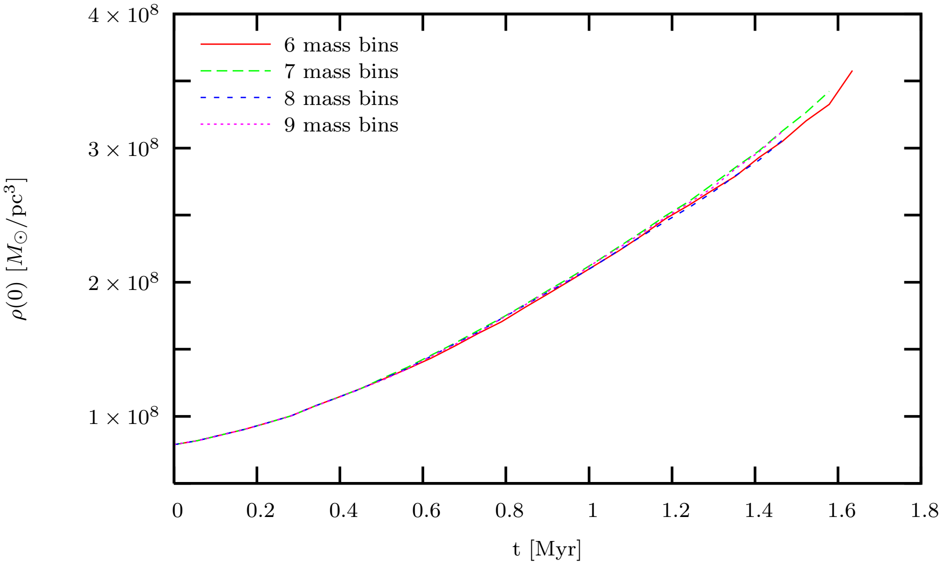

Figure 4.26: Comparison of central density for runs in which the stellar mass spectrum is split into the number of geometrically-spaced mass bins shown. All are model E4B with 50 points in the r^2 grid, a 34x34 (I, J) grid, and an overall mass range of $1 - 32 M_\odot$. The IMF is Kroupa-like, similar to that used in the later full simulations. Rotation and a stellar bar are included, with a bar mass fraction of 1% and an initial rotation parameter of $\lambda_0 = 0.016$.

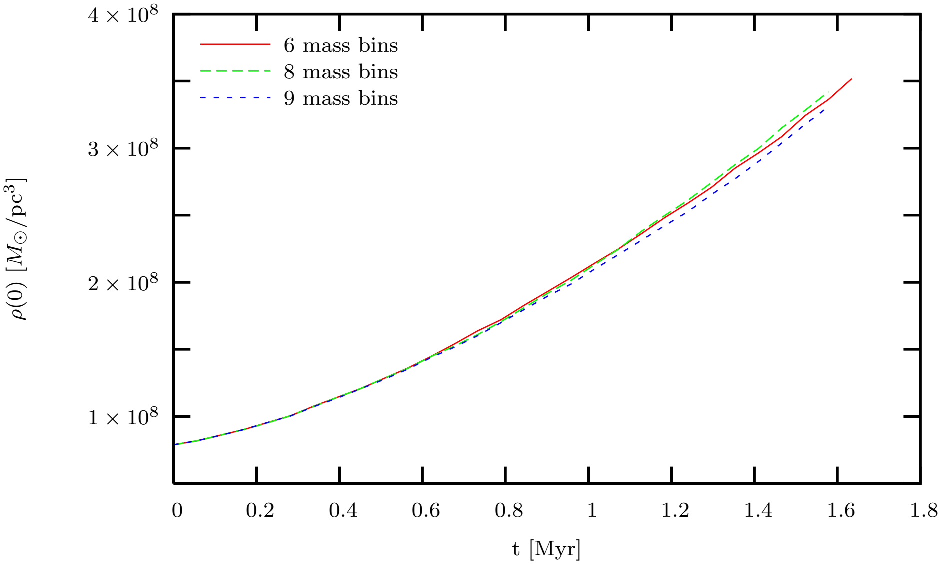

Figure 4.27: Similar to Fig. 4.26 but with stellar mergers included in the simulation. (The 9-bin case shows slightly lower central density than the others due to it following a somewhat different series of values for the innermost point of its dynamic r^2 grid; this was unusual and did not occur during the full simulations presented in Chapter 5 except where noted.)

$32M_\odot$ stellar mass bin only. Although the curves of the central density in this artifical case are somewhat rough, a choice of 7 or 8 mass bins seems to best show the full evolution of the system while avoiding the numerical instability displayed by the 9-bin run. For a range of $1 - 32M_\odot$, 7 bins corresponds to a mass ratio between adjacent bins of 1.78, and 8 bins gives a ratio of 1.64. The full simulation runs shown in Chapter 5 used adjacent-bin mass ratios of 1.71 for the Kroupa IMF and 1.675 for the Arches-style IMF. This is similar to and consistent with the findings of Amaro-Seoane [41] who found that an average mass-bin ratio of $\simeq 1.72$ was sufficient to model the Fokker-Planck evolution of clusters of $0.2 - 100M_\odot$ stars which contained already-formed or primordial massive central objects.

Figures 4.26 and 4.27 show the central density evolution using a more-realistic Kroupa-style IMF and incorporating a stellar bar and a moderate amount of rotation. It can be seen that both with and without stellar mergers, the choice of number of stellar mass bins does not strongly affect the simulation results; as seen in Fig. 4.27 the dynamic r^2 grid can sometimes have a greater, but still small, effect.

4.8 Binary Heating

As binary stars' orbits harden as a result of encounters with other stars, the system is effectively heated by the energy tranferred from the binary system to the cluster as a whole, and potentially disrupted somewhat. The physics of binary heating is not incorporated in the simulation code; however, an upper limit on its possible effect was calculated for a selection of models using a direct application of the binary-heating formulation of Quinlan and Shapiro [7]. The extreme assumption that all stars in the cluster are in binaries and that all binaries harden forever (*i.e.,* none are soft binaries that do not harden, and none are

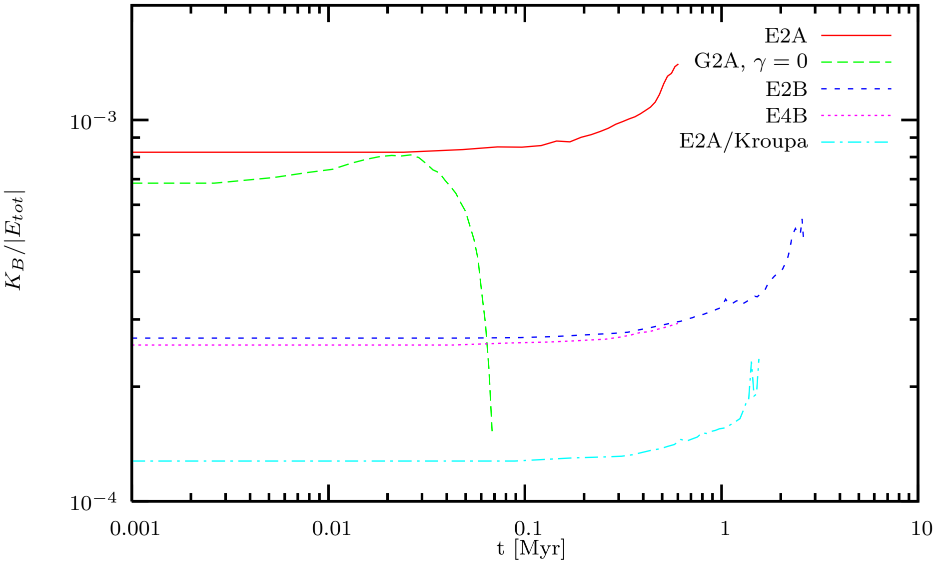

Figure 4.28: Ratio of maximum possible amount of binary heating K_B to overall (gravitational + kinetic) system energy $|E_{tot}|$ for a variety of models. All used the Arches-style IMF except for model "E2A/Kroupa", and all had an initial Plummer-sphere distribution except for "G2A", which started as a $\gamma = 0$ sphere. A logarithmic scale is used so that different models' timescales can be plotted together; exactly flat lines at the start of each model's plot indicate the first timestep in the calculation, before which binary heating is zero.

themselves disrupted by interactions with other stars) is used to give an absolute limit on how much binary heating could possibly have occured, had it been included in the dynamical calculations.

The results are shown in Fig. 4.28, from which it can be seen that at all times (except at the very end of one simulation) the upper limit of possible binary heating remains a small fraction, 10^{-3} or less, of the overall system energy. Given that not all stars can be in binaries, that not all binaries will harden and heat the system, and that some binaries will be disrupted by encouters and by internal merging, the assumption that binary heating can be neglected holds.

Chapter 5

Results

5.1 General Considerations

5.1.1 Initial Models

The progression of initial potential-density models studied starts with the E2A and E2B Plummer spheres which, as listed in Table 2.1 and described in §2.6.1, can represent the nuclei of dwarf elliptical or bulgeless spiral galaxies, as well as very dense globular clusters. Model E4B is somewhat more massive and more dense with a higher initial velocity dispersion, along the lines of the core of a giant elliptical galaxy. Model E4A starts out even more dense than E4B, but was found to be numerically unstable and so could not be simulated.

An alternative to the Plummer sphere is an initial "$\gamma = 0$ sphere", which can represent galactic spheroids. When mapping the masses and core radii of the above Plummer models onto $\gamma = 0$ spheres, only the resulting model G2A (analogous to Plummer model E2A) has an initial central density and velocity dispersion within the ranges required for modeling dense astrophysical systems without assuming an unreasonably high central density. However, a model G3C, which starts with the highest initial density studied, was run for comparison purposes, as was intermediate model G3A; these are described later in this Chapter.

For each of the various potential-density models the distribution of stellar masses is given by the initial mass function (IMF). As developed in §2.6.3 the base IMF used here is the Kroupa IMF, which is a good fit for the observed galactic stellar population in general. The second is an Arches-style IMF which has recently been determined to specifically describe the stellar distributions in dense clusters observed at the centers of galaxies, *i.e.,* the objects being simulated here. Unsurpringly the Kroupa IMF is more weighted towards lower-mass stars while the Arches IMF has a flatter distribution of masses, but still with an upper cutoff of $m \lesssim 150 M_\odot$.

5.1.2 Physical Effects

For sufficient buildup of the cluster's density to occur so that a massive object can form as the product of collisional mergers of stars, rotational support against the infall of stellar material to the central region must be overcome. As described in Chapter 1, given infinite time such collapse would be expected even in a nonrotating system if it is sufficiently dense. However, if it is to occur before the majority of stars reach the end of their main-sequence lifetimes – only 3 Myr for the largest stars in the IMFs considered – something in addition to interactions of individual stars is expected to be required. Fortunately, even though they were not intentionally constructed to be so, when given a reasonable value for the rotation parameter λ the models chosen all satisfy the criterion for being unstable against the creation of a bar-like perturbation in the stellar potential; this stellar bar can (hopefully) then transport angular momentum in bulk from the system's center to its outer regions.

Thus the problem is essentially a race against time: can gravitational effects, aided by the transport of angular momentum afforded by a stellar bar, allow the system's central regions to condense quickly enough that collisional mergers can in turn produce an object of $M > 250M_{\odot}$ which will evolve into a massive seed black hole, all before the stars outlive their main-sequence lifetimes. Simulations for each model and IMF are performed incorporating all of: an overall rotation of the system, a stellar bar, and collisional mergers of stars. In order to examine the various effects some simulations are also presented in which the system is not rotating, it does not contain a stellar bar, and/or the stars are not allowed to collide and merge to produce larger stars.

After a brief summary of what constituted successful simulations and which models yielded them, the remainder of this chapter reports results for each choice of initial model and IMF.

5.1.3 Overview

Of the five possible initial models whose density-potential pairs are listed in Table 2.1 and described in §2.6.1 (or eight, when models G2A, G3A and G3C from §2.6.1 are included), only some yielded simulations which satisfied all the criteria for stability. The criteria included:

- giving consistent results with different choices of timestep size Δt;
- there existing a value of the cutoff $l_{\rm max}$ of the potential expansion high enough that terms $l > l_{\rm max}$ did not contribute significantly (as indicated by the central density's increase with time), but low enough to avoid the numerical instability higher values of $l_{\rm max}$ exhibited; and
- giving consistent results across the Sun, Alpha, and AMD/x86 computer architectures.

The models which met the above criteria are listed in Table 5.1. For each model a possible astronomical system to which it best corresponds (if any) is listed; this expands on the descriptions given in §2.6.1, which did not yet take choice of IMF into consideration. *Table 5.1 can serve as a convenient reference for the model-specific sections that comprise the remainder of this chapter.*

Simulations with a Kroupa IMF used a mass range of $1 - 125M_{\odot}$; the Arches IMF has a smaller contribution from low-mass stars (*cf.* §2.6.3) and so employed a range of $2-125M_{\odot}$. The set of possible variations of each model include: whether the system has initial rotation (with the rotation parameter set near the canonical value $\lambda \simeq 0.05$ as described in §2.6.2) or not ($\lambda = 0$); if rotating, whether there is a stellar bar or not; and whether stellar mergers

Initial Profile	Model Set	IMF	Corresponding Astronomical System
Plummer sphere	E2A/E2B	Kroupa	dwarf elliptical or bulgeless spiral galaxy nucleus
		Arches	nuclear cluster
	E4B	Kroupa	core of a giant elliptical galaxy
		Arches	*none – inappropriate IMF*
$\gamma = 0$ sphere	G2A	Kroupa	galactic spheroid or spiral bulge
		Arches	*none – inappropriate IMF*
	G3A	Kroupa	galactic spheroid or spiral bulge
	G3C	Kroupa	*none – initial central density too high*

Table 5.1: Models which yielded stable results, along with possible analogous astronomical systems.

are allowed or not. For both the Kroupa and the Arches-style IMF, two models produced stable results in at least some cases: E2A and E4B. The E2B model was stable as well but only when using the Arches IMF. The results for each model are described in detail below, and are summarized in Tables 5.2 and 5.3 for the "full" runs which included rotation and stellar mergers.

5.2 Plummer-sphere Models

5.2.1 Model E4B, with Kroupa IMF: Core of a Giant Elliptical Galaxy

With $M_{\text{tot}} = 3.1 \times 10^7 M_\odot$ and an initial $\rho(0) = 10^8 M_\odot/\text{pc}^3$, Model E4B was the largest overall system studied with the highest central density. (Model E4A had an initial $\rho(0) = 3 \times 10^8 M_\odot/\text{pc}^3$ but did not yield stable simulations.) Figure 5.1 shows the change in central density $\rho(0)$ with time for the four cases: (1) nonrotating, (2) rotating, (3) rotating, with a stellar bar and (4) rotating, with a stellar bar and star-star mergers. The stellar bar clearly influences the system evolution greatly as it produces an increase in $\rho(0)$ of more than an order of magnitude in the 3 Myr main-sequence lifetime of the largest stars in the system, compared to an increase of $< 5\%$ for the cases without a stellar bar. Contrasting this is the effect of stellar mergers, which is minimal: at the end of the simulation that included mergers, $< 0.1\%$ of the lowest-mass stars had undergone a collision and merger, while fewer than 30 of the highest-mass ($125 M_\odot$) stars had been created from collisional mergers – although this increase is still visible in Figure 5.2. Table 5.2 shows that G_{250}, the calculated rate of producing $250 M_\odot$ stars by stellar mergers, is effectively zero even after 3 Myr; similarly, the production rate for all stellar masses larger than $200 M_\odot$ was also found to be zero.

Looking at the system as a whole, Fig. 5.3 shows the range of $\rho(r)$ over the full radial extent of the stellar cluster at the starting time, the midpoint, and the end of the simulation. As the core contracts and gets more dense, the outer regions lose stellar density accordingly. Interestingly, when the model is run without including a stellar bar no such change in density is observed in the outer regions; this behavior is seen in general for all the Plummer-sphere models studied and is attributable to the stellar bar transporting angular momentum outwards, which rarifies the outer regions while allowing the inner regions to lose rotational support and contract.

Model	$l_{\max}$	bar	$\rho(0)$ at t_0	$\rho(0)$ at t_1	t_1	$G_{250}(t_0)$	$G_{250}(t_1)$	$S(t_1)$	$S'(t_1)$
E2A	4	5.3%	2.5×10^7	3.0×10^8	1.6	0.	0.	0.82	0.83
E4B	2	none	6.3×10^7	7.0×10^7	2.5	0.	0.	1.09	1.11
	”	none	”	7.2×10^7	3.0	0.	0.	1.09	1.11
	”	10%	”	3.8×10^8	2.5	0.	6×10^{-5}	1.006	1.02
G2A	4	none	1.7×10^7	1.3×10^8	.42	0.	0.	0.95	0.96
G3A	3	7.7%	4.4×10^7	1.1×10^9	.46	0.	0.	0.87	0.87
G3C	2	7.7%	1.4×10^8	3.0×10^9	.16	0.	0.	1.04	1.04

Table 5.2: List of models which yielded stable simulations using the Kroupa IMF, with central-density and merger-rate results shown at various times t_1 in each simulation (usually the endpoint, or chosen for comparison with another). Times are in Myr, central densities in $M_\odot/\text{pc}^3$, and each simulation had a rotation parameter $\lambda \simeq 0.05$. G_{250} is the calculated rate at which $250 M_\odot$ stars would be expected to be produced via collisional mergers of $125 M_\odot$ stars, in Myr^{-1}. S is the factor by which the central density of the most-massive stars increased as a result of mass segregation, relative to the overall increase in central density; S' is the same factor but also including the effects of stellar mergers. For the G2A and G3C models, the presence or absence of a stellar bar had little effect on the system's evolution. (Model G3A was not run without a stellar bar but it was also expected to have little effect.) Note that the central densities listed here are for the initial models with rotation, and so are lower than those given in Table 2.1. Starting time $t_0 = 0$.

