Loss cone refilling by flyby encounters
A numerical study of massive black holes in galactic centres

Mimi Zhang
Wolfson College
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy
in the subject of Theoretical Astrophysics

Trinity Term 2008
Loss cone refilling by flyby encounters
A numerical study of massive black holes in galactic centres

Mimi Zhang
Wolfson College
Rudolf Peierls Centre for Theoretical Physics
University of Oxford

Abstract

A gap in phase-space, the loss cone (LC), is opened up by a supermassive black hole (MBH) as it disrupts or accretes stars in a galactic centre. If a star enters the LC then, depending on its properties, its interaction with the MBH will either generate a luminous electromagnetic flare or give rise to gravitational radiation, both of which are expected to have directly observable consequences. A thorough understanding of loss-cone refilling mechanisms is important for the prediction of astrophysical quantities, such as rates of tidal disrupting main-sequence stars, rates of capturing compact stellar remnants and timescales of merging binary MBHs. If a galaxy were isolated and perfectly spherical, the only refilling mechanism would be diffusion due to weak two-body encounters between stars. This would leave the LC always nearly empty. However, real galaxies are neither perfectly spherical nor isolated. In this thesis, we use N-body simulations to investigate how noise from accreted satellites and other substructures in a galaxy’s halo can affect the LC refilling rate.

Any N-body model suffers from Poisson noise which is similar to, but much stronger than, the two-body diffusion occurring in real galaxies. To lessen this spurious Poisson noise, we apply the idea of importance sampling to develop a new scheme for constructing N-body realizations of a galaxy model, in which interesting regions of phase-space are sampled by many low-mass particles. This scheme minimizes the mean-square formal errors of a given set of projections of the galaxy’s phase-space distribution function. Tests show that the method works very well in practice, reducing the diffusion coefficients by a factor of $\sim 100$ compared to the standard equal-mass models and reducing the spurious LC flux in isolated model galaxies to manageably low levels.

We use multimass N-body models of galaxies with centrally-embedded MBHs to study the effects of satellite flybys on LC refilling rates. The total mass accreted by the MBH over the course of one flyby can be described, using a simple empirical fitting formula that depends on the satellite’s mass and orbit. Published large-scale cosmological simulations yield predictions about the distribution of substructure in galaxy halos. We use results of these together with our empirical fitting formula to obtain an upper bound on substructure-driven LC refilling rates in real galaxies. We find that although the flux of stars into the initially emptied LC is enhanced, but the fuelling rate averaged over the entire subhalos is increased by only a factor 3 over the rate one expects from the Poisson noise due the discreteness of the stellar distribution.

A thesis submitted for the degree of Doctor of Philosophy
Trinity Term 2008
Contents

Title Page ................................................................. i 
Abstract ........................................................................ iii
Table of Contents ........................................................ iv
Acknowledgments .......................................................... viii

1 Introduction ............................................................... 1
  1.1 Supermassive black holes in galactic centres .................... 3
    1.1.1 Energy arguments for MBHs in AGNs ....................... 3
    1.1.2 Dynamical evidence for MBHs in inactive galaxies .......... 4
  1.2 Fuelling MBHs .......................................................... 7
    1.2.1 Loss of stars in the loss cone ................................. 8
    1.2.2 How to feed stars to a loss cone? ............................ 9
  1.3 Outline and declaration ................................................ 11

2 Dynamical modelling .................................................... 13
  2.1 Equilibrium model .................................................. 14
    2.1.1 DFs for collisionless stellar system .......................... 14
    2.1.2 Collisionless Boltzmann equation ............................ 15
    2.1.3 Preliminaries ..................................................... 17
  2.2 N-body modelling ................................................... 21
    2.2.1 Monte Carlo integration ....................................... 23
    2.2.2 The CBE and N-body simulations ............................ 25
  2.3 Departure from equilibrium ......................................... 32
    2.3.1 Collision terms .................................................. 32
    2.3.2 Fokker-Planck equation ....................................... 33
    2.3.3 Relaxation time .................................................. 34
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Multi-mass schemes for collisionless simulations</td>
</tr>
<tr>
<td>3.1</td>
<td>Collisionless ( N )-body simulations</td>
</tr>
<tr>
<td>3.1.1</td>
<td>( N )-body simulations and the CBE</td>
</tr>
<tr>
<td>3.1.2</td>
<td>How to fight against small ( N ) limitations?</td>
</tr>
<tr>
<td>3.2</td>
<td>Formulation</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Observables</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Optimal sampling scheme</td>
</tr>
<tr>
<td>3.2.3</td>
<td>ICs for ( N )-body model</td>
</tr>
<tr>
<td>3.3</td>
<td>Comparisons of a bulge model [no MBH]</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Formal errors</td>
</tr>
<tr>
<td>3.3.2</td>
<td>How well is the acceleration field reproduced?</td>
</tr>
<tr>
<td>3.3.3</td>
<td>How well are integrals of motion conserved?</td>
</tr>
<tr>
<td>3.4</td>
<td>Simulations of a bulge + MBH model</td>
</tr>
<tr>
<td>3.4.1</td>
<td>( N )-body realizations</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Results: diffusion in ( J^2 )</td>
</tr>
<tr>
<td>4</td>
<td>Loss cone refilling by flyby encounters</td>
</tr>
<tr>
<td>4.1</td>
<td>Loss cone refilling by two-body relaxation</td>
</tr>
<tr>
<td>4.2</td>
<td>Flyby encounters</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Tests: calibrating noise-driven LC refilling</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Experiments: measuring perturber-driven LC refilling</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Results: an empirical ( M_{\text{lc}} ) formula</td>
</tr>
<tr>
<td>4.2.4</td>
<td>Discussion</td>
</tr>
<tr>
<td>4.3</td>
<td>Where do LC stars come from?</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Angular-momentum transport by satellite torque</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Resonances from perturbation theory</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Test-particle experiments of a periodic perturber</td>
</tr>
<tr>
<td>4.4</td>
<td>Summary</td>
</tr>
<tr>
<td>5</td>
<td>Rates of MBH feeding in bulge-halo systems</td>
</tr>
<tr>
<td>5.1</td>
<td>Modelling a Milky-way sized halo and its substructures</td>
</tr>
<tr>
<td>5.1.1</td>
<td>The halo density profile</td>
</tr>
<tr>
<td>5.1.2</td>
<td>The subhalo mass function</td>
</tr>
<tr>
<td>5.1.3</td>
<td>The subhalo spatial distribution</td>
</tr>
<tr>
<td>5.1.4</td>
<td>The subhalo distribution function</td>
</tr>
<tr>
<td>5.2</td>
<td>LC flux driven by noisy DM clumps</td>
</tr>
<tr>
<td>5.2.1</td>
<td>LC flux from individual clumps</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Averaged LC refilling rate</td>
</tr>
<tr>
<td>5.3</td>
<td>Discussions</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Some modelling uncertainties</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Some theoretical investigations</td>
</tr>
<tr>
<td>6</td>
<td>Conclusions and future studies</td>
</tr>
<tr>
<td>6.1</td>
<td>Constructing collisionless galaxy models</td>
</tr>
<tr>
<td>6.2</td>
<td>Rates of loss cone refilling by MBHs</td>
</tr>
</tbody>
</table>
Acknowledgments

I wish to thank Dr. John Magorrian for providing me with a complicated jumbo holiday, for supporting me with constant entertainments, and for sharing with me the rewarding experience. My acknowledgments go to this fearless leader who draws me into galactic dynamics and gets me into \( N \)-body simulations. No matter where I go, I will live with the courage he implants in me. All along the way, I have had good fortune to learn \textit{Galactic Dynamics} with its author James Binney.

I owe a debt to Andy, who is instrumental in getting my programing started as a project. The “Beecroft institute” of Ralf, Rachel, Sarah, Andy, Michael, Ben, Callum have been great about sharing experience. And it is my duty to thank Dorothy Hodgkin Postgraduate PPARC-BP Awards for the financial support.

As seeming to languish for the last four years, I have challenged my willingness to see if I can lose my assumptions. Replace my empty mind with an open one, and return with my empty mind in the end. What a failure and what an achievement!

I dedicate this thesis to my mother and father who cherish me more than I do, and vice versa.

\( \text{M}^2\text{Z} \)

Oxford, June, 2008
In the high-redshift universe, massive black holes (MBHs) have been implicated as the powerhouses for quasars in the Active Galactic Nuclei (AGN) paradigm; in the nearby galaxies, the existence of dead quasar engines has also been supported by modern stellar-dynamical searches (e.g., Ferrarese & Ford 2005). Over three dozens of local detections have unveiled the demographics of MBH populations such as a tight correlation between MBH mass and galactic velocity dispersion (e.g., Gebhardt et al. 2000; Ferrarese & Merritt 2000; Merritt & Ferrarese 2001), and a less tight correlation between MBH mass and the mass/luminosity of the bulge (e.g. Kormendy & Richstone 1995; Magorrian et al. 1998; Marconi & Hunt 2003). While efforts to build a larger and statistically significant sample continue, it is time to understand the origin, evolution and cosmic relevance of these fascinating objects.

The subject is large and my space is limited. In this thesis, I concentrate on developing and using $N$-body models to learn more about the stellar dynamics in galactic centres, especially around the MBHs. I begin immediately below by reviewing the “journey” to the MBHs: from the AGN paradigm-mandated MBHs as power sources to the observation-
supported MBHs. Then, I describe two astrophysical ingredients: electromagnetic flares by tidal disruption of main-sequence stars and gravitational wave signals by capture of compact stars. Next, I explain the emptiness of “loss cone” and the starvation of the MBH which motivate the work in thesis. Finally, I outline briefly $N$-body modelling procedures developed in the subsequent chapters.
1.1 Supermassive black holes in galactic centres

A black hole (of mass $M_\bullet$) is a **compact** object whose gravity is so strong even light cannot escape. That is, the escape velocity at the surface is greater than the speed of light $\sqrt{2GM_\bullet/r} > c$. Therefore, a non-rotating object (of mass $M$) is a black hole if it is smaller than its Schwarzschild radius $r < r_{sch} = 2GM/c^2$ (Schwarzschild 1916). In astrophysics, small black holes of ordinary-mass $\sim 10 M_\odot$ are the evolutionary end points of some massive stars. This thesis deals with the super-massive $10^6 M_\odot \leq M_\bullet \leq 10^9 M_\odot$ ones that might power the quasars and their weak kin, active galactic nuclei (AGN).

1.1.1 Energy arguments for MBHs in AGNs

In the 1960s, energy released from gravitational source was invoked (Salpeter 1964; Zel’Dovich 1964) to explain the enormous luminosities of the newly discovered 3C273 in Virgo constellation (Schmidt 1963). This quasar, contraction of quasi-stellar radio source, has an average apparent magnitude of 12.8 and an absolute magnitude of $-26.7$. From a distance of $\sim 749$ Mpc ($z \simeq 0.158$), this quasar produces a luminosity $^1$ of $L \sim 10^{39}$ W, about $2 \times 10^{12}$ the luminosity of our sun ($L_\odot$) or about 10 times that of the brightest galaxies. For this class of objects, some are observed to have radio jets and lobes where the minimum stored energy is $E \sim 10^{53-57}$ W, the energy-equivalent mass is $M = E/c^2 \sim 10^{6-10}$ solar mass ($M_\odot$), and the mass-associated horizon scale is $r_{sch} \sim 10^{9-13}$ m.

Quasars are found to vary in luminosity on a variety of time scales, from few months, weeks to days; a few objects change their brightness in minutes. The causal light-speed limit places an upper bound on the size of the radiation region, e.g., $R \sim c\tau \sim 10^{11}$ m. Thus, the required quasar mass must be confined to a region not much larger than the MBH.

---

$^1$Quasar luminosities are measured from the power output in optical or UV light after bolometric correction, assuming that the redshift of quasars is due to expansion of the space and radiation is isotropic. Thus, they are likely to be affected by gravitational redshift and Doppler boosting somewhat in either direction.
event horizon, mandating gravity as the energy source.

The argument for gravity power was sharpened by Lynden-Bell (1969, 1978). He firstly showed that MBHs of mass \( > 10^7 \text{M}_\odot \) must normally lie at the centre, because dynamical friction drags them to the bottom of the potential well rapidly. Lynden-Bell further argued that the accretion discs around central MBHs can convert \( \sim 10\% \) of the rest mass of an object into energy, much more efficient than thermonuclear reactions with a typical energy production rate of \( \sim 0.7\% \) in sun-like stars.

By the early 1980s, accreting gas onto a central MBH became accepted as the general source of nuclear activities especially for the most powerful AGN members namely quasars, high-luminosity Seyferts and strong radio galaxies (Rees 1984; Blandford et al. 1990; see, e.g., Krolik 1999). Despite the popularity of the paradigm, the rigor stops here due to the lack of any compelling dynamical evidence.

1.1.2 Dynamical evidence for MBHs in inactive galaxies

I. Dead quasar engines in nearby galaxies

Because quasars are populous in the distant youthful universe but mostly die out at redshift \( z < 2 \), the nearby universe should be populated with dead quasar engines, relic MBHs. Under the assumption that quasar luminosity is produced by gas accretion onto central MBHs, Soltan (1982) was first to argue that the average mass density of MBHs \( \rho_u \) should match or exceed the mass-equivalent energy density in quasar light \( u \) as

\[
    u = \int_0^{\infty} \int_0^{\infty} \Phi(L, z) L dL \frac{dt}{dz} dz = 1.3 \times 10^{-16} \text{ J m}^{-3},
\]  

(1.1)
where $\Phi(L, z)$ is the comoving density of quasars with luminosity $L$ at redshift $z$, $t$ is cosmic time (Rees 1988; Marconi et al. 2004). For a radiative energy conversion efficiency $\epsilon$, the present-day MBH mass density is $\rho_u = u/(\epsilon c^2) = 2.2 \times 10^4 \epsilon^{-1} M_\odot \text{Mpc}^{-3}$. If compared with the galaxy luminosity density $j = 1.1 \times 10^8 L_\odot \text{Mpc}^{-3}$ (for $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$; Loveday et al. 1992), we obtain the ratio of the MBH mass to the galaxy bolometric luminosity, as

$$\Upsilon = \frac{\rho_u}{j} = 1.8 \times 10^{-3} \left(\frac{0.1}{\epsilon}\right) \left(\frac{M_\odot}{L_\odot}\right).$$ (1.2)

Popular geometrically-thin optically-thick accretion disk models give an efficiency of $\epsilon \sim 0.1$. Eq. (1.2) then predicts a $\sim 10^7 M_\odot$ MBH in a typical $L_\star \sim 10^{10} L_\odot$ galaxy and $\sim 10^6 M_\odot$ MBHs for dwarf ellipticals. For the brightest quasars with $L \sim 10^{40} \text{ W}$, the Eddington luminosity places an upperbound of $\sim 10^9 M_\odot$ in MBH masses. All these suggest that MBHs with mass $M_\bullet \sim 10^{6-9.5} M_\odot$.

### II. Stellar dynamical search for candidate MBHs

Aside from checking and further developing the AGN paradigm, modern spectroscopical studies performed with the refurbished Hubble space telescope and the largest ground-based telescope have targeted almost exclusively quiescent or weakly active nearby galaxies, where “dormant” MBHs are expected to reside.

At small radii, the MBH’s deep potential well speeds up stars and causes the RMS speed of stars/gas to rise. However, the degeneracy between the MBH mass and velocity anisotropy makes it difficult to distinguish whether the rise in dispersion is associated with the MBH or radial anisotropy of stellar orbits. By fitting dynamical models to both the photometry and the line-of-sight velocity dispersion (LOSVD) measured at different radii especially on parsec scales, several groups have successfully hunted galactic MBH candidates.
Chapter 1: Introduction

(e.g., Richstone & Tremaine 1985; van der Marel et al. 1998; Gebhardt et al. 2003). In brief, all dynamical techniques rely on a determination of mass enclosed within a radius $r$. We refer to the object as a “massive dark object” (MDO). To date, we have irrefutable dynamical evidence for the presence of $10^{6-9} M_\odot$ MDOs in at least 30 nearby galaxies.

III. Are MDOs really MBHs?

Dynamical data of MDOs alone cannot provide rigorous proof of MBHs unless relativistic velocities are detected at a few Schwarzschild radii $^2$, $r_{sch} \simeq 4.8 \times 10^{-6} M_*/(10^8 M_\odot) \text{ pc}$. Only in our own Milky Way, one can measure the orbits of stars passing at $1300 r_{sch}$; in almost all other cases, current measurements cannot probe regions closer than $\sim 10^5 r_{sch}$. In the absence of relativistic signatures, we have to seek for indirect physical arguments to eliminate the possibility of contrived astrophysical alternatives such as dense clusters of stellar-mass BHs.

Maoz (1998) considered the most plausible MBH alternatives: clusters of low-luminosity gravitating objects, from very low-mass objects of cosmic composition ($< 10^{-3} M_\odot$) and brown dwarfs ($\simeq 0.09 M_\odot$) to stellar remnants including white dwarfs ($\simeq 1.4 M_\odot$), neutron stars ($1.4-3 M_\odot$) and stellar-mass black holes ($> 3 M_\odot$). He derived their maximum possible lifetime against evaporation and collisions. When all MDO detections claimed to date are considered, MBHs are verified in the case of the Milky Way, NGC 4258 and Circinus (Ferrarese & Ford 2005). There, a sparse cluster of massive stellar remnants would evaporate completely in $\lesssim 10^8$ yr, while a dark cluster of numerous brown dwarfs would collide, merge and finally turn into stars.

Complementary to Maoz, Miller (2006) used observational constraints to rule out

\footnote{Only observations of the Fe Kα emission line in Type 1 AGNs might give us a change of peering within the relativistic regime of a MBH (Pounds et al. 1990), although this is still considered to be a controversial issue irrelevant to inactive galaxies.}
Chapter 1: Introduction

binary heating and derived a more stringent estimate of lifetime against core collapse, since it is a factor of $\sim 20$ less than the evaporation time. He also showed that the existence of stellar-mass black holes would lead to the production of MBHs in many specific sources, even if the MBHs has not formed in other ways. Using these criteria, one further rules out dense dark clusters as an alternative explanation for the inferred MDO in M31.

1.2 Fuelling MBHs

MBHs now fit in with the broad picture of galaxy evolution as follows: the energy radiated from the agglomeration of gas onto the MBH manifests itself as a large-$z$ quasar, whereas activity in low-$z$ galaxies may be correlated with a tidal merger of stars (Rees 1984, 1998). Galactic spheroids have ample supplies of gas and stars. The gas can be driven by torques from non-axisymmetric potential perturbations such as stellar bars or transient distortions of potential during merger or accretion events (Shlosman et al. 1990). In contrast, the stellar feeding is always of minor importance, because the rate at which stars are scattered into the maw of MBH is too low to reproduce observed AGN luminosities (Young et al. 1977; Frank 1978).

Despite the difficulty of channelling stars to a MBH, debris from tidally disrupted stars is an inevitable source of fuel whenever a MBH was present. Motions of stars are always gravitational and dissipationless. Based on models of the stellar distribution and velocities, estimating rates of feeding stars to a MBH is therefore a relatively “clean” problem. This becomes the theme of the thesis

How often do MBHs consume stars if present at the centre of every galaxy?

What is the stellar feeding rate to the MBH?

\ldots
Chapter 1: Introduction

Below, I introduce one definition of “loss cone” and discuss challenges existed in feeding stars to a MBH.

1.2.1 Loss of stars in the loss cone

Consider a MBH of mass $M_*$ residing at the centre of a galaxy. For simplicity, let us suppose the galaxy is otherwise composed entirely of stars with mass $m_*$ and radius $r_*$. For each star there is a point where the surface gravity of the star equals the tidal force from the MBH across its diameter,

$$r_t = g \left( \frac{M_*}{m_*} \right)^{1/3} r_* \simeq g \times 10^{-5} \text{ pc} \left( \frac{M_*}{10^8 M_\odot} \right)^{1/3} \left( \frac{m_*}{M_\odot} \right)^{-1/3} \left( \frac{r_*}{R_\odot} \right), \quad (1.3)$$

the so-called tidal radius\(^3\). Neglecting relativistic effects, a star that comes within a tidal radius is either tidally disrupted in the case of main-sequence (MS) stars \(^4\) (Rees 1988), or swallowed whole in the case of compact stellar remnants (Hills 1975). Follow Magorrian & Tremaine (1999), the “loss cone” (LC) is defined to consist of all orbits with pericentres less than a characteristic radius $r_{lc} = r_t$. I shall call $r_{lc}$ the effective “LC radius”, call the sphere of radius $r_{lc}$ the “LC sphere”. I shall also say a star has been “consumed” by the MBH once it enters into the LC sphere.

- Tidal disruption of stars provides an inevitable source of fuel to the (dead) quasar engines. After several orbital periods, much of the debris gets ejected but a portion settles into an accretion disk. As material gradually falls into the MBH, intense electromagnetic radiation, UV or soft X-ray “flare” is expected to emerge from the innermost rings, lasting a few months to a year (Rees 1988, 1998; Lee 1999; Bogdanović et al.

---

\(^3\)For stars modelled as $n = 1.5$ polytropes, Ivanov & Chernyakova (2006) find that $g$ varies between 1.1 and 2.4 as $M_*$ varies from $10^6 M_\odot$ to $4 \times 10^7 M_\odot$.

\(^4\)For main-sequence stars, there is a critical MBH mass above which tidal radius occurs well inside the MBH horizon (i.e., Schwarzschild radius $r_{sch} \equiv 2GM_*/c^2$) and ordinary stars are not subjected to tidal disruption but swallowed whole; for the disruption of solar-type stars $M_\text{crit} \sim 10^8 M_\odot$. 
Plausible models (Ulmer 1999) predict $V$-band luminosities of $\sim 10^9 L_\odot$ (see Rees 1988 for theoretical discussion and Komossa et al. 2004 for observations). Successful detections of such disruption flares would constitute robust diagnostics of MBHs in quiescent galaxies; they could also put conceivable constraints on MBH masses and spins (Rees 1998).

- Compact stars, such as white dwarfs, neutron stars and stellar mass BHs, have very small geometrical cross sections. Once captured, they can orbit around a MBH at a distance well inside the tidal radius for ordinary stars of the same mass. As its orbit decays, a burst of gravitational radiation is emitted near perisapse. The typical radiation frequency is determined by the orbital frequency, hence the MBH mass. Inside the tidal radius of a MS star around a $\sim 10^6 M_\odot$ MBH, the inspiral frequency of compact stars, $10^{-4} - 10^{-2}$Hz (Sigurdsson & Rees 1997), well matches the expected sensitivity of the planned space-based gravitational-wave detector LISA (Danzmann 2003). They become one of the most likely gravitational radiation events to be detected by LISA.

1.2.2 How to feed stars to a loss cone?

A brief history of loss cone studies

The study of loss cone dynamics has been addressed by numerous researchers at different situations. Starting with Frank & Rees (1976), Lightman & Shapiro (1977) and Cohn & Kulsrud (1978), classical LC theory was directed towards understanding the observational consequences of MBHs at the centres of globular clusters. Globular clusters are many relaxation times old; under the influence of gravitational encounters, the stellar phase-space density near the MBH has reached an approximate steady-state.
Recently, the LC paradigm has been utilized to estimate rates of astrophysical events in the \textbf{galactic centres}, including tidal disruption of stars (Syer & Ulmer 1999; Magorrian & Tremaine 1999; Wang & Merritt 2004), capture of compact stellar remnants (Sigurdsson & Rees 1997), and scattering stars into the capture zone of MBH binary (Yu 2002). The fact that galactic centres are not collisionally relaxed has implications for a more detailed form of the phase-space density near the LC boundary, and hence for the feeding rate.

\textbf{Two mechanisms}

In spherical galaxies, stars on disruptable orbits are removed by the MBH in less than an orbital time. The rate of all the observable events is set by the rate at which stars can enter the loss cone (Frank & Rees 1976). How to bring stars into the LC? Two basic dynamical mechanisms are generally invoked (Sec. 7.5.9 BT08):

1. In spherical galaxies, stars can enter the LC on disruptable orbits by two-body relaxation and be removed by the MBH in less than an orbital time (Frank & Rees 1976; Lightman & Shapiro 1977).

2. In non-spherical galaxies, torques from the overall mass distribution can carry stars into the LC (e.g., Magorrian & Tremaine 1999).

The most promising way to bring stars to the LC is to break the symmetry. Other more efficient but less general refilling mechanisms have been studied, such as (1) chaotic orbits in triaxial potentials (Norman & Silk 1983; Gerhard & Binney 1985; Merritt & Poon 2004; Holley-Bockelmann & Sigurdsson 2006); however, the presence of a MBH may destroy the triaxiality near the center (Merritt & Quinlan 1998; Holley-Bockelmann et al. 2002; Sellwood 2002); (2) increased fraction of low angular-momentum orbits in non-spherical potentials.
(MT99; Berczik et al. 2006); (3) accelerated resonant relaxation of angular-momentum near the MBH where the orbits are Keplerian (Rauch & Tremaine 1996; Rauch & Ingalls 1998; Hopman & Alexander 2006; Levin 2007); (4) perturbations by a massive accretion disk or an intermediate mass black hole (IMBH) companion (Polnarev & Rees 1994; Zhao et al. 2002; Levin et al. 2005), and massive perturbers (mainly giant molecular clouds) in the centre of the Galaxy (Perets et al. 2007; Perets & Alexander 2008). Most of these mechanisms require special circumstances to work (e.g. specific asymmetries in the potential), or are short-lived (e.g. the IMBH will eventually coalesce with the MBH).

