

The formation of the Galactic halo

The formation of the Galactic halo

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Dr. W.A. Wagenaar,
hoogleraar in de faculteit der Sociale Wetenschappen,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 28 juni 2000
te klokke 14:15 uur

door

Amina Helmi
geboren te Bahía Blanca, Argentinië in 1970

PROMOTIECOMMISSIE:

Promotor: Prof. dr. P. Tim de Zeeuw
Prof. dr. Simon D.M. White (Max Planck für Astrophysik, Garching)

Referent: Prof. dr. K.C. Freeman (Australian National University)

Overige leden: Prof. dr. W. B. Burton
Prof. dr. M. Franx
Prof. dr. K. H. Kuijken (Rijksuniversiteit Groningen)
Prof. dr. G. K. Miley

A mis padres

Front cover: *Leiden in the Halo*, by Mariano Méndez.

The galaxy visible here is NGC 1232, in the constellation Eridanus. Similar to our Galaxy in its spiral structure, it is almost twice as large. It lies about 100 million light-years away from us. This image was obtained with the ESO Very Large Telescope (VLT) of 8.2 m located in Paranal (Chile). Credit: European Southern Observatory.

Contents

1	Introduction and outline	1
1.1	Galaxy formation in a hierarchical Universe	2
1.2	The Milky Way galaxy	4
1.2.1	The Galactic components	4
	The thin disk	4
	The thick disk	5
	The bulge	6
1.2.2	The Galactic halo	6
	Some observational characteristics of halo field stars	6
	Globular clusters	7
	Key observables and formation scenarios	8
1.3	Outline of this thesis	10
1.4	A possible scenario for the formation of our Galaxy	12
1.5	Next steps	14
2	Building up the stellar halo of the Galaxy	19
2.1	Introduction	20
2.2	The simulations	21
2.2.1	Model	21
2.3	Properties of the debris: Simulations	24
2.3.1	Entropy as a measure of the phase-mixing	24
2.3.2	Configuration space properties	26
2.3.3	Velocity space properties	27
2.4	Properties of the debris: Analytical approach	28
2.4.1	Action-Angle variables and Liouville’s theorem	28
	The evolution of the distribution function	30
	The distribution function in observable coordinates	32
2.4.2	Spherical potential	34
	Analytic predictions	34
	Comparison to the simulations	36
2.4.3	Axisymmetric case	38
2.4.4	What happens if there is phase-mixing	39
	The filling factor	40
	Properties of an accreted halo in the Solar neighbourhood	42
2.5	An observational application	44
2.6	Discussion and conclusions	45

Appendix A	Spherical potential	48
Appendix B	Axisymmetric Eddington potential	52
Appendix C	Axisymmetric Stäckel potential	53
3	Debris streams in the Solar neighbourhood as relicts from the formation of the Milky Way	57
4	Mapping the substructure in the Galactic halo with the next generation of astrometric satellites	65
4.1	Introduction	66
4.2	Building up a stellar halo	67
4.2.1	Initial conditions for the satellites	67
	Orbital properties	67
	Internal properties of the satellites	68
4.2.2	Galactic potentials	69
4.2.3	Numerical methods	70
4.2.4	Generating catalogues of halo stars	70
4.3	Finding disrupted galaxies	71
4.3.1	Integrals of motion space	71
4.3.2	Method: FOF in integrals of motion space	75
4.3.3	Results	75
4.4	Clumpiness in the kinematics of halo stars	76
4.5	Discussion	80
5	Simple dynamical models of the Sagittarius dwarf galaxy	83
5.1	Introduction	84
5.2	Method	84
5.3	Results	86
5.3.1	Constant mass-to-light ratio: A purely stellar model	86
5.3.2	Varying mass-to-light ratio: A model with a dark halo	88
5.4	Discussion	90
6	The phase-space structure of a cluster halo: Insights into our Galactic halo	95
6.1	Introduction	96
6.2	Methodology	98
6.2.1	The simulations	98
6.2.2	The cosmology	98
6.2.3	Scaling to a Milky Way halo	99
6.3	Results	100
6.3.1	Phase-space evolution of halo debris	100
6.3.2	The Solar neighbourhood	104
6.3.3	The filling factor	107
6.4	Discussion	109
	Summary	113
	Nederlandse samenvatting	117
	Curriculum Vitae	121
	Acknowledgments	123

List of Figures

1.1	Schematic representation of the growth of a halo in time as a merger tree	3
1.2	The luminosity function of the Milky Way from semi-analytic modelling	5
1.3	Kinematics of a sample of metal-poor red giants and RR Lyrae stars in the Solar neighbourhood	8
1.4	Galaxy distribution around a “Milky Way” halo from a high-resolution simulation of a cluster combined with semi-analytic modelling of galaxy formation . .	13
2.1	Projections of the different orbits of the satellites for our six principal simulations	23
2.2	Evolution of the entropy of the system for different experiments	24
2.3	Isodensity surfaces encompassing most of the satellite’s mass, 14 Gyr after infall	26
2.4	Sky projection of satellite debris after 14 Gyr of evolution	27
2.5	Scatter plots of the different velocity components for particles at various locations in the Galaxy, 14 Gyr after infall	29
2.6	1-D graphical interpretation of the evolution of the system in phase-space	32
2.7	Time evolution of the velocity ellipsoid and the density of satellite debris moving in a spherical potential	36
2.8	Time evolution of the density along a stream for a satellite moving in a spherical potential: Predictions vs. Simulations	37
2.9	Time evolution of the velocity ellipsoid and the density of satellite debris moving in an axisymmetric potential	39
2.10	Time evolution of the filling factor for a satellite with an initial velocity dispersion of 25 km s^{-1} and 1 kpc size moving in the Galactic potential	42
3.1	Scatter plots of the different velocity components for our sample of nearby halo stars and for one Monte Carlo realization	60
3.2	The distribution of nearby halo stars in the plane of angular momentum components, J_z vs. $J_\perp$, compared to one Monte Carlo realization	61
3.3	The distribution of nearby halo stars in velocity space and in the $J_z - J_\perp$ plane with the candidates for our detected substructure highlighted	62
3.4	The distribution of particles in the vicinity of the Sun in velocity space and in the $J_z - J_\perp$ plane for one of our simulations of a disrupted satellite	63
4.1	Circular velocity profile for two different Galactic potentials	70
4.2	Number counts profile $N(r) = r^2 \rho(r)$ for the simulated stellar halo resulting from the superposition of 33 disrupted satellite galaxies, 12 Gyr after their infall . . .	71
4.3	Initial distribution of the satellites particles in the integrals of motion space . . .	73

4.4	Final distribution of the satellites particles in the integrals of motion space 12 Gyr after infall, after convolution with the errors expected for GAIA for the original potential	73
4.5	Final distribution of particles in the $L_z - E$ space after error convolution for FAME and for DIVA, using the original potential	73
4.6	Final distribution of particles in the $L_z - E$ space after GAIA error convolution for the alternative potential.	73
4.7	Lumps detected with our FOF algorithm using two different Galactic potentials, in the cases of simulated GAIA and FAME catalogues	77
4.8	Velocity space distribution for particles in a cubic volume of 2 kpc on a side centred on the Sun for one realization of the GAIA catalogue	77
4.9	Two-point correlation function in velocity space for “giant stars” inside spheres of 1 kpc radius around the Sun	79
5.1	Orthogonal projections of a possible orbit of Sagittarius	86
5.2	Distribution of particles in distance from the Sun as a function of latitude	87
5.3	Surface isodensity contours for the remnant system	88
5.4	Mean heliocentric radial velocity and dispersion as a function of Galactic latitude across the remnant system	89
5.5	Mass profile for Model II and for its “luminous” component	90
5.6	Sky distribution, heliocentric distance and heliocentric radial velocity as functions of Galactic latitude for the particles in our constant mass-to-light ratio model of Sagittarius, 12.5 Gyr after infall. Different colours indicate material stripped off in different passages	91
5.7	Number counts in $5 \times 5 \text{ deg}^2$ fields in different directions, normalized to the main body of Sagittarius	93
6.1	The circular velocity $V_c(r) = \sqrt{GM(r)/r}$ of the cluster for our simulation with 13 million particles.	99
6.2	A schematic representation of the merger history of the cluster.	100
6.3	Ratio of the mass of halos directly accreted with respect to the mass of the cluster main progenitor at the time of identification.	101
6.4	The phase-space behaviour of a debris stream in our S3 simulation	102
6.5	The phase-space behaviour of a debris stream in a static NFW potential consistent with the present day properties of the S3 cluster	103
6.6	Principal axes projections of the velocity of particles located in a box of 2 kpc on a side at the Solar radius	105
6.7	Same as previous figure but colours here indicate particles originating in the same halo 11 Gyr ago.	105
6.8	The mass-weighted number of streams in boxes of 2.72 kpc on a side for our simulation S4 as a function of distance from the centre of the Galactic halo. . . .	108
6.9	Mass-weighted velocity dispersions of streams inside cubic volumes of 2.72 kpc on a side, located in the inner Galactic halo.	109
6.10	Number of particles in boxes of 1.36 kpc on a side, containing a ten times smaller mass-weighted number of streams; the mass-weighted velocity dispersion of streams in units of the 3D velocity dispersion; and the same velocity dispersion in km s^{-1}	110

List of Tables

2.1	Orbital parameters for the different experiments of our six principal simulations	22
2.2	Characteristic densities of the debris at different times	28
4.1	Estimated precision in parallax and proper motion as a function of V magnitude for DIVA, FAME and GAIA	72
5.1	Properties of Sagittarius	84
6.1	Parameters of the high-resolution simulations	98

*And what is this unknown island you
want to go in search of, If I could tell
you that, it wouldn't be unknown ...*
J. Saramago: "The tale of the unknown island"

Chapter 1

Introduction and outline

*"G*ALAXY formation is simple in principle. One starts with a mass distribution and an associated velocity field at some time in the early Universe, and evolves it. The system will collapse around local over-dense regions, will cool, mix, form stars and evolve" (Gilmore 1993). Every now and then, an astronomy professor will tell his or her graduate student that this is the right time to embark into the seemingly simple but yet ambitious task of trying to unveil how our own Galaxy, the Milky Way, formed. With similar words, this is what Simon White told me during a meeting in Munich in March 1997. Three years later I present this thesis which is the result of one such attempt, certainly not the first and most likely not the last one.

Over the last twenty years a standard paradigm has emerged for the formation and evolution of structure in the Universe (Peebles 1974; White & Rees 1978). Fundamental cornerstones in this paradigm are the discovery of the (Hubble) expansion of the Universe in the early twenties, and of the cosmic microwave background radiation in the sixties (Penzias & Wilson 1965; Dicke et al. 1965), which along with observations of large-scale structure show that the Universe is fairly homogeneous and isotropic on very large scales. Put together, these ingredients lead to a standard cosmological model: an isotropic and homogeneous Universe that evolved from an initially hot and dense state. We may understand the homogeneity on very large scales if the entire observable Universe was causally connected at some early epoch. This would be possible, for example if the Universe went through an inflationary phase in its early expansion (Guth 1981; Linde 1982). Inflation models through quantum fluctuations provide a Gaussian field of density fluctuations: the seeds from which all the structure we observe today formed. Indeed, such small fluctuations are imprinted on the cosmic microwave background radiation (CMB), as demonstrated by the COBE satellite (Smoot et al. 1992). COBE not only confirmed the high degree of homogeneity of the CMB, it also detected temperature fluctuations $\Delta T/T \sim 10^{-5}$ on a scale of 10° consistent with a power law spectrum of Gaussian initial density fluctuations. However, at the moment the different CMB experiments are not yet able to determine, for example, whether the density inhomogeneities had their origin in quantum fluctuations during inflation or in topological defects, such as cosmic strings. We may have to wait for new satellite missions such as MAP and Planck, to measure the precise values of all cosmological parameters and the spectrum of fluctuations that gave rise to the present day structure.

Another key ingredient of the standard paradigm is that most of the mass in the Universe is dark. This is based on the fact that the luminous (baryonic) matter in the Universe alone cannot account for its Hubble expansion rate, nor for the motions of stars and gas in the outer parts of galaxies, or of galactic systems in clusters. Probably favoured by the particle physics Standard Model, most likely candidates are, yet to be detected, nonbaryonic weakly interacting elementary particles.

The paradigm described sets the framework to tackle the problem of how galaxies form from

very definite initial conditions: one only needs to specify the shape of the power spectrum, the cosmological parameters and the dominant form of dark matter. Currently preferred are the “hierarchical” models. These assume that the spectrum of density fluctuations, as produced during inflation, is close to scale invariant, and that the dominant matter component in the Universe is in the form of *massive* weakly interacting particles (Peebles 1982). Because of their mass, these particles had non-relativistic motions at the time they decoupled from the expansion of the Universe, which is why they are known as cold dark matter particles (CDM). The actual shape of the CDM power spectrum is such that structures on the small scales collapse first. Therefore, in hierarchical models structure is built bottom-up, as dark matter aggregates into larger and larger clumps (halos) when the gravitational forces amplify the initially small density fluctuations. Galaxies then arise as the gas associated with these dark halos cools and collects in their cores, where it can form stars, which heat and enrich the remaining gas present in a disk or halo (White & Frenk 1991).

The dynamics and evolution of the dark matter halos can be studied in these models in a fairly straightforward fashion, either through analytic approximations or numerical simulations. These now have very high resolution, and the structure of halos can be determined in great detail (Navarro, Frenk & White 1996), as can be their substructure (Klypin et al. 1999; Moore et al. 1999) and internal dynamics (Ghigna et al. 1998; Tormen, Diaferio & Syer 1998). The nonlinear growth of structure on scales greater than or of the order of a kiloparsec is thus well-studied, and reasonably well understood being simply dominated by gravitational dynamics. Below these scales it is clear that other nonlinear processes such as star formation and feedback are dominant. Understanding these processes, which affect the behaviour of the baryonic components, is thus extremely important for making a sensible comparison of the outcome of the hierarchical models with the observational data. Unfortunately, this is not the case and thus such processes are modelled only crudely. Numerically it is very difficult to simulate the formation of a galactic system in the proper cosmological context because of the large dynamic range on which the different relevant processes are acting at the same time (but see some attempts by Steinmetz & Navarro 1999; and Pearce et al. 1999). In recent years considerable progress has been made through a phenomenological approach, which combines simple, physically motivated prescriptions for the behaviour of the baryonic components, with analytic or numerical methods that describe the dark matter evolution (e.g. Kauffmann, White & Guiderdoni 1993; Baugh, Cole & Frenk 1996). In what follows we shall discuss the basic ingredients and relevant processes in the build-up of galactic systems in hierarchical models, thus setting the scene for tackling the problem of how our own Galaxy formed.

1.1 Galaxy formation in a hierarchical Universe

When an overdense region decouples from the expanding Universe, turns around and collapses onto itself it forms a bound structure: a halo. To a great extent the evolution of galaxies is dominated by the dark halos dynamics and histories. Therefore we need to know, for example, the rates of halo merging and how these rates evolve in time and depend on environment. Assuming an initial random Gaussian field of fluctuations, Press & Schechter (1974) derived an analytic approximation for the mass distribution of nonlinear objects at a given time. Starting from the Press–Schechter formalism (or any of its extensions e.g. Bond et al. 1991; Bower 1991), one may derive merger histories, the distribution of objects of mass M at time t , the survival times, etc. (Lacey & Cole 1993). All this information is contained in what are called “merger trees”, such as the one shown in Figure 1.1.

Once we specify the cosmological parameters: the density of the Universe (in units of the

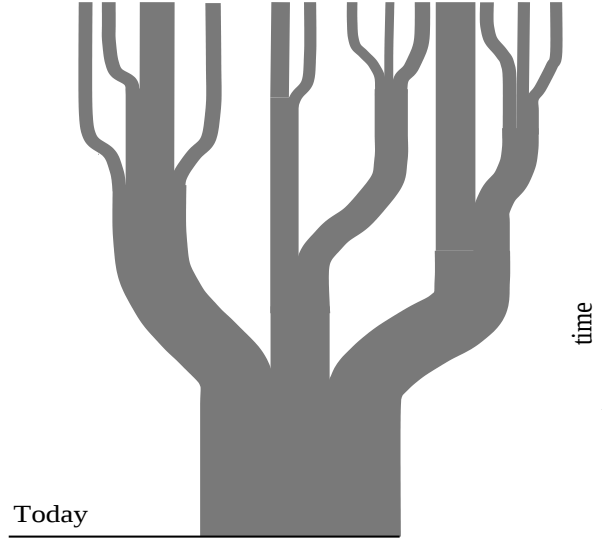


FIGURE 1.1— Schematic representation of a “merger tree” showing the growth of a halo in time. Time increases from top to bottom, and the widths of the branches are proportional to the masses of the progenitor halos (based on Lacey & Cole, 1993).

critical density) Ω_0 , the Hubble constant H_0 , the cosmological constant Λ , the baryon content of the Universe Ω_b , and the initial fluctuation power spectrum (shape and normalization), we may use the Press & Schechter theory (or some of its extensions) to make realizations of the merging histories of dark matter halos. Within each dark matter halo present at any time we may follow the evolution of the baryonic components: hot gas, cold gas and stars.

In the phenomenological approach mentioned previously (also known as semi-analytic modelling) one assumes that when a halo forms, its gas (a certain fraction of the total mass of the halo, which is mostly fixed by Ω_b) will relax to a distribution identical to that of the dark halo, at a temperature related to the circular velocity of the halo. Some fraction of this gas may be able to cool, and will thus collapse to form a rotationally supported disk at the centre of the halo.

In this rotationally supported disk star formation will take place, and as observations of nearby spirals seem to indicate (Kennicutt 1989), at a rate determined by the local surface density and the dynamical timescale of the disk. Some of the stars formed will later explode as supernova and blow material back into the hot halo. Winds from very massive stars may also be important in reheating cold gas and driving it back to the hot halo, or even beyond, especially in very active star forming regions. These mechanisms, which transform cold gas into hot gas, are known as feedback, and are currently some of the most poorly known processes affecting the formation and evolution of a galactic system.

Another process affecting the evolution of a galaxy is the chemical enrichment of its interstellar medium. As stars evolve they enrich their surrounding medium, and in a closed box model, all the metals remain in the galaxy halo. In the case of strong feedback one can imagine that some fraction of the processed material could be mixed with gas at very large radii, even outside the virial radius of the halo.

In semi-analytic models one also needs to specify what happens to the baryonic components once two halos merge. For example, in a newly formed halo one can imagine that the hot gas

present before the merger will remain in the hot phase, but with a virial temperature now given by the circular velocity of the new halo. If one of the halos is much smaller than the other, it may be considered a “satellite”. One can imagine that after the satellite has crossed the virial radius of the main galaxy halo it will cease accreting gas, and thus as soon as its cold gas supply present at the time of infall is exhausted, it will stop forming stars. This satellite will be accreted by the main galaxy on a certain timescale, which in these models is generally assumed to be a dynamical friction timescale (e.g. Navarro, Frenk & White 1995).

In the case of large mass ratios (“galaxy + satellite”) the morphology of the main galaxy will not change appreciably, and will thus maintain its disk-like configuration. If on the other hand, the masses of the halos are similar, then in these equal mass mergers the disks will be destroyed, and an elliptical system will be formed (e.g. Barnes 1999). Any cold gas present will be turned into stars on a very short timescale, so that new ellipticals formed through a merger with a large gas content will be starburst systems. And if the reservoir of gas in the hot halo is sufficiently large, a disk may grow later on, thus turning an elliptical galaxy into a spiral galaxy bulge.

Each of the processes mentioned here depends on a series of parameters whose values may be fixed, for example, from observations of galaxies in the local Universe. After this has been done, one can make realizations of the formation of a galaxy for different merger trees. For plausible choices of these free parameters, semi-analytic models can reproduce many of the observed characteristics of the galaxy population, such as luminosity functions, trends of galaxy colour with morphology and environment, the Tully–Fisher relation for spirals, etc. (van den Bosch 1998; Somerville & Primack 1999; Kauffmann et al. 1999). Examples of some early work trying to fit the luminosity function of the Milky Way and its satellites are shown in Figure 1.2.

1.2 The Milky Way galaxy

From the work described above we know, at least to some extent, how a typical galaxy forms and evolves in the hierarchical paradigm, starting almost from first principles. Naturally, we want to determine if the Milky Way was formed in this way. For our Galaxy, not only can we test this scenario by constructing the evolutionary history of a typical spiral and making a rough, global comparison between the predictions and the observations, but we can and should go farther. The very detailed knowledge of the kinematics, distances, ages and chemical compositions of individual stars should be useful to unravel how our Galaxy formed, and in this way to put constraints on the hierarchical structure formation scenario. The simplest example of this idea is the estimation of the age of the Universe from the ages of globular clusters (e.g. Sandage 1961; Carretta et al. 2000).

1.2.1 The Galactic components

The Milky Way is a large disk galaxy. Its main components are the thin disk, the thick disk, the bulge, the stellar halo and the dark matter halo. We describe here the first three components in some detail, and leave the Galactic halo (dark and stellar) for a separate section.

The thin disk

The thin disk contains most of the stars in the Solar neighbourhood. The stars, gas and dust in the thin disk follow an approximately exponential distribution, both in the radial and z -direction, with scale heights $h_R \sim 3.5$ kpc and $h_z \sim 0.3$ kpc respectively. The thin disk rotates rapidly with a circular velocity at the Solar radius of 215 ± 20 km s⁻¹ (where the quoted uncertainty is largely due to modelling uncertainties, e.g. Dehnen & Binney 1998). The thin disk contains stars of a wide range of ages, and there is also evidence for bursts in its star formation

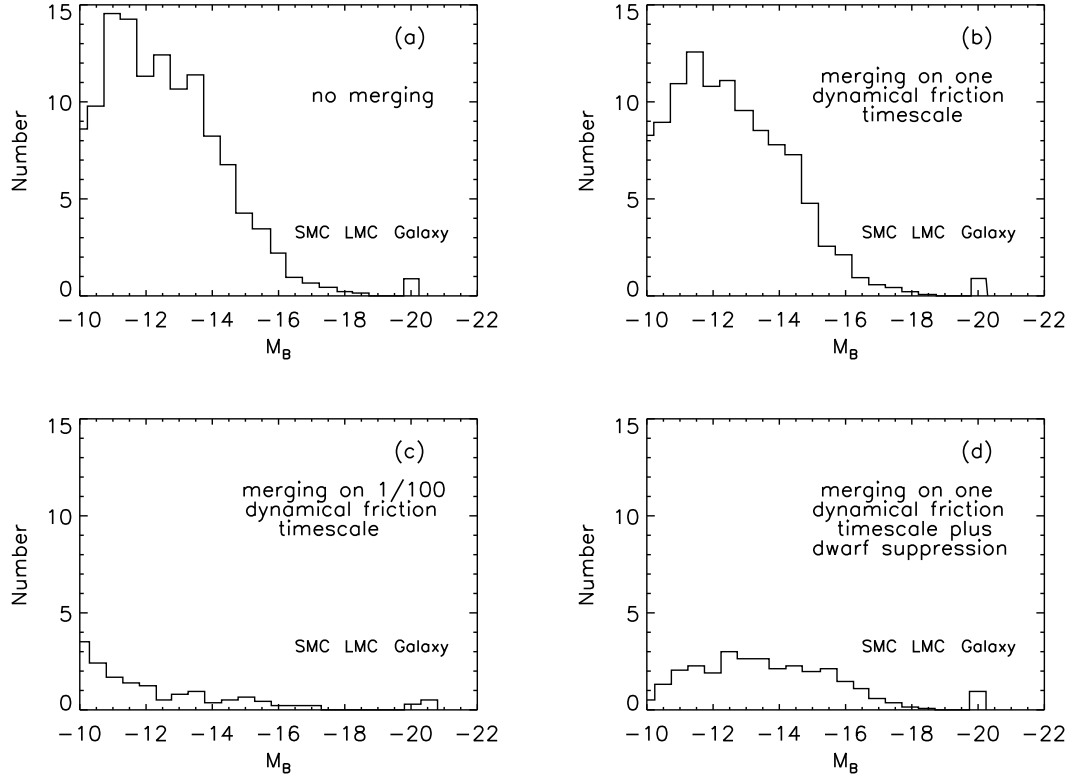


FIGURE 1.2— An early example of the Milky Way luminosity function in a SCDM cosmology with $\Omega_b = 0.1$ resulting from the semi-analytic modelling by Kauffmann et al. (1993). The different panels correspond to different values of one of the free parameters involved in the modelling, which in this case is the timescale on which the merging of two halos takes place. In (a) this timescale is infinite; for (b) it is the dynamical friction timescale as derived from the Chandrasekhar formula and for (c) it is 100 times shorter. In these first attempts it is already clear, as shown in panel (b), that there is an excess of faint galaxies in comparison to the observations (recall that the Milky Way has no more than about thirteen satellite galaxies). Even a reduction of a factor of 100 in the merging timescale is unsatisfactory, as it eliminates almost all satellites similar to the SMC and LMC. Inhibiting the formation of dwarf galaxies seems, however, to do much better. This can be achieved by invoking, for example, a strong UV background reionising field which would inhibit the cooling of gas in very small halos. The resulting luminosity function is shown in panel (d). In this way, the comparison to the observed luminosity function helps fixing the values of the free parameters for later modelling.

history (cf Rocha-Pinto et al. 2000; Hernández, Valls-Gabaud & Gilmore 2000). The oldest stars have ages of 8 ± 1.5 Gyr (e.g. Leggett, Ruiz & Bergeron 1998). Most of the stars in the Sun's vicinity have Solar metallicities, with $[\text{Fe}/\text{H}] \gtrsim -0.2$ dex (Edvardsson et al. 1993).

The thick disk

The thick disk has a larger vertical scale length, of $\sim 1 - 1.5$ kpc, a vertical velocity dispersion of 45 km s^{-1} , a typical stellar metal abundance of $1/4$ of the Solar metallicity, and a mean asymmetric drift (i.e. the lag with respect to the thin disk) of $30 - 50 \text{ km s}^{-1}$ (Gilmore, Wyse & Kuijken 1989). Being older (it is about 12 Gyr old) and more metal-poor than the thin disk, the thick disk must contain useful information on the early stages of the evolution of the Galaxy. Possible formation mechanisms for the thick disk include (cf Majewski 1993 for a good discussion):

- A slow, pressure-supported collapse phase after the formation of the stellar halo;

- The dynamical heating of the early thin disk by satellite accretion (e.g. Quinn, Hernquist & Fullagar 1993);
- A mixture of satellite debris and a heated early thin disk (Wyse 2000);
- Stochastic heating from a rapid increase in the star formation rate due to enhanced cooling (of halo gas) once the metallicity reached -1 dex (Wyse & Gilmore 1988).

The bulge

The bulge has about $1/3$ of the mass of the disk, and there is considerable evidence that it contains a bar, though its precise orientation and structure are still somewhat uncertain (Binney, Gerhard & Spergel 1997). The bulge is a component which remains a bit of a “mystery”, since much of what we know is based on the properties of stars in Baade’s window, a small area in the sky which is almost free of obscuring dust. Nevertheless, considerable progress on the determination of its structural properties has been achieved through observations in the infrared, in particular with IRAS, COBE and ISO observations. The bulge abundance distribution is rather broad, with a mean of $[\text{Fe}/\text{H}] \sim -0.25$, and a spread that goes from -1.25 to $+0.5$ dex (McWilliam & Rich 1994, but see also Minniti et al. 1995). There is no general agreement on the age and the spread in ages for bulge stars; and while most determinations agree that the bulge ought to be older than $5 - 8$ Gyr, it is yet unclear whether it might be as old as 15 Gyr (cf Gilmore, Wyse & Franx 1997). The bulge has a scale height of about 0.4 kpc, a velocity dispersion of about 100 km s^{-1} in the radial direction, with little indication that the bar plays a major dynamical role. The bulge may have been formed before the old thin disk in an inside-out fashion (Larson 1990), through mergers of disks (as proposed in the semi-analytic models described previously), simultaneously with the disk from the accretion of satellite galaxies, or after the formation of the disk, due to bar-like instabilities (Norman, Sellwood & Hasan 1995; see also Wyse & Gilmore 1992; and Bouwens, Cayón & Silk 1999 for interesting discussions of possible formation mechanisms).

1.2.2 The Galactic halo

The Galactic halo may be thought of as a two-component system. The dark halo, whose existence may be derived, for example, from the motions of the satellites of the Milky Way (e.g. Leo I, Zaritsky et al. 1989) or from timing arguments of the binary system M31 – Milky Way (Kahn & Woltjer 1959; Einasto & Lynden-Bell 1982; Peebles 1995), extends to over 200 kpc and has a total mass of the order of $10^{12} M_{\odot}$ (Zaritsky 1999; Wilkinson & Evans 1999). The stellar halo is much less important dynamically, but because it contains the most metal-poor stars in the Galaxy and possibly some of the oldest ones, it provides us with a picture of the Milky Way in its early stages of evolution. The very metal-poor halo stars are thus fossils, whose chemical abundance and motions contain information of their sites of origin. As Eggen, Lynden-Bell & Sandage (1962) nicely put it: “... a study of these subsystems allows us partially to reconstruct the Galactic past because the time required for stars in the Galactic system to exchange their energies and momenta is very long compared with the age of the Galaxy. Hence knowledge of the present energy and momenta of individual objects tells us something of the initial dynamic conditions under which they were formed”.

Some observational characteristics of halo field stars

The stellar halo has a total luminosity in the range $8 \times 10^8 - 10^9 L_{\odot}$ (e.g. Freeman 1995). Its luminosity profile may be fit with an $r^{-3.5}$ law (e.g. Kinman, Suntzeff & Kraft 1994), thus being very centrally concentrated, considerably more than the disk. Its half-light radius lies probably

around 3 kpc from the Galactic centre. It is difficult to determine whether this steep profile continues all the way to the Galactic centre, or whether it has a break. It is also unclear whether the stellar halo is the extension of the bulge (Morrison & Harding 1993; Minniti 1996). It should be kept in mind, however, that their chemical properties are rather different: the bulge is metal-rich whereas, as we shall see, the stellar halo is metal-poor.

The stellar halo shows no systemic rotation, with $v_\phi \sim 3 \pm 21 \text{ km s}^{-1}$, $v_R \sim 16 \pm 18 \text{ km s}^{-1}$ and $v_z \sim -10 \pm 12 \text{ km s}^{-1}$ (Chiba & Yoshii 1998), but there is some evidence that it may have a net retrograde motion in the outer parts (e.g. Majewski 1992; Carney et al. 1996). The velocity ellipsoid in the Solar neighbourhood is radially elongated with $v_\phi = 115 \pm 7 \text{ km s}^{-1}$, $v_r = 161 \pm 18 \text{ km s}^{-1}$, and $v_z = 108 \pm 7 \text{ km s}^{-1}$ (see Figure 1.3). There are some indications that this velocity ellipsoid becomes tangentially elongated farther away from the Galactic centre. This duality (radial vs tangential anisotropy) has been taken as evidence that the inner halo formed during a dissipative collapse, whereas the outer halo formed through the accretion and merging of satellites (e.g. Sommer-Larsen et al. 1997). Over the years, several groups have detected moving groups in the stellar halo, further indication that accretion and merging have played an important role in shaping the Galaxy (e.g. Majewski, Munn & Hawley 1994; see also references in Chapter 2). Some of these detections have not survived further tests nor have they been confirmed by later work. In general such claims have not been convincingly associated with accreted satellites, in part due to the lack of detailed theoretical modelling of the characteristics of satellite debris, which has only become available recently (e.g. Johnston, Hernquist & Bolte 1996; Helmi & White 1999).

The stellar halo contains stars as metal-poor as $[\text{Fe}/\text{H}] \sim -3.5$ dex (Ryan, Norris & Beers 1996), and extends to $[\text{Fe}/\text{H}] \sim -1$ dex in the metal-rich end (Nissen & Schuster 1997). Its characteristic metallicity range is $[\text{Fe}/\text{H}] \lesssim -1.6$ dex. Halo stars typically show α -element enhancement $[\alpha/\text{Fe}] > +0.3$ dex when compared to the Sun or other thin disk stars. This enhancement is indicative of metal enrichment by Type II supernovae only, which are the result of the evolution of very massive stars that explode on rather short timescales, only about 0.5 Gyr after their formation. On timescales longer than 1 Gyr, supernova Type I explosions will contribute with iron-peak elements, thereby decreasing $[\alpha/\text{Fe}]$ towards Solar values. The chemical properties of the stellar halo thus imply that star formation occurred on timescales shorter than 1 Gyr, after which it stopped (Gilmore & Wyse 1998). However, there is also evidence that some metal-rich halo stars with $[\text{Fe}/\text{H}] \sim -1$ dex have low α -values, probably indicating enrichment by Type Ia SN. These stars have been interpreted as having been accreted from dwarf galaxies (for an alternative explanation see Gilmore & Wyse, 1998). Another important property of the halo is the absence of a metallicity gradient in the halo field population.

Halo stars are amongst the oldest stars in the Galaxy (VandenBerg 1999). It is also interesting to note here that there have been detections of a few metal-rich A stars in the halo (Rodgers, Harding & Sadler 1981), and some metal-poor A and F main sequence stars with halo kinematics (Preston, Beers & Schectman 1994). These stars, being of intermediate age, may have originated from a satellite galaxy that was accreted in the past 10 Gyr.

Globular clusters

The globular clusters of the Milky Way are amongst the oldest objects in the Universe. They have metallicities in the range $-2.4 \leq [\text{Fe}/\text{H}] \leq 0.2$, absolute magnitudes $-10 \leq M_V \leq -3$, and Galactocentric distances that range from 1 kpc up to 100 kpc. Their radial density distribution follows an $r^{-3.5}$ profile, as do halo field stars, all the way from 3 kpc to 100 kpc (Zinn 1985).

The metallicity distribution appears to be bimodal, with peaks located around $[\text{Fe}/\text{H}] \sim -1.6$ and $[\text{Fe}/\text{H}] \sim -0.6$ dex (Zinn & West 1984; see however Sarajedini 1999). The metal-

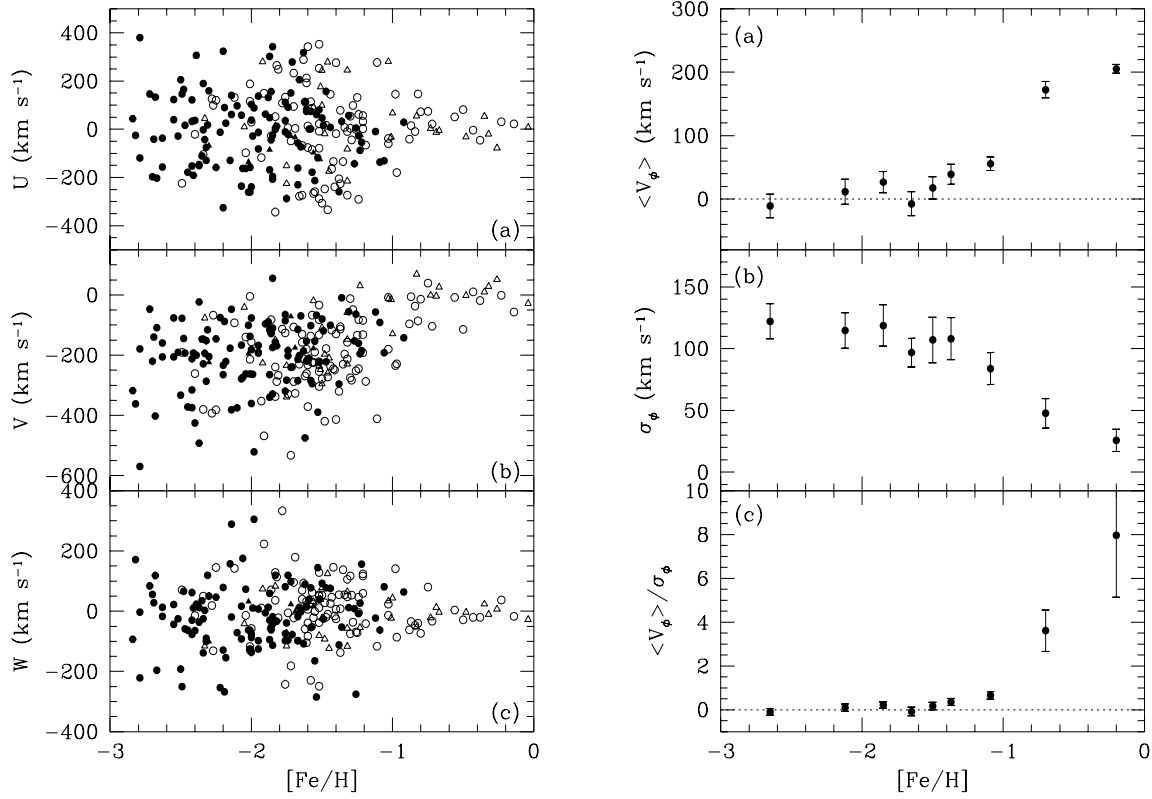


FIGURE 1.3— The left panel shows the kinematics of a sample of metal-poor red giants (filled symbols) and RR Lyrae (open symbols) in the Solar neighbourhood, with proper motions measured by the HIPPARCOS satellite. Circles denote stars with small relative errors in the proper motion, while triangles correspond to stars with large errors. In the right panel the rotational properties vs $[Fe/H]$ for the same sample are shown. The transition from the halo (low metallicities, zero mean rotational velocity) to the disk is very evident in the two top panels. Noticeable as well in the bottom panel is the transition from pressure supported to rotationally supported kinematics which takes place at a metallicity $[Fe/H] \sim -1$ dex. (From Chiba & Yoshii 1998.)

poor population is distributed in a roughly spherical halo around the Galactic centre with high velocity dispersion and no rotation. The metal-rich population appears to be in a flattened disk-like distribution with significant rotation, and may be associated with the thick disk or the bulge, instead of the stellar halo (Zinn 1985).

Whereas there is a general consensus that the globular clusters are very old objects, an interesting issue is whether there is an age range among them, or equivalently, what was their formation timescale. Other relevant questions are whether there is a correlation between age and metallicity, and what is the relation between age and Galactocentric distance (cf Stetson, Vandenberg & Bolte 1996; Sarajedini, Chaboyer & Demarque 1997). There are some indications that there is indeed an age range of about 2 Gyr among all globular clusters (Sarajedini 1999). The range is practically zero for the metal-poor population, and seems to increase as the metal abundance increases (Vandenberg, Bolte & Stetson 1990; Salaris & Weiss 1997; Saviane, Rosenberg & Piotto 1999). No correlations of age and Galactocentric distance have yet been found.

Key observables and formation scenarios

The difficulty in deriving a formation scenario for the Milky Way does not lie in the lack of observational data, as is the case for the majority of the galactic systems in the Universe, but on

the contrary, in the large amounts of detailed data that are available. Trying to put it all together may seem an impossible task if we attempt to use everything we know about our Galaxy! It is thus useful to define a list of key observables to start sketching possible formation scenarios for our Galaxy. Following Majewski (1993) we define the following set:

1. The Galactic halo and thick disk appear to be coeval and old.
2. The globular clusters formed over a few Gyr.
3. The old thin disk is discrete in age from the thick disk and halo, and several Gyr younger.
4. The halo lacks a metallicity gradient.
5. The thick disk may have a mild metallicity gradient (0 to -0.1 dex kpc $^{-1}$), and a large metallicity spread, reaching $[\text{Fe}/\text{H}] \leq -1.6$.
6. There is circumstantial evidence suggesting that at least part of the halo has been accreted.

In the early sixties, the first model for the formation of our Galaxy was put together by Eggen, Lynden-Bell & Sandage (1962, hereafter ELS), who analysed the kinematics and abundances of dwarf stars in the Solar neighbourhood. They found that *“correlations exist between the chemical composition of the individual stars, the eccentricity of their orbits, their angular momenta, and the height they reach above the Galactic plane. These correlations are interpreted by a sequence of events starting with the protogalaxy condensing out of a pregalactic medium, collapsing toward the fundamental plane, and shrinking in diameter until the present equilibrium state is reached where the gravitational attraction is just balanced by the centrifugal acceleration”*. The timescale they proposed was a few times 10^8 years, and they also concluded that *“the Galaxy had not settled down ... at the time of its first star formation but was in its initial gravitational contraction from a larger protogalaxy”*.

It is unclear whether the ELS original picture implied a metallicity gradient in the stellar halo, because the proposed collapse timescale would be too short for this to happen. The *a posteriori* discovered age spread in the globular cluster population led to the idea that either (i) the ELS collapse lasted about 1 Gyr (Sandage & Fouts 1987); or (ii) the ELS scenario was not correct. But if (i) were true, then one would expect a metallicity gradient, because this timescale is longer than that required for SN Type II enrichment, and such a metallicity gradient has not been observed as already mentioned. Moreover, even though ELS do not discuss whether the collapse was homogeneous, the *“protogalaxy condensing out of a pregalactic medium”* seems to imply that any type of substructure (if not erased during the very short timescale of the collapse), would have been on the scales of molecular clouds or star forming regions, and thus inconsistent with the observational evidence of moving groups in the outer halo.