Model	$l_{\max}$	bar	$\rho(0)$ at t_0	$\rho(0)$ at t_1	t_1	$G_{250}(t_0)$	$G_{250}(t_1)$	$S(t_1)$	$S'(t_1)$
E2A	3	5.3%	2.5×10^7	3.7×10^8	.63	3.9	6.5	1.	1.004
E2B	2	none	8.2×10^6	1.4×10^7	3.0	2.2	2.7	1.	1.01
	”	5.3%	”	6.1×10^7	3.0	2.6	3.3	1.	1.01
	3	”	”	9.4×10^7	3.0	2.6	3.7	1.	1.01
	”	8%	”	5.3×10^8	2.7	2.3	8.3	1.	1.006
E4B	2	8%	6.3×10^7	1.0×10^9	2.2	35.	73.	1.	1.015
	”	10%	”	1.5×10^9	2.0	33.	69.	1.	1.025
G2A	3	none	1.7×10^7	1.3×10^8	.07	1.2	(1.9)*	1.	1.
	”	7.7%	1.7×10^7	1.3×10^8	.07	1.1	1.6	1.	1.

Table 5.3: Similar to Table 5.2, but for simulations using an Arches-style IMF. [*] indicates that the value of $G_{250}(t_1)$ calculated for model G2A without a bar is considered to be numerically suspect.

Model	$l_{\max}$	bar	$\rho(0)$ at t_0	$\rho(0)$ at t_1	t_1	$G_{250}(t_0)$	$G_{250}(t_1)$	$S(t_1)$	$S'(t_1)$
E2B	2	5.3%	8.2×10^6	2.9×10^8	4.3	2.6	5.1	1.	1.01
	3	”	”	4.9×10^8	4.3	2.6	8.7	1.	1.01

Table 5.4: Addendum to Table 5.3, showing results of simulations allowed to proceed beyond the 3 Myr lifetime of the largest main-sequence stars.

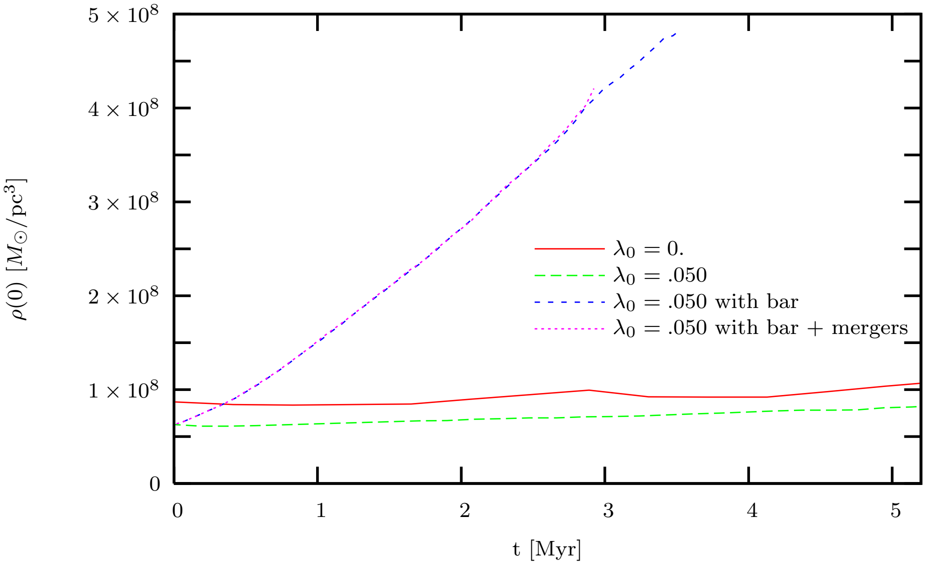

Figure 5.1: Central density v. time for model E4B using the Kroupa IMF with a stellar mass range of $1 - 125 M_\odot$ and $l_{\max} = 2$. Note that the nonrotating ($\lambda_0 = 0$) case has an initial $\rho(0)$ slightly higher than the cases that include rotation; this is a product of the method for introducing rotation into the initial distribution function of the system.

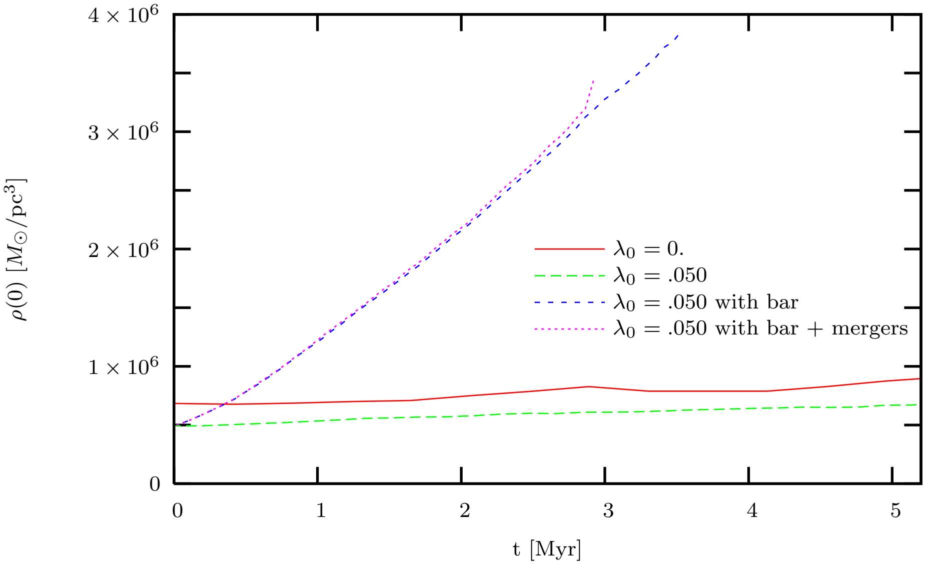

Figure 5.2: Similar to Fig. 5.1 but showing only the highest-mass stars ($m_Q = 125 M_\odot$). The upturn at the end of the "with mergers" curve is for the final timestep only and is a result of a numerical instability which casued the simulation to end early.

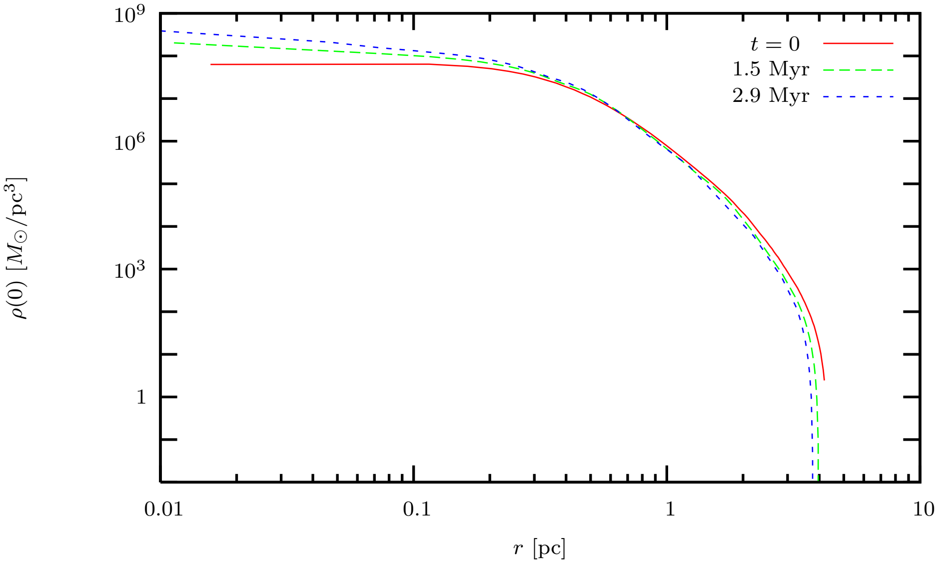

Figure 5.3: Density $\rho(r)$ v. radial distance r from the cluster center for the start, midpoint and end of the simulation for model E4B with a Kroupa IMF. Stellar mergers were enabled, as was a 10% stellar bar. Rotation parameter $\lambda = 0.05$ and $l_{\max} = 2$. (This corresponds to the uppermost plot of Fig. 5.1.) Note that the decrease in $\rho(r)$ with time for large r is only exhibited when a stellar bar is present.

The quantity S in Table 5.2 gives the relative increase in the central density of the highest-mass (125 $M_\odot$) stars relative to that of the overall system, and so indicates how much mass segregation has occured. S' is a similar measure but also includes the effects of stellar mergers on the increased density of high-mass stars. For model E4B without a stellar bar a moderate ($\lesssim$ 10%) relative increase in high-mass stars is seen, with mergers contributing another couple percent. (The nonrotating $\lambda = 0$ case, plotted in Figures 5.1 and 5.2 but not listed in Table 5.2, displayed similar behavior.) In this measure the presense of a stellar bar appears to strongly damp the relative mass segregation to a large extent, but that is merely due to how much more the overall central density increases with the bar (by a factor of $\simeq 6$ with the bar as compared to $\simeq 1.1$ without). Thus a small relative mass segregation in the with-bar case still corresponds to a larger absolute increase in the central density of the highest-mass stars. Still, it also means that the bar is dominating the dynamics and the different stellar-mass populations evolve less distinctly than in the no-bar case.

5.2.2 Model E4B, Arches-style IMF

When studying model E4B with the Kroupa IMF replaced by an Arches-style IMF the situation changes quite a bit. As shown by Figures 5.4 and 5.5 both the overall central density and that of the largest-mass stars increase much more rapidly even in simulations without a stellar bar: by 3 Myr, $\rho(0)$ is up by a factor of $\gtrsim 3-4$ for the nonrotating and rotating cases, even without the presence of a stellar bar. More dramatically, the systems which included a bar perturbation showed an increase in $\rho(0)$ of $\simeq 20\times$ or more, and in a shorter amount of simulated time (by 2 Myr when using a bar of strength 10%, by 2.2

Myr for an 8% bar[1]), after which the simulations were not able to track the evolution stably. This even more rapid increase is consistent with the Arches IMF being weighted more towards larger stellar masses than is the Kroupa IMF. The figures show that unlike in the Kroupa-IMF case, with an Arches IMF the stellar system achieves core collapse before the 3 Myr stellar-lifetime limit, with the central density $\rho(0)$ increasing exponentially: the run with a 10% bar and mergers was able to track somewhat into the core collapse stage before aborting, while the other with-bar runs reached the turning point of $\rho(0)$. Once core collapse begins, the central core of the cluster decouples dynamically from the outer regions – a situation the Fokker-Planck model is not constructed to deal with.

The stellar merger situation is also qualitatively different with an Arches IMF: as seen in Table 5.3, even at $t = 0$ the rate at which $250M_\odot$ stars were being produced from collisions of $125M_\odot$ stars was $\simeq 33\,\mathrm{Myr}^{-1}$, increasing to $\simeq 69\,\mathrm{Myr}^{-1}$ by the end of the simulation with a 10% stellar bar. Thus the Arches IMF was sufficient to produce $250M_\odot$ stars from the outset, although the presence of a stellar bar enhanced the production. It was also found that by $t = 2.0$ Myr, approximately 500 stars of mass $125M_\odot$ had been produced by collisional mergers of lower-mass stars.

Interestingly, the flatter Arches IMF seems to strongly damp any mass segregation when compared to the Kroupa-IMF case above: the relative increase in central density $\rho_Q(0)$ of the highest-mass stars tracks that of the overall system $\rho(0)$ to several significant figures, regardless of the presence or strength of the stellar bar. And again, the nonrotating $\lambda = 0$ case plotted in Figures 5.4 and 5.5 (but not listed in Table 5.2) displayed a similar complete lack of mass segregation; $S \equiv 1$ for all E4B/Arches cases. There is only a slight relative increase in $\rho_Q(0)$ due to stellar mergers, as shown by $S'(t_1) = 1.025$.

Figure 5.6 plots the stellar mass density $\rho(r)$ from the center of the cluster outwards. As in the Kroupa-IMF case, the stellar bar transports angular momentum from the inner regions to the outer, allowing core contraction and a corresponding decrease in stellar density at large r. A similar simulation but which does not include a stellar bar is plotted in Fig. 5.7. Without a bar to transport angular momentum outwards, there is very little decrease in stellar density in the outer parts of the cluster. Note that Figures 5.6 and 5.7 are plotted on the same scale for easier comparison.

Other quantities: bar composition, velocity dispersion and relaxation time

For the "with bar" runs presented in this chapter, the default was to populate the bar with stars from the lowest-mass stars considered in the simulation, *e.g.,* $2M_\odot$ stars when using the Arches IMF. As a test, the E4B/Arches "with bar" simulation was also run with a bar comprised of $1M_\odot$ stars. It gave identical results to the default case, which is expected since for purposes of its perturbing potential the bar is treated as a single bulk object and not as a collection of individual stars.

In addition to stellar density, several other quantities are tracked by the simulation code. Two of the most interest physically are the velocity dispersion σ_1 and the relaxation time t_r. Model E4B/Arches can be used as an example to show the results for these properties, which evolved similarly for all models.

---

[1] In model E4B the strength-10% bar case used the default 1% mass-fraction bar, as described in §3.6; the strength-8% bar corresponded to a mass fracton of 0.8%, below which the code had difficulty converging.

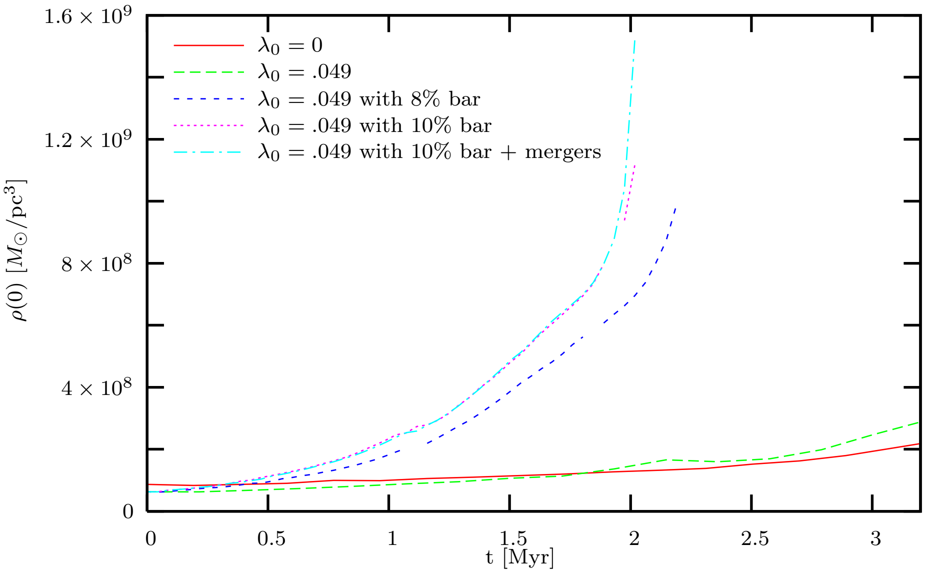

Figure 5.4: Central density v. time for model E4B using the Arches IMF with a stellar mass range of $2 - 125 M_\odot$ and $l_{\max} = 2$.

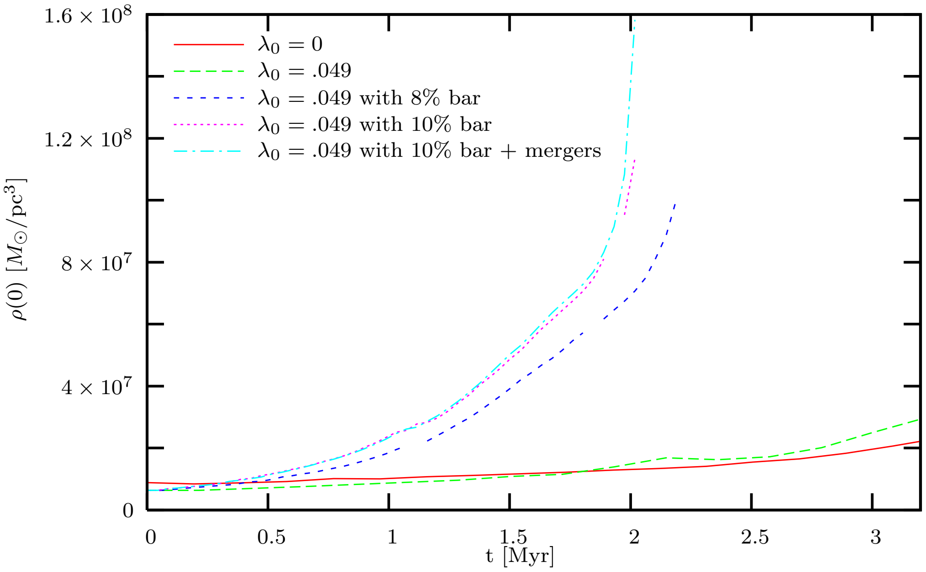

Figure 5.5: Similar to Fig. 5.4 but showing only the highest-mass stars ($m_Q = 125 M_\odot$).