In this thesis, I focus on one of the many possible non-spherical geometries by introducing a low-mass flyby perturber. This is inspired by the fact that galaxies typically inhabit noisy neighborhoods, primarily due to satellites or continuing infall and inhomogeneity in the dark matter distribution (Murali 1999). For relatively weak encounters, Vesperini & Weinberg (2000) find that low-mass interloping galaxies can give rise to significant asymmetries in the primaries. If the non-axisymmetric time-dependent central potentials are strong enough, they may lead to the refilling of a LC on dynamical time-scales.

1.3 Outline and declaration:

In the remainder of this thesis, I set up numerical experiments to model flyby encounters.

- In order to understand the particle-based $N$-body method, we review the theories involved in dynamical modelling collisionless galaxies, followed by numerical techniques for $N$-body modelling (Chapter 2).

- To make the best use of simulations, we develop a general scheme to construct high-
resolution $N$-body models (Chapter 3).

- I use a suit of optimized model and code, both having high-resolution at the centre, to follow the evolution of flyby encounters for a broad range of satellite parameters (Chapter 4).

- I apply the findings to predict a MBH fuelling rate driven by orbiting substructures within the dark matter halo (Chapter 5).

- Finally, I summarize the findings and provide some future work (Chapter 6).

Declaration

The work presented in this thesis is all my own. Large portions of Chapter 3, as well as some of Sections 2.2.1 and 2.2.2 have appeared in the following paper:

Galaxies, as self-gravitating systems, can be idealized as configurations of point mass fully described by a phase-space distribution function (DF) \( f(x,v,t) \). The mass distribution determines the gravitational field through Poisson’s equation. Over their lifetime, galaxies are to a high degree collisionless; the DFs therefore satisfy the Collisionless Boltzmann equation (Sec. 2.1). With the help of Monte Carlo methods (Sec. 2.2.1), the evolution of the DF can be followed by \( N \)-body integrations (Sec. 2.2.2). When encounters are taken into account, one writes the collision term in Master equation form (Sec. 2.3.1) and expands in a Taylor series to derive the Fokker-Planck equation (Sec. 2.3.2).
2.1 Equilibrium model

When making dynamical models, the dynamics of the gas is usually ignored because only about 10% of the mass of an elliptical galaxy is in the form of gas or dust. What is more important, the so-called interstellar medium (ISM) is very responsive to non-gravitational influences such as thermal, magnetic or radiation pressure gradients and stellar winds. In contrast, the stellar dynamical problem is relatively “clean” because motions of stars are always gravitational.

2.1.1 DFs for collisionless stellar system

For most purposes, the gravitational field can be described as the sum of a dominating smooth component (the mean-field $\Phi_{\text{smooth}}$) and a small “granular” part (some short-term, small-scale fluctuations $\delta\Phi$).

$$\Phi = \Phi_{\text{smooth}} + \delta\Phi.$$  \hspace{1cm} (2.1)

The fluctuating $\delta\Phi$ causes a star to change its actual orbit in $\Phi_{\text{smooth}}$ slowly, a process known as “two-body relaxation”. The relaxation time can be estimated as (see eq.1.38, BT08)

$$t_{\text{relax}} \simeq \frac{0.1N}{\ln N} t_{\text{cross}},$$  \hspace{1cm} (2.2)

where it takes $t_{\text{cross}}$—one crossing time—for a typical star to cross the galaxy once. After many crossings, the velocity will have changed completely. Galaxies typically have $N \approx 10^{11}$ stars and $t_{\text{cross}} \approx 100\text{ Myr}$, so in general the effects of gravitational encounters can be neglected over the lifetime $\sim 10\text{ Gyr}$.

The insignificant effects of two-body encounters allow us to idealize a galaxy as a
continuous mass distribution. We therefore describe the system in terms of distribution functions (DF), such that the average number of stars at time $t$ in the range $(x, x + d^3x)$ and $(v, v + d^3v)$ is given by $N f d^3x d^3v$ where $N$ is the total number of stars. In the continuous limit, $f$ can be interpreted as a mass probability density if all stars have mass $1/N$. Integrating over all phase-space gives the total mass

$$
\frac{1}{N} \cdot N \int d^3x d^3v f(x, v, t) = 1 \tag{2.3}
$$

to be 1.

2.1.2 Collisionless Boltzmann equation

Any given star can be viewed as some point in a 6-dimension phase-space

$$(x, v) \equiv w \equiv (w_1, ..., w_6), \tag{2.4}$$

and its motion defines a phase-flow

$$\dot{w} = (\dot{x}, \dot{v}) \equiv (v, -\nabla \Phi). \tag{2.5}$$

For the pure dynamical effects considered here, if stars can neither die nor be born and the flow through phase-space is incompressible, so that

$$f(w(t); t) = f(w_0; 0), \tag{2.7}$$

---

1This description assumes all stars to be identical and requires the six-dimensional volume $d^3x d^3v$ around the point $x$ and velocity $v$ to be infinitesimal.

2If we treat $w = (q, p)$ as an arbitrary system of canonical coordinates, then $\dot{w} = (\dot{q}, \dot{p}) = (\partial H/\partial p, -\partial H/\partial q)$

$$\frac{d\dot{w}}{dw} = \frac{\partial \dot{q}}{\partial q} + \frac{\partial \dot{p}}{\partial p} = \frac{\partial^2 H}{\partial q\partial p} - \frac{\partial^2 H}{\partial p\partial q} \equiv 0. \tag{2.6}$$
the phase-space density is conserved. The time evolution of the DF is governed by the

**Collisionless Boltzmann equation** (CBE eq. 4.4 BT08)

$$\frac{\partial f}{\partial t} + \frac{\partial (\dot{w}f)}{\partial w} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} - \frac{\partial \Phi}{\partial \mathbf{x}} \cdot \frac{\partial f}{\partial \mathbf{v}} = 0,$$

where the gravitational potential

$$\Phi = \Phi_{\text{self}} + \Phi_{\text{ext}},$$

is obtained from some possible external potential $\Phi_{\text{ext}}$ and the DF through Poisson equation

$$\nabla^2 \Phi_{\text{self}} = 4\pi G \int f(x, v) \, d^3v.$$  \hspace{1cm} (2.10)

The CBE can be rewritten in different forms. For example, if we treat $w = (q, p)$ as an arbitrary system of canonical coordinates, the convective or Lagrangian derivative of DF can be written (eq.4.9 BT08)

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + [f, H] = \frac{\partial f}{\partial t} + \dot{q} \frac{\partial f}{\partial q} + \dot{p} \frac{\partial f}{\partial p}$$ \hspace{1cm} (2.11)

where the square bracket is a Poisson bracket (e.g., eq.D.65 BT08); or

$$\left. \frac{df}{dt} \right|_H = 0.$$ \hspace{1cm} (2.12)

The DF is constant along any trajectory in phase-space. In the inertial Cartesian coordinates where the Hamiltonian $H = \frac{1}{2} \mathbf{v}^2 + \Phi(x)$, eq. (2.12) recovers the expression in eq. (2.8).

An integral of motion is any function of the phase-space coordinates alone that is constant along an orbit. For example, the orbital binding energy $\mathcal{E}$ is always conserved
in any stationary potential; if assuming spherical symmetry, three components of angular momentum $J = \mathbf{x} \times \mathbf{v}$ ($J_x, J_y, J_z$) are also conserved. Jeans theorem tells us the DF of any steady-state galaxy can depend on $(\mathbf{x}, \mathbf{v})$ only through $n$ integrals of motion. Unfortunately, this is almost intractable without further help from the strong Jeans theorem, which states: the regular part of a DF may be represented by a function of only three independent isolating integrals\(^3\) (see Sec. 4.2 BT08). So in a spherical potential, the steady-state DF can be a function of angular-momentum. If with no preferred directions of rotation, it can be fully described by $f(\mathcal{E}, J^2)$, where $J^2 = J_x^2 + J_y^2 + J_z^2$.

### 2.1.3 Preliminaries

#### A reference model

Throughout the thesis, we always adopt a reference spherical isotropic stellar system with Hernquist’s (1990) profile\(^4\)

$$\rho_H(r) = \frac{M_s r_s}{2\pi r(r + r_s)^3}, \quad (2.13)$$

where the total mass $M_s$ and scale-radius $r_s$ are unity, and the half mass radius is $r_{1/2} \approx 2.4 r_s$. It possesses a central MBH, where the mass ratio of the MBH to galaxy $\mu = M_s/M_*$ is set to be 0.01. We scale the model to a small galaxy with total mass $M_*=10^{10} \, M_\odot$ and scale-radius $r_*=1 \, \text{kpc}$. The mass ratio is rather larger than the typical ratio, $\sim 10^{-3}$ observed between the MBH mass and bulge mass (Häring & Rix 2004). But as will be seen in Sec. 4.2.1, such a large MBH mass is chosen in order

---

\(^3\)An isolating integral confines orbits and reduces the dimensionality of the phase-space volume that orbits are allowed to occupy by one.
to minimize the rate of relaxation-driven LC refilling, which occurs more rapidly for smaller $M_\bullet$ (e.g., Berczik et al. 2005, 2006). Throughout the text, when we quote a system unit, we may follow it by the physical values. For example, the MBH has a mass of $M_\bullet = 0.01(10^8 M_\odot)$.

In numerical orbital integration, particles move only at discretized time-steps. This requires the simulated density profiles to be non-singular, because the dynamical time becomes formally zero at the origin of a point-mass profile. Hence, the MBH is added as a softened particle with a Plummer profile

$$\rho_p(r) = \frac{3M_\bullet}{4\pi \epsilon_\bullet^3} \left(1 + \frac{r^2}{\epsilon_\bullet^2}\right)^{-5/2},$$

the scale-radius $\epsilon_\bullet = 0.003$ (3 pc). It reduces to a point mass when $\epsilon_\bullet = 0$.

### Units conversion and scalings

For a typical small galaxy with total mass $M_0 = 10^{10} M_\odot$ and scale radius $r_0 = 1$ kpc, we choose a velocity unit to be $V_0 = 100$ km s$^{-1}$, these give the time unit $T_0 = r_0/V_0 \simeq 10^7$ yr; It follows that the gravitational constant in the $N$-body model, where the total mass and the scale radius are unity, is

$$G_N = \frac{GM_0}{r_0^2} / V_0^2 = 4.31,$$

where $G = 6.672 \times 10^{-11}$ m$^3$ kg$^{-1}$ s$^{-2}$.

For another galaxy with total mass $M_1$ and scale radius $r_1$, physical units can
Table 2.1: Model to physical units conversions

<table>
<thead>
<tr>
<th>Model parameter</th>
<th>Physical units</th>
</tr>
</thead>
<tbody>
<tr>
<td>unity $G_N = 4.31$</td>
<td></td>
</tr>
<tr>
<td>Mass</td>
<td>$10^{10} M_\odot$  $10^{12} M_\odot$</td>
</tr>
<tr>
<td>Length</td>
<td>1 kpc</td>
</tr>
<tr>
<td>Time</td>
<td>$10^7$ yr</td>
</tr>
<tr>
<td>Velocity</td>
<td>100 km s$^{-1}$</td>
</tr>
</tbody>
</table>

be scaled as follows:

\[
T_1 = 10^7 \text{ yr} \left( \frac{M_1}{10^{10} M_\odot} \right)^{-\frac{1}{2}} \left( \frac{r_1}{1 \text{ kpc}} \right)^{\frac{3}{2}},
\]  

(2.16)

\[
V_1 = 100 \text{ km s}^{-1} \left( \frac{M_1}{10^{10} M_\odot} \right)^{\frac{1}{2}} \left( \frac{r_1}{1 \text{ kpc}} \right)^{-\frac{1}{2}}.
\]  

(2.17)

Table 2.1 summarizes some fiducial values and their implied N-body scalings.

A MBH dominates the central gravity up to its gravitational influence radius, where the enclosed stellar mass equals the MBH mass

\[
M_{*}(r \leq r_{\text{infl}}) = M_\bullet.
\]  

(2.18)

In principle it is straightforward to add the MBH’s contribution to the potential field

\[
\psi(r) = \frac{GM_\bullet}{r + r_*} + \frac{GM_\bullet}{\sqrt{r^2 + \epsilon_*^2}}
\]  

(2.19)

where $\psi(r)$ remains a function of $r$. An ergodic DF $f(\mathcal{E})$ then follows by inverting

\[
\rho_H(r) = 4\pi \int_0^{\psi(r)} f(\mathcal{E}) \sqrt{2(\psi(r) - \mathcal{E})} d\mathcal{E}.
\]  

(2.20)
We solve eq. (2.20) for $f(E)$ using Eddington’s formula

$$f(E) = \frac{1}{\sqrt{8\pi^2}} \left[ \int_0^E \frac{d\psi}{\sqrt{E - \psi}} d^2\rho + \frac{1}{\sqrt{E}} \left( \frac{d\rho}{d\psi} \right)_{\psi=0} \right]. \quad (2.21)$$

The numerical method is described in Appendix A.1. As shown on the left panel of figure 2.1, the presence of a central MBH enables stars with large binding energy to populate the galaxy; this is not possible in the self-consistent finite potential of stars alone. The velocity dispersion given by

$$\sigma^2 = \frac{\int d^3v \, v^2 f(x, v)}{\int d^3v \, f(x, v)} \quad (2.22)$$

is just the second momentum of DF. As shown on the right panel of figure 2.1, there is a significant increase in $\sigma^2$ within the MBH’s sphere of influence $r_{\text{infl}} \approx 0.16(160 \text{ pc})$. Below, we present some terminology and equations.

The radial period, the time taken for a star of energy $E$ and angular-momentum $J$ to travel from pericentre to apocentre distances $r_{\text{peri}}$ and $r_{\text{apo}}$ and back is (eq. 3.17 BT08)

$$T_r(E, J^2) = 2 \int_{r_{\text{peri}}}^{r_{\text{apo}}} \frac{dr}{\sqrt{2[\psi(r) - E] - J^2/r^2}}. \quad (2.23)$$

Similarly, the azimuthal period is (eqs. 3.18b and 3.19 BT08)

$$T_\phi = \frac{2\pi}{|\Delta\phi|} T_r, \quad \Delta\phi = 2J \int_{r_{\text{peri}}}^{r_{\text{apo}}} \frac{dr}{r^2 \sqrt{2[\psi(r) - E] - J^2/r^2}}. \quad (2.24)$$

The radial period is mainly determined by the orbital binding energy $E$, so for highly eccentric radial orbits, eq. (2.23) becomes

$$T_r(E) = 2 \int_0^{r_E} \frac{dr}{\sqrt{2[\psi(r) - E]}}. \quad (2.25)$$
where \( r_E \) is the energy labeled radius at which \( \psi(r_E) = E \), and \( r_{apo} < r_E \). \( T_r(E) \) is shown on the right panel of figure 2.2.

The **density of states** \( g(E, J^2) \) plays an important role in the statistical description of stellar dynamical system. In the limit of classical physics,

\[
g(E, J^2) = 8\pi^2 \int_{r_{peri}}^{r_{apo}} \frac{dr}{\sqrt{2[\psi(r) - E] - J^2/r^2}},
\]

(2.26)

\( g(E, J^2) \) is the volume of phase-space per unit \( E \) per unit \( J^2 \) (eq. 4.81 of BT08). Compared to the radial period (\( T_r \) in eq. 2.23), we have \( g(E, J^2) = 4\pi^2 T_r(E, J^2) \). The fraction of stars inside any small volume of phase-space \( dE dJ^2 \) around \( (E, J^2) \) is

\[
N(E, J^2) dE dJ^2 = 4\pi^2 f(E, J^2) T_r(E, J^2) dE dJ^2.
\]

(2.27)

With an isotropic DF \( f(E) \), the fraction of stars that have energies in the range \( (E, E + dE) \) is given by

\[
N(E) dE = 4\pi^2 f(E) dE \int_{0}^{J^2(E)} T_r(E, J^2) dJ^2,
\]

(2.28)

where \( J^2_c(E) \) is the angular-momentum of a circular orbit with energy \( E \). \( N(E) \) is also known as the **differential energy distribution** (e.g., Sec. 4.3.1 BT08).

### 2.2 N-body modelling

In the collisionless ideal, the evolution of DF satisfies the CBE. In principle, the coupled CBE and Poisson equations can be solved by finite-difference methods, but this would require impractically large grids in six-dimensional phase-space. In practice, the DF is followed by the so-called collisionless \( N \)-body simulations. These \( N \) particles do not represent real stars. Instead, they provide a Monte Carlo realization of the smooth
Figure 2.1: Distribution function (left panel) and velocity dispersion (right panel) for three models. They all have Hernquist’s density profile of $M_\ast = 1(10^{10} M_\odot)$ and scale length $r_\ast = 1(1 \text{ kpc})$, and an isotropic velocity distribution. The self-consistent Hernquist (1990) model is shown in thick-solid curves. The other two models both harbour a Plummer’s MBH of mass $M_\bullet = 0.01(10^8 M_\odot)$, one with a scale length of $\epsilon_\bullet = 0.003(3 \text{ pc})$ (thin-solid) the other with an unsoftened $\epsilon_\bullet = 0$ MBH (thin-dashed).

Figure 2.2: Density of states (left panel) and radial period (right panel). Line styles are the same as in figure 2.1.
underlying DF from which one can estimate the potential $\Phi(x, t)$. By integrating the orbits, one is solving the CBE by the method of characteristics (Hernquist & Ostriker 1992; Leeuwin Combes & Binney 1993). After reviewing the basics of Monte Carlo integration, we explain the connection between the CBE and $N$-body simulations, together with some $N$-body techniques.

### 2.2.1 Monte Carlo integration

For later reference we recall some of the basic ideas (e.g., Press 1992) in using Monte Carlo methods to evaluate integrals, such as

$$I = \int_D f \, dV = \langle f \rangle,$$  \hspace{1cm} (2.29)

of a known function $f$ over a domain $D$. We first consider the case where $D$ has unit volume: $\int_D dV = 1$. Then, given $N$ points, $x_1 \ldots x_N$, drawn uniformly from $D$, the estimate for the integral is given by

$$E = \frac{1}{N} \sum_{i=1}^{N} f(x_i).$$  \hspace{1cm} (2.30)

In the large $N$ limit, $E \to I$. The variance in the estimate (eq. 2.30), which measures the averaged squared-distance of the estimator from its true value,

$$\text{Var}E = \langle (E - I)^2 \rangle = \langle E^2 \rangle - I^2,$$  \hspace{1cm} (2.31)
is asymptotically related to the variance of the function $f$, $\text{Var}(f) = \langle (f - I)^2 \rangle = \int_D |f - I|^2 \, dV$, by the relation 5

$$\text{Var}E = \frac{1}{N} \text{Var}(f) = \frac{1}{N} \left[ \int_D f^2 \, dV - I^2 \right]. \quad (2.33)$$

Now let us relax the assumption that $D$ has unit volume and, instead of drawing points uniformly from $D$, let us take $N$ points drawn from a sampling distribution $f_s$. We assume that $f_s$ is normalized: $\int_D f_s \, dV = 1$. Making a change of variables from $dV$ to $dV' = f_s \, dV$, the integral (2.29) can be estimated as

$$I = \int_D \frac{f}{f_s} \cdot dV' \simeq \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{f_s(x_i)}. \quad (2.34)$$

The variance of the integrated function $f/f_s$ becomes

$$\text{Var}(f/f_s) = \int_D \left[ \frac{f}{f_s} - I \right]^2 \, dV', \quad (2.35)$$

and the variance in the estimate (eq. 2.34) can be approximated as

$$\frac{1}{N} \text{Var}(f/f_s) = \frac{1}{N} \left[ \int_D \left( \frac{f}{f_s} \right)^2 \, dV' - I^2 \right] = \frac{1}{N} \left[ \int_D f^2 \, dV - I^2 \right]. \quad (2.36)$$

---

5We detail the calculation of $\text{Var}E = \langle E^2 \rangle - I^2$ in eq. 2.31

$$\langle E^2 \rangle = \left\langle \frac{1}{N^2} \sum_{i,j=1}^{N} f(x_i) f(x_j) \right\rangle = \frac{1}{N^2} \left( \sum_{i=1}^{N} f^2(x_i) + 2 \sum_{1 \leq i < j \leq N} f(x_i) f(x_j) \right).$$

Since $x_1, \ldots, x_N$ are $N$ independent variables drawn from $V$, we have $\langle f(x_i) f(x_j) \rangle = \langle f(x_i) \rangle \cdot \langle f(x_j) \rangle = I^2$. Therefore, $\langle E^2 \rangle = \frac{1}{N} \langle f^2 \rangle + \frac{N(N-1)}{N^2} I^2$ and eq. (2.31) becomes

$$\text{Var}E = \frac{1}{N} \langle f^2 \rangle + \left[ \frac{N(N-1)}{N^2} - 1 \right] I^2 = \frac{1}{N} \left[ \langle f^2 \rangle - I^2 \right] = \frac{1}{N} \left[ \int_D f^2 \, dV - I^2 \right] = \frac{1}{N} \text{Var}(f). \quad (2.32)$$
2.2.2 The CBE and \(N\)-body simulations

The CBE is of central importance for collisionless \(N\)-body simulations. First, an initial equilibrium distribution of stars is idealized as time-independent solution to the CBE; it can be founded when the partial derivative of the DF satisfies \(\partial f / \partial t = 0\). Secondly, the time-dependent first-order partial differential equation (PDE 2.8) can be solved by integrating the characteristic\(^6\) equations,

\[
\frac{dt}{1} = \frac{d\mathbf{x}}{v} = \frac{dv}{a},
\]

(2.40)

together with using Monte Carlo integration to estimate the acceleration

\[
a(x; t) \equiv -\frac{\partial \Phi}{\partial \mathbf{x}} = -G \nabla \int f(\mathbf{w}'; t) \frac{\mathbf{w}}{|\mathbf{x} - \mathbf{w}'|} d^6 \mathbf{w}'.
\]

(2.41)

From eq. (2.34) it follows that

\[
a(x; t) \simeq -G \nabla \sum_{i=1}^{N} \frac{m_i}{|\mathbf{x} - \mathbf{x}_i|},
\]

(2.42)

corresponding to a distribution of \(N\) point particles with masses

\[
m_i = \frac{1}{N} \frac{f(w_i; t)}{f_s(w_i; t)}.
\]

(2.43)

\(^6\)In a mathematical language, the first-order partial differential equation of \(f\) (PDE 2.8) is solved by characteristics. The goal is to change coordinates from \((x(t), v(t), t)\) to a new coordinate system \((x(s), v(s), s)\) in which the PDE becomes an ordinary differential equation (ODE) along characteristic curve in the plane

\[
\frac{df}{ds} = \frac{\partial f}{\partial s} + \frac{dx}{ds} \frac{\partial f}{\partial x} + \frac{dv}{ds} \frac{\partial f}{\partial v} = 0.
\]

(2.37)

If we regard \(a = -\frac{\partial \Phi}{\partial \mathbf{x}}\) as given, the new variable \(s\) will vary but \(f\) equals a constant following a particle’s trajectory

\[
\frac{dt}{ds} = 1, \quad \frac{dx}{ds} = v, \quad \frac{dv}{ds} = a.
\]

(2.38)

Or along the characteristics

\[
\frac{dt}{1} = \frac{d\mathbf{x}}{v} = \frac{dv}{a},
\]

(2.39)

in phase-space.
I. Initial conditions

Eq. (2.42) requires a set of $N$ bodies, each possessing a mass $m_i$, a position $x_i$ and a velocity $v_i$ ($i = 1 \ldots N$), to be initialized. These $m_i$ clearly depend on the choice of sampling DF $f_s$. The simplest choice is $f_s(w; t) = f(w; t) = f(w_0; 0)$, in which case all particles have equal masses $m_i = 1/N$. However, one is free to tailor the choice of $f_s$ to suit the particular problem under study. With the notable exception of some heuristic multi-mass schemes (e.g., Sigurdsson, Hernquist & Quinlan 1995, hereafter SHQ95; Weinberg & Katz 2007; Sellwood 2008 and Zemp et al. 2008), most other IC-generation schemes have used equal-mass particles. The construction of “multi-mass schemes for collisionless $N$-body simulations” becomes the theme of Chapter 3.

Generally speaking, an initial condition is a representation of the sampling DF $f_s(x, v, t_0)$ at some instant time $t_0$. In effect, the continuous sampling DF in eq. (2.43) is replaced with a set of delta functions

$$f_s(x, v) \rightarrow \sum_{i=1}^{N} m_i \delta^3(x - x_i) \delta^3(v - v_i). \quad (2.44)$$

For this substitution to work, the expected mass of the bodies within any phase-space volume $V$ must be equal to the integral of the sampling DF over that volume,

$$\int_V d^3x d^3v f_s(x, v) \simeq \left\langle \sum_{(x_i, v_i) \in V} m_i \right\rangle, \quad (2.45)$$

where the angle brackets indicate an average over statistically equivalent realizations. Since bodies are selected independently, the actual number within any given volume $V$ will have a Poissonian distribution about the mean. More fundamentally, a coarse-grained representation of a continuous distribution of mass introduces numerical artifacts, such as non-zero
multipoles for a spherical model, which in general can only be reduced by using a large \( N \).