The age spread amongst globular clusters and the absence of a metallicity gradient in their distribution, led Searle & Zinn (1978, hereafter SZ) to *“think of the process of halo formation as the merging of a number of distinct subsystems. Such subsystems, which we shall call fragments, may perhaps be thought of as having the character of small gas-rich irregular galaxies”*. Support for this picture comes also from the outer halo moving groups, although this may only be circumstantial, as it is difficult to link such groups to disrupted “fragments”. The smooth kinematic distribution of halo stars in the *inner* Galaxy was taken as evidence that the SZ picture was, at most, correct in the *outer* Milky Way, and many would agree that the Galaxy is the result of an ELS collapse + later SZ accretion (e.g. Norris 1994; but see new evidence of moving groups also in the Solar neighbourhood by Helmi et al. 1999).

In parallel to the development of models for the formation of the Milky Way in the late seventies and early eighties, cosmologists were starting to put together the hierarchical paradigm described in the beginning of this chapter. In the past few years both ends of the galaxy formation research community have come together, and we may now be able to envisage a scenario of the formation of our Galaxy in its proper cosmological framework.

We will start our journey of unraveling the formation history of the Milky Way in this cosmological context by focusing on its stellar halo, since, as we already stated, this contains most metal-poor and possibly oldest stars in the Galaxy. In the context of hierarchical structure formation, the stars of the structures that assembled to form the Milky Way should now make up much or all of the Galactic bulge and halo. Galaxies that merged to form our Milky Way would have left trails of stars along their orbits, so that an ensemble of such would give rise to the spheroidal components we observe today.

1.3 Outline of this thesis

The goal of the work presented in Chapter 2 is to assess whether a merger origin for the stellar halo would leave observable fossil structure in the phase-space distribution of nearby stars. With this in mind, we perform numerical simulations of satellite galaxy disruption in a potential resembling that of the Milky Way. We find that after 10 Gyr of evolution few obvious asymmetries remain in the distribution of particles in configuration space, but that strong correlations are still present in velocity space. Based on a linearised treatment in action-angle variables, we develop a simple analytic description of these effects which shows how the kinematic and density structure of debris streams changes with time. By applying this description we find that a single dwarf elliptical-like satellite of current luminosity $10^8 L_{\odot}$ disrupted 10 Gyr ago from an orbit circulating in the inner halo, would contribute about ~ 30 streams with internal velocity dispersions below 5 km s^{-1} to the local stellar halo. If the whole stellar halo were built by such disrupted satellites, it should consist locally of 300 – 500 such streams. Clear detection of all these structures would require a sample of a few thousand stars with 3-D velocities accurate to better than 5 km s^{-1} . Even with velocity errors several times worse than this, the expected clumpiness should be quite evident.

Thus, if hierarchical models are correct, we may infer that the stars of the structures that assembled to form the Milky Way, should be visible as “fossil” streams. However *direct* evidence that the bulk of the Milky Way’s population of old stars was built up from mergers has so far been lacking. In Chapter 3 we use kinematic data from the HIPPARCOS satellite to demonstrate that about ten per cent of the metal-poor stellar halo outside the Solar radius comes from a single coherent object which was disrupted during or soon after the Milky Way’s formation. This object had a highly inclined orbit with an apogalacticon of about 16 kpc, and probably resembled the Fornax and Sagittarius dwarf spheroidal galaxies.

In Chapter 4 we investigate what the next generation of astrometric satellites will reveal by observing the halo of the Milky Way, under the assumption that this was built exclusively from disrupted galaxies. Starting from numerical simulations of the infall of satellite galaxies onto a fixed potential, we generate artificial DIVA, FAME and GAIA halo catalogues, in which we look for the signatures left by the accreted satellites. We find that a method based on the standard Friends-of-Friends algorithm applied to the space of integrals of motion can recover about 50% of the different accretion events, when the observational uncertainties expected for GAIA are taken into account, and even when the exact form of the Galactic potential is unknown. The recovery rate for DIVA and FAME is much smaller, but these missions, like GAIA, should be able to test the hierarchical formation paradigm on our Galaxy by measuring the amount of

halo substructure in the form of nearby kinematically cold streams with for example, a two-point correlation function in velocity space.

Accretion and merging are still continuing in our Galaxy, as evidenced by the Sagittarius dwarf galaxy (Ibata, Gilmore & Irwin 1994), which is currently being torn apart by the Milky Way. With the aim of understanding the survival and characteristics of Sagittarius, we performed N-body simulations of the progressive disruption of a satellite galaxy orbiting for 12.5 Gyr within a realistic Galactic potential. In Chapter 5 we present two simple dynamical models for Sagittarius, both having observable properties similar to those of current outlying dwarfs; in one case it is purely stellar while in the other it is embedded in an extended massive halo. We find that a possible progenitor of Sagittarius could be a satellite with a core radius of 0.56 kpc and a total stellar velocity dispersion of about 18 km s^{-1} . Both these models are consistent with all published data on the current Sagittarius system: they match not only the observed properties of the main body of Sagittarius, but also those reported for unbound debris at larger distances. At present, the data available can only weakly constrain the total initial extent either of the light or of the mass. Further work on the debris streams of Sagittarius is needed to constrain better its initial total luminosity, and to distinguish between purely stellar or dark matter dominated progenitors. Sagittarius will most likely have disappeared in the next pericentric passage: Its stars will have spread out in the Galaxy and become part of the Galactic halo, much like, as the accumulated evidence seems to indicate, has happened to many other systems in the past.

Our first attempt to determine whether the merging history of the Milky Way may be imprinted in the phase-space structure of its stellar halo (Chapter 2) was worked out under the assumption of a fixed Galactic potential, onto which small galaxies fall in. Clearly, this assumption is not correct for the hierarchical build up of structure, which is ultimately the scenario that we want to test here. In the fully hierarchical regime, the concept of a background galaxy is meaningless, in the case of a spiral galaxy probably until $z \sim 2 - 3$. Moreover, the potential may be varying very violently, with large numbers of clumps orbiting the centre of what will become a virialised halo. Thus one of the goals of Chapter 6 is to determine to what extent our previous analysis may have been correct (or not), and also understand why this may be the case. We tackle these issues by analysing a high-resolution simulation of the formation of a cluster in a Λ CDM cosmology (Springel et al. 2000). As in Chapter 2, we measure how much phase-space substructure is left over from the halos that merged to build up the cluster we observe at the present day. We study debris streams originating in such halos and find that their evolution can be explained simply in terms of the conservation of phase-space density. We also find that the properties of these streams are consistent with having a phase-mixing origin. This result suggests that the merging process is not fully chaotic, and even though the coarse-grained distribution function may be the same for all halos, as the universality of the Navarro, Frenk & White (1996) profile seems to imply, the memory of their history is still preserved in the fine-grained distribution function, even after a Hubble time. By scaling the cluster to a galactic size halo, we find that the *dark matter* velocity ellipsoid in the equivalent of the “Solar vicinity” is formed by roughly a thousand streams, with internal velocity dispersions less than a few tens of km s^{-1} . These results indicate that current and future experiments designed to determine the properties of dark matter particles in our immediate neighbourhood need not take into account the “graininess” of the velocity distribution. It is unlikely that the statistics would be dominated by the presence of just one stream of particles. By appropriately scaling the number of *dark* streams to derive the number of *star* streams, we find that the stellar velocity ellipsoid in the Solar neighbourhood should consist of roughly five hundred streams, thus confirming our earlier analytic estimates.

1.4 A possible scenario for the formation of our Galaxy

The hierarchical paradigm for the formation of structure in the Universe is quite different conceptually from either the ELS or the SZ models, even though all of these aim at reproducing, for example, the observational characteristics of the Milky Way, and they do so, at least to some extent, quite successfully. It is remarkable that the pioneering simple model set forward by Eggen, Lynden-Bell & Sandage has survived almost forty years, without being completely ruled out. The main difference between the ELS and the hierarchical models lies in the monolithic formation against a gradual assembly from smaller systems. The hierarchical scenario is also quite different from the proposed SZ model. A large galaxy like the Milky Way does not come about from the merging of a thousand or more small fragments of $10^8 M_\odot$ as in SZ case, but more likely from no more than ten to twenty quite large systems each having their own identity.

The combination of semi-analytic modelling of galaxy assembly with high-resolution simulations of the formation of a cluster scaled down to a galactic size halo (the same as that used in Chapter 6), shows that the oldest stars in the Milky Way should lie in the central bulge, which is expected to be metal-rich (Springel et al. 2000), as illustrated in the bottom left panel in Figure 1.4. This result is closely linked to the build up of structure in hierarchical models.

The earliest stars formed in the first collapsed halos at $z \sim 20$ should clearly be metal-poor. The gas in these small halos should have been easily expelled in winds – because of the shallow potential wells – and so star formation halted. Thus, on average stars formed more quickly in the more massive halos that collapsed slightly later than in the small ones (Miralda-Escudé 2000). The interstellar medium of these comparatively large halos should then have been rapidly enriched (Cen & Ostriker 1999). In most small halos, in turn, the gas may have not been able to cool until fairly low redshift because of the UV reionising background produced by quasars. Thus small systems may have kept their original (almost) metal-free gas and may form stars of extremely low metallicity even at the present day (an example of which is the Milky Way satellite galaxy Leo I, e.g. Gallart et al. 1999).

In a merger large systems decay by dynamical friction against their mutual halos. The stars already present thus end up in the very centre of the newly formed galaxy, which therefore will generally be metal-rich. While such galaxies are still falling against each other, some of their stars are released. These stars should have less tight orbits, and thus may contribute to a stellar halo. On the other hand, some fraction of halo stars will come from smaller objects which did not directly sink towards the centre, and were in general more metal-poor. A metallicity gradient in the ELS sense would not be expected in the hierarchical formation of a galaxy like the Milky Way. Age and metallicity are not necessarily correlated for the bulk of the stellar population.

During these mergers a globular cluster population may be formed. Because of the fairly high redshift in which these mergers take place – somewhere between $z = 5$ and $z = 3$ for a spiral galaxy like our own, i.e. 12.6 and 11.6 Gyr ago, respectively, in a Λ CDM cosmology – each system will have a large (cold and hot) gas content, but which is already polluted by the first population of stars. The more metal-rich and slightly younger component of the globular cluster system of the Milky Way could well have formed in such mergers (e.g. Ashman & Zepf 1992). This system should be rapidly rotating, reflecting the total orbital angular momentum of the merged galaxies, a “prediction” consistent with the observational data. The rest of the globular clusters of the Milky Way may have come with smaller galaxies that merged at different times (Côté et al. 2000), and so be associated with the stellar halo rather than with the bulge. Their distribution would then reflect the orbits of their parent galaxies, and thus naturally be more extended. No signature of a net rotation would be expected and the velocity dispersion of the

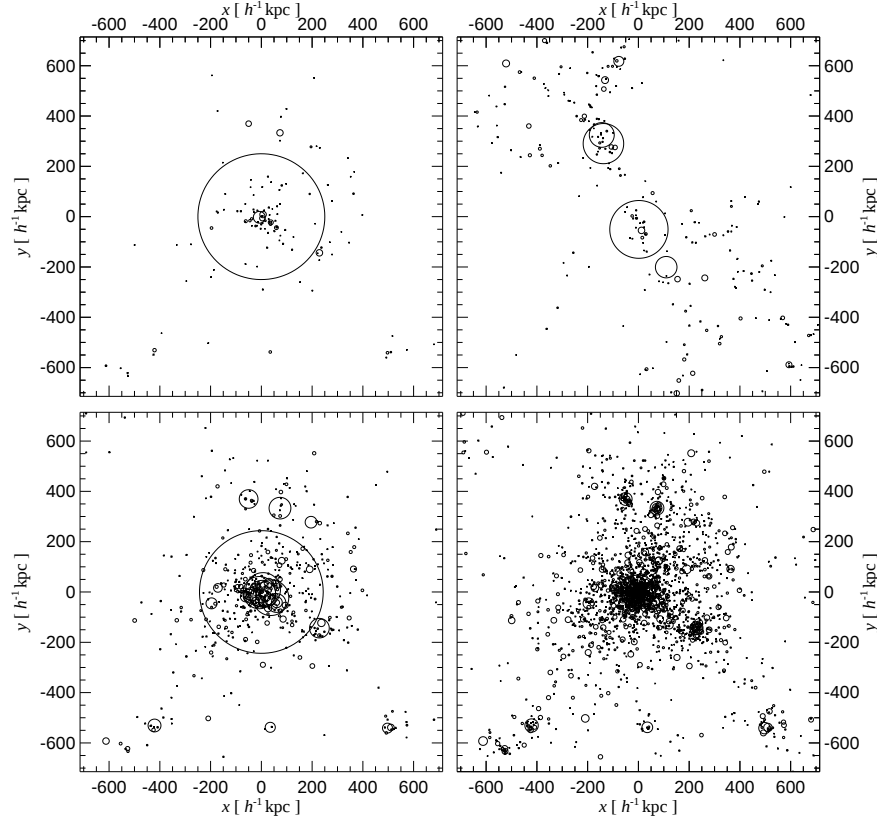


FIGURE 1.4— Projections of the galaxy distribution in a cube of comoving side $1.4h^{-1}$ Mpc around a “Milky Way” halo from a high-resolution simulation of a cluster combined with semi-analytic modelling of galaxy formation. The upper right image is at $z = 6.9$ (which should be shifted to $z \sim 10$ because of the later assembly of a cluster halo). At this redshift 1% of the final stars have formed. Note that the initial distribution of these “first” stars is very inhomogeneous, in a handful of progenitor galaxies (here the symbol area is proportional to stellar mass). The other three panels correspond to $z = 0$. The top left panel shows the present-day stellar distribution, which is dominated by the disk where 80% of all stars in the “Milky Way” reside. The 1% “first” stars are even more centrally concentrated as shown in the lower left panel. Finally, the lower right panel corresponds to the 1% “lowest metallicity” stars. In these bottom panels symbol size is proportional to the mass of the relevant stellar population. (From White & Springel, 2000.)

whole system should be quite large, as indeed seems to be the case (e.g. Dinescu, Girard & van Altena 1999).

The 1-Gyr timescale quoted for the formation of the whole Galactic stellar halo based on the chemical enrichment by Type I supernovae, can also be explained in this context. The first galaxies from which the Milky Way was assembled did only have a small amount of time to retain their identity (very likely no more than 1 Gyr, i.e. until $z \sim 7$), after which they were strongly affected by the tidal forces of the rest of the clumps orbiting what would become a virialised halo. Thus star formation in such systems probably stopped quite quickly, and almost immediately after these systems coalesced and contributed to the stellar population in the halo with old stars. These stars would then be both metal-poor and show an α -element enhancement as observed. Note for example, that by $z \sim 10$, White & Springel (2000) find that their simulated Milky Way had formed 1% of its stars, i.e., of the order of 60% of the stellar halo. Therefore the kind of argumentation often used against the hierarchical formation of the Milky Way by comparing present-day satellite galaxies with the inner Galactic stellar halo is not valid (Unavane, Wyse & Gilmore 1996). Stars in systems like the Carina dwarf spheroidal were probably formed

about the same time or may be even much later than the stars that were already in place in the Galaxy 10 Gyr ago. Clearly, if Carina had fallen in roughly 10 Gyr ago it would have looked quite different from how it does today!

The stars already present in the systems that merged to build up the Milky Way contributed both to the bulge and stellar halo, and we should expect to observe these stars to be distributed in kinematically cold streams, as shown in Chapter 2 of this thesis. The smooth distribution generally claimed to support the ELS scenario is also expected in hierarchical models, if the samples are small and the velocity uncertainties are very large, as usually has been the case thus far. Therefore the properties of the bulge and stellar halo are consistent with the hierarchical build-up of the Milky Way. Some fraction (or even most) of the Milky Way's bulge may have formed through instabilities of the central disk (Norman, Sellwood & Hasan 1996). Such a mechanism would also enhance the metal-rich character of the bulge against that of the metal-poor halo.

What about the disk components of the Milky Way? The thick disk appears to be only 1 Gyr younger than the stellar halo, meaning that its stars must have formed only slightly later, though not necessarily in the same place. A possible scenario for the formation of the disk components would imagine that, after the formation of a good fraction of the bulge, the (thin disk) precursor of the thick disk could have been formed by the cooling of some of the hot gas in the halo (around 12 Gyr ago, i.e. $z = 3.5$). In the meantime some fraction of the stellar halo was being accreted through smaller mergers, objects that may have fallen in earlier but, due to their smaller mass did not sink to the inner regions very quickly. In one of these minor mergers, an object probably having about 20% of the mass of the (proto)Milky Way, heated up precursor disk material and gave rise to the thick disk as we know it today. The thick disk, therefore, should also contain debris from the disrupted satellite. There is some observational evidence (based on data obtained using the 2-degree-field spectrograph) that this may be the case, but at this stage this evidence is still weak (Wyse 2000), and detailed modelling is required to determine the clear signatures, for example in the kinematics, of such a population. The thin disk formed later, from more gas cooling from the halo, and from material that at this stage had already been completely polluted by supernovae Type I, which thus explains the dichotomy between the halo and the disk populations.

1.5 Next steps

It seems from the previous discussion that the best way to test the hierarchical formation our Galaxy would be by obtaining large samples of halo and bulge stars with good kinematics. This will be possible in the next years, both with astrometric satellite missions such as FAME and GAIA (see Chapter 4 for references), and also with large, complete, ground-based samples of halo stars with good proper motions and radial velocities, some of which are being currently constructed (e.g. Beers et al. 2000). Detailed estimates of the ages of individual stars may also be useful, although more difficult to obtain reliably. Element abundances for bulge stars are required to understand how and on what timescale the central component of the Milky Way formed. Also important will be further theoretical modelling on subgalactic scales. In particular a good understanding of star formation and feedback is required to constrain better any formation scenario for the build up of galactic systems. A more robust phenomenological description of the behaviour of the baryonic components of small galaxies combined with an analytic modelling of the phase-space evolution of satellite debris (such as that developed in Chapter 2) may be useful for understanding the outcome of large surveys like 2dF and the Sloan Digital Sky Survey. For example, the Sloan commissioning data has shown that a large number

of the RR Lyrae stars in the outer parts of the halo are distributed in substructures (Ivezic et al. 2000), which can be linked almost entirely to the streams of material stripped off the Sagittarius dwarf (Ibata et al. 2000). In fact, their positions in the sky and their magnitudes are in fairly good agreement with the model shown in Chapter 5 of this thesis. Another interesting constraint that may be derived from these data sets is the number and properties of the satellites that have merged with the Milky Way to form the outer stellar halo. Their debris should be visible as coherent structures in the sky, although may not be as obvious as the Sagittarius streams. These structures would be useful, for example, to determine more accurately the potential of the Milky Way, and its total extent (e.g. Johnston et al. 1999).

In 1962, Oort wrote that “...the subject of structure and evolution of galaxies is in the unsatisfactory stage where we have observed many intriguing phenomena, but have no good theories to explain them...”. Forty years later, we may confidently say that we have made the first steps towards such an understanding. And the wealth of data that will become available in the coming ten to fifteen years, and theoretical progress that will surely be made in modelling the formation of galactic systems, promise a very exciting next decade for cosmology, galaxy evolution and, in particular, for unraveling the formation history of our Galaxy.

References

- Ashman K.M., Zepf S.E., 1992, *ApJ*, 384, 50
 Barnes J., 1999, in “Galaxy Dynamics”, ed. Merritt D.R. et al., ASP Conf. Series 182, p. 463
 Baugh C., Cole S., Frenk C.S., 1996, *MNRAS*, 283, 1361
 Beers T.C., Chiba M., Yoshii Y., Platais I., Hanson R.B., Fuchs B., Rossi S., 2000, *AJ*, in press (astro-ph/0003103)
 Binney J., Gerhard O., Spergel D., 1997, *MNRAS*, 288, 365
 Bond J.R., Cole S., Efsthathiou G., Kaiser N., 1991, *ApJ*, 379, 440
 Bouwens R., Cayón L., Silk J., 1999, *ApJ*, 516, 77
 Bower R.J., 1991, *MNRAS*, 248, 332
 Carney B.W., Laird J.B., Latham D.W., Aguilar L.A., 1996, *AJ*, 112, 668
 Carretta E., Gratton R.G., Clementini G., Fusi Pecci F., 2000, *ApJ*, 533, 215
 Cen R., Ostriker J.P., 1999, *ApJ*, 519, L109
 Chiba M., Yoshii Y., 1998, *AJ*, 115, 168
 Côté P., Marzke R.O., West M.J., Minniti D., 2000, *ApJ*, 533, 869
 Dehnen W., Binney J., 1998, *MNRAS*, 294, 429
 Dicke R.H., Peebles P.J.E., Roll P.G., Wilkinson D.T., 1965, *ApJ*, 142, L415
 Dinescu D. I., Girard T.M., van Altena W.F., 1999, *AJ*, 117, 1972
 Edvardsson B., Andersen J., Gustafsson B., Lambert D.L., Nissen P.E., Tomkin J., 1993, *A&A*, 275, 101
 Eggen O.J., Lynden-Bell D., Sandage A.R., 1962, *ApJ*, 136, 748 (ELS)
 Einasto J., Lynden-Bell D., 1982, *MNRAS*, 199, 67
 Freeman K.C., 1995, in “New light on galaxy evolution”, Proc. IAU Symp. 171, eds. Bender R. & Davies R.L., Kluwer (The Netherlands), p. 3
 Gallart C., Freedman W.L., Aparicio A., Bertelli G., Chiosi C., 1999, *AJ*, 118, 2245
 Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 1998, *MNRAS*, 300, 146
 Gilmore G., Wyse R.F.G., Kuijken K., 1989, *ARA&A*, 27, 555
 Gilmore G., 1993, in “The Minnesota lectures on the structure and dynamics of the Milky Way”, ed. Humphreys R., ASP Conf. Series 39, p. 66
 Gilmore G., Wyse R.F.G., Franx M., 1997, *ARA&A*, 35, 637
 Gilmore G., Wyse R.F.G., 1998, *AJ*, 116, 748

- Guth A.H., 1981, *Phys. Rev. D*, 23, 347
- Helmi A., White S.D.M, 1999, *MNRAS*, 307, 495 (Chapter 2)
- Helmi A., White S.D.M, de Zeeuw P.T., Zhao H.S., 1999, *Nature*, 402, 53 (Chapter 3)
- Hernández X., Valls-Gabaud D., Gilmore G., 2000, *MNRAS*, in press (astro-ph/0003113)
- Ibata R., Gilmore G., Irwin M., 1994, *Nature*, 370, 194
- Ibata R., Irwin M., Lewis G., Stolte A., 2000, submitted to *ApJL* (astro-ph/0004255)
- Ivezic Z., Goldston J., Finlator K., et al. (for the SDSS), 2000, submitted to *AJ* (astro-ph/0004130)
- Johnston K.V., Hernquist L., Bolte M., 1996, *ApJ*, 465, 278
- Johnston K.V., Zhao H.S., Spergel D.N., Hernquist L., 1999, *ApJ*, 512, L109
- Kahn F.D., Woltjer L., 1959, *ApJ*, 130, 705
- Kauffmann G., White S.D.M, Guiderdoni B., 1993, *MNRAS*, 264, 201
- Kauffmann G., Colberg J.M., Diaferio A., White S.D.M., 1999, *MNRAS*, 303, 188
- Kennicutt R.C., 1989, *ApJ*, 344, 685
- Kinman T.D., Suntzeff N.B., Kraft R.P., 1994, *AJ*, 108, 1722
- Klypin A., Kravtsov A., Valenzuela O., Prada F., 1999, *ApJ*, 522, 82
- Lacey C., Cole S., 1993, *MNRAS*, 262, 627
- Larson R.B., 1990, *PASP*, 102, 790
- Leggett S.K., Ruiz M.T., Bergeron P., 1998, *ApJ*, 497, 294
- Linde A., 1982, *Phys. Lett. B*, 108, 389
- Majewski S.R., 1992, *ApJS*, 78, 87
- Majewski S.R., 1993, *ARA&A*, 31, 575
- Majewski S.R., Munn J.A., Hawley S.L., 1994, *ApJ*, 427, L37
- McWilliam A., Rich R.M., 1994, *ApJS*, 91, 749
- Minniti D., Olszewski E., Liebert J., White S.D.M., Irwin M., 1995, *MNRAS*, 277, 1293
- Minniti D., 1996, in "Formation of the Galactic Halo...Inside and Out", eds. Morrison H. & Sarajedini A., *ASP Conf. Series* 92, p. 92
- Miralda-Escudé J., 2000, in *Proc. of the MPA/ESO workshop "The First Stars"*, ed. Weiss A. et al., in press (astro-ph/9911214)
- Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, *ApJ*, 524, L19
- Morrison H.L., Harding P., 1993, *PASP*, 105, 977
- Navarro J.F., Frenk C.S., White S.D.M., 1995, *MNRAS*, 275, 56
- Navarro J.F., Frenk C.S., White S.D.M., 1996, *ApJ*, 462, 563
- Nissen P.E., Schuster W.J., 1997, *A&A*, 326, 751
- Norman C., Sellwood J.A., Hasan H., 1996, *ApJ*, 462, 114
- Norris J.E., 1994, *ApJ*, 431, 645
- Oort J.H., 1962, in "Problems of extra-galactic research", *IAU Symp.* 15, ed. McVittie G.C., p. 179
- Pearce F.R., et al. (The Virgo Consortium), 1999, *ApJ*, 521, L99
- Peebles P.J.E., 1974, *ApJ*, 189, L51
- Peebles P.J.E., 1982, *ApJ*, 263, L1
- Peebles P.J.E., 1995, *ApJ*, 449, 52
- Penzias A.A., Wilson R.W., 1965, *ApJ*, 142, L419
- Press W.H., Schechter P., 1974, *ApJ*, 187, 425
- Preston G., Beers T.C., Sheckman S., 1994, *AJ*, 108, 538
- Quinn P.J., Hernquist L., Fullagar D., 1993, *ApJ*, 403, 74
- Rocha-Pinto H.J., Scalo J., Maciel W.J., Flynn C., 2000, submitted to *A&A* (astro-ph/0001383)
- Rodgers A.W., Harding P., Sadler E., 1981, *ApJ*, 244, 912
- Ryan, S.G., Norris J.E., Beers T.C., 1996, *ApJ*, 471, 254
- Salaris M., Weiss A., 1997, *A&A*, 327, 107

- Sandage A.R., 1961, *AJ*, 66, 54
- Sandage A.R., Fouts G., 1987, *AJ*, 93, 74
- Sarajedini A., 1999, in "The Galactic Halo", *Proc. of the 3rd. Stromlo Symposium*, ed. Gibson B. et al., *ASP Conf. Series* 165, p. 295
- Sarajedini A., Chaboyer B., Demarque P., 1997, *PASP*, 109, 1321
- Saviane I., Rosenberg A., Piotto G., 1999, in "The Galactic Halo", *Proc. of the 3rd. Stromlo Symposium*, ed. Gibson B. et al., *ASP Conf. Series* 165, p. 308
- Searle L., Zinn R., 1978, *ApJ*, 225, 357 (SZ)
- Smoot G.F., et al., 1992, *ApJ*, 396, L1
- Somerville R.S., Primack J.R., 1999, *MNRAS*, 310, 1087
- Sommer-Larsen J., Beers T.C., Flynn C., Wilhelm R., Christiansen P.R., 1997, *ApJ*, 481, 775
- Springel V., White S.D.M., Tormen G., Kauffmann G., 2000, in preparation
- Steinmetz M., Navarro J.F., 1999, *ApJ*, 513, 555
- Stetson P.B., Vandenberg D.A., Bolte M., 1996, *PASP*, 108, 569
- Tormen G., Diaferio A., Syer D., 1998, *MNRAS*, 299, 728
- Unavane M., Wyse R.F.G., Gilmore G., 1996, *MNRAS*, 278, 727
- Vandenberg D.A., 1999, in "The Galactic Halo", *Proc. of the 3rd. Stromlo Symposium*, ed. Gibson B. et al., *ASP Conf. Series* 165, p. 46
- Vandenberg D.A., Bolte M., Stetson P.B., 1990, *AJ*, 100, 445
- van den Bosch F.C., 1998, *ApJ*, 507, 601
- White S.D.M., Rees M.J., 1978, *MNRAS*, 183, 341
- White S.D.M., Frenk C.S., 1991, *ApJ*, 379, 52
- White S.D.M., Springel V., 2000, in *Proc. of the MPA/ESO workshop "The First Stars"*, ed. Weiss A. et al., in press (astro-ph/9911378)
- Wilkinson M.I., Evans N.W., 1999, *MNRAS*, 310, 645
- Wyse R.F.G., Gilmore G., 1988, *AJ*, 95, 1404
- Wyse R.F.G., Gilmore G., 1992, *AJ*, 104, 144
- Wyse R.F.G., 2000, in "The Galactic Halo : from Globular Clusters to Field Stars", *Proc. of the 35 Liège Astroph. Coll.*, in press (astro-ph/9911358)
- Zaritsky D., Olszewski E., Schommer R.A., Peterson R.C., Aaronson M., 1989, *ApJ*, 345, 759
- Zaritsky D., 1999, in "The Galactic Halo", *Proc. of the 3rd. Stromlo Symposium*, ed. Gibson B. et al., *ASP Conf. Series* 165, p. 34
- Zinn R.J., West M.J., 1984, *ApJS*, 55, 45
- Zinn R.J., 1985, *ApJ*, 293, 424

Chapter 2

Building up the stellar halo of the Galaxy*

ABSTRACT

We study numerical simulations of satellite galaxy disruption in a potential resembling that of the Milky Way. Our goal is to assess whether a merger origin for the stellar halo would leave observable fossil structure in the phase-space distribution of nearby stars. We show how mixing of disrupted satellites can be quantified using a coarse-grained entropy. Although after 10 Gyr few obvious asymmetries remain in the distribution of particles in configuration space, strong correlations are still present in velocity space. We give a simple analytic description of these effects, based on a linearised treatment in action-angle variables, which shows how the kinematic and density structure of the debris stream changes with time. By applying this description we find that a single dwarf elliptical-like satellite of current luminosity $10^8 L_{\odot}$ disrupted 10 Gyr ago from an orbit circulating in the inner halo (mean apocentre ~ 12 kpc) would contribute about ~ 30 kinematically cold streams with internal velocity dispersions below 5 km s^{-1} to the local stellar halo. If the whole stellar halo were built by such disrupted satellites, it should consist locally of 300 – 500 such streams. Clear detection of all these structures would require a sample of a few thousand stars with 3-D velocities accurate to better than 5 km s^{-1} . Even with velocity errors several times worse than this, the expected clumpiness should be quite evident. We apply our formalism to a group of stars detected near the North Galactic Pole, and derive an order of magnitude estimate for the initial properties of the progenitor system.

*Based on: Amina Helmi and Simon D.M. White, MNRAS, 307, 495 (1999)

2.1 Introduction

There have been two different traditional views on the formation history of the Milky Way. The first model was introduced by Eggen, Lynden-Bell & Sandage (1962) to explain the kinematics of metal poor halo field stars in the Solar neighbourhood. According to their view the Galaxy formed in a monolithic way, by the free fall collapse of a relatively uniform, star-forming cloud. After the system became rotationally supported, further star formation took place in a metal-enriched disk, thereby producing a correlation between kinematics and metallicity: the well-known disk-halo transition. In later studies Searle & Zinn (1978) noted the lack of an abundance gradient and a substantial spread in ages in the outer halo globular cluster system. This led them to propose an alternative picture in which our Galaxy's stellar halo formed in a more chaotic way through merging of several protogalactic clouds. (See Freeman 1987 for a complete review).

This second model resembles more closely the view of the current cosmological theories of structure formation in the Universe. These theories postulate that structure grows through the amplification by the gravitational forces of initially small density fluctuations (Peebles 1970; White 1976; Peebles 1980, 1993). In all currently popular versions small objects are the first to collapse; they then merge forming progressively larger systems giving rise to the complex structure of galaxies and galaxy clusters we observe today. This hierarchical scenario is currently the only well-studied model which places galaxy formation in its proper cosmological context (see White 1996 for a comprehensive review). Numerical simulations of large-scale structure formation show a remarkable similarity to observational surveys (e.g. Jenkins et al. 1997, and references therein; and Efstathiou 1996 for a review). For galaxy formation, the combination of numerical and semi-analytic modelling has proved to be very powerful, despite the necessarily schematic representation of a number of processes affecting the formation of a galaxy (Katz 1992; Kauffmann, White & Guiderdoni 1993; Cole et al. 1994; Navarro & White 1994; Steinmetz & Muller 1995; Kauffmann 1996; Mo, Mao & White 1998; Somerville & Primack 1999; Steinmetz & Navarro 1999). This general framework, where structure forms bottom-up, provides the background for our work.

We are motivated, however, not only by this theoretical modelling, but also by the increasing number of observations which suggest substructure in the halo of the Galaxy (Eggen 1965; Rodgers, Harding & Sadler 1981; Rodgers & Paltoglou 1984; Ratnatunga & Freeman 1985; Sommer-Larsen & Christensen 1987; Doinidis & Beers 1989; Arnold & Gilmore 1992; Preston, Beers & Sackett 1994; Majewski, Munn & Hawley 1994, 1996). Detections of lumpiness in the velocity distribution of halo stars are becoming increasingly convincing, and the recent discovery of the Sagittarius dwarf satellite galaxy (Ibata, Gilmore & Irwin 1994) is a dramatic confirmation that accretion and merging continue to affect the Galaxy.

There have been a number of recent studies of the accretion and disruption of satellite galaxies (Quinn, Hernquist & Fullagar 1993; Oh, Lin & Aarseth 1995; Johnston, Spergel & Hernquist 1995; Velázquez & White 1995, 1999; Sellwood, Nelson & Tremaine 1998). Much of this work has been limited to objects which remain mostly in the outer parts of the Galaxy, which may be well represented by a spherical potential plus a small perturbation due to the disk (Johnston, Hernquist & Bolte 1996; Kroupa 1997; Klessen & Kroupa 1998). In this situation simple analytic descriptions of the disruption process, of the properties of the debris, etc. are possible (Johnston 1998). However, it is questionable whether such descriptions can be applied to most of the regions probed by past or current surveys of the halo, which are quite local: in this case the influence of the disk cannot be disregarded or treated as a small perturbation.

Since formation models for the Galaxy should address the broader cosmological setting, we

are naturally led to ask what should be the signatures of the different accretion events that our Galaxy may have suffered through its lifetime. Should this merging history be observable in star counts, kinematic or abundance surveys of the Galaxy? How prominent should such substructures be? How long do they survive, or equivalently, how well-mixed today are the stars which made up these progenitors? What can we say about the properties of the accreted satellites from observations of the present stellar distribution? Our own Galaxy has a very important role in constraining galaxy formation models, because we have access to 6-D information which is available for no other system. Observable structure which could strongly constrain the history of the formation of galaxies is just at hand.

In this chapter we will try to answer some of the questions just posed. We focus on the growth of the stellar halo of the Galaxy by disruption of satellite galaxies. We have run numerical simulations of this process, and have studied the properties of the debris after many orbits, long after the disruption has taken place. We analyse how the debris phase-mixes by following the growth of its entropy and the variations of the volume it fills in coordinate space. We also study the evolution of its kinematical properties. In order to model the characteristic properties of the disrupted system, such as its size, density and velocity dispersion, we develop a simple analytic prescription based on a linearised Lagrangian treatment of its evolution in action-angle variables. We apply our results to derive the observable properties of an accreted halo in the Solar neighbourhood. We also analyse the clump of halo stars detected near the North Galactic Pole (NGP) by Majewski et al. (1994), and obtain an order of magnitude estimate for the initial properties of the progenitor system.

This chapter is organized as follows. Section 2.2 presents our numerical simulations. In Section 2.3 we analyse the characteristics of the debris in these models, and in Section 2.4 we develop an analytic formalism to understand their properties. We apply this formalism to describe the characteristics of an accreted halo in this same section. In Section 2.5 we compare our modelling with the observations of Majewski et al. (1994). We leave for the last section the discussion of the results, their validity, and the potential of our approach for understanding the formation of our Galaxy.

2.2 The simulations

To study the disruption of a satellite galaxy of the Milky Way, we carry out N-body simulations in which the Galaxy is represented by a fixed, rigid potential and the satellite by a collection of particles. The self-gravity of the satellite is modelled by a monopole term as in White (1983) and Zaritsky & White (1988).

2.2.1 Model

The Galactic potential is represented by two components: a disk described by a Miyamoto–Nagai (1975) potential,

$$\Phi_{\text{disk}} = - \frac{GM_{\text{disk}}}{\sqrt{R^2 + (a + \sqrt{z^2 + b^2})^2}}, \quad (2.1)$$

where $M_{\text{disk}} = 10^{11} M_{\odot}$, $a = 6.5$ kpc, $b = 0.26$ kpc, and a dark halo with a logarithmic potential,

$$\Phi_{\text{halo}} = v_{\text{halo}}^2 \ln(r^2 + d^2), \quad (2.2)$$

with $d = 12$ kpc and $v_{\text{halo}} = 131.5 \text{ km s}^{-1}$. This choice of the parameters gives a circular velocity at the Solar radius of 210 km s^{-1} , and of 200 km s^{-1} at ~ 100 kpc.

TABLE 2.1— Orbital parameters for the different experiments of our six principal simulations.

Experiment	pericentre (kpc)	apocentre (kpc)	$z_{\max}$ (kpc)	period (Gyr)
1	10.9	51.5	25.0	0.69
2	13.5	93.1	69.1	1.23
3	5.0	51.5	5.1	0.64
4	9.2	96.5	12.0	1.24
5	0.5	45.5	30.1	0.56
6	6.0	37.0	24.8	0.48

We have taken two different initial phase-space density distributions for our satellites: *i*) two spherically symmetric Gaussian distributions in configuration and velocity space of 1 kpc (5 kpc) width and $5 - 25 \text{ km s}^{-1}$ (20 km s^{-1}) velocity dispersion, corresponding to masses of $\sim 5.9 \times 10^7 - 1.5 \times 10^9 M_{\odot}$ ($4.7 \times 10^9 M_{\odot}$); and *ii*) a Plummer profile (1911)

$$\rho(r) = \frac{\rho_0}{(r^2 + r_0^2)^{5/2}}, \quad (2.3)$$

with $\rho_0 = 3M/4\pi r_0^3$, M being the initial mass of the satellite and r_0 its scale length. In this second case, the distribution of initial velocities is generated in a self-consistent way with the density profile. For the characteristic parameters we chose $M = 10^7 - 10^9 M_{\odot}$ and $r_0 = 0.53 - 3.0 \text{ kpc}$, giving a one-dimensional internal velocity dispersion $\sigma_{1D} = 2.9 - 11.3 \text{ km s}^{-1}$.

The force on particle i due to the self-gravity of the satellite is represented by

$$\mathbf{F}(\mathbf{x}_i) = -\frac{GM_{\text{in}}}{(r_i^2 + \epsilon^2)^{3/2}} \mathbf{r}_i, \quad (2.4)$$

where M_{in} is the mass of the satellite inside $r_i = |\mathbf{x}_i - \mathbf{x}_c|$, $\mathbf{x}_c$ being the position of the expansion centre defined by a test particle with the same orbital properties as those of the satellite. The value for the softening ϵ is $0.25 r_0$. The approximation for the self-gravity of the satellite may not be very accurate during the disruption process, where tidal forces are strong and elongations in the bound parts of the satellite are expected. However, because we are interested in what happens after many perigalactic passages, well after the satellite has been tidally destroyed, our conclusions on the whole process are unaffected by details of the disruption process.

In total we ran sixteen different simulations, six of which we analyse and describe in full detail in Section 2.3. Some of the remaining simulations are used in Section 2.4 for comparison with the analytic predictions and the rest are briefly mentioned in the discussion. The characteristic properties of our six principal simulations are summarised in Table 2.1. They differ only in their orbital parameters and all initially have a Plummer profile and a mass of $10^7 M_{\odot}$. We have imposed the restriction that the orbits pass close to the Solar circle in order to be able to compare the results of the experiments with the known properties of the local stellar halo. In all cases the satellite was represented by 10^5 particles of equal mass.

In Figure 2.1 we show projections of orbits 1–6 in three orthogonal planes, where XY always coincides with the plane of the Galaxy. Notice that the plane of motion of a test particle on these orbits changes orientation substantially showing that the non-sphericity induced by the disk significantly affects the motion of the satellite.