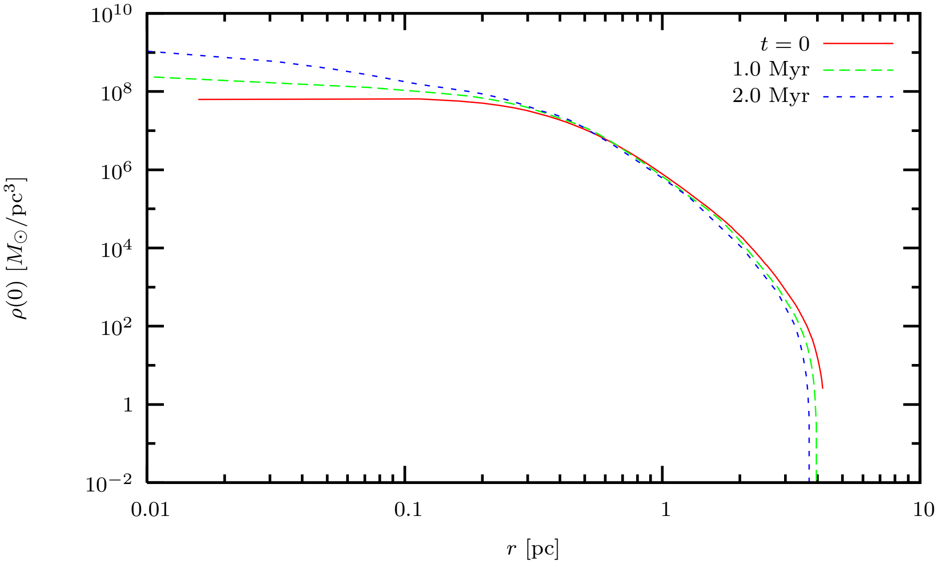

Figure 5.6: Density $\rho(r)$ v. radial distance r from the cluster center for the start, midpoint and end of the simulation for model E4B with an Arches-style IMF. Stellar mergers were enabled, as was a 10% stellar bar; rotation parameter $\lambda = 0.049$ and $l_{\max} = 2$. (This corresponds to the leftmost plot of Fig. 5.4.)

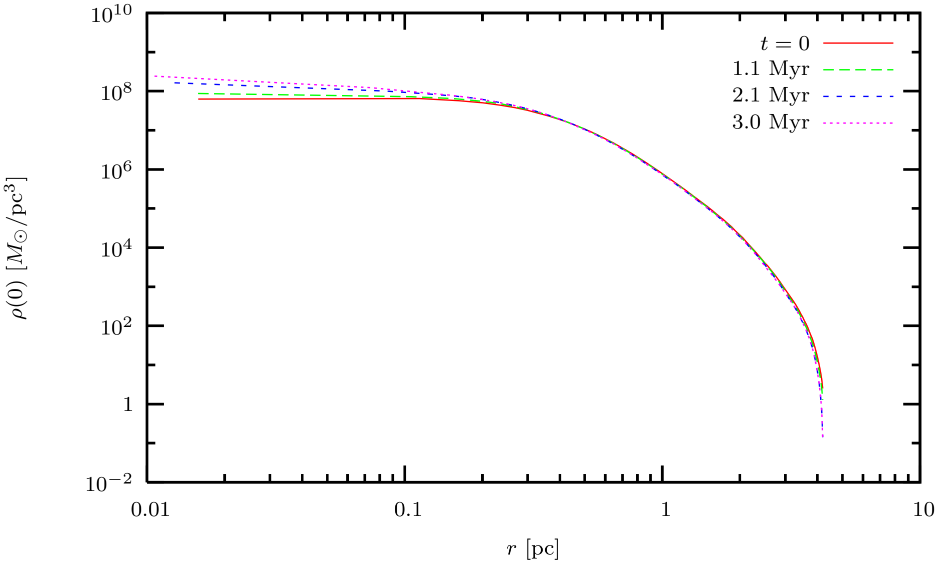

Figure 5.7: Density $\rho(r)$ v. radial distance r from the cluster center for the start and end of the simulation for model E4B/Arches, as well as at two similar times as those shown in Fig. 5.6 for comparison. Neither stellar mergers nor a stellar bar were enabled (in contrast with Fig. 5.6). Rotation parameter $\lambda = 0.049$ and $l_{\max} = 2$.

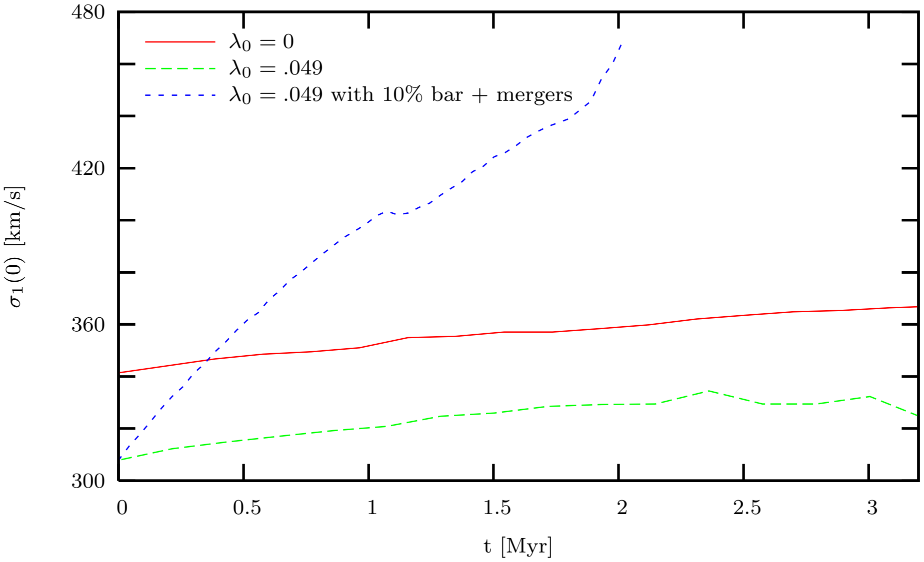

Figure 5.8: Central velocity dispersion $\sigma_1(0)$ v. time of the lowest-mass ($2M_\odot$) stars in model E4B with an Arches IMF, for nonrotating, rotating and with-bar cases. (The case shown with a stellar bar also had stellar mergers enabled, but a run without mergers was similar.)

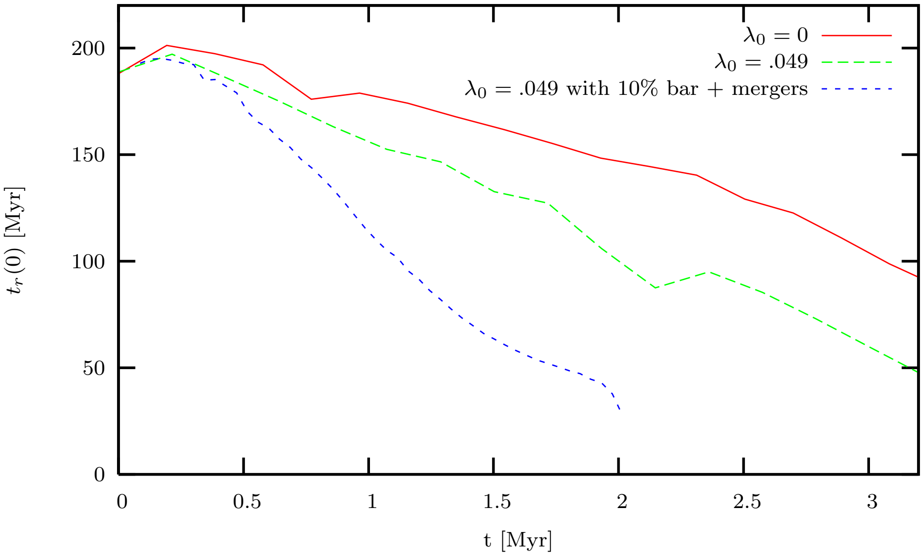

Figure 5.9: Median central relaxation time $t_r(0)$ v. time of the lowest-mass ($2M_\odot$) stars in model E4B with an Arches IMF, for nonrotating, rotating and with-bar cases. (The case shown with a stellar bar also had stellar mergers enabled, but a run without mergers was similar.)

Figure 5.8 shows the change with time of $\sigma_1(0)$ in the various cases studied. The systems which are initially rotating start at a lower value of $\sigma_1(0)$ because the method of incorporating rotation into the initial distribution function $f(I, J, 0)$ necessarily "cools" the system by increasing the amount of order present in the distribution of angular momentum J. The two runs which do not include a stellar bar have a slowly increasing $\sigma_1(0)$ as the system contracts and the central core gains entropy. Contrasting this is the case with a stellar bar, which collapses much more quickly and has a corresponding rapid increase in the central velocity dispersion. Of note, the expected initial value of $\sigma_o(0)$ for model E4B without rotation as given by Table 2.1 is 400 km/s; all models had a calculated initial $\sigma_1(0)$ which was similarly lower than would be expected if isotropy of the velocity dispersion (as given by (2.39) with $\delta \equiv 0$) strictly held. To make the calculated initial $\sigma_1(0)$ values agree with those of $\sigma_o(0)$ in Table 2.1, an anisotropy parameter $\delta \simeq -0.3$ is required; fortunately no derived quantities in the calculation depend strongly on an assumption of isotropy.

The median central relaxation time $t_r(0)$ for the same set of cases is shown in Fig. 5.9. The initial value of $t_r(0)$ is also slightly different than that listed in Table 2.1, which is expected as the full calculation of t_r applies to a given test star, not to a specific point in space (*cf.* the footnote in §3.1). Thus only a median value of t_r across all stars which traverse the system center can be determined, and any slight error in $\sigma_1(0)$ does propogate to the deduced $t_r(0)$, via its dependence on the cube of σ_1. This being the case, it can be seen that the presence of a stellar bar does cause a rapid decrease in the central relaxation time of the system, in this case by close to an order of magnitude within 2 Myr.

5.2.3 Model E2A, Arches-style IMF: Nuclear Cluster

Model E2A is somewhat different than E4B: the total mass, initial central density and velocity dispersion are all lower for E2A, as is the initial relaxation time (*cf.* Table 2.1). When using the Arches IMF the simulations which did not include a stellar bar were not numerically consistent for different choices of the timestep size, and so only the "full" case that included a stellar bar and collisional mergers is shown in Fig. 5.10.

Despite these differences, the increase in $\rho(0)$ by a factor of $\simeq 14$ before the simulation ended was similar to that of model E4B prior to its core collapse. Model E2A did not reach core collapse before aborting – although for E2A the simulation ended much earlier, at $t = 0.63$ Myr, consistent with the shorter initial value for t_r and smaller σ_1. And similarly to as was found for model E4B, there was sufficient merging of $125 M_\odot$ stars initially to create $250 M_\odot$ stars – in this case, at a rate of $3.9\,\mathrm{Myr}^{-1}$, which increased to $6.5\,\mathrm{Myr}^{-1}$ by the simulation's end. Approximately 15 stars of mass $125 M_\odot$ had been created via collisional mergers of lower-mass stars as well by the simulation's end, a small number attributable to the short amount of simulation time. The mass segregation behavior, with no segregation observed (*i.e.,* $S = 1$) due to dynamical effects and only a small amount ($S' \gtrsim 1$) from stellar mergers, was also similar to that of the E4B/Arches model, as was to be expected given the short simulated time of evolution of the system.

5.2.4 Model E2B, Arches-style IMF: Nuclear Cluster

Model E2B, when given an Arches-style IMF, shares many properties with model E4B/Arches described in §5.2.2. Figure 5.11 shows that the nonrotating case evolves very slowly, as does

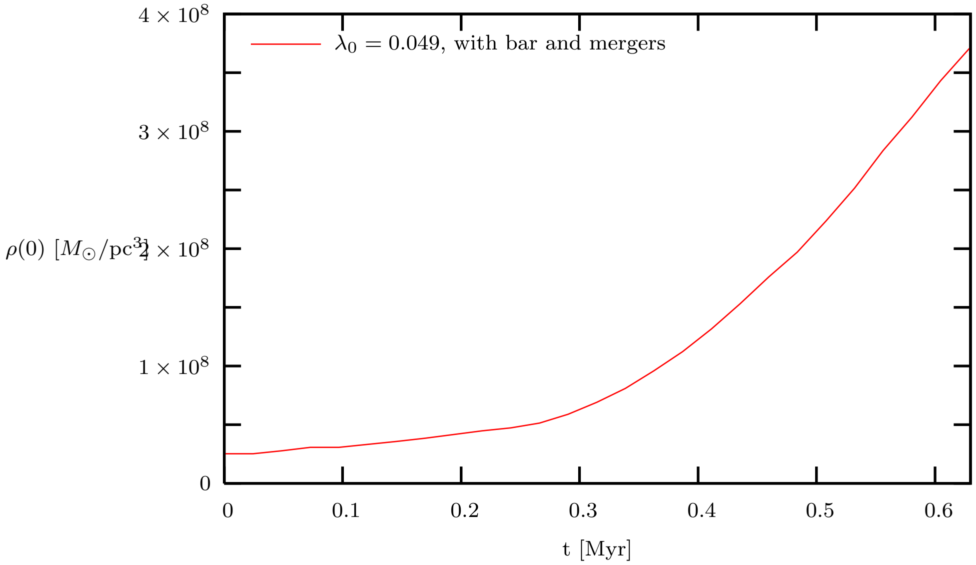

Figure 5.10: Central density v. time for model E2A with Arches-style IMF, $l_{\max} = 3$ and a stellar mass range of $2 - 125M_\odot$.

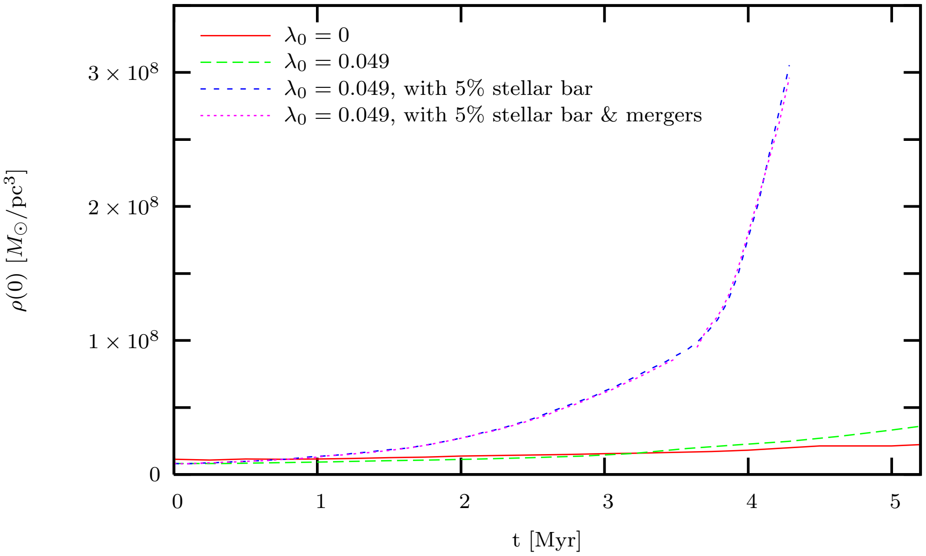

Figure 5.11: Central density v. time for model E2B with $l_{\max} = 2$ and using the Arches IMF with a stellar mass range of 2-125$M_\odot$. Note that the main-sequence lifetime of the most massive stars is 3 Myr and so the region of the graph beyond $t = 3$ is nonphysical.

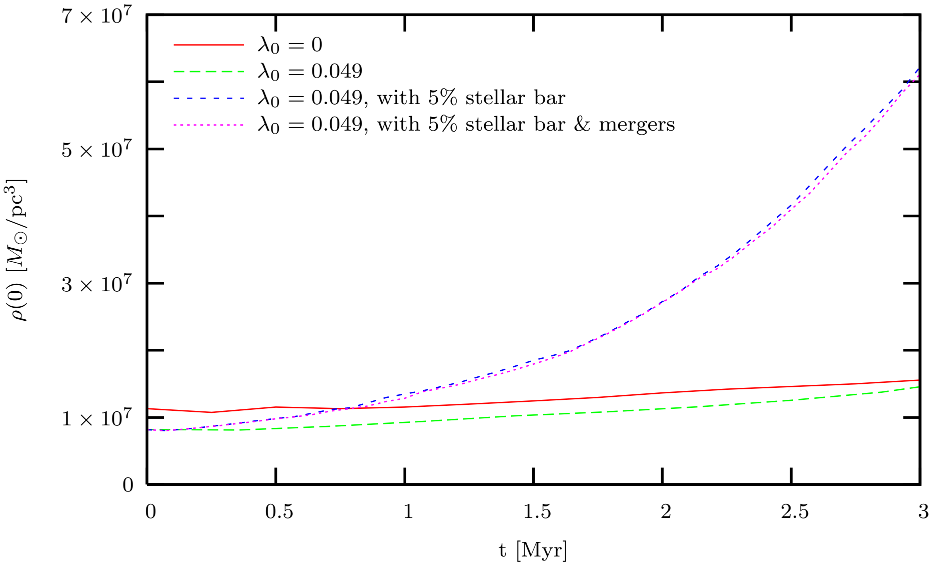

Figure 5.12: Detail of the first 3 Myr of Fig. 5.11.