II. Poisson solvers

Eq. (2.42) provides only the most simple-minded estimate of the integral (2.41), and one has some leeway to reconstruct \( f(w; t) \) from the discrete realization furnished by the \( N \) particles. Of course, the reliability of any sensible reconstruction will be wholly dependent on how well the DF is sampled. To obtain \( a(x) \), collisionless simulations use techniques:

- Direct summation – softened force kernels

The most straight-forward evaluation of the forces on particle \( i \) is by simply calculating all pairwise interactions between all the particles; together with a force softening kernel \( S_F(x_k - x_i) \) via

\[
a_i = -G \sum_{k \neq i} m_k S_F(x_k - x_i) \frac{x_k - x_i}{|x_k - x_i|},
\]

(2.46)

For example \( S_F = (r + \epsilon)^{-2} \), where \( r = |x_k - x_i| \). This modification to the Newtonian gravity avoids the Keplerian divergence in the potential field, and suppresses artificial small scale fluctuations caused by two-body encounters which would otherwise require very careful and CPU intensive numerical integration of the equations of motion. Among all the approaches, the direct summation method is robust, accurate and completely general. But it requires \( N - 1 \) calculations for each of the \( N \) particles, and \( \mathcal{O}(N^2) \) operations for a complete force calculation. Improvements such as efficiently using individual time-steps (e.g., Ahmad & Cohen 1973; Aarseth 1985) and implementing in specialised hardware (the GRAvity PipE, or GRAPE computers, Kawai et al. 2000), have kept direct summation
remarkably competitive.

- Hierarchical evaluation

Hierarchical force calculation algorithms provide fast, general and reasonably accurate approximation for gravity. They exploit the fact that higher-order multipoles of an object’s gravitational field decay rapidly with respect to the dominant monopole term. Hence the long-range gravitational potential of a region can be approximated by $1/r$. By doing a $O(\log N)$ splitting, all hierarchical codes partition the mass distribution into a tree structure, where each node of the tree provides a concise description of the matter within some spatial volume. Tree-codes replace the sum over $N-1$ bodies in direct summation methods with a sum over only $O(\log N)$ regions, and therefore can be viewed as hierarchical variations on direct summation with a reduced complexity $O(N \log N)$ (Barnes & Hut 1986). Tree structures may be created either by hierarchically grouping particles or by recursively subdividing space. Despite the merit of employing no grid and parallelizability, it lacks error bounds and automatic ways of adjusting representations.

- Field expansion

Field methods represent the potential and density as series expansion:

$$\Phi(x) = \sum_k A_k \Phi_k(x), \quad \rho(x) = \sum_k A_k \rho_k(x),$$

(2.47)

where $A_k$ are the expansion coefficients, and the basis functions $\Phi_k$ and $\rho_k$ are related by Poisson’s equation,

$$\nabla^2 \Phi_k = 4\pi G \rho_k,$$

(2.48)
The basic procedure is to determine the expansion coefficients by fitting the density to the mass distribution, and then to obtain forces by differentiating the expansion of the potential field. There are many ways to do this, for example:

(1) Particle-Mesh codes (PM) or grid methods employ a mesh or grid to estimate the density at a set of discrete points. For the simplest unchanging Cartesian grid, Poisson equation (2.48) can be solved using a fast Fourier transform (FFT) routine (Sellwood 1987). In this case, the computational cost scales as $O(N + N_g \log N_g)$, where $N_g$ is the number of grid points. Note that in the limit of large $N \gg N_g$, the complexity scales as $O(N)$; the superior efficiency of PM techniques enables a much larger particle number that can be used. The potential in simulations is smoothed by the effective softening length implied by the grid size. Variations include P$^3$M (particle-particle-particle-mesh) codes and AP$^3$M (adaptive-P$^3$M) codes.

(2) Self-Consistent Field methods (SCF Hernquist & Ostriker 1992) tailor a series of orthogonal basis functions to fit the basis geometry of the system and calculate overlap integrals. The angular dependence is expanded in the familiar $Y_{lm}(\theta, \phi)$ spherical harmonics while one carefully designs radial basis sets. By deleting selected terms in the series expansion, one can enforce various symmetries and conserve angular momentum to high accuracy; but not linear momentum. Computational cost scales as $\sim O(N n_{\text{max}} l_{\text{max}}^2)$ where $n$ is the radial and $l$ is the angular “quantum” number. However the coordinate singularity at the origin makes these codes unreliable tools for investigating the dynamics close to the centre of the stellar system. It also suffers from the inflexibility in adaptation to different applications and resolutions; basically for every new application an appropriate set
of basis functions to evaluate the potential has to be researched, and the attempt to increase small-scale resolution, e.g., in angular direction, increases the CPU time nearly as prohibitively as for direct \( N \)-body simulations, since CPU time goes roughly with \( N_t^2 \), where \( N_t \) is the number of terms in the series evaluation for tangential resolution, much like a force calculation scales with \( N^2 \) where \( N \) is the total particle number.

For \textit{collisionless} systems where the effects of two-body encounters are approximated away, the objective then is to suppress the relaxation caused by \( \sqrt{N} \)-type fluctuations in the particle distributions.

\textbf{III. Leap-frog integrator}

To follow the characteristic equation (2.40), an ideal algorithm should enforce exactly certain conservation laws characteristic of Hamiltonian dynamics. The time-centred leapfrog integrator is an optimal choice since

- It treats \( x \) and \( v \) symmetrically, so its numerical solution is time-reversible if computed with a constant time-step \( \tau \).

- The phase-space flow should preserve its symplectic structure\(^7\) and prohibits numerical dissipation (see Chapter 3.4 BT08).

- Second-order leapfrog integrator also conserves linear momentum, provided the accelerations are obtained using a potential solver that respects Newton’s third law.

\(^7\)The Newtonian gravitational \( N \)-body system is a Hamiltonian system in which the phase-space is a smooth manifold. Birkhoff’s ergodic theorem (e.g., Cornfeld 1982) states that the symplectic structure is invariant under the Hamiltonian flow, where the symplectic structure \( dp \wedge dq \) is just a measure on the phase-space with the canonical coordinates \((q,p)\).
Most materials below are adapted from Magorrian (2007). Saha & Tremaine (1992) give a nice review on the standard leapfrog integrator, focusing on the integrability of Hamiltonians of the form (e.g., Wisdom & Holman 1991)

\[ H = T + \sum_{k=-\infty}^{\infty} \delta_\epsilon \left( k - \frac{1}{2} \right) V(x_1, \ldots, x_N) \]  

(2.49)

where \( T \equiv \frac{1}{2} \sum_i m_i v_i^2 \) is the kinetic energy of all the particles and \( \delta_\epsilon(x) \equiv \frac{1}{2}(\delta(x-\epsilon) + \delta(x+\epsilon)) \) with \( 0 < \epsilon \ll 1 \). The periodic comb of delta functions turns on the potential energy \( V(x_1, \ldots, x_N) \) only at times \( t = (k \pm \epsilon)\tau \) for integer \( k \). Integrating the resulting equations of motion from time \( t = k\tau \) to \( t = (k+1)\tau \) yields

\[
\begin{align*}
v_i(k + \frac{1}{2}) &= v_i(k) + \frac{1}{2}\tau a_i(k), \\
x_i(k + 1) &= x_i(k) + \tau v_i(k + \frac{1}{2}), \\
v_i(k + 1) &= v_i(k + \frac{1}{2}) + \frac{1}{2}\tau a_i(k + 1),
\end{align*}
\]  

(2.50)

where the accelerations \( a_i(k) \equiv -\frac{\partial V}{m_i \partial x_i} \) evaluated at time \( t = k\tau \). This reflects the fact that assuming a smooth potential, the fixed leapfrog step \( \tau \) is an exact solution for motions in a time-dependent Hamiltonian.

There are two versions of the leapfrog, depending on the starting point, the sequence of steps for either the kick-drift-kick (eqs. 2.50) form or the drift-kick-drift form. The latter can be obtained by adding \( \frac{1}{2} \) to the argument of the delta functions or, alternatively, by integrating the equations of motion from \( (k - \frac{1}{2})\tau \) to \( (k + \frac{1}{2})\tau \) instead. Another way is to consider each of these versions as compositions of the two time-asymmetric first-order symplectic integrators (each applied left to right), \( K(\tau/2)D(\tau/2) \) and \( D(\tau/2)K(\tau/2) \), whose first-order error terms cancel (Saha & Tremaine 1992).

**Summary**

The principle of an \( N \)-body simulation can be summarized as
1. Setting up initial conditions (ICs), assigning a mass, position and velocity to each particle.

2. Finding the forces, using a Poisson solver to efficiently calculate the gravitational forces on each particle.

3. Moving particles, using an integrator to advance the position and momentum of each particle for a short time.

2.3 Departure from equilibrium

The gravitational field of real galaxies cannot be smooth on small scales. As already shown in eq. 2.1, \( \Phi = \Phi_{\text{smooth}} + \delta \Phi \); the fluctuation term \( \delta \Phi \) causes stars to diffuse in phase-space away from their original orbits. With some modifications, the CBE can still provide satisfactory descriptions of the behavior of collisionless stellar system in the presence of encounters.

2.3.1 Collision terms

When encounters are taken into account, particles begin to influence each other individually as opposed to collectively. The phase-space probability around a star changes with time and the convective derivative of the DF does not vanish anymore. This can be taken into account by introducing an interaction term, namely the encounter operator \( \Gamma[f] \), to the RHS of CBE (in eq. 2.12) as

\[
\frac{d f}{d t} = \Gamma[f]. \tag{2.51}
\]

Below, we follow the treatment presented in Sec. 7.4 of BT08 to develop an expression for \( \Gamma[f] \).

Let \( \Psi(w, \Delta w) d^6 \Delta w \Delta t \) be the probability that a particle at the phase-space coordinates \( w \) is perturbed (through forces not derived from \( \psi_{\text{smooth}} \) as already accounted for in the CBE) to \( w + \Delta w \) during short time interval \( \Delta t \). Stars are scattered out of an element of phase-space around

---

A mathematical precise definition of the encounter operator is given in (eq. 7.47) Sec. 7.3 of BT08. In words it is a measure of the rate of change of the phase-space density around a given star, driven by the correlation between particles in phase-space.
Chapter 2: Dynamical modelling

\( w \) at a rate

\[
\Gamma_- = -f(w) \int d^6(\Delta w) \Psi(w, \Delta w), \tag{2.52}
\]

while stars from other phase-space position \((w - \Delta w)\) scatter into this element at a rate

\[
\Gamma_+ = \int d^6(\Delta w) \Psi(w - \Delta w, \Delta w)f(w - \Delta w). \tag{2.53}
\]

The collision term is thus \( \Gamma = \Gamma_- + \Gamma_+ \). Hence, we arrive at a modified CBE with the collision term— the master equation.

2.3.2 Fokker-Planck equation

In the case of stellar dynamics, at least two critical approximations can be made to simplify the study of the evolution of stellar system due to encounters.

First, the dominance of weak encounters guarantees typical changes \( \Delta w \) are small and the transition probability \( \Psi \) is sufficiently smooth. The Fokker-Planck treatment develops \( \Psi \) and \( f \) around \( w \) in a Taylor series to second order in \( \Delta w \) (eq. 7.65 BT08), so the encounter operator becomes

\[
\Gamma[f] = -\sum_{i=1}^{6} \frac{\partial}{\partial w_i}\{D[\Delta w_i]f(w)\} + \frac{1}{2} \sum_{i,j=1}^{6} \frac{\partial^2}{\partial w_i \partial w_j}\{D[\Delta w_i \Delta w_j]f(w)\}, \tag{2.54}
\]

where

\[
D[\Delta w_i] = \int d^6(\Delta w) \Psi(w, \Delta w) \Delta w_i, \tag{2.55}
\]

is called the drift vector and

\[
D[\Delta w_i \Delta w_j] = \int d^6(\Delta w) \Psi(w, \Delta w) \Delta w_i \Delta w_j, \tag{2.56}
\]

is called the diffusion tensor. Both diffusion coefficients are functions of \((w, t)\) that measure the rate at which collisional effects cause stars to drift in \( w_i \) and diffuse in \( w_i w_j \) respectively.

Secondly, the dominance of local encounters affects velocity only. Neglecting the changes in position and choosing the canonical phase-space coordinates to be Cartesian coordinates \((x, v)\),
\[ \Psi(w, \Delta w) \] is zero unless \( \Delta x = 0 \). This leads to the governing Fokker-Planck equation

\[
\frac{df}{dt} = -\sum_{i=1}^{3} \frac{\partial}{\partial v_i} \{ D[\Delta v_i] f(w) \} + \frac{1}{2} \sum_{i,j=1}^{3} \frac{\partial^2}{\partial v_i \partial v_j} \{ D[\Delta v_i \Delta v_j] f(w) \}. \tag{2.57}
\]

Computations of these velocity diffusion coefficients \( D[\Delta v_i], D[\Delta v_i \Delta v_j] \) are well documented in Sec. 7.4 of BT08.

### 2.3.3 Relaxation time

Diffusion coefficients can be used to improve the crude estimate of the speed of relaxation processes (eq. 2.2). Based on the diffusion coefficient for the parallel component of the velocity \( D[(\Delta v_{||})^2] \), BT08 defines the relaxation time of a subject star to be (see eq. 7.105)

\[
t_{\text{relax}} = \frac{v^2}{D[(\Delta v_{||})^2]}. \tag{2.58}
\]

In this approach, relaxation is reduced to the cumulative effects of a large number of uncorrelated two-body encounters, which can be treated like local Keplerian small-angle hyperbolic velocity deflections due to objects with a density and velocity distribution identical to the local ones. To measure the characteristic relaxation time for a population of identical stars with mass \( m_* \), BT08 assume that the velocity distribution of the field stars is Maxwellian with dispersion \( \sigma \). We have

\[
t_{\text{relax}} = \frac{0.34 \sigma^3}{G^2 m_* \rho \ln \Lambda}
= 0.95 \times 10^{10} \text{yr} \left( \frac{\sigma}{200 \text{ km s}^{-1}} \right)^3 \left( \frac{\rho}{10^6 \text{ M}_\odot \text{ pc}^{-3}} \right)^{-1} \left( \frac{m_*}{\text{M}_\odot} \right)^{1} \left( \frac{\ln \Lambda}{15} \right)^{-1}. \tag{2.59}
\]

The fudge factor \( \ln \Lambda \), called the Coulomb logarithm, is determined through \( \Lambda \approx b_{\text{max}}/b_{90} \), the ratio between the largest impact radius and the so called 90\(^{\circ}\) deflection radius \( b_{90} \equiv G(m_* + m)/V_{\text{typ}} \). Essentially, a locally Maxwellian velocity distribution will ultimately be established beyond \( t_{\text{relax}} \).

Applying this formula to estimate relaxation time around the central MBH, one simply takes the maximum impact parameter to be the influence radius of the MBH \( b_{\text{max}} = r_{\text{infl}} \) (eq. 2.18) and \( b_{\text{min}} = 2Gm_*/\sigma^2 \). For a sample of early-type galaxies known to host MBHs, Merritt, Mikkola
Figure 2.3: Relaxation time versus the central stellar velocity dispersion for a sample of early-type galaxies known to host MBHs (Côté et al. 2004). Merritt, Mikkola & Szell (2007) use eq. (2.59) to measure $t_{\text{relax}}$ at MBHs’ influence radius, assuming solar-type stars. Filled symbols are galaxies in which the MBH’s influence radius is resolved; the star is the Milky Way bulge.

& Szell (2007) use eq. (2.59) to measure $t_{\text{relax}}$ at MBHs’ influence radius, assuming solar-type stars. This is shown in figure 2.3. Considering both the (relative) low stellar density $\rho < 10^6 \text{M}_\odot \text{pc}^{-3}$ and the presence of a MBH which increases $\sigma$ (Faber et al. 1997; Ferrarese et al. 2006), stellar collisions are probably always of minor importance in our study of galactic centre, at least on a scale much larger than the event horizon with $r \gg r_{\text{sch}}$. 
Chapter 3

Multi-mass schemes
for collisionless $N$-body simulations

A general scheme for constructing Monte Carlo realizations of equilibrium, collisionless galaxy models with known distribution function (DF) $f_0$ is established. It uses importance sampling to find the sampling DF $f_s$ that minimizes the mean-square formal errors in a given set of projections of the DF $f_0$. The result is a multi-mass $N$-body realization of the galaxy model in which “interesting” regions of phase-space are densely populated by lots of low-mass particles, increasing the effective $N$ there, and less interesting regions by fewer, higher-mass particles.

The chapter is organized as follows. After recapping in the connection between $N$-body simulations and the CBE (Sec. 3.1.1) and presenting two possible weapons to fight against small $N$ limitation (Sec. 3.1.2), we explain our multi-mass formulation in Sec. 3.2. In Sec. 3.3 we give two examples of using our scheme to suppress fluctuations in the monopole component of acceleration in spherical bulge models with or without a central MBH. We calculate formal estimates of the noise in $N$-body models constructed using equal-mass scheme and SHQ95’s method, and compare them to our own scheme (Sec. 3.3.2). In Sec. 3.3.3 and 3.4 we test how well our realizations behave in practice when evolved using a real $N$-body code. For the bulge + MBH model in Sec. 3.4, much care has been taken to resolve the dynamics around the MBH.
3.1 Collisionless $N$-body simulations

3.1.1 $N$-body simulations and the CBE

When modelling galaxies, the number of stars is large enough and the two-body relaxation time is long enough (greatly exceeds the age) that discreteness effects, such as two-body relaxation, are usually unimportant. In the limit of a very large number $N$ of bodies, stars and dark matter particles move in a smooth mean-field potential $\Phi(x; t)$ and behave as a collisionless fluid in six-dimensional phase-space, the (mass) density at any point $(x, v)$ being given by the distribution function $f(x, v; t)$. The time-evolution of the DF is described by the Collisionless Boltzmann Equation (2.8).

3.1.2 How to fight against small $N$ limitations?

In reality, no simulation is perfectly collisionless because Poisson noise in the estimates of $\Phi(x, t)$ inevitably leads to numerical error in particles’ accelerations and hence diffusion in their orbits. To reduce this noise, it is important to make $N$, the number of particles being used, as large as possible. Unfortunately, the cost of running an $N$-body code scales at least linearly with $N$, so increasing $N$ also makes the simulation more expensive to run. The good news is that alternative, more sophisticated weapons are available for use in the fight against small $N$ limitations.

Improve estimates of the acceleration field

Given a discrete $N$-particle realization of the underlying DF $f$, the most sophisticated approaches to estimate the accelerations (e.g., Dehnen 2001 and references therein) have focused on finding softening kernels that minimize the errors in the acceleration field given a static distribution of $N$ equal-mass particles.

By softening the force at inter-particle separations, one reduces the fluctuations due to close encounters, that is to improve the estimate of the derivative of gravitational potential locally. However, most of the relaxation is due to the variation of the potential field (noise) on larger scales.
(Theis 1998) and therefore, softening alone does not much reduce the artificial \(^1\) two-body relaxation.

**Improve sampling from DFs**

A collisionless \(N\)-body code is essentially a Monte Carlo method and so should be amenable to well-known variance-reduction methods such as **importance sampling** (Press 1992).

In this chapter, we concentrate on developing a generally-applicable, essentially model-independent method for constructing \(N\)-body realizations of isolated model galaxies in equilibrium, suitable for use as initial conditions in collisionless simulations. We do not investigate how different softening lengths or softening kernels affect \(N\)-body models; the bottom line is that the reliability of any sensible reconstruction will be wholly dependent on how well the DF is sampled. Our scheme uses importance sampling to find a sampling DF \(f_s\) that minimizes the mean-square uncertainty in a chosen set of projections of the DF \(f_0\). The trajectories are then sampled randomly from the sampling DF \(f_s\). This is very different from the so-called quiet start technique (Sellwood 1983), in which one reduces the noise by a careful rather than random choice of the trajectories.

For example if modelling bar evolution, one might be most interested in following the detailed evolution of the DF around the strongest resonances. It is natural then to try to increase the sampling density near these regions by populating them with lots of low-mass particles. Outside these interesting regions, however, one must also have enough particles to maintain accurate estimations of the force field which governs the evolution of the system as a whole.

### 3.2 Formulation

#### 3.2.1 Observables

What constitutes a “good” choice of sampling density \(f_s\)? The DF \(f\) is a high-dimensional probability density and itself is not measurable. We are usually only interested in coarse-grained

\(^1\)In collisionless simulations, the bodies are just a representation of the one-particle DF. Consequently, binaries as well as two-body encounters are entirely artificial (Dehnen 2001).
projections of the DF,
\[ \langle Q_i \rangle \equiv \int f(w)Q_i(w) \, d^6 w, \]
where the kernels \( Q_i(w) \) are some functions of the phase-space co-ordinates \((x, v)\). For the purposes of the present paper, we consider a “good” sampling scheme to be one that minimizes the uncertainty in the estimates of some given set of \( \langle Q_i \rangle \). Apart from some general guidance, we do not address the question of how best to choose these \( Q_i \), which usually requires some experience of the particular problem at hand.

We now give some examples. It is helpful to introduce the indicator function
\[ 1_V(w) = \begin{cases} 1, & \text{if } w \in V \\ 0, & \text{otherwise}. \end{cases} \]
Then a particularly simple but important choice of kernel is
\[ Q_i(w) = 1_{V_i}(w), \]
for which \( \langle Q_i \rangle \) measures the mass inside a volume \( V_i \). For many problems one might choose some of the \( V_i \) to surround important resonances in phase-space, so that \( \langle Q_i \rangle \) measures the phase-space density around the resonances. With appropriate choices of projection kernel \( Q_i \), the expression (3.1) includes quantities such as the galaxy’s density profile, its velocity moments or even its projected line-of-sight velocity distributions.

More fundamentally, an N-body model should provide a good estimate of the galaxy’s acceleration field. Therefore we recommend that many of the \( \langle Q_i \rangle \) be used to measure at least the monopole component of the galaxy’s acceleration field at a range of points. This can be achieved using \( 1_V \) with spherical volumes \( V_i \) centred on \( x = 0 \) for a range of radii \( r_i \), encompassing all velocity space for \(|x| < r_i\). Similarly, one can include higher-order multipole components of the galaxy’s acceleration field by choosing a slightly more complicated projection kernel \( Q_i \) (see equation 3.23 below).
3.2.2 Optimal sampling scheme

The problem we address in this work is the following. We wish to construct an equilibrium \(N\)-body realization of a galaxy model with some known DF \(f_0\). Specifically, we seek ICs that faithfully represent some projections,

\[
\langle Q_i \rangle = \int f_0 Q_i \, d^6 w,
\]  

(3.4)

of this DF, for a set of \(n_Q\) kernels \(Q_i(w)\). What is the “best” choice of sampling DF \(f_s\) given this \(f_0\) and choice of kernels \(Q_i\)?

More formally, from (2.36) the uncertainty in a Monte Carlo estimate of \(\langle Q_i \rangle\) obtained using \(N\) particles drawn from the sampling distribution \(f_s\) is given by

\[
\text{Var} \langle Q_i \rangle = \frac{1}{N} \left[ \int f_0^2 Q_i^2 \, d^6 w - \langle Q_i \rangle^2 \right].
\]  

(3.5)

Notice that, unlike most introductory textbook examples of Monte Carlo methods, we have \(n_Q\) such estimates but just one \(f_s\). We seek a normalized sampling DF \(f_s\) that minimizes the mean-square fractional uncertainty

\[
S = \sum_{i=1}^{n_Q} (\delta Q_i)^2
\]  

(3.6)

where \(\delta Q_i\), the formal relative uncertainty in a measurement of \(Q_i\), is given by

\[
(\delta Q_i)^2 = \frac{\text{Var} \langle Q_i \rangle}{\langle Q_i \rangle^2}.
\]  

(3.7)

Of course there are many other possible measures of the “goodness” of some choice of \(f_s\).

One can immediately use the Euler–Lagrange equation to show that choosing

\[
f_s^2(w) \propto f_0^2(w) \sum_{i=1}^{n_Q} \frac{Q_i^2(w)}{\langle Q_i \rangle^2}
\]  

(3.8)

extremizes (3.6), the proportionality constant being set by the constraint that \(f_s\) should be normalized, \(\int f_s \, d^6 w = 1\). This direct solution is flawed, however, since for most interesting choices of \(Q_i\)}
the resulting $f_s$ depends on orbit phase; using this $f_s$ the masses of particles sampling a given orbit would vary along the orbit! Therefore in practice we use a slightly less direct approach.

We partition phase space into $n_f$ cells and write $\tau_j$ for the phase-space volume enclosed by the $j$th cell $^2$. We parametrize $f_s$ as

$$f_s(w) = \sum_{j=1}^{n_f} \frac{\tau_j}{a_j} f_0(w),$$

(3.9)

so that within the $j$th phase-cell $f_s$ is given by $f_0(w)/a_j$. For the equilibrium models considered, it is natural to choose $\tau_j$ to be cells in integral space. Substituting this $f_s$ into (3.7) yields

$$(\delta Q_i)^2 = \frac{1}{N} \left[ \sum_{j=1}^{n_f} a_j H_{ij} - 1 \right],$$

(3.10)

where

$$H_{ij} = \frac{\int_{\tau_j} f_0 Q_i^2 d^6 w}{\langle Q_i \rangle^2}.$$  

(3.11)

If we further define

$$H_j \equiv \sum_{i=1}^{n_Q} H_{ij},$$

(3.12)

then the mean-square fractional uncertainty (3.6) becomes

$$S = \frac{1}{N} \left[ \sum_{j=1}^{n_f} a_j H_j - n_Q \right].$$

(3.13)

Our goal is to find the coefficients $a_j$ that minimize this $S$, subject to the constraint that the resulting $f_s$ be normalized. The normalization constraint is that

$$\int f_s d^6 w = \sum_{j=1}^{n_f} \frac{I_j}{a_j} = 1,$$

(3.14)

$^2$Note that we use $V$ to denote subvolumes of phase-space used in calculating the projections (3.1) of the DF $f_0$, and $\tau$ for the subvolumes used in the discretization of the sampling DF $f_s$. 