While orbiting the Galaxy, the satellite loses all of its mass. As expected, the most dramatic effects take place during pericentric passages. The satellites do not survive very long, being

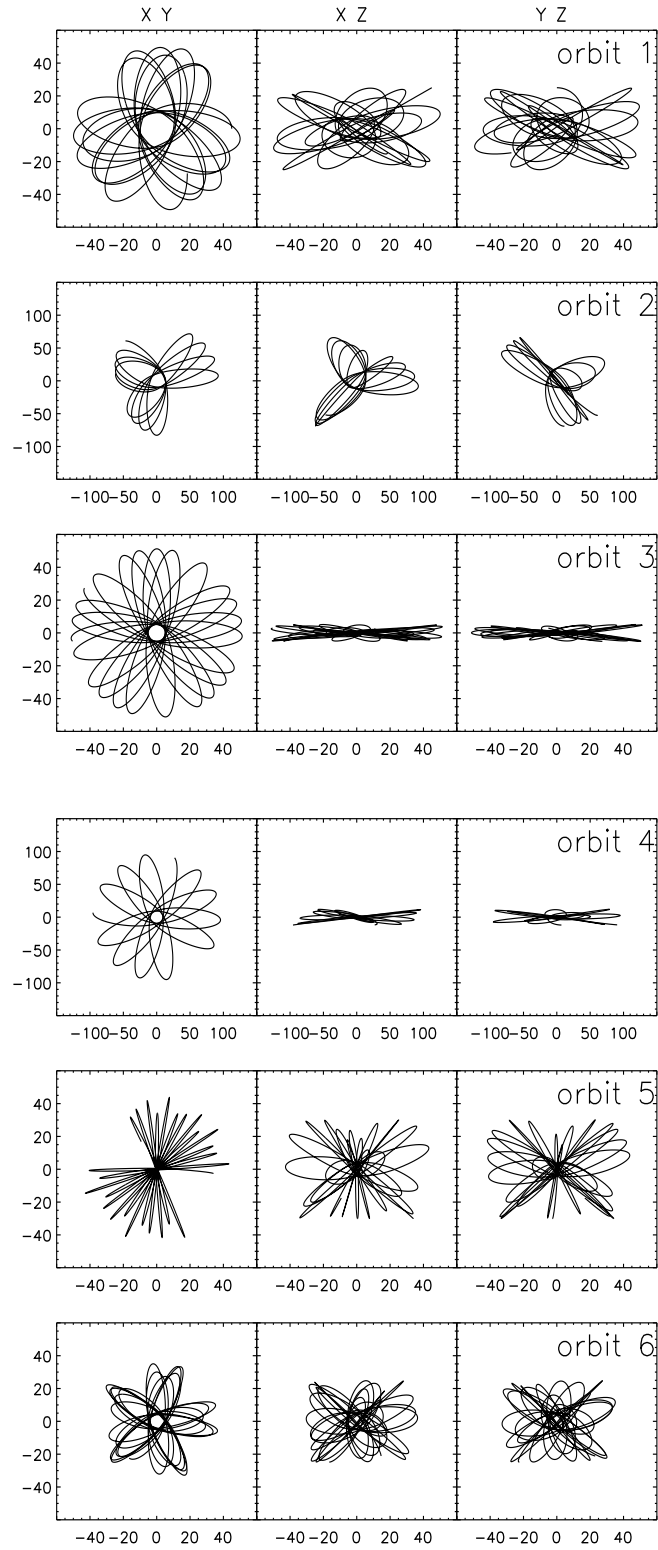


FIGURE 2.1— Projections of the orbits of our six principal simulations on different orthogonal planes, where XY coincides with the plane of the Galaxy. All distances are in kpc.

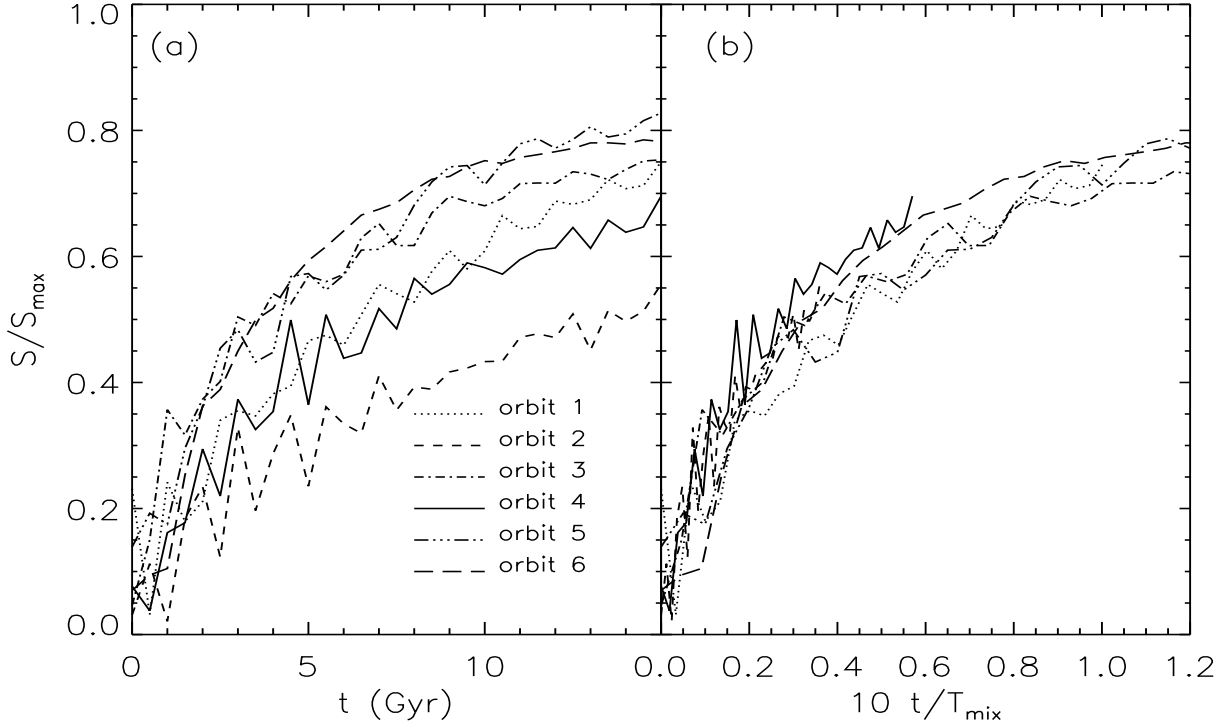


FIGURE 2.2— Evolution of the entropy of the system for the different experiments, as a function of time in (a), and scaled with the mixing timescale in (b). The dispersion in the scaled entropy is of the order of 0.06.

disrupted completely after 3 passages. This means that for our experiments, for any relatively low density satellite on an orbit which plunges deeply into the Galaxy with a period of 1 Gyr or less, the disruption itself occupies only a relatively small part of the available evolution time.

2.3 Properties of the debris: Simulations

2.3.1 Entropy as a measure of the phase-mixing

The state of a collisionless system is completely specified by its distribution function $f(\mathbf{x}, \mathbf{v}, t)$. In making actual measurements, it is often more useful to work with the coarse-grained distribution function $\langle f \rangle$, which is the average of f over small cells in phase-space. An interesting property of the coarse-grained distribution function is that it can yield information about the degree of mixing of the system (Tremaine, Hénon & Lynden-Bell 1986; Binney & Tremaine 1987).

In statistical mechanics the entropy is defined as

$$S = - \int d^3x d^3v f(\mathbf{x}, \mathbf{v}, t) \ln f(\mathbf{x}, \mathbf{v}, t). \quad (2.5)$$

Since the coarse-grained distribution function decreases as the system evolves towards a well-mixed state, an entropy calculated using $\langle f \rangle$ will increase, whereas one calculated using f will remain constant, a consequence of the collisionless Boltzmann equation: $Df/Dt = 0$. We therefore quantify the mixing state of the debris by calculating its coarse-grained entropy as a function of time. We represent the coarse-grained distribution function by taking a partition in the 6-dimensional phase-space and counting how many particles fall in each 6-D box. Naturally the size chosen for the partition and the discreteness of the simulations will affect the result. We can quantify the expected discreteness noise in the following way. The uncertainty in the entropy

can be attributed to fluctuations in the number counts, which we can estimate as Poissonian, $\propto \sqrt{N_i}$ in each occupied cell. Therefore, the uncertainty in the entropy in each cell is

$$\Delta S_i \approx \frac{\Delta N_i}{N} \left(1 + \ln \frac{N_i}{N} \right) \approx \frac{\sqrt{N_i}}{N} \ln N$$

for $N \gg 1$. The total uncertainty is thus

$$\Delta S \approx \frac{\ln N}{\sqrt{N}}, \quad (2.6)$$

which, for experiments with 10^5 particles is 0.04. In order to have a normalized measure of the mixing properties of the debris, we also computed the entropy of points equidistant in time along the corresponding orbit. After a very long integration, the orbit will fill the available region in phase-space, whose shape and size are determined by its integrals of motion. In this way, by comparing the entropy calculated for the debris with the “entropy of the orbit”, we have a measure of how well mixed the debris is. We plot this “normalized” entropy in Figure 2.2(a) as a function of time. Note that the orbits which have the shortest periods show the most advanced state of mixing, but that this is not complete after a Hubble time.

The degree of mixing basically depends on the range of orbital frequencies in the satellite, essentially as $(\Delta\nu)^{-1}$ (Merritt 1999). This means, for example, that a small satellite will disperse much more slowly than a larger one on the same orbit. On the other hand a satellite set close to a resonance will mix on a much longer time scale. One can also imagine that if there are fewer isolating integrals than degrees of freedom so that chaos might develop, a satellite located initially in a chaotic region will have a large spread $\Delta\nu$ because of the extreme sensitivity to the initial conditions. Therefore the mixing timescale (no longer a *phase*-mixing timescale) will be very short, since the neighbouring orbits diverge exponentially, instead of like power laws. If indeed the mixing rate is set by the spread in the orbital frequencies ν of the satellite, by normalizing the time variable with this timescale we should be able to derive a unique curve for the entropy evolution $S = S_{\max} f(t/T_{\text{mix}})$.

In what follows we shall assume that the behaviour of the system is regular as seems to be the case for our experiments. Let us recall that any regular motion can be expressed as a Fourier series in three basic frequencies (Binney & Spergel 1984; Carpintero & Aguilar 1998). The motion is therefore a linear superposition of waves of the basic frequencies with different amplitudes. Terms in this expansion which have the largest amplitude will be the dominant terms and may be used to define three independent (basic) frequencies. By performing a spectral dynamics analysis as outlined by Carpintero & Aguilar (1998) for ten randomly selected particles in our satellites in each experiment, we compute the frequencies associated with the largest amplitude terms in the x - (or y , since the problem is axisymmetric) and z -motions, and their dispersion around the mean. We then define

$$T_{\text{mix}}^{-1} = \min\{\sigma(\nu_x^{(1)}), \sigma(\nu_x^{(2)}), \sigma(\nu_z^{(1)}), \sigma(\nu_z^{(2)})\}, \quad (2.7)$$

where, for example, $\sigma(\nu_x^{(1)})$ denotes the dispersion in the frequency corresponding to the largest amplitude term in expansion describing the x -motion. The curves obtained by scaling time with T_{mix} are shown in Figure 2.2(b) and they can be well fitted with the function

$$\frac{S}{S_{\max}} = 0.78 - 0.69 \exp(-27.03 \frac{t}{T_{\text{mix}}}). \quad (2.8)$$

The good fit and small dispersion confirms that mixing is governed primarily by the spread in frequency.

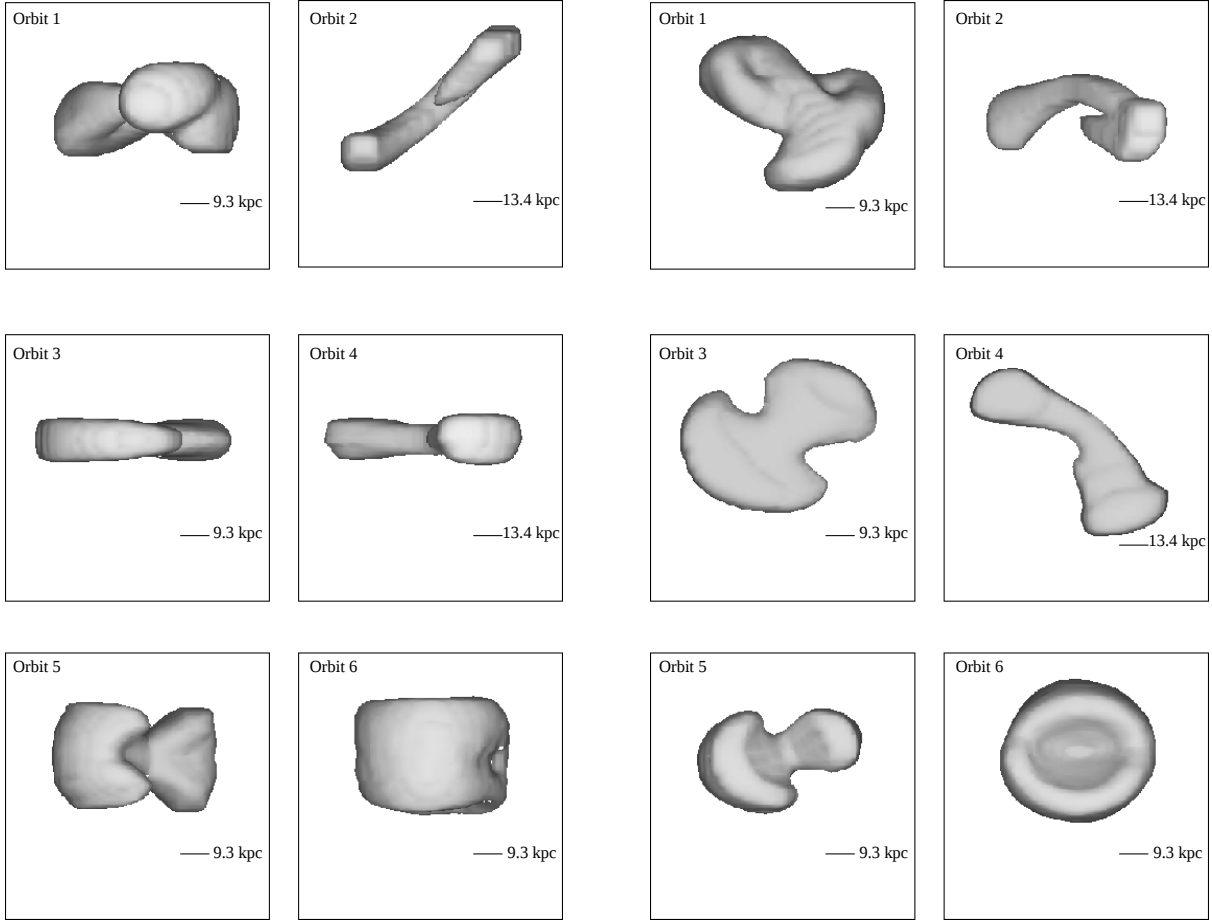


FIGURE 2.3— Isodensity surface of $10^{-6}\rho_0$ after 14 Gyr, seen from the Galactic plane (left) and from the Galactic pole (right), for the different experiments. These surfaces encompass most of the satellite’s mass.

2.3.2 Configuration space properties

To analyse the spatial properties of the debris several Gyr after disruption, we have plotted smoothed isodensity surfaces and calculated different characteristic densities. In Figure 2.3 we show the density surface at approximately 10^{-6} times the initial density of the satellite. This encompasses most of the satellite’s mass.

This density surface practically does not change over the last 2 Gyr for experiments 3, 5, 6, showing that the system has reached a stage where it fills most of its available 3-D coordinate space. The shape of this isodensity surface also gives a measure of how advanced the disruption is. The form of the accessible 3-D configuration volume is basically a torus, defined by the apocentre, pericentre and the maximum height the orbit reaches above the Galactic plane. In Figure 2.3 we clearly see that shape for experiment 6. Experiments 3 and 5 are in an intermediate state and still need to fill part of their tori. In the opposite limit, experiment 2 has filled only a small fraction of its available volume. All this is consistent with what was found using the entropy in the previous subsection. The characteristic extent of the debris is much larger than the initial size of the satellite. Moreover, debris with these properties may well span a very large solid angle on the sky as shown in Figure 2.4, and so be poorly described as a stream in coordinate space. This is the principal difference between our own experiments and those in which the Galaxy is represented by a spherical potential. In the latter the plane of motion of the

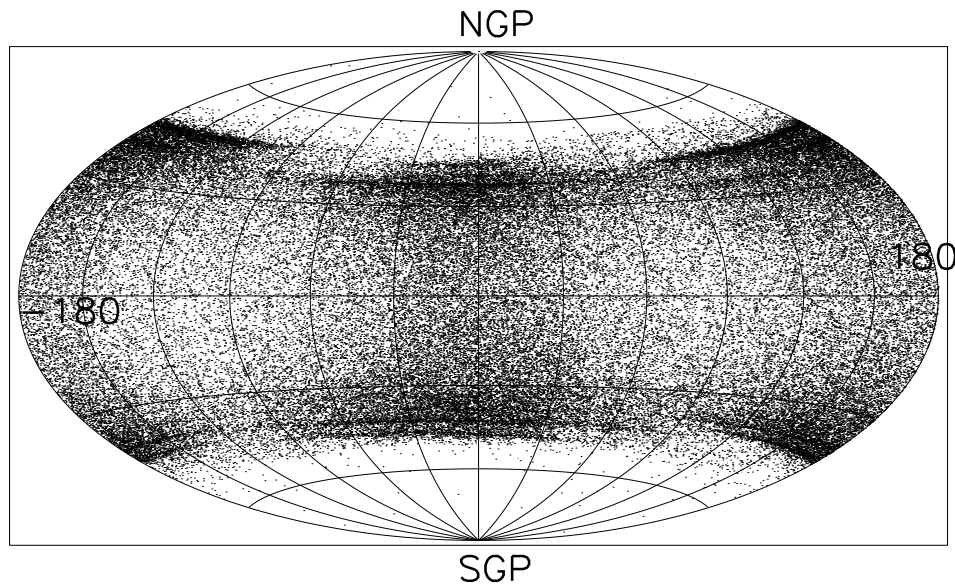


FIGURE 2.4— Sky projection of the debris from a satellite on orbit 6, 14 Gyr after infall.

satellite has a fixed orientation, and therefore all the particles have to remain fairly close to this plane, naturally giving a stream-like configuration. Late accretion events in the outer halo of the Galaxy will plausibly have this characteristic, as shown in Johnston et al. (1996) and Johnston (1998). However, similar behaviour should not be expected in the Solar neighbourhood, or even as far as 10 – 15 kpc from the Galactic centre since at such radii no strong correlations are left in the spatial distribution of satellite particles. Any method which attempts to find moving groups purely by counting stars will probably fail in this regime.

In Table 2.2, we present a summary of characteristic densities at different times which were calculated by counting particles within spheres of 0.5 kpc radii. The maximum density is achieved at the pericentre of the orbit, though most of the mass is distributed closer to the apocentre. In all cases the maximum density is between three and four orders of magnitude lower than the initial density of the satellite, and the mean density of the debris is between four and five orders of magnitude lower. These values give another estimate of the degree of mixing of the debris. Note that, in accordance with the entropy computation, experiment 6 has the smallest characteristic densities, meaning that it has reached a rather evolved state, whereas experiment 2 has high densities in comparison to the rest. The maximum density in all of the experiments is roughly comparable (similar or an order of magnitude lower) to the local density of the Milky Way’s stellar halo, though the sizes of regions where this density is reached get fairly small, a few kpc^3 , as the evolution proceeds.

2.3.3 Velocity space properties

Let us now focus on the characteristics of the debris in velocity space. We divided the 3-D coordinate space into boxes and analysed the kinematical properties of the particles inside each box. Figure 2.5 shows an example. The scatter diagrams indicate that there is a strong correlation between the different components of the velocity vector inside any given box. Notice also the large velocity range in each component when close to the Galactic centre. This shows that the debris can appear kinematically hot. As we shall see this results from a combination of multiple streams within a given box (clearly visible in Figure 2.5) and of strong gradients along each

TABLE 2.2— Characteristic densities of the debris of our principal experiments at different times.

Experiment	time Gyr	ρ_{mean} $10^2 M_{\odot} \text{ kpc}^{-3}$	ρ_{max} $10^2 M_{\odot} \text{ kpc}^{-3}$
1	5.0	67.0	886.2
	10.0	14.6	223.5
	12.5	7.0	152.8
	15.0	6.8	181.4
2	5.0	84.7	857.5
	10.0	26.5	376.2
	12.5	11.5	202.4
	15.0	9.7	288.4
3	5.0	41.5	437.4
	10.0	8.9	72.6
	12.5	8.7	181.4
	15.0	6.9	177.6
4	5.0	40.8	446.9
	10.0	5.9	99.3
	12.5	5.7	171.9
	15.0	5.1	156.6
5	5.0	36.4	996.9
	10.0	10.9	210.1
	12.5	6.1	183.3
	15.0	5.7	213.9
6	5.0	13.8	403.0
	10.0	4.3	82.1
	12.5	4.3	95.5
	15.0	3.4	63.0

stream. At a given point on a particular stream the dispersions are usually very small.

2.4 Properties of the debris: Analytical approach

In this section we will develop an analytic formalism to understand and describe the spatial and kinematical properties of the stream. Let us recall that because the disruption of the satellite occurs very early in its history, the stars that were once part of it behave as test particles in a rigid potential for most of the evolution. One of the distinguishing properties of this ensemble of particles is that it initially had a very high density in phase-space, and by virtue of Liouville's theorem, this is true at all times. At late times, however, this is no longer reflected by a strong concentration in configuration space. This evolution can be understood in terms of a mapping from the initial configuration to the final configuration, which we will describe by using the adiabatic invariants, namely the actions.

2.4.1 Action-Angle variables and Liouville's theorem

Let $H = H(\mathbf{q}, \mathbf{p})$ be the (time-independent) Hamiltonian of the problem and $(\mathbf{q}, \mathbf{p})$ a set of canonical coordinates. We wish to transform the initial set $(\mathbf{q}, \mathbf{p})$ to one in which the evolution of the system is simpler, for example, where all the momenta P_i are constant. To meet this last condition, it is sufficient to require that the new Hamiltonian be independent of the new coordinates

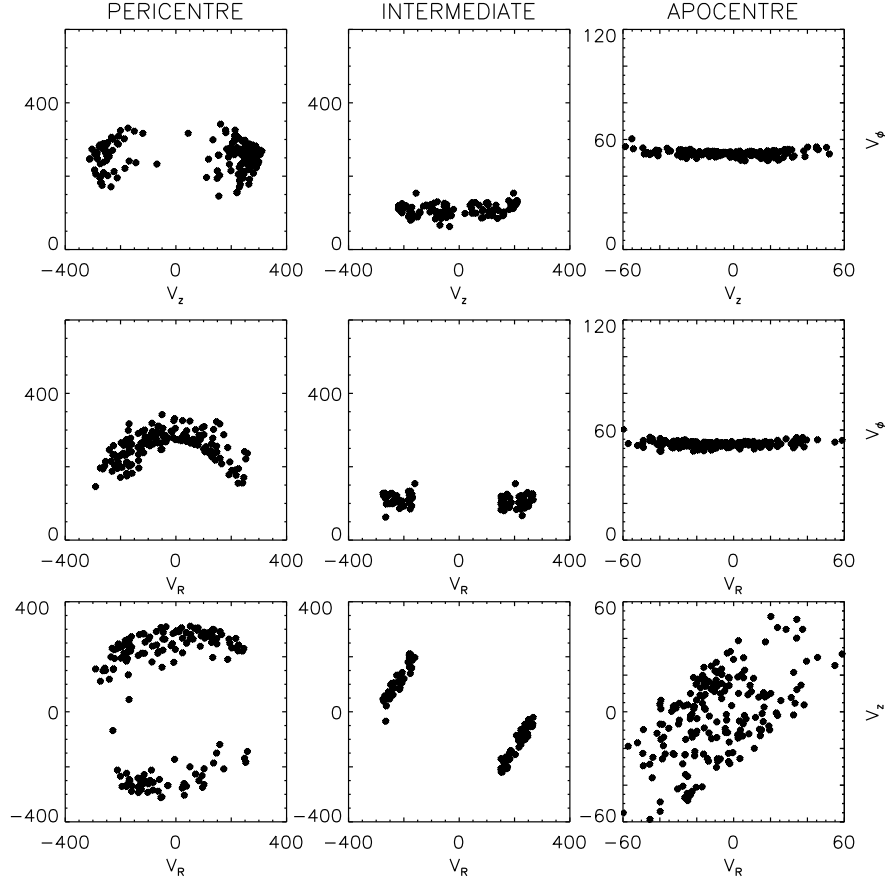


FIGURE 2.5— Scatter plots of the different velocity components (in km s^{-1}) for stars in boxes of ~ 3 kpc on a side at different locations in the Galaxy for experiment 6, 14 Gyr after infall. Similar characteristics are observed in all our experiments.

Q_i : $H = H(\mathbf{P}) = E$. The equations of motion then become

$$\dot{Q}_i = \nu_i, \quad \dot{P}_i = 0,$$

with solutions

$$Q_i = Q_i^0 + \nu_i t, \quad P_i = P_i^0.$$

The generating function that produces this transformation is known as Hamilton's Characteristic function $W(\mathbf{q}, \mathbf{P})$, and satisfies the Hamilton–Jacobi partial differential equation:

$$H(q, \frac{\partial W}{\partial q}) = E.$$

The solution to this equation involves N constants of integration α_i (including E) for a system with $2N$ degrees of freedom. Therefore, the new momenta $\mathbf{P}$ may be chosen as functions of these N constants of integration. A particularly simple situation occurs if the potential is separable in the original coordinate set $(\mathbf{q}, \mathbf{p})$. The characteristic function may then be expressed as $W = \sum_i W_i(q_i, \alpha_1 \dots \alpha_N)$, and the Hamilton–Jacobi equation breaks up into a system of N independent equations of the form:

$$H_i \left(q_i, \frac{\partial W_i}{\partial q_i}, \alpha_1 \dots \alpha_N \right) = \alpha_i,$$

each of which involves only one coordinate and the partial derivative of W_i with respect to that coordinate. The transformation relations between the original and new sets of variables are

$$p_i = \frac{\partial W}{\partial q_i}, \quad Q_i = \frac{\partial W}{\partial P_i},$$

and each component of the characteristic function is given by

$$W_i(q_i, \alpha_1 \dots \alpha_N) = \int dq'_i p_i(q'_i, \alpha_1 \dots \alpha_N). \quad (2.9)$$

(For more details, e.g. Goldstein 1953).

The actions and angle variables are a set of coordinates that describe simply the evolution of a system of particles. They are particularly useful in problems where the motion is periodic. The actions are functions of the constants α_i and are defined for a set of coordinates $(\mathbf{q}, \mathbf{p})$ as

$$J_i = \frac{1}{2\pi} \oint dq_i p_i, \quad (2.10)$$

and their conjugate coordinates, the angles, are

$$\phi_i = \frac{\partial W}{\partial J_i}. \quad (2.11)$$

The evolution of the dynamical system thus becomes:

$$\begin{aligned} \phi_i &= \phi_i^0 + \mathbf{\Omega}_i(\mathbf{J})t, \\ J_i &= J_i^0 = \text{constant}. \end{aligned} \quad (2.12)$$

The evolution of the distribution function

Let us assume that the initial distribution function of the ensemble of particles is a multivariate Gaussian in configuration and velocity space

$$f(\mathbf{x}, \mathbf{v}, t^0) = f_0 \exp \left[- \sum_{i=1}^3 \frac{(x_i - \bar{x}_i^0)^2}{2\sigma_x^2} \right] \exp \left[- \sum_{j=1}^3 \frac{(v_j - \bar{v}_j^0)^2}{2\sigma_v^2} \right],$$

which we can also express using matrices as

$$f(\mathbf{x}, \mathbf{v}, t^0) = f_0 \exp \left[- \frac{1}{2} \Delta_{\varpi}^0{}^\dagger \sigma_{\varpi}^0 \Delta_{\varpi}^0 \right]. \quad (2.13)$$

Here t^0 denotes the initial time. Δ_{ϖ}^0 is a 6-dimensional vector, with three spatial and three velocity components, and $\Delta_{\varpi}^0{}^\dagger$ is obtained by transposing Δ_{ϖ}^0 . Explicitly $\Delta_{\varpi}^0{}_i = x_i - \bar{x}_i^0$ for $i = 1..3$ and $\Delta_{\varpi}^0{}_i = v_j - \bar{v}_j^0$ for $i = j + 3 = 4..6$ in a Cartesian coordinate system. The matrix σ_{ϖ}^0 is diagonal with $\sigma_{\varpi}^0{}_{ii} = 1/\sigma_x^2$ for $i = 1..3$, and $\sigma_{\varpi}^0{}_{ii} = 1/\sigma_v^2$ for $i = 4..6$. As we shall see the matrix formulation is particularly useful to study the evolution of the distribution of particles of the system.

At the initial time, we perform a coordinate change from Cartesian to action-angle variables. Since the particles are initially strongly clustered in phase-space, a linearised transformation

can be used to obtain the distribution function of the whole system in the $(\phi, \mathbf{J})$ variables. We express this coordinate transformation as

$$\Delta_{\varpi}^0 = \mathbf{T}^0 \Delta_w^0, \quad \text{with} \quad T_{ij}^0 = \left. \frac{\partial \varpi_i}{\partial w_j} \right|_{\bar{\mathbf{x}}^0, \bar{\mathbf{v}}^0}, \quad (2.14)$$

where $\varpi = (\mathbf{x}, \mathbf{v})$, $w = (\phi, \mathbf{J})$ and the elements of matrix $\mathbf{T}^0$ are evaluated at the central point of the system, around which the expansion is performed. By substituting this in Eq.(2.13), and by defining $\sigma_w^0 = \mathbf{T}^{0\dagger} \sigma_{\varpi}^0 \mathbf{T}^0$ the distribution function in action-angle coordinates becomes

$$f(\phi, \mathbf{J}, t^0) = f_0 \exp \left[-\frac{1}{2} \Delta_w^{0\dagger} \sigma_w^0 \Delta_w^0 \right], \quad (2.15)$$

that is, it is also a multivariate Gaussian, but with dispersions now given by σ_w^0 .

The deviation of any individual orbit from the mean orbit, defined by the centre of mass or the central particle of the system, $\Delta_{w_i} = w_i - \bar{w}_i(t)$ may in turn be expressed in terms of the initial action-angle variables as

$$J_i - \bar{J}_i = J_i^0 - \bar{J}_i^0, \quad (2.16)$$

and

$$\phi_i - \bar{\phi}_i(t) = \phi_i^0 - \bar{\phi}_i^0 + \left. \frac{\partial \Omega_i}{\partial J_k} \right|_{\mathbf{J}} (J_k - \bar{J}_k) t, \quad (2.17)$$

where we expanded the difference in the frequencies to first order in $J_k - \bar{J}_k$. Eqs.(2.16) and (2.17) can also be written as

$$\Delta_w(t) = \Theta^{-1}(t) \Delta_w^0, \quad (2.18)$$

where $\Theta(t)$ is the blockmatrix:

$$\Theta(t) = \begin{bmatrix} \mathcal{I}_3 & -\Omega' t \\ \mathbf{0} & \mathcal{I}_3 \end{bmatrix}. \quad (2.19)$$

$\mathcal{I}_3$ here is the identity matrix in 3-D, and Ω' represents a 3×3 matrix whose elements are $\partial \Omega_i / \partial J_j$. The distribution function in action-angle space in the neighbourhood of the central particle at any point of its orbit $(\bar{\phi}(t), \bar{\mathbf{J}})$ is then

$$f(\phi, \mathbf{J}, t) = f_0 \exp \left[-\frac{1}{2} \Delta_w^\dagger(t) \sigma_w(t) \Delta_w(t) \right], \quad (2.20)$$

with $\Delta_w(t) = (\phi - \bar{\phi}(t), \mathbf{J} - \bar{\mathbf{J}})$ and

$$\sigma_w(t) = \Theta(t)^\dagger \sigma_w^0 \Theta(t), \quad (2.21)$$

or in terms of the original coordinates $\sigma_w(t) = (\mathbf{T}^0 \Theta(t))^\dagger \sigma_{\varpi}^0 (\mathbf{T}^0 \Theta(t))$.

Example: 1-D Case. To understand more clearly what the distribution function in Eq.(2.20) tells us with respect to the evolution of the system, we consider the 1-D case. The initial distribution function becomes:

$$f(\phi, J, t^0) = f_0 \exp \left[-\frac{(\phi - \bar{\phi}^0)^2}{2\sigma_\phi^2} - \frac{(J - \bar{J})^2}{2\sigma_J^2} - (\phi - \bar{\phi}^0)(J - \bar{J})C_{\phi J} \right],$$

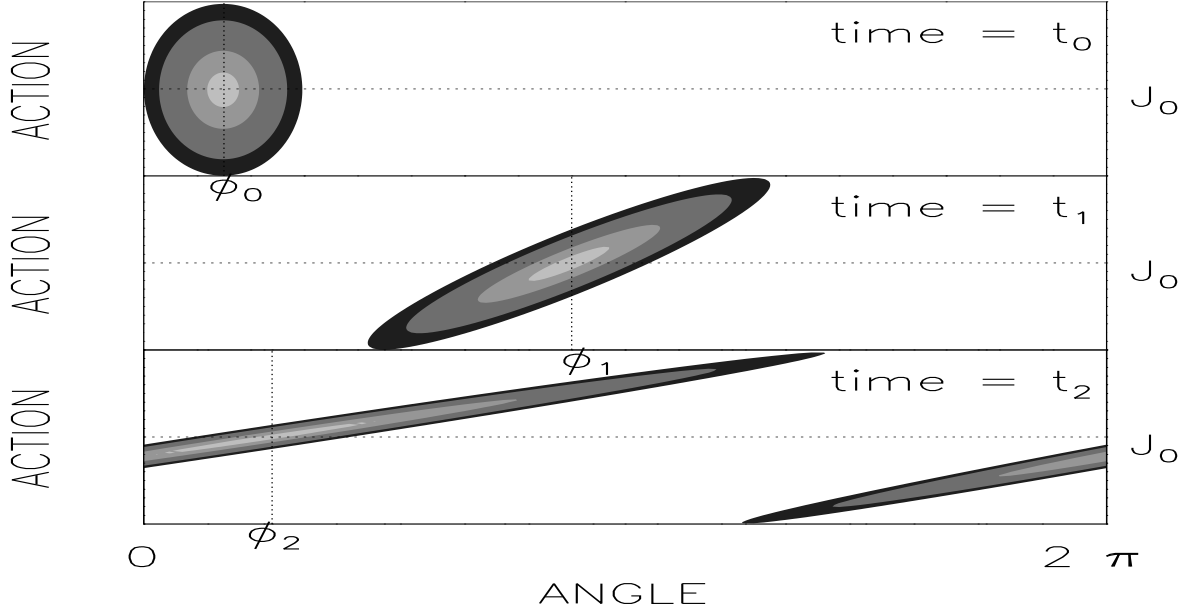


FIGURE 2.6— 1-D graphical interpretation of Liouville's theorem and the evolution of the system in phase-space. The system is initially a Gaussian in action-angle space, with no correlations between ϕ and J . As time passes by, the system evolves into an ellipsoidal configuration, with principal axes that are no longer aligned with the action or the angle directions. After a some time, the system wraps around in the angles, giving rise to phase-mixing: at the same phase we observe more than one stream, each with a small variance in the action due to the conservation of the area in phase-space.

where $C_{\phi J}$ denotes the initial correlation[†] between ϕ and J . After considering the time evolution of the system (as in Eq.(2.17)) we find

$$f(\phi, J, t) = f_0 \exp \left[-\frac{(\phi - \bar{\phi}(t))^2}{2\sigma_\phi^2} - (J - \bar{J})^2 \left(\frac{1}{2\sigma_J^2} + \frac{\Omega'^2 t^2}{2\sigma_\phi^2} \right) - (\phi - \bar{\phi}(t))(J - \bar{J}) \left(C_{\phi, J} + \frac{\Omega' t}{\sigma_\phi^2} \right) \right],$$

where $\Omega' = d\Omega/dJ$. This means that the dispersion in the J -direction effectively decreases in time and the covariance between ϕ and J increases with time. The system becomes an elongated ellipsoid in phase-space as time passes by as a consequence of the conservation of the local phase-space density. This evolution is illustrated in Figure 2.6.

The distribution function in observable coordinates

To compute the characteristic scales of a system that evolved from an initial clumpy configuration, such as satellite debris, we have to relate the dispersions in action-angle variables to dispersions in a set of observable coordinates. The transformation from the action-angle coordinate system to the observable $(\mathbf{x}, \mathbf{v})$ has to be performed locally since we generally cannot express in a simple way the global relation between the two sets of variables. Because the system has expanded so much along some directions in phase-space, the transformation from $(\phi, \mathbf{J})$ to $(\mathbf{x}, \mathbf{v})$ has to be done point to point along the orbit. This transformation is given by the inverse of $\mathbf{T}$ at time t :

$$T_{ij}^{-1} = \left. \frac{\partial w_i}{\partial \varpi_j} \right|_{\mathbf{x}, \mathbf{v}}, \quad (2.22)$$

[†] $C_{\phi J}$ is not the correlation coefficient, usually denoted as ρ . They are related through $\rho = \frac{-C_{\phi J} \sigma_\phi^2 \sigma_J^2}{1 - C_{\phi J} \sigma_\phi^2 \sigma_J^2}$.

where the derivatives are now evaluated at the particular point of the orbit around which we wish to describe the system in $(\mathbf{x}, \mathbf{v})$ coordinates. In particular, if the expansion is performed around $(\bar{\phi}(t), \bar{\mathbf{J}})$ then

$$\Delta_w(t) = \mathbf{T}^{-1} \Delta_{\varpi}(t), \quad (2.23)$$

and the distribution function may be expressed in the region around $\bar{\varpi} = (\bar{\mathbf{x}}, \bar{\mathbf{v}})$ as

$$f(\mathbf{x}, \mathbf{v}, t) = f_0 \exp \left[-\frac{1}{2} \Delta_{\varpi}(t)^{\dagger} \sigma_{\varpi}(t) \Delta_{\varpi}(t) \right], \quad (2.24)$$

with

$$\Delta_{\varpi_i}(t) = \begin{cases} x_i - \bar{x}_i(t), & i = 1..3, \\ v_j - \bar{v}_j(t), & i = j + 3 = 4..6, \end{cases} \quad (2.25)$$

and

$$\sigma_{\varpi}(t) = (\mathbf{T}^0 \Theta(t) \mathbf{T}^{-1})^{\dagger} \sigma_{\varpi}^0 (\mathbf{T}^0 \Theta(t) \mathbf{T}^{-1}). \quad (2.26)$$

We find once more that, locally, the distribution function is a multivariate Gaussian, where the variances and covariances depend on their initial values, on the time evolution of the system and on the position along the orbit where the system centre is located at time t .

If we wish to describe the properties of a group of particles that are located at a different point $\tilde{\mathbf{w}}$ than the central particle (i.e. the expansion centre does not coincide with the satellite centre at time t) a slightly different approach must be followed. The region of interest is then $\Delta_w(t) = w' - \bar{w}(t) = (w' - \tilde{w}) + (-\tilde{w}(t) + \tilde{w}) = \Delta'_w + \tilde{\mathbf{D}}(t)$. We replace this in Eq.(2.20) and write

$$f(\phi, \mathbf{J}, t) = f_0 \exp \left[-\frac{1}{2} (\Delta'_w - \tilde{\mathbf{D}}(t))^{\dagger} \sigma_w(t) (\Delta'_w - \tilde{\mathbf{D}}(t)) \right], \quad (2.27)$$

or equivalently

$$f(\phi, \mathbf{J}, t) = f'_0(t) \exp \left[-\frac{1}{2} \Delta'^{\dagger}_w \sigma_w(t) \Delta'_w - \frac{1}{2} \Delta'^{\dagger}_w \sigma_w(t) \tilde{\mathbf{D}}(t) - \frac{1}{2} \tilde{\mathbf{D}}(t)^{\dagger} \sigma_w(t) \Delta'_w \right], \quad (2.28)$$

where $f'_0(t) = f_0 \exp [-1/2 \tilde{\mathbf{D}}(t)^{\dagger} \sigma_w(t) \tilde{\mathbf{D}}(t)]$. We may now express $\Delta'_w = \mathbf{T}'^{-1} \Delta'_{\varpi}$, since the transformation is local again. The distribution function becomes

$$f(\mathbf{x}', \mathbf{v}', t) = \tilde{f}_0(t) \exp \left[-\frac{1}{2} (\Delta'_{\varpi} - \delta(t))^{\dagger} \sigma_{\varpi'}(t) (\Delta'_{\varpi} - \delta(t)) \right], \quad (2.29)$$

with

$$\delta(t) = \mathbf{T}' \tilde{\mathbf{D}}(t), \quad \sigma_{\varpi'}(t) = (\mathbf{T}'^{-1})^{\dagger} \sigma_w(t) \mathbf{T}'^{-1}, \quad (2.30)$$

and $\tilde{f}_0(t) = f'_0(t) \exp [-1/2 (\mathbf{T}'^{-1} \delta(t))^{\dagger} \sigma_w(t) \mathbf{T}'^{-1} \delta(t)]$. This means that the local distribution function is Gaussian centred around $\mathbf{x}_m = \tilde{\mathbf{x}} + \delta(t)$, which in general will not be very different from $\tilde{\mathbf{x}}$, with variances given by the elements of $\sigma_{\varpi'}(t)$. Thus the same type of behaviour as derived for the region around the system centre holds also if far from it.