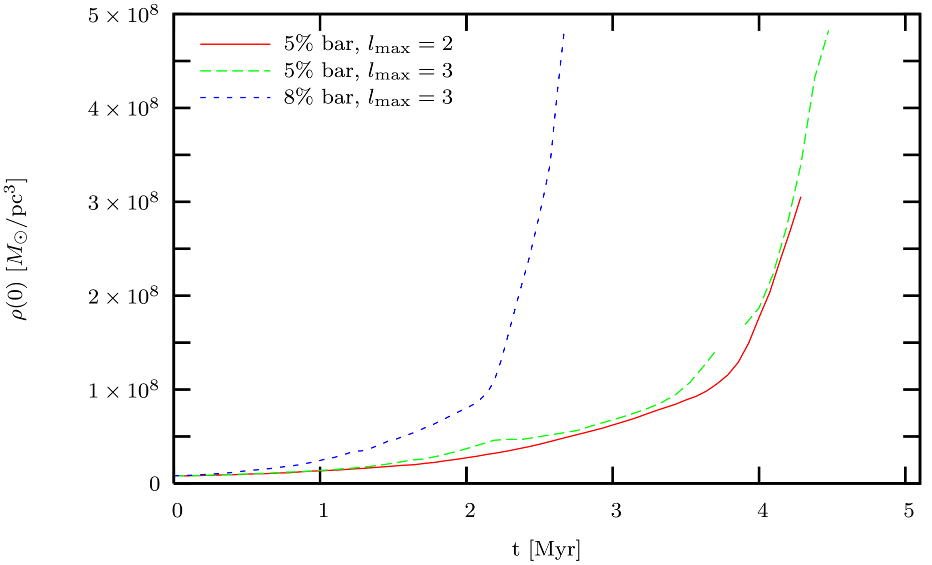

Figure 5.13: Similar to Fig. 5.11 but including $l_{\max} = 3$ and an 8% bar case. All have $\lambda = 0.049$ and do not include stellar mergers. (Simulations performed with mergers enabled yielded almost-identical runs of $\rho(0)$ to those shown in this figure.) Note that the "5% bar" plot here is the same as that shown in Fig. 5.11. The gap in one plot is due to a glitch in the output routine, as described in Chapter 4.

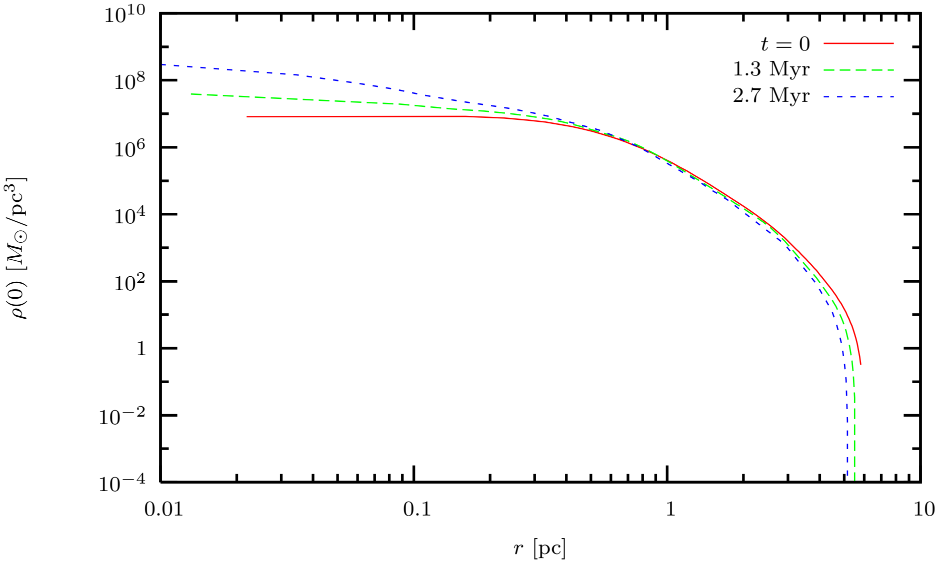

Figure 5.14: Density $\rho(r)$ v. radial distance r from the cluster center for the start, midpoint and end of the simulation for model E2B with an Arches-style IMF. Stellar mergers were enabled, as was an 8% stellar bar; rotation parameter $\lambda = 0.049$, and $l_{\max} = 3$. This corresponds to the leftmost plot of Fig. 5.13. The decrease in $\rho(r)$ with time for large r is only seen when a stellar bar is present.

the rotating case which does not have a stellar bar; in contrast, a modest 5.3% stellar bar induces core collapse and a $\simeq 60\times$ increase in central density $\rho(0)$ – but only at $t \simeq 4$ Myr of simulation time, well beyond the 3 Myr main sequence lifetime of the largest stars. As seen in Fig. 5.12, the evolution prior to $t = 3$ Myr is much more modest. Stellar mergers have little affect on $\rho(0)$, and Table 5.3 shows that with a 5.3% bar, by $t = 3$ Myr the rate of massive star production increases modestly from $G_{250} = 2.6$ Myr^{-1} to 3.7 Myr^{-1}. (*i.e.,* a small number of massive stars are produced collisionally at the start, and a small number are produced at the end of the stellar lifetimes.) Table 5.4 shows that if the system was able to continue on to $t \simeq 4$ Myr, core collapse would result in a somewhat larger increase in the rate of massive-star production.

A rise over time in the rate of producting $250M_\odot$ stars via mergers can be attributed either to an increase in overall density ρ, or to the presence of a greater number of massive stars due to previous collisional mergers. A comparison to determine how much each of the above factors contributed was performed by running a simulation in which stellar mergers were shut off at first and then turned on at $t = 3$ Myr. The result was a value of $G_{250} \simeq 3.4$ Myr^{-1} at $t = 3$, compared to 2.6 Myr^{-1} at $t = 0$, and to 3.3 Myr^{-1} at $t = 3$ when mergers were allowed from the start. Hence, at least for model E2B with a 5.3% bar, there is little cumulative effect of collisional mergers on the later merger rate of larger stars.

Allowing for a somewhat stronger bar gives a significant effect: Figure 5.13 shows that with an 8% bar (still well within the typical range of $5 - 10\%$) core collapse is accelerated and shifted earlier to $t \simeq 2.5$ Myr; likewise the rate of $250M_\odot$ star production increases somewhat, now to 8.2 Myr^{-1} – still not as high as for model E4B/Arches but notably more than at $t = 0$ in this model. As well, ultimately ~ 60 stars of $125M_\odot$ were created by mergers

IMF	*factor*	$\rho(0)$ at $t=0$	$\rho(0)$ at t_1	t_1	$G_{250}(t=0)$	$G_{250}(t_1)$	S	S'
Kroupa	1	6.3×10^7	4.2×10^8	2.9	0.	0.	0.98	1.04
	5	”	4.4×10^8	2.9	”	0.	”	1.12
	”	”	4.8×10^8	3.0	”	0.	”	1.13
Arches	1	”	4.0×10^8	1.6	33.	58.	1.	1.01
	5	”	5.0×10^8	1.6	”	68.	”	1.07
	”	”	9.1×10^8	2.1	”	87.	”	1.10

Table 5.5: Effect of artificially increasing the collisional merger rate by a factor of 5 in order to account for the effect of tidal-capture binaries, as described in §5.2.6. Listed are results for model E4B with $l_{\max}=2$ and a 10% stellar bar. The $factor=1$ entry has slight differences from the values shown in Table 5.3 due to different timestep sizes being used for the simulations listed here.

of lower-mass stars by the time the simulation ended at $t=2.7$ Myr. And also similarly to model E4B/Arches, no mass segregation is seen before the simulations terminate – which again was also the case for a simulation performed with $\lambda=0$ (*i.e.,* without rotation).

Finally, the full range of stellar mass density $\rho(r)$ plotted at various times in Fig. 5.14 again shows the core of the cluster contracting and becoming denser, while the outer regions lose density over time. As was also seen for model E4B, a simulation performed without a stellar bar did not display any similar effects of angular momentum transport to the very outer regions of the cluster as did one which included a bar. (The non-bar case is not plotted here, but compares similarly to Fig. 5.14 as do Figures 5.13 and 5.6 for E4B.)

5.2.5 Model E2A, Kroupa: Bulgeless Spiral or Dwarf Elliptical Nucleus

The simulations of model E2A with the Kroupa IMF behaved somewhat strangely: for very similar cases – *e.g.,* two runs both with rotation and a stellar bar, but only one with stellar mergers (which had a small effect when using the Kroupa IMF) – the dynamic r^2 grid would evolve rather differently for the various cases, making direct comparisons difficult. Also, finding a satifactory value for the potential-expansion parameter $l_{\max}$ proved elusive: $l_{\max}=5$ appeared to be required to capture the bulk of the interaction strength, but was even more unstable than $l_{\max}\leq 4$. Still, as shown in Table 5.2 this model gave results consistent with model E4B/Kroupa: the stellar bar produced an increase in $\rho(0)$ of an order of magnitude in a short time compared to the main-sequence lifetime of the most massive stars, but the rate of collisional mergers of those massive stars was negligible. The "reverse mass segregation" indicated by $S<1$ is unexpected. However, test runs using a lower value of $l_{\max}=2$, while not capturing the full extent of the model's evolution, also showed a similar trend towards $S<1$ through $t=1.6$ Myr – but at later times S turned around and began increasing again. Even so, why there would be a delay in mass segregation being exhibited for this model remains unexplained.

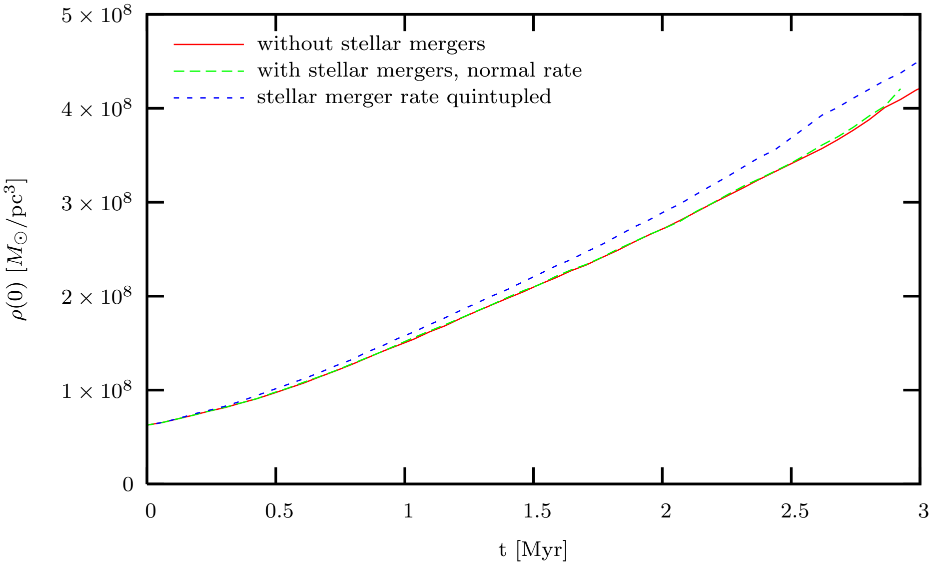

Figure 5.15: Effect of artificially increasing the collisional merger rate. Plotted is model E4B using a Kroupa IMF with $l_{\max} = 2$ and a 10% stellar bar. Also see Table 5.5.

5.2.6 Effect of Collisional Merger Rates

The development of the collisional merger rates in §3.8 assumed only direct 2-body collisions. However, indirect collisions – due to the creation of tight binaries formed via tidal capture which then merge on a short timescale – can enhance the merger rate by a factor of 3 to 5 in some situations [88]. (Other factors which may enhance the expected rate include that low mass $[5 - 10M_\odot]$ stars do not to contract to their long-term main sequence radius immediately and so have a larger cross section initially, and a second-order effect in that when those lower-mass stars collide with more massive stars they form a disk, which increases the new star's effective cross section.) To test how an enhanced merger rate would affect the overall results, model E4B was re-run with a the loss and gain coefficients of (3.44) and (3.50) increased by a factor of 5, the effect of which is shown in Fig. 5.15 and Table 5.5. (Tabulated but not plotted are the results for model E4B with the Arches IMF; while stable, the r^2 grid of that case's $factor = 5$ run sufficiently diverged from that of the base $factor = 1$ simulation to preclude a direct visual comparison.)

While in both the Kroupa- and Arches-IMF cases an enhanced merger rate resulted in a faster evolution of the overall system, little qualitative difference was observed. The central density $\rho(0)$ and merger-induced mass segregation ratio S' both increased more quickly with the additional factor of 5 in the merger rate – as did the rate of production of massive stars $G_{250}(t_1)$ when using the Arches IMF. But still no massive stars were created when using the Kroup IMF, and core collapse did not occur any more quickly with the Arches IMF. As model E4B has the highest initial $\rho(0)$ and is already the fastest-evolving of the various models studied, other models are expected to also show little qualitative change when given an enhanced merger rate.

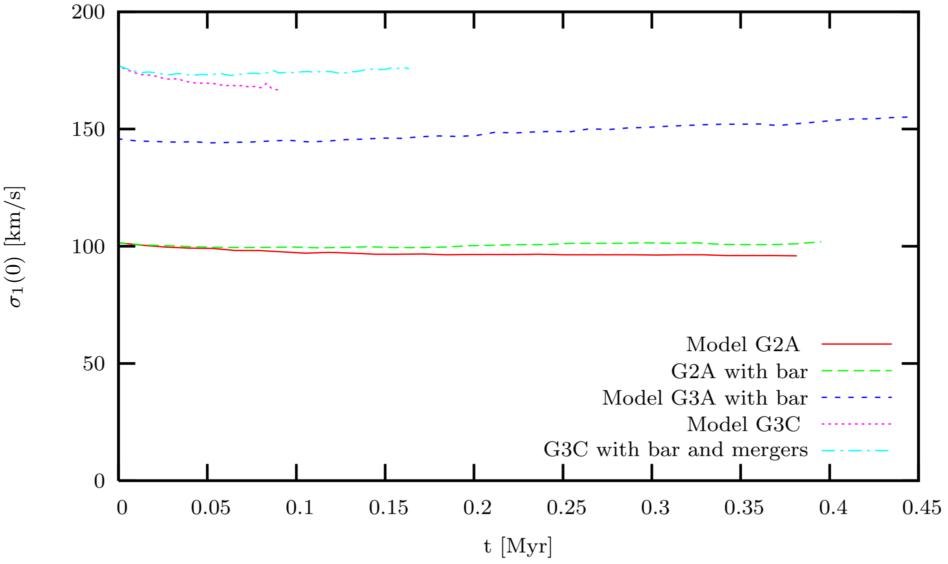

Figure 5.16: Central velocity dispersion v. time for the lowest-mass ($1M_\odot$) stars in the $\gamma = 0$ sphere models with a Kroupa IMF. All have $\lambda \simeq 0.05$. Plots of runs with and without stellar mergers were identical to each other; differences within each model are due to the presence or absence of a stellar bar.

5.3 $\gamma = 0$ Sphere Models

As described in §2.6.1, of the $\gamma = 0$ models which could be made using same values of total mass M and core radius r_{core} to the Plummer-sphere models described above, only the "G2A" model has reasonable values for the initial central density $\rho(0)$ and velocity dispersion $\sigma_1(0)$. An extreme model "G3C" in which the half-mass radius was the same as that for model E2B was also run and is described here, as is an intermediate model "G3A".

Figures 5.16 and 5.17 show $\sigma_1(0)$ and the central relaxation time $t_r(0)$ for various cases of models G2A, G3A and G3C, all with a Kroupa IMF. Unlike as was seen for the Plummer sphere models, the stellar bar had little effect on the $\gamma = 0$ models' evolution, as can be seen from the contrast of Fig. 5.16 with Fig. 5.9. This trend was common to all the $\gamma = 0$ sphere results and can at least partially be attributed to the short amount of simulation time before the models ended, as indicated by the small values on the plots' horizontal axes. Relaxation times were much shorter than for the Plummer sphere models, and so the bar had little time to transport angular momentum – and mergers had little time to produce larger stars as well. Even so, in some cases the short $t_r(0)$ values did allow for rapid overall evolution as is described in the following sections for each specific model.