---

*Chapter 3: Multi-mass schemes for collisionless simulations*
where
\[ I_j = \int_{\tau_j} f_0 \, d^6w. \] (3.15)

Using the method of Lagrange multipliers, the coefficients of the “best” sampling DF obtained by minimizing (3.13) subject to the constraint (3.14) are simply
\[ a_j = \sqrt{\frac{I_j}{H_j}} \sum_{k=1}^{n_f} \sqrt{I_k H_k}. \] (3.16)

which is just the direct solution (3.8) in disguise, but averaged over the phase-space cells \( \tau_j \) and correctly normalized. This averaging means that the resulting \( f_s \) will be smooth, provided that none of the kernels \( Q_i \) pick out specific regions of integral space.

Substituting the \( f_s \) given by (3.9) into (2.43), we have that particles in phase-space cell \( \tau_j \) have masses \( m_j = a_j/N \). One can therefore easily impose additional, direct constraints on the masses of particles within a subset of the phase-space cells \( \tau_j \); simply repeat the minimisation of (3.13) subject to (3.14) while holding the relevant subset of the \( a_j \) fixed at the desired values. For example, when generating an \( N \)-body realization of a dark-matter halo model inside which one intends to embed a disk of light particles, one might want to ensure that those halo particles passing through the disk have the same mass as the disk particles. Of course, a more pedestrian approach would be to introduce additional kernels \( Q_i \) to pick out the relevant parts of integral space. We caution, however, that we have not tested how well such a “bumpy” \( f_s \) would work in practice; the tests we present later all involve smoothly varying sampling distributions.

### 3.2.3 ICs for \( N \)-body model

Together with \( f_0 \), the coefficients \( a_j \) completely determine the sampling DF \( f_s \) of the form (3.9). In particular, it reduces to the conventional equal-mass case when all \( a_j = 1 \).

We apply the following sequence of steps \( N \) times to draw particles from this \( f_s \), thereby constructing an \( N \)-body realization of the galaxy model:

1. Choose one of the \( n_f \) cells at random, the probability of choosing the \( j^{th} \) cell being given by \( I_j/a_j \). Let \( i \) be the index of the chosen cell.
(2). Assign a mass \( m_i \equiv f_0(w_i)/N f_s(w_i) = a_i/N \) to the particle.

(3). Within the \( i^{th} \) cell, draw \( x_i \) from its density distribution, \( \rho_i = \int f_0 \delta_0, d^6 w \). For the special case of a spherical galaxy, one can precompute the cumulative mass distribution \( M_i(r) \) for each of the \( n_f \) cells and use this to draw a radius \( r_i \), followed by angles \( \theta_i \) and \( \phi_i \).

(4). Use an acceptance-rejection method to draw \( v_i \) from \( f_0(x_i, v) \) at this fixed value of \( x_i \).

### 3.3 Comparisons of a bulge model (no MBH)

In this section we use a simple bulge model to demonstrate our scheme. Our bulge model is spherical and isotropic, with density profile (Hernquist 1990)

\[
\rho(r) = \frac{M_\star}{2\pi r(r + r_\star)^3},
\]

(3.17)

total mass \( M_\star \) and scale radius \( r_\star \). By Jeans’ theorem, the model DF \( f(x, v) \) depends on \( (x, v) \) only through the binding energy per unit mass \( \mathcal{E} \). Hernquist (1990) gives an expression for \( f(\mathcal{E}) \).

We want to construct an \( N \)-body realization of this model that minimizes the mean-square error in the monopole component of the acceleration averaged over many decades in radius, from \( r_{\text{min}} = 10^{-4} r_\star \) up to \( r_{\text{max}} = 10^2 r_\star \). To achieve this we choose kernels \( Q_i = \mathbf{1}_V(r) \) that measure the mass enclosed within a sequence of 25 spheres centred on the origin, with radii \( r_i \) spaced logarithmically between \( r_{\text{min}} \) and \( r_{\text{max}} \). We use \( (3.10) \) to calculate the formal uncertainty \( \delta M_i \) in the enclosed mass for a range of discretized sampling densities of the form \( (3.9) \), including \( (3.16) \).

To implement this, we first of all partition integral space \( (\mathcal{E}, J^2) \) onto a regular \( n_f = n_\mathcal{E} \times n_X \) grid. The \( n_\mathcal{E} \) energies \( \mathcal{E}_j \) are chosen to match the potential \( \mathcal{E}_j = \psi(r_j) \) with \( r_j \) logarithmically spaced between \( 10^{-6} r_\star \) and \( 10^3 r_\star \). At each \( \mathcal{E}_j \), there are \( n_X \) values of \( X_{jk} \) running linearly from 0 to 1, where \( X_{jk} = J_k(\mathcal{E}_j)/J_c(\mathcal{E}_j) \) is the orbital angular momentum normalized by the circular angular momentum at energy \( \mathcal{E}_j \). These choices ensure that our \( f_s \) samples well the interesting parts of phase-space. For the calculations below we take \( n_\mathcal{E} \times n_X = 200 \times 100 \), although a coarser grid (e.g.,

\[ X = J/J_c \] is a measure of an orbit’s circularity; orbits with \( X = 1 \) are perfectly circular, while those with \( X = 0 \) are perfectly radial.
3.3.1 Formal errors

Before applying our method, we study two other schemes: the conventional equal-mass scheme and the multi-mass scheme of SHQ95.

I. The conventional equal-mass scheme

The most common (albeit implicit) choice of sampling density is \( f_s = f_0 \), which corresponds to setting all \( a_j = 1 \) in our equation (3.9). All particles then have the same \( 1/N \) mass. For our example Hernquist model the fraction of particles within radius \( r \) is \( r^2/(r + r_*)^2 \), so that less than 1% of the particles are within \( 0.1r_* \). As shown in figure 3.1 for this \( f_s \) the formal uncertainty...
$\delta M(r)$ rises steeply towards the centre. Although this scheme produces accurate estimates of the bulge’s potential outside the scale radius $r_*$, it performs poorly in the interesting $r^{-1}$ central density cusp.

II. Sigurdsson et al.'s multi-mass scheme

SHQ95 have used an interesting heuristic scheme to improve the resolution of $N$-body models near bulge centre. In effect, they use an anisotropic sampling function of the form

$$a(\tau) \equiv B \times \begin{cases} \left( \frac{r_{\text{peri}}(\tau)}{r_*} \right)^\lambda & \text{if } r_{\text{peri}} < r_*, \\ 1 & \text{otherwise}, \end{cases}$$

(3.18)

where $r_{\text{peri}}(\tau)$ is the smallest pericentre radius of any orbit from the phase-space cell $\tau$, and the constant $B$ is chosen to normalize $f_s$. When the parameter $\lambda = 0$, then $a_j = 1$ and the sampling DF $f_s$ is identical to $f_0$. Increasing $\lambda$ improves the sampling of the cusp by increasing the number density of particles having pericentres $r_{\text{peri}} < r_*$. Consequently, as $r \to 0$ the number density of particles rises more rapidly than the mass density, permitting better resolution in the centre. To balance this increase in number density, each particle is assigned a mass $f_0/N f_s = a_j/N$ so that the phase-space mass density is still given by the desired $f_0$.

The dashed curve in figure 3.1 shows the formal error $\delta M(r)$ in our implementation of their scheme for $\lambda = 1$. Their scheme does much better than the conventional equal-mass scheme at small radii $r \ll r_*$, at the cost of a slightly noisier monopole at $r \gtrsim r_*$.  

III. Our scheme

It is encouraging to see that SHQ95’s multi-mass scheme does, to some extent, improve mass resolution at small radius. However, as shown in figure 3.1, $\delta M$ at $r = 10^{-4} r_*$ is still almost two orders of magnitude larger than at $r = r_*$. Can we achieve even better results by carefully designing an $f_s$ that generates a flat $\delta M(r)$ across a large range of radii?

The $f_s$ given by our optimal choice of coefficients is plotted in figure 3.2. It is qualitatively similar to SHQ95’s results, in the fact that it samples densely the low-angular momen-
Chapter 3: Multi-mass schemes for collisionless simulations

Figure 3.2: Contour map of $\log_{10}(f_s/f_0)$ for the optimal multi-mass sampling scheme of section 3.3.1 notice the strong enhancement in low angular momentum, $X^2(\mathcal{E}) = J^2/J_c^2(\mathcal{E})$ and large energy (small $r_\mathcal{E}$) region. These correspond to orbits with small peri-centre radii.

The momentum parts of phase-space. The detailed shape of the function is different, however, and the thick solid curve in figure 3.1 shows that our scheme provides much better estimates of the monopole components of the acceleration at small radii; in fact, the formal error $\delta M(r)$ varies by only a factor $\sim 4$ over six decades in radius.

$N$-body realizations

Figure 3.3 shows the spectrum of masses obtained using the algorithm detailed in §3.2.3 to draw $N = 10^6$ particles from this optimal $f_s$. Unlike the conventional scheme which would give all particles the same $10^4M_\odot$ mass if we assume the Hernquist galaxy has a total mass $M_* = 10^{10}M_\odot$, our multi-mass scheme assigns a range of masses between $10^{-2}M_\odot$ and $10^6M_\odot$ (8 decades), with many low-mass particles in the central region.

As a simple sanity check of our formal estimates of the errors in the monopole, we count the mass of particles within the same spheres $V_i$ used to calculate $\delta M_i$. The deviations from the mass profile of the target Hernquist model are consistent (figure 3.4) with the expected values of $\delta M$ from equation 3.10.
Figure 3.3: Histogram of particle masses from an $N = 10^6$ multi-mass realization of a Hernquist bulge, scaled to a total mass $M_\star = 10^{10}$ $M_\odot$. The span of 8 decades in mass gives sub-solar mass resolution in “interesting” regions of phase space. In contrast, in an equal-mass realization all particles would have mass $10^4$ $M_\odot$ (thick solid line).

Figure 3.4: RMS fractional deviations in acceleration (solid line), in mass (circle-solid line) together with its analytical value $\delta M$ in dashed line; left panel for equal-mass model and right panel for multi-mass model. For this single acceleration calculation, we include 16 levels of refinements and therefore all the values should be believable outside $\epsilon_{\text{min}} \approx 10^{-4}$.
Ultimately, the purpose of our sampling scheme is to improve the numerical modelling of collisionless galaxies close to equilibrium using full $N$-body integrations. To test how well our scheme succeeds at this task, we use the particle-multiple-mesh code GROMMET (Magorrian 2007) to compare the evolution of our multi-mass models against equal-mass ones. Below, we adopt $N$-body units $G = 4.31, M = a = 1$.

3.3.2 How well is the acceleration field reproduced?

As already mentioned in Chapter 3.1 in the present work we do not investigate how different softening lengths or softening kernels affect our multi-mass models. We simply adopt a nested series of boxes with boundaries at $|x| = 100 \times 2^{-i}$ with $i = 0, \ldots, 16$, each box covered by a $60^3$ mesh. As one moves to smaller length scales the effective softening length decreases, with $\epsilon_{\text{min}} = 200/60 \times 2^{-16} \approx 10^{-4}$.

For the multi-mass ICs, figure 3.5 shows the RMS fractional deviations in the radial component of the acceleration field (solid curve). This can be quantified by considering a test star (at a distance $r$ from the centre) randomly scattered by its close neighbours. Here, we calculate the averaged acceleration felt by one particle due to the presence of its neighbours. We bin particles into shells centred on the origin to get a mean mass $\overline{m_*}(r)$ and a mean number density $\overline{n_*}(r)$ around the particle; the mean distance between two particles is approximated as $d_*(r) \approx \left[\overline{n_*}(r)\right]^{-1/3}$, and the mean acceleration becomes

$$a(r) = G \frac{\overline{m_*}}{(d_*)^2} = G \frac{\overline{m_*}}{\left(\overline{n_*}\right)^{2/3}}$$

(3.19)

Comparing to the true value, we get the fractional errors in the mean acceleration, and show them in figure 3.5. In the central region, the RMS fractional deviations is consistent with the noise estimated from neighbour stars. In the outer region where the enclosed mass is already large, the acceleration error becomes less sensitive to noise from a much more distant neighbour particle.

Figure 3.4 compares the fractional error in the enclosed mass (solid lines) and the one in the radial acceleration field for both equal- and multi-mass realizations. Between them, there is an
Figure 3.5: For the multi-mass ICs, RMS fractional deviations in the radial acceleration (solid curve) and the relative error in the mean inter-particle acceleration (eq. 3.19 in dashed curve). They are consistent in the central region, where the enclosed mass is small and the error is very sensitive to noise from a single neighbour particle.

approximately constant offset. Since our multi-mass ICs have been tailored to minimize the variance in the radial mass distribution, hence the variance in the monopole component of the acceleration field, the detailed arrangement of particles within a given sphere must have effect on the overall acceleration field. So, how important is our neglect of the higher-order multipoles in the mass distribution?

In terms of multipole moments, the radial component of the acceleration is (e.g., BT87)

\[
a_r(r, \theta, \phi) = 4\pi G \sum_{lm} \frac{Y_l^m(r)}{2l + 1} \times \left[ -\frac{l + 1}{r^{l+2}} \int_0^r \rho_{lm}(r') r'^{l+2} dr' + \int_r^\infty \rho_{lm}(r') \frac{dr'}{r'^{l-1}} \right]
\]

where

\[
\rho_{lm}(r) = \int_0^{2\pi} d\phi \int_0^\pi \sin \theta Y_l^m*(\theta, \phi) \rho(r, \theta, \phi).
\] (3.20)
This can be rewritten as

\[ a_r(r, \theta, \phi) = \frac{4\pi G}{r^2} \sum_{lm} \langle M_{lm}(r) \rangle Y_l^m(\theta, \phi), \]  

(3.21)

where \( \langle M_{lm}(r) \rangle \) are given by

\[ \langle M_{lm}(r) \rangle = \int f_0(w') M_{lm}(r, w') \, d^6w' \]  

(3.22)

with projection kernels

\[ M_{lm}(r, w') = Y_l^m(\theta', \phi') \left[ \frac{l+1}{2l+1} \frac{r^l}{r} V(r) \langle w' \rangle - \frac{l}{2l+1} \frac{r^{l+1}}{r} V(r) \langle w' \rangle \right], \]  

(3.23)

where \( V(r) \) encompasses all phase-space points having radii less than the (real-space) radius \( r \), and \( V^c(r) \) is its complement. For our spherical bulge,

\[ \langle M_{lm}(r) \rangle = \int f_0(w') M_{lm}(r, w') \, d^6w' = \begin{cases} M(r), & \text{if } l = m = 0 \\ 0, & \text{otherwise.} \end{cases} \]  

(3.24)

The corresponding variance in \( a_r(r) \) for an \( N \)-body realization drawn from some choice of \( f_s \) is

\[ \text{Var} \langle a_r(r) \rangle = \frac{4\pi G}{r^2} \sum_{lm} \text{Var} \langle M_{lm}(r) \rangle Y_l^m(\theta, \phi), \]  

(3.25)

where, from (3.5),

\[ \text{Var} \langle M_{lm}(r) \rangle = \frac{1}{N} \left[ \int \frac{f_0^2(w')}{f_s(w')} M_{lm}(r, w')^2 \, d^6w' - \langle M_{lm}(r) \rangle^2 \right]. \]  

(3.26)

Similarly, the variance in the tangential component of acceleration field can be achieved by using projection kernels

\[ \text{Var} \langle a_{\theta,\phi}(r) \rangle = \frac{4\pi G}{r^2} \sum_{lm} \text{Var} \langle M_{lm}^t(r) \rangle \left| Y_l^m(\theta, \phi) \right|^2_{\theta,\phi}, \]  

(3.27)
where

\[
M_{lm}^t(r, \mathbf{w}') = Y_{lm}^m(\theta', \phi') \left[ \frac{1}{2l + 1} \frac{r^l}{r} \mathbf{1}_V(r)(\mathbf{w}') + \frac{1}{2l + 1} \frac{r^{l+1}}{r} \mathbf{1}_V(r)(\mathbf{w}') \right].
\]

(3.28)

So, given any choice of \(f_s\), we can use the expressions above to calculate the contribution of the higher-order multipole moments to the formal errors in the acceleration. We find that, as we progressively include more terms, our estimate of the formal variance \(\text{Var}(a_r(r))\) approaches the actual errors observed in the \(N\)-body realization.

Alternatively, we can find our optimal sampling DF \(f_s\) by minimizing

\[
S \equiv \sum_{i=1}^{n_Q} \sum_{l=0}^{l_{\text{max}}} \sum_{m=-l}^{l} \frac{\text{Var}(M_{lm}^t)}{(M_i)^2},
\]

(3.29)

truncated at say \(l_{\text{max}} = 2\). Notice that this new \(S\) reduces to the old one in eq. (3.6) when \(l_{\text{max}} = 0\), but otherwise includes additional terms with \(l > 0\), each weighted by the monopole component \(l = 0\). On increasing \(l_{\text{max}}\) from 0 to 2, the formal variance \(\text{Var}(a_r(r))\) increases but the shape of the curve remains approximately unchanged and there are no noticeable differences in the resulting \(f_s\). Therefore, our neglect of higher-order multipole moments is justified, at least in the present case, provided one bears in mind that the errors in the full acceleration field are going to be larger by an approximately constant factor than what one would estimate from the monopole component alone. The variance in mass within one grommet cell is constantly larger than that in the monopole component.

### 3.3.3 How well are integrals of motion conserved?

This paragraph describes the details of a full \(N\)-body implementation. Using both equal and multi-mass schemes, we draw \(10^6\) particles with radii between \(10^{-3} < r < 10^2\). In order to suppress slight deviations from symmetry (the odd terms of higher-order multipoles) and remove any intrinsic transient in linear momentum (see also McMillan & Dehnen 2005), ICs \((\mathbf{x}, \mathbf{v})\) are extended to include the mirror distribution by reflecting each of the \(10^6\) particles with \((\mathbf{x}, \mathbf{v}) \rightarrow (-\mathbf{x}, -\mathbf{v})\). The full ICs then have \(N = 2 \times 10^6\) particles. Taking the efficiency of integration into consideration, only a 12-level nested series of boxes each covered by a 60³ mesh is used, together with a single
time-step of $2 \times 10^{-4}$. Therefore, we expect our numerical results to be trustworthy at radii greater than a few times $10^{-3}$.

Figure 3.6 plots the inner density profiles of both realizations after evolving each for 200 time-units (or 300 circular orbit periods at $r = 0.01$). The lack of particles at small radius $r \sim 10^{-2}$ in the equal-mass realization means that the initial model is out of exactly-detailed equilibrium and causes the central density profiles to flatten. In contrast, the density profile in the multi-mass case is always much better behaved there.

It is interesting to examine what is going on at the level of individual orbits. Both realizations begin with spherical symmetry and remain spherical, apart from the effects of Poisson noise. The amount of diffusion in the angular momentum $J$ of each particle’s orbit serves as a strong gauge of relaxation effects. This is complicated by the fact that many particles in isotropic models being considered here have $J(t = 0) \simeq 0$. In such cases, even a small change in $J(t > 0)$ would yield a large fractional change when measured in respect to its initial value. To circumvent this artificial problem, for each particle we measure the change in angular momenta relative to its circular value.
at $t = 0$ using

$$\Delta X_i^2 = \left[ \frac{J_i(t) - J_i(0)}{J_i(\mathcal{E}_i)} \right]^2 \frac{1}{t}.$$  \hspace{1cm} (3.30)

Binning particles by energy and calculating the mean $\Delta X_i^2$ within each energy bin gives us the time-averaged diffusion rate $\delta X^2(\mathcal{E})$. As shown on the left panel of figure 3.7, both models suffer diffusion, but due to the enhancement of particle numbers and hence the smoothness of potential field in the central region, diffusion in the multi-mass scheme is suppressed by two orders of magnitude across the whole system.

As a further test of the robustness of our multi-mass scheme, we have evolved our multi-mass ICs using the tree code FALCON (Dehnen 2000) with a single interparticle softening radius of $10^{-3}$, comparable to the finest mesh size used in the GROMMET runs. The dashed curve on the left panel of figure 3.7 plots the resulting $\delta X^2$; our scheme works just as well for tree codes as it does for mesh codes, although the variable softening in GROMMET does slightly decrease the amount of diffusion. This is not surprising, since the only difference between the two runs is the approximations used to estimate the accelerations.

In any model with a broad spectrum of particle masses, a natural question is what happens if heavy bodies from the outskirts visit the centre full of light mass elements and vice versa. To address this issue, we have measured the $\delta X^2(\mathcal{E})$ of equation (3.30) but, instead of considering all particles of a given $\mathcal{E}$, we compare the diffusion of particles on radially-biased orbits with $X^2 < 0.1$ to those on nearly-circular orbits with $X^2 > 0.9$. As shown on the right panel of figure 3.7, there are no systematic differences between them. The reason for this is simply that particles with $X \approx 0$ spend most of their time at apocentre, the apocentre radius being only a factor $\sim 2$ larger than the radius of a circular orbit of the same energy. Nevertheless, a particle with $X \approx 0$ will affect all of the more tightly bound orbits as it plunges through the centre of the bulge, but our measured diffusion rates account for this.
Chapter 3: Multi-mass schemes for collisionless simulations

Figure 3.7: Left panel: Time-average diffusion rate $\delta X^2$ (eq. 3.30) against energy labelled radius $r_\text{E}$ measured between $t = 0$ and $t = 200$ for a multi-mass realization evolved using GROMMET (thick curve), falcON (dashed) and for an equal-mass realization evolved using GROMMET (thin curve). Right panel: $\delta X^2$ in multi-mass simulations for particles with $X^2 < 0.1$ (dashed curve) and $X^2 > 0.9$ (solid).

3.4 Simulations of a bulge + MBH model

In this section, our treatment of self-consistent galaxy models is extended to models containing a central MBH. We use the reference model introduced on page 17. The density-potential pair has the form:

$$
\rho(r) = \frac{M_*}{2\pi r(r + r_*)^3}, \quad \psi(r) = \frac{GM_*}{(r + r_*)} + \frac{GM_*}{\sqrt{r^2 + \epsilon_*^2}},
$$

(3.31)

where stars have total mass $M_* = 1$ and scale radius $r_* = 1$, the MBH has mass $M_\bullet = 0.01$ and scale radius $\epsilon_\bullet = 0.003$.

3.4.1 N-body realizations

To model galaxies containing MBHs is numerically challenging: (1) it requires a large number of equal-mass particles to make statistically significant assertions about the central regions.
of a realistic bulge model; (2) it requires a tiny time-step to track accurately the evolution of stars near the centre, despite the large range in intrinsic timescales across the whole bulge. Below, we explain how to construct multi-mass ICs and how to move particles.

I. ICs

This is a straight-forward generalization of section 3.3 to a bulge + MBH model. Find the isotropic DF $f_0$ which reproduces the density in the combined potential $\psi(r)$ (section 2.1.3; also see Appendix A.1).

Figure 3.8 shows the phase-space number density. The bulge + MBH model (right panel) much like the previous bulge model (left panel), samples densely the high-energy and low-angular-momentum part of phase-space. For example, in phase-space 48% particles are with $J^2 \leq 10^{-3}$. This is in contrast with less than 0.1% in the equal-mass case (left panel) and we gain a 480-fold boost in phase-space resolution. Similarly, 62% of particles are inside the MBH’s sphere of influence $\sim 0.16(160 \text{ kpc})$ compared to 2%. The contrast between the mass of the MBH and the mean mass of surrounding particles within a sphere of radius $r = 0.01(10 \text{ pc})$, is $10^7$ compared to $10^4$; see table 3.1.
Chapter 3: Multi-mass schemes for collisionless simulations

Figure 3.9: The sequence of steps for motion in the Hamiltonian $K^D$ with two levels of timestep refinement. For any given timestep level $l$, the $K$ operation “kicks” particles inside any boxes having that timestep level, applying to each an impulse $\frac{1}{2}\tau_l \cdot (-\partial V_l(x)/\partial x)$, where the timestep $\tau_l = 2^{-l}\tau_0$. These impulses change the particles’ velocities, but not their positions. They conserve the particles’ total linear momentum. The $D$ operation “drifts” all particles for a time $\frac{1}{2}\tau_l$, changing their positions but not their velocities (adapted from Magorrian 2007).

for a summary. Better resolution in the low-angular-momentum part of phase-space and better space/mass resolution within the MBH’s sphere of influence simply suggest that this multi-mass model will be valuable for the study of dynamics around the MBH.

<table>
<thead>
<tr>
<th>Equal-mass</th>
<th>Multi-mass</th>
<th>Boost factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N(r \leq r_{\text{infl}})$</td>
<td>0.020</td>
<td>0.62</td>
</tr>
<tr>
<td>$N(J^2 \leq 10^{-3})$</td>
<td>0.0009</td>
<td>0.48</td>
</tr>
<tr>
<td>$M_\bullet/\bar{m}_*$†</td>
<td>$2.0 \times 10^4$</td>
<td>$2.6 \times 10^7$</td>
</tr>
</tbody>
</table>

† $\bar{m}_*$ is the mean mass of particles within a central sphere $r < 0.01(10 \text{ pc})$.