The formalism here developed is completely general, but the actions will not always be easy to compute. As we mentioned briefly in the beginning of this section, this depends mainly on whether the potential is separable in some set of coordinates. We focus on the spherical case and a simple axisymmetric potential in the next section to show how this procedure can be used to describe the characteristic scales of the debris. We refer the reader to the Appendix for details of the computation.

2.4.2 Spherical potential

Analytic predictions

For a spherical potential $\Phi(r)$, the Hamiltonian is separable in spherical coordinates and depends on the actions J_φ and J_θ only through the combination $J_\varphi + J_\theta = L$. This means that the problem can be reduced to 2-D, and so we may choose a system of coordinates which coincides with the plane of motion of the satellite centre. The position of a particle is given by its angular (ψ) and radial (r) coordinates on that plane. Thus

$$\begin{aligned} L &= J_\psi = p_\psi, \\ J_r &= \frac{1}{\pi} \int_{r_1}^{r_2} dr \frac{1}{r} \sqrt{2(E - \Phi(r)) r^2 - L^2}, \end{aligned} \quad (2.31)$$

where L is the total angular momentum of the particle, E its energy and r_1 and r_2 are the turning points in the radial direction of motion. The frequencies of motion and their derivatives needed to compute the matrix $\Theta(t)$ and to obtain the time evolution of the distribution function, can be obtained by differentiating the implicit function $g = g(E, L, J_r) \equiv 0$ defined by Eq.(2.31).

Let us assume that the variance matrix[†] in action-angle variables is diagonal at $t = 0$. This simplifies the algebraic computations and, since we are only trying to calculate late-time behaviour, this assumption does not have a major influence on our results. As shown in the previous section, the evolution of the system in action-angles is obtained from $\sigma_w(t) = \Theta(t)^\dagger \sigma_w^0 \Theta(t)$. We find the properties of the debris in configuration and velocity space by transforming the action-angle coordinates $w = (\phi, \mathbf{J})$ locally to the separable $\omega = (\mathbf{x}, \mathbf{p})$, and then by transforming from $\omega = (\mathbf{x}, \mathbf{p})$ to $\varpi = (\mathbf{x}, \mathbf{v})$. That is $\sigma_\varpi(t) = \mathbf{T}'^\dagger \sigma_w(t) \mathbf{T}'$, with the $T' = T_{w \rightarrow \omega} T_{p \rightarrow v}$.

The diagonalization of the variance matrix $\sigma_\varpi(t)$ yields the values of the dispersions along the principal axes and their orientation. It can be shown that *two of the eigenvalues increase with time*, whereas the other *two decrease with time*. This is directly related to what happens in action-angle variables: as we have shown for the 1-D case, the system becomes considerably elongated along an axis which, after a very long time, is parallel to the angle direction. For 2-D (3-D), the evolution in action-angles can also be divided into two (three) independent motions (whether or not the Hamiltonian is separable), so that along each of these directions this same effect can be observed. The directions of expansion and contraction are linear combinations of the four axes ($\xi_\psi, \xi_r, \xi_{v_\psi}, \xi_{v_r}$) and, generally, none is purely spatial or a pure velocity direction.

To understand the properties of the debris in observable coordinates, we will examine what happens around a particular point in configuration space. This is equivalent to studying the velocity part of the variance matrix: $\sigma_\varpi(v)$. For example, by diagonalizing the matrix $\sigma_\varpi(v)$ we obtain the principal axes of the velocity ellipsoid at the point $\bar{\mathbf{x}}$. Its eigenvalues are the roots of $\det[\sigma_\varpi(v) - \lambda \mathcal{I}] = 0$. For $t \gg t_{\text{orb}}$

$$\begin{aligned} \lambda_1 \lambda_2 &= t^4 \left(\Omega'_{11} \Omega'_{33} - \Omega_{13}'^2 \right)^2 r^2 \frac{p_r^2}{\Omega_r'^2} \sigma_{11} \sigma_{33}, \\ \lambda_1 + \lambda_2 &= t^2 r^2 \left[\sigma_{11} \left(\Omega'_{11} - \frac{\Omega'_{13}}{\Omega_r'} \left(\Omega_\psi - \frac{L}{r^2} \right) \right)^2 + \sigma_{33} \left(\Omega'_{13} - \frac{\Omega'_{33}}{\Omega_r'} \left(\Omega_\psi - \frac{L}{r^2} \right) \right)^2 \right] \\ &\quad + t^2 \left[\sigma_{11} \Omega_{13}'^2 + \sigma_{33} \Omega_{33}'^2 \right] \frac{p_r^2}{\Omega_r'^2}, \end{aligned}$$

[†]Strictly speaking σ is the inverse of the covariance matrix. However we will loosely refer to σ as the variance matrix.

where the subindices 1 and 3 represent ψ and r respectively, and $\sigma_{ii} = 1/\sigma_{\phi_i}^2$, the initial variance in the angles. Since $\sigma(v_i) = \sqrt{1/\lambda_i}$ both directions in velocity space have decreasing dispersions on the average.

So far we did not describe how the debris is spread along the transverse direction to the plane of motion: ξ_ϑ and ξ_{v_ϑ} . This is because we reduced the problem to 2-D in configuration space. However, the problem is not really 2-dimensional since the system has a finite width in the direction transverse to the plane of motion. Now that we have understood the dynamics of the reduced problem, the generalisation to 3-D is straightforward. If the variance matrix initially is diagonal in action-angle variables, then the dispersions along ϕ_ϑ and J_ϑ do not change because the frequency of motion in the transverse direction is zero. Thus the velocity dispersion and width of the stream also remain unchanged in the direction perpendicular to the orbital plane.

By integrating Eq.(2.24) with respect to the velocities, we compute the density at the point $\bar{\mathbf{x}}$

$$\rho(\bar{\mathbf{x}}, t) = \int_{\Delta v_r} \int_{\Delta v_\varphi} \int_{\Delta v_\theta} dv_\theta dv_\varphi dv_r f(\bar{\mathbf{x}}, \mathbf{v}, t). \quad (2.32)$$

For $t \gg t_{\text{orb}}$,

$$\rho(\bar{\mathbf{x}}, t) = \frac{(2\pi)^{3/2} f_0 \sigma_{\phi_3}}{|\Omega'_{11} \Omega'_{33} - \Omega'_{13}{}^2|} \left[\sqrt{\left(\frac{1}{\sigma_{\phi_1}^2} + \frac{1}{\sigma_{\phi_2}^2} \right) \left(\frac{1}{\sigma_{J_1}^2} + \frac{1}{\sigma_{J_2}^2} \right)} \right]^{-1} \frac{\Omega_r L}{r^2 \sin \theta |p_r p_\theta|} \frac{1}{t^2}, \quad (2.33)$$

where σ_X is the initial dispersion in the quantity X . This equation shows that the density at the central point of the system decreases, on the average, as $1/t^2$. It tends to be larger near pericentre since it depends on radius as $1/r^2$; moreover it diverges at the turning points of the orbit. Even though the system evolves smoothly in action-angle variables, when this behaviour is projected onto observable space, singularities arise associated with the coordinate transformation. In action-angle variables the motion is unbounded, whereas in configuration space the particle finds itself at a “wall” near the turning points. This divergence shows up in the elements of the transformation matrix $T_{w \rightarrow \varpi}$ (Eq.(A3)), some of which tend to zero, while others diverge keeping the matrix non-singular. Because of the secular evolution of the dispersions, the intensity of the spikes will decrease with time. They are generally stronger at the pericentre of the orbit than at the apocentre, because of the $1/r^2$ dependence of the density.

A direct consequence of the secular evolution is that the characteristic sizes of the system, the width and length of the stream, will increase linearly with time, reflecting the conservation of the full 6-D phase-space density. At the turning points one of these scales becomes extremely small. In Figure 2.7 we plot the predicted behaviour of the dispersions along the principal axes of the velocity ellipsoid as a function of time. We have chosen for the initial conditions a spherically symmetric Gaussian in configuration and velocity space. We follow the evolution of the variance matrix and, in particular, of the velocity dispersions along the three principal axes at the positions of the central particle. In all panels we can clearly see the periodic behaviour associated with the orbital phase of the central particle, superimposed on the secular behaviour related to the general expansion of the system along the two directions in the orbital plane. The dispersion in the third panel is on average constant: it is in the direction perpendicular to the plane of motion. Its periodic behaviour is due to the fact that we did not start with a diagonal matrix in action-angles. The initial transformation from $(\mathbf{x}, \mathbf{v})$ to $(\phi, \mathbf{J})$ produces cross terms between all three directions. As the system evolves, and we project again onto configuration space, our 6-D ellipsoid rotates continually, producing a contribution in the

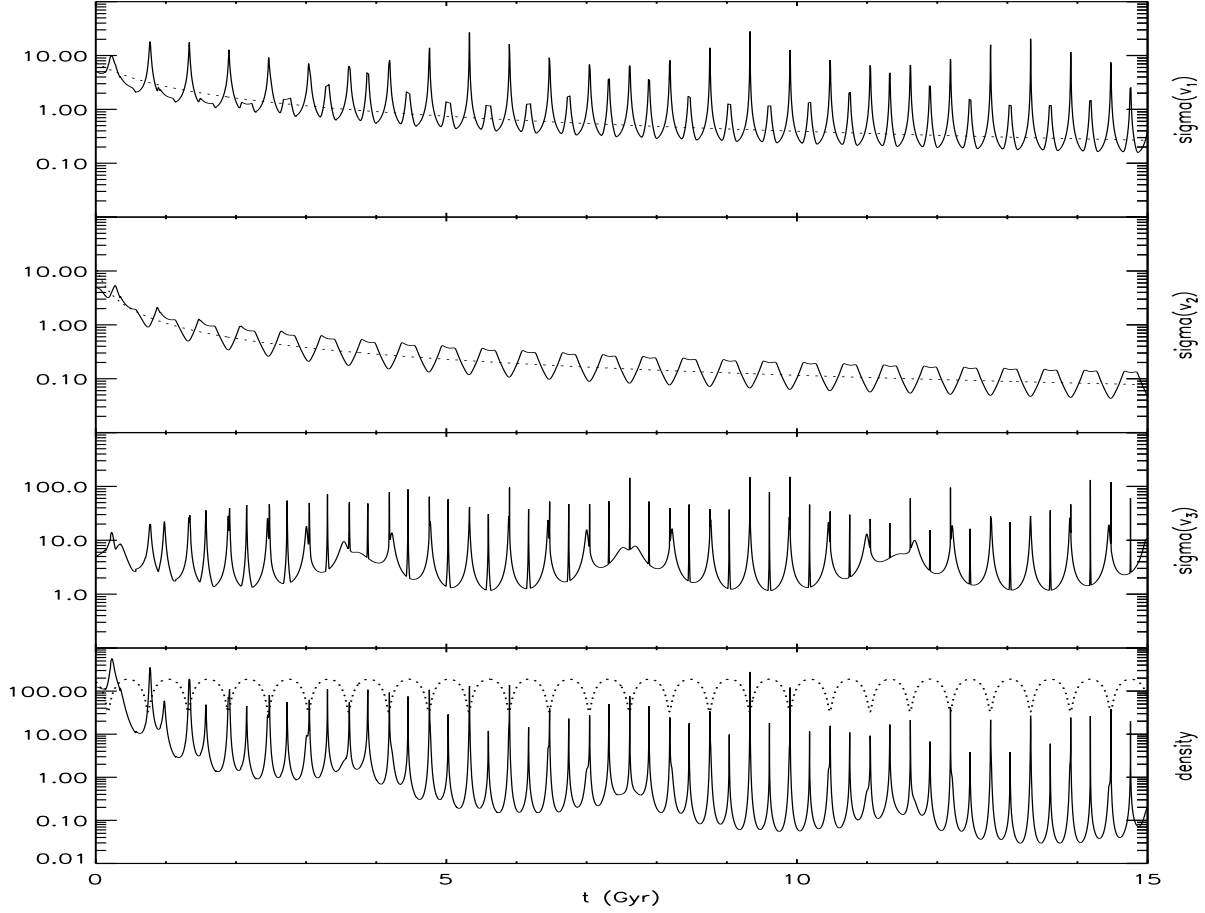


FIGURE 2.7— Time evolution of the velocity dispersions (in km s^{-1}) along the principal axes, computed as outlined in Section 2.4.2, for the logarithmic spherical potential of Eq.(2.2). Two of the dispersions decrease with time as $1/(1 + t/t_0)$ (dotted curve), whereas the third one is constant on the average. The periodic variations are due to the combination of the radial and angular oscillations, as described in the text. The last panel shows the product of the three dispersions which is proportional to the density (full curve). The radial oscillation is shown (dotted curve) so that the occurrence of density spikes can be compared with the location of the turning points of the orbit.

direction perpendicular to the orbital plane which varies with the frequencies Ω_r and Ω_θ . By fitting $\sigma(v)/\sigma_0(v) = a/(1 + t/t_0)$, we find for the velocity dispersion in the first panel $a = 1.5$ and $t_0 = 0.6$ Gyr, whereas for the dispersion in the second panel $a = 2.6$ and $t_0 = 0.1$ Gyr.

In the last panel we show the behaviour of the product of the three dispersions, which is proportional to the density (see Eq.(A10)). Note that, since two of the velocity dispersions have decreased approximately a factor of ten, the density has done so by a factor of hundred. Note also the decrease in the amplitude of the spikes and the good correlation of these with the turning points of the orbit.

Comparison to the simulations

In order to assess the limitations of our approach, we will compare our predictions with simulations of satellites with and without self-gravity. We first consider what happens to a satellite with no self-gravity moving in a spherical logarithmic potential. We take two different sets of initial properties for the satellites: 1 kpc width and $\sigma_{1D} = 5 \text{ km s}^{-1}$, corresponding to an initial mass of $\sim 5.9 \times 10^7 M_\odot$; and 5 kpc width and $\sigma_{1D} = 20 \text{ km s}^{-1}$, corresponding to

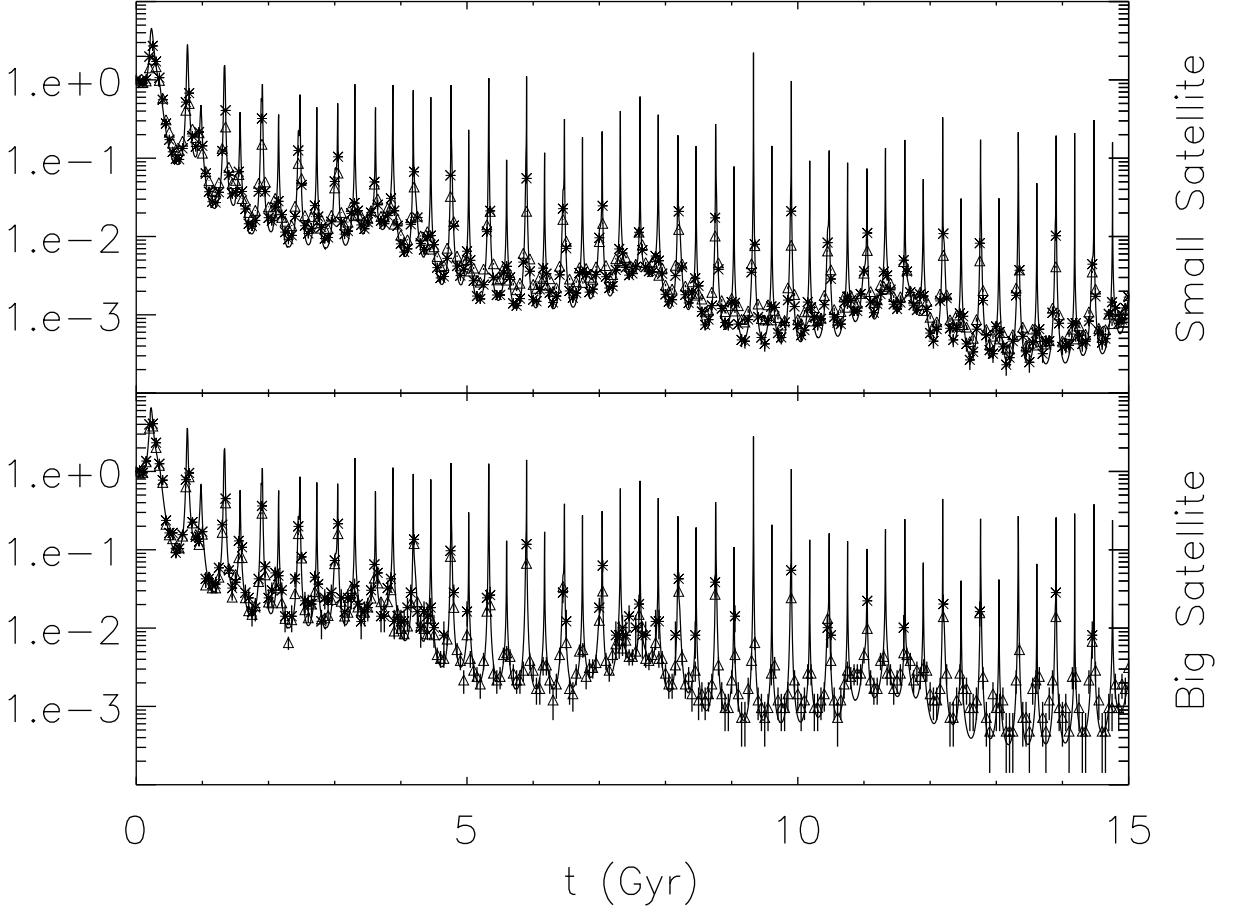


FIGURE 2.8— Time evolution of the density for a satellite moving in a spherical potential (Eq.(2.2)), with similar orbital parameters as those of Experiment 6 in Table 2.1. The full line represents our prediction, normalized to the initial density. In the upper panel we plot the density behaviour for the $\sim 5.9 \times 10^7 M_\odot$ satellite (see main text), whereas the lower panel corresponds to the $\sim 4.7 \times 10^9 M_\odot$ satellite. The stars indicate the number of particles that fall in a volume of 1 kpc on a side around the central particle of the system, and the triangles represent the number of particles in a cubic volume of twice the side, both normalized to the initial value. The spike-like behaviour occurs at the turning points of the orbit (see main text – Eq.(2.33)).

$M \sim 4.7 \times 10^9 M_\odot$ for the larger satellite. Both begin as spherically symmetric Gaussians in coordinate and velocity space. We launch them on the same orbit so that we can directly study the effects of the change in size.

What observers measure are not the velocity dispersions or densities of a stream at a particular point, but mean values given by a set of stars in a finite region. We can estimate the effects of this smoothing by comparing our analytic predictions with the simulations. In the upper panel of Figure 2.8 we show the time evolution of the density (normalized to its initial value) for the small satellite. The full line represents our prediction and the stars correspond to the simulation. We simply follow the central particle of the system as a function of time, and count the number of particles contained in a cube of 1 kpc on a side surrounding it. Triangles represent the number density from an 8 times larger volume (2 kpc on a side). The agreement between the predictions and the estimated values from the simulations is very good. The representation of a continuous field with a finite number of particles introduces some noise which, together with the smoothing, is responsible for the disagreement. Note, however, how well the simulated density spikes agree with those predicted at the orbital turning points. The overall agreement

is slightly better for the small cube than for the large one. This is due to the smoothing which inflates some of the dispersions as a result of velocity gradients along the stream.

In the lower panel of Figure 2.8 we show a similar comparison for the large satellite. In general the prediction does very well here too. Note for the small boxes and at late times, we only have simulation points at the spikes (i.e. when the density is strongly enhanced). This is because the satellite initially has a larger velocity dispersion and therefore spreads out more rapidly along its orbit.

We tested the effect of including self-gravity in the small satellite simulation, and found no significant qualitative or quantitative difference in the behaviour.

2.4.3 Axisymmetric case

As an illustrative example of the main characteristics of the axisymmetric problem, let us consider the class of Eddington potentials $\Phi(r, \theta) = \Phi_1(r) + \eta(\beta \cos \theta)/r^2$ (Lynden-Bell 1962, 1994) which are separable in spherical coordinates. The third integral for this type of potentials is $I_3 = \frac{1}{2}L^2 + \eta(\beta \cos \theta)$. The actions are computed from:

$$J_\varphi = L_z, \quad (2.34)$$

$$J_\theta = \frac{1}{2\pi} \oint d\theta \sqrt{2(I_3 - \eta(\theta)) - \frac{J_\varphi^2}{\sin^2 \theta}}, \quad (2.35)$$

$$J_r = \frac{1}{2\pi} \oint dr \sqrt{2(E - \Phi_1(r)) - \frac{2I_3}{r^2}}. \quad (2.36)$$

Since the frequencies of motion are all different and non-zero, the system has the freedom to spread along three directions in phase-space. The conservation of the local phase-space density will force the dispersions along the remaining three directions to decrease in time.

Following a similar analysis as for the spherical case we derive for the density at the central point $\bar{\mathbf{x}}(t)$ of the system at time t

$$\rho(\bar{\mathbf{x}}, t) = \frac{(2\pi)^{3/2} f_0}{\sqrt{\det \sigma_\phi^0}} \frac{1}{|\det \Omega'|} \frac{\partial I_3}{\partial J_\theta} \frac{\Omega_r}{r^2 \sin \theta} \frac{1}{|p_r p_\theta|} \frac{1}{t^3}, \quad (2.37)$$

where σ_ϕ^0 is the angle submatrix of the initial variance matrix in action-angle variables. Therefore the density at the central point of the system decreases as t^{-3} , because of the extra degree of freedom that the rupture of the spherical symmetry introduces (see Appendix B), and so after a Hubble time, the density decreases by approximately a factor of a thousand.

In Figure 2.9 we plot the time evolution of the components of the velocity ellipsoid for a system on an orbit with the same initial conditions as for the spherical case, in the potential

$$\Phi(r, \theta) = v_h^2 \log(r^2 + d^2) + \frac{\beta^2 \cos^2 \theta}{r^2}, \quad (2.38)$$

where $v_h = 123 \text{ km s}^{-1}$, $d = 12 \text{ kpc}$ and $\beta = 950 \text{ kpc km s}^{-1}$. This choice of parameters produces a reasonably flat potential which is physical (giving a positive density field) outside 7 kpc. All velocity dispersions now decrease as $1/t$.

The analytic formalism developed here can be applied to any separable potential in a straightforward manner, using the definitions and results of Sec. 2.4.1. This includes, of course, the set of Stäckel potentials which may be useful in representing the Milky Way (Batsleer & Dejonghe 1994), or any axisymmetric elliptical galaxy (de Zeeuw 1985; Dejonghe & de Zeeuw 1987). The

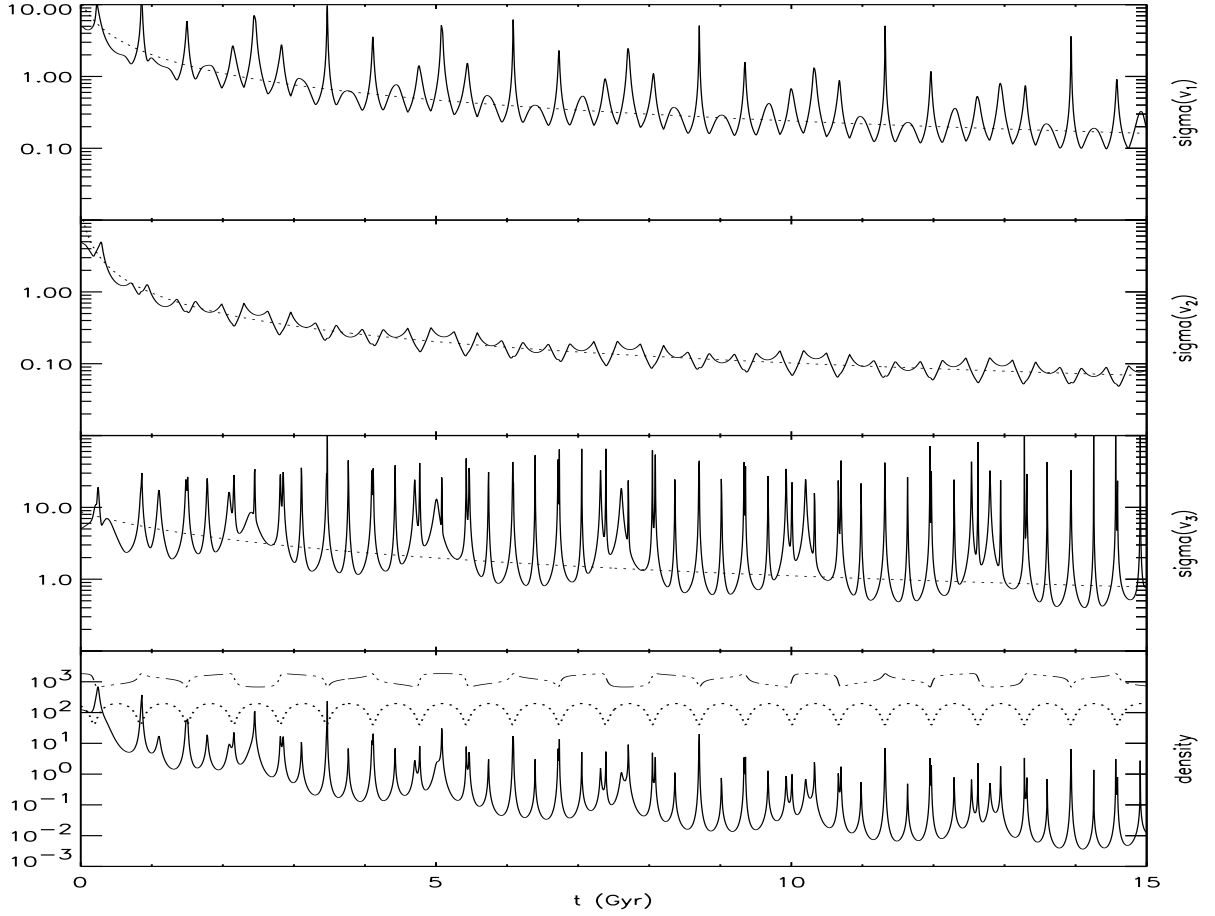


FIGURE 2.9— Time evolution of the velocity dispersions (in km s^{-1}) along the principal axes, computed as outlined in Section 2.4.2 and 2.4.3, for the simple axisymmetric potential of Eq.(2.38). Now all the dispersions decrease with time as $1/t$ (dotted curve). The periodic time behaviour is due to the combination of the radial and angular oscillations, as described in the text. The last panel shows the product of the three dispersions which is proportional to the density. The radial and θ -oscillations are also plotted to indicate the position of the turning points.

only difference is that the matrix $\mathbf{T}$ of the transformation from the usual coordinates $(\mathbf{x}, \mathbf{v})$ to the action-angle variables should be first multiplied by the matrix of the mapping from $(\mathbf{x}, \mathbf{v})$ to the ellipsoidal coordinates $(\lambda, \mu, \varphi, p_\lambda, p_\mu, p_\varphi)$, since this is the system in which the problem is separable. We discuss some of the properties Stäckel potentials and derive, for a particular model for our Galaxy, the explicit form for the density in Appendix C. Even if the potential is not separable our general results on the evolution of the system remain valid provided most orbits remain regular. In the general case the frequencies and their derivatives with respect to the actions will have to be computed through a spectral dynamics analysis similar to that used in Section 2.3.1 (Carpintero & Aguilar 1998).

2.4.4 What happens if there is phase-mixing

The procedure outlined above assumes that only one stream of debris from the satellite is present in any volume which is analysed. When phase-mixing becomes important we may find more than one kinematically cold stream near a given point. The velocity dispersions of the debris in such a region would then appear much larger than predicted naively using our formalism. We can make a rough estimate for the velocity dispersions also in this case by using

the following simple argument.

If the system is (close to) completely phase-mixed, then the coarse-grained distribution function that describes it will be uniform in the angles and therefore will only depend on the adiabatic invariants, i.e. $f(\mathbf{x}, \mathbf{v}) = f(\mathbf{J}(\mathbf{x}, \mathbf{v}))$. Since these are conserved, the moments of the coarse-grained distribution function will be given by the moments of the initial distribution function. Therefore $f(\mathbf{J})$ is completely determined by the initial properties of the system in the adiabatic invariants space. If the initial distribution function is Gaussian in action-angles then $f(\mathbf{J})$ will be Gaussian with mean and dispersion given by their values at $t = t^0$.

As an example, let us analyse the velocity dispersion in the φ -direction in a particular region in which there is a multistream structure:

$$\sigma^2(v_\varphi) = \frac{\int d^3x d^3v (v_\varphi - \bar{v}_\varphi)^2 f(\mathbf{J}(\mathbf{x}, \mathbf{v}))}{\int d^3x d^3v f(\mathbf{J}(\mathbf{x}, \mathbf{v}))} = \frac{\int d^3x d^3J \left(\frac{J_\varphi}{R} - \frac{\bar{J}_\varphi}{\bar{R}} \right)^2 f(\mathbf{J})}{\int d^3x d^3J f(\mathbf{J})},$$

where we used that $v_\varphi = J_\varphi/R$. By expanding to first order we find

$$\sigma^2(v_\phi) = \sigma^2(J_\varphi)/\bar{R}^2 + \Delta_x^2 \bar{J}_\varphi^2 / \bar{R}^4. \quad (2.39)$$

Here we replaced $\sigma(R)$ by Δ_x (the size of the region in question) which is justified by our previous result that the spatial dimensions of streams grow with time, and neglected the correlation between J_φ and R . The first term in Eq.(2.39) estimates the dispersion between streams, while the second estimates the contribution from the velocity gradient along an individual stream. For the experiments of Table 2.1 the values of the dispersions range from 50 to 150 km s⁻¹. These dispersions increase in proportion to those of the initial satellite.

The filling factor

We can use the results of our previous section to quantify the probability of finding more than one stream at a given position in space. This probability is measured by the filling factor. We define this by comparing the mass-weighted spatial density of individual streams with a mean density estimated by dividing the mass of the satellite by the total volume occupied by its orbit. The first density can be calculated formally through an integral over the initial satellite:

$$\langle \rho(t) \rangle = \frac{1}{M} \int dm(\mathbf{x}, \mathbf{v}) \rho(\mathbf{x}, \mathbf{v})(t) = \frac{1}{M} \int d^3x d^3v f(\mathbf{x}, \mathbf{v}, t^0) \rho(\mathbf{x}, \mathbf{v})(t),$$

where $\rho(\mathbf{x}, \mathbf{v})(t)$ is the density at time t of the individual stream in the neighbourhood of the particle which was initially at $(\mathbf{x}, \mathbf{v})$. The filling factor is then

$$F(t) = \frac{M}{V_o} \frac{1}{\langle \rho(t) \rangle},$$

where V_o is the volume filled by the satellite's orbit. An estimate of the filling factor can be obtained by approximating $\langle \rho(t) \rangle$ by $\rho(\bar{\mathbf{x}}, t)/(2\sqrt{2})$ taken from Eqs.(2.33), (2.37) or (C12) for spherical, axisymmetric Eddington or Stäckel potentials, respectively. The factor $1/2\sqrt{2}$ is the ratio of the central to mass-weighted mean density for a Gaussian satellite. We approximate $V_o = 4\pi r_{\text{apo}}^3 \cos \theta_f / 3$, where r_{apo} and θ_f are the maximum distance and maximum angular height above the plane reached by the satellite centre in its orbit. Since we are interested in deriving an estimate for the filling factor for the Solar neighbourhood, we focus on the Stäckel potential

described in Appendix C, which produces a flat rotation curve resembling that of the Milky Way. Thus

$$F(t) = \frac{6\sqrt{2}M\sqrt{\det\sigma_\phi^0}\langle R\rangle\langle|\nu-\lambda|v_\lambda v_\nu\rangle}{2(2\pi)^{5/2}f_0}\frac{|\det\Omega'|}{r_{\text{apo}}^3\cos\theta_f}\frac{1}{\left|\Omega_\nu\frac{\partial I_3}{\partial J_\lambda}-\Omega_\lambda\frac{\partial I_3}{\partial J_\nu}\right|}t^3, \quad (2.40)$$

where λ, ν are spheroidal coordinates (for which the potential is separable), J_λ and J_ν are the corresponding actions, and Ω_λ and Ω_ν the frequencies; and I_3 is the third integral of motion. If we approximate $\langle v_\lambda v_\nu \rangle \sim v_{\text{circ}}^2/4$ and replace $f_0 = M/(2\pi\sigma(x)\sigma(v))^3$ then

$$F(t) \sim C_{\text{orbit}}C_{\text{IC}}\left(\frac{\sigma(x)}{r_{\text{apo}}}\right)^2\frac{\sigma(v)}{v_{\text{circ}}}(\Omega_\lambda t)^3, \quad (2.41)$$

where

$$C_{\text{orbit}} = \frac{3\sqrt{\pi}\langle|\nu-\lambda|\rangle\langle R\rangle v_{\text{circ}}^5|\det\Omega'|}{2\cos\theta_f\left|\Omega_\nu\frac{\partial I_3}{\partial J_\lambda}-\Omega_\lambda\frac{\partial I_3}{\partial J_\nu}\right|\Omega_\lambda^3}, \quad (2.42)$$

depends on the orbital parameters of the satellite, and

$$C_{\text{IC}} = \frac{h_\lambda h_\nu}{\left|\Omega_\nu\frac{\partial I_3}{\partial J_\lambda}-\Omega_\lambda\frac{\partial I_3}{\partial J_\nu}\right|}\frac{\lambda-\nu}{P^3Q^3}\frac{R}{r_{\text{apo}}v_{\text{circ}}^2}\bigg|_{\bar{x}^0, \bar{v}^0}, \quad (2.43)$$

with

$$h_\tau = 2p_\tau\frac{\partial p_\tau}{\partial \tau}, \quad \tau = \lambda, \nu,$$

is a function of its initial position on the orbit. (See Appendix C for further details and definitions). This last expression holds if the satellite is initially close to a turning point of its orbit.

For example, a satellite of 10 km s^{-1} velocity dispersion and 0.4 kpc size on an orbit with an apocentric distance of 13 kpc, a maximum height above the plane of 5 kpc and an orbital period of $\sim 0.2 \text{ Gyr}$, gives an average of 0.4 streams of stars at each point in the inner halo after 10 Gyr. A satellite of 25 km s^{-1} dispersion and 1 kpc size on the same orbit would produce 5.9 streams on the average after the same time. Let us compare this last prediction with a simulation for the same satellite and the same initial conditions in the Galactic potential described in Section 2.2. In Figure 2.10 we plot the behaviour of the filling factor from the simulation, computed as

$$F(t) = \frac{N}{V_o}\frac{1}{n(t)},$$

where N is the total number of particles, $n(t) = N^{-1}\sum_i \rho_i$ with ρ_i the density of the stream where particle i is, which we calculate by dividing space up into 2 kpc boxes and counting the number of particles of each stream in each box. Note that the filling factor increases as t^3 at late times as we expect for any axisymmetric potential. Our prediction is in good agreement with the simulations, showing also that it is robust against small changes in the form of the Galactic potential.

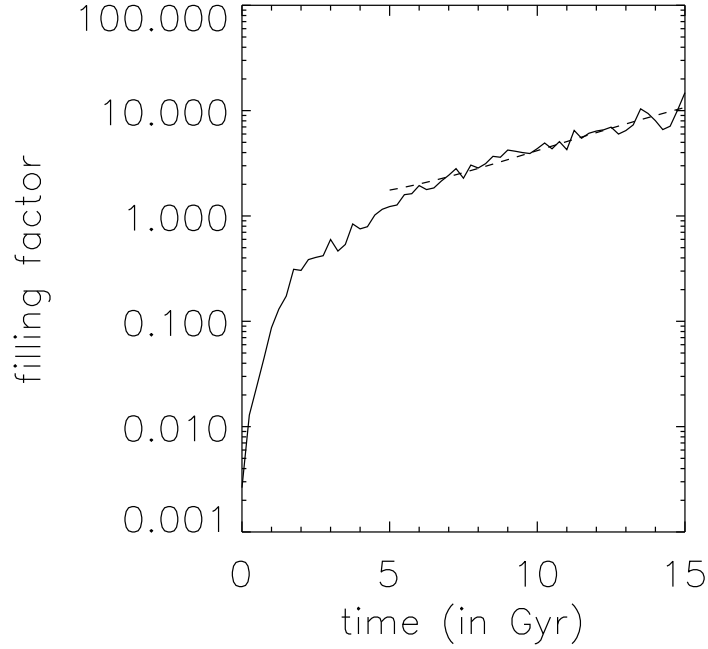


FIGURE 2.10— Time evolution of the filling factor for a satellite with an initial velocity dispersion of 25 km s^{-1} and size of 1 kpc, moving in the Galactic potential described in Section 2.2. Its orbital parameters resemble those of halo stars in the Solar neighbourhood. The dashed-curve indicates a $\gamma_0 + \gamma_1 t^3$ fit for late times.

Properties of an accreted halo in the Solar neighbourhood

To compare with the stellar halo it is more useful to derive the dependence of the filling factor on the initial luminosity of a satellite. We shall assume that the progenitor satellites are similar to present-day dwarf ellipticals, and satisfy both a Faber–Jackson relation:

$$\log \frac{L}{L_\odot} - 3.53 \log \frac{\sigma(v)}{\text{km s}^{-1}} \sim 2.35, \quad (2.44)$$

for $H_0 \sim 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and a scaling relation between the effective radius ($R_e \sim \sigma(x)$) and the velocity dispersion $\sigma(v)$:

$$\log \frac{\sigma(v)}{\text{km s}^{-1}} - 1.15 \log \frac{R_e}{\text{kpc}} \sim 1.64, \quad (2.45)$$

both as given by Guzmán, Lucey & Bower (1993) for the Coma cluster. Expressed in terms of the luminosity of the progenitor, the filling factor then becomes

$$F(t) \sim C_{\text{orbit}} C_{\text{IC}} \left(\frac{L}{L_n} \right)^{0.776} (\Omega_\lambda t)^3, \quad (2.46)$$

where L_n is a normalization constant that depends on the orbit and on the properties of the parent galaxy as:

$$L_n = 3.75 \times 10^{11} L_\odot \left(\frac{r_{\text{apo}}}{10 \text{ kpc}} \right)^{2.58} \left(\frac{v_{\text{circ}}}{200 \text{ km s}^{-1}} \right)^{1.29}. \quad (2.47)$$

If the whole stellar halo had been built from disrupted satellites, we can derive the number of streams expected in the Solar neighbourhood by adding their filling factors using the appropriate orbital parameters in Eq.(2.41) or Eq.(2.46): $F_\odot(t) = N_{\text{sat}} F(t)$. For a sample of giant

stars located within 1 kpc from the Sun with photometric distances and radial velocities measured from the ground (Carney & Latham 1986; Beers & Sommer-Larsen 1995; Chiba & Yoshii 1998), and proper motions measured by HIPPARCOS, we estimate $C_{\text{orbit}} \times C_{\text{IC}} \sim 1.29 \times 10^{-3}$. The median pericentric (apocentric) distance is 3.7 (11.6) kpc, and the median Ω_λ is 26.6 Gyr^{-1} (equivalent to a period of $\sim 0.24 \text{ Gyr}$). Thus using Eq.(2.41)

$$F_\odot(t) \sim 0.9 N_{\text{sat}} \left(\frac{\sigma(x)}{\text{kpc}} \right)^2 \frac{\sigma(v)}{\text{km s}^{-1}} \left(\frac{t}{10 \text{ Gyr}} \right)^3.$$

If now we assume that the progenitor systems are similar to present-day dwarf ellipticals, then using Eq.(2.46) we find for the whole $10^9 L_\odot$ stellar halo

$$F_\odot(t) \sim \left(\frac{t}{10 \text{ Gyr}} \right)^3 \times \begin{cases} 5.1 \times 10^2, & 100 \times 10^7 L_\odot \text{ satellites,} \\ 3.0 \times 10^2, & 10 \times 10^8 L_\odot \text{ satellites.} \end{cases} \quad (2.48)$$

For $t \sim 10 \text{ Gyr}$, the *number of streams* expected in the Solar neighbourhood is therefore in the range

$$F_\odot \sim 300 - 500. \quad (2.49)$$

Fuchs & Jahreiß (1998) have obtained a lower limit for the local mass density of spheroid dwarfs of $1 \times 10^5 M_\odot \text{ kpc}^{-3}$. We may use this estimate to derive the mass content in subdwarfs of an individual stream in a volume of 1 kpc^3 centred on the Sun:

$$F_M(t) \sim \frac{M_{\text{local halo (in 1 kpc}^3)}}{F(t)}. \quad (2.50)$$

Thus with our previous estimate for the filling factor

$$F_M(t) \sim \left(\frac{10 \text{ Gyr}}{t} \right)^3 \times \begin{cases} 1.9 \times 10^2 M_\odot, & \text{for } 10^7 L_\odot \text{ satellites,} \\ 3.3 \times 10^2 M_\odot, & \text{for } 10^8 L_\odot \text{ satellites.} \end{cases} \quad (2.51)$$

Therefore, after 10 Gyr, each stream contains $F_M \sim (200 - 350) M_\odot$ in subdwarf stars, depending on the orbital parameters of the progenitors and their initial masses.