5.3.1 Model G2A, Kroupa IMF: Galactic Spheroid / Spiral Bulge

Model G2A with a Kroupa IMF showed only an almost linear increase in $\rho(0)$ before ending due to numerical issues; Figure 5.18 shows that this was the situation for all cases regardless of whether either mergers or a bar were included. (The slight upturn in the "with mergers" line in the plot was not reproducable and is likely a numeric artifact.) Figure 5.19 gives the

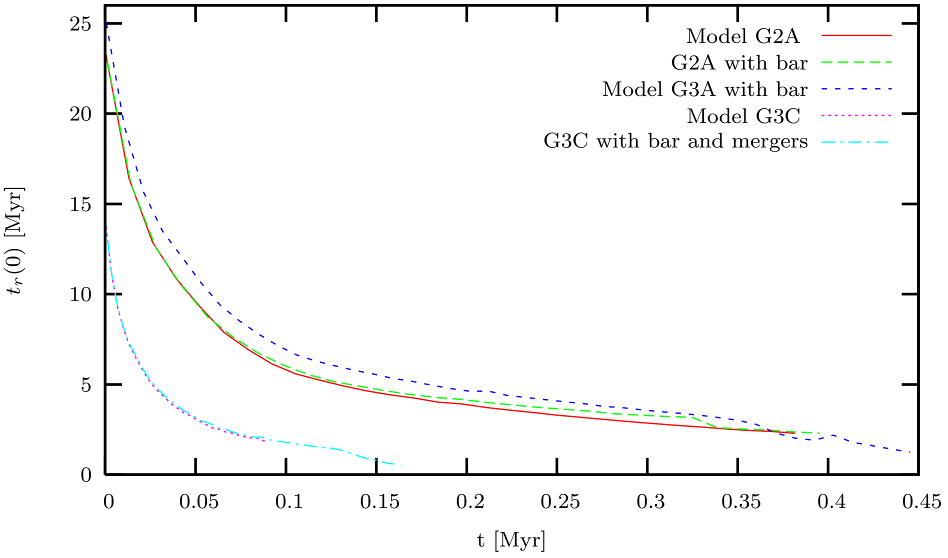

Figure 5.17: Central relaxation time for the lowest-mass ($1M_\odot$) stars in the $\gamma = 0$ sphere models with a Kroupa IMF. All have $\lambda \simeq 0.05$.

full density run $\rho(r)$ at various times in the simulation; interestingly the stellar density rises with time at both small and large r, and drops slightly with time in between; this compares to the Plummer sphere behavior in which the stellar density dropped with time for the largest values of r, *e.g.,* as in Fig. 5.6. The difference could be due to the $\gamma = 0$ sphere having relatively more mass to absorb angular momentum from the bar at intermediate values of r, to the slower pattern speed of the bar ($\Omega_b \simeq 160$ Myr^{-1}, compared to $\Omega_b \simeq 400$ Myr^{-1} for model E2A), or simply to the short evolution time of the simulation not allowing for angular momemtum to be transported all the way to the cluster's edge.

As was the case for all the Plummer sphere models, the calculated rate of $250M_\odot$ star production remained nill by the end of the simulation, and no stars of mass $125M_\odot$ had been created via collisional mergers. Table 5.2 shows that model G2A exhibited reverse mass segregation ($S < 1$); a test run with $l_{\max}$ lowered from 3 to 2 showed that S initially decreased below 1 and then increased again at times later than were reached by the "full" $l_{\max} = 4$ simulation, indicating that the apparent reverse mass segregation is a temporary phenomenon. Again, this behavior was also as was found for model E2A.

5.3.2 Models G3A (Larger Galactic Spheroid) and G3C, Kroupa IMF

Model G3A started with a somewhat higher density than model G2A did, and had a corresponding faster and greater amount of evolution of the system; model G3C continued this trend. In Figures 5.20 and 5.22 the runs without a stellar bar were either unstable or ended early, but the case with a bar case continued long enough to reach the beginnings of core collapse; these were the only Kroupa-IMF models studied to do so. Even so, the density of high-mass stars remained low enough, and the total simulated time was short enough, that no $125M_\odot$ stars were created by the simulations' end. (The calculated value was ~ 0.6 new stars of mass $125M_\odot$ for both models, and the Fokker-Planck code ignores

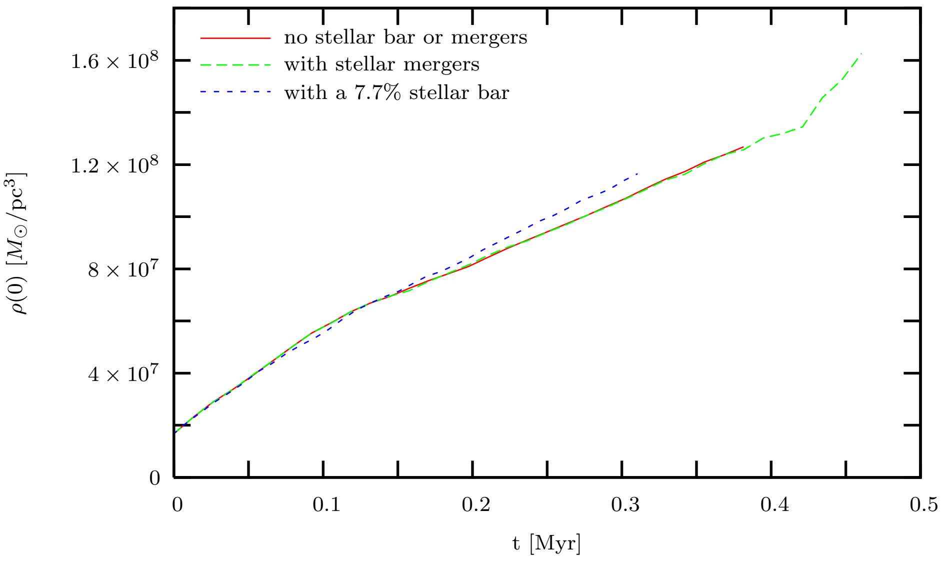

Figure 5.18: Central density v. time for $\gamma = 0$ sphere model G2A with a Kroupa IMF, $l_{\max} = 4$ and $\lambda = 0.051$. The nonrotating ($\lambda = 0$) case gave similar results but was less stable and not plotted.

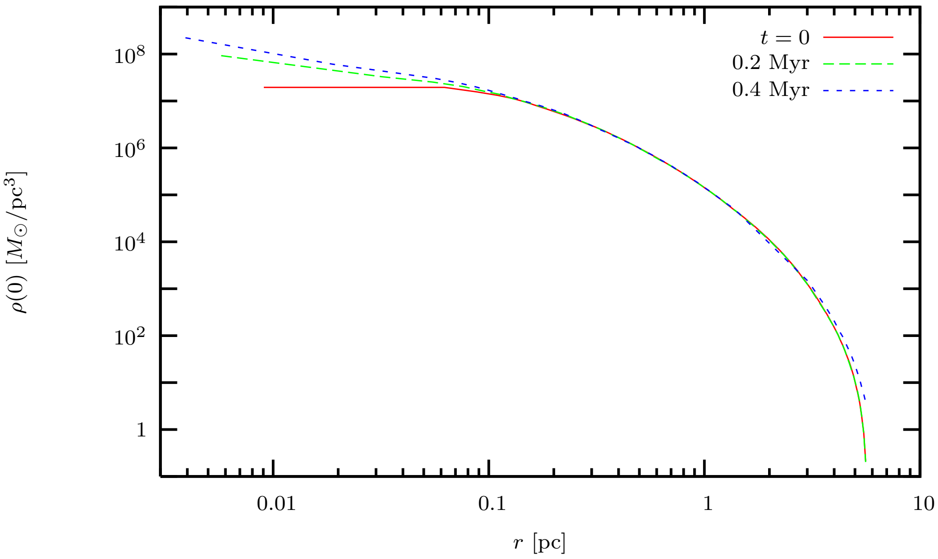

Figure 5.19: Density $\rho(r)$ v. radial distance r from the cluster center for the start, midpoint and end of the simulation for model G2A with a Kroupa IMF. Neither rotation nor stellar mergers were enabled, and $l_{\max} = 4$; this corresponds to the solid-line plot of Fig. 5.18. The behavior of $\rho(0)$ v. r for the other two cases from Fig. 5.18 is almost identical to that plotted here and are so not shown.

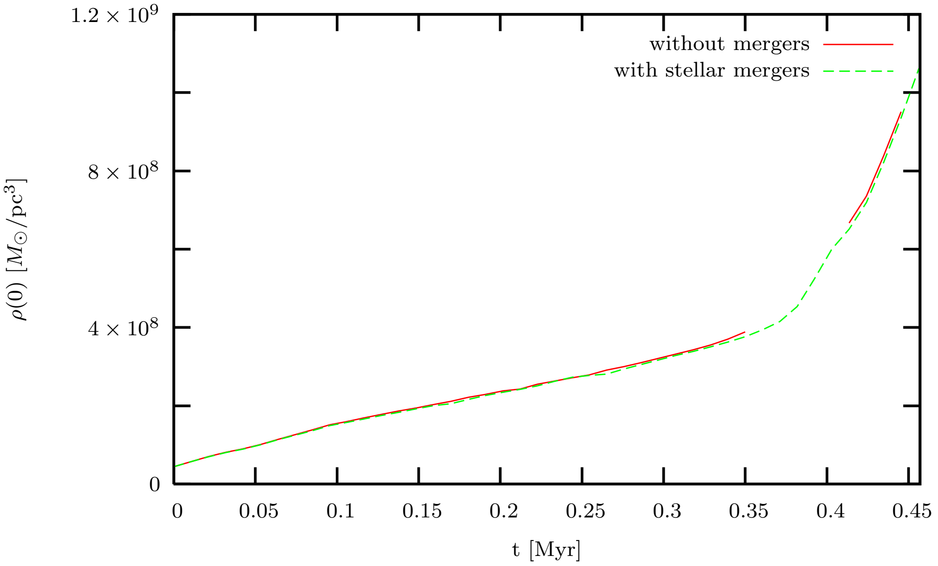

Figure 5.20: Central density v. time for $\gamma = 0$ sphere model G3A with a Kroupa IMF, $l_{\max} = 3$, $\lambda = 0.046$ and a 7.7% stellar bar. The gap in the no-merger plot is due to a glitch in the output routine. Runs without a stellar bar did not yield stable simulations and are not plotted.

stellar population numbers that are less than 1.) Similarly the calculated rate of $250 M_\odot$ star production remained nill, as given in Table 5.2. A modest amount of mass segregation did take place even in the short simulation time.

Despite the shorter simulation time, the run of $\rho(r)$ given in Fig. 5.23 displays an increase in density for model G3C at small and large r values (and a decrease for intermediate r) more clearly than does Fig. 5.19 for model G2A or Fig. 5.21 for G3A. This is consistent with G3C's much-shorter relaxation time, *e.g.,* as shown in Fig. 5.17. However, while model G3A may be seen as representing a larger version of the galactic spheroid modeled by G2A, it is doubtful that model G3C corresponds to a realistic astronomical system, and is more useful here as a demonstration case.

5.3.3 Model G2A, Arches IMF

Model G2A with an Arches-style IMF behaved much like model G2A/Kroupa but with a more rapid evolution. Figure 5.24 shows the central density $\rho(0)$ increasing roughly linearly with time but only a factor of $\simeq 8$ before the simulation ended due to numerical difficulties. The presence or absence of either a stellar bar or collisional mergers again had very little effect on the overall evolution of the system. Table 5.3 shows that the system starts with just a high enough rate of collisions of massive stars $G_{250} = 1.1 \mathrm{Myr}^{-1}$ to produce a small number of $250 M_\odot$ stars, with the rate increasing modestly bebfore the simulation ends. And in the short simulation time exactly one star of mass $125 M_\odot$ was created through collisional mergers of smaller stars.

This chapter has presented the simulation results for each model studied in turn. In the next chapter these results will be discussed with respect to what astrophysical effects were

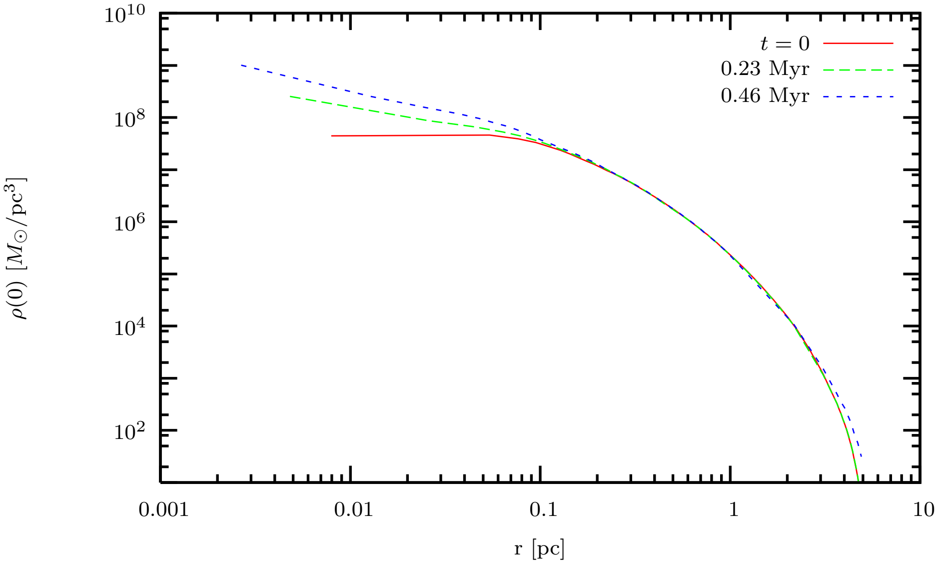

Figure 5.21: Density $\rho(r)$ v. radial distance r from the cluster center for the start, midpoint and end of the simulation for model G3A with a Kroupa IMF, $\lambda = 0.046$, a 7.7% stellar bar, and stellar mergers enabled. This corresponds to the dashed plot of Fig. 5.20.

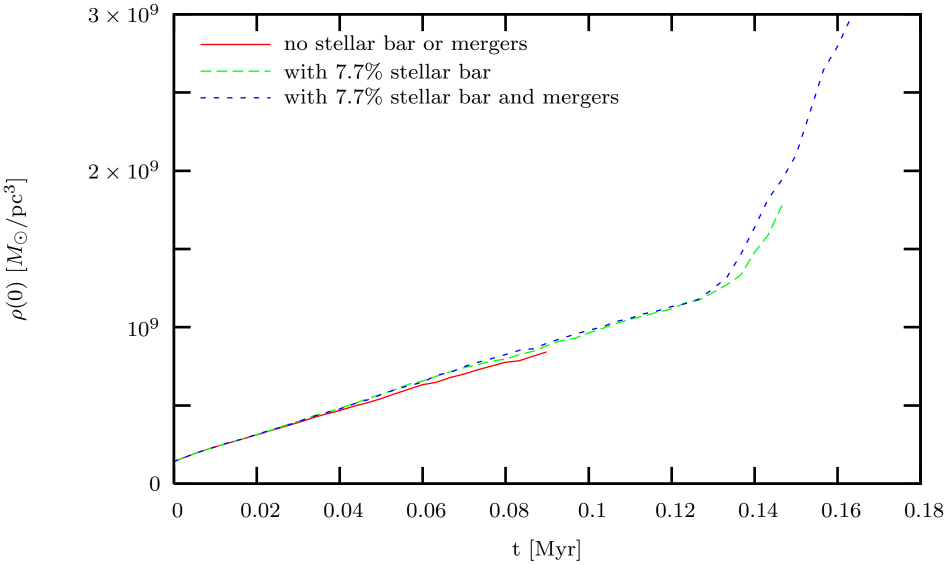

Figure 5.22: Central density v. time for $\gamma = 0$ sphere model G3C with a Kroupa IMF, $l_{\text{max}} = 2$, and $\lambda = 0.046$. The nonrotating ($\lambda = 0$) case did not yield a stable simulation and is not plotted.

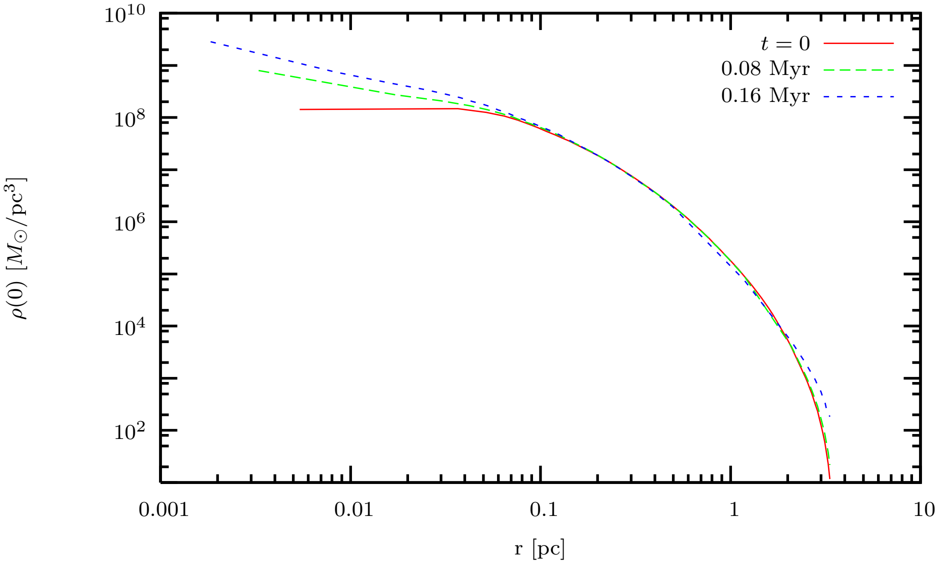

Figure 5.23: Density $\rho(r)$ v. radial distance r from the cluster center for the start, midpoint and end of the simulation for model G3C with a Kroupa IMF, $\lambda = 0.046$, a 7.7% stellar bar, and stellar mergers enabled. This corresponds to the short-dashed plot of Fig. 5.22.