II. Block leapfrog time-stepping

The size of a time-step is a compromise between speed and physics: the former argues for a large value while the latter requires the size of a time-step should be many times smaller than the local dynamical time scale. Figure 3.10 shows the circular orbital period

$$t_{\text{cir}}(r) = \frac{2\pi r}{v_{\text{cir}}} = 2\pi \sqrt{\frac{r^3}{GM(r)}},$$

(3.32)

for the MBH (thin) and bulge (thick) component. The range of the local dynamical time, from $10^{-2} - 10^4$, is huge. In order to deal with inhomogeneous galaxies, GROMMET employs an efficient block-time-step leapfrog.
GROMMET employs a series of nested boxes with different spatial refinement levels. The outermost box is associated with a timestep $\tau_0$ and timestep level $l = 0$. Each subbox has a timestep $\tau_l = 2^{-l} \tau_0$ with timestep level $l$ either equal to that of its parent or larger by one. Let $V_{(l)}(\mathbf{x})$ be the contribution to the potential energy from boxes having timestep level $l$. Recall the time-dependent Hamiltonian expression in eq. (2.49) and let us consider the alternative Hamiltonian

$$H = T + \sum_{l=0}^{l_{\text{max}}} \sum_{k=-\infty}^{\infty} \delta_e \left( k - \frac{t}{2^{-l} \tau_0} \right) V_{(l)}(\mathbf{x}_1, \ldots, \mathbf{x}_N), \quad (3.33)$$

where $l_{\text{max}}$ is the maximum time-step refinement level. Instead of turning on the full potential $V = \sum_{l} V_{(l)}$ at every time-step, the split potential $V_{(l)}$ is turned on only at $t = 2^{-l} k \tau_0$.

Integrating the equations of motion for this new Hamiltonian results in a nested sequence of KDDK leapfrog steps, as shown in figure 3.9 (Magorrian 2007).

The essence is that when a star approaches the centre and crosses the boundary of subboxes with different spatial refinement levels, the time-step of the integration either decreases or stays the same, thereby allowing better resolution of the pericentre passage. We use the same nested series of boxes as in section 3.3.2 with boundaries at $|\mathbf{x}| = 100 \times 2^{-i}$ for $i = 0, \ldots, 12$ each box covered by a $60^3$ mesh. We implement the block time-step integrator with six levels of refinement which starts from a time-step $6.4 \times 10^{-3}$ for particles with $|\mathbf{x}| > 100 \times 2^{-7} \approx 0.78$. The time-step halves at each subsequent boundary of boxes $i = 7 \ldots 12$, so that the innermost ($|\mathbf{x}| < 0.024$) box has a time-step $1 \times 10^{-5}$.

Working with GROMMET and the block-time-step algorithm, a safe and formal way of including the effects of an external potential is to add an extra term

$$\sum_{k=-\infty}^{\infty} \delta_e \left( k - \frac{t}{2^{-l_{\text{max}}} \tau_0} \right) V_{\text{ext}}(\mathbf{x}_i, t), \quad (3.34)$$

to the Hamiltonian. This ensures that the perturbation is turned on at the appropriate
Figure 3.10: Local circular orbital period $t_{\text{cir}}(r)$: a Hernquist model with $(M_*, r_*) = (1, 1)$ is given by the thick curve and a plummer model with $(M_*, \epsilon_*) = (10^{-2}, 3 \times 10^{-3})$ is given by the thin curve. The desired time-step should resolve the lower value of two.

For the MBH with a Plummer’s profile,

$$V_{\text{ext}}(x_i, t) = \sum_{i=1}^{N} \frac{GM_\bullet m_i}{(|x_i - x_\bullet|^2 + \epsilon_\bullet^2)^{1/2}},$$

(3.35)

where subscripts $\bullet$ and $i$ denote properties of the MBH and stars respectively. Note the force on one particle is balanced by an equal and opposite force on another particle, the system does conserve linear momentum. The MBH’s acceleration

$$a_\bullet = -\frac{\partial V_{\text{ext}}}{M_\bullet \partial x_\bullet} = \sum_{i=1}^{N} \frac{Gm_i (x_i - x_\bullet)}{M_\bullet (|x_i - x_\bullet|^2 + \epsilon_\bullet^2)^{3/2}},$$

(3.36)

is updated at time $t = k2^{-i_{\text{max}}\tau_0}$. The mesh is recentred on the MBH after every coarse block-time-step $t = k\tau_0$.

The PM force calculation and the integration scheme explicitly conserve linear momentum. They are suitable for use in studying $l = 1$ perturbations without contamination of numerical
artifacts due to centring (Weinberg 1994; Magorrian 2007).

### 3.4.2 Results: diffusion in \( J^2 \)

Figure 3.11 shows the density, velocity and anisotropy profiles, \( \beta = 1 - \sigma_k^2/2\sigma_r^2 \), both in the ICs and after evolving the system for 100 time units. They are consistent to within a few smoothing lengths of the centre. To gauge the relaxation effects, we use eq. (3.30) to calculate the time-averaged diffusion rate \( \delta X^2(\mathcal{E}) \) and show the results with thin curve in figure 3.12. We also compare the diffusion of particles on radially-biased orbits with \( X^2 < 0.1 \) to those on nearly-circular orbits with \( X^2 > 0.9 \). As shown in figure 3.12 there are no systematic differences between them.

It would be interesting to compare results between the bulge model (Section 3.3) and bulge + MBH model (here in Section 3.4), since they realize the same mass distribution but with very different dynamics. Figure 3.12 shows that \( \delta X^2 \), in the central region, the presence of a MBH would provide an extra restoring force on the core stars and effectively improve the particles’ force resolution there. As more light-mass particles are distributed inside \( r_{\text{infl}} \), the relative small number of particles in the outskirt gives rise to a larger diffusion \( \delta X^2 \) there and explains the drop of \( \delta X^2 \) within the MBH’s sphere of influence, \( r \leq r_{\text{infl}} \approx 0.16(160 \text{ pc}) \). Due to the softening of the MBH at \( \epsilon \approx 0.003 \), \( \delta X^2 \) converges to the bulge case there.
Figure 3.11: Mass density (top left), velocity dispersion of radial (top right) and tangential component (bottom right) and velocity anisotropy profiles (bottom left) of the unperturbed MBH-Hernquist model. The solid curves show the ICs and the dash curves show the profiles after 100 time units (or 3000 circular orbit periods at $r = 0.01$, 50 at $r = 1$). They have not changed by any significant amount, suggesting that the ICs are already close to equilibrium.
Chapter 3: Multi-mass schemes for collisionless simulations

Figure 3.12: Time-average diffusion rate $\delta X^2$ (eq. 3.30) against energy labelled radius $r_E$ measured between $t = 0$ and $t = 100$ evolved using GROMMET for multi-mass realizations. Left panel compares values between a bulge model (thick curve) and a bulge + MBH model (thin curve). Right panel compares multi-mass particles with $X^2 < 0.1$ (dashed curve) and $X^2 > 0.9$ (solid) in the MBH case.
In the last chapter, we have developed a general scheme for constructing multi-mass collisionless galaxy models. It has been successfully applied to set up a bulge + MBH model, whose long-time behavior has also been verified by full $N$-body experiments. In this chapter, the MBH-embedded model is extended to include a loss cone (LC): a consumption sphere centred on the MBH. First of all, control tests have been directed to calibrate (artificial) noise-driven flux into the LC. Then starting from this near-equilibrium galaxy, satellites on flyby orbits are added as transient perturbers. For satellites with different mass and orbital parameters, I carry out extensive experiments to measure the mass of stars captured into the LC after a single flyby encounter. Given a set of satellite parameters, an empirical relationship is found to predict the LC refilling mass. Finally, to understand the interactions between the perturber and stars, especially to disentangle effects of resonance coupling from velocity impulse, I run extra test-particle experiments of orbiting satellites.
4.1 Loss cone refilling by two-body relaxation

What is a loss cone?

In Sec. 1.2.1, we knew that stars venturing into a LC sphere of effective ( tidal LC) radius will be consumed by the MBH. There, the geometrical LC is characterized by its effective radius, which is on the order of tidal radius \( r_{lc} \approx r_t \) (eq 1.3). In a system where the steady-state gravitational potential from the MBH and the surrounding stars is spherical \( \psi(r) = -\Phi(r) \), orbits with pericentre distance \( r_{peri} < r_{lc} \) occupy a specific region in energy-angular-momentum space. The loss cone (LC) can be redefined through the orbital binding energy per unit mass \( E = \psi(r) - v^2/2 \), and angular-momentum per unit mass \( J = |x \times v| \) of a star as

\[
J^2 \leq J^2_{lc}(E) = 2[\psi(r_{lc}) - E] r_{lc}^2 \approx 2GM_\bullet r_{lc} \quad (E \ll GM_\bullet/r_{lc}),
\]

(4.1)

the second approximation comes from the fact that most stars are consumed from nearly radial orbits, \( E \approx -GM_\bullet/r \). The LC boundary is almost independent of \( E \).

How many stars are inside a full a loss cone?

One may approximate the radial periods of almost radial LC orbits as \( T_r(E) \equiv T_r(E, 0) \). So for a spherical galaxy with isotropic DF \( f(E) \), the fraction of stars within a full LC per unit energy interval is (cf. eq. 2.27)

\[
N_{lc}(E) dE \approx 4\pi^2 f(E) J^2_{lc}(E) T_r(E) dE.
\]

(4.2)

For a \( 10^8 M_\odot \) MBH embedded in a \( 10^{10} M_\odot \) Hernquist galaxy, a full LC of effective radius \( r_{lc} = 10^{-5} \) pc contains about \( 10 M_\odot \) mass, if all stars are of solar mass and radius.
What is the size of a loss cone?

To measure an orbit’s circularity, I use $X^2 = J^2 / J^2_c$, the ratio between the orbital angular-momentum to the circular angular-momentum at $\mathcal{E}$. Then, the relative size of the LC can be described by

$$X^2_{lc} = \frac{J^2_{lc}}{J^2_c} \simeq \frac{2GM_\bullet r_{lc}}{J^2_c(\mathcal{E})}.$$  \hfill (4.3)

In my bulge + MBH model (on page 63), figure 4.1 shows $X^2_{lc}$ for a point mass MBH with $\epsilon_\bullet = 0$. Compared to a LC with $r_{lc} = 10^{-8}(10^{-5}\text{ pc})$ one with $r_{lc} = 0.003(3\text{ pc})$ occupies a larger domain in phase-space, reflecting the fact that $X^2_{lc}$ is proportional to $r_{lc}$. Taking the reliability of $N$-body implementations into consideration, I shall always use $r_{lc} = 0.003$ which is comparable to the $N$-body softening length. For the physically motivated tidal disruption events, I will extrapolate results to the value of MBH’s tidal radius, $r_{lc} = r_t = 10^{-8}(10^{-5}\text{ pc})$. 

Figure 4.1: The phase-space LC boundary $X^2_{lc}$ (eq. 4.3) for the bulge + MBH model: $r_{lc} = 0.003(3\text{ pc})$ in solid curve and $r_{lc} = 10^{-8}(10^{-5}\text{ pc})$ in dashed curve. On the right panel I zoom in between $0 \leq X^2_{lc} \leq 0.01$ to show that all but the most tightly bound stars, $X^2_{lc}(\mathcal{E})$ is almost independent of $\mathcal{E}$. The region bounded by the curve and axes describes the ratio of LC to the whole orbital family.
How to refill a loss cone in spherical galaxies?

In a precisely spherical potential, the LC at a given $\mathcal{E}$ would be emptied in one orbital period and no more stars would be consumed. In realistic systems, however, through encounters with other stars fresh stars may diffuse into the LC. This can be summarized as: stars diffuse in both $\mathcal{E}$ and $J^2$ with the same characteristic timescale $t_{\text{relax}}$:

- by losing orbital \textbf{energy}, stars can shrink their orbits down to the effective size of the LC. However, for all but the most tightly bound stars, the phase-space LC boundary $X_{\text{lc}}^2(\mathcal{E})$ is almost independent of $\mathcal{E}$ (see figure 1). No steep gradient in $\mathcal{E}$ causes the characteristic $\mathcal{E}$-relaxation time $\sim t_{\text{relax}}$ to be on the same order as the two-body relaxation time, which can be much longer than the Hubble time. Therefore, the effect of $\mathcal{E}$-diffusion is likely to be only subdominant.

- by losing orbital \textbf{angular-momentum} so that orbits become nearly radial and interact strongly with the MBH at their periapse visits. Outside the MBH’s sphere of influence (small $\mathcal{E}$), LC boundary $X_{\text{lc}}^2(r \gg r_{\text{infl}}) \ll 1$. This implies that $J$-relaxation can be $\ll t_{\text{relax}}$, much shorter than the two-body relaxation time. Hence, diffusion in $J^2$ is the dominant contributor to the refilling rate.

Classical LC theory

Assuming the distribution of stars has evolved to a \textit{steady-state} where the LC consumption of stars is balanced by the encounter-driven supply, Cohn & Kulsbrud (1978) numerically integrated the Fokker-Planck equation in energy-angular-momentum space to study the evolution of star clusters containing a central MBH. Magorrian & Tremaine (1999, hereafter MT99) generalized this Fokker-Planck treatment to non-Keplerian potentials and compute the steady-state flux of stars into the LC. Below is a summary of Sec. 3 in MT99. Let $f(\mathcal{E}, J^2; r)d\mathcal{E}dJ^2$ be the probability of finding a star within $d\mathcal{E}dJ^2$ of $(\mathcal{E}, J^2)$ at a given radius $r$. Neglecting the effects of large-angle scatterings, the evolution of $f(\mathcal{E}, J; r)$ near the LC can be approximated by the local Fokker-Planck
Chapter 4: Loss cone refilling by flyby encounters

equation 1 (eq. 2.57)
\[
\frac{\partial f}{\partial t} + v_r \frac{\partial f}{\partial r} = \frac{\partial}{\partial X^2} \left[ -\langle \Delta X^2 \rangle f + \frac{1}{2} \frac{\partial}{\partial X^2} \left( \langle (\Delta X^2)^2 \rangle f \right) \right],
\] (4.4)

where radial velocity \( v_r = [2(\psi(r) - \mathcal{E})]^{1/2} \); the diffusion coefficients \( \langle \Delta X^2 \rangle \) and \( \langle (\Delta X^2)^2 \rangle \), both functions of \( (\mathcal{E}, X^2; r) \) measure the rate at which two-body encounters cause stars to diffuse in \( X^2 \).

The steady-state diffusion rate of stars into the LC is then given by the steady-state solutions to eq. (4.4):
\[
F_{2\text{body}}^{\text{LC}}(\mathcal{E}) d\mathcal{E} = \frac{F_{\text{max}}^{\text{LC}}(\mathcal{E}) d\mathcal{E}}{\ln^{-1} X_0^2(\mathcal{E})},
\] (4.5)

where
\[
F_{\text{max}}^{\text{LC}}(\mathcal{E}) \equiv 4\pi^2 \bar{f}(\mathcal{E}) J_c^2(\mathcal{E}) \mu(\mathcal{E}) T_r(\mathcal{E}),
\]

\[
X_0^2(\mathcal{E}) \equiv X_{lc}^2(\mathcal{E}) \times \begin{cases} 
\exp(-q) & \text{for } q(\mathcal{E}) > 1 \\
\exp(-0.186q - 0.824\sqrt{q}) & \text{for } q(\mathcal{E}) < 1
\end{cases}.
\] (4.6)

MT99 derived expressions for the local angular-momentum diffusion coefficients (see Appendix A.2), from which the orbit-averaged diffusion coefficient \( \bar{\mu}(\mathcal{E}) \) can be calculated as
\[
\bar{\mu}(\mathcal{E}) \equiv \frac{1}{T_r(\mathcal{E})} \int \frac{dr}{v_r} \lim_{X^2 \to 0} \frac{\langle (\Delta X^2)^2 \rangle}{2X^2},
\] (4.7)

and
\[
q(\mathcal{E}) \equiv \frac{1}{X_{lc}^2(\mathcal{E})} \int \frac{dr}{v_r} \lim_{X^2 \to 0} \frac{\langle (\Delta X^2)^2 \rangle}{2X^2} = \frac{\bar{\mu}(\mathcal{E}) T_r(\mathcal{E})}{X_{lc}^2(\mathcal{E})}.
\] (4.8)

\( q(\mathcal{E}) \) can be treated as the ratio between the orbital period \( T_r(\mathcal{E}) \) and the orbit-averaged time scale for two-body diffusion to refill the consumption zone.

\( X_0^2 \) is the value of \( X^2 \) at which \( f \) falls to zero due to the removal of stars that enter into the LC and it is usually it is not equal to its phase-space boundary value \( X_{lc}^2 \neq X_{lc}^2 \), because the

---

1Because of the presence of the LC, the steady-state distribution of scatters is not quite isotropic, \( f(\mathcal{E}, X^2) \). It is, however, reasonable to calculate the diffusion coefficients using the “isotropized” DF defined by \( \bar{f}(\mathcal{E}) \equiv \int_0^1 f(\mathcal{E}, X^2) dX^2 \).
scattering of stars into LC orbits permits $f$ to be nonzero even for $X^2 < X_{lc}^2$.

**How to model a loss cone in $N$-body simulations?**

When using $N$-body modelling to mimic tidal disruption events, I simply monitor stars that venture into a LC sphere and set their mass to be zero. This model ignores the tiny growth in the MBH’s mass.

Starting with a spherical isotropic galaxy, it is reasonable to assume that stars on low-angular-momentum orbits have already been cleared, that is orbits with $J^2 \leq J_{lc}^2$ are unpopulated. Figure 4.2 shows the modified stellar density and velocity dispersion profiles. Owing to the depletion of low angular-momentum stars which will otherwise populate the central region, the original cuspy density profile is turned into a small hole. Left with large angular-momentum or large-$v_t$ stars, the central velocity field becomes strongly tangentially biased $\beta = 1 - \sigma_t^2/2\sigma_r^2 \leq 0$.

### 4.2 Flyby encounters

Here, I make use of the multi-mass bulge + MBH model from Sec. 3.4 and allow the LC consumption sphere to eat/remove stars.

**Data analysis**

From simulations, I directly measure $M_{lc}(t)$, the mass of stars captured into the LC as a function of time. By smoothing the $M_{lc}(t)$ curve with an averaging window of width $\Delta t = 0.4$, I estimate the flux $\dot{M}_{lc}(t)$. This window is large enough to remove the sharp edges and small enough not to affect the global shape of the envelope. However, the arbitrary choice of $\Delta t$ does make $\dot{M}_{lc}(t)$ a less robust quantity when compared to $M_{lc}(t)$.

#### 4.2.1 Tests: calibrating noise-driven LC refilling

When modelling LC dynamics, it is of primary importance to ensure the empty nature of the LC is not compromised by Poisson noise in $N$-body realizations. Below, I test the sensitivity
Figure 4.2: Mass density (top left), mass-weighted velocity dispersion of radial (top right) and tangential component (bottom right) and velocity anisotropy profiles $\beta = 1 - \sigma_t^2 / 2\sigma_r^2$ (bottom left) of the bulge + MBH model. The solid curves show the profiles after depleting particles inside a LC sphere of radius $r_{lc} = 0.003(3\,\text{pc})$. For comparison, un-modified profiles are shown in dashed curves.
of noise-driven LC fluxes for different $N$-body parameters.

- **The particle mass spectrum**
  I compare results between a multi-mass realization and its equivalent equal-mass model in figure 4.3. The equal-mass returned results (thick-solid curves) are considerably larger than the multi-mass results (thin curves), reflecting its larger degree of numerical diffusion. Another by-product of the multi-mass realization is its smoother LC stream, owing to its increased resolution in low-angular-momentum phase-space. Also shown in figure 4.3 are two multi-mass realizations generated from two random seeds (thin-solid and thin-dashed curves), they are practically identical.

- **The number of particles**
  I construct a series of multi-mass realizations with $N/10^6$ ranging from 0.05, 0.1, 0.2, 0.4, 0.8 to 2. Refilling rates found in figure 4.4 exhibit a strong $N$-dependence. On the basis of LC theory (Milosavljević & Merritt 2003a), the $N$-dependent behavior can be described quantitatively as: stars removed are replaced on a (numerical) two-body relaxation time-scale which increases roughly as $N$.

- **The size of the LC**
  I set up another LC model with a smaller radius $r_{lc} = 0.001\, (1\, \text{pc})$. As shown in figure 4.5, a larger consumption zone helps the LC to interact with more stars and captures a larger mass.

The noise-driven LC refilling is very sensitive to numerical resolutions. With the same $N$, a multi-mass realization returns a smaller and smoother LC stream compared to a conventional equal-mass realization. This is because the multi-mass model has a better mass resolution in high-$\mathcal{E}$ and low-$J$ phase-space. Or, it is low-mass stars on eccentric orbits that interact with the LC sphere most frequently. For different multi-mass realizations, a large $N$ further suppresses the phase-space diffusion and the noise-driven flux. To the first order, the captured mass and its flux scale linearly with the linear size of the LC, if $r_{lc} \ll 1$. 

![Image](image-url)
Chapter 4: Loss cone refilling by flyby encounters

Figure 4.3: Noise-driven LC refilling in the bulge + MBH model, using $N = 2 \times 10^6$ particles. The LC of radius $r_{lc} = 0.003$ is emptied when simulations start. Two panels show the cumulative mass of stars fed into LC sphere $M_{lc}$ (left) and its value per unit time $\dot{M}_{lc}$ (right). After the initial transient period ($t < 2$), all $\dot{M}_{lc}$ converges to some typical values. Two multi-mass models generated from different random seeds (thin-dashed and thin-solid), are practically identical. Equal-mass results (thick-solid curves) are considerably larger than the multi-mass ones; what is more, multi-mass models evolve much smoother with time because low-mass stars are put on orbits which take them to the centre and interact with the LC sphere.

Figure 4.4: Same as figure 4.3 but for realizations with various $N$, from top to bottom are: $N/10^6 = 0.05$ (dashed), 0.1 (dashed), 0.2 (solid), 0.4 (solid), 0.8 (solid), 2 (thick-solid) respectively. $M_{lc}$ decreases with increasing $N$, implying that the evolution is limited to collisional mechanism known as (numerical) two-body relaxation. $\dot{M}_{lc}$ from small-$N$ realizations (dashed curves) are both large and noisy.
Chapter 4: Loss cone refilling by flyby encounters

4.2.2 Experiments: measuring perturber-driven LC refilling

The satellite model

I now introduce a low-mass satellite to the otherwise (almost) spherical symmetric system. The satellite is modelled as a rigid Plummer sphere (eq. 2.14) with mass \( M_s \) and scale length \( \epsilon_s \). It is sent on a hyperbolic orbit with pericentre parameter \( b \) (the distance at which a satellite passes the MBHs) and pericentre velocity \( V \) (the relative velocity at its closest approach).

N-body realizations

The satellite contributes an external potential energy to eq. (3.34) as

\[
V_{\text{ext}}(x_i, t) = - \sum_{i=1}^{N} m_i \left[ \frac{GM_\bullet}{(|x_i - x_\bullet|^2 + \epsilon_\bullet^2)^{1/2}} + \frac{GM_s}{(|x_i - x_s|^2 + \epsilon_s^2)^{1/2}} - \frac{GM_\bullet M_s}{(|x_i - x_s|^2)^{1/2}} \right]
\]  (4.9)

where subscripts \( \bullet, i \) and \( s \) denote properties of the MBH, stars and satellite respectively. Similar to eq. (3.35), contributions to the force on one particle is balanced by an equal and opposite force on another. Together with the GROMMET code, the system does respect Newton’s third law. I use

Figure 4.5: Same as figure 4.3 but for different LC sizes: \( r_{lc} = 0.003(3\text{ pc}) \) (in thick curve) and \( r_{lc} = 0.001(1\text{ pc}) \) (in thin curve).
the same boxes and time-step refinement levels as stated on page 57, and use the same time-step as for the finest box to move both the MBH and the satellite.

I. Toy perturbers

What combinations of \((M_s, b, V)\) can give rise to a LC flux that is significantly above the threshold of noise-driven value?

To answer this question, I use a toy model to survey \((M_s, b)\) parameter space. That is, I pin the MBH together with its consumption sphere at the origin in space, but change its mass or distance to the MBH in order to find out the physically interesting part of \((M_s, b)\)-parameter space. This model is unrealistic from an astrophysical point of view, but is justified since I am only interested in the limiting case: either small \(M_s\) or large \(b\) where the induced perturbations are small. As described below, I change \(M_s\) or \(b\) one at a time as:

- \(b = 1\) and \(M_s = 0.001, 0.003, 0.01, 0.03, 0.1, 0.3\).

Figure 4.6 shows the resultant \(M_{lc}(t)\) and \(\dot{M}_{lc}(t)\). The fact that a \(M_s = 0.001\) perturber gives a signal just above the noise level allow me to use it as the smallest satellite mass in the flyby case. In \(M_s = 0.3, 0.1\) cases, \(M_{lc}(t)\) converges very quickly at some small values. The reason is that instead of being a small-mass perturber, too massive ones can destroy the underlying stellar system and become the new gravitational centre of the system.

- \(M_s = 0.1\) and \(b = 1, 2, 8, 32\).

Results are shown in figure 4.7. The fact that a perturber outside a sphere of \(r = 10\) gives no observable signals even after \(t = 20\), suggests the largest and physically meaningful value of \(b\) be studied is \(b = 10\); it also suggests that introducing a satellite outside a sphere of \(r = 10\) causes no serious transient effects to the otherwise (almost) equilibrium model.