Since the halo stars in the Solar neighbourhood have one-dimensional dispersions $\sigma_{\text{obs}}(v) \sim 100 - 150 \text{ km s}^{-1}$, in order to distinguish kinematically whether their distribution is really the superposition of $\sim 300 - 500$ individual streams of velocity dispersion $\sigma_{\text{st}}(v)$ we might require that

$$\sigma_{\text{st}}^3(v) < \frac{1}{27} \frac{\sigma_{\text{obs}}^3(v)}{F_\odot}, \quad (2.52)$$

where the factor $1/27$ would ensure a $\sim 3\sigma$ distinction between streams. Using our previous estimate of $F_\odot$ this condition becomes $\sigma_{\text{st}}(v) < \sigma_{\text{obs}}(v)/(20 - 24)$, and thus $\sigma_{\text{st}}(v) < 5 \text{ km s}^{-1}$. Currently the observational errors in the measured velocities of halo stars are of order 20 km s^{-1} , and thus there is little hope to distinguish at the present day all the individual streams which may make up the stellar halo of our Galaxy. Since intrinsic velocity dispersions for streams originating from $10^7 - 10^8 L_\odot$ objects are of the order of $3 - 5 \text{ km s}^{-1}$ after 10 Gyr, it should be possible to distinguish such streams with the astrometric missions SIM and GAIA, if they reach their planned accuracy of a few km s^{-1} . Even with an accuracy of 15 km s^{-1} per velocity component, streams are predicted to be marginally separated. The clumpy nature of the distribution should thus be easily distinguishable in samples of a few thousand stars. One way of identifying streams which are debris from the same original object, is through clustering in action or integrals of motion space (Helmi, Zhao & de Zeeuw 1998).

2.5 An observational application

Majewski et al. (1994) discovered a clump of nine halo stars in a proper motion survey near the NGP (Majewski 1992), which appeared separated from the main distribution of stars in the field. They measured proper motions, photometric parallaxes, F magnitudes and $(J - F)$ colours for all nine stars and radial velocities for six of them. For these six stars we find for the mean velocity $\bar{v}_\varphi = -152 \pm 23 \text{ km s}^{-1}$, $\bar{v}_R = -260 \pm 18 \text{ km s}^{-1}$ and $\bar{v}_z = -76 \pm 18 \text{ km s}^{-1}$, and for the velocity dispersions $\sigma(v_\varphi) = 99 \pm 33 \text{ km s}^{-1}$, $\sigma(v_R) = 100 \pm 24 \text{ km s}^{-1}$ and $\sigma(v_z) = 35 \pm 24 \text{ km s}^{-1}$. If the dispersions are computed along the principal axes, we find $\sigma(v_1) = 29 \pm 20 \text{ km s}^{-1}$, $\sigma(v_2) = 68 \pm 94 \text{ km s}^{-1}$, $\sigma(v_3) = 125 \pm 5 \text{ km s}^{-1}$.

Since the mean velocities are significantly different from zero, the group of stars can not be close to any turning point of their orbit. The only way to understand the large observed dispersions, in particular of $\sigma(v_3)$, if the stars come from a single disrupted satellite, is for the group to consist of more than one stream of stars. We believe that this may actually be the case. By computing the angular momenta of the stars we find they cluster into two clearly distinguishable subgroups: $\bar{L}_z^{(1)} = -784$ and $\sigma^{(1)}(L_z) = 299$, and $\bar{L}_z^{(2)} = -2180$ and $\sigma^{(2)}(L_z) = 313$ in kpc km s^{-1} . If we accept the existence of two streams as a premise, we may compute the velocity dispersions in each of them. We find for the stream with 4 stars

$$\sigma^{(1)}(v_1) = 25 \pm 25, \sigma^{(1)}(v_2) = 43 \pm 62, \sigma^{(1)}(v_3) = 100 \pm 45,$$

while for the stream with 2 stars

$$\sigma^{(2)}(v_1) = 3 \pm 4, \sigma^{(2)}(v_2) = 25 \pm 21, \sigma^{(2)}(v_3) = 89 \pm 64,$$

all in km s^{-1} . These results are consistent at a 2σ level with very small 3-D velocity dispersions, as expected, if indeed these are streams from a disrupted satellite.

With this interpretation of the kinematics of this group, we can estimate the mass of the progenitor and its initial size and velocity dispersion. Galaxies today obey scaling laws of the Faber-Jackson or Tully-Fisher type. If we assume that the original satellite was similar to present-day dwarf ellipticals, then we may use Eq.(2.45) to derive a relation between the initial dispersion in the z -component of the angular momentum and initial velocity dispersion of the progenitor

$$\sigma_i^2(L_z) = \sigma_i^2(v) R_{\text{apo}}^2 + 0.0375^2 \frac{L_z^2}{R_{\text{apo}}^2} \sigma_i^{1.74}(v), \quad (2.53)$$

where R_{apo} is the apocentric distance of its orbit. Under the assumption that L_z is conserved, we can derive $\sigma_i(v)$ by replacing in the previous equation the observed values of L_z , $\sigma(L_z)$ and an estimate of R_{apo} . We obtain the latter by orbit integration in a Galaxy model, which includes a disk, bulge and halo and find $R_{\text{apo}} \sim 12 \text{ kpc}$. Our estimate for the initial velocity dispersion of the progenitor is then

$$\sigma_i(v) \sim 48 \text{ km s}^{-1}, \quad (2.54)$$

which in Eqs.(2.44) and (2.45) yields for its initial luminosity and size

$$L \sim 2 \times 10^8 L_\odot, \quad R \sim 1 \text{ kpc}. \quad (2.55)$$

We estimate that the relative error-bars in these quantities are of order 50%, if measurement errors and a 50% uncertainty in the apocentric distance are included.

In summary, if indeed these stars come from a single disrupted object, we must accept that the first six stars that were detected (Majewski et al. 1994) are part of at least two independent streams. This seems reasonable, since two streams can be indeed be distinguished, and

the velocity dispersions, in each stream are very small. Moreover, a disrupted object with the properties just derived (luminosity, initial size and velocity dispersion), would fill its available volume rapidly, producing a large number of streams. In view of our explanation, a number of stars from the same disrupted object but with positive z -velocities should also be present in the same region, since phase-mixing allows streams to be observed with opposite motion in the R and/or z directions. Candidates for such additional debris should have similar v_ϕ , since L_z is conserved during phase-mixing. By simple inspection of Figure 1(a) in Majewski et al. (1994), other stars can be indeed found, with similar v_ϕ but opposite v_R and v_z .

2.6 Discussion and conclusions

We have studied the disruption of satellite galaxies in a disk + halo potential and characterized the signatures left by such events in a galaxy like our own. We developed an analytic description based on Liouville's theorem and on the very simple evolution of the system in action-angle variables. This is applicable to any accretion event if self-gravity is not very important and as long as the overall potential is static or only adiabatically changing. Satellites with masses up to several times $10^9 M_\odot$ are likely to satisfy this adiabatic condition if the mass of the Galaxy is larger than several times $10^{10} M_\odot$ at the time of infall and if there are no other strong perturbations. Even though we have not studied how the system gets to its starting point, it seems quite plausible that in this regime dynamical friction will bring the satellites to the inner regions of the Galaxy in a few Gyr, where they will be disrupted very rapidly. Their orbital properties may be similar to those found in CDM simulations of the infall structure onto clusters, where objects are mostly on fairly radial orbits (Tormen, Diaferio & Syer 1998); this is consistent with the dynamics of Solar neighbourhood halo stars. Their masses range from the low values estimated observationally for dwarf spheroidals to the much larger values expected for the building blocks in hierarchical theories of galaxy formation.

We summarise our conclusions as follows. After 10 Gyr we find no strong correlations in the spatial distribution of a satellite's stars, since for orbits relevant to the bulk of the stellar halo this is sufficient time for the stars to fill most of their available configuration volume. This is consistent with the fact that no stream-like density structures have so far been observed in the Solar neighbourhood. On the contrary, strong correlations are present in velocity space. The conservation of phase-space density results in velocity dispersions at each point along a stream that decrease as $1/t$. On top of the secular behaviour, periodic oscillations are also expected: at the turning points of the orbit the velocity dispersions, and thus the mean density of the stream, can be considerably enhanced. Some applications of this density enhancement deserve further study. For example, the present properties of the Sagittarius dwarf galaxy seem difficult to explain, since numerical simulations show that it could have been disrupted very rapidly given its current orbit (Johnston, Spergel & Hernquist 1995; Velázquez & White 1995). This puzzle has led to some unconventional suggestions to explain its survival, like a massive and very dense dark matter halo (Ibata & Lewis 1998) or a recent collision with the Magellanic Clouds (Zhao 1998). However, since the densest part of Sagittarius seems to be near its pericentre, it could be located sufficiently close to a "caustic" to be interpreted as a transient enhancement. Sagittarius could simply be a galaxy disrupted several Gyr ago (c.f. Kroupa 1997).

If the whole stellar halo of our Galaxy was built by merging of N_{sat} similar smaller systems of characteristic size $\sigma(x)$ and velocity dispersion $\sigma(v)$, then after 10 billion years we expect the stellar distribution in the Solar neighbourhood to be made up of $F_\odot$ streams, where

$$F_\odot \sim 0.9 N_{\text{sat}} \left(\frac{\sigma(x)}{\text{kpc}} \right)^2 \frac{\sigma(v)}{\text{km s}^{-1}}.$$

For satellites which obey the same scaling relations as the dwarf elliptical galaxies, this means 300 to 500 streams. Individually, these streams should have extremely small velocity dispersions, and inside a 1 kpc^3 volume centred on the Sun, each should contain a few hundred stars. Since the local halo velocity ellipsoid has dispersions of the order of 100 km s^{-1} , 3-D velocities with errors smaller than 5 km s^{-1} are needed to separate unambiguously the individual streams. This is better by a factor of four than most current measurements, which would, however, be good enough to give a clear detection of the expected clumpiness in samples of a few thousand stars. The combination of a strongly mixed population with relatively large velocity errors yields an apparently smooth and Gaussian distribution in smaller samples. Since the intrinsic dispersion for a stream from an LMC-type progenitor is of the order of $3 - 5 \text{ km s}^{-1}$ after a Hubble time, one should aim for velocity uncertainties below 3 km s^{-1} . With the next generation of astrometric satellites (in particular GAIA, e.g. Gilmore et al. 1998) we should be able to distinguish almost all streams in the Solar neighbourhood originating from disrupted satellites.

Our analytic approach is based on Liouville's Theorem and the very simple evolution of the system in action-angle variables. Although the latter is likely to fail in the full merging regime, the conservation of local phase-space density will still hold. It will be interesting to see how this conservation law influences the final phase-space distribution in the merger of more massive disk-like systems. These are plausible progenitors for the bulge of our Galaxy in hierarchical models.

Acknowledgments

A.H. wishes to thank the hospitality of MPA, HongSheng Zhao for many very useful discussions, Tim de Zeeuw for comments on earlier versions of this manuscript and Daniel Carpintero for kindly providing the software for the spectral analysis used in Section 2.3.1. EARA has provided financial support for A.H. visits to MPA.

References

- Arnold R., Gilmore G., 1992, *MNRAS*, 257, 225
- Batsleer P., Dejonghe H., 1994, *A&A*, 287, 43
- Beers T.C., Sommer-Larsen J., 1995, *ApJS*, 96, 175
- Binney J., Spergel D.N., 1984, *MNRAS*, 206, 159
- Binney J., Tremaine S., 1987, *Galactic Dynamics*, Princeton University Press (Princeton, NJ)
- Carney B.W., Latham D.W., 1986, *AJ*, 92, 60
- Carpintero D.D., Aguilar L.A., 1998, *MNRAS*, 298, 1
- Chiba M., Yoshii Y., 1998, *AJ*, 115, 168
- Cole S., Aragón-Salamanca A., Frenk C.S., Navarro J.F., Zepf S.E., 1994, *MNRAS*, 271, 781
- Dejonghe H., de Zeeuw P.T., 1988, *ApJ*, 333, 90
- de Zeeuw P.T., 1985, *MNRAS*, 216, 273
- Doinidis S.P., Beers T.C., 1989, *ApJ*, 340, L57
- Efstathiou G., 1996, in "Cosmology and Large Scale Structure", Les Houches, Session LX, ed. Schaeffer R. et al., Elsevier (Amsterdam), p. 133
- Eggen O.J., 1965, in *Galactic Structure, Stars and Stellar Systems*, volume 5, eds. Blaauw A. & Schmidt M., University of Chicago Press, p. 111
- Eggen O.J., Lynden-Bell D., Sandage A.R., 1962, *ApJ*, 136, 748
- Freeman K.C., 1987, *ARA&A*, 25, 603
- Fuchs B., Jahreiß H., 1998, *A&A*, 329, 81

- Gilmore G.F., Perryman M.A., Lindegren L., Favata F., Hoeg E., Lattanzi M., Luri X., Mignard F., Röser S., de Zeeuw P.T., 1998, *SPIE*, 3350, 541
- Goldstein H., 1953, *Classical Mechanics*, Addison-Wesley (Cambridge, Mass.)
- Guzmán R., Lucey J.R., Bower R.G., 1993, *MNRAS*, 265, 731
- Helmi A., Zhao H.S., de Zeeuw P.T., 1998, in “The Galactic Halo”, *Proc. of the 3rd. Stromlo Symposium*, ed. Gibson B. et al., *ASP Conf. Series* 165, p. 125 (see also Chapter 4)
- Ibata R., Gilmore G., Irwin M., 1994, *Nature*, 370, 194
- Ibata R., Lewis G.F., 1998, *ApJ*, 500, 575
- Jenkins A., Frenk C.S., Thomas P.A., Colberg J.M., White S.D.M., Couchman H.M.P., Peacock J.A., Efstathiou G., Nelson A.H., 1997, *ApJ*, 499, 20
- Johnston K.V., Spergel D.N., Hernquist L., 1995, *ApJ*, 451, 598
- Johnston K.V., Hernquist L., Bolte M., 1996, *ApJ*, 465, 278
- Johnston K.V., 1998, *ApJ*, 495, 297
- Katz N., 1992, *ApJ*, 391, 502
- Kauffmann G., White S.D.M., Guiderdoni B., 1993, *MNRAS*, 264, 201
- Kauffmann G., 1996, *MNRAS*, 281, 487
- Klessen R., Kroupa P., 1998, *ApJ*, 498, 143
- Kroupa P., 1997, *NewA*, 2, 77
- Lynden-Bell D., 1962, *MNRAS*, 124, 9
- Lynden-Bell D., 1994 in “The Formation and Evolution of Galaxies, V Canary Islands Winter School of Astrophysics”, eds. Muñoz-Tuñón C. & Sánchez F., Cambridge University Press, p. 85
- Majewski S.R., 1992, *ApJS*, 78, 87
- Majewski S.R., Munn J.A., Hawley S.L., 1994, *ApJ*, 427, L37
- Majewski S.R., Munn J.A., Hawley S.L., 1996, *ApJ*, 459, L73
- Merritt D., 1999, *PASP*, 111, 129
- Miyamoto M., Nagai R., 1975, *PASJ*, 27, 533
- Mo H.J., Mao S., White S.D.M., 1998, *MNRAS*, 295, 319
- Navarro J.F., White S.D.M., 1994, *MNRAS*, 267, 401
- Oh K.S., Lin D.N.C., Aarseth S.J., 1995, 442, 1420
- Peebles P.J.P., 1970, *AJ*, 75, 13
- Peebles P.J.P., 1980, *The Large-Scale Structure of the Universe*, Princeton University Press (Princeton, NJ)
- Peebles P.J.P., 1993, *Principles of Physical Cosmology*, Princeton University Press (Princeton, NJ)
- Plummer H.C., 1911, *MNRAS*, 71, 460
- Preston G., Beers T.C., Shectman S., 1994, *AJ*, 108, 538
- Quinn P.J., Hernquist L., Fullagar D., 1993, *ApJ*, 403, 74
- Ratnatunga K., Freeman K.C., 1985, *ApJ*, 291, 260
- Rodgers A.W., Harding P., Sadler E., 1981, *ApJ*, 244, 912
- Rodgers A.W., Paltoglou G., 1984, *ApJ*, 283, L5
- Searle L., Zinn R., 1978, *ApJ*, 225, 357
- Sellwood J.A., Nelson R.W., Tremaine S., 1998, *ApJ*, 506, 509
- Sommer-Larsen J., Christiansen P.R., 1987, *MNRAS*, 225, 499
- Somerville R., Primack J., 1999, *MNRAS*, 310, 1087
- Steinmetz M., Muller E., 1995, *MNRAS*, 276, 549
- Steinmetz M., Navarro J., 1999, *ApJ*, 513, 555
- Tormen G., Diaferio A., Syer D., 1998, *MNRAS*, 299, 728
- Tremaine S., Hénon, M., Lynden-Bell D., 1986, *MNRAS*, 219, 285

- Velázquez H., White S.D.M., 1995, MNRAS, 275, 23L
 Velázquez H., White S.D.M., 1999, MNRAS, 304, 254
 White S.D.M., 1976, MNRAS, 177, 717
 White S.D.M., 1983, ApJ, 274, 53
 White S.D.M., 1996, in “Cosmology and Large Scale Structure”, Les Houches, Session LX, ed. Schaeffer R. et al., Elsevier (Amsterdam), p. 349
 Zaritsky D., White S.D.M., 1988, MNRAS, 235, 289
 Zhao H.S., 1998, ApJ, 500, L149

Appendix A Spherical potential

2-D case

For a spherical potential $\Phi(r)$, the Hamiltonian is separable in spherical coordinates and depends on the actions J_ϕ and J_θ only through the combination $J_\phi + J_\theta = L$. We therefore may choose a system of coordinates which coincides with the plane of motion of the system, reducing the problem to 2-D. The position of a particle is given by its angular ψ and radial r coordinates on that plane. In that case, we have

$$L = J_\psi = p_\psi, \quad J_r = \frac{1}{\pi} \int_{r_1}^{r_2} dr \frac{1}{r} \sqrt{2(E - \Phi(r))r^2 - L^2} \quad (\text{A1})$$

where L is the total angular momentum of the particle, E its energy and r_1 and r_2 are the turning points in the radial direction of motion. The action J_r cannot be computed analytically in general for an arbitrary potential. However, Eq.(A1) defines an implicit function $g = g(E, L, J_r) \equiv 0$, which we can differentiate to find the frequencies of motion and their derivatives. These are needed to compute the elements of the matrix $\Theta(t)$ (Eq.(2.19)) and to obtain the time evolution of the distribution function.

To simplify the computations, we assume that the variance matrix[§] in action-angle variables is diagonal at $t = 0$: $\sigma_{w_{ij}}^0 = \sigma_{ii}\delta_{ij}$. The evolution of the system in phase-space is obtained through the product $\Theta(t)^\dagger \sigma_w^0 \Theta(t)$, which yields the following variance matrix $\sigma_w(t)_{ij} = \{i, j\}$ at time t

$$\sigma_w(t) = \begin{bmatrix} \sigma_{11} & 0 & -\sigma_{11}\Omega'_{11}t & -\sigma_{11}\Omega'_{12}t \\ \{1,2\} & \sigma_{22} & -\sigma_{22}\Omega'_{12}t & -\sigma_{22}\Omega'_{22}t \\ \{1,3\} & \{2,3\} & \sigma_{11}\Omega'_{11}{}^2t^2 + \sigma_{22}\Omega'_{12}{}^2t^2 + \sigma_{33} & \Omega'_{11}\sigma_{11}\Omega'_{12}t^2 + \Omega'_{12}\sigma_{22}\Omega'_{22}t^2 \\ \{1,4\} & \{2,4\} & \{3,4\} & \sigma_{11}\Omega'_{12}{}^2t^2 + \sigma_{22}\Omega'_{22}{}^2t^2 + \sigma_{44} \end{bmatrix}$$

in action-angle variables, with $\Omega'_{ij} = \partial\Omega_i/\partial J_j$. Subindices $\{1\}$ and $\{3\}$ refer to directions associated to ψ , such as for example ϕ_ψ and J_ψ , whereas $\{2\}$ and $\{4\}$ are related to r .

We find the properties of the debris in configuration and momenta space by transforming the action-angle coordinates locally around $\bar{\mathbf{x}}$ with the matrix $\mathbf{T}^{-1}$. Its elements are the second derivatives of the characteristic function $W(\mathbf{q}, \mathbf{J})$:

$$\mathbf{T}^{-1} = \begin{bmatrix} \mathbf{W}_{JJ}\mathbf{J}_q + \mathbf{W}_{Jq} & \mathbf{W}_{JJ}\mathbf{J}_p \\ \mathbf{J}_q & \mathbf{J}_p \end{bmatrix} \quad (\text{A2})$$

[§]As we mentioned in Section 2.4.2, σ is in fact the inverse of the covariance matrix. However we refer to σ as the variance matrix.

with $\mathbf{J}_q = -\mathbf{W}_{qJ}^{-1}\mathbf{W}_{qq}$ and $\mathbf{J}_p = \mathbf{W}_{qJ}^{-1}$, and has the following form for a spherical potential in 2-D

$$\mathbf{T}^{-1} = \begin{bmatrix} 1 & t_{12} & t_{13} & t_{14} \\ 0 & t_{22} & t_{23} & t_{24} \\ 0 & 0 & 1 & 0 \\ 0 & t_{42} & t_{43} & t_{44} \end{bmatrix} \quad (\text{A3})$$

with

$$\begin{aligned} t_{12} &= -\frac{h(r)}{\Omega_r} \frac{\partial^2 W}{\partial L \partial J_r} + \frac{1}{p_r} \left(\Omega_\psi - \frac{L}{r^2} \right), & t_{13} &= \frac{\partial^2 W}{\partial L^2} + \frac{\partial^2 W}{\partial L \partial J_r} t_{43}, & t_{14} &= \frac{\partial^2 W}{\partial L \partial J_r} \frac{p_r}{\Omega_r}, \\ t_{22} &= -\frac{h(r)}{\Omega_r} \frac{\partial^2 W}{\partial J_r^2} + \frac{\Omega_r}{p_r}, & t_{23} &= \frac{\partial^2 W}{\partial L \partial J_r} + \frac{\partial^2 W}{\partial J_r^2} t_{43}, & t_{24} &= \frac{\partial^2 W}{\partial J_r^2} \frac{p_r}{\Omega_r}, \\ t_{42} &= -\frac{h(r)}{\Omega_r}, & t_{43} &= -\frac{1}{\Omega_r} \left(\Omega_\psi - \frac{L}{r^2} \right), & t_{44} &= \frac{p_r}{\Omega_r} \end{aligned}$$

and

$$h(r) = -\Phi'(r) + \frac{L^2}{r^3}, \quad p_r = \sqrt{2(E - \Phi(r)) - \frac{L^2}{r^2}},$$

where all functions are evaluated at $\bar{\mathbf{x}}$. Therefore the variance matrix in $(\mathbf{x}, \mathbf{p})$ is

$$\sigma_\omega = (\Theta(t)\mathbf{T}^{-1})^\dagger \sigma_\omega^0 (\Theta(t)\mathbf{T}^{-1}), \quad (\text{A4})$$

so that, by substituting

$$\sigma_\omega = \begin{bmatrix} \sigma_{11} & \sigma_{11}A & \sigma_{11}B & \sigma_{11}C \\ \{1,2\} & \sigma_{11}A^2 + \sigma_{22}D^2 + \sigma_{44}t_{42}^2 & A\sigma_{11}B + D\sigma_{22}E + t_{42}\sigma_{44}t_{43} & A\sigma_{11}C + D\sigma_{22}F + t_{42}\sigma_{44}t_{44} \\ \{1,3\} & \{2,3\} & \sigma_{11}B^2 + \sigma_{22}E^2 + \sigma_{33} + \sigma_{44}t_{43}^2 & B\sigma_{11}C + E\sigma_{22}F + t_{43}\sigma_{44}t_{44} \\ \{1,4\} & \{2,4\} & \{3,4\} & \sigma_{11}C^2 + \sigma_{22}F^2 + \sigma_{44}t_{44}^2 \end{bmatrix}$$

and where

$$\begin{aligned} A &= t_{12} - \Omega'_{12}t_{42}t, & B &= t_{13} - \Omega'_{11}t - \Omega'_{12}t_{43}t, & C &= t_{14} - \Omega'_{12}t_{44}t, \\ D &= t_{22} - \Omega'_{22}t_{42}t, & E &= t_{23} - \Omega'_{12}t - \Omega'_{22}t_{43}t, & F &= t_{24} - \Omega'_{22}t_{44}t. \end{aligned}$$

In general, one is more interested in the characteristics of the debris in velocity space, rather than in momenta space. Thus we transform the variance matrix according to $\sigma_\omega = T_{p \rightarrow v}^\dagger \sigma_\omega T_{p \rightarrow v}$, with

$$\mathbf{T}_{p \rightarrow v} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & v_\psi & r & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (\text{A5})$$

The diagonalization of the variance matrix σ_ω yields the values of the dispersions along the principal axes and their orientation: two of its eigenvalues increase with time, whereas the other two decrease with time. To understand the directly observable properties of the debris we examine what happens around a particular point $\bar{\mathbf{x}}(t)$ in configuration space located on the mean orbit of the system. This is equivalent to studying the velocity submatrix

$$\sigma_\omega(v) = \begin{bmatrix} r^2(\sigma_{11}B^2 + \sigma_{22}E^2 + \sigma_{33} + \sigma_{44}t_{43}^2) & r(B\sigma_{11}C + E\sigma_{22}F + t_{43}\sigma_{44}t_{44}) \\ \{1,2\} & \sigma_{11}C^2 + \sigma_{22}F^2 + \sigma_{44}t_{44}^2 \end{bmatrix}.$$

For example, by diagonalizing the matrix $\sigma_{\varpi}(v)$ we obtain the directions of the principal axes of the velocity ellipsoid at the point $\bar{\mathbf{x}}(t)$, and their dispersions. Its eigenvalues are the roots of $\det[\sigma_{\varpi}(v) - \lambda I] = 0$. For $t \gg t_{\text{orb}}$

$$\lambda_i = \frac{t^2}{2} \left\{ r^2 \left[\sigma_{11} \left(\Omega'_{11} - \frac{\Omega'_{12}}{\Omega_r} \left(\Omega_{\psi} - \frac{L}{r^2} \right) \right)^2 + \sigma_{22} \left(\Omega'_{12} - \frac{\Omega'_{22}}{\Omega_r} \left(\Omega_{\psi} - \frac{L}{r^2} \right) \right)^2 \right] + \left(\frac{p_r}{\Omega_r} \right)^2 \left[\sigma_{11} \Omega_{12}'^2 + \sigma_{22} \Omega_{22}'^2 \right] \pm \sqrt{R} \right\}, \quad (\text{A6})$$

for $i = 1, 2$, and where

$$R = \left\{ r^2 \left[\sigma_{11} \left(\Omega'_{11} - \frac{\Omega'_{12}}{\Omega_r} \left(\Omega_{\psi} - \frac{L}{r^2} \right) \right)^2 + \sigma_{22} \left(\Omega'_{12} - \frac{\Omega'_{22}}{\Omega_r} \left(\Omega_{\psi} - \frac{L}{r^2} \right) \right)^2 \right] - \left(\frac{p_r}{\Omega_r} \right)^2 \left[\sigma_{11} \Omega_{12}'^2 + \sigma_{22} \Omega_{22}'^2 \right] \right\}^2 + 4r^2 \left(\frac{p_r}{\Omega_r} \right)^2 \left[\sigma_{11} \Omega'_{12} \left(\Omega'_{11} - \frac{\Omega'_{12}}{\Omega_r} \left(\Omega_{\psi} - \frac{L}{r^2} \right) \right) + \sigma_{22} \Omega'_{22} \left(\Omega'_{12} - \frac{\Omega'_{22}}{\Omega_r} \left(\Omega_{\psi} - \frac{L}{r^2} \right) \right) \right]^2. \quad (\text{A7})$$

Therefore

$$\begin{aligned} \lambda_1 \lambda_2 &= \left[t^2 (\Omega'_{11} \Omega'_{22} - \Omega_{12}'^2) t_{44} \right]^2 \sigma_{11} \sigma_{22} r^2, \\ \lambda_1 + \lambda_2 &= t^2 r^2 \left[\sigma_{11} (\Omega'_{11} + \Omega_{12}' t_{43})^2 + \sigma_{22} (\Omega'_{12} + \Omega_{22}' t_{43})^2 \right] + t^2 \left[\sigma_{11} \Omega_{12}'^2 + \sigma_{22} \Omega_{22}'^2 \right] t_{44}^2. \end{aligned}$$

Since $\sigma(v_i) = \sqrt{1/\lambda_i}$ both velocity dispersions decrease on the average as $1/t$. The principal axes of the ellipsoid rotate as time passes by, not being coincident with any particular direction.

3-D treatment

As we discussed in Section 2.4.2, the problem of the disruption of the system and its evolution in phase-space is really a 3-D problem, since our initial satellite had a finite width in all directions. Since we just discussed in great detail what happens in the 2-D case and the way of proceeding once more dimensions are added is the same, we will simply outline our main results, focusing on what happens to the velocity submatrix.

If we assume that the system had initially a diagonal variance matrix in action-angle variables, the velocity submatrix at time t is

$$\sigma_{\varpi}(v) = \mathbf{T}_{p \rightarrow v}^\dagger \mathbf{S}(t) \mathbf{T}_{p \rightarrow v}, \quad (\text{A8})$$

with

$$\mathbf{S}(t) = (\mathbf{W}_{JJ} \mathbf{W}_{qJ}^{-1} - t \Omega' \mathbf{W}_{qJ}^{-1})^\dagger \sigma_\phi^0 (\mathbf{W}_{JJ} \mathbf{W}_{qJ}^{-1} - t \Omega' \mathbf{W}_{qJ}^{-1}) + \mathbf{W}_{qJ}^{-1 \dagger} \sigma_J^0 \mathbf{W}_{qJ}^{-1}, \quad (\text{A9})$$

where $\mathbf{W}_{JJ}$ is the matrix whose elements are the second derivatives of the characteristic function with respect to the actions, $\mathbf{W}_{qJ}$ the matrix that contains the second derivatives of W with respect to the coordinates $\mathbf{q}$ and the actions $\mathbf{J}$, and $\Omega'_{ij} = \partial \Omega_i / \partial J_j$. Note that, since the potential is spherical Ω' and $\mathbf{W}_{JJ}$ have two equal rows. The initial variance matrix in action-angle space

$$\sigma_w^0 = \begin{bmatrix} \sigma_\phi^0 & \mathbf{0} \\ \mathbf{0} & \sigma_J^0 \end{bmatrix}.$$

We can compute the density at a later time at the point $\bar{\mathbf{x}}(t)$ located on the mean orbit of the system by integrating

$$f(\bar{\mathbf{x}}, \mathbf{v}, t) = f_0 \exp \left[-\frac{1}{2} \Delta_{\varpi}^{\dagger}(t) \sigma_{\varpi}(t) \Delta_{\varpi}(t) \right],$$

with respect to the velocities using the submatrix $\sigma_{\varpi}(v)$

$$\rho(\bar{\mathbf{x}}, t) = \int_{\Delta v_r} \int_{\Delta v_{\theta}} \int_{\Delta v_{\varphi}} dv_{\varphi} dv_{\theta} dv_r f(\bar{\mathbf{x}}, \mathbf{v}, t).$$

In the principal axes frame

$$\rho(\bar{\mathbf{x}}, t) = f_0 (2\pi)^{3/2} \sigma_{v_1}(t) \sigma_{v_2}(t) \sigma_{v_3}(t) \text{Erf} \left[\frac{a_1}{\sqrt{2} \sigma_{v_1}(t)} \right] \text{Erf} \left[\frac{a_2}{\sqrt{2} \sigma_{v_2}(t)} \right] \text{Erf} \left[\frac{a_3}{\sqrt{2} \sigma_{v_3}(t)} \right], \quad (\text{A10})$$

with a_1, a_2, a_3 the boundaries of the integration volume. For $t \gg t_{\text{orb}}$ the error function tends to 1, and therefore

$$\rho(\bar{\mathbf{x}}, t) = (2\pi)^{3/2} f_0 \sigma_{v_1}(t) \sigma_{v_2}(t) \sigma_{v_3}(t), \quad (\text{A11})$$

that is equivalent to

$$\rho(\bar{\mathbf{x}}, t) = (2\pi)^{3/2} f_0 / \sqrt{\lambda_1 \lambda_2 \lambda_3}, \quad (\text{A12})$$

where the λ 's are the eigenvalues of $\sigma_{\varpi}(v)$. With simple algebra it can be shown that

$$\lambda_1 \lambda_2 \lambda_3 = \det \sigma_{\varpi}(v), \quad (\text{A13})$$

which is readily computable from Eqs.(A8) and (A9)

$$\det \sigma_{\varpi}(v) = (\det \mathbf{T}_{p \rightarrow v})^2 (\det \mathbf{W}_{qJ}^{-1})^2 \det[(\mathbf{W}_{JJ} - t \mathbf{\Omega}')^{\dagger} \sigma_{\phi}^0 (\mathbf{W}_{JJ} - t \mathbf{\Omega}') + \sigma_J^0], \quad (\text{A14})$$

where

$$\det \mathbf{T}_{p \rightarrow v} = r^2 \sin \theta \quad (\text{A15})$$

and

$$\det \mathbf{W}_{qJ}^{-1} = \frac{p_r}{\Omega_r} \frac{p_{\theta}}{L}. \quad (\text{A16})$$

The remaining determinant in Eq.(A14) for $t \gg t_{\text{orb}}$ is

$$\sigma_{33}(\sigma_{11} + \sigma_{22})(\sigma_{44} + \sigma_{55})(\Omega'_{11} \Omega'_{33} - \Omega'_{13})^2 t^4,$$

so that finally

$$\rho(\bar{\mathbf{x}}, t) = \frac{(2\pi)^{3/2} f_0}{|\Omega'_{11} \Omega'_{33} - \Omega'_{13}|^2} \frac{\Omega_r L}{\sqrt{\sigma_{33}(\sigma_{11} + \sigma_{22})(\sigma_{44} + \sigma_{55})}} \frac{1}{r^2 \sin \theta |p_r p_{\theta}|} \frac{1}{t^2}. \quad (\text{A17})$$

Let us recall that $\sigma_{ii} = 1/\sigma_{\phi_i}^2$ for $i = 1..3$ and $\sigma_{ii} = 1/\sigma_{j_j}^2$ for $i = j + 3 = 4..6$.

Appendix B Axisymmetric Eddington potential

To exemplify and understand how the rupture of the spherical symmetry affects the characteristic scales of the system, we take a very simple Eddington potential $\Phi(r, \theta) = \Phi_1(r) + \eta(\beta \cos \theta)/r^2$ (Lynden-Bell 1962; 1994) which is separable in spherical coordinates. The third integral for this class of potentials is $I_3 = \frac{1}{2}L^2 + \eta(\beta \cos \theta)$. The actions are computed from:

$$J_\varphi = L_z, \quad (\text{B1})$$

$$J_\theta = \frac{1}{2\pi} \oint d\theta \sqrt{2(I_3 - \eta(\theta)) - \frac{J_\varphi^2}{\sin^2 \theta}}, \quad (\text{B2})$$

$$J_r = \frac{1}{2\pi} \oint dr \sqrt{2(E - \Phi_1(r)) - \frac{2I_3}{r^2}}. \quad (\text{B3})$$

The procedure outlined in Section 2.4.1 and Appendix A can also be applied to a system moving in this type of potentials. In particular we are interested in the behaviour of the density. By virtue of the previous discussion we only need to find the determinant of the variance matrix as in Eq.(A14), for this potential. Since Eqs.(A15) and (A16) remain unchanged, we only focus on $\det[(\mathbf{W}_{JJ} \mathbf{W}_{qJ}^{-1} - t \boldsymbol{\Omega}' \mathbf{W}_{qJ}^{-1})^\dagger \sigma_\phi^0 (\mathbf{W}_{JJ} \mathbf{W}_{qJ}^{-1} - t \boldsymbol{\Omega}' \mathbf{W}_{qJ}^{-1}) + \sigma_J^0]$. For $t \gg t_{\text{orb}}$ the term with $\boldsymbol{\Omega}'$ will dominate with respect to $\mathbf{W}_{JJ}$, and the product $t^2 (\boldsymbol{\Omega}' \mathbf{W}_{qJ}^{-1})^\dagger \sigma_\phi^0 \boldsymbol{\Omega}' \mathbf{W}_{qJ}^{-1}$ will dominate over σ_J^0 [¶]. Therefore

$$\det \sigma_\omega(v) = (\det \mathbf{T}_{p \rightarrow v})^2 (\det \mathbf{W}_{qJ}^{-1})^2 (\det \boldsymbol{\Omega}' t)^2 \det \sigma_\phi^0, \quad (\text{B4})$$

and so the density at the point $\bar{\mathbf{x}}$ at time t is

$$\rho(\bar{\mathbf{x}}, t) = \frac{(2\pi)^{3/2} f_0}{\sqrt{\det \sigma_\phi^0}} \frac{1}{|\det \boldsymbol{\Omega}'|} \frac{\partial I_3}{\partial J_\theta} \frac{\Omega_r}{r^2 \sin \theta} \frac{1}{|p_r p_\theta|} \frac{1}{t^3}. \quad (\text{B5})$$

This expression is valid for a satellite described initially by a Gaussian distribution. The variance matrix at $t = t^0$ may be

1. diagonal in action-angle variables:

$$\det \sigma_\phi^0 = 1/(\sigma_{\phi_1} \sigma_{\phi_2} \sigma_{\phi_3})^2,$$

2. diagonal in configuration-velocity space:

$$\begin{aligned} \det \sigma_\phi^0 &= \frac{p_\theta^2 p_r^2}{\Omega_r^2 (\partial I_3 / \partial J_\theta)^2} \frac{1}{\sigma_\varphi^2 \sigma_v^2} \left\{ \frac{1}{\sigma_\theta^2} \left[W_{rr}^2 + \frac{v_\varphi^2 + v_\theta^2}{r^2} \right] \right. \\ &\quad \left. + \frac{1}{\sigma_v^2} \left[W_{rr}^2 \frac{W_{\theta\theta}^2}{r^2} + v_\varphi^2 \left(\frac{\cos^2 \theta}{\sin^2 \theta} W_{rr}^2 + \left(\frac{v_\theta \cos \theta}{r \sin \theta} - \frac{W_{\theta\theta}}{r^2} \right)^2 \right) \right] \right\}, \end{aligned}$$

where all functions are evaluated at $(\bar{\mathbf{x}}^0, \bar{\mathbf{v}}^0)$, and

$$W_{\theta\theta} = \frac{h_\theta}{p_\theta}, \quad h_\theta = -\eta'(\theta) + J_\varphi^2 \frac{\cos \theta}{\sin^3 \theta},$$

and

$$W_{rr} = \frac{h_r}{p_r}, \quad h_r = -\Phi_1'(r) + \frac{2I_3}{r^3}.$$

[¶]This does not hold for the spherical case because $\det[(\boldsymbol{\Omega}' \mathbf{W}_{qJ}^{-1})^\dagger \sigma_\phi^0 \boldsymbol{\Omega}' \mathbf{W}_{qJ}^{-1}] \propto \det \boldsymbol{\Omega}' \equiv 0$

The expression for the determinant of the angle submatrix at $t = 0$ may be simplified if the satellite is initially close to a turning point of the orbit. In this case the term $W_{\theta\theta}^2 W_{rr}^2$ will be dominant and

$$\det \sigma_\phi^0 = \left[\frac{h_\theta(\theta^0) h_r(r^0)}{\Omega_r r^0} \left(\frac{\partial I_3}{\partial J_\theta} \right)^{-1} \frac{1}{\sigma_\varphi \sigma_v^2} \right]^2.$$

Note that the main differences with the spherical case are

- the time dependence: t^3 instead of t^2 because of the increase in the dimensionality of the problem;
- the dependence on the derivatives of the basic frequencies of motion: the same functional dependence $\det \Omega'$, but now with three independent frequencies and derivatives;
- the inclusion of the term $\partial I_3 / \partial J_\theta$, which for the spherical case is simply L ;
- the form of $p_\theta = \sqrt{2(I_3 - \eta(\theta)) - J_\varphi^2 / \sin^2 \theta}$, which also includes the angular dependence of the potential.

Appendix C Axisymmetric Stäckel potential

In this section we collect some basic properties of Stäckel potentials and derive the density behaviour as a function of time, as in previous sections, from Liouville's Theorem and the evolution of the system in action-angle variables. Further details on Stäckel potentials can be found in de Zeeuw (1985).