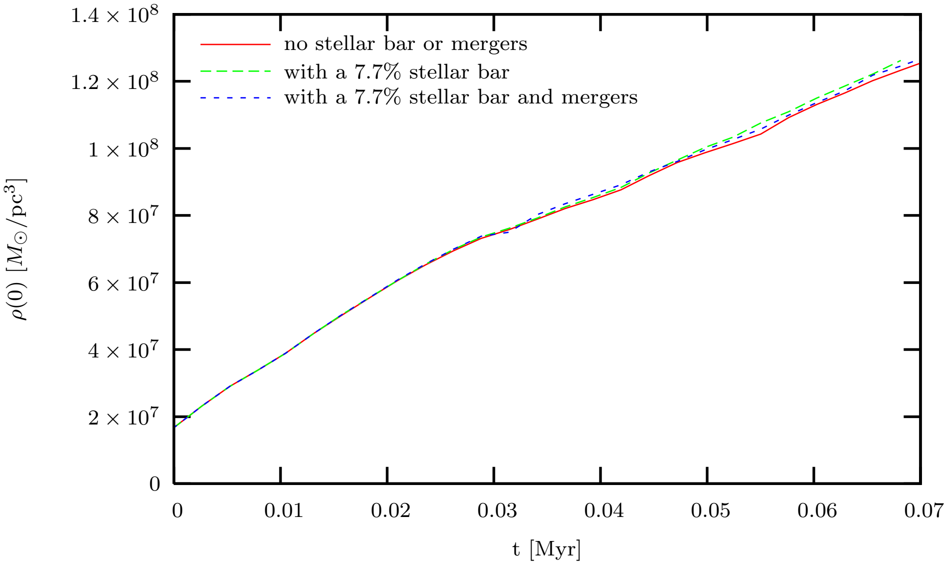

Figure 5.24: Central density v. time for $\gamma = 0$ sphere model G2A with an Arches-style IMF, $l_{\max} = 3$ and $\lambda = 0.051$. Two of the runs continued beyond $t = 0.07$ Myr but became unstable; those portions are not shown.

observed for which models, how that compares to what would be expected from timescale and density arguments, and what conclusions can be reached regarding the production of massive central objects in actual stellar systems.

Chapter 6

Discussion

According to Böker, nuclear clusters – massive, dense stellar clusters in the nuclei of galaxies – are observed to be nearly ubiquitous and share a similar relation to the host galaxy as do active galactic nuclei [11]. He closes with the following quote:

> It has recently been proposed by [Ferrearse *et al.* 2006] that nuclear clusters extend the well-known scaling relation between the mass of a galaxy and that of its central super-massive black hole (SMBH) to lower masses. This has triggered speculation about a common formation mechanism of nuclear clusters and SMBHs, being governed mostly by the mass of the host galaxy. The idea put forward is that nuclear clusters and SMBHs are two incarnations of a central massive object which forms in every galaxy. In galaxies above a certain mass threshold ($\simeq 10^{10} M_{\odot}$), galaxies form predominantly SMBHs while lower mass galaxies form nuclear clusters.
>
> . . . Is a nuclear cluster possibly a pre-requisite for the formation of a SMBH? Is the formation of a BH (not necessarily a super-massive one) a logical consequence of the high stellar densities present in nuclear clusters? Progress along these lines will require a better understanding of the formation of pure disk galaxies in the early universe, as well as improved models for the evolution of extremely dense stellar systems.

This study is an attempt at just such an improved model of dense stellar systems, linking what the above reference calls nuclear clusters with the eventual supermassive black holes in the centers of active galaxies. This chapter begins with two brief discussions, first of the timescale on which mass segregation alone would be expected to result in core collapse, and then of what central density is required before either runaway accretion or runaway collisional mergers might occur. The main body of the chapter then discusses the physical effects exhibited by the various simulations and what they imply for astronomical systems. Finally a brief closing considers what future observations may support the assumptions and results of these simulations, and what further improvements of the simulation method are indicated.

6.1 Timescale Arguments and Expectations

6.1.1 Mass Segregation

For mass segregation to occur, energy equipartition across different mass populations must be an unstable situation. Spitzer in 1969 formulated an analytic condition for preventing the mass-segregation instability: for a system of total mass $M_1 + M_2$, with M_1 contained in stars of individual mass m_1 and M_2 in stars of individual mass $m_2 > m_1$ as

$$\left(\frac{M_2}{M_1}\right)\left(\frac{m_2}{m_1}\right)^{3/2} < 0.16 \tag{6.1}$$

which for realistic mass spectra is never satisfied, *i.e.,* there are enough heavy stars that they decouple from the dynamics of the system as a whole and experience mass segregation [54]. For a system with multiple masses, the analogous condition has been numerically determined to be [41]

$$\left(\frac{N_h}{N_l}\right)\left(\frac{m_h}{m_l}\right)^{2.4} < 0.16 \tag{6.2}$$

in which N_l is now the total number of stars of individual mass m_l and N_h is the total number mass $m_h > m_l$. For the Kroupa IMF used here, the value of the above relation is 13; for the Arches IMF, 209. Even if the Kroupa IMF is extended down to $0.2 M_\odot$, the value is still 1.4. So indeed for all IMF studied here there are sufficient heavy stars so that mass segregation should at least be possible.

Given the plausibility of mass segregation, one can ask if it would be expected to foster core collapse. The timescale for segregation-driven core collapse has been variously estimated to be $\sim 10\%$ of the initial half-mass relaxation time [41] or $\sim 15\%$ of the initial mean-mass central relaxation time [54], both independent of the IMF used. For the models studied only the extreme model G3C has a resulting expected segregation-collapse timescale which falls under the 3 Myr stellar-lifetime limit, being in the $1 - 2$ Myr range for the two conditions given above. (Figure 5.17 shows model G3C having a very short relaxation time in general.) The runs of model G3C all ended before 0.5 Myr of simulated time and so would not be expected to exhibit segregation-driven collapse. In addition, other models were allowed to run beyond the 3 Myr limit as a test but none reached its expected segregation-collapse time – the closest being model E4B with an Arches IMF and no stellar bar, which had an expected collapse time of $t_c \simeq 15$ Myr and which ended at $t \simeq 10$ Myr showing no signs of collapse.

6.1.2 Critical Density

Whether or not core collapse is achieved, the central question is whether, given the $\sim 150 M_\odot$ upper mass limit of the IMF, a more-massive object can be formed which will evolve to be a seed black hole. Once a critical local density of $\rho \gtrsim \rho_{\rm crit} = 5 \times 10^9 (\frac{\sigma_1}{100 \rm km/s}) M_\odot/{\rm pc}^3$ is achieved, a $\sim 100 M_\odot$ star can directly accrete $10^3 M_\odot$ of material in ~ 5 Myr, creating an IMBH [55]. As all models used here have a central velocity dispersion $\sigma_1 \gtrsim 100$ km/s, Tables 5.2 and 5.3 show that none of the models meet that requirement. Model G3A/Kroupa may eventually reach it once deeper into core collapse than the simulation was able to track: Fig.

5.16 shows its $\sigma_1 \simeq 150$ km/s and remaining roughly steadly even as core collapse starts at $t \simeq 0.4$ Myr in Fig. 5.20. Model E4B/Arches reaches an even higher central density during its core collapse (Fig. 5.4), but it has a central velocity dispersion approaching 500 km/s and still increasing (Fig. 5.8). However neither model is expected to represent a typical astronomical system. As described in §2.6, G3A has a somewhat high density for the galactic spheroids that the $\gamma = 0$ spheres best model. Likewise the Arches IMF used in model E4B/Arches is determined from stellar clusters, while E4B's velocity dispersion places it in the range of galaxy cores, not individual clusters.

Other than accretion, collisional mergers are another avenue for forming massive objects from smaller stars. For massive ($125M_\odot$) stars, a local density of $\rho_{125} \simeq 1.2 \times 10^8 M_\odot/\mathrm{pc}^3$ is required in order for the average star to experience 1 collision per main sequence lifetime [88]. None of the models started with such a high central density, and only model E4B/Arches with the maximally strong bar achieved it by the end of the simulation. While by no means is it necessary for a collision rate of 1 per $125M_\odot$ star be reached in order for a massive object to form – such a rate would produce an abundance of $250M_\odot$ objects, while previous simulations of young clusters with an upper mass limit of $120M_\odot$ have found that more than one very massive star is never formed [56] – it does give an indication of the density range required. As will be seen in the following discussion of detailed simulation results, for a broad range of initial models using the Arches IMF (observed in stellar clusters), $250M_\odot$ objects are calculated to have been formed, while with the Kroupa IMF (representing the field star population) no such massive objects are created.

6.2 Summary of Simulation Results

The models studied fall into two main classes. The majority are Plummer-sphere initial models, which to be consistent with Quinlan and Shapiro [7] are labeled with an initial "E". Models E2A and E2B can represent either dense globular clusters or the nuclei of dwarf elliptical or bulgeless spiral galaxies, and E4B the core of a giant elliptical galaxy.

The other class of model is the "$\gamma = 0$ sphere", labeled with a "G", which fit the surface brightness of galactic spheroids. Model G2A has identical mass and core radius as model E2A, while models G3A and G3C have the same mass as model E2B but with smaller core radii. G3C is an extreme case which does not likely represent a realistic astronomical system.

For each model two possible initial mass functions were employed: a Kroupa IMF based on general field star populations, and an Arches-style IMF similar to that observed in dense clusters specifically. Note that the Arches IMF may not be a realistic IMF for either model E4B or the $\gamma = 0$ spheres, as those models best represent galactic centers while the Arches IMF is derived from observations of individual stellar clusters. However, the simulations of those models serve to complete the overall picture of the physical evolution of stellar systems at galactic centers. More complete details of the individual models and IMFs, and how rotation is introduced into them, is given in §2.6. The remainder of this section describes the physical effects exhibited by the various simulations, first for the Plummer models and then for the $\gamma = 0$ spheres.

6.2.1 Plummer-Sphere Simulations

The simulations which started with a Plummer Sphere distribution (the "E" models) give a clear picture. As each model's results show, the presence of a moderate-strength stellar bar can dramatically raise the rate of central density increase of a dense stellar cluster, by a factor of 1-2 orders of magnitude. However, the IMF also plays a crucial role: with the standard Kroupa IMF the rate of increase of central density $\rho(0)$ remains roughly linear with time through the main-sequence lifetime of the largest stars in a cluster, even for the most-dense model studied (E4B) – as shown in Fig. 5.1 – whereas with the cluster-specific Arches-style IMF core collapse can take place near to the 3 Myr stellar lifetime limit (*e.g.,* model E2B with a 5% bar, shown in Fig. 5.11), or even within it (model E4B with an 8 or 10% bar, Fig. 5.4).

More detailed discussion of the various physical aspects is given below.

Rotating v. Non-rotating Clusters

As described in §6.1.1, none of the Plummer models was expected to experience mass-segregation-driven core collapse, and none was observed in the simulations. For all of models E4B/Kroupa (Fig. 5.1), E4B/Arches (Fig. 5.4) and E2B/Arches (Fig. 5.11) both the nonrotating and rotating cases showed only slow growth in the central density by simulation's end, if no stellar bar was incorporated. Interestingly, rotation was not a strong impediment to this essentially secular increase in $\rho(0)$, *e.g.,* the E4B/Kroupa cases' density plots track each other with the only difference being the lower initial $\rho(0)$ value of the rotating cluster, while in the E4B/Arches and E2B/Arches models the rotating case actually has a somewhat larger rate of central density increase. While counterintuitive, rotation serving to enhance the rate of collapse has been observed before in Fokker-Planck simulations [39], possibly due to individual stars being able to interact with a larger (or at least different) range of neighboring stellar orbits.

Effects of a Stellar Bar Perturbation

When a moderate stellar bar – 5% to 10% strength, in line with observed bars – was included, however, all models showed rapid increase in central density. Core collapse was achieved in both E4B/Arches and E2B/Arches, and in less than the 3 Myr stellar-lifetime limit when the bar strength was $\simeq$ 8% or greater. Figure 5.13 shows the dramatic effect a greater bar strength can have on reducing the time to core collapse, going from $t \simeq 4$ Myr with a 5% bar to slightly over 2 Myr with an 8% bar. Comparing Figures 5.10 and 5.11 it appears that bar-driven core collapse would also be expected in model E2A/Arches had its simulation been able to remain numerically stable for a longer amount of simulated time.

Collisional Stellar Mergers: Effect on Cluster Dynamics

In general stellar mergers had little effect on the overall dynamics in the Plummer-sphere simulations. Figures 5.1 for E4B/Kroupa and 5.11 for E2B/Arches show only a very slight extra increase in $\rho(0)$ when stellar mergers are enabled, as is also born out in comparing values of S and S' in Tables 5.2 and 5.3. Even in the most extreme Plummer-sphere model studied, E4B/Arches with a 10% stellar bar, the main effect was that core collapse was achieved slightly earlier with stellar mergers, as shown in Fig. 5.4. That figure does hint that as core collapse progresses, collisional mergers may begin to drive the evolution of the central regions. However, this computational model cannot track into that regime, and it remains that the core collapse that is seen here is consistently bar-driven.

Collisional Stellar Mergers: Massive Star Formation

When the buildup of massive ($250 M_\odot$) stars via collisional stellar mergers is considered, the accelerated central density increase brought about by the bar perturbation is insufficient to produce even a single massive star within 3 Myr using the Kroupa IMF, as seen in Table 5.2. But the Arches IMF is already sufficiently populated with high-mass (up to $125 M_\odot$) stars to allow for the collisional production of massive $250 M_\odot$ stars, and the accelerated collapse induced by the bar only increases the rate of production, as seen in the $G_{250}(t_1)$ column in Table 5.3. So as far as the formation of massive stars is concerned it is the IMF which determines whether or not $250 M_\odot$ or larger objects are created at all. The overall rate of formation due to collisional mergers is primarily set by the initial conditions of the cluster, *i.e.,* its size and density, as can be seen by comparing the $G_{250}(0)$ values for models E2A, E2B and E4B in Table 5.3. However, the table also shows that the presence and strength of any stellar bar can have a large effect on how much the rate of massive star formation G_{250} increases in the period before core collapse occurs.

Complementing the above discussion of the calculated rate of massive object ($250 M_\odot$) formation, one can look at the numbers for the largest stars in the IMFs used. The physical upper limit was $150 M_\odot$, which was implemented numerically as a mass bin centered on and associated with stars of mass $125 M_\odot$. In the Kroupa IMF models, no simulation yielded an increase in the number of $125 M_\odot$ stars due to collisional mergers, which is consistent with a null rate for $250 M_\odot$ object formation. As detailed in Chapter 5, the Arches-IMF models ranged from creating 15 stars of mass $125 M_\odot$ stars in model E2A with a 5.3% bar, to 500 stars of mass $125 M_\odot$ stars in model E4B with a 10% bar. Noting the short amount of simulation time for which model E2A ran, these numbers are commensurate with the calculated rates of $250 M_\odot$ object formation.

Collisional Stellar Mergers: Possibility of Runaway Merging

The question of how massive-star production due to collisional mergers might be self-reinforcing is more difficult to answer. It would be expected that as more low-mass stars merge and so produce more intermediate-mass products, then more intermediate-mass merger candidates would be available to again collide and produce even more-massive stars. As discussed in §5.2.4 model E2B/Arches was allowed to run until time $t = 3$ Myr without stellar collisions enabled, and the rate of massive-star production due to mergers was at that point calculated to be within 3% of the value found when mergers had been allowed *ab initio* (and both were $\gtrsim 25\%$ larger than the rate at $t = 0$). Thus the amplification of the rate of massive star production due to the collisional merging of smaller stars is not significant when compared to the increase in the merger rate which results from the overall central density increase afforded by the stellar bar, at least in the pre-core-collapse period.

6.2.2 “$\gamma = 0$ Sphere” Simulations

The $\gamma = 0$ sphere models have similar density profiles to their Plummer sphere counterparts, although they are somewhat less centrally concentrated. A major distinction is that the $\gamma = 0$ spheres have a much shorter initial relaxation time, as seen by comparing Figures 5.17 and 5.9. This results in a very different physical evolution, as described next.

Rotating v. Non-rotating Clusters, Effect of Bar Perturbations

Contrasting the situation for the Plummer-sphere clusters, simulations of initial $\gamma = 0$ spheres showed no significant differences in evolution between nonrotating systems, rotating systems without a stellar bar, and rotating systems with a bar. Figures 5.18 for model G2A/Kroupa and 5.24 for G2A/Arches show this with respect to the central density $\rho(0)$. In model G2A/Kroupa the with-bar case exhibits a somewhat more rapid increase in $\rho(0)$, but it is extremely slight compared to the effect the bar has in the Plummer sphere simulations. In Fig. 5.24 model G2A/Arches shows even less difference, although it ran for a very short simulated time.

Even though the simulations of $\gamma = 0$ sphere models were not able to track the system as far into their evolution as some of the Plummer-sphere simulations were able to do, Figures 5.18 (for G2A/Kroupa) and 5.22 (for G3C/Kroupa) show a trend for the $\gamma = 0$ sphere models with no stellar bar to evolve almost identically to the same models with a bar, in stark contrast to how the Plummer sphere models behaved. Thus for the $\gamma = 0$ sphere models, overall system relaxation dominates over any bar- or collision-driven collapse, and again rotation does not obviously support the system against an increase in the central density.

Collisional Stellar Mergers

Stellar mergers played a similar role in $\gamma = 0$ sphere simulations as they did for the Plummer-sphere models. Figures 5.18 and 5.20 show that the central density $\rho(0)$, in cases in which stellar mergers are enabled, exactly tracks $\rho(0)$ in the corresponding no-mergers simulations for models G2A and G3A with a Kroupa IMF. Mergers have only a slight effect on $\rho(0)$ for model G2A/Arches (Fig. 5.24). Only in the extreme model G3C of Fig. 5.22 do stellar mergers affect the evolution, and then only by producing a slightly more rapid start to core collapse.

In terms of numbers of high-mass stars formed through collisions, the $\gamma = 0$ situation is slightly different, in that for both models G3A/Kroupa and G3C/Kroupa the calculated number of $125 M_\odot$ stars formed via collisions of lower-mass stars was between 0.5 and 1. Given the short simulation time of each model, a few additional $125 M_\odot$ would be expected to be created within the 3 Myr limit. While not interesting astrophysically, this contrasts with the situation for Plummer spheres in which no Kroupa IMF simulation produced any $125 M_\odot$ stars. However, Table 5.2 still shows that no $250 M_\odot$ objects at all would be created through mergers even in these cases – and Table 5.3 shows that only a small number are formed even in model G2A with a more top-heavy Arches IMF.