In all, I choose \(1 \leq b \leq 10\) and \(0.001 \leq M_s \leq 0.3\), and illustrate this in figure 4.8. \(b\) is chosen logarithmically at 1, 2, 4, 8 up to 32; \(V\) is chosen at 1, 2, 4, 8, 16, larger than the circular velocity at \(b\). This gives \(b\) the pericentre radius of an orbit. The sample is complete in the sense that it (sparsely) covers all the regions where the perturber-driven LC refilling can beat the noise-driven value.
Figure 4.6: Fixed perturber-driven LC refilling. A constant perturbing potential is introduced at $b = 1$ to the system in which the MBH and its LC sphere are pinned at the centre. All simulations use a total number of $N = 2 \times 10^6$ particles. The LC of radius $r_{lc} = 0.003$ is emptied when simulations start. The panels show the cumulative mass of stars fed into LC sphere $M_{lc}$ (left) and its value per unit time $\dot{M}_{lc}$ (right). Different curves are for perturbers with different masses: the thick solid curves are for $M_s = 0.3$, thick-dashed curves are for $M_s = 0.1$, and the set of thin curves from top to bottom are for $M_s = 0.03, 0.01, 0.003$, and $M_s = 0.001$. The dotted curves indicate the numerical noise level.

Figure 4.7: Same as figure 4.6 but adding a $M_s = 0.1$ perturber at different distance $b$ from the galactic centre: thick curves are for $b = 1$, the set of thin curves from top to bottom are $b = 2, 8, 32, 128$. Note that cases with $b = 32, 128$ are almost indistinguishable from the noise-driven LC stream as shown with light-dotted curves. Too massive perturbers ($M_s = 0.3, 0.1$) can destroy the underlying stellar system simply because they become the new gravitational centres; $M_{lc}(t)$ converge very quickly at some small values.
Figure 4.8: Setup for flyby encounters (asterisks). A pericentre radius \( b \) is chosen from the scale length outwards \( b \geq 1 \) (kpc); a pericentre velocity \( V \) is then chosen to be larger than the circular velocity at \( b \) and \( V > V_{\text{cir}}(b) \), ensuring \( b \) to be the pericentre.

The coordinates are chosen so that the MBH is at the origin and the satellite starts with \( r_s^0 = (0, y^0, 10) \) and \( V^0 = (0, V_y^0, V_z^0) \). Each \( (b, V) \) pair corresponds to a \( (r_s, V) \) pair where \( r_s = (0, b, 0) \) and \( V = (0, 0, V) \). Assuming a test particle orbits in the combined potential of the MBH and stars, I move the satellite back to its “initial” conditions. Having \( r_s^0 \) and \( V^0 \), the only thing left is to choose a satellite mass \( M_s \), from 0.01, 0.03, 0.1 (default) to 0.3.

II. An example of flyby encounters

Now, I detail the numerical experiments of one satellite model well after it recedes from the galaxy’s central region. The chosen satellite has \( (M_s, b, V) = (0.3, 2.0, 2.0) \), and the simulation has been run for 20 time units.

Figure 4.9 shows the results for a \( N = 2 \times 10^6 \) multi-mass realization. The distance \( D(t) \) is calculated from the satellite to the MBH, where the MBH follows the centre-of-mass (COM) of the stars closely. \( M_{lc}(t) \) remains the same as the Poisson noise threshold (dashed curves) for \( t < 2 \) and \( t > 8 \). The former suggests the satellite is introduced without raising any noticeable transient
effects; the latter suggests that the system is followed well after the satellite leaves the central region.

Above the noise level, there is always an enhanced LC refilling signal as the satellite approaches the LC sphere; $\dot{M}_{lc}(t)$ reaches its peak value shortly after the satellite’s peri-centre passage (indicated by the asterisk). For convenience, I denote $M_{lc}(t)$ to be the net LC refilling mass calculated by subtracting the noise-driven value from the total flux.

Figure 4.10 compares results between models using $N = 2 \times 10^5$ and $2 \times 10^6$ multi-mass particles; $M_{lc}(t)$ is essentially the same. The absence of $N$-dependence implies that the process is determined by the overall structure. The evolution of $M_{lc}(t)$ is also smoother when $N$ is larger.

I compare results from generating multi-mass models with different seeds in figure 4.11, with different mass-species in figure 4.12, and between introducing satellite at different epochs in figure 4.13. Among all these experiments, the results are comparable.

In all, although various realizations have very different noise levels, the net LC streams (given by subtracting the numerical noise value from the total flux) converge towards $2.1 \times 10^{-4} (2.1 \times 10^6 M_\odot)$, suggests it is the correct answer. The $N$-independent behavior indicates that the supply of stars to the LC is not limited by collisional mechanism and results can be robustly scaled to real galaxies.

4.2.3 Results: an empirical $M_{lc}$ formula

Here and below, I denote $M_{lc}$ to be the maximum net LC refilling mass but ignore its detailed evolution. It has been measured against the satellite’s pericentre radius $b$ and pericentre velocity $V$ for a wide range of masses $M_*$. The $N$-body results are listed in table 4.1. For a LC model with $(M_*, n_{lc}) = (0.01, 0.003)$, I summarize the data by fitting them with a formula

$$\log M_{lc} = \alpha \log M_* - \beta \log b - \gamma \log V + C,$$

(4.10)

where $\alpha, \beta, \gamma$ and $C$ are evaluated by carrying out standard least-squares fitting which yield

$$\alpha = 1.0, \quad \beta = 2.0, \quad \gamma = 1.3, \quad C = -1.85.$$  

(4.11)
Chapter 4: Loss cone refilling by flyby encounters

Figure 4.9: The short-term evolution during a flyby encounter with a satellite of $M_s = 0.1$ passing at pericentre $b = 1$ with velocity $V = 4$ (solid curves). Four panels are: the cumulative refilling flux $M_{lc}^a$ (top left), its rate $\dot{M}_{lc}^a$ (top right), net refilling mass $M_{lc}$ (bottom left) and distance $D$ to the MBH. For comparison, the noise-driven values are shown in dashed-curves.
Figure 4.10: Similar to figure 4.9 except the right bottom panel gives the filling rate. Different numbers of particles have been used: $2 \times 10^5$ in thin-solid curves and $2 \times 10^6$ in thick-solid curves. For comparison, noise-driven evolution is plotted in dashed curves. When the spherical model is perturbed by a flyby encounter, not only the net filling rate is high and significant, but also there is no systematic dependence on $N$. 
Figure 4.11: Same model as in figure 4.9 using different random seeds.

Figure 4.12: Same model as in figure 4.9 using different particle mass-species: equal-mass (thin curves) vs multi-mass (thick curves).

Figure 4.13: Same model as in figure 4.9 but introducing satellites at different epochs. The outcome is insensitive to the phase of the flyby orbit, as long as it is introduced outside a sphere of $r_s = 10$. 
And the formula (4.10) becomes

$$M_{lc} = 1.4 \times 10^{-2} \left( \frac{M_s}{1} \right)^{1.0} \left( \frac{1}{b} \right)^{2.0} \left( \frac{1}{V} \right)^{1.3}. \quad (4.12)$$

Supplying it with the satellites’ parameters ($M_s, b, V$), eq. (4.12) gives the function-fitting $M_{lc}^f$. Plotting them against the $N$-body observed values in figure 4.14, 4.15 and 4.16 one can see that the formula fits data well at high-$M_s$, low-$V$ and small-$b$ end where signals are strong.

I carry out additional simulations using ($M_\bullet, r_{lc}$) = (0.01, 0.001). Results are listed in table 4.2. The sampled satellite parameters are much limited by the competition between weak signals and artificial numerical noise. I assume $M_{lc}$ maintains its current form while add $r_{lc}$, the LC effective radius to the fitting formula as

$$M_{lc} = 4.7 \left( \frac{M_s}{1} \right)^{1.0} \left( \frac{r_{lc}}{1} \right)^{1.0} \left( \frac{1}{b} \right)^{2.0} \left( \frac{1}{V} \right)^{1.3}. \quad (4.13)$$

For small LC, the LC flux is proportional to its linear size $M_{lc} \propto J_{lc}^2 \simeq 2GM_\bullet r_{lc}$; this gives $M_{lc} \propto r_{lc}$.

In terms of scaling,

$$M_{lc} = 4.7 \cdot M_{\text{bulge}} \left( \frac{r_{lc}}{r_{\text{bulge}}} \right) \left( \frac{M_s}{M_{\text{bulge}}} \right) \left( \frac{b}{r_{\text{bulge}}} \right)^{-2.0} \left( \frac{V}{V_{\text{bulge}}} \right)^{-1.3}, \quad (4.14)$$

where $M_{\text{bulge}}$ and $r_{\text{bulge}}$ are the total mass and scale length for a Hernquist bulge, and $V_{\text{bulge}}$ is determined through

$$V_{\text{bulge}} = 100 \text{ km s}^{-1} \left( \frac{M_{\text{bulge}}}{10^{10} \text{M}_\odot} \right)^{1/2} \left( \frac{r_{\text{bulge}}}{1 \text{ kpc}} \right)^{-1/2}. \quad (4.15)$$

**The validity of $M_{lc}$ formula**

The derived $M_{lc}$-formula (eq. 4.13) can be extrapolated to values beyond the regime covered by the $N$-body simulations, however, much care should be taken.
Chapter 4: Loss cone refilling by flyby encounters

Figure 4.14: LC refilling mass $M_{lc}$ versus satellite's pericentre velocity $V$. Asterisks are values from $N$-body integrations. The dashed lines indicate the least-square fits to four groups of orbits with similar pericentre parameter $b$, from top to bottom $b = 1, 2, 4, 8$; they have the same slopes as $V^{-1.3}$.

Figure 4.15: Same as fig. 4.14 but for $M_{lc}$ versus satellite pericentre radius $b$. The dashed lines indicate four groups of orbits with similar $V$; from top to bottom are $V = 1, 2, 4, 8, 16$. They have the same slopes as $b^{-2.0}$.

Figure 4.16: Same as fig. 4.14 but for $M_{lc}$ versus satellite mass $M_s$. The dashed lines indicate three groups of orbits with similar $(b, V)$; from top to bottom are $(1, 2), (1, 4), (2, 2)$. They have the same slope as $M_s^{1}$. 
Chapter 4: Loss cone refilling by flyby encounters

Table 4.1: Satellite parameters and LC flux from N-body simulations

<table>
<thead>
<tr>
<th>$M_s$</th>
<th>$b$</th>
<th>$V$</th>
<th>$M_{lc}$</th>
<th>$M_s$</th>
<th>$b$</th>
<th>$V$</th>
<th>$M_{lc}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.9</td>
<td>2.1</td>
<td>$1.7 \times 10^{-3}$</td>
<td>0.1</td>
<td>1.1</td>
<td>1.7</td>
<td>$5.8 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.3</td>
<td>1.0</td>
<td>4.1</td>
<td>$5.9 \times 10^{-4}$</td>
<td>0.1</td>
<td>1.0</td>
<td>3.9</td>
<td>$2.3 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.3</td>
<td>1.0</td>
<td>8.0</td>
<td>$2.9 \times 10^{-4}$</td>
<td>0.1</td>
<td>1.0</td>
<td>7.9</td>
<td>$1.0 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.3</td>
<td>1.0</td>
<td>16.0</td>
<td>$1.1 \times 10^{-4}$</td>
<td>0.1</td>
<td>1.0</td>
<td>2.7</td>
<td>$2.7 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.3</td>
<td>1.9</td>
<td>2.0</td>
<td>$4.2 \times 10^{-4}$</td>
<td>0.1</td>
<td>2.1</td>
<td>1.9</td>
<td>$1.5 \times 10^{-4}$</td>
</tr>
<tr>
<td>0.3</td>
<td>2.0</td>
<td>4.0</td>
<td>$2.1 \times 10^{-4}$</td>
<td>0.1</td>
<td>2.0</td>
<td>3.9</td>
<td>$6.4 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.3</td>
<td>2.0</td>
<td>8.0</td>
<td>$7.7 \times 10^{-5}$</td>
<td>0.1</td>
<td>2.0</td>
<td>7.9</td>
<td>$2.5 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.3</td>
<td>2.0</td>
<td>16.0</td>
<td>$1.5 \times 10^{-5}$</td>
<td>0.1</td>
<td>2.0</td>
<td>6.6</td>
<td>$6.6 \times 10^{-8}$</td>
</tr>
<tr>
<td>0.3</td>
<td>4.1</td>
<td>0.9</td>
<td>$2.6 \times 10^{-4}$</td>
<td>0.1</td>
<td>4.4</td>
<td>0.9</td>
<td>$6.9 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.3</td>
<td>4.0</td>
<td>2.0</td>
<td>$1.2 \times 10^{-4}$</td>
<td>0.1</td>
<td>4.1</td>
<td>1.9</td>
<td>$3.5 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.3</td>
<td>4.0</td>
<td>4.0</td>
<td>$5.4 \times 10^{-5}$</td>
<td>0.1</td>
<td>4.0</td>
<td>4.0</td>
<td>$1.6 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.3</td>
<td>4.0</td>
<td>8.0</td>
<td>$1.7 \times 10^{-5}$</td>
<td>0.1</td>
<td>4.0</td>
<td>8.0</td>
<td>$3.4 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.3</td>
<td>8.3</td>
<td>1.0</td>
<td>$6.1 \times 10^{-5}$</td>
<td>0.1</td>
<td>8.6</td>
<td>1.0</td>
<td>$1.4 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.3</td>
<td>8.1</td>
<td>2.0</td>
<td>$2.7 \times 10^{-5}$</td>
<td>0.1</td>
<td>8.2</td>
<td>2.0</td>
<td>$7.0 \times 10^{-6}$</td>
</tr>
<tr>
<td>0.3</td>
<td>8.1</td>
<td>4.0</td>
<td>$1.0 \times 10^{-5}$</td>
<td>0.1</td>
<td>8.1</td>
<td>4.0</td>
<td>$3.3 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.03</td>
<td>1.3</td>
<td>1.5</td>
<td>$1.6 \times 10^{-4}$</td>
<td>0.01</td>
<td>1.0</td>
<td>1.7</td>
<td>$5.0 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.03</td>
<td>1.0</td>
<td>3.8</td>
<td>$6.0 \times 10^{-5}$</td>
<td>0.01</td>
<td>1.0</td>
<td>3.8</td>
<td>$1.4 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.03</td>
<td>1.0</td>
<td>7.9</td>
<td>$2.4 \times 10^{-5}$</td>
<td>0.01</td>
<td>1.0</td>
<td>7.9</td>
<td>$4.8 \times 10^{-6}$</td>
</tr>
<tr>
<td>0.03</td>
<td>1.0</td>
<td>16.0</td>
<td>$3.6 \times 10^{-6}$</td>
<td>0.01</td>
<td>1.0</td>
<td>16.0</td>
<td>$5.0 \times 10^{-7}$</td>
</tr>
<tr>
<td>0.03</td>
<td>2.2</td>
<td>1.8</td>
<td>$4.7 \times 10^{-5}$</td>
<td>0.01</td>
<td>2.2</td>
<td>1.8</td>
<td>$1.5 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.03</td>
<td>2.1</td>
<td>3.9</td>
<td>$2.0 \times 10^{-5}$</td>
<td>0.01</td>
<td>2.1</td>
<td>3.8</td>
<td>$6.5 \times 10^{-6}$</td>
</tr>
<tr>
<td>0.03</td>
<td>2.0</td>
<td>7.9</td>
<td>$5.8 \times 10^{-6}$</td>
<td>0.01</td>
<td>2.0</td>
<td>7.9</td>
<td>$2.6 \times 10^{-6}$</td>
</tr>
</tbody>
</table>

($M_* = 0.01, r_* = 0.003; r_{lc} = 0.003$)

- The scaling does not hold at arbitrarily small pericentre radius where the assumption of a non-fatal flyby encounter may fail. When $b$ is too small, $M_s$ would become comparable to the mass of the host system that lies interior to its orbit; even a low-mass satellite will shatter the host system. Therefore, there should be a cut off in $b$ where $M_s \leq M_{\text{bulge}}(r < b)$.

- Similarly, the scaling does not hold in the very low velocity regime where $M_{lc}$ would diverge. There, I expect $M_{lc}$ to increase more slowly and eventually converge to a constant value. For a parabolic orbit considered, there is a natural physical constraint: the pericentre velocity should be larger than the circular velocity at pericentre radius, $V > V_{\text{cir}}(b)$.
### Table 4.2: LC flux cont.

<table>
<thead>
<tr>
<th>$r_{lc}$</th>
<th>$b$</th>
<th>$V$</th>
<th>$M_{lc}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>1.1</td>
<td>1.7</td>
<td>$9.8 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.001</td>
<td>0.8</td>
<td>4.0</td>
<td>$7.5 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.001</td>
<td>2.1</td>
<td>1.9</td>
<td>$4.1 \times 10^{-5}$</td>
</tr>
<tr>
<td>0.001</td>
<td>2.0</td>
<td>4.0</td>
<td>$1.7 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

($M_* = 0.01, \epsilon_* = 0.003; M_s = 0.1, \epsilon_s = 0.5$.)

- For the satellite model, there is a minimal softening length $\epsilon_s = 2GM_s/V^2$ which guarantees a small-angle deflection. To test the sensitivity of $M_s$ on the choice of $\epsilon_s$, I choose two different scale length $\epsilon_s = 0.5$ (default) and 0.8. In general, a more extended satellite gives a smaller LC flux; however, the change is less than 10% and should not affect the fitting parameters.

In all, this fitting formula can be used to predict the LC captured mass during a flyby event, provided

$$M_s < M_{\text{bulge}}(r < b), \quad V > V_{\text{cir}}(b), \quad b > \epsilon_s > 2GM_s/V^2.$$  \hspace{1cm} (4.16)

#### 4.2.4 Discussion

**The rigid satellite model**

I have made use of a rigid satellite model, which explicitly ignores the possibility of tidal stripping. This is justified, since one does not expect a significant mass loss for a compact perturber during a single flyby. In any case, most of the mass loss is likely to occur at and shortly after the peri-centre passage, when the primary galaxy has already been significantly perturbed (Vesperini & Weinberg 2000).

**Dynamical friction**

In my self-consistent $N$-body experiments, a satellite particle experiences substantial dynamical friction. Figure 4.17 shows the simulation returned $(b, V)$-pair, for different satellite masses (see the caption for details). In short, a larger frictional force experienced by a heavier satellite
Chapter 4: Loss cone refilling by flyby encounters

Figure 4.17: N-body returned pericentre radius \((b)\) and pericentre velocity \((V)\) parameters for flyby encounters, for different satellite mass: \(M_s = 0.3\) (circle), 0.1 (asterisk), 0.03 (square), 0.01 (triangle). In the full self-consistent N-body experiments, a satellite particle experiences substantial dynamical friction. A more massive satellite \((M_s \uparrow)\) experiences a larger frictional force; it is dragged further into the potential well \((b \downarrow)\) and gains a larger kinetic energy \((V \uparrow)\).

\((M_s \uparrow)\) drags it further into the MBH’s potential well, together with a larger kinetic energy gain \((V \uparrow)\).

**Density profiles**

My dynamical model also provides information on the galaxy’s structural properties that a flyby encounter leaves. I take the density profile as an example.

Firstly, I have to assess the stability of the galaxy profile in the absence of a satellite flyby. Because both removing particles with \(J^2 \leq J^2_{lc}\) in the ICs and eliminating particles that drift into the LC sphere during the flyby encounter, would have changed the gravitational potential. The results are that the removal of \(M_{lc} \sim 10^{-4}\) mass can not perturb the system far away from an equilibrium. As shown in figure 4.18 the innermost density profile hardly changes after evolving for \(t = 50\).

Now, in the satellite flyby case, figure 4.19 shows an example of the central mass and density profiles after a \(M_s = 0.3, b = 1.9\) and \(V = 2.0\) satellite passing through. In response to
Chapter 4: Loss cone refilling by flyby encounters

the external perturbation, more stars are reshuffled onto LC orbits and removed. Hence, the core continues to grow as stars captured into the LC, counteracted by some fresh stars which find their way to the centre. As the satellite departs the system, stars settle down to a new equilibrium configuration within a few crossing times. Therefore, if perturbed by cumulative perturbers, one can simply accumulate the total consumed mass.

4.3 Where do LC stars come from?

At the high-resolution limit of multi-mass simulations, I examine the phase-space distribution of the LC star progenitors.

I first check out in spherical potential, the progenitors of the noise-driven LC flux. The left panel of fig. 4.20 shows that stars with low-angular-momentum are in the first place to be sent to the consumption zone (also see caption for details). In the case of satellite flyby, however, the lack of an axis of symmetry implies that stellar orbits need not conserve any component of their angular-momentum. The right panel of fig. 4.20 shows the results: after subtracting the two-body values (left panel) out of the total flux, I see that stars initially distributed far from the LC boundary ($J^2 \gg J^2_{lc}$) can be captured into the consumption zone.

4.3.1 Angular-momentum transport by satellite torque

In a frame centred on the orbit of the satellite, any field star could be viewed as approaching the satellite with a speed $V_s (\gg v_*, \text{ much larger than the stars’ internal velocity})$ at a distance $d_s (\geq \epsilon_s)$. By approximating the stars’ trajectory as a straight line or in the limit $d_s > 2GM_s/V_s^2$ (eq. 4.16), I can immediately use equation (1.30) of BT08 to show that the impulse from the satellite mainly deflects the perpendicular component of the stars’ velocity by an amount

$$
\Delta v^\perp_s \equiv |\Delta v^\perp_*| = \frac{2GM_s}{d_s V_s},
$$

(4.17)
Chapter 4: Loss cone refilling by flyby encounters

Figure 4.18: The long-time behavior of the bulge + MBH model. A LC sphere of radius $r_{lc} = 0.003$ is emptied when simulations start. The panels show the inner mass (left) and its density (right) profiles. The LC removes stars that drift into the LC sphere. Compared to the ICs (thick-solid curves), profiles at $t = 50$ (dashed curves) hardly evolve. Thin-solid curves indicate the original Hernquist model.

Figure 4.19: Same as figure 4.18, but after a $M_s = 0.3, b = 1.9$ and $V = 2.0$ satellite passing through. In response to the external perturbation, more stars are reshuffled onto LC orbits and removed. Hence, at $t = 30$, the central density profile drops and develops a core up to $r \sim 0.5 \text{(0.5 kpc)}$. 
much larger than the changes in the parallel component $\Delta v_\parallel^* \gg \Delta v_\parallel^*$. It follows that the change in angular-momentum can be estimated by

$$|\Delta J^*| \leq r^*_s \cdot \Delta v^*_\perp. \quad (4.18)$$

In words, $|\Delta J^*|$ is roughly equal to the acceleration at closest approach $GM_s/d^*_s$, times the duration of this interaction $2d^*_s/V_s$, times its distance to the central MBH $r^*_s = |r_s + d^*_s|$. $|\Delta J^*|$ varies as $d^{-1}_s$, thus stars with the smallest impact parameter $d^*_s = \epsilon_s (\ll r^*_s)$ get the strongest effect

$$\Delta J^*_s = f_{J} \cdot \frac{2Gr^*_\parallel M_s}{d^*_s V_s}, \quad |f_{J}| \leq 1. \quad (4.19)$$

This crude estimation shows that $\Delta J^*_s$ excited in an encounter falls off as $V_s^{-1}$ and proportional to $M_s$. During a flyby encounter, the directions between the impulse and the star relative to the centre of mass (COM) of the bulge are usually uncorrelated. This is because the former depends on the relative orientations between stars and the perturber while the latter depends on orbital phase of stars alone. Therefore, it is reasonable to assume that half of the stars gain angular-momentum and move to larger slower orbits while the other half lose angular-momentum and move inwards; the latter may find their way into the consumption zone.

### 4.3.2 Resonances from perturbation theory

Many theoretical investigations have explored the effects of weak encounters by means of linear perturbation theory. I review some relevant results.

In a spherical system, the orbital motion of stars is confined to a plane which can be characterized by two frequencies: a radial frequency and an azimuthal frequency. The radial frequency is defined through radial period (eq. 2.23), as

$$\Omega_r = \frac{2\pi}{T^*_r}, \quad T^*_r = 2 \int_{r_{peri}}^{r_{apo}} \frac{dr}{\sqrt{2[\psi(r) - \mathcal{E}] - J^2/r^2}}. \quad (4.20)$$
Figure 4.20: The initial phase-space density of particles (mass) that drift into the LC up to $t = 20$. Left panel: Poisson noise-driven LC stars in an isolated spherical potential; Right panel: LC in the presence of a satellite with $M_s = 0.1, b = 1, V = 4$, after subtracting the two-body values (left panel); the thin curve indicates the largest-possible, circular angular-momentum a star with energy $E$ can have. The thick line shows the LC boundary $J_{lc}^2 = 2.6 \times 10^{-4}$ corresponding to a consumption zone of $r_{lc} = 3 \times 10^{-3}$.

Figure 4.21: In the rest frame of the satellite, a close encounter between the satellite ($M_s$) and a field star ($m_*$); the MBH is indicated by $M_\bullet$. The impulse from the satellite mainly deflects the perpendicular component of the stars’ velocity much larger than the changes in the parallel component $\Delta v_\perp^* \gg \Delta v_\parallel^*$. tars with the smallest impact parameter $d_* = \epsilon_*$ get deflected most.
Similarly, the azimuthal frequency is defined through azimuthal period (eq. 2.24) as

$$\Omega_\phi = \frac{2\pi}{T_\phi} = \frac{\Delta \phi}{T_r}, \quad \Delta \phi = 2J \int_{r_{peri}}^{r_{apo}} \frac{dr}{r^2 \sqrt{2[\psi(r) - E] - J^2/r^2}}$$

(4.21)

Figure 4.22 shows the radial (left panel) and azimuthal (right panel) frequencies for orbital motions of stars in the bulge + MBH model (see caption for details).