Let us first introduce spheroidal coordinates (λ, ν, φ) , where φ is the azimuthal angle in the usual cylindrical coordinates (R, z, φ) , and λ and ν are the two roots for τ of

$$\frac{R^2}{\tau - a^2} + \frac{z^2}{\tau - c^2} = 1, \quad (\text{C1})$$

where $c^2 \leq \nu \leq a^2 \leq \lambda$. A potential is of Stäckel form if it can be expressed as

$$V = - \frac{(\lambda - c^2)G(\lambda) - (\nu - c^2)G(\nu)}{\lambda - \nu}, \quad (\text{C2})$$

where $G(\tau)$ is an arbitrary function ($\tau = \lambda, \nu$). In this case, the Hamiltonian becomes

$$H = \frac{p_\lambda^2}{2P^2} + \frac{p_\nu^2}{2Q^2} + \frac{p_\varphi^2}{2R^2} + V(\lambda, \nu) \quad (\text{C3})$$

where the functions P and Q are

$$P^2 = \frac{\lambda - \nu}{4(\lambda - a^2)(\lambda - c^2)}, \quad Q^2 = \frac{\nu - \lambda}{4(\nu - a^2)(\nu - c^2)}. \quad (\text{C4})$$

Three isolating integrals of motion can be found (E, I_2, I_3), and the system is separable since the equations of motion can be written as

$$p_\tau^2 = \frac{1}{2(\tau - a^2)} \left[G(\tau) - \frac{I_2}{\tau - a^2} - \frac{I_3}{\tau - c^2} \right], \quad \tau = \lambda, \nu, \quad (\text{C5})$$

and

$$p_\varphi = L_z = \sqrt{2I_2}. \quad (C6)$$

To represent the Galaxy we may choose a superposition of two Stäckel potentials: a disk plus a halo component

$$V = kV_{\text{disk}} + (1 - k)V_{\text{halo}}, \quad (C7)$$

where k represents the mass fraction of the disk with respect to the total mass of the Galaxy. Since the coordinates used for the halo and the disk have to be the same, this introduces a relation between the characteristic parameters (a_d, c_d) and (a_h, c_h) of the Stäckel potentials. It can be shown that the potential

$$V(\lambda, \nu, q) = -GM \left[\frac{k}{\sqrt{\lambda} + \sqrt{\nu}} + \frac{1 - k}{\sqrt{\lambda + q} + \sqrt{\nu + q}} \right] \quad (C8)$$

where q is related to the flattening of the halo component, provides a good description yielding a flat rotation curve with similar properties to that of our Galaxy (Batsleer & Dejonghe 1994). The function $G(\tau)$ in Eq.(C2) is

$$G(\tau) = GM \left[\frac{k}{\sqrt{\tau} + c} + \frac{1 - k}{\sqrt{\tau + q} + c} \right] \quad (C9)$$

For the characteristic parameters we choose $a_d = 2$, $c_d = 1$, $a_h/c_h = 1.01$ (giving a rather spherical halo), $k = 0.12$ and $M = 5 \times 10^{11} M_\odot$.

In order to obtain the evolution of the mean density of debris as a function of time in a Stäckel potential we use the results of Section 2.4 and of Appendices A and B. From Eqs.(A12) and (A13) the density is proportional to the determinant of the velocity submatrix. Since the Hamiltonian is separable in spheroidal coordinates, to obtain the density in cylindrical (or spherical) coordinates we need to multiply Eq.(B4) by the determinant of the matrix that performs the transformation between the two sets of coordinates. Thus

$$\det \sigma_\omega(v) = \left[\det \mathbf{T}_{p \rightarrow v} \det \mathbf{T}_{p_\tau \rightarrow p_{cyl}} \det \mathbf{W}_{qJ}^{-1} \det \boldsymbol{\Omega}'_t \right]^2 \det \sigma_\phi^0, \quad (C10)$$

where

$$\det \mathbf{T}_{p \rightarrow v} = R, \quad \det \mathbf{T}_{p_\tau \rightarrow p_{cyl}} \det \mathbf{W}_{qJ}^{-1} = \frac{(\nu - \lambda)v_\lambda v_\nu}{\Omega_\nu \frac{\partial I_3}{\partial J_\lambda} - \Omega_\lambda \frac{\partial I_3}{\partial J_\nu}}. \quad (C11)$$

The mean density at time t at the point $\bar{\mathbf{x}}$ on the mean orbit of the system becomes

$$\rho(\bar{\mathbf{x}}, t) = \frac{(2\pi)^{3/2} f_0}{\sqrt{\det \sigma_\phi^0}} \frac{1}{|v_\lambda v_\nu| |\lambda - \nu| R} \frac{\left| \Omega_\nu \frac{\partial I_3}{\partial J_\lambda} - \Omega_\lambda \frac{\partial I_3}{\partial J_\nu} \right|}{|\det \boldsymbol{\Omega}'|} \frac{1}{t^3}. \quad (C12)$$

This expression is valid for a satellite described initially by a Gaussian distribution. The variance matrix at $t = t^0$ may be

1. diagonal in action-angle variables:

$$\det \sigma_\phi^0 = 1/(\sigma_{\phi_1} \sigma_{\phi_2} \sigma_{\phi_3})^2,$$

2. diagonal in configuration-velocity space. If the satellite is initially close to a turning point of the orbit then

$$\det \sigma_\phi^0 = \left[\frac{h_\lambda h_\nu}{\left| \Omega_\nu \frac{\partial I_3}{\partial J_\lambda} - \Omega_\lambda \frac{\partial I_3}{\partial J_\nu} \right|} \frac{\lambda - \nu}{P^3 Q^3} \frac{1}{\sigma_\phi \sigma_\nu^2} \right]^2, \quad (\text{C13})$$

where all functions are evaluated at $(\bar{\mathbf{x}}^0, \bar{\mathbf{v}}^0)$, and

$$h_\tau = 2p_\tau \frac{\partial p_\tau}{\partial \tau}, \quad \tau = \lambda, \nu.$$

*Uno busca lleno de esperanzas,
el camino que los sueños,
prometieron a sus ansias,
Sabe que la lucha es cruel y es mucha,
pero lucha y se desangra,
por la fe que lo empecina.*

E. S. Discépolo: "Uno"

Chapter 3

Debris streams in the Solar neighbourhood as relicts from the formation of the Milky Way*

ABSTRACT

It is now generally believed that galaxies were built up through gravitational amplification of primordial fluctuations and the subsequent merging of smaller precursor structures. The stars of the structures that assembled to form the Milky Way should now make up much or all of its bulge and halo, in which case one hopes to find "fossil" evidence for those precursor structures in the present distribution of halo stars. Confirmation that this process is continuing came with the discovery of the Sagittarius dwarf galaxy (Ibata, Gilmore & Irwin 1994), which is being disrupted by the Milky Way, but direct evidence that this process provided the bulk of the Milky Way's population of old stars has so far been lacking. Here we show that about ten per cent of the metal-poor stars in the halo of the Milky Way, outside the radius of the Sun's orbit, come from a single coherent structure that was disrupted during or soon after the Galaxy's formation. This object had a highly inclined orbit about the Milky Way at a maximum distance of ~ 16 kpc, and it probably resembled the Fornax and Sagittarius dwarf spheroidal galaxies.

*Based on: Amina Helmi, Simon D.M. White, P. Tim de Zeeuw and HongSheng Zhao, Nature, 402, 53 (1999)

Early studies treated the formation of the Milky Way's spheroid as an isolated collapse, argued to have been either rapid and "monolithic" (Eggen, Lynden-Bell & Sandage 1962), or inhomogeneous and slow compared to the motions of typical halo stars (Searle & Zinn 1978). A second dichotomy distinguished "dissipationless" galaxy formation, in which stars formed before collapse (Gott 1977), from "dissipative" models in which the collapsing material was mainly gaseous (Larson 1975). Aspects of these dichotomies remain as significant issues in current theories (White & Frenk 1991), but they are typically rephrased as questions about whether small units equilibrate and form stars before they are incorporated into larger systems, and about whether they are completely disrupted after such incorporation. Stars from Galactic precursors should be visible today either as "satellite" galaxies, if disruption was inefficient, or as part of the stellar halo and bulge, if it was complete.

Recent work examined the present-day distribution expected for the debris of a precursor which was disrupted during or soon after the formation of the Milky Way (Helmi & White 1999). Objects which could contribute substantially to the stellar halo near the Sun must have had relatively short orbital periods. Ten Gyr after disruption their stars should be spread evenly through a large volume, showing none of the trails characteristic of currently disrupting systems like Sagittarius (Johnston, Hernquist & Bolte 1996). In any relatively small region, such as the Solar neighbourhood, their stars should be concentrated into a number of coherent "streams" in velocity space, each showing an internal velocity dispersion of only a few km s^{-1} . Objects initially similar to the Fornax or Sagittarius dwarf galaxies should give rise to a few streams in the vicinity of the Sun.

The high quality proper motions provided by the HIPPARCOS satellite allow us to construct accurate three-dimensional velocity distributions for almost complete samples of nearby halo stars. Drawing on two recent observational studies (Beers & Sommer-Larsen 1995; Chiba & Yoshii 1998), we define a sample containing 97 metal deficient ($[\text{Fe}/\text{H}] \leq -1.6$ dex) red giants and RR Lyrae within 1 kpc of the Sun and with the following properties:

1. HIPPARCOS proper motions are available for 88 of them (ESA 1997); for the remaining stars there are ground-based measurements (Röser & Bastian 1988); in all cases accuracies of a few mas yr^{-1} are achieved.
2. Radial velocities have been measured from the ground, with accuracies of the order of 10 km s^{-1} . Metal abundances have been determined either spectroscopically or from suitable photometric calibrations (Norris, Bessell & Pickles 1985; Beers et al. 1990; Anthony-Twarog & Twarog 1994).
3. Calibrations of absolute magnitude M_V against $[\text{Fe}/\text{H}]$ for red giants (Norris, Bessell & Pickles 1985; Beers et al. 1990; Anthony-Twarog & Twarog 1994) and RR Lyrae (Layden 1994), allow photometric parallaxes to be derived to an accuracy of roughly 20%. These are more accurate than the corresponding HIPPARCOS trigonometric parallaxes, but still remain the largest source of uncertainty in the derived tangential velocities.
4. We estimate the completeness to be of the order of $\sim 92\%$, based on the fact that there are eight known giants which satisfy our selection criteria but do not have measured proper motions.

We look for substructure in our set of halo stars by studying the entropy S of the sample, defined as:

$$S = - \sum_i \frac{N_i}{N} \log \frac{N_i A_P}{N}, \quad (3.1)$$

where the sum is over the A_P elements of the partition, the i -th element contains N_i stars, and N is the total number of stars. In the presence of substructure the measured entropy will be smaller than that of a smooth distribution, and will depend on the details of the partition; some

partitions will enhance the signal, whereas others will smear it out. If there is no substructure then all partitions will yield similar S values, and no significant minimum value will be found.

We implement this entropy test initially by partitioning velocity space into cubic cells 70 km s⁻¹ on a side. This choice is a compromise. It leaves a large number of cells in the high velocity range empty, but in the regions containing most of the stars, there are at least a few stars per cell.

It is necessary to quantify the significance of any observed low entropy value relative to the distribution expected in the absence of substructure. Here we do this by generating Monte Carlo realizations which test whether the kinematics of the sample are consistent with a multivariate Gaussian distribution (Sommer-Larsen et al. 1997). Our synthetic data sets have the same number of stars and the same spatial distribution as the observed sample. The characteristic parameters of the multivariate Gaussian used to describe the kinematics are obtained by fitting to the observed mean values and variances after appropriate convolution with the observational errors. We then generate 10000 “observed” samples as follows. A velocity is drawn from the underlying multivariate Gaussian; it is transformed to a proper motion and radial velocity (assuming the observed parallax and position on the sky); observational “errors” are added to the parallax, the radial velocity and the proper motion; these “observed” quantities are then transformed back to an “observed” velocity. Finally we calculate entropies for these 10000 Monte Carlo samples on the same partition as the real data; only for 5.6% do we find values of S smaller than observed. We have repeated this test for many partitions, finding a large number with probabilities as low or lower than this. In particular, for a partition with a 250 km s⁻¹ bin in v_ϕ , and 25 km s⁻¹ bins in v_R and v_z , (v_R , v_ϕ and v_z are the velocity components in the radial, azimuthal and z -directions respectively), only 0.06% of Monte Carlo simulations have S smaller than observed. In general cubic cells yield lower significance levels, suggesting that the detected structure may be elongated along v_ϕ . We conclude that a multivariate Gaussian does not properly describe the distribution of halo star velocities in the Solar neighbourhood.

At this point the main problem is to identify the structure which makes the observed data incompatible with a smooth velocity distribution. A comparison of the three principal projections of the observed distribution to similar plots for our Monte Carlo samples reveals no obvious differences, as shown in Figure 3.1. To better identify streams we turn to the space of adiabatic invariants. Here clumping should be stronger, as all stars originating from the same progenitor have very similar integrals of motion, resulting in a superposition of the corresponding streams. We focus on the plane defined by two components of the angular momentum: J_z and $J_\perp = \sqrt{J_x^2 + J_y^2}$, although $J_\perp$ is not fully conserved in an axisymmetric potential. In Fig. 3.2(a) we plot J_z versus $J_\perp$ for our sample. For comparison, Fig. 3.2(b) gives a similar plot for one of our Monte Carlo samples. For $J_\perp \leq 1000$ kpc km s⁻¹ and $|J_z| \leq 1000$ kpc km s⁻¹, the observed distribution appears fairly smooth. In this region we find stars with relatively low angular momentum and at all inclinations. In contrast, for $J_\perp \geq 1000$ kpc km s⁻¹, there are a few stars moving on retrograde low inclination orbits, an absence of stars on polar orbits, and an apparent “clump” on a prograde high inclination orbit.

To determine the significance of this clumping, and to confirm it as the source of the signal detected by our entropy test, we compare the observed star counts in this plane to those for our Monte Carlo data sets. We count how many stars fall in each cell of a given partition of this angular momentum plane and compare it to the expected number in the Monte Carlo simulations. We say that the i -th cell has a significant overdensity if there is less than 1% probability of obtaining a count as large as the observed N^i from a Poisson distribution with mean

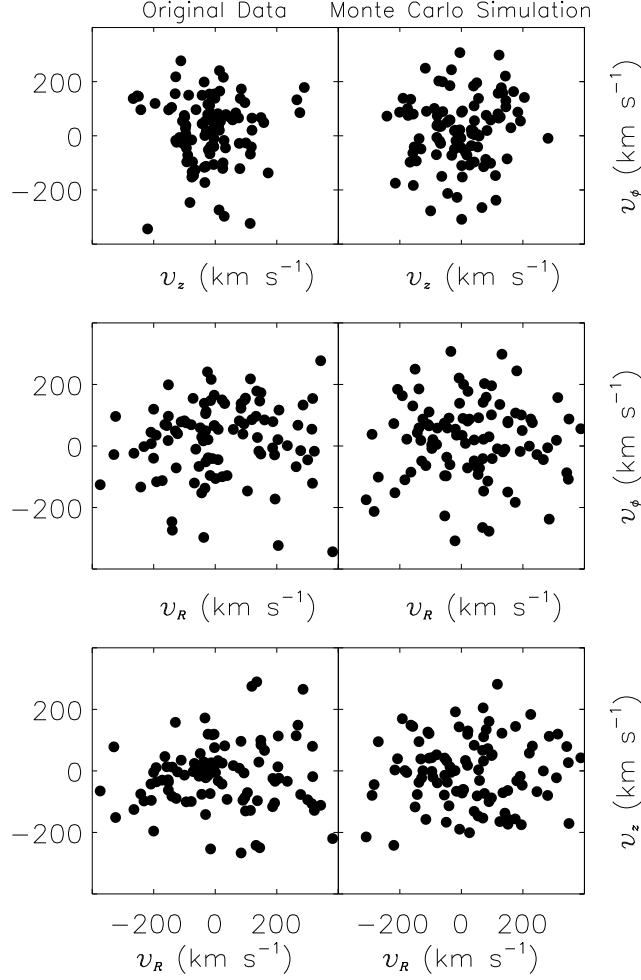


FIGURE 3.1— Scatter plots of the different velocity components. The left column corresponds to the stars in our sample, the right column represents one realization of the Monte Carlo set of simulations. Velocities are referred to the Galactic Centre; we adopt 8 kpc as the distance to the Galactic Centre and 220 km s^{-1} towards Galactic longitude $l = 0$ and Galactic latitude $b = 0$ as the velocity of the Local Standard of Rest.

$\langle N^i \rangle = N_{\text{sim}}^{-1} \sum_{j=1}^{N_{\text{sim}}} N_j^i$, where N_j^i is the count in the i -th cell in the j -th simulation, and N_{sim} is the

number of simulations. We repeat this test for a series of regular partitions of $p \times q$ elements, with p, q ranging from 3 to 20, thus allowing a clear identification of the deviant regions. We find a very significant deviation in most partitions for cells with $J_{\perp} \sim 2000 \text{ kpc km s}^{-1}$ and $500 < J_z < 1500 \text{ kpc km s}^{-1}$; the probabilities of the observed occupation numbers range from 0.03% to 0.98%, depending on the partition, and in some partitions more than one cell is significantly overdense.

Given this apparently significant evidence for substructure in the local halo, we study what happens if we relax our metallicity and distance selection criteria. We proceed by including in our sample all red giants and RR Lyrae stars studied by Chiba and Yoshii (1998) with metallicities less than -1 dex and distances to the Sun of less than 2.5 kpc. This new sample contains 275 giant stars and adds 5 new stars to the most significant clump in our complete sample. Of the 13 members of the clump, 9 have $[\text{Fe}/\text{H}] \leq -1.6$, whereas the remaining 4 have

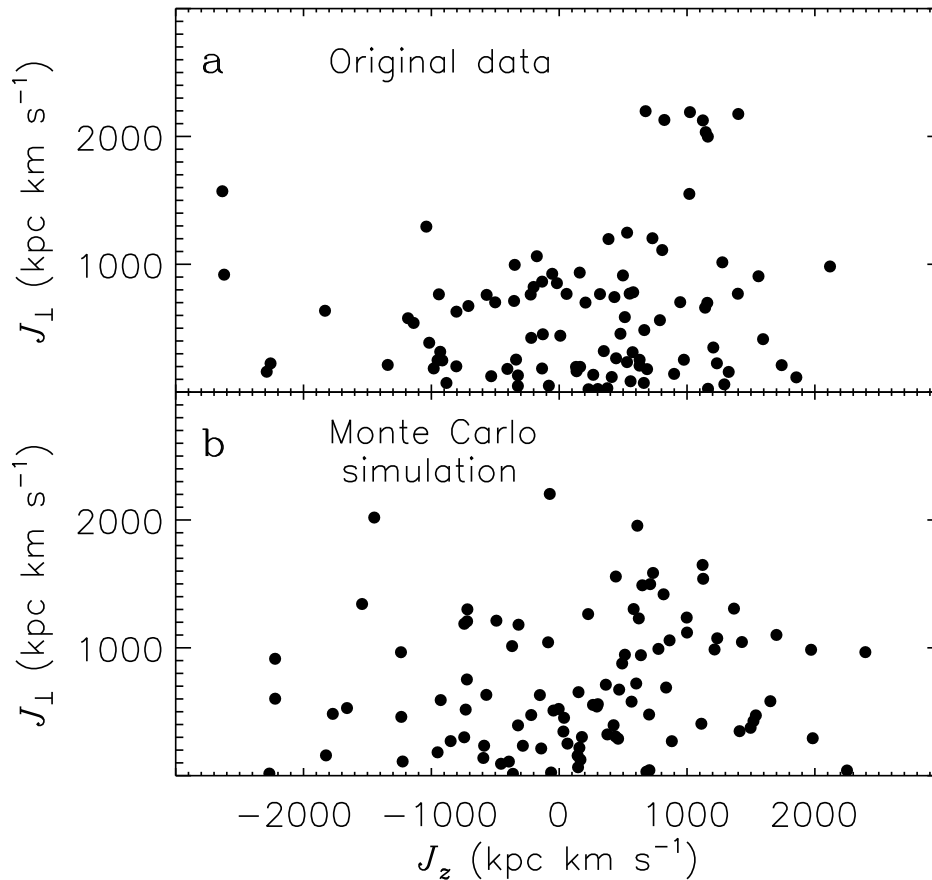


FIGURE 3.2— The distribution of nearby halo stars in the plane of angular momentum components, J_z vs. $J_\perp = \sqrt{J_x^2 + J_y^2}$, for our near complete sample (a) and for one Monte Carlo realization (b).

$\langle [\text{Fe}/\text{H}] \rangle \sim -1.53 \pm 0.12$, indicating that they are also very metal-poor. These stars are distributed all over the sky with no obvious spatial structure.

In Figure 3.3 we highlight the kinematic structure of the clump in the extended sample. The clump stars are distributed in two streams moving in opposite directions perpendicular to the Galactic Plane, with one possible outlier. This star has $v_R = 285 \pm 21 \text{ km s}^{-1}$, and we exclude it because its energy is too large to be consistent with the energies of the other members of the clump. The velocity dispersions for the stream with negative v_z (9 stars) are $\sigma_\phi = 30 \pm 17$, $\sigma_R = 105 \pm 16$, $\sigma_z = 24 \pm 28$ in km s^{-1} , whereas for the stream with positive v_z (3 stars) these are $\sigma_\phi = 49 \pm 22$, $\sigma_R = 13 \pm 33$, $\sigma_z = 31 \pm 28$ in km s^{-1} . An elongation in the v_R -direction is expected for streams close to their orbital pericentre (compare with other plots of simulated streams, cf. Helmi & White 1999).

The orbit of the progenitor system is constrained by the observed positions and velocities of the stars. The orbital radii at apocentre and pericentre are $r_{\text{apo}} \sim 16 \text{ kpc}$ and $r_{\text{peri}} \sim 7 \text{ kpc}$, the maximum height above the plane is $z_{\text{max}} \sim 13 \text{ kpc}$, and the radial period is $P \sim 0.4 \text{ Gyr}$, for a Galactic potential including a disk, bulge and dark halo (as in Johnston, Hernquist & Bolte 1996). We run numerical simulations of satellite disruption in this potential to estimate the initial properties of the progenitor. After 10 Gyr of evolution, we find that the observed properties

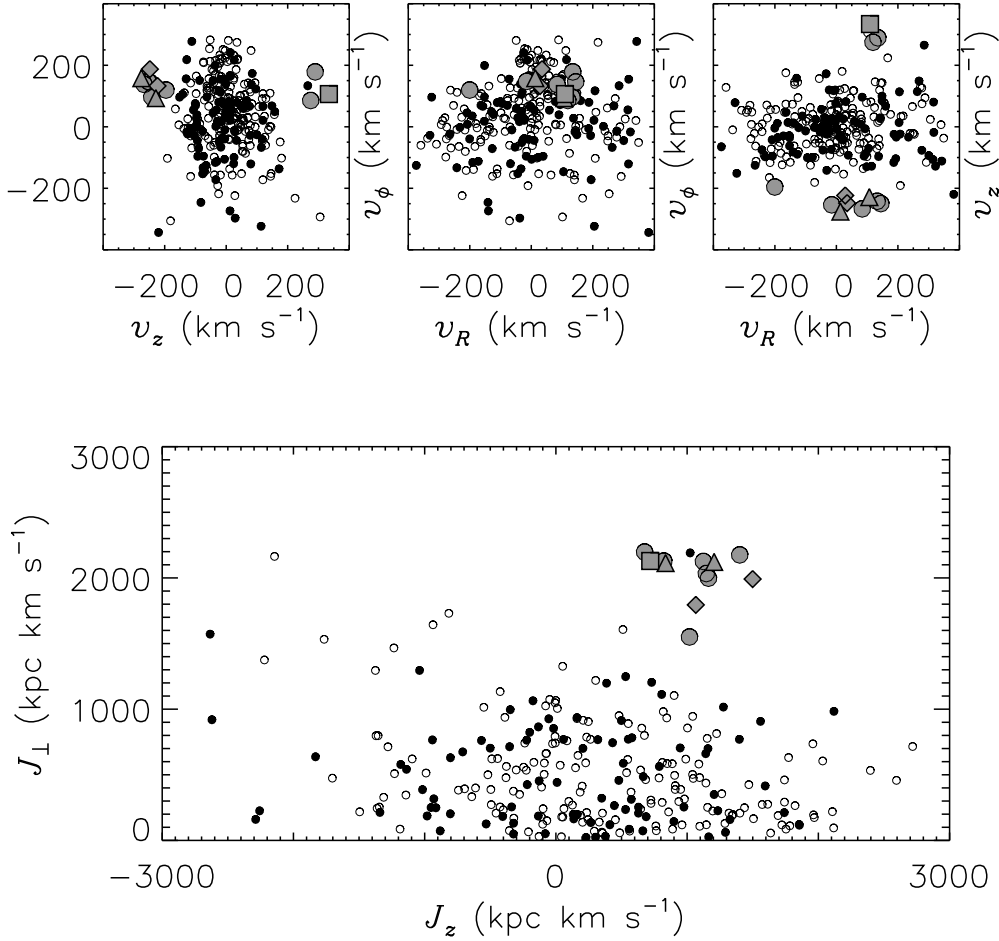


FIGURE 3.3— The distribution of nearby halo stars in velocity space and in the $J_z - J_\perp$ plane. Data are shown for our original sample (filled circles) and for the extended sample of more metal-rich and more distant giants (Chiba & Yoshii 1998) (open circles). Candidates for our detected substructure are highlighted in grey: triangles indicate more metal-rich giant stars at distances > 1 kpc, diamonds more metal-rich giants at ≤ 1 kpc, squares metal-poor giants at > 1 kpc, and circles metal-poor giants at ≤ 1 kpc.

of the streams detected can be matched by stellar systems similar to dwarf spheroidals with initial velocity dispersions σ in the range $12 - 18 \text{ km s}^{-1}$ and core radii R of $0.5 - 0.65$ kpc. An example is shown in Figure 3.4. We also analysed whether the inclusion of an extended dark halo around the initial object would affect the structures observed and found very little effect. We derive the initial luminosity L from $L = L^*/(f^{\text{giant}} \times C^* \times f^{\text{sim}})$, where $L^* = 350 L_\odot$ is the total luminosity of the giants in the clump in our near-complete sample, $f^{\text{giant}} \sim 0.13$ is the ratio of the luminosity in giants with M_V and $(B - V)$ in the range observed to the total luminosity of the system for an old metal-poor stellar population (Bergbusch & Vandenberg 1992), $C^* = 0.92$ is our estimated completeness, and $f^{\text{sim}} \sim 1.9 \times 10^{-4}$ is the fraction of the initial satellite contained in a sphere of 1 kpc radius around the Sun as determined from our simulations. This gives $L \sim 1.5 \times 10^7 L_\odot$, from which we can derive, using our previous estimates of the initial velocity dispersion and core radii, an average initial core mass-to-light ratio $M/L \sim 3 - 10 Y_\odot$, where $Y_\odot$ is the mass-to-light ratio of the Sun. A progenitor system with these characteristics would be similar to Fornax. Moreover, the mean metal abundance of the stars is consistent with

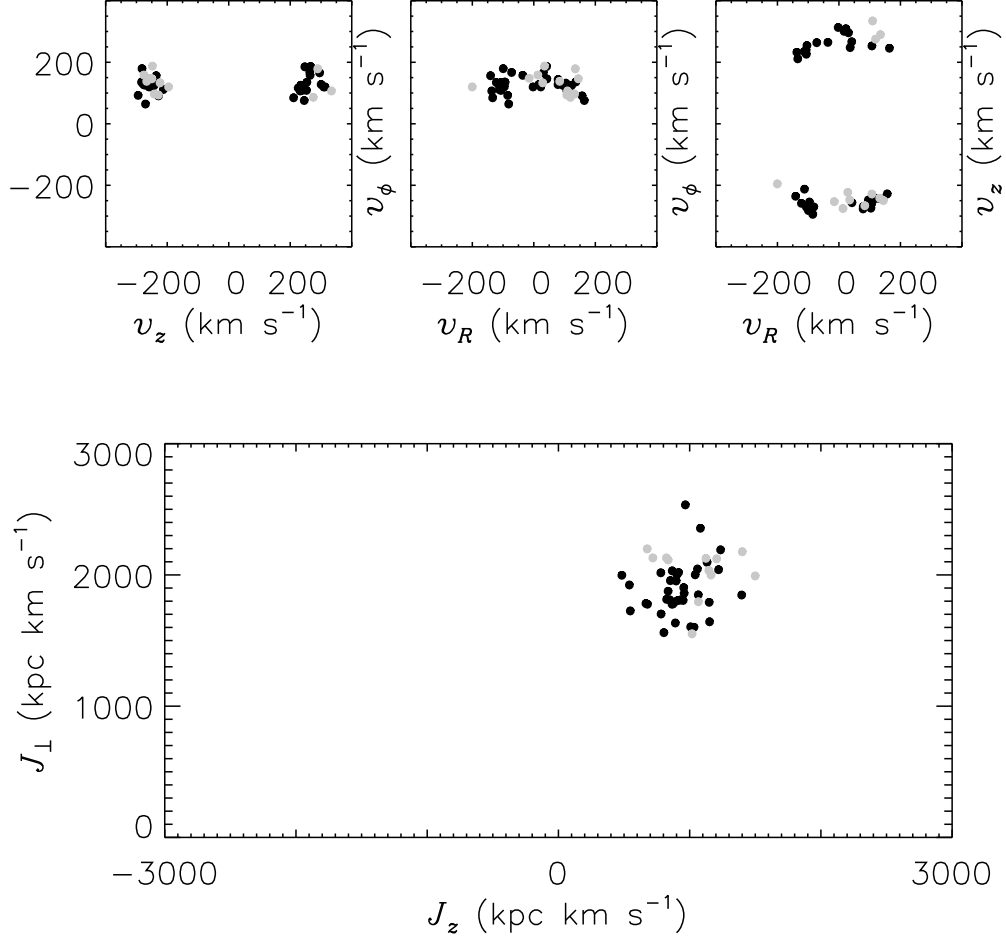


FIGURE 3.4— The distribution of particles in the vicinity of the Sun from one of our simulations of a disrupted satellite, 12 Gyr after infall, in velocity space and in the $J_z - J_\perp$ plane. In this case, the initial core radius of the progenitor system is 0.5 kpc and its initial velocity dispersion is 16.5 km s⁻¹. The candidate stars for our detected substructure are highlighted in grey for comparison.

the derived luminosity, if the progenitor follows the known metallicity-luminosity relation of dwarf satellites in the Local Group (Mateo 1998).

The precursor object was apparently on an eccentric orbit with relatively large apocentre. Given that it contributes 7/97 of the local halo population, our simulations suggest that it should account for 12% of all metal-poor halo stars outside the Solar circle. Figure 3.3 shows that there are few other halo stars on high angular momentum polar orbits in the Solar neighbourhood, just the opposite of the observed kinematics of satellites of the Milky Way (Lynden-Bell & Lynden-Bell 1995). The absence of satellite galaxies on eccentric non-polar orbits argues that some dynamical process preferentially destroys such systems; their stars should then end up populating the stellar halo. As we have shown, the halo does indeed contain fossil streams with properties consistent with such disruption.

Acknowledgments

A.H. wishes to thank the Max Planck Institut für Astrophysik for hospitality during her visits. We made use of the Simbad database (maintained by Centre de Donnée astronomiques de Strasbourg) and of the HIPPARCOS online facility at the European Space Research and Technology Centre (ESTEC) of the European Space Agency (ESA).

References

- Anthony-Twarog B.J., Twarog B.A., 1994, *AJ*, 107, 1577
Beers T.C., Preston G.W., Shectman S.A., Kage J.A., 1990, *AJ*, 100, 849
Beers T.C., Sommer-Larsen J., 1995, *ApJS*, 96, 175
Bergbusch P.A., Vandenberg D.A., 1992, *ApJS*, 81, 163
Chiba M., Yoshii Y., 1998, *AJ*, 115, 168
Eggen O.J., Lynden-Bell D., Sandage A.R., 1962, *ApJ*, 136, 748
ESA: “The Hipparcos and Tycho Catalogues”, 1997, SP-1200, ESA Publications Division, ESTEC (Noordwijk)
Gott J.R. III, 1977, *ARA&A*, 15, 235
Helmi A., White, S.D.M., 1999, *MNRAS*, 307, 495 (Chapter 2)
Ibata R., Gilmore G., Irwin M., 1994, *Nature*, 370, 194
Johnston K.V., Hernquist L., Bolte M., 1996, *ApJ*, 465, 278
Larson R.B., 1975, *MNRAS*, 173, 671
Layden A.C., 1994, *AJ*, 108, 1016
Lynden-Bell D., Lynden-Bell R.M., 1995, *MNRAS*, 275, 429
Mateo M., 1998, *ARA&A*, 36, 435
Norris J., Bessell M.S., Pickles A.J., 1985, *ApJS*, 58, 463
Röser S., Bastian, U., 1988, *A&AS*, 74, 449
Searle L., Zinn R., 1978, *ApJ*, 225, 357
Sommer-Larsen J., Beers T.C., Flynn C., Wilhelm R., Christensen P.R., 1997, *ApJ*, 481, 775
White S.D.M., Frenk C.S., 1991, *ApJ*, 379, 52

Chapter 4

Mapping the substructure in the Galactic halo with the next generation of astrometric satellites*

ABSTRACT

We run numerical simulations of the disruption of satellite galaxies in a Galactic potential to produce a stellar halo, in order to investigate what the next generation of astrometric satellites will reveal by observing the halo of the Milky Way. We generate artificial DIVA, FAME and GAIA halo catalogues, in which we look for the signatures left by the accreted satellites. We develop a method based on the standard Friends-of-Friends algorithm applied to the space of integrals of motion. We find this simple method can recover about 50% of the different accretion events, when the observational uncertainties expected for GAIA are taken into account, even when the exact form of the Galactic potential is unknown. The recovery rate for DIVA and FAME is much smaller, but these missions, like GAIA, should be able to test the hierarchical formation paradigm on our Galaxy by measuring the amount of halo substructure in the form of nearby kinematically cold streams with for example, a two-point correlation function in velocity space.

*Based on: Amina Helmi and P. Tim de Zeeuw, submitted to MNRAS

4.1 Introduction

Hierarchical theories of structure formation in the Universe propose that galaxies are the result of mergers and accretion of smaller building blocks (White & Rees 1978). Detailed studies of the properties of a galaxy built in this way have shown that such events leave fossil signatures in the present day components, which for a galaxy like our own would be clearly detectable with future astrometric missions (Helmi & White 1999). In particular the stellar halo would be the natural place to look for such substructures, since a spheroidal component is formed by the trails of stars left by disrupted satellite galaxies. Moreover, recent observations have shown that indeed considerable structure is still present in Milky Way's halo, indicating that accretion events have had some role in its formation history (e.g. Ibata, Gilmore & Irwin 1994; Majewski, Munn & Hawley 1996; Helmi et al. 1999).

In the next ten years, several satellite missions will be devoted to measure with very high accuracy the motions of thousands to many millions of stars in our Galaxy. NASA's Space Interferometry Mission (SIM) is a targeted mission which will obtain parallaxes and proper motions for about 10000 stars. With somewhat different goals, and more similar to the HIPPARCOS satellite, the Full-sky Astrometric Mapping Explorer (FAME, Horner et al. 1999) promises to measure positions and parallaxes for stars brighter than $V \sim 5$ to better than $50 \mu\text{as}$ and proper motions to $50 \mu\text{as yr}^{-1}$. At $V \sim 15$ these accuracies will be degraded by an order of magnitude. The resulting astrometric database will have 4×10^7 stars, and may be combined with the radial velocities from the Sloan Digital Sky Survey or from other ground based catalogues to obtain full phase-space information. Less ambitious but still an improvement over HIPPARCOS is the German DIVA mission (Röser 1998). If launched it will observe of the order of 3.5×10^7 stars, at four times the precision of HIPPARCOS ($\sigma_\pi = 0.25 \text{ mas}$ and $\sigma_\mu = 0.4 \text{ mas yr}^{-1}$ at $V = 10$), thereby completing the knowledge of nearby stars. Like FAME, DIVA will not measure radial velocities. On the other hand, the proposed ESA astrometric satellite GAIA (Gilmore et al. 1998) will provide very precise astrometry ($<10 \mu\text{as}$ in parallax and $<10 \mu\text{as yr}^{-1}$ in proper motion at $V \sim 15$, increasing to 0.2 mas yr^{-1} at $V \sim 20$) and multicolour photometry, for all 1.3 billion objects to $V \sim 20$, and radial velocities with accuracies of a few km s^{-1} for most stars brighter than $V \sim 17$, so that full and homogeneous six-dimensional phase-space information will be available. These satellite missions will thereby provide a very large and statistically reliable sample of stars, from which the fundamental questions concerning the origin and evolution of the Galaxy may finally be answered.

In this paper we shall focus on what GAIA will tell us about the history and formation of the stellar halo of the Milky Way. We will also discuss the impact of DIVA and FAME, and leave aside SIM as this mission will not provide a survey but a hand-picked catalogue of stars. Even though we focus on the stellar halo, the method that we shall propose for finding substructures in phase-space may also be extended to find, for example, disk moving groups (e.g. de Zeeuw et al. 1999; Chereul, Cr     & Bienaym   1999).

There are several methods for detecting moving groups. The Great Circle Counts method (G3C) proposed by Johnston, Hernquist & Bolte (1996) uses the position on the sky, and employs the fact that satellite galaxies in orbits that probe only the outer (spherical) halo conserve the orientation of their plane of motion, thereby leaving their debris along great circles on the sky, if observed from the Galactic centre. The methods used in the Solar neighbourhood for detection of disk moving groups and open clusters use also proper motions (and sometimes parallax), and assume that all the stars belonging to the same system have the same velocity vector (e.g. Hoogerwerf & Aguilar 1998; de Bruijne 1998). Lynden-Bell & Lynden-Bell's method (1995) needs the position on the sky and the radial velocity, and has been used, for example, to

link globular clusters which lie in the same plane to some of the (disrupted) dwarf companions of our Galaxy (see also Lynden-Bell, 1999). The applicability of the above mentioned methods is questionable in the inner parts of the halo. In this regime, the Galactic potential is significantly flattened so that the debris does not remain on a fixed plane, the situation where G3C works. As noted by Helmi & White (1999) no spatial correlations should be expected for satellites disrupted several Gyr ago. On the other hand, even though the velocity dispersions in a stellar stream do decrease with time, and therefore, very strong correlations are to be expected, in the inner halo strong phase-mixing takes place. For example in the Solar neighbourhood several hundred (mainly) cold streams originating in disrupted satellites may be present, but it may in practice be difficult to resolve each one of such moving groups completely. Clearly, before exploring the full capabilities of the next generation of astrometric satellite missions, we first need to identify where the clustering that is characteristic of a satellite manifests itself in the debris that we observe after many galactic orbits. As shown in Helmi et al. (1999) a method based on the lumpiness in integrals of motion space seems to be a promising tool for unveiling the merger history of our Galaxy.

4.2 Building up a stellar halo

Our main goal is to test whether with the next generation of astrometric satellites, we would be able to find the signatures left by merger events in the Galactic stellar halo. We will assume that the whole stellar halo is the result of the superposition of several disrupted satellite galaxies which fell onto the Milky Way about 10 Gyr ago. We shall here discuss the initial conditions and the numerical methods used to generate this version of the stellar halo.

4.2.1 Initial conditions for the satellites

Orbital properties

The stellar halo has a density profile (Kinman, Suntzeff & Kraft 1994)

$$\rho_*(r) = \rho_0 \left(\frac{r}{r_0} \right)^{-3.5}, \quad (4.1)$$

a total luminosity of about $10^9 L_\odot$, and a half light radius which probably lies around 3 kpc from the Galactic centre. For $r = r_0 = 8$ kpc (the distance to the Galactic Centre from the Sun), ρ_0 corresponds to the local stellar halo density, for which we take $\rho_0 = 1.5 \times 10^4 M_\odot \text{ kpc}^{-3}$ (Fuchs & Jahreiß 1998).

The initial orbital conditions of our satellites should be drawn from the Galactic halo distribution function (DF), which we assume to be a function of energy E and angular momentum L : $f(E, L)$. For simplicity here we shall assume that the stellar halo is a power-law tracer population embedded in a singular isothermal sphere, representing the dark matter halo of the Milky Way. Following van den Bosch et al. (1999), we assume that

$$f(E, L) = g(E)h_\alpha(\eta), \quad \text{where} \quad \eta = L/L_c(E),$$

and $L_c(E)$ is the angular momentum of a circular orbit with energy E : $L_c(E) = r_c(E)V_c$, where $r_c(E) = e^{-1/2} \exp [E/V_c^2]$. The function $h_\alpha(\eta)$ is known as the circularity, and determines the degree of anisotropy of the DF. We choose a simple parametrization of $h_\alpha(\eta)$ (Gerhard 1991):

$$h_\alpha(\eta) = \begin{cases} \tanh(\frac{\eta}{\alpha}) / \tanh(\frac{1}{\alpha}) & \alpha > 0 \\ 1 & \alpha = 0 \\ \tanh(\frac{1-\eta}{\alpha}) / \tanh(\frac{1}{\alpha}) & \alpha < 0 \end{cases} \quad (4.2)$$

so that for $\alpha = 0$, the DF is isotropic, for $\alpha < 0$ it is radially anisotropic and for $\alpha > 0$ is tangentially anisotropic. We shall take $\alpha = -0.5$, since the halo appears to be radially anisotropic.