What remains unclear is whether model G3A/Kroupa will reach a regime in which massive star formation becomes plausible: as seen in Fig. 5.20 it does achieve the beginnings of core collapse, unlike any Plummer model with a Kroupa IMF. It does so in a short amount of simulation time, leaving ~ 2.5 Myr for post-core collapse evolution to occur before the stellar population begins to evolve off the main sequence. Model G3C/Kroupa achieves core collapse even more rapidly, but it is doubtful that it represents a realistic astronomical system.

6.2.3 Summary

Before performing the simulations it was expected that rotational support against collapse would be countered by the outward transport of angular momentum afforded by a stellar bar, allowing core collapse to proceed even in a rotating system. While the effect of the bar was observed in all Plummer sphere simulations it did not result in the anticipated core collapse in all models, nor in an ensuing runaway of collisional mergers creating a massive object. For the $\gamma = 0$ spheres the overall system relaxation dominated over the effect of the bar.

In the end the initial mass function turned out to be the determining factor: with the possible exception of model G3A no simulation performed using a Kroupa IMF, representative of a general galactic stellar population, reached core collapse, and none obtained a large enough stellar merger rate to create a massive object. In contrast, all models which employed the Arches-style IMF, which is specific to dense stellar clusters, started with a merger rate already sufficient to produce massive $250 M_\odot$ objects – although the presence of a stellar bar did make the difference in whether or not core collapse was reached, as well as in producing a large increase in the massive-object formation rate.

The above description is fairly robust:

- with a Kroupa IMF only the most extreme models (G3A and G3C) reached core collapse;
- with an Arches IMF all models reached core collapse (or for E2A, showed indications of it);
- no $\gamma = 0$ sphere model was dominated by the effect of a bar;
- even with an Arches IMF, a bar was required in order for any Plummer model to reach core collapse; and
- runaway mergers were not observed in any simulation, although a large increase in high-mass merger rates followed from core collapse in barred models.

Of note is that a cluster's evolution did not change qualitatively even when the merger rate was artificially increased by a factor of 5, as described in §5.2.6 – so it does not appear that a collisional-merger runaway is likely to occur for any pre-core-collapse system of main sequence stars.

Thus the results indicate two possible paths for formation of a massive object within a dense stellar system, one clear and one somewhat tentative. For cluster-sized systems similar to models E2A or E2B – including dwarf elliptical galaxy nuclei or bulgeless spiral galaxies –

with an Arches-style IMF, the expected degree of rotation is sufficient to support a stellar bar perturbation which in turn leads to core collapse through the outward transport of angular momentum. However, even without core collapse, through stellar mergers the system will produce massive objects in less than the main-sequence stellar lifetime sufficient to seed supermassive black hole growth. For larger systems, *e.g.,* galactic spheroids similar to model G3A, core collapse can be obtained even with a Kroupa IMF, although massive object formation through stellar mergers is not expected to be achieved within main-sequence lifetimes and so will require interactions of collapsed relativistic objects, which was outside the scope of this study.

6.3 Future Observations, Future Work

A basic assumption in this study is that stars preceded quasar black holes in the early universe. Two future space missions currently being considered by the European Space Agency in its *Cosmic Vision 2015-2025* program will go a long way towards addressing this assumption by observing the earliest luminous objects whose surrounding clouds of dense gas and dust obscure them from current instruments. First, the Far Infrared Interferometer (FIRI) will perform high-resolution imaging spectroscopy to resolve the creation of the first luminous objects in the universe, separating out the formation of stars and the growth of black holes [89]. Complementing FIRI, the ESA's XEUS X-ray observatory will detect the earliest quasars as they form. Together these two missions, or similar ones, could answer the question of whether stars or quasar black holes formed first [90]. They or similar observations may also address the main question remaining about the current study's results: whether or not an Arches-style IMF, with its weighting towards higher stellar masses, is in fact a better representation of the actual IMF that existed in early-universe stellar clusters than the Kroupa IMF is. If so, then the main result of these simulations – that individual nuclear stellar clusters are likely to form seed objects for massive black holes – holds. If not, then one must resort to post-core-collapse evolution to obtain the seed objects, and likely in galactic spheroids instead of in nuclear clusters.

As with all simulations, the results presented here have some limitations that could be addressed in future work.

- **Finite stellar lifetimes.** A conceptually straightforward but technically nontrivial extension would be to track the system's dynamical behavior as the more-massive stars move off the main sequence, which would involve the appropriate moving of stars from their present distribution functions to new ones representing white dwarfs, neutron stars and stellar-mass black holes – as represented by terms B_q and R_q in the general Fokker-Planck equation of 3.5, but not yet implemented in the simulations. Corresponding different collisional cross sections would also be required, as has been done by Quinlan and Shapiro in modeling one-dimensional, nonrotating Plummer sphere systems ([7], [1]).
- **Rotation-induced flattening.** As described in §2.5, the provision for a non-spherical gravitational potential is already implemented in the simulation code, but it is not yet debugged and shown to be numerically stable; once that is achieved it may be possible to add another degree of realism to the simulations of rotating systems. For

globular clusters at least, flattening of the system has been found to correlate with rotation [25], and to increase merger rates somewhat [26], which could strengthen the results presented here.

- **Dark Matter.** No provision has been made for there to be a background component in the gravitational potential, *e.g.,* due to the presense of collionless dark matter. This is consistent with all previous Fokker-Planck studies of stellar systems – in recent works Fiestas *et al.* ([25]), Kim *et al.* ([39], [24]) and Takahashi ([45], [46]) all ignore dark matter in their globular cluster simulations, while Arabadjis [26] and Quinlan & Shapiro ([1], [7]) look to the collisionally-produced massive objects in their simulations as *being* dark matter. Amaro-Seoane [41] points out that the current state of astrophysical modeling “comes to its limits” in embedding simulated systems in collisionless dark halos. Still, it is conceivable that introducing an *ad hoc* offset to the overall gravitational potential, or an additional collisionless component to the system’s matter-density profile, could serve as a toy model of the effect of a smooth dark matter background.

The early evolution of quasars and host galaxies is not simple – we have only observed the final $\simeq 30\%$ of quasar evolution, but even that is enough to indicate galactic spheroids and massive black holes do not grow together [90]. The results presented here show that even when rotation is accounted for, dense clusters of main-sequence stars in early galaxy cores can provide a possible mechanism for forming the seed objects that eventually become quasar black holes.

Appendix A

Non-Spherical Gravitational Potentials

A.1 Potential Calculations in Homeoidal Coordinates

For the more general, but rarely needed case in which there is an overall and variable ellipticity $e(a)$ to the potential, a somewhat more complex procedure than that given in §2.5 is required. One still starts with the inverted Poisson equation as given in (2.40) but now the derivation of the inverted Laplacian $\mathcal{L}^{-1}$ is more involved, as developed below.

To describe the non-spherical potential Φ when the isodensity surfaces of ρ are ellipsoids, homeoidal coordinates are called for [42]. Consider the potential $d\Phi$ of a thin ellipsoid of mass dM, semimajor axis a and ellipticity e at homeoidal radius u_o. Defining $\Delta \equiv ae$, the homeoidal radius u is related to the standard cylindrical coordinates by

$$\frac{R^2}{\cosh^2 u} + \frac{z^2}{\sinh^2 u} = \Delta^2. \tag{A.1}$$

(There is also an angle-like coordinate v which does not come into play.) Then the potential $d\Phi(u)$ at u due to the ellipsoid at u_o is

$$d\Phi(u) = -\frac{G\,dM}{\Delta} \left\{ \begin{array}{ll} \sin^{-1} e, & u < u_o \\ \sin^{-1}(\operatorname{sech} u), & u \geq u_o \end{array} \right. \tag{A.2}$$

where the mass element $dM = \rho dV$ for the ellipsoidal shell[1] is

$$\rho dV = 4\pi \rho a^2 (1 - e^2)^{1/2}\, da \tag{A.3}$$

and the total potential is found by integrating over mass shells at all values of u_o as will now be shown. Let us define $d\Phi_e \equiv -\frac{G\,dM}{\Delta} \sin^{-1} e$ and $d\Phi_u \equiv -\frac{G\,dM}{\Delta} \sin^{-1}(\operatorname{sech} u)$. Having

---

[1] Here it is implictly assumed that the ellipticity e is a slowly-enough varying function of radial coordinate a so as to not affect the differential dV; this approximation is examined in the Appendix.

no explicit dependence on u, the integration of $d\Phi_e$ is straightforward:

$$\Phi_e(u) = -4\pi G \int_{a(u)}^{\infty} (1-e^2)^{1/2} \rho \, \frac{\sin^{-1} e}{ae} \, da. \tag{A.4}$$

For the contribution to $\Phi(u)$ due to shells with $u_o < u$ we need a relation for $u(a)$. Noting that in the $z = 0$ plane, $R = a$ by definition, we accomplish this by solving (A.1) for u; the resulting quadratic equation in $\operatorname{sech}^2 u$ yields

$$\operatorname{sech}^2 u = \frac{1}{2R^2} \left[(R^2 + z^2 + \Delta^2) \mp \left((R^2 + z^2 + \Delta^2)^2 - 4R^2\Delta^2 \right)^{-1/2} \right] \tag{A.5}$$

where the necessity for $\operatorname{sech}^2 u \leq 1$ implies that only the upper sign is allowed (*e.g.,* taking $R = z$ makes this clear). Then the contribution to the total potential due to mass shells interior to u is obtained by integrating (A.2) using (A.3):

$$\Phi_u(u) = -4\pi G \int_0^{a(u)} (1-e^2)^{1/2} \rho \, \frac{\sin^{-1}(\operatorname{sech} u)}{ae} \, da. \tag{A.6}$$

Equation (A.5) may be expanded, yielding approximations necessary for use when $R \to 0$. (When $\Delta \to 0$ the simpler spherically-symmetric procedure can be used.) The total gravitational potential at an arbitrary value point $u(R, z)$ is thus given by

$$\Phi^{\text{new}}(u) = \Phi_u + \Phi_e. \tag{A.7}$$

Note that with use of (A.5) in (A.6), the only reference to the homeoidal coordinate u in the calculation of $\Phi(u)$ is in locating the transition from Φ_u to Φ_e, despite u_o being the (implicit) variable of integration. Note also that the coordinate transformation of (A.1) is referenced to the integration variable u_o, and not to the location of measurement u; thus each homeoidal transform is defined by the mass shell responsible for the element of $d\Phi_u$ currently contributing to $\Phi_u = \int d\Phi_u$. Hence $a = a(u_o)$ and $e = e(u_o)$; neither a nor e has an explicit dependence on u. Likewise, the Δ in (A.5) is $\Delta = a(u_o)e(u_o)$ even though the coordinate being transformed is u and not u_o, and so (A.1) constitutes a "running coordinate transformation" which is continuously evolving across the range of integration of (A.6).

A.2 Laplacian Solution in Ellipsoidal Coordinates: Numerical Aspects

A novel aspect of this study is that e is allowed to vary over the range of a. This is different from the fixed-ellipiticy methods (*e.g.,* as described by Tremaine and Weinberg [30]) but was found to be both advantageous in that it allows potentially more accurate modeling, and necessary to avoid a pitfall of the fixed-ellipiticy schemes: physically interesting interactions will primarily occur in high-density regions, *i.e.,* the core of the distribution, but these tend not to be the high-ellipticity regions which are usually far from the center of the cluster – and so taking a single value of e for the entire cluster can artifically increase the effects of non-sphericity. This was found to be the case in practice, and was what motivated the allowance for a variable e. (Provision for fixing e at a predetermined value is also built in however.)

One may worry that a changing $e(a)$ may induce troublesome shell-crossing. Numerically, this is only a concern if the a^2 grid is so fine and $e(a)$ changes so rapidly that the semiminor axes $b = a(1-e)^{1/2}$ of adjacent shells on the grid cross (*i.e.*, if $a_{r+1}(1-e_{r+1})^{1/2} < a_r(1-e_r)^{1/2}$ for gridpoints a_r^2 and $a_{r+1}^2 > a_r^2$). In practice the variability of $e(a)$ is not large enough for this to occur, given the courseness of the a^2 grids. And in principle, possible shell-crossing is not a fundamental concern anyway, as the ellipsoidal isodensity surfaces are merely smoothed representations of the underlying discrete stellar distribution.

To see how much effect the variability of $e(a)$ has on the entire ρ—Φ scheme, consider what effect it has on the mass element dM, which was not taken into account in (A.3). This is not a strict calculation of what effect variable e has on the calculus, but it does give an idea of the size of its effect:

$$dM = \rho\, dV = \rho\, d\left[\frac{4\pi}{3}a^3(1-e^2)^{1/2}\right] = 4\pi\rho a^2(1-e^2)^{1/2}\left[1 - \frac{ae}{3(1-e^2)^{1/2}}\frac{de}{da}\right] da \qquad \text{(A.8)}$$

which shows that the effect of the dependence of e on a is of $O(e^2)$ smaller than the dominant term. Using the more straightforward form of (A.6) greatly simplifies the numerical calculation without explicitly affecting the results, as the $e(a)$ run is still allowed to converge on a consistent set of values along with $\rho(a)$ and $\Phi(a)$ during the iterative solution for all of these at each new timestep, with resulting new f.

Thus the assumption of ellipsoidal isodensity surfaces, along with f_n, determines the potential Φ. The required integrals described above, however, are too computationally expensive to be performed each time the knowledge of potential is needed in the Fokker-Planck coefficient calculation. After testing interpolation schemes and analytic approximations, only the Clutton-Brock self-consistent field (“SCF”) method as described by Hernquist and Ostriker [91] proved adequate in both accuracy and speed. For the SCF “core radius” r_{core} the value at which $\Phi(r_{\text{core}}) = \Phi(0)/\sqrt{2}$ is used. This is also the core radius of the Plummer potential with central potential $\Phi(0)$. The field method is tested at each timestep, and if sufficient accuracy cannot be achieved with a reasonable number of expansion terms, the code falls back on direct integration. For interpolating over the other, one-dimensional grids of quantities discussed in Chapter 2, simple polynomial or cubic spline schemes are employed.

Appendix B

Tables of Symbols

	Quantity	**Symbol**	**Example**
§2.2	radial component	r or 1	v_r, I_1
	tangential component	T or 2	v_T, I_2
	azimuthal compoment	z or 3	J_z, w_3
	orbital endpoints	p and a	r_p, $\mathcal{R}_a$
	dummy component variable	i or $j = 1$, 2, or 3	Ω_j, D_{ij}
§2.2.2	"circular" value or component	c	r_c
§2.3.2	maximum dynamically-allowed value	max	$J_{\max}$, $\beta_{\max}$
§2.3.3	velocity-space coordinates	v	v, θ_v, ϕ_v
	a specific mass "bin"	q; later also q' & $\mathcal{Q}$	m_q, $f_{q'}$
§2.4.1	total value of a dynamical quantity	tot	J_{tot}
	minimum dynamically-allowed value	min	$J_{\min}$
§2.4.4	average over a given quantity x, possibly with a weighting function	$\langle\cdot\rangle_x$	see text for definitions
§2.6	core (radius)	core	r_{core}
§2.6.2	amount of a dynamical quantity due to rotation	rot	J_{rot}, T_{rot}
§3.2.1	spherical harmonic indices	l, m	Y_{lm}
	resonance indices	knm; later $\ell_1\ell_2\ell_3$	Ψ_{knm}
	an individual object (whether a single star or the stellar bar)	$*$	Φ_*
§3.3	the stellar bar	B	f_{B}, Ω_{B}
§3.4	repeated summation index	[any repeated index]	$\ell_p\Omega_p \equiv \sum_p \ell_p\Omega_p$

Table B.1: List of frequently-used subscripts.

	Quantity	**Symbol**
§2.1	Newton's constant of gravitation	G
§2.2	radial component of orbital action	I_1 or I; later also $\mathcal{I}$
	tangential component of orbital action	I_2 or J or J; later also $\mathcal{J}$
	vertical component of orbital action	$J_z = J\cos\beta$
	orbital energy per unit mass	E
	orbital position coordinates	r , θ, ϕ
	radial & tangential orbital frequencies	Ω_1, Ω_2
	gravitational potential per unit mass	Φ
	canonical orbital angle coordinates	w_1, w_2, w_3; later also $\mathcal{W}_1$ *etc.*
	orbital inclination (Euler angle)	β
	orbital azimuth (Euler angle)	ψ
	orbital radial endpoints	r_p, r_a; later also $\mathcal{R}_p$, $\mathcal{R}_a$
§2.2.1	radial component of orbital velocity	$v_r \equiv \sqrt{[2(E-\Phi) - J^2/r^2]}$
§2.2.2	"circular" orbital radius	r_c where $\partial v_r(r_c)/\partial r = 0$
§2.3.1	individual stellar mass	m
§2.3.3	tangential component of orbital velocity	$v_T = J/r$
§2.4	average rotational velocity	$\varpi(r)$
§2.4.1	orbital polar angle	$\alpha \equiv \frac{\pi}{2} - \beta$
§2.4.3	"ellipsoidal" radial coordinate	$a(r, e(r))$
	cylindrical radial and vertical coordinates	R, z
§2.6.1	one-dimensional (nonrotating) distribution function for stars of mass m_q	$\bar{F}_q(E)$
§2.6.2	total cluster gravitational potential energy	$W_{\rm grav}$
	total cluster rotational angular momentum	$J_{\rm rot}$
	cosmological rotation parameter	$\lambda = \frac{J_{\rm rot}\lvert W_{\rm grav}\rvert^{\frac{1}{2}}}{GM^{2.5}}$

Table B.2: Symbols first introduced in Chapter 2, listed by section in which each was first introduced. These denote properties of individual stars in the system, not bulk properties. Bulk properties based on summing over the individual stars' values are denoted by subscripts, for example total energy of the stellar system $E_{\rm tot}$ or the rotational contribution to the total angular momentum $J_{\rm rot}$. Only symbols that are used in more than one section are included.