Now, let us consider the gravitational potential arising from the flyby satellite and estimate its instantaneous orbital frequency as

$$\Omega_s = \frac{V_s}{r_s}$$

(4.22)

$\Omega_s$ is constantly changing as the satellite passes through the centre of the bulge. It lies in a range between $\Omega_{s,\text{min}}$ and $\Omega_{s,\text{max}}$; the minimum orbital frequency is non-zero so long as the system has a finite extent and the maximum orbital frequency is $\Omega_{s,\text{max}} = V/b$ when the satellite approaches its pericentre. As an example, the right panel of figure 4.23 shows $\Omega_s$ varies between 0.05 and 2 for a flyby satellite with $b = 1, V = 2$.

Tremaine & Weinberg (1984) investigate the torque on a satellite which revolves in a spherical system. They show that the angular-momentum in a single star changes as it passes through some confined islands in phase-space. With the triple of integers $m_1, m_2, m_3$, the regions corresponding to resonances can be identified as

$$m_1 \Omega_r + m_2 \Omega_\phi - m_3 \Omega_s = 0, \quad (-\infty < m_1, m_2 < \infty, 0 < m_3 < \infty).$$

(4.23)

e.g., if $m_1 = 0$ the (only) corotation resonance occurs at $\Omega_\phi = \Omega_s$; a resonant star is orbiting at the same angular speed as the satellite. For the bulge + MBH model studied, I show the five lowest order resonance in figure 4.23: the inner Lindblad resonance (LR) $(1, -1, 1), (-1, 2, 2)$, the corotation resonance $(0, 1, 1)$ and the outer LR $(1, +1, 1), (+1, 2, 2)$. 

Figure 4.22: Radial (left panel) and azimuthal (right panel) orbital frequencies in the bulge + MBH model. The set of five curves distinguish orbits with very different circularities $X^2 = J_2^2/J_c^2$, from the most eccentric orbits 0.0039 in thick curve to more circular orbits with $X^2 = 0.0156, 0.0625, 0.25, 1$ in thin curves. Notice on the left panel the five curves overlap each other, reflecting the fact that the radial frequency is mainly determined by the orbital energy and insensitive to its orbital circularity.

Figure 4.23: Left panel: Resonance quantity $\Omega_s = \frac{m_1}{m_3} \Omega_r + \frac{m_2}{m_3} \Omega_\phi$ in eq. (4.23). The set of five curves locate the lowest order resonances. From left to right, the solid-thin curves label the inner Lindblad resonance (LR) $(m_1, m_2, m_3) = (-1, 2, 2)$ and the outer LR $(+1, 2, 2)$; the dashed curves label the inner LR $(1, -1, 1)$ and $(1, +1, 1)$ the outer LR; the thick curve is the corotation resonance $(0, 1, 1)$. Right panel: An example of satellite’s instantaneous orbital frequency. A $(b, V) = (1, 2)$ flyby gives a $\Omega_s(t)$ (eq. 4.22) varies between 0.05 and 2.
Chapter 4: Loss cone refilling by flyby encounters

Figure 4.24: Sampled phase-space points in the test-particle experiments (dots), overlaid with locations of the lowest-order corotation resonances \((m_1, m_2, m_3) = (0,1,1)\) (solid curves), from left to right are: \(\Omega_s = \Omega_\phi = 8, 4, 2, 1\) (thick), 0.5, 0.25. Each point in \((E, J^2)\)-space gives a unique pericentre and pericentre velocity pair \((p, v_p)\), for which \(N_p = 10^4\) stars are smeared out in orbital phase uniformly. My sample covers \(r_E\) space uniformly, especially have sufficient phase-space coverage to locate resonances around the expected \(\Omega_s = 1\) (thick) resonance where \(r_E \sim 2\).

4.3.3 Are resonances important?

Test-particle experiments of a periodic perturber

The linear Hamiltonian perturbation theory suggests that the transfer of angular-momentum is mediated by orbits in resonance with quasi-periodic perturbers (Tremaine & Weinberg 1984; Vesperini & Weinberg 2000). For flyby encounters of interest, could resonances be the culprit for redistributing angular-momentum and eventually send stars to the LC?

To answer this question, I conduct additional experiments to track resonant dynamical processes pertain to a periodic driving, modelled as an orbiting satellite with a well-defined frequency \(\Omega_c\).
Model setup

In order to avoid any subtle numerical artifacts such as $N$-body fluctuation noise that mask or swamp resonances, I design a suite of test-particle experiments:

- In a rigid bulge + MBH model, I treat particle satellite encounters one at a time. Massless particles preclude any artificial orbital scattering.

- To keep a perturber with some given frequency $\Omega_c$ to orbit at $r_c$, a centripetal acceleration $\Omega_c^2 r_c$ is added by hand. I grow the satellite

$$M_s(t) = M_s^f \times \begin{cases} 3 \left( \frac{t}{t_g} \right)^2 - 2 \left( \frac{t}{t_g} \right)^3 & \text{if } t < t_g \\ 1 & \text{otherwise,} \end{cases} \quad (4.24)$$


to its final mass $M_s^f = 0.001$ within a time $t_g = 10$. Hence transient effects are suppressed.

- To have sufficient phase-space coverage, I choose $22 r_j$ logarithmically from $10^{-1}$ and $10^2$; at each $E_j = \psi(r_j)$, I choose $5 X_k^2 = J_k^2 / J_{\text{cir}}^2$ logarithmically from $10^{-5}$ to 1. As shown in figure 4.24, this gives about 100 $(E, J^2)$ pairs. $N_p = 10^4$ stars are sampled from each point and smeared out in orbital phase uniformly. I calculate the averaged fractional change in the angular-momentum between $t$ and 0 as

$$\eta = \frac{1}{N_p} \sum_{i=1}^{N_p} \left[ \frac{J_i(t) - J_i(t = 0)}{J_i(t = 0)} \right]^2. \quad (4.25)$$

My sample has sufficient phase-space coverage to locate resonances, because it well covers $r_E$ space, especially around $r_E \sim 2$ where the expected $\Omega_s = 1$ (thick) resonance would be. What is more important, the densely sampled orbital phase secures the change in any individual orbit is represented by contribution from orbits at many phases (Weinberg & Katz 2006).

- Although free of interparticle noise, the perturber still destroys stable orbits including ones in resonance: the perturber’s gravitational field scatters orbits, therefore, any otherwise con-

---

2On a circular orbit, the perturber can not be introduced from the distant location but can be slowly turned on in the distant past.
served quantities would drift and orbital phases would jump. In order to detect any resonant effects, I deliberately keep the satellite away from those orbits of interest. So for a driving frequency $\Omega_s = 1$, I first locate its lowest order resonance in energy (equivalent $r_E$) space, e.g., the prescribed corotation resonances would take place at $r_E = 2$. Then, I put the satellite orbiting at $r_c = 20$ such that stars in strong resonances would never have a chance to come close to the satellite since their apocentre radius $r_{apo} \leq r_E \ll r_c$. The phase-space width of resonant responses is proportional to the strength of the perturbation (e.g., Holley-Bockelmann et al. 2005).

Results

The behavior of $\eta$ (eq. 4.25), as shown on the left panel of figure 4.25, simply demonstrates that low-order resonances are responsible for effectively transferring angular-momentum. However, it requires a finite time to “turn on” resonances. Or, stars in resonance gradually (after many revolutions) build up a coherent response to the periodic perturbing potential (also see caption for detail).

This is in agreement with the prediction of linear perturbation theory (Tremaine & Weinberg 1984): (1) An off-resonance orbit precesses rapidly with respect to the orbiting perturber where the net change in the orbit cancels. During adiabatic changes, potential variations are slow compared to the typical orbital frequency and stars’ angular-momentum keep constant. (2) An in-resonance orbit is closed in the satellite frame, adiabatic invariance is broken. Lingering at resonance for many orbits, the response becomes non-linear and scales as the square root of the perturbation potential.

4.4 Summary

Repopulating an (nearly) empty LC is typically assumed to be driven by uncorrelated gravitational encounters between stars, where stars are scattered onto low angular-momentum orbits (Cohn & Kulsbrud 1978; MT99; Yu 2002; Wang & Merritt 2004). The so-called two-body relaxation
Figure 4.25: The fractional change in angular-momentum $\eta$ (eq. 4.25). The three groups are results at different times, from bottom to top $t = 20, 40, 60$. A satellite of mass $M_s = 0.001$ orbits at $r_c = 20$ with a driving frequency $\Omega_s = 1$. The low-order resonances are labeled on the right panel; they are nearly vertical lines on the $(E, J^2)$ diagram; the short vertical lines on both panels only indicate their locations. After about 10 revolutions, there are enhanced changes in $\eta$ at and around the corotation resonances (thick short vertical line). Stars in resonance gradually (after many revolutions) build up a coherent response to the periodic perturbing potential. Different circularities (initial $J$) do not affect $\eta$, as long as orbits have similar $E$.

Figure 4.26: Phase-space distribution of stars after evolving the system for $t = 60$, corresponding to 10 revolutions of the orbiting satellite. For each group of dark dots ($N_p = 10^4$), the over-plotted light dot indicates their same initial condition. After perturbed by an orbiting satellite, stars initially close to the LC boundary will have a good chance to be knocked into the LC.
is inherent to all stellar systems and ensures a minimal LC refilling rate. In this chapter, I explore another dynamical refilling mechanism: an interloping satellite.

The flyby orbit is very short-lived in time; uncertainty principle requires that it has a broad distribution in frequency. It can be regarded as superpositions of periodic orbits with their power being spreaded over different frequencies combined at different phases. There is a possibility of resonant coupling between the satellite and the stellar orbits. (1) However, stars can only pick up a limited amount of power, since the power of any single frequency is weak. (2) The main limitation is that a flyby encounter does not repeat its pattern; this prevents stars in instantaneous resonance from building up a significant response.

Sufficiently close to the MBH, the potential is nearly Keplerian. There, the radial and azimuthal orbital frequencies are equal, the coherent torques between stars lead to resonant angular-momentum relaxation (Rauch & Tremaine 1996; Rauch & Ingalls 1998; Hopman & Alexander 2006; Levin 2007). When scaling Hopman & Alexander’ (2006) results to the adopted bulge model, the resonant angular-momentum relaxation becomes unimportant at distances $\geq 10$ pc from the MBH. Therefore, this mechanism is irrelevant in current experiments, since most of the captured stars come from larger distance with $r_x \geq 100$ pc (fig. 4.20).

All lead to the conclusion that resonant coupling, either among stars or between stars and the satellite, are unimportant in explaining the enhanced LC flux caused by flyby encounters. It is the gravitational encounter with the flyby satellite that pumps an impulse onto the surrounding stars, transports angular-momentum and ultimately sends stars to the LC.
Rates of MBH feeding in bulge-halo systems

In this chapter, I study the prospects for feeding the central MBH by orbiting halo substructures. I start by reviewing substructure properties observed in high-resolution cosmological N-body simulations, then summarize information on the subhalo spatial distribution, radial density profile, mass function and concentration parameters into a DF of the form $f_{\text{sub}}(\mathcal{E}; M)$. In Sec. 5.2, I use the subhalo DF together with the knowledge of how much mass can be fed to a MBH after a single encounter to estimate the perturber-driven LC refilling rate for a typical small galaxy.
5.1 Modelling a Milky-way sized halo and its substructures

For a typical small galaxy, the bulge + MBH system in Sec. 4.2 fully determines the perturber-driven LC captured mass, given a satellite’s parameters. Neglecting the contribution from the disk component, the bulge is now assumed to be a spherical model of the Milky Way galaxy, however, harbouring a $10^8 \, M_\odot$ MBH which is $\sim 25$ times the mass of SgrA* at the Galactic centre, $3.7 \times 10^6 \, M_\odot$ (Ghez et al. 2005).

The bulge + MBH is believed to be encompassed by a galactic dark matter (DM) halo. High resolution cosmological experiments indicate that a large fraction of the mass within a collision-less DM halo is in the form of virialized, continuing infalling subhalos. Precise modelling of both the smooth DM distribution (the diffuse galactic component) and the subhalo population within (the clumpy component) is therefore mandatory to assess the rates of perturber-driven MBH feeding.

5.1.1 The halo density profile

For a Milky Way-sized DM halo of interest, I resort to “Via Lactea” (Diemand et al. 2007a), to date the highest resolution simulation of DM substructures. The spherically-averaged halo density profile of the model is consistent within 10 per cent with the benchmark NFW model (Navarro Frenk and White 1996)

$$\rho_{\text{NFW}}(r) = \frac{\rho_{\text{NFW}}}{r/r_s \left(1 + r/r_s \right)^2}.$$  (5.1)

where the estimated scale radius $r_s$ is related to the scale density $\rho_{\text{NFW}}$ by the virial mass $^1M_{200}$ such that

$$M_{200}^{\text{NFW}} = \frac{4}{3} \pi \rho_{\text{NFW}} r_s^3.$$  (5.3)

From an astrophysical point of view, the NFW model has a serious defect: its logarithmic

---

$^1$The virial mass, a measure of the total mass of the halo, is the mass inside a sphere of virial radius $r_{200}$, within which the mean density equals 200 times the critical density

$$M_{200} = \frac{4}{3} \pi \left(200 \rho_{\text{crit}} \right) r_{200}^3.$$  (5.2)

Inside $r_{200}$ the halo is assumed to be in virial equilibrium.
Figure 5.1: Left panel: Density profile of the NFW model (eq. 5.1) in thick curves, the Hernquist model (eq. 5.4) model in thin curves. I scale the Hernquist model $\rho_s^H = 1.6 \rho_s^{\text{NFW}}$ so that the mass (right panel) inside scale radius $r_s$ matches the value for the NFW model; this gives $M^H_{200} = 2.4 M^{\text{NFW}}_{200}$.

As shown in figure 5.1, to the inner parts $r < r_s$, $\rho_H$ provides the same $r^{-1}$ cusp as in an NFW halo; to the outer parts $r > r_s$, $\rho_H \propto r^{-4}$ leads to a finite mass $M^H_{200} = 2\pi \rho_s^H r_s^3$. I adopt here

$$\rho_{\text{halo}}(r) = \rho_H(r) = \frac{\rho_s^H}{r/r_s (1 + r/r_s)^3}; \quad \rho_s^H = 1.6 \rho_s^{\text{NFW}}.$$

The mass inside $r_s$ matches the value in the NFW model and $M^H_{200} = 2.4 M^{\text{NFW}}_{200}$. I adopt $M^H_{200} = 10^{12} M_\odot$ and $r_s = 20 \text{kpc.}$
Chapter 5: Rates of MBH feeding in bulge-halo systems

5.1.2 The subhalo mass function

Simulations show that DM halos are not smooth but contain a wealth of virialized substructures in all resolved mass scales. For virialized subhalos, there exists an “universal” mass function ($MF$) (e.g., Gao et al. 2004)

$$f_{\text{MF}}^{\text{MF}}(M; r) \propto M^{-(\beta+1)} \cdot \theta [r - r_{\text{min}}(M)], \quad M_{\text{min}} \leq M \leq M_{\text{max}}, \quad (5.6)$$

with $\beta = 1$, independently of the host halo mass from $M_{\text{max}} \sim 10^{10} M_\odot$ down to the smallest resolved mass $\sim 10^6 M_\odot$ at the present epoch (Jenkins et al. 2001; Gao et al. 2004; Diemand et al. 2007a). When using the scale-invariant halos-in-halos arguments, eq. (5.6) can be extrapolated to the lightest super symmetric particle at a cutoff mass about $M_{\text{min}} \sim 10^{-6} M_\odot$. They are assumed to be the very first gravitational bound objects containing no substructures because no smaller mass halos have collapsed in the hierarchy (Green et al. 2004; Green et al. 2005; Diemand et al. 2005). The effect of tidal disruption is accounted for by the Heaviside step function $\theta [r - r_{\text{min}}(M)]$ where the tidal radius $r_{\text{min}}(M)$ will be determined in the next section.

5.1.3 The subhalo density profile and spatial distribution

At $z = 0$, Diemand et al. (2007b) find that the radial density profile of subhalos above $4 \times 10^6 M_\odot$ in the “Via Lactea” simulation are well-fitted by the NFW model. This result is also valid for much smaller substructures with masses in the range $[10^{-6} M_\odot, 4 \times 10^{-3} M_\odot]$; they populate a parent halo of $0.014 M_\odot$ at $z = 86$ (Diemand et al. 2006). A large fraction of these small substructures may (Moore et al. 2005; Berezinsky et al. 2006) or may not (Zhao et al. 2005) survive gravitational disturbances during early merger process and late tidal disruption from stellar encounters. The survivors suffer from significant mass loss which presumably modifies the outer part of their original NFW density profile to an exponential cutoff. However, the innermost region seems to preserve its power-law $\rho \propto r^{-1}$ profile.

Yet these constraints from numerical experiments do not uniquely define the subhalo density profiles. I simply assume that subhalos were “born” with the same NFW density profile as
their massive host, but with different virial concentration parameters defined to be the ratio between the virial radius and the scale radius

$$c = \frac{r_{200}}{r_s}. \quad (5.7)$$

The median virial concentration decreases with growing mass, consistent with the assertion that small-mass halos are mostly assembled at much earlier epoch when the universe was more dense, and scale with \((1 + z)\) as the mean density \(\rho_{\text{crit}}\) decreases with Hubble expansion between \(z = 0\), corresponding to \(r_s(z) \sim \text{constant}\) (NFW96; Bullock et al. 2001). From Neto et al. (2007), I adopt the concentration-mass relation as

$$c(M_{200}, z = 0) = 5.26(M_{200}/10^{14} h^{-1} \text{M}_\odot)^{-0.10}, \quad 10^{12} h^{-1} \text{M}_\odot \leq M_{200} \leq 10^{15} h^{-1} \text{M}_\odot. \quad (5.8)$$

A naive low-mass extrapolation gives a very large over-estimate \(^2\), e.g., \(c(10^{-6} \text{M}_\odot, z = 0) \sim 500\).

As far as the spatial distribution of subhalos inside a Milky Way-like halo is concerned, I follow the indications of the numerical experiment of (Reed et al. 2005) and assume that the subhalo distribution traces that of the underlying host mass from \(r_{200}\) down to \(r_{\text{min}}(M)\), within which subhalos are efficiently destroyed by gravitational interactions. I explicitly assume spherical symmetry, follow the Roche criterion and compute \(r_{\text{min}}(M)\) as the minimum distance to the centre where the self-gravity of the subhalo at \(r_s\) equals the tidal field of the halo host computed at the orbital radius of subhalo (eq.8.92 BT08)

$$\rho_{\text{sub}} = 3 \frac{M_{\text{halo}}(r_{\text{min}})}{4\pi r_{\text{min}}^3} \quad (5.9)$$

As a result, \(r_{\text{min}}(M)\) is an increasing function of the subhalo mass, implying that no subhalos survive within \(r_{\text{min}}(10^{-6} \text{M}_\odot) \sim 200\) pc.

---

\(^2\)Assuming all existing subhalos with mass \(10^{-6} \text{M}_\odot\) form at the 5\(\sigma\) peaks of the density field, Pieri et al. (2008) gives \(c(10^{-6} \text{M}_\odot, z = 0) = 400\).
5.1.4 The subhalo distribution function

Folding these indications together, the subhalo DF is defined such that the number of subhalos with mass in the range \((M, M + dM)\) within some phase-space volume \(d^3x d^3v\) around \((x, v)\) is \(f_{\text{sub}}(x, v; M) d^3x d^3v dM\). It can be further separated as

\[
f_{\text{sub}}(x, v; M) = f_{\text{sub}}(x, v) f_{\text{MF}}^{\text{sub}}(M; r),
\]

where the MF is given in eq. (5.6).

I make an extreme and dubious assumption that the total mass of the DM is all in the form of virialized subhalos, although in reality they only contain a sizable fraction. For the adopted halo with an isotropic velocity field, I use the ergodic Hernquist (1990) DF

\[
f_{\text{sub}}(x, v) = f_{\text{Hsub}}^H(E) = \frac{1}{\sqrt{2} (2\pi)^{3/2} (GM_{200}^H r_s)^3} \left( \frac{1}{1 - q^2} \right)^{5/2} \times \left[ 3 \sin^{-1} q + q(1 - q^2)^{1/2}(1 - 2q^2)(8q^4 - 8q^2 - 3) \right],
\]

where \(q = E r_s / GM_{200}^H\). In the absence of the bulge + MBH, \(f_{\text{sub}}(x, v)\) yields a Hernquist profile.

Putting eqs (5.6), (5.10) and (5.11) together, the detailed subhalo DF has the form

\[
f_{\text{sub}}(E; M) = f_{\text{Hsub}}^H(E) f_{\text{MF}}^{\text{sub}}(M; r),
\]

and should be normalized such that

\[
M_{200}^H = \int_{M_{\text{min}}}^{M_{\text{max}}} f_{\text{MF}}^{\text{sub}}(M; r) M dM \cdot \int_{r_{\text{min}}(M)}^{\infty} 4\pi r^2 dr \cdot \left[ 4\pi \int_0^{\psi(r)} f_{\text{Hsub}}^H(E) \sqrt{2(\psi(r) - E)} dE \right].
\]

5.2 LC flux driven by noisy DM clumps

So far, I have built up a galaxy model which incorporates a bulge + MBH \((M_{\text{bulge}} = 10^{10} M_\odot, r_{\text{bulge}} = 1 \text{kpc}, M_* = 10^8 M_\odot)\), a Hernquist halo \((M_{\text{halo}} = 10^{12} M_\odot, r_{\text{halo}} = 20 \text{kpc})\), and the central MBH has a LC of effective radius \(r_{lc} = 1 \times 10^{-5} \text{pc}\). The rates of feeding MBH due to the
infalling subhalo clumps, can be factorized into a term depending on the properties of an individual subhalo \( M_{lc}(b, V; M) / T_r(b, V) \), and a term depending on the subhalos’ distribution \( N_{sub}(b, V; M) \):

\[
\frac{dF_{sub}^{lc}}{db \, dV \, dM} = \frac{M_{lc}}{T_r} \times N_{sub}(b, V; M). \tag{5.14}
\]

In words, LC refilling rate is the orbit-averaged captured mass for a single event averaged over the number of such events divided by their orbit periods.

### 5.2.1 LC flux from individual clumps

In Chapter 4 I have used \( N \)-body experiments to calibrate how much mass of stars can be deflected into the LC, if the bulge + MBH is perturbed by a satellite of mass \( M \) passing at a pericentre radius \( b \) with a pericentre velocity \( V \). The captured mass can be fitted by the scaling relation \( 4.14 \):

\[
M_{lc}(b, V; M) = 4.7 \cdot M_{\text{bulge}} \left( \frac{r_{lc}}{r_{\text{bulge}}} \right) \left( \frac{M}{M_{\text{bulge}}} \right) \left( \frac{b}{r_{\text{bulge}}} \right)^{-2.0} \left( \frac{V}{V_{\text{bulge}}} \right)^{-1.3}. \tag{5.15}
\]

The orbit-averaged LC refilling rate is

\[
\frac{M_{lc}(b, V; M)}{T_r(E, J^2)}, \tag{5.16}
\]

where \( J = bV \) and \( E = \psi(b) - \frac{1}{2}V^2 \).

### 5.2.2 Averaged LC refilling rate

The number of subhalos, that have mass in the range \((M, M + dM)\), pericentre in the range \((b, b + db)\) and pericentre velocity in the range \((V, V + dV)\), is

\[
N_{sub}(b, V; M) \, db \, dV \, dM = \left[ N_{sub}(E, J^2; M) \frac{\partial(E, J^2)}{\partial(b, V)} \, dM \right] \, db \, dV
\]

\[
N_{sub}(E, J^2; M) \, dE \, dJ^2 \, dM = 4\pi^2 f_{sub}(E; M) T_r(E, J^2) \, dE \, dJ^2 \, dM. \tag{5.17}
\]
To evaluate eq. (5.17), I change the integration variables from \((\cal{E}, J^2)\) to \((b, V)\), using 
\[ |J| \equiv bV \] and 
\[ \cal{E} \equiv \psi(b) - \frac{1}{2} V^2; \] it is then a straightforward generalization of eq. (2.27) for single-mass systems to a system with a spectrum of (subhalo) mass. The physical constraint that \(b\) be the pericentre of an orbit requires \(V^2 > V_{\text{cir}}^2(b) = -b \cdot \partial \psi / \partial b\), means that the Jacobian determinant 
\[ \frac{\partial (\cal{E}, J^2)}{\partial (b, V)} = 2bV \left( \frac{\partial \psi}{\partial b} + V^2 \right) > 0. \] (5.18)

Substituting eqs. (5.16) and (5.17) into eq.(5.14) gives the rate of mass increase – the LC refilling rate as
\[ F_{\text{lc}}^{\text{sub}} = 4\pi^2 \int \int \int_{f_{\text{sub}}(\cal{E}; M) \frac{\partial (\cal{E}, J^2)}{\partial (b, V)}} \cdot M_{\text{lc}}(b, V; M) \, dB \, dV \, dM. \] (5.19)

Working in the \(G_N = 4.31\) halo system with mass unit \(M_{\text{halo}} = 10^{12} M_{\odot}\), length unit \(r_{\text{halo}} = 20 \text{kpc}\), velocity unit \(V_{\text{halo}} = 223 \text{ km s}^{-1}\) and time unit \(T_{\text{halo}} = 8.7 \times 10^7 \text{ yr}\) (see table 2.1), eq. (5.19) has the form
\[ \dot{M}_{\text{lc}}(\tilde{b}, \tilde{V}; \tilde{M}) = A \dot{M} \tilde{b}^{-2.0} \tilde{V}^{-1.3}, \quad A = 1.5 \times 10^{-11}. \] (5.20)

All the dimensionless quantities have a tilde on their symbols.