For a singular isothermal sphere

$$\rho(r) = \frac{V_c^2}{4\pi G r^2}, \quad \phi(r) = V_c^2 \ln \frac{r}{r_s}. \quad (4.3)$$

The corresponding DF is

$$g(E) = \frac{e}{16\pi^2 G V_c \kappa} \exp \left[-\frac{2E}{V_c^2} \right], \quad (4.4)$$

(Gerhard 1991) where

$$\kappa = \int_0^\infty du e^{-u} \int_0^{\eta_{\max}} h_a(\eta) \frac{\eta d\eta}{\sqrt{\eta_{\max}^2 - \eta^2}}. \quad (4.5)$$

Here $\eta_{\max} = \sqrt{2e}\sqrt{u}e^{-u}$, with $u = (E - \phi)/V_c^2$. Since the density profile may be derived from the initial distribution function as

$$\rho(r) = \frac{4\pi}{r} \int_{\phi(r)}^\infty dE g(E) L_c(E) \int_0^{\eta_{\max}} h_a(\eta) \frac{\eta d\eta}{\sqrt{\eta_{\max}^2 - \eta^2}}, \quad (4.6)$$

the joint probability distribution of E and η at a given radius r_p is

$$P(E, \eta) = \frac{4\pi}{r_p \rho(r_p)} g(E) L_c(E) \frac{\eta h(\eta)}{\sqrt{\eta_{\max}^2 - \eta^2}},$$

(van der Marel, Sigurdsson & Hernquist 1997). Using Eq.(4.4), we find that the normalized cumulative probability distribution of E is

$$\hat{P}(< E) = 1 - \exp \left[-\frac{E - \phi}{V_c^2} \right]. \quad (4.7)$$

We may derive the initial positions of the satellites by assuming the profile given in Eq.(4.1), and using, from Eq.(4.7), the energy as $E = \phi(r_p) - V_c^2 \ln(1 - \mathcal{R})$, with $\mathcal{R}$ a uniform random variable. With the energy (or $\mathcal{R}$) we can compute $\eta_{\max} = \sqrt{2e}\sqrt{\gamma}(1 - \mathcal{R})^\gamma \sqrt{-\ln(1 - \mathcal{R})}$, for $\gamma = 1/2.5$ to mimic the stellar halo power law. Using $\eta_{\max}$ and the probability distribution for η we may derive the non-circularity of the orbits, and in this way fully determine the phase-space initial position of a satellite.

Internal properties of the satellites

For the satellites we assume they initially have King profiles, as do most of the satellites in the Local Group. Their present day total luminosity is fixed to be that of the stellar halo, and we assume the initial number of satellites to be 33. The luminosity of each satellite is drawn from a Gaussian distribution with mean $2.5 \times 10^7 L_\odot$ and dispersion $10^7 L_\odot$. We assume that the satellites follow the scaling relations (Burstein et al. 1997)

$$\begin{aligned} \log L &= 5.35 + 1.80 \log \sigma_v, \\ \log R &= -0.82 + 0.51 \log \sigma_v. \end{aligned}$$

These relations allow us to derive from the luminosity L , the core radius R , and the central velocity dispersion σ_v . Our King models have a concentration parameter $c = \log r_t / r_K \sim 0.72$, where r_t and r_K are the tidal and King radii respectively. The initial mass of the satellite is now also fully determined.

4.2.2 Galactic potentials

We will consider two different Galactic potentials. In both cases, our Galaxy has three components: a dark halo, a disk and a bulge, but we take different functional forms for the potential. In Model I, we take a dark logarithmic halo

$$\Phi_{\text{halo}} = v_{\text{halo}}^2 \ln(r^2 + d^2), \quad (4.8)$$

a Miyamoto–Nagai disk

$$\Phi_{\text{disk}} = -\frac{GM_{\text{disk}}}{\sqrt{R^2 + (a + \sqrt{z^2 + b^2})^2}}, \quad (4.9)$$

and a spherical Hernquist bulge

$$\Phi_{\text{bulge}} = -\frac{GM_{\text{bulge}}}{r + c}, \quad (4.10)$$

where $d = 12$ kpc and $v_{\text{halo}} = 131.5$ km s⁻¹; $M_{\text{disk}} = 10^{11} M_{\odot}$, $a = 6.5$ kpc and $b = 0.26$ kpc; $M_{\text{bulge}} = 3.4 \times 10^{10} M_{\odot}$ and $c = 0.7$ kpc. This choice of parameters gives a flat rotation curve with an asymptotic circular velocity of 186 km s⁻¹.

In Model II, we represent the disk density profile with a double exponential (Quinn & Goodman 1986)

$$\rho_D(R, z) = \frac{M_D}{4\pi R_D^2 z_o} e^{-R/R_D} e^{-\beta|z|},$$

where $R_D = 3.5$ kpc is the disk scale length, z_o is its scale height, $\beta = 1/z_o$ and $M_D = 5.5 \times 10^{10} M_{\odot}$ the total disk mass. The associated potential is

$$\Phi_D(R, z) = -\frac{GM_D}{R_D^3} \times \int_0^\infty \frac{dk k J_0(kR)}{k^2 + 1/R_D^2} \frac{\beta^2}{\beta^2 - k^2} \left\{ \frac{e^{-k|z|}}{k} - \frac{e^{-\beta|z|}}{\beta} \right\}. \quad (4.11)$$

For the halo we choose (Hernquist 1993)

$$\rho_h(r) = \frac{M_h}{2\pi^{3/2}} \frac{\alpha_q}{r_c} \frac{e^{-r^2/r_c^2}}{r^2 + \gamma_q^2},$$

where $M_h = 1.5 \times 10^{12} M_{\odot}$, $\alpha_q = \left(1 - \sqrt{\pi} q e^{q^2} (1 - \text{Erf}[q])\right)^{-1}$ with $q = \gamma_q/r_c$ and $r_c = 200$ kpc is the cutoff radius. The corresponding potential is

$$\Phi_h(r) = -\frac{GM_h(r)}{r} + \frac{GM_h}{\sqrt{\pi} r_c} \text{Ei} \left[-\left(\frac{r}{r_c}\right)^2 - q^2 \right], \quad (4.12)$$

where

$$M_h(r) = \frac{2M_h \alpha_q}{\sqrt{\pi}} \int_0^{r/r_c} dx \frac{x^2 e^{-x^2}}{x^2 + q^2} \quad (4.13)$$

and $\text{Ei}(x)$ is the exponential integral (e.g. Gradshteyn & Ryzhik 1965). For the bulge we use Eq.(4.10) but we take $M_b = 1.1 \times 10^{10} M_{\odot}$ and $c = 0.525$ kpc (following Velázquez & White 1995). Figure 4.1 shows the circular velocity curves produced by each of the two potentials.

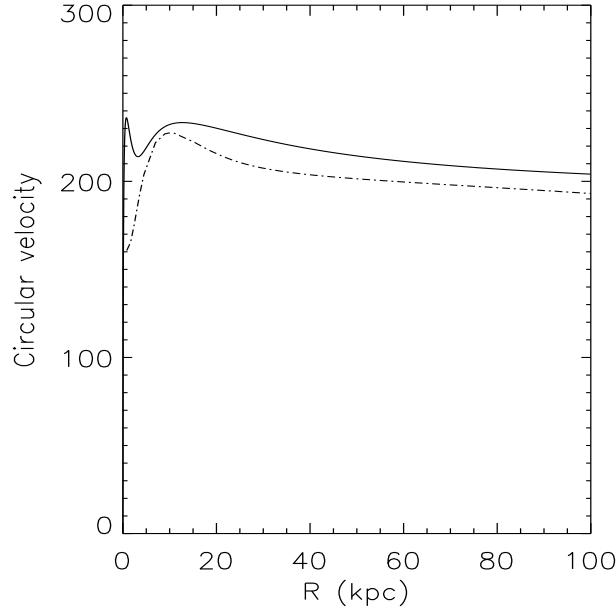


FIGURE 4.1— The circular velocity profile as a function of distance from the Galactic centre. The solid curve represents the potential used in the simulations. The dashed curve corresponds to the alternative potential.

4.2.3 Numerical methods

In our numerical simulations, we use the potential of Model I for our Galaxy. We represent the satellite galaxy by a collection of 10^5 particles and model their self-gravity by a multipole expansion of the internal potential to fourth order (White 1983; Zaritsky & White 1988). This type of code has the advantage that a large number of particles can be followed in a relatively small amount of computer time. In this quadrupole expansion, higher than monopole terms are softened more strongly. We choose $\epsilon_1 \sim 0.2 - 0.25R$ for the monopole term (R is the core radius of the system) and $\epsilon_2 = 2\epsilon_1$ for dipole and higher terms and for the centre of expansion. The centre of expansion is a particle which, in practice, follows the density maximum of the satellite closely at all times.

After letting our satellite relax in isolation, we integrate each simulation for ~ 12 Gyr. In Figure 4.2 we show the final particle counts in radial bins $N(r) = r^2\rho(r)$ as a function of distance from the Galactic centre resulting from the superposition of all our experiments. For guidance, we also plot the expected $r^{-1.5}$, arbitrarily shifted. We see that within the range of 3 to 30 kpc, our simulations follow relatively well the profile. Outside this range we see a sharp drop, due to the fact that we are (intentionally) not populating the outer halo. Since the properties of the inner stellar halo are not so well constrained, we do not worry about the fact that we find a shallower slope in the inner few kiloparsecs (this is also the result of our initial conditions). More important is the fact that our simulations can reproduce very well the regime where the astrometric missions promise to give accurate six dimensional phase-space information.

4.2.4 Generating catalogues of halo stars

To generate an artificial catalogue for the Galaxy we assume that each particle in our simulations represents a giant star of absolute magnitude $M_V = 1$. The total number of particles in our simulations corresponds well to the expected number of giant stars in the Galactic halo (based on the luminosity function, derived for the age and metallicity characteristic of halo stars). We

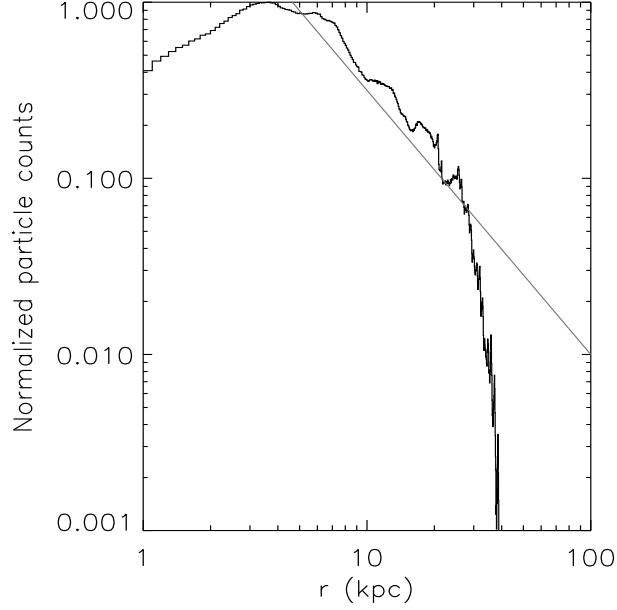


FIGURE 4.2— Number counts profile $N(r) = r^2 \rho(r)$ for the simulated stellar halo resulting from the superposition of 33 disrupted satellite galaxies, after 12 Gyr of evolution. The straight line represents the expected $r^{-1.5}$ law, arbitrarily shifted.

prefer to take only giant stars at this stage because they are bright enough to be easily observable from the Sun. We need to determine a limiting magnitude V_{lim} for our “artificial” catalogue, which we define so that all giant stars brighter than V_{lim} have accurate full 6-dimensional phase-space information. In the case of GAIA, we take $V_{\text{lim}} = 15$, a limit set by the accuracy in the radial velocity. For FAME, $V_{\text{lim}} = 12.5$ as all stars brighter than this magnitude will have relative parallax errors σ_π/π smaller than (or of the order of) 25%. For DIVA, we take $V_{\text{lim}} = 11$ for which $\sigma_\pi/\pi \sim 0.3$. Our GAIA, FAME and DIVA catalogues have 386144, 12497 and 1742 “stars” with $M_V = 1$ respectively.

The positions and velocities of each particle are first transformed into the observables (α, δ, π) and $(\mu_\alpha, \mu_\delta, v_r)$; the expected observational “errors” are then added to the parallax, the radial velocity and the proper motion, according to Table 4.1. For GAIA the precision in the radial velocity is taken to be 5 km s^{-1} for $V < 14$, and to vary like $\sigma_v = 10(V - 14) + 10 \text{ km s}^{-1}$ up to $V = 15$. Since FAME and DIVA will not measure radial velocities on board, for these we estimate the error $\sigma_v = 15 \text{ km s}^{-1}$, as achievable from the ground for such large samples. These “observed” quantities are then transformed back to “observed” positions and velocities. We repeat this procedure 5 times to obtain 5 different realizations of the data.

4.3 Finding disrupted galaxies

4.3.1 Integrals of motion space

Our satellites disrupt relatively quickly, in only a few pericentric passages. Therefore we may consider each of the 33 satellites as an ensemble of particles with very similar integrals of motion (energy, angular momentum). As we show in Figure 4.3, initially satellites are both clumps in configuration and velocity space, as they are in (E, L, L_z) space. If these are conserved quantities, or evolve only slightly, this initial clumping should be present even after the system has phase-mixed completely. Thus the space of integrals or adiabatic invariants seems to be the

TABLE 4.1— Estimated precision in parallax (σ_π , in μas) and proper motion (σ_μ , in $\mu\text{as yr}^{-1}$) as a function of V magnitude. For FAME and DIVA we assume $\sigma_\pi = \sigma_\mu$ (based on Horner (1999) and Röser (1998), respectively). In the case of GAIA, the estimated precisions correspond to a K3 III star with no reddening, and increase to 0.2 mas at $V \sim 20$ (Gilmore et al. 1998).

		9	10	11	12	13	14	15
GAIA	σ_π	3.65	3.65	3.65	3.65	4.83	7.05	10.8
	σ_μ	2.74	2.74	2.74	2.74	3.62	5.28	8.10
FAME	σ_π	24	36	56	90	146		
DIVA	σ_π	200	250	300				

natural space to look for the substructure produced by an accreted satellite.

There are a few issues we should address here before fully discussing a method based on clumping in the integrals of motion space. To compute the energy of the particles (or stars that will be observed by GAIA for example) we need to assume a Galactic potential. To determine the success of such a method we need to understand how our lack of knowledge on the precise form of the Galactic potential influences our results. We shall therefore proceed in two steps. In the first step, we take the same potential as that used in the simulations, Model I. In the final step we use the alternative potential introduced in Sec. 4.2.2, our Model II. This last step, in which we do not know the exact form of the Galactic potential but we make a reasonable guess, is most likely to represent the real situation.

Secondly, even though the total angular momentum is not fully conserved for an axisymmetric potential (only L_z is), it evolves preserving a certain degree of coherence. The advantage of using the integrals of motion space is that the number of clumps detected in this way will represent well the total number of accretion/merging events, since unlike other methods which are only local, it singles out all the stars from a given accreted object, independently of how different their phases and velocities might be. We choose to make use of all three integrals to reduce the chances of overlap amongst different lumps, since this probability clearly depends on the dimensionality of the space.

The analogue of Figure 4.3 for particles “brighter than 15th magnitude” (roughly within 6 kpc from the Sun) in the simulations, after 12 Gyr of evolution and for the original potential, shows that, even though there is some degree of evolution, clumping remains in the integrals of motion space. In Figure 4.4 we plot the integrals of motion space for one realization of the GAIA catalogue, i.e. after error convolution. A number of substructures are clearly visible, many of which can be directly related to the initial distribution, even with the GAIA observational uncertainties taken into account. This shows that the expected observational errors for GAIA will not affect the chances of detecting such substructures. In the case of FAME the situation is not as good, as illustrated in the left panel of Figure 4.5, where the different lumps are less populated (because of the magnitude limit) and considerably more smeared out (because of the larger observational errors). For DIVA the clumping has disappeared almost completely, as shown in the right panel of the same figure.

Figure 4.6 corresponds to the same realization of the GAIA catalogue as used before, but with the energies calculated using the case of the potential of Model II. Clearly, even though the two considered potentials are different, the substructure remains. The uncertainty in the precise form of the Galactic potential therefore does not affect the chances of finding disrupted satellites.

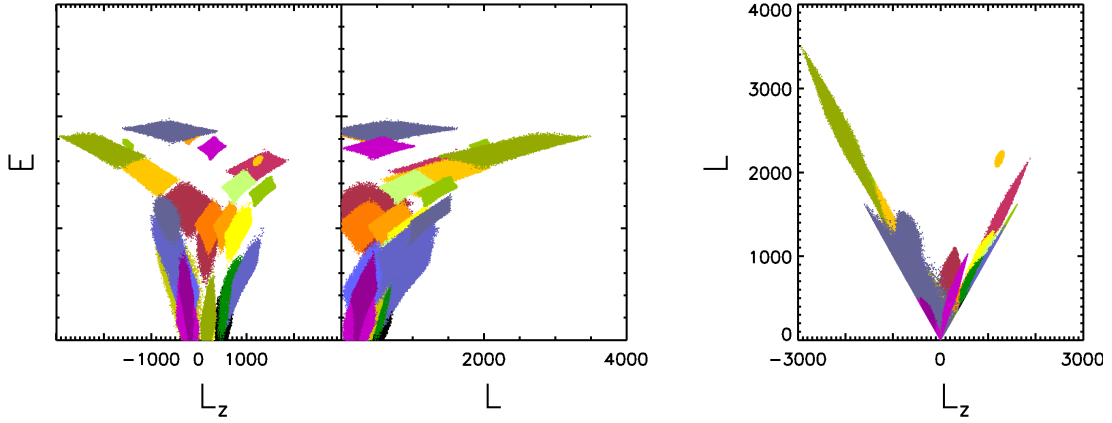


FIGURE 4.3— Initial distribution of particles in the integrals of motion space. The different colours represent different satellites.

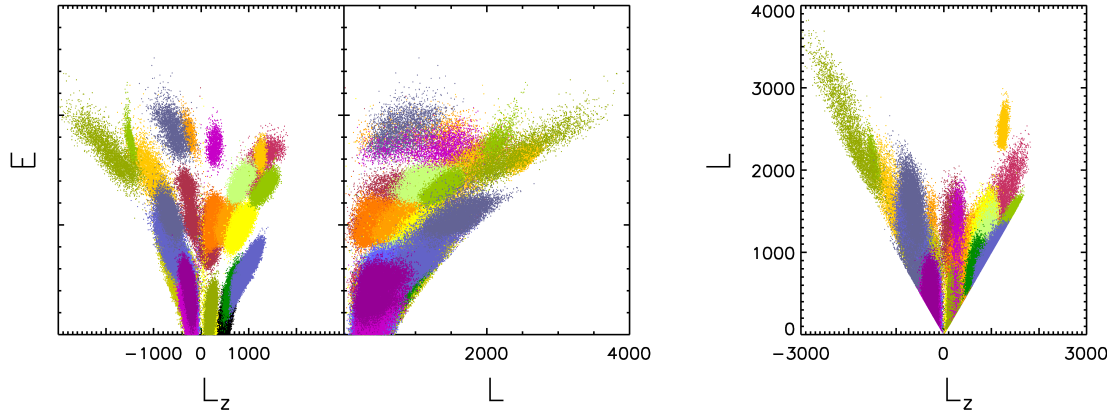


FIGURE 4.4— Final distribution of particles in the integrals of motion space after 12 Gyr, after convolution with the errors expected for GAIA for the original potential. Here we include all particles brighter than $V = 15$ (i.e. within roughly 6 kpc from the Sun).

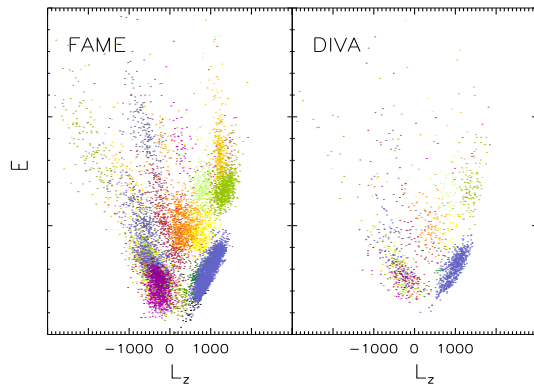


FIGURE 4.5— Final distribution of particles in the $L_z - E$ space after error convolution for FAME (left panel) and for DIVA (right panel), with energies computed using the original potential. A comparison to the left panel of Figs. 4.3 and 4.4 shows that the expected errors for these missions tend to erase much of the substructure left in the integrals of motion space.

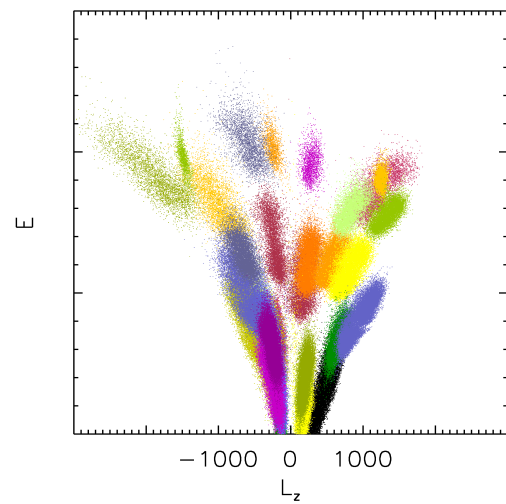


FIGURE 4.6— Final distribution of particles in the $L_z - E$ space after GAIA error convolution for the alternative potential. Compare to left panel of Fig.4.4.

4.3.2 Method: FOF in integrals of motion space

We use a Friends-of-Friends (FOF) algorithm to find clumps in the integrals of motion space. This method has been used frequently to find bound halos in cosmological N-body simulations. The basic idea is that all particle pairs separated by less than a fraction ℓ of the mean interparticle distance are linked. Disjoint sets of connected particles are then identified as halos (Efstathiou et al. 1988). These halos correspond approximately to the regions interior to isodensity contours at an overdensity of $2/\ell^3$. This FOF procedure allows a rapid identification of halos, and moreover, all members of a given halo found for a particular value of ℓ are members of the same halo in any list generated for a larger value of ℓ . In the case of cosmological simulations, the linking distance is defined so that the mean density of a halo is about 200 times the density of the Universe at the time of identification.

In our case, it is less clear how we should define the interparticle distance, or the linking length. Because the energy and angular momentum have, by definition, different scales, it seems natural to try to reduce everything to the same scale, or equivalently, to use instead of spheres an ellipsoidal configuration. Even though the angular momentum and its z -component have the same scale, lumps are generally elongated in the L -direction with a 2:1 ratio, as can be seen from Figure 4.3. We therefore search for lumps whose characteristic size would be defined as:

$$\Delta L \sim 2\Delta L_z, \quad \Delta E \sim 20\Delta L_z,$$

where now ΔL_z would be related to the linking length. This implies that we re-scale the variables according to

$$E \rightarrow E/20, \quad L \rightarrow L/2, \quad L_z \rightarrow L_z.$$

The factor 20 in the energy scaling may be derived (heuristically) from the fact that the typical energy range in the Solar neighbourhood is $1.6 \times 10^5 (\text{km s}^{-1})^2$, whereas the range of L_z is 8000 kpc km s^{-1} (from -4000 to $4000 \text{ kpc km s}^{-1}$).

We will apply the FOF algorithm for two different linking lengths, to allow for different characteristic sizes of the halos and resolutions in the algorithm. Note that there are particular regions in this space which are occupied by more than one satellite, even in this 3-dimensional space (this is even worse if only the $L_z - L$ plane is used), so that not each of the lumps found may correspond to only one satellite, but may have contributions of a few.

4.3.3 Results

We apply the FOF algorithm to our GAIA catalogue, including error convolution for all particles brighter than 15th magnitude, and using the original potential. We take two different values for the FOF linking length: $\ell = 16$ and $\ell = 30$, where $\Delta L_z = 5\ell$. We consider groups with at least 500 “stars”. We combine the two group catalogues to obtain a new group catalogue which contains all lumps detected. If some particles are found to belong to two different clumps (one from each catalogue) we keep the lump which has the smallest size. We now iterate one more time on the catalogue defined by the particles that do not belong to any of the lumps found by our FOF, again for the two values of ℓ and with a minimum of 250 “stars”. Some of the newly found lumps can be related to those previously detected, and some others are found to resolve some of the largest lumps in our initial group catalogue. In the left panel of Figure 4.7 we show the distribution of energy E and L_z for our final group catalogue.

We find 17 different groups with this method. Not all the groups may be associated exclusively with one of our original satellites. As can be seen from Fig.4.3, there is quite a bit of superposition in this three-dimensional space, and so not all the original satellites can be recovered, or equivalently, not all lumps can be resolved with just two iterations. If we analyse how

the particles in the different lumps can be related to particles in the initial satellites we find that, out of the 17 groups discovered, 14 can be associated almost uniquely to one satellite[†]. This means that our simple method is capable of finding more than 40% of all the satellites that were accreted by our “Galaxy”.

Similarly, we apply the FOF algorithm to the same GAIA catalogue but now compute the energy E of the particles with the alternative potential. In this case we again find 17 different lumps (after two iterations, and combining the results of the two different values of the linking length). Of these 17 groups, 14 can be uniquely associated to one satellite. This is shown in the central panel of Figure 4.7. Our method is thus quite successful in identifying disrupted satellite galaxies in integrals of motion space, even when we only have a guess for the Galactic potential and when the observational uncertainties are taken into account.

When we apply the same method for the original potential on the FAME catalogue we are able to find 6 different groups. For 5 of these a unique correspondence with an accreted satellite exists, as shown in the rightmost panel of Figure 4.7. In the case of DIVA we find only 1 group (with at least 20 particles), which can be easily identified visually from the right panel in Figure 4.5. In the cases of DIVA and FAME we used slightly larger linking lengths to take into account the smearing out of the lumps caused by the larger observational uncertainties. Because the samples are also smaller (because of the limiting magnitude), we consider groups with at least 50 particles in the case of FAME, and 20 particles for DIVA.

4.4 Clumpiness in the kinematics of halo stars

In Figure 4.8 we show the velocities of all the particles contained in a volume of 2 kpc on a side for one realization of the GAIA catalogue. There is considerable substructure, which is visible thanks to the great precision that GAIA will achieve. From the upper panels it is clear that, as discussed in the introduction, distinguishing the satellites that gave rise to each one of the different moving groups is a non-trivial task in this space. In the lower panels we have coloured the different contributions from the 14 groups detected by our FOF algorithm in the case of the alternative potential. A comparison between upper and lower panels also shows how successful our method is.

The kinematically cold streams visible in Figure 4.8 remain as coherent structures for longer than a Hubble time. This is true even when mergers, rather than simple satellite accretion, are dominant (Helmi, White & Springel 2000). The clumpiness in the kinematics of halo stars should thus be a distinct feature of the hierarchical formation of our Galaxy. It is therefore also interesting to determine the degree of the clumpiness and whether it will be measurable with future astrometric missions. We determine this clumpiness using the two-point correlation function ξ in velocity space for a sphere of 1 kpc radius around the Sun. We estimate ξ from

$$\xi = \frac{\langle DD \rangle \langle RR \rangle}{\langle DR \rangle^2} - 1 \quad (4.14)$$

(e.g. Hamilton 1993) where $\langle DD \rangle$ is the normalized number of pairs of particles with velocities in a given velocity range (or bin), i.e.

$$\langle DD \rangle = \frac{\sum \text{pairs of particles } i, j \text{ with } v < |\mathbf{v}_i - \mathbf{v}_j| < v + \Delta}{N_D(N_D - 1)} \quad (4.15)$$

and N_D is the number of particles in the sphere. $\langle RR \rangle$ is defined analogously but for N_R random points. The random variates are drawn from a trivariate Gaussian distribution determined from

[†]We say that a group is almost uniquely associated to one satellite if more than 70% of the particles in the group belong to only that satellite.

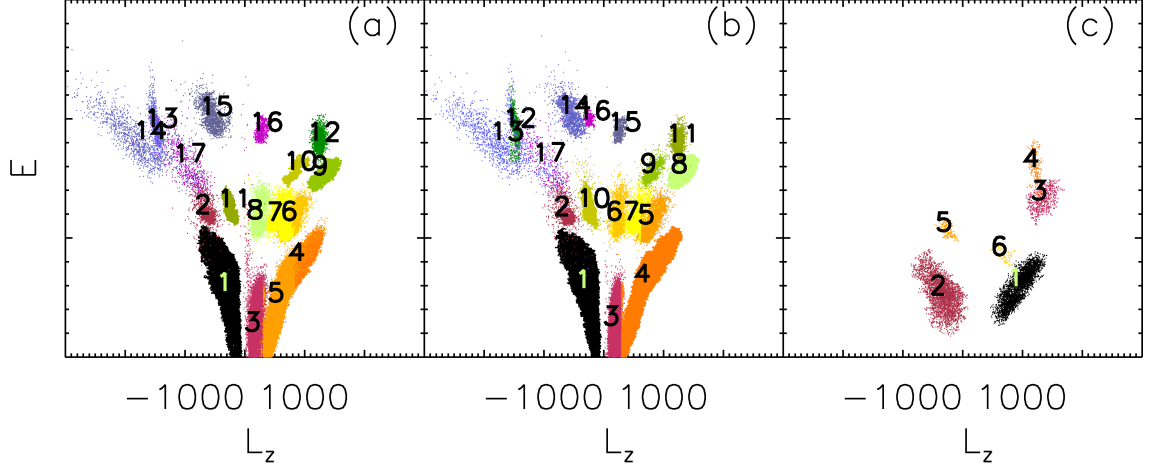


FIGURE 4.7— Lumps detected with our FOF algorithm. In panel (a) we show the final group catalogue for the original potential used in the simulations after convolving with the observational errors expected for GAIA. Panel (b) corresponds to our alternative potential and also to the GAIA catalogue. In both cases the recovery rate is about 50%. Panel (c) shows the lumps recovered by our FOF applied to the FAME catalogue generated as described in the text and for energies computed with the original potential. Compare to Fig. 4.4 in the case of GAIA and to the left panel in Fig.4.5 for FAME.

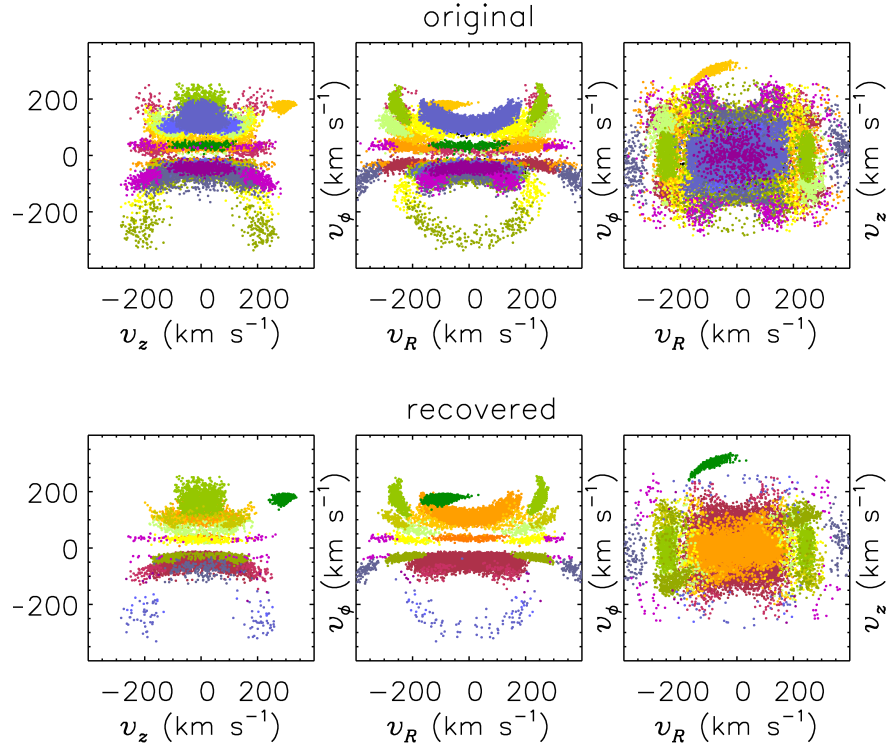


FIGURE 4.8— The velocity space distribution for particles in a cubic volume of 2 kpc on a side centred on the Sun for one realization of the GAIA catalogue. In the upper panels different colours indicate particles associated with different satellites (using the same colour coding as in Figure 4.3). In the lower panels, the colours are used to show particles associated to the lumps recovered by our FOF algorithm applied to the GAIA catalogue in the case of the alternative Galactic potential. (Here the colour coding corresponds to that used in panel (b) of Figure 4.7.)

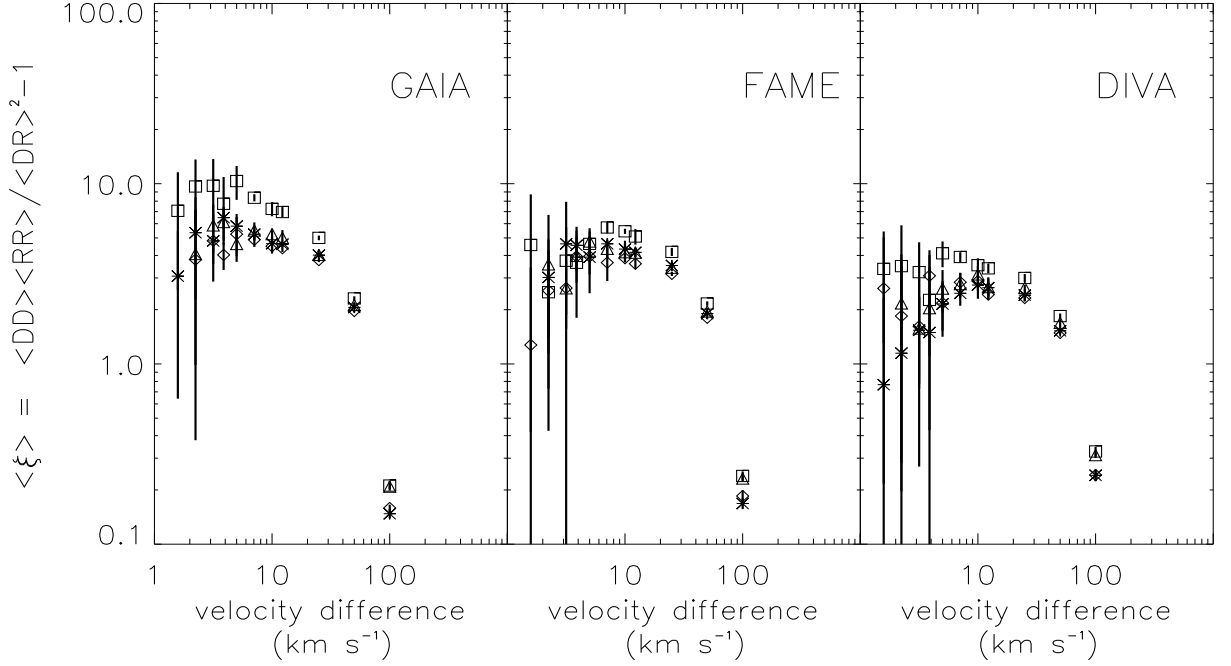


FIGURE 4.9— The two-point correlation function for “giant stars” inside spheres of 1 kpc radius around the Sun (defined as 8 kpc from the Galactic centre on the Galactic disk) computed as the weighted average over five realizations of the DIVA, FAME and GAIA catalogues. The different symbols correspond to ξ measured inside spheres at different locations of the “Sun” on the Solar circle.

the “data” in the principal axes velocity frame. Here we take $N_R = 10N_D$. Finally $\langle DR \rangle$ are the normalized counts for “data”–random pairs. We estimate the uncertainty from

$$\Delta_\xi = (1 + \xi) \sqrt{\frac{2}{N_D(N_D - 1)\langle DD \rangle}}. \quad (4.16)$$

If the sample contains kinematically cold streams, we should find an excess of pairs in the bins corresponding to small velocity differences, i.e. the correlation function should be significantly different from zero (which corresponds to the absence of correlations). We proceed by measuring ξ for our GAIA, FAME and DIVA catalogues including error convolution as described in Sec. 4.2.4. We also vary the position of the 1 kpc sphere around the “Sun”, keeping the same distance from the “Galactic centre”. That is, we place the Sun at $(x, y) = \{(8, 0), (0, -8), (-8, 0), (0, 8)\}$ kpc and $z = 0$. This allows us to account for the natural variations one may have from volume to volume. We then make five realizations for each “Sun” position for each catalogue. In Figure 4.9 we show the correlation function obtained by averaging over all the realizations, for each 1 kpc sphere and for each catalogue. The average ξ for each volume is the weighted mean, where the weights are given by $1/\Delta_\xi^2$, and the error bars indicate the (weighted) dispersion around the (weighted) mean. We find an excess of pairs of stars with similar motions, the signature indicating the presence of cold streams as expected for a stellar halo built by disrupted satellites. Note that it will even be possible to determine that the halo is not a smooth distribution in the Solar neighbourhood even with velocity errors of the order of 20 km s^{-1} such as those expected for DIVA for a star with $M_V = 1$ at 500 pc from the Sun.

4.5 Discussion

We simulated the stellar halo of the Galaxy starting from disrupted satellite galaxies, and “observed” it with the next generation of astrometric satellites (DIVA, FAME and GAIA). We analysed the observations with the aim of recovering the different accretion events our “Galaxy” experienced over its lifetime. We used a FOF algorithm to find clumps in the integrals of motion space, which we expected would correspond to the disrupted satellites. Our integrals of motion space is defined by energy E , total angular momentum L and its z -component L_z , even though strictly speaking these are not fully conserved quantities (because of interaction of the stars while still bound to the satellite, and because of the axisymmetry of our Galaxy). We have shown that the initial clumping in this space is maintained to a great extent even after 12 Gyr of evolution.

After using our FOF algorithm we find that we can only recover a couple of accreted satellites (in our analysis just one) for the DIVA catalogue, whereas for FAME we recover about 15% of all satellites. In both cases we assume that the astrometry is complemented by ground based radial velocity measurements. The situation is significantly different in the case of the GAIA catalogue, for which we recover almost half of all disrupted satellites with this simple algorithm. The improvement generally lies in the larger volume for which full 6-D information is available, in particular when comparing FAME and GAIA. The use of 6-D information appears to be essential to recover all the events, as there is a large fraction of phase-space where these are superposed. This is particularly clear from Figure 4.4 (rightmost panel), where angular momentum alone cannot be used to distinguish the different satellites. Whereas by eye inspection in the (E, L, L_z) space we may recover five or six events, for the space (L, L_z) this is reduced to one or two events.

The evolution of the Galactic potential may be the most crucial simplification in our analysis. In hierarchical cosmologies the number of objects that form a galaxy like our own is in the range of 5 – 20, with comparable masses. The process of formation is likely to be very violent and the potential is surely not static, quite probably not axisymmetric, and therefore the initial clumping of the system may not be reflected in clumping in our defined integrals of motion space. However, if this happened during the first few Gyr, any object infalling later ought to have perceived a fairly static (or adiabatically changing) Galaxy, and then our method would still be useful. Indeed, some preliminary analysis of the formation of a halo in a Λ CDM cosmology indicates that particles from different satellites may be recovered as lumps in this space (Helmi et al. 2000), though the structure is less evident than in the plots shown here, where even by simple eye inspection one may recover about 1/5 of all satellites.

What will anyway remain as signatures of the merger history of our Galaxy will be the kinematically cold streams originating in disrupted halos. An interesting observational test is the comparison of the kinematics of a smooth, possibly Gaussian, distribution (which may be expected in the case of a monolithic collapse) to the kinematics observed in the stellar halo built by disrupted satellites. Our analysis of the correlation function in velocity space indicates the presence of a larger number of streams with very small velocity dispersions in a sphere of 1 kpc radius around the Sun. This test will be feasible even for DIVA. The key to the success of this test lies in the complete and large sample of stars with 3-D velocities which will be available.

In this paper we have focused on determining the merger history of the Milky Way, rather than the precise form of the Galactic potential or to what extent it may have varied. However, these are key questions that will be solved very likely by SIM and GAIA (e.g. Johnston et al. 1999). We may add here that after finding the different satellites we will be able to determine the conditions and characteristics of objects that fell onto the Milky Way more than 10 Gyr ago.

Acknowledgments

We wish to thank Volker Springel for his FOF algorithm, Simon White for many useful discussions, and Anthony Brown for comments on an earlier version of this manuscript.

References

- Burstein D., Bender R., Faber S.M., Nolthenius R., 1997, *AJ*, 114, 1365
- Chereul E., Cr     M., Bienaym   O., 1999, *A&AS*, 135, 5
- de Bruijne J.H.J., 1999, *MNRAS*, 306, 381
- de Zeeuw P.T., Hoogerwerf R., de Bruijne J. H. J., Brown A. G. A., Blaauw A., 1999, *AJ*, 117, 354
- Efstathiou G., Frenk C.S., White S.D.M., Davis M., 1988, *MNRAS*, 235, 715
- Fuchs B., Jahre     H., 1998, *A&A*, 329, 81
- Gerhard O., 1991, *MNRAS*, 250, 812
- Gilmore G. F., Perryman M. A., Lindegren L., Favata F., Hoeg E., Lattanzi M., Luri X., Mignard F., R     S., de Zeeuw P. T., 1998, *SPIE*, 3350, 541
- Gradshteyn I.S., Ryzhik I.M., 1965, *Table of integrals, series and products*, Academic Press
- Hamilton A.J., 1993, *ApJ*, 417, 19
- Helmi A., White S.D.M., 1999, *MNRAS*, 307, 495 (Chapter 2)
- Helmi A., White S.D.M., de Zeeuw P.T., Zhao H.S., 1999, *Nature*, 402, 53 (Chapter 3)
- Helmi A., White S.D.M., Springel V., 2000, in preparation (Chapter 6)
- Hernquist L., 1993, *ApJS*, 86, 389
- Hoogerwerf R., Aguilar L., 1999, *MNRAS*, 306, 394
- Horner S.D., Germain M.E. et al., 2000, in "Working on the Fringe: An International Conference on Optical and IR Interferometry from Ground and Space", *ASP Conf. Series* (in press)
- Ibata R., Gilmore G., Irwin M., 1994, *Nature*, 370, 194
- Johnston K.V., Hernquist L., Bolte M., 1996, *ApJ*, 465, 278
- Johnston K.V., Zhao H.S., Spergel D.N., Hernquist L., 1999, *ApJ*, 512, L109
- Kinman T.D., Suntzeff N.B., Kraft R.P., 1994, *AJ*, 108, 1722
- Lynden-Bell D., Lynden-Bell R.M., 1995, *MNRAS*, 275, 429
- Lynden-Bell D., in "The Galactic Halo", *Proc. of the 3rd. Stromlo Symposium*, ed. Gibson B. et al., *ASP Conf. Series* 165, p. 17
- Majewski S.R., Munn J.A., Hawley S.L., 1996, *ApJ*, 459, L73
- Quinn P.J., Goodman J., 1986, *ApJ*, 309, 475
- R     S., 1998, in *Jahrestagung der Astronomischen Gesellschaft, Heidelberg: "DIVA - Beyond HIPPARCOS and Towards GAIA"* (<http://www.aip.de/groups/DIVA/>)
- van den Bosch F.C., Lewis G.F., Lake G., Stadel J., 1999, *ApJ*, 515, 50
- van der Marel R.P., Sigurdsson S., Hernquist L., 1997, *ApJ*, 487, 153
- Vel         H., White S.D.M., 1995, *MNRAS*, 275, 23L
- White S.D.M., Rees M.J., 1978, *MNRAS*, 183, 341
- White S.D.M., 1983, *ApJ*, 274, 53
- Zaritsky D., White S.D.M., 1988, *MNRAS*, 235, 289

*Thus, the task is, not so much to see what no one
has yet seen; but to think what nobody has yet
thought, about that which everybody sees.*

E. Schrödinger

Chapter 5

Simple dynamical models of the Sagittarius dwarf galaxy*

ABSTRACT

We present two simple dynamical models for Sagittarius based on N-body simulations of the progressive disruption of a satellite galaxy orbiting for 12.5 Gyr within a realistic Galactic potential. In both models the satellite initially has observable properties similar to those of current outlying dwarfs; in one case it is purely stellar while in the other it is embedded in an extended massive halo. The purely stellar progenitor is a King model with a total velocity dispersion of 18.1 km s^{-1} , a core radius of 0.56 kpc and a tidal radius of 3.8 kpc. The initial stellar distribution in the other case follows a King profile with the same core radius, a similar total velocity dispersion and a smaller extent. Both these models are consistent with all published data on the current Sagittarius system, they match not only the observed properties of the main body of Sagittarius, but also those reported for unbound debris at larger distances.

*Based on: Amina Helmi and Simon D.M. White, submitted to MNRAS

5.1 Introduction

The Sagittarius dwarf galaxy is the closest satellite of the Milky Way (Ibata, Gilmore & Irwin 1994, 1995, hereafter IGI95). Soon after its discovery, several groups carried out simulations to see if its properties are consistent with the disruption of an object similar to the other dwarf companions of the Milky Way, but none produced a model in full agreement with both the age and the structure of the observed system (Johnston, Spergel & Hernquist 1995; Velázquez & White 1995; Edelson & Elmegreen 1997; Ibata et al. 1997, hereafter I97). All groups assumed light to trace mass and an initial system similar to observed dwarf spheroidals. All found the simulated galaxy to disrupt after one or two orbits whereas the observed system has apparently completed ten or more. Most considered this to be a problem (but cf Velázquez & White 1995). As a result, several unconventional models were proposed to explain the survival and structure of Sagittarius. In an extensive numerical study, Ibata & Lewis (1998) concluded that Sagittarius must have a stiff and extended dark matter halo if it is to survive with 25% of its initial mass still bound today. Since an extended halo cannot remain undistorted in the Galaxy's tidal field for any conventional form of dark matter, it is unclear how this idea should be interpreted. Furthermore, it produces an uncomfortably large mass-to-light ratio (~ 100), it cannot reproduce the observed elongation, and it suggests that little tidal debris will be liberated, in apparent conflict with the observations of Mateo, Olszewski & Morrison (1998), and Majewski et al. (1999) (see also Johnston et al. 1999). A somewhat less unorthodox model was proposed by Zhao (1998), where Sagittarius was scattered onto its current tightly bound orbit by an encounter with the Magellanic Clouds about 2 Gyr ago. This appears physically possible but requires careful tuning of the orbits of the two systems (see Ibata & Lewis 1998; and Jiang & Binney 2000).

Driven by this apparent puzzle, we decided to search more thoroughly for a self-consistent model of the disruption of Sagittarius, which, after a Hubble time, has similar characteristics to those observed. (See Table 5.1 for a summary of the observed properties of the system.) Below we present two models which meet these requirements.

TABLE 5.1— Properties of Sagittarius (IGI95, I97)

Orbital properties	
distance from the Sun d	25 ± 2 kpc
heliocentric radial velocity v_r^{sun}	140 ± 2 km s $^{-1}$
proper motion in b μ_b	250 ± 90 km s $^{-1}$
gradient along the orbit dv_r/db	< 3 km s $^{-1}/\text{deg}$
angular position in the sky (l, b)	$(5.6^\circ, -14^\circ)$
Internal properties	
luminosity	$\gtrsim 10^7 L_\odot$
velocity dispersion $\sigma(v_r)$	11.4 ± 1 km s $^{-1}$
angular extent in (l, b)	$8^\circ \times 3^\circ$
half-mass radius	0.55 kpc
mean metallicity $\langle [\text{Fe}/\text{H}] \rangle$	~ -1 dex

5.2 Method

In our numerical simulations, we represent the Galaxy by a fixed potential with three components: a dark logarithmic halo

$$\Phi_{\text{halo}} = v_{\text{halo}}^2 \ln(r^2 + d^2), \quad (5.1)$$

a Miyamoto–Nagai disk

$$\Phi_{\text{disk}} = -\frac{GM_{\text{disk}}}{\sqrt{R^2 + (a + \sqrt{z^2 + b^2})^2}}, \quad (5.2)$$

and a spherical Hernquist bulge

$$\Phi_{\text{bulge}} = -\frac{GM_{\text{bulge}}}{r + c}, \quad (5.3)$$

where $d = 12$ kpc and $v_{\text{halo}} = 131.5$ km s⁻¹; $M_{\text{disk}} = 10^{11} M_{\odot}$, $a = 6.5$ kpc and $b = 0.26$ kpc; $M_{\text{bulge}} = 3.4 \times 10^{10} M_{\odot}$ and $c = 0.7$ kpc. This choice of parameters gives a flat rotation curve with an asymptotic circular velocity of 186 km s⁻¹.

We represent the satellite galaxy by a collection of 10^5 particles and model their self-gravity by a multipole expansion of the internal potential to fourth order (White 1983; Zaritsky & White 1988). This type of code has the advantage that a large number of particles can be followed in a relatively small amount of computer time. Hence a substantial parameter space can be explored while retaining considerable detail on the structure of the disrupted system. In this quadrupole expansion, higher than monopole terms are softened more strongly. We choose $\epsilon_1 \sim 0.2 - 0.25r_c$ for the monopole term (r_c is the core radius of the system) and $\epsilon_2 = 2\epsilon_1$ for dipole and higher terms and for the centre of expansion. The centre of expansion is a particle which, in practice, follows the density maximum of the satellite closely at all times.

For the stellar distribution of the pre-disruption dwarf we choose a King model (King 1966), since this is a good representation of the distant dwarf spheroidals. King models are defined by a combination of three parameters: $\Psi(r=0)$ (depth of the potential well of the system), σ^2 (measure of the central velocity dispersion), and ρ_0 (central density) or r_0 (King radius). The ratio $\Psi(r=0)/\sigma^2$ defines how centrally concentrated the system is, and for any value of this parameter, a set of homologous models with different central densities and core (or King) radii may be found. We assume that the progenitor of Sagittarius obeys the known metallicity–luminosity relation for the Local Group dSph (Mateo 1998). The metallicity determinations for Sagittarius (I97) indicate $\langle [\text{Fe}/\text{H}] \rangle \sim -1$, corresponding to a *total* luminosity in the range $3.5 \times 10^7 - 3.5 \times 10^8 L_{\odot}$. To obtain an initial guess for the mass of the system, we transform this luminosity into a mass assuming a mass-to-light ratio ~ 2 . The relevant initial stellar mass interval is then $7 \times 10^7 - 7 \times 10^8 M_{\odot}$.

Note that our choice of a fixed potential to represent our Galaxy means that we neglect any exchange of energy between the satellite and the Galactic halo. This is an excellent approximation for the range of orbits and satellite masses that we consider, since these imply dynamical friction decay times substantially in excess of the Hubble time. The orbits are also sufficiently large that impulsive heating during disk passages can be neglected.

The orbit of Sagittarius is relatively well constrained (I97). The heliocentric distance $d \sim 25 \pm 2$ kpc and position $(l, b) = (5.6^\circ, -14^\circ)$ of the galaxy core are well-determined; the heliocentric radial velocity $v_r^{\text{sun}} \sim 140 \pm 2$ km s⁻¹, and its variation across the satellite are also accurately measured. Outside the main body ($b < -20^\circ$) the radial velocity shows a small gradient $dv_r/db \lesssim 3$ km s⁻¹ deg⁻¹, but no gradient is detected across the main body itself. The proper motion measurements are not very accurate; $\mu_b \sim 2.1 \pm 0.7$ mas yr⁻¹, and no measurement is available in the l -direction. On the other hand the strong North-South elongation of the system suggests that it has little motion in the l -direction. We generate a range of possible orbits satisfying these constraints and concentrate on those with relatively long periods in order to maximise the survival chances of our satellite. We begin all our simulations half a radial period after the Big Bang to allow for the initial expansion. We place the initial satellite at apocentre, then we integrate forward until ~ 13 Gyr. The orbits are chosen so at this time the position and velocity

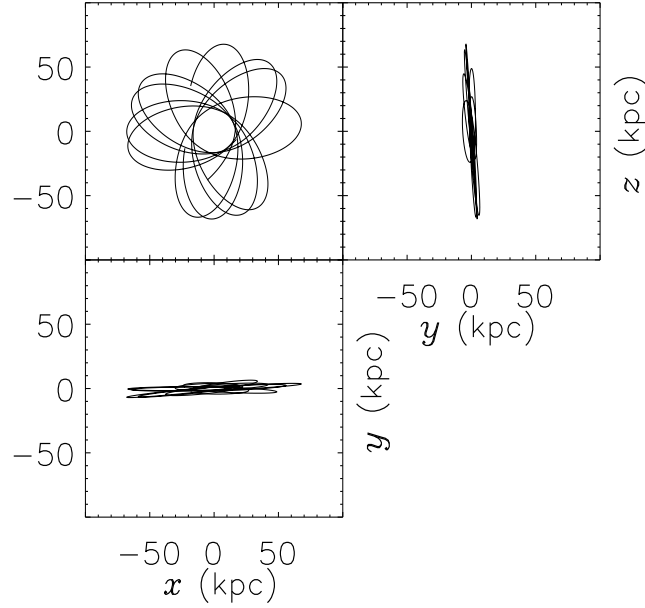


FIGURE 5.1 — Projections of a possible orbit of Sagittarius on different orthogonal planes, where xy coincides with the plane of the Galaxy.

of the satellite core correspond to those observed. We allow ourselves some slight freedom in choosing the final time in order to fit the observed data as well as possible.

5.3 Results

Figure 5.1 gives an example of an orbit which is consistent with all the current data on Sagittarius. It has a pericentre of 16.3 kpc, an apocentre of 68.3 kpc, and a radial period of ~ 0.85 Gyr.

After letting our satellite relax in isolation, we integrate each simulation for ~ 13 Gyr. In practice we needed to run a large number of simulations, and test each to see if it satisfies the observational constraints at the present time. Since it remains uncertain whether dwarf spheroidals have extended dark halos (e.g Klessen & Kroupa 1998), we have considered both purely stellar models and models in which the initial stellar system is embedded in a more massive and more extended dark halo.

5.3.1 Constant mass-to-light ratio: A purely stellar model

Our preferred purely stellar model (Model I) initially has a core radius of $r_c = 0.56$ kpc, a total velocity dispersion of 18.1 km s^{-1} , and a concentration parameter $c = \log_{10}(r_t/r_c) \sim 0.83$. This implies a total mass of $M = 5.74 \times 10^8 M_\odot$. For a satellite to survive for about 10 Gyr on an orbit with pericentre ~ 15 kpc, apocentre ~ 70 kpc, and period ~ 1 Gyr (for which the observational constraints are satisfied) its initial central density has to be $\rho_0 \geq 0.25 - 0.3 M_\odot \text{ pc}^{-3}$. Satellites with significantly smaller initial densities do not survive long enough.

In Figure 5.2 we plot heliocentric distance as a function of galactic latitude for particles projected near the main remnant 12.5 Gyr after infall. Streams of particles are visible at all latitudes over a broad range in distance. Sagittarius has been orbiting long enough for its debris streams to be wrapped several times around the Galaxy. (See also Figure 5.6.)

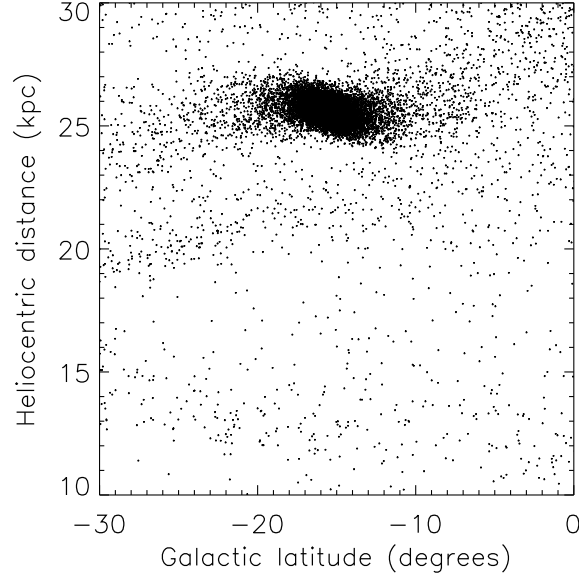


FIGURE 5.2— Distribution of particles in distance from the Sun as a function of latitude. For direct comparison see Figure 4 of I97.

The remnant galaxy, i.e. the central region of the satellite’s debris, is similar to the real system. In Figure 5.3 we plot its mass surface density. The transformation from observed surface brightness to mass surface density (which is what the simulations give us) can be done as follows. The observed mass surface density Σ for an assumed mass-to-light ratio Υ is

$$\Sigma = \frac{N_X L_X}{f_X} \Upsilon \left[\frac{\text{M}_\odot}{\text{deg}^2} \right], \quad (5.4)$$

where N_X is the number of observed stars of type X per square degree, L_X is their luminosity, and f_X is the fraction of the total luminosity in stars of type X . In IGI95 the spatial structure of Sagittarius was determined from the excess of counts at apparent magnitude of the horizontal branch. Uncertainties in the result are due primarily to contamination by sources in the Galactic bulge. Their lowest isodensity contour is at $\Sigma_{\min} \sim 5 \times 10^5 \text{ M}_\odot \text{ deg}^{-2}$, assuming $\Upsilon \sim 2.25$ and $[\text{Fe}/\text{H}] \sim -1$ (Bergbusch & Vandenberg 1992), and has an extent of $7.5^\circ \times 3^\circ$. This same isodensity contour is shown in Figure 5.3 as a thick line. It has an extent of $\sim 8^\circ \times 4^\circ$, in reasonable agreement with the observations given the uncertainties. In I97 isodensity contours were derived from counts of main sequence stars close to the turn-off, roughly one magnitude above their plate limit. The minimum contour in this case corresponds to $\Sigma_{\min} \sim 10^5 \text{ M}_\odot \text{ deg}^{-2}$, and has an extent of roughly $15^\circ \times 7^\circ$. In Figure 5.3 this contour is shown as a dashed-line, and has an extent of $15^\circ \times 5.5^\circ$, also in good agreement with the observations. Note that the isophotes (or isodensity contours) become rounder towards the centre of the satellite. Its angular core radius is $R_c \sim 1.24^\circ$, which for a distance of 25.6 kpc (derived from the simulations) corresponds to 0.55 kpc, again in good agreement with the observations.

The kinematic properties of the remnant galaxy are more difficult to compare with observations because a substantial amount of mass from debris streams is projected on top of the main body. Like I97, we measure the radial velocity across the system considering only particles for

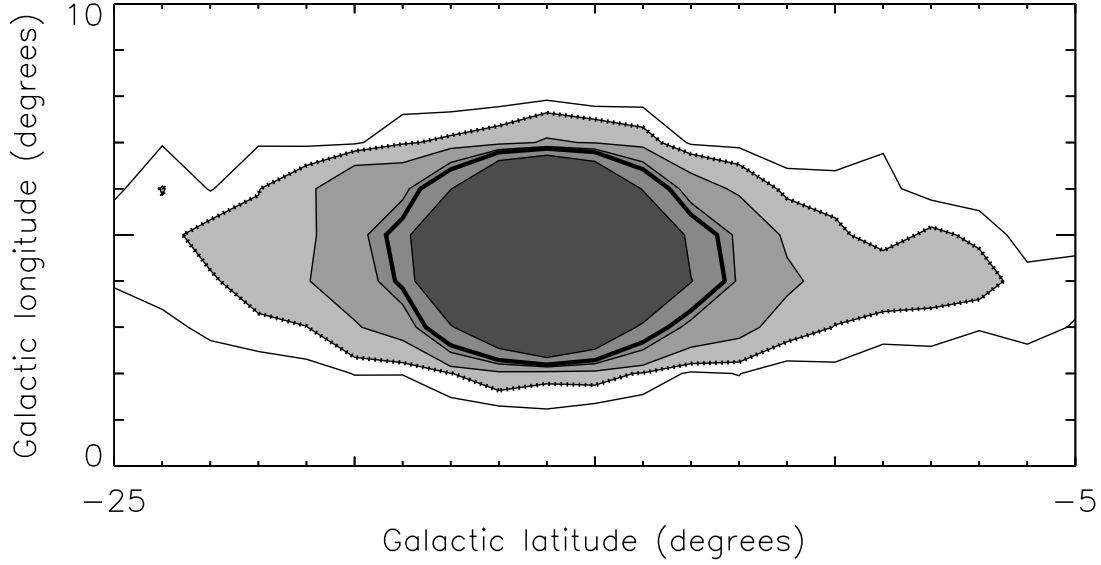


FIGURE 5.3— Surface isodensity contours for the remnant system. The thick and dashed lines indicate the contours that, for $M/L = 2.25$, would correspond to the minimum contours plotted in 1994 and in 1997 respectively by Ibata and collaborators. Each succeeding contour has half the mass surface density of the previous one.

which $100 \text{ km s}^{-1} \leq v_r^{\text{sun}} \leq 180 \text{ km s}^{-1}$. In the left panel of Figure 5.4 we plot the heliocentric radial velocity, and in the right panel we plot its dispersion as a function of Galactic latitude. For comparison, we analysed the observations of I97 at Cerro Tololo in the same way (their Table 2b); these data have a precision of a few km s^{-1} (triangles in Figure 5.4). Our model is consistent with the observed kinematics; we obtain a heliocentric radial velocity of 138 km s^{-1} and an internal velocity dispersion in the radial direction of 10.3 km s^{-1} for the main body. However, when the radial velocity restrictions for inclusion in this calculation are relaxed, we find much larger velocity dispersions because of the contribution of stars from other streams. It is important to consider this problem when determining which stars should be considered members of Sagittarius.

5.3.2 Varying mass-to-light ratio: A model with a dark halo

The observational data for Sagittarius mainly refer to the current remnant system, which corresponds to the innermost regions of the progenitor satellite. As a consequence, models that are initially dark matter dominated in their outskirts are relatively poorly constrained.

As an example we focus on a progenitor with a mass distribution which is similar to that of Model I in its inner regions but is more extended. We choose the mass-to-light ratio of satellite material to be a decreasing function of binding energy, so that the most bound particles have near “stellar” mass-to-light ratios, whereas weakly bound particles are almost entirely “dark”. We take the initial mass distribution to be a King model with $r_c = 0.7 \text{ kpc}$ and $r_t = 13.2 \text{ kpc}$. For an orbit like that of Model I this produces a suitable remnant after 12 Gyr for an initial total velocity dispersion of 17.5 km s^{-1} , giving a total initial mass of $M = 1.26 \times 10^9 M_\odot$. The mass distribution of this remnant satisfies many of the observational constraints of Table 5.1. Its angular core radius is slightly larger $R_c \sim 1.24^\circ$, and the radial velocity dispersion in the main body is 12.5 km s^{-1} . We construct a two-component satellite with this mass distribution

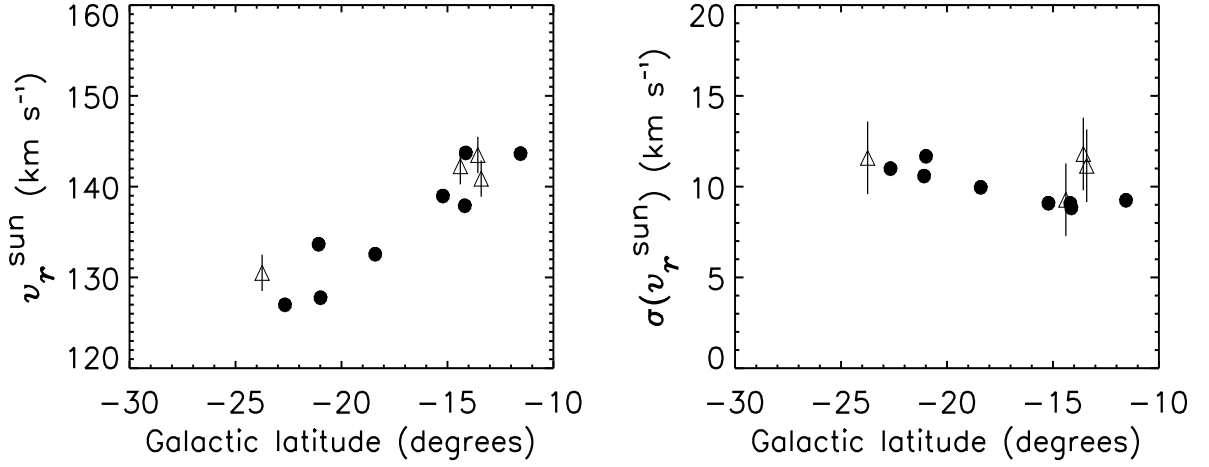


FIGURE 5.4— In the left panel we plot mean heliocentric radial velocity as a function of Galactic latitude, for bins of $\sim 2.5^\circ \times 2.5^\circ$ across the remnant system. The right panel shows the heliocentric radial velocity dispersion in the same bins. To determine variations across the main body of the galaxy, we have taken bins centered on the same Galactic latitude but offset in Galactic longitude. The triangles correspond to data from I97, error bars indicate 2 km s^{-1} uncertainty.

by solving for the dependence of mass-to-light ratio on initial binding energy that produces an initial light profile which is approximately a King model with $r_c = 0.56 \text{ kpc}$ and $r_t = 2.44 \text{ kpc}$. If we identify the central mass-to-light ratio of this system as the “stellar” value, its total to stellar mass ratio is 6. In fact this is a lower limit, since there must be some dark matter also at the centre of the system.

The actual value of the central mass-to-light ratio is chosen so that the central surface brightness of the remnant agrees with that of Sagittarius. This requires $\Upsilon = 3.25$. The total luminosity of the initial model is then $6.87 \times 10^7 L_\odot$, implying a total mass-to-light ratio of 18.3. We shall refer to this model as Model II. We can calculate its “observable” properties by weighting each simulation particle by $(M/L)^{-1}$. Thus we find its initial velocity dispersion to be 17.5 km s^{-1} , and the velocity dispersion of the remnant to be 12.8 km s^{-1} . The visible extent of the remnant is also slightly smaller than the extent of its mass leading to properties which are almost identical to those of Model I and in good agreement with the observations.

Figure 5.5 compares the initial mass distribution of Model II with the distribution of its light, and with the mass distribution of Model I. Light has been converted to mass using the central mass-to-light ratio; the result thus overestimates the “actual” stellar mass as noted above. The two initial models clearly have very similar mass distributions in their inner regions, a consequence of the requirement that these regions should remain (just) bound after 12.5 Gyr. Although, by construction, the main body of the remnants is very similar in the two cases, there is a significant difference in the properties of their debris streams. In Model I the unbound debris streams are predicted to contain 7.9 times the light in the main body of the remnant ($M_V \sim -13.8$), as defined by the dotted contour in Figure 5.3, whereas in Model II ($M_V \sim -14.5$) this ratio is 0.24. If we had chosen Model II to be a constant mass-to-light ratio model with the same central value as before, we would have got an almost equally good fit to the main body of Sagittarius, but would have predicted this ratio to be 8.2. In this last case, Sgr would have contributed $5 \times 10^8 L_\odot$ to the Galactic stellar halo in the form of debris stars (for $\Upsilon = 2.25$). Thus we see that the observed properties of the main remnant do not usefully constrain the number

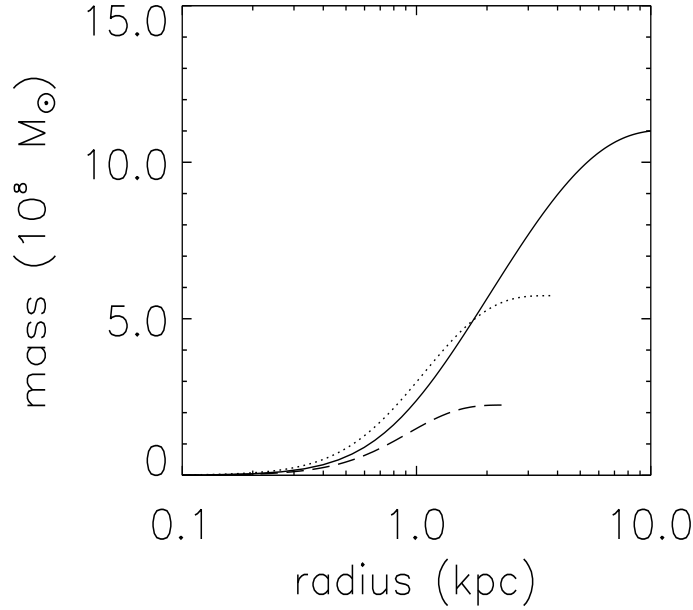


FIGURE 5.5— Mass profile for Model II (solid curve) and for its “luminous” component (dashed curve). Also shown is the mass profile for Model I (dotted curve).

of stars that may be present in the debris streams, but that the different models can be better constrained from the properties of their debris streams.

5.4 Discussion

In this section we focus for simplicity on Model I. We can use it to predict where streams originating in different mass-loss events should be found. This is illustrated in Figure 5.6 where different colours indicate material lost at different pericentric passages. Note that since the surface brightness of the unbound material decreases with time, material lost in early passages is considerably more difficult to detect than recent mass-loss (for an axisymmetric potential the time dependence is $1/t^2$, but if the potential may be considered as nearly spherical the surface density will effectively decrease as $1/t$; see Helmi & White 1999). The central panel (latitude vs. heliocentric distance) explains why Sagittarius streams have been detected below the Galactic plane but never above it. From the left panel, $-90^\circ \leq l \leq 90^\circ$, we see that the stream of stars lost in the previous pericentric passage (shown in blue) becomes more distant as we go north. For example, at $b = 40^\circ$, the stream is located approximately 50 kpc from the Sun. The red giant clump visual magnitude at this distance would be roughly 19.3^m , compared to the 17.85^m observed in the main body of Sagittarius. Unfortunately, the observations reported in the literature at this Galactic latitude, either do not reach this magnitude limit, or are offset by a few degrees from the expected location. Thus, for example, Majewski et al. (1999) have a limiting magnitude of 21 at $b = 41^\circ$ and $l = -6^\circ$, but the actual stream in our model is predicted to go through $l \sim 2^\circ$ and to be about 2° wide. Note that the width prediction is more secure than the location since the motion of Sagittarius in the l -direction is poorly constrained at present. It is also interesting to note that Majewski and collaborators claimed to detect a Sagittarius stream at $b = -40^\circ$ and $l = 11^\circ$, at a slightly smaller heliocentric distance of 23 kpc and with a radial velocity of the order of 30 km s^{-1} . As they discuss, this velocity may be strongly affected by contamination by other Galactic components. We note, however, that we would predict a stream of stars (shown in blue) going through this latitude and longitude with roughly the observed

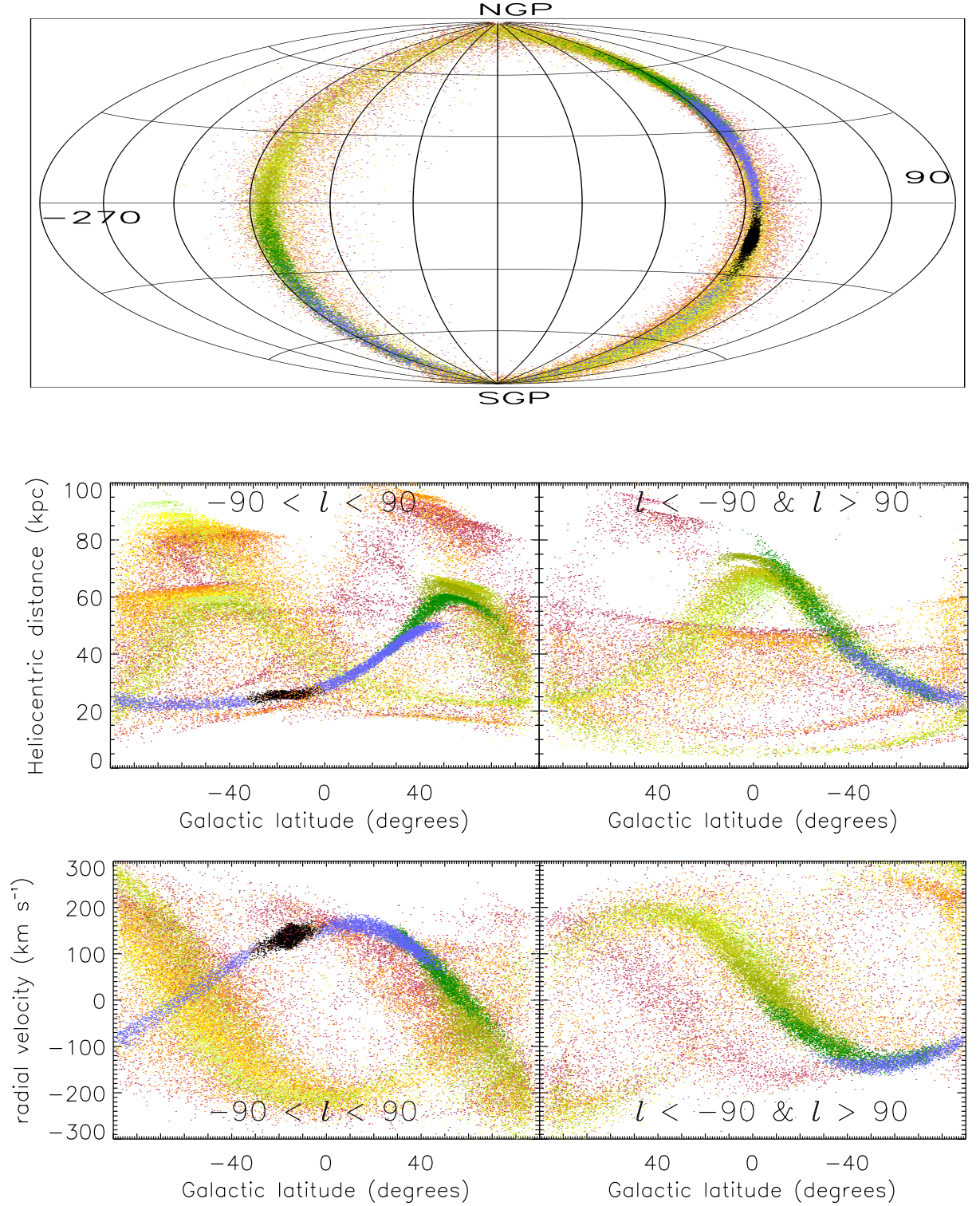


FIGURE 5.6— **Top panel:** Distribution in the sky (l, b) of the particles for our constant mass-to-light ratio model of Sagittarius after 12.5 Gyr. Different colours indicate material stripped off in different passages. **Central panel:** Heliocentric distance as a function of Galactic latitude, at the same time as the top panel, and with the same colour coding. Note that “streams” formed early on are wider than the more recent ones. **Bottom panel:** Heliocentric radial velocity as a function of Galactic latitude, at the same time and using the same colour coding as before.

distance, and with a radial velocity of 54 km s^{-1} . (See the central and bottom left panels of Figure 5.6, $-90^\circ \leq l \leq 90^\circ$.) This agreement is encouraging. As mentioned above, this stream is formed by material lost in the previous pericentric passage and not three passages ago, as in the model of Johnston et al. (1999). This difference reflects the different orbital timescales in the two models. The surface density of stars may be able to distinguish between them; it is predicted to be higher in our case.

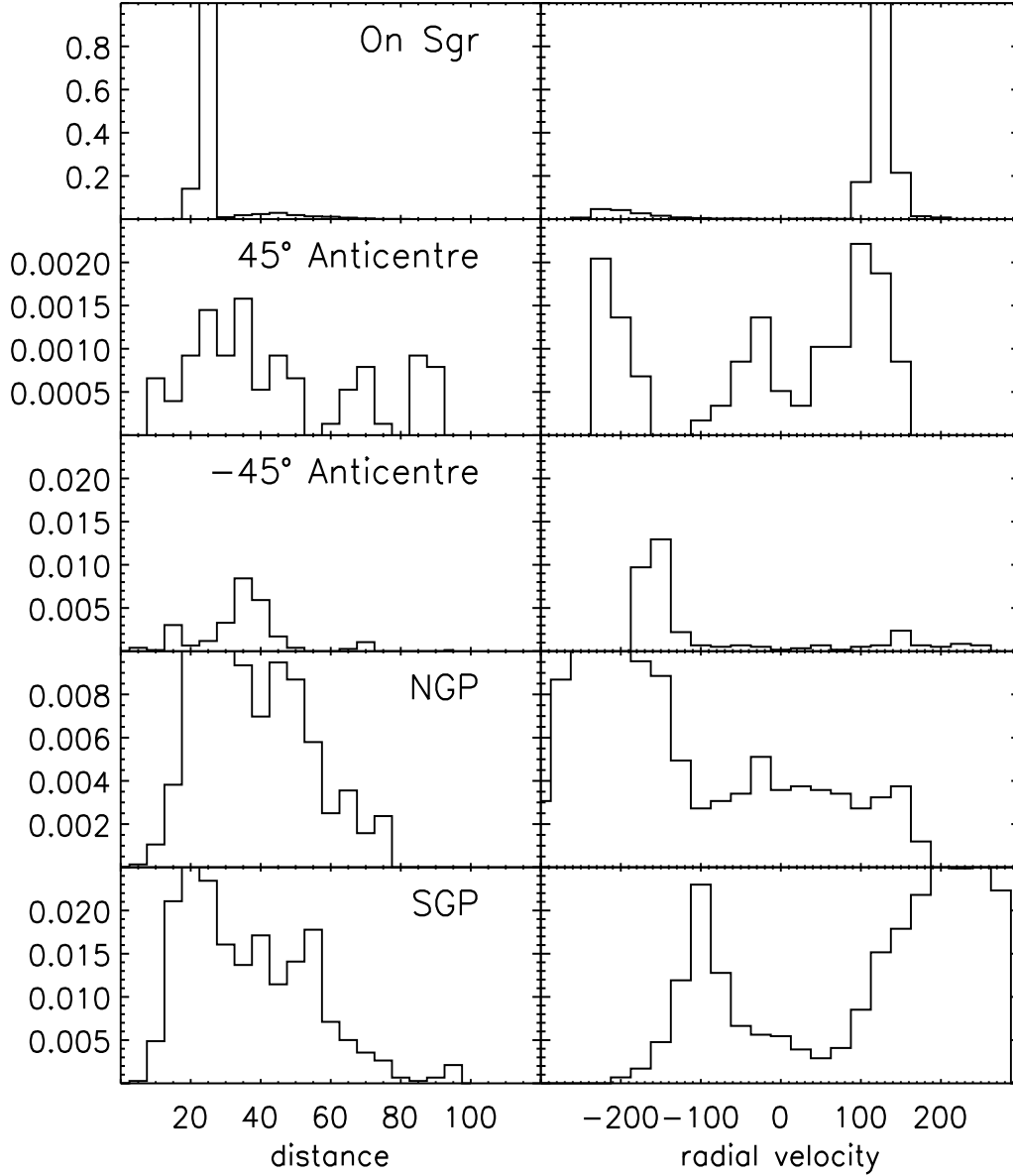


FIGURE 5.7— Number counts in $5 \times 5 \text{ deg}^2$ along various lines of sight normalized to the main body of Sagittarius, which is shown in the top row. Distance bins are 5 kpc, and radial velocity bins are 25 km s^{-1} . All quantities are heliocentric.

Our model can also be used to predict star counts as a function of distance and radial velocity at different points on the sky. This is illustrated in Figure 5.7, where the number counts are normalized to their values on the main body of our simulated Sagittarius, as shown in the first row. We assume fields which are $5^\circ \times 5^\circ$. For the distance, we use 5 kpc bins, whereas for the radial velocity we take 25 km s^{-1} bins. Note that the contrast of structures in the radial velocity counts are generally larger than in the distance counts, indicating that it should be easier to detect streams in velocity space rather than as density inhomogeneities (see also Helmi & White 1999). This is particularly true considering the much greater relative precision of the velocity measurements. Space density enhancements often occur near the orbital turning points; several are seen as sharp features in the central panel of Figure 5.6.

A possible progenitor of Sagittarius could thus be a satellite with a core radius of about 0.56 kpc, a central M/L ratio of 2.25 and a total stellar velocity dispersion of about 18 km s^{-1} . We have found viable models with a wide range of total luminosities and masses, and both with and without extended dark halos. The data available at present only weakly constrain the total initial extent either of the light or of the mass. The observed metallicity data, for example, are consistent with an initial galaxy similar to either of our detailed models, both of which would lie within the scatter of the luminosity–size–velocity dispersion–metallicity distribution for more distant dwarf spheroidal galaxies in the Local Group. Thus we see no indication that Sagittarius is in any way anomalous. Further work on the debris streams of Sagittarius is needed to constrain better its initial total luminosity, and to distinguish between purely stellar or dark matter dominated progenitors.

Acknowledgments

We wish to thank Pavel Kroupa and James Binney for comments on an earlier version of this manuscript.

References

- Bergbusch P.A., Vandenberg D.A., 1992, *ApJS*, 81, 163
 Edelson D.J., Elmegreen B.G., 1997, *MNRAS*, 290, 7
 Helmi A., White S.D.M., 1999, *MNRAS*, 307, 495 (Chapter 2)
 Ibata R., Gilmore G., Irwin M., 1994, *Nature*, 370, 194
 Ibata R., Gilmore G., Irwin M., 1995, *MNRAS*, 277, 781 (IG195)
 Ibata R., Wyse R.F.G., Gilmore G., Irwin M., Suntzeff N.B., 1997, *AJ*, 113, 634 (I97)
 Ibata R., Lewis G.F., 1998, *ApJ*, 500, 575
 Jiang I.G., Binney J., 2000, *MNRAS