	Quantity	**Symbol**
§2.2	gravitational potential per unit mass	Φ
§2.3.1	stellar mass density	ρ
	total cluster mass	M
	stellar distribution function in position space	$F(\mathbf{r}, \mathbf{v})$
	stellar distribution function in energy/angular momentum space	$F(E, J)$
2.3.2	orbit-averaged stellar distribution function	$f(I, J)$
2.4.1	stellar distribution function parameterized in terms of orbital inclination	$f_\beta(I, J; \beta) \equiv h(\alpha(\beta))f(I, J)$
	even (g) and odd (Θ) components of inclination parametization	$h(\alpha) \equiv g(\lvert\alpha\rvert) + \Theta(\alpha)$
	total angular momentum of all stars of energy E	$J_{\mathrm{tot}}(E)$
§2.4.3	ellipticity of stellar mass distribution	$e(r)$
§2.4.4	one-dimensional *"rms"* stellar velocity dispersion	σ_1
	anisotropy parameter	δ
	generalized stellar velocity dispersion	$\sigma_o \equiv \sqrt{(1-\delta)\sigma_1^2}$

Table B.3: Symbols first introduced in Chapter 2, listed by section in which each was first introduced. These are bulk properties of the system or of a subpopulation thereof. Only symbols that are used in more than one section are included.

	Quantity	Symbol
§3.1	Fokker-Planck drift coefficient	$C_j(I, J)$
	Fokker-Planck diffusion coefficient	$D_{ij}(I, J)$
	ad hoc merger loss and gain terms in F-P equation	$L_q(I, J)$, $G_q(I, J)$
§3.2.1	gravitational potential of individual perturber (whether a single star or the stellar bar)	Φ_*
	spherical harmonics	$Y_{lm}(\theta, \phi)$
	radial coefficients of perturbation potential expansion	$\Phi_{lm}(r)$
	coefficients of perturbation potential expansion in action space	$\Psi_{knm}(I, J)$
	Slater rotation coefficient	$V_{lnm}(\beta)$
	resonance strength coefficient	$W_{klnm}(I, J)$
	resonance indices	(k, n, m); later (ℓ_1, ℓ_2, ℓ_3)
§3.2.2	any quantity $y(\beta)$, weighted-averaged over orbital inclination	$\langle y\rangle_{h(\beta)} = \frac{\pi}{\beta_{\max}} \int d\beta\, h(\beta) y(\beta)$
	any quantity $y(\beta)x(\theta)$, weighted-averaged over polar angle $\theta(\psi, \beta)$	$\langle y; x\rangle_\theta = \frac{2}{\beta_{\max}} \int d\beta\, h(\beta) y(\beta)\, d\psi\, x(\theta)$
	average-squared perturbation coefficient strength	$\overline{\Psi^2}_{knm} \equiv \left\langle \Psi_{knm}(\beta)^2 \right\rangle_{h(\beta)}$
§3.4	perturbation ”turning on” parameter	$\eta \gtrsim 0$
	(complex) perturbation orbital frequency	$\omega = m\Omega_* + i\eta$
§5.2.1	Relative density increase of highest-mass stars	$S = [\rho_Q(0)/\rho(0)]_{t_1}/[\rho_Q(0)/\rho(0)]_{t_0}$
	As above, but including effect of stellar mergers	$S' = \ldots$
	Rate of massive ($250 M_\odot$) star creation	G_{250}

Table B.4: Symbols first introduced in Chapters 3, 4 and 5, listed by section in which each was first introduced. Only symbols that are used in more than one section are included.

Bibliography

[1] G. D. Quinlan and S. L. Shapiro, Astrophys. J. **343**, 725 (1989).

[2] P. Hut, S. Mineshige, D. C. Heggie, and J. Makino, Progress of Theoretical Physics **118**, 187 (2007).

[3] M. Umemura, in *Coevolution of Black Holes and Galaxies, Carnegie Observatories Centennial Symposium I, Pasadena 2002*, Vol. 1 of *Carnegie Observatories Astrophysics Series*, edited by L. C. Ho (Carnegie Observatories, Pasadena, 2004).

[4] G. Cusumano *et al.*, Nature **440**, 164 (2006).

[5] S. Naoz, S. Noter, and R. Barkana, Mon. Not. Roy. Astr. Soc. Lett. **373**, 98 (2006).

[6] S. M. Koushiappas, J. S. Bullock, and A. Dekel, Mon. Not. Roy. Astr. Soc. **354**, 292 (2004).

[7] G. D. Quinlan and S. L. Shapiro, Astrophys. J. **356**, 483 (1990).

[8] R. H. David, L. P. Durisen and H. Cohn, Astrophys. J. **316**, 505 (1987).

[9] S. F. Portegies Zwart and S. L. W. McMillan, Astrophys. J. **576**, 899 (2002).

[10] M. C. Begelman and M. J. Rees, Mon. Not. Roy. Astr. Soc. **185**, 847 (1978).

[11] T. Böker, in *41st ESLAB Symposium: The impact of HST on European Astronomy, Astrophysics and Space Science Proceedings*, edited by G. D. Marchi and F. D. Macchetto (Springer, Berlin, 2007).

[12] D. F. Figer, in *The Initial Mass Function 50 Years Later*, Vol. 327 of *Astrophysics and Space Science Library*, edited by E. Corbelli, F. Palla, and H. Zinnecker (Springer, Dordrecht; New York, 2005).

[13] S. P. Oh, K. M. Nollett, P. Madau, and G. J. Wasserburg, Astrophys. J. Lett. **562**, 1 (2001).

[14] T. Ebisuzaki, J. Makino, T. G. Tsuru, Y. Funato, S. Portegies Zwart, S. McMillan, S. Matsushita, H. Matsumoto, and R. Kawabe, Astrophys. J. Lett. **562**, 19 (2001).

[15] M. Coleman Miller and E. J. M. Colbert, Intl. J. Modern Phys. D **13**, 1 (2004).

[16] C. Tyler, B. Janus, and D. Santos-Noble, Bull. Am. Astr. Soc. **36.2**, 42.08 (2004).

[17] N. Deuchting, Phys. Rev. D **70**, 064015 (2004).

[18] F. A. Rasio, M. Freitag, and M. A. Gürkan, in *Coevolution of Black Holes and Galaxies*, edited by L. Ho (Cambridge University Press, Cambridge, 2004).

[19] M. Freitag, M. A. Gürkan, and F. A. Rasio, in *Massive Stars in Interacting Binaries*, Vol. 367 of *ASP Conference Series*, edited by T. Moffat and N. S. Louis (Astronomical Society of the Pacific, San Francisco, 2005), p. 707.

[20] J. C. L. Jr., J. S. Warren, F. A. Rasio, A. Sills, and A. R. Warren, Astrophys. J. **568**, 939 (2002).

[21] J. Barnes and G. Efstathiou, Astrophys. J. **319**, 575 (1987).

[22] E. Kim, C. Einsel, H. M. Lee, R. Spurzem, and M. G. Lee, Mon. Not. Roy. Astr. Soc. **334**, 310 (2002).

[23] A. Meza, Astron. & Astrophys. **395**, 25 (2002).

[24] E. Kim, H. M. Lee, and R. Spurzem, Mon. Not. Roy. Astr. Soc. **351**, 220 (2004).

[25] J. Fiestas, R. Spurzem, and E. Kim, Mon. Not. Roy. Astr. Soc. **373**, 677 (2007).

[26] J. S. Arabadjis, Ph.D. thesis, University of Michigan, 1997.

[27] H. Dejonghe, Phys. Rept. **133**, 218 (1986).

[28] I. Shlosman, in *The Central Kiloparsec of Starbusts and AGN: The La Palma Connection*, Vol. 249 of *ASP Conference Series*, edited by J. H. Knapen, J. E. Beckman, I. Shlosman, and T. J. Mahoney (Astronomical Society of the Pacific, San Francisco, 2001), p. 55.

[29] R. F. G. Wyse, Astrophys. J. Lett. **612**, 17 (2004).

[30] S. Tremaine and M. Weinberg, Mon. Not. Roy. Astr. Soc. **209**, 729 (1984).

[31] D. Lynden-Bell and A. J. Kalnajs, Mon. Not. Roy. Astr. Soc. **157**, 1 (1972).

[32] E. Athanassoula, in *Nonlinear Dynamics in Astronomy and Physics: In Memory of Henry Kandrup*, Vol. 1045 of *Annals New York Acad. Sci.*, edited by S. T. Gottesman, J.-R. Buchler, and M. E. Mahon (The New York Academy of Sciences, New York, 2005), p. 168.

[33] M. D. Weinberg and N. Katz, Astrophys. J. **580**, 627 (2002).

[34] M. Weinberg, Mon. Not. Roy. Astr. Soc. **213**, 451 (1985).

[35] I. Shlosman, in *Disks of Galaxies: Kinematics, Dynamics and Perturbations*, Vol. 275 of *ASP Conference Series*, edited by E. Athanassoula, A. Bosma, and R. Mujica (Astronomical Society of the Pacific, San Francisco, 2002), p. 231.

[36] W. Maciejewski, P. J. Teuben, L. S. Sparke, and J. M. Stone, Mon. Not. Roy. Astr. Soc. **329**, 502 (2002).

[37] W. Maciejewski, in *Disks of Galaxies: Kinematics, Dynamics and Perturbations*, Vol. 275 of *ASP Conference Series*, edited by E. Athanassoula, A. Bosma, and R. Mujica (Astronomical Society of the Pacific, San Francisco, 2002), p. 251.

[38] F. Wernli, E. Emsellem, and Y. Copin, Astron. & Astrophys. **396**, 73 (2002).

[39] E. Kim, I. Yoon, H. M. Lee, and R. Spurzem, Mon. Not. Roy. Astr. Soc. **383**, 2 (2008).

[40] E. A. Spiegel and J.-L. Thiffeault, in *Stellar Astrophysical Fluid Dynamics: Proceedings of the Chateau de Mons meeting in honour of Douglas Gough's 60th birthday* (Cambridge University Press, Cambridge, 2003), p. 377.

[41] P. Amaro-Seoane, Ph.D. thesis, University of Heidelberg, 2004.

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

[43] J. Goodman, Ph.D. thesis, Princeton University, 1983.

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

[45] K. Takahashi, Proc. Astr. Soc. Japan **47**, 561 (1995).

[46] K. Takahashi, Proc. Astr. Soc. Japan **49**, 547 (1997).

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

[48] H. Dejonghe, Astrophys. J. **320**, 477 (1987).

[49] J. F. Thompson, Z. U. A. Warsi, and C. W. Mastin, *Numerical Grid Generation: Foundations and Applications* (North-Holland, New York, 1985).

[50] P. Henrici, *Elements of Numerical Analysis* (Wiley, New York, 1964).

[51] L. Spitzer, *Dynamical Evolution of Globular Clusters* (Princeton University Press, Princeton, NJ, 1987).

[52] W. Dehnen, Mon. Not. Roy. Astr. Soc. **265**, 250 (1993).

[53] H. Cohn, Astrophys. J. **242**, 765 (1980).

[54] M. A. Gürkan, M. Freitag, and F. A. Rasio, Astrophys. J. **604**, 632 (2004).

[55] S. McMillan and S. Portegies Zwart, in *Massive Stars in Interacting Binaries*, Vol. 367 of *ASP Conference Series*, edited by T. Moffat and N. S. Louis (Astronomical Society of the Pacific, San Francisco, 2005), p. 697.

[56] M. Freitag, F. A. Rasio, and H. Baumgardt, Mon. Not. Roy. Astr. Soc. **368**, 121 (2006).

[57] M. Freitag, M. A. Gürkan, and F. A. Rasio, Mon. Not. Roy. Astr. Soc. **368**, 141 (2006).

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

[59] W. Jaffe, Mon. Not. Roy. Astr. Soc. **202**, 995 (1983).

[60] L. Hernquist, Astrophys. J. **356**, 359 (1990).

[61] B. Cervantes-Sodi, X. Hernandez, C. Park, and J. Kim, Mon. Not. Roy. Astr. Soc. **388**, 863 (2008).

[62] P. Kroupa, C. A. Tout, and G. Gilmore, Mon. Not. Roy. Astr. Soc. **262**, 545 (1993).

[63] D. F. Figer, in *The Formation and Evolution of Massive Young Star Clusters*, Vol. 322 of *ASP Conference Series*, edited by H. Lamers, L. Smith, and A. Nota (Astronomical Society of the Pacific, San Francisco, 2004), p. 49.

[64] M. S. Oey and C. J. Clarke, Astrophys. J. Lett. **620**, 43 (2005).

[65] B. G. Elmegreen, in *Massive Stars: From Pop III and GRBs to the Milky Way, Conference held at the Space Telescope Science Institute, May 8-11, 2006*, edited by M. Livio and E. Villaver (Cambridge Univ. Press, Cambridge, 2007).

[66] J. Van Vleck, *Quantum Principles and Line Spectra* (National Research Council, Washington, D.C., 1926).

[67] J. M. Fregeau, K. J. Joshi, S. F. Portegies Zwart, and F. A. Rasio, Astrophys. J. **570**, 171 (2002).

[68] S. McMillan, H. Baumgardt, S. Portegies Zwart, P. Hut, and J. Makino, in *Formation and Evolution of Massive Young Star Clusters*, Vol. 322 of *ASP Conference Series*, edited by H. J. G. L. M. Lamers, L. J. Smith, and A. Nota (Astronomical Society of the Pacific, San Francisco, 2004), p. 449.

[69] H. Baumgarft, G. D. Marchi, and P. Kroupa, Astrophys. J. **685**, 247 (2008).

[70] F. Combes and B. Elmegreen, Astron. & Astrophys. **271**, 391 (1993).

[71] V. G. Gurzadyan, in *IAU XXVth General Assembly, Sydney, July 2003*, Vol. 13 of *Highlights of Astronomy*, edited by O. Engvold (Astronomical Society of the Pacific, San Francisco, 2005).

[72] E. Athanassoula, in *Papers from the XXXV Annual Meeting of the Division of Dynamical Astronomy of the American Astronomical Society. Cannes (France), 19-23 April 2004*, Vol. 91 of *Celestial Mechanics & Dynamical Astronomy*, edited by A. Morbidelli and S. Ferraz-Mello (Springer, Dordrecht, 2005), p. 9.

[73] A. Toomre, in *Structure and Evolution of Normal Galaxies*, edited by S. M. Fall and D. Lynden-Bell (Cambridge University Press, Cambridge, 1981), p. 111.

[74] J. A. Sellwood, in *The Shapes of Galaxies and Their Halos, Proceedings of the Yale Cosmology Workshop*, edited by P. Natarajan (World Scientific, Singapore, 2002).

[75] D. A. Gadotti and R. E. de Souza, Astrophys. J. Lett. **583**, 75 (2003).

[76] I. Christodoulou, D.M. Shlosman and T. J.E., Astrophys. J. **443**, 551 (1995).

[77] F. Combes and R. Sanders, Astron. & Astrophys. **96**, 164 (1981).

[78] D. Ceverino and A. Klypin, Mon. Not. Roy. Astr. Soc. **379**, 1155 (2007).

[79] I. Berentzen, I. Shlosman, I. Martinez-Valpuesta, and C. H. Heller, Astrophys. J. **666**, 189 (2007).

[80] P. O. Lindblad, Astron. & Astrophys. Rev. **9**, 221 (1999).

[81] M. Weinberg and S. Tremaine, Astrophys. J. **264**, 364 (1983).

[82] *Standard Mathematical Tables*, 19th ed., edited by S. M. Selby (The Chemical Rubber Company, Cleveland, OH, 1971).

[83] G. Arfken, *Mathematical Methods for Physicists*, 3rd ed. (Academic Press, London, 1984).

[84] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge University Press, Cambridge, 1992).

[85] J. Chang and G. Cooper, J. Comp. Phys. **6**, 1 (1970).

[86] J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations* (Wadsworth & Brooks/Cole, Pacific Grove, CA, 1989).

[87] H. M. Lee, Astrophys. J. **319**, 801 (1987).

[88] M. Freitag, in *Massive Star Formation: Observations Confront Theory*, Vol. 387 of *ASP Conference Series*, edited by H. Beuther, H. Linz, and T. Henning (Astronomical Society of the Pacific, San Francisco, 2008), p. 49.

[89] F. Helmich and R. Ivison, Experimental Astronomy **23**, 245 (2009).

[90] M. J. Page, in *Proceedings of the 39th ESLAB Symposium*, Vol. 588 of *ESA-SP series*, edited by F. Favata and A. Gimenez (European Space Agency Publications, Noordwijk, The Netherlands, 2005), p. 105.

[91] L. Hernquist and J. Ostriker, Astrophys. J. **386**, 375 (1992).