Further substituting the detailed subhalo DF eq. (5.12) into (5.14), I produce the LC flux as 
\( F_{\text{lc}}^{\text{sub}} = A \tilde{K} \) where 
\[ \tilde{K} = 4\pi^2 \int_{\tilde{M}_{\text{min}}}^{\tilde{M}_{\text{max}}} \int_{f_{\text{MF}}(\tilde{M}) \bar{M} \bar{d}} \int_{bb(\bar{M})} \int_{V_{\text{cir}}(\tilde{b})} \tilde{b} \cdot f_{\text{sub}}(\tilde{\cal{E}}) \frac{\partial (\tilde{\cal{E}}, \tilde{J}^2)}{\partial (\tilde{b}, \tilde{V})} \left( \tilde{b}^{-2.0} \tilde{V}^{-1.3} \right), \] (5.21)

\( \tilde{\cal{E}} = \psi(\tilde{b}) - \frac{1}{2} \tilde{V}^2, \) \( \tilde{J} = \tilde{b} \tilde{V}, \) \( \bar{b}_0(\bar{M}) \) is the tidal radius of a subhalo of mass \( \bar{M}, \) and \( V_{\text{cir}}(\tilde{b}) \) is the circular velocity at \( \tilde{b}. \)

As seen in eq. (5.19), the fuelling rate is proportional to the subhalo mass density, 
\( F_{\text{lc}}^{\text{sub}} \propto \int \dot{M} \mathcal{M} f_{\text{MF}}^{\text{sub}}, \) which ignores how exactly the clumpiness is distributed in mass. For subhalo MF

---

This is a direct consequence of \( M_{\text{lc}}\)-scaling relation: the mass captured into a LC is proportional to the mass of the perturber. And it is in contrast with the \( \gamma\)-ray annihilation rate calculation, which is proportional...
Chapter 5: Rates of MBH feeding in bulge-halo systems

5.3 Discussions

5.3.1 Some modelling uncertainties

The actual LC refilling rates depends on a couple of assumptions made on the infalling clumps:

\[ f_{\text{sub}}^\text{MF} \text{, I simply choose } \beta = 1 \text{ in eq. 5.6.} \]

In order to optimize the prospects of feeding the MBH, besides assuming the subhalo population has a total mass of \(10^{12} M_\odot\), I also adopt a very low mass cutoff, \(M_{\text{min}} = 10^{-6} M_\odot\), together with a very low value for the tidal cutoff radius \(b_{\text{min}}(M) = b_0\). The numerically integrated \(\tilde{K}\) (eq. 5.21) is shown on the right panel of figure 5.2. In physical units, \(F_{\text{lc}}^\text{sub} = F_{\text{lc}}^\text{sub} \cdot (M_{\text{halo}}/T_{\text{halo}}) = 5.2 \times 10^{-5} M_\odot \text{ yr}^{-1}\), is shown with an asterisk in figure 5.3.

Figure 5.2: Left panel: Median virial concentration-mass relation at \(z = 0\) (eq. 5.8): high mass halos with mass \(10^{11-15} h^{-1} M_\odot\) is shown in thick curve (Neto et al. 2007); a naive extrapolation at low-mass end (thin curve) gives an overestimate, e.g., \(c(10^{-6} M_\odot, z = 0) \approx 500\). Right panel: results for integration of \(\tilde{K}\) in eq. 5.21.

An overall underestimate of the cutoff radius \(r_{\text{min}}(M)\) allows massive perturbers with low concentration to come arbitrarily close to the centre. This effectively adds more noise to the density squared and depends sensitively on the detailed MF(Diemand et al. 2007a).
central region.

[ii] Virtually the total DM halo mass is assumed to be in the form of substructures. This gives an over-optimistic amount of noise-driven flux.

[iii] The subhalo spatial distribution is assumed to follow its underlying smooth component, with a $r^{-1}$ cusp in the centre. However, a shallower inner profile (Reed et al. 2005) is more favored, and it would give a smaller LC flux.

[iv] An isotropic velocity field is a reasonable assumption, although simulated halos typically have an velocity anisotropy parameter $\beta \simeq 0.6$ (van den Bosch et al. 1999). On one hand, subhalos on more radial orbits tend to spend less time at their pericentres which would suppress the LC refilling. On the other hand, unlike the ergodic Hernquist (1990) DF, more subhalos would be put onto eccentric orbits; this may increase the averaged LC refilling rate.

[v] The total mass of the DM halo, e.g., the Galactic halo, can be as small as $2 \times 10^{11} M_\odot$ or as large as $5 \times 10^{12} M_\odot$ (Sec. 1.1 BT08); this can change the fuelling rates by a factor $0.5 - 2$.

To sum up, points (i-iii) only lead to an overestimate of $F_{\text{sub}}^{\text{lc}}$; points (iv,v) may vary the rates by a factor of unity.

### 5.3.2 Some theoretical investigations

Below, I discuss some typical LC refilling rates.

(a) $F_{\xi}^{\text{lc}}$ for a full LC

The flux of stars per unit energy interval streaming through a surface of constant $J^2 = J_{\text{lc}}^2$, were the stars full upto some maximum energy $\mathcal{E}$, is given by

$$F_{\xi}^{\text{lc}} = \int_0^\mathcal{E} 4\pi^2 J_{\text{bulge}}^2(\mathcal{E}') J_{\text{lc}}^2(\mathcal{E}') d\mathcal{E}',$$

(5.22)

and is shown with the solid curve in figure 5.3.

(b) $F_{\text{2body}}^{\text{lc}}$ by two-body relaxation

In Sec. 4.1 I have reviewed the local Fokker-Planck treatment of diffusion in $X^2$. Now, I calculate $\Pi$ (eq. 4.7) using the numerical method described in Appendix A.2. It is then
Figure 5.3: MBH fuelling rate driven by different mechanisms: solid curve measures the flux of stars through a surface of constant $J^2(J_{lc}^2)$ in eq. (5.22), assuming stars are populated up to an energy level $E(r_E)$; the asterisk indicates the DM clumps driven $F_{lc \text{sub}}$ and the square indicates the steady-state two-body relaxation driven $F_{lc \text{2body}}$. $F_{lc \text{sub}} = 3F_{lc \text{2body}}$ can be interpreted as DM noise is just as competitive as relaxation in terms of causing stars to diffuse in phase-space and feeding the LC. One can see both mechanisms keep the LC full up to an energy level inside the MBH ($10^8 M_\odot$) sphere of influence (vertical line at $r_E/r_{\text{bulge}} = 0.16$).

From results presented in figure 5.3, one learns that LC refilling driven by gravitational perturbations from subhalos: (1) populates the LC at a higher rate than the two-body relaxation effects by a factor of 3; (2) keeps the LC full up to an energy level well inside the MBH sphere of influence (thick vertical line). (3) However, the slightly enhanced LC flux should be interpreted as: $F_{lc \text{sub}}$ can be comparable to $F_{lc \text{2body}}$. Therefore, the scarcity of subhalo population near the galactic centre makes this mechanism less interesting in the context of feeding single stars to the MBH.

The most relevant work on massive perturbers (MP) driven LC refilling rate would be Perets et al. (2007), who extended the Fokker-Planck LC formalism to approximately account for relaxation by rare encounters with massive perturbers. By compiling the MP mass function from
published observations they showed that, relative to stellar two-body relaxation alone, MPs (such as giant molecular clouds or open clusters of masses $10^{3-8} M_\odot$ outside the inner few pc) dominate and accelerate relaxation in the inner $\sim 100$ pc of the Galaxy centre. MPs will not contribute much to the disruption (LC refilling) rate of single stars, since stellar two-body encounters are efficient enough to replenish the LC. Also, due to the uncertainties in determining the MP distribution on the smallest scales, the enhancement of the event rates is very uncertain.

Compared to their work, I investigate subhalo perturbers that exist much further from the Galactic centre ($\geq 200$ pc); I also adopt a $\sim 25$ times heavier MBH in the Galaxy model. Nevertheless, my subhalo-enhanced LC refilling rate $5.2 \times 10^{-5} M_\odot \text{yr}^{-1}$, is comparable with Perets et al.’s (2007) MPs-induced disruption rate of single stars, ranging between $2.6 \times 10^{-8} M_\odot \text{yr}^{-1}$ and $1.1 \times 10^{-4} M_\odot \text{yr}^{-1}$. 
Chapter 6

Conclusions and future studies

The motivation behind this thesis work was to improve our understanding of the internal dynamics at galactic centres, in particular, around super massive black holes. The best available tool on the dynamical study is undoubtedly full self-consistent $N$-body simulations. We have (1) developed a multi-mass scheme for constructing collisionless $N$-body models and (2) used the multi-mass bulge + MBH model to study loss cone dynamics. Below, I summarize results in earlier chapters and suggest some future work.
6.1 Constructing collisionless galaxy models

In Chapter 2, I presented the basic dynamical equations that govern the evolution of collisionless systems. Beginning with the collisionless Boltzmann equation that is of essential importance for system (Sec. 2.1.2), I show explicitly how the CBE is related to the N-body modelling (Sec. 2.2). When encounters are taken into account, one writes the collision term in Master equation form and expands in a Taylor series to derive the Fokker-Planck equation (Sec. 2.3).

Rather technical, Chapter 3 forms the backbone of the thesis. In it we saw how to design a sampling distribution function (DF) $f_s$ from some known DF $f_0$ (Sec. 3.2.2) and how to generate a multi-mass collisionless model from it (Sec. 3.2.3). In two applications, a bulge model and a bulge + MBH model, we aim at minimizing the shot noise in estimates of the acceleration field. Models constructed using our multi-mass scheme easily yield a factor $\sim 100$ reduction in the variance at the central acceleration field when compared to a traditional equal-mass model with the same number of particles. When evolving both models with an N-body code, the diffusion coefficients in our model are reduced by a similar factor. Therefore, for certain types of problems, our scheme is a practical method for reducing the two-body relaxation effects, thereby bringing the N-body simulations closer to the collisionless ideal. We note the following preparation for the applications of multi-mass modelling scheme:

- For successful application, a system should be in a steady state, or close to one.

- The DF $f_0$ should be quick and cheap to evaluate, either numerically or analytically. Finding $f_0$ for axisymmetric or triaxial galaxies is a longstanding and nontrivial problem since one rarely has sufficient knowledge of the underlying potential’s integrals of motion, but suitable flattened DFs do exist, including the standard axisymmetric two-integral $f(\mathcal{E}, L_z)$ models and also rotating triaxial models such as those used in, e.g., Berczik et al. (2006) An alternative way of constructing flattened multi-mass realizations would be to apply Holley-Bockelmann et al. (2002)’s adiabatic sculpting scheme to a spherical N-body model constructed using our scheme.

- The general multi-mass scheme uses importance sampling to find the tailored sampling DF $f_s$
that minimizes the sum of mean-square uncertainties in \(Q_i\) (of the form eq. 3.4). As long as \(f_s\) is smooth in integral space, Monte Carlo realizations of \(f_s\) should work for any reasonably general collisionless \(N\)-body code. The utility of our multi-mass scheme, therefore, depends critically on the selection of the projection kernels \(Q_i(w)\).

The last point is new to this field. It is probably best addressed by experimenting with different sets of kernels, especially since it is easy to test the consequences of modifying them. Nevertheless, there are cases in which modest physical insight offers some guidance on choosing the \(Q_i\). Besides the loss cone problems addressed in Chapter 4, I here give another example for future work:

\(\heartsuit\)1. Sinking satellites Kazantzidis et al. (2004) demonstrate the significance of using equilibrium \(N\)-body realizations of satellite models when investigating the effect of tidal stripping of DM substructure halos (satellites) orbiting inside a more massive host potential. Besides the shape of the background potential and the amount of tidal heating, the mass-loss history is very sensitive to the detailed density profile of the satellite itself. One can therefore make one step further from equal-mass realizations by designing kernels to pick out orbits that pass through the tidal radius, while again maintaining an accurate estimate of the satellite’s acceleration field.

### 6.2 Rates of loss cone refilling by MBHs

Chapter 4 describes applications of a multi-mass bulge + MBH model to study the loss cone (LC) dynamics. In Sec. 4.1 I introduced a loss cone consumption sphere to mimic tidal disruption events. The “empty” nature of the LC was then scrutinized by calibrating artificial LC refilling (Sec. 4.2.1). Toy models in Sec. 4.2.2 taught us that only a fairly massive satellite \(\left(\frac{M_s}{b}\right)\) passing through the centre \((b)\) with low speed \((V)\) can give rise to a LC flux well above the threshold of numerical noise-driven LC refilling. From the reduced satellite parameter space, about 100 experiments have been carried out to measure the maximum net mass \(\left(\frac{M_{lc}}{b}\right)\) captured into an initially empty LC. The collected data were summarized by a fitting \(M_{lc}\)-formula \(4.18\) in Sec. 4.2.3.
• Stellar orbits become strongly perturbed during a satellite flyby. The LC-captured mass reaches a peak value soon after the perturber visits its pericentre.

• $M_{lc}$ is strongly correlated with the satellite parameters: scaling linearly with the satellite mass ($M_s$) and loss cone size ($r_{lc}$), but inversely with $b^{2.0}$ and $V^{1.3}$.

• As the satellite departs, stars settle down to a new equilibrium configuration within a few crossing times. Therefore, if perturbed by cumulative perturbers, one can simply accumulate each single event to get the total consumed mass.

The rest of the Chapter 4 is devoted to understanding numerical results: where do LC stars come from? Linear perturbation theory claims that resonances are important if not most important in determining the response of a stellar system to any external perturbation (Sec. 4.3.2). But for the flyby encounters of interest, could resonances be the culprit for redistributing angular-momentum in stars and eventually send stars to the LC? This was addressed in Sec. 4.3.3 where test-particle experiments were carried out to track resonant dynamics relevant to a periodic driver. I make the following summaries:

• Close encounters with the satellite effectively change angular-momentum of stars. The ones that lose enough angular-momentum can be effectively deflected into the loss cone and destroyed by falling into the MBH.

• A flyby encounter does not repeat its pattern, which prevents stars in instantaneous resonance from building up a significant response. Therefore, resonance coupling between stars and the external perturber is unimportant in terms of refilling the emptied loss cone.

• The accelerated angular-momentum relaxation within a nearly Keplerian is also unimportant here.

In Chapter 5 I introduced an outer halo as a reservoir to continuously inject orbiting subhalos into the inner bulge. One further step was made by describing the subhalo properties in terms of phase-space probability density, containing all the dynamical information (Sec. 5.1). Finally in Sec. 5.2 I put things together to predict a LC refilling rate. The conclusion is:
• The flux of stars into the loss cone is **enhanced** when the loss cone is initially emptied, but due to the scarcity of subhalo population near the galactic centre, the LC refilling rate averaged over the entire orbiting dark halo substructures is **not** strongly affected.

**The future**

For the galaxy model being used in the entire thesis, I have introduced several simplifications. At this point, I bring some notes about possible improvements and extensions of the current model.

** وليس** Better galaxy models The galaxy model used in this thesis is not ideal: (a) the mass of the MBH is considerably larger than those observed in galactic centres and the effective radius \( r_{\text{lc}} \) of the LC is many orders of magnitude larger than the tidal radius in real galaxies. Unfortunately, it is still hard for simulations to determine accurately the flux of stars into a small MBH or a tiny LC. The only remedy is to increase the number of particles, the more the better. (b) The adopted Hernquist profile has a \( r^{-1} \) central cusp and no alternatives, such as core (shallower) or nucleus (steeper) profiles, have been explored here. It would be useful to extend the work in this thesis by building \( N \)-body models with different mass profiles, MBH masses and LC sizes.

** وليس** Evolution of binary supermassive black holes Larger elliptical galaxies and bulges grow through mergers. If more than one of the progenitors contains a MBH and the in-spiral time is less than a Hubble time, the MBHs will form a bound system. Begelman, Blandford & Rees (1980) showed that the evolution of binary MBHs in gas-poor galaxies can be divided into three phases: (a) As galaxies merge, the core undergoes violent relaxation. Via dynamical friction (Chandrasekhar 1943), the captured MBH sinks towards the center of the common gravitational potential where they form a bound binary. (b) The binary continues to decay via gravitational slingshot interactions (Saslaw, Valtonen & Aarseth 1974): stars on orbits intersecting the binary are ejected at much higher velocities comparable to the binary’s orbital velocity. (c) Finally **if** the binary’s separation decreases to the point at which gravitational radiation becomes the dominant dissipative force to carry away the last remaining
energy and angular momentum, the binaries coalesce rapidly (Peters 1964). The long-term evolution of MBH binaries is always related to a long standing problem called the “final parsec problem” (Milosavljević & Merritt 2003b): is super-elastic scattering off individual stars in the background efficient enough to transit a binary MBH from dynamical friction regime on the order of 1pc scale, to the gravitational radiation dominated regime with a separation \( \leq 10^{-3}\text{pc} \) and finally merge within a Hubble time? Uncertainties about the resolution have been a major impediment to predicting the frequency of MBH mergers in galactic nuclei, and hence to computing event rates for proposed gravitational wave interferometers like LISA.\(^1\)

As a natural continuation of the work in Chapter 5, I intend to further explore the use of \( M_{\text{lc}} \)-fitting-formula for the evolution of binary MBHs.

\(\triangledown\)4. Response of galaxy to perturbations

Our conclusion that resonances are unimportant in redistributing angular-momentum within stars, is in disagreement with Vesperini & Weinberg (2000), who explore the effects of relatively weak encounters by computing the response of a spherical stellar system to the perturbation induced by a low-mass system during a flyby. They determine the contribution to the total self-consistent response from three sources: (a) the perturbation applied by the external perturber; (b) the reaction of the system to its own response to this perturbation; (c) the excitation of discrete damped modes of the primary system; and conclude that it is the resonances (point c) that lead to the excitation of patterns in the primary system. The main discrepancies probably stem from the galaxy model under study, since all their conclusions are based on a very strong assumption: discrete modes are weakly damped in \( \text{King} \) models with different concentrations. On the theoretical side, I intend to explore the response of galaxy to external perturbations by means of linear perturbation theory, hopefully to find out whether weakly-damped modes exist in every galaxy.

\(^1\)http://lisa.jpl.nasa.gov/


[23] Danzmann K. 2003, Advances in Space Research, 32, 1233


Bibliography

[34] Ferrarese L., & Ford H. 2005, Space Science Reviews, 116, 523


[63] Kormendy J. 2004, Coevolution of Black Holes and Galaxies, 1


[82] Merritt D., & Milosavljević, M., 2005, Living Reviews in Relativity 8, 8


[99] Rauch K. P., & Tremaine S., 1996, New Astronomy, 1, 149


[139] Zel’Dovich Y. B. 1964, Soviet Physics Doklady 9, 195


Appendix A

Numerical solutions

A.1 Spherical isotropic DFs for density-potential pair

To recover an ergodic DF that generates a model with the given density $\rho(x) = \int f \, d^3v$ but in a non-self-consistent spherical potential $\nabla^2 \Phi \neq 4\pi G \rho$, we use Eddington’s formula (eq. 4.46 BT08)

$$f(\mathcal{E}) = \frac{1}{\sqrt{8\pi^2}} \left[ \int_{0}^{\mathcal{E}} \frac{d\psi}{\sqrt{\mathcal{E}-\psi}} \frac{d^2 \rho}{d\psi^2} + \frac{1}{\sqrt{\mathcal{E}}} \left( \frac{d\rho}{d\psi} \right)_{\psi=0} \right].$$

(A.1)

At large radii, the second term on the RHS of eq. (A.1) vanishes for any sensible behavior of $\psi$ and $\rho$. However, the second order differentiation $d^2 \rho/d\psi^2$ in the first term would be difficult to deal with numerically. Instead, I calculate its alternative

$$\frac{d^2 \rho}{d\psi^2} = \left[ -\left( \frac{d\psi}{dr} \right)^{-3} \frac{d^2 \psi}{dr^2} \right] \frac{d\rho}{dr} + \left( \frac{d\psi}{dr} \right)^{-2} \frac{d^2 \rho}{dr^2},$$

(A.2)

where all the derivatives can be evaluated analytically (or numerically). Thus, the construction of isotropic distribution function in eq. (A.1) reduces to a numerical quadrature with no differentiation required.

For the bulge + MBH model where the bulge is a Hernquist model of mass $M_\ast = 1$ and $^1\psi = 0$ when $r \to \infty$, on the other hand as $r \to \infty$, $\rho \propto r^{-4} \propto \psi^4$. So we have $\frac{d\rho}{dr} \propto \psi^3 \mid_{\psi=0} = 0$.  

\[1\]
scale length $r_\star = 1$, the MBH is a Plummer model of mass $M_\bullet = 0.01$ and scale length $\epsilon_\bullet = 0.003$, the mass ratio between MBH mass to the galaxy mass $\mu = M_\bullet / M_* = 0.01$. The density and potential pair are:

$$
\rho(r) = \frac{1}{2\pi r(r+1)^3}, \quad \psi(r) = \frac{G_N}{r+1} + \frac{G_N \mu}{\sqrt{r^2 + \epsilon_\bullet^2}}.
$$

where $G_N = 4.31$ is the chosen gravitational constant in the model system. The derivatives in eq. (A.2) have the following forms:

$$
\frac{d\rho}{dr} = \frac{1}{2\pi} \frac{4r+1}{r^2(1+r)^4}, \quad \frac{d^2\rho}{dr^2} = \frac{1}{2\pi} \frac{20r^2 + 10r + 2}{r^3(1+r)^5},
$$

$$
\frac{d\psi}{dr} = -G_N \left[ \frac{\mu}{r^2} + \frac{1}{(r+1)^2} \right],
$$

$$
\frac{d^2\psi}{dr^2} = -4\pi G_N \rho(r) - 2 \frac{d\psi}{dr},
$$

$$
\frac{d^2r}{d\psi^2} = -\left( \frac{d\psi}{dr} \right)^{-3} \frac{d^2\psi}{dr^2}.
$$

### A.2 Diffusion coefficients

adapted from Appendix B MT99

In this Appendix we derive expressions for the diffusion coefficients used in equation 4.4 in the limit $X^2 \to 0$. Because of the presence of the loss cone, the steady-state distribution of scatterers is not quite isotropic. It is, however, reasonable to calculate the diffusion coefficients using the isotropized distribution function

$$
\bar{f}(E) \equiv \int_0^1 f(E, X^2) dX^2.
$$

For the bulge + MBH model, I simply use the numerically founded DF $f(E)$.

We make the reasonable approximation that all encounters take place instantaneously and so change the scattered star’s velocity but not its position. In addition, we make the usual (though more dubious) assumption that the distribution of scatterers is homogeneous in space.
Appendix A: Numerical solutions

Since \( X^2 = r^2 v_t^2 / J_e^2(\mathcal{E}) \), where \( v_t^2 = v_\phi^2 + v_\theta^2 \), we can immediately use equation (7.88) of Binney & Tremaine (2008) to show that

\[
\langle \Delta X^2 \rangle = \frac{32\pi^2 r^2 G^2 m_\star^2 \ln \Lambda}{3 J_e^2} \left[ 3 I_{\Delta} - I_{\Delta} + 2I_0 \right] + O(X^2),
\]

where

\[
I_0 = \int_0^\mathcal{E} \bar{f}(\mathcal{E}') \, d\mathcal{E}'.
\]

\[
I_{\Delta} = \left[ 2 (\psi(r) - \mathcal{E}) \right]^{- \frac{1}{2}} \int_0^{\psi(r)} \left[ 2 (\psi(r) - \mathcal{E}') \right]^{- \frac{1}{2}} \bar{f}(\mathcal{E}') \, d\mathcal{E}',
\]

and \( \ln \Lambda \) is the usual Coloumb logarithm. We follow Spitzer & Hart (1971) and take \( \Lambda = 0.4 M_\star / m_\star \).

The second-order diffusion coefficient is \( \langle (\Delta X^2)^2 \rangle = r^4 (\Delta v_t^2)^2 / J_e^4(\mathcal{E}) \). Since

\[
(\Delta v_t^2)^2 = 4 v_\phi^2 (\Delta v_\phi)^2 + 4 v_\theta^2 (\Delta v_\theta)^2 + 8 v_\theta v_\phi (\Delta v_\theta)(\Delta v_\phi) + O((\Delta v)^3),
\]

it follows from equations (7.88) and (7.89) of Binney & Tremaine (2008) that

\[
\langle (\Delta X^2)^2 \rangle = X^2 \frac{64\pi^2 r^2 G^2 m_\star^2 \ln \Lambda}{3 J_e^2} \left[ 3 I_{\Delta} - I_{\Delta} + 2I_0 \right] + O(X^4).
\]

Notice that \( \langle \Delta X^2 \rangle = \frac{1}{4} \partial \langle (\Delta X^2)^2 \rangle / \partial X^2 \), the orbit-averaged version of which holds generally whenever the scattering is done by an external perturbation (Landau 1937; Binney & Lacey 1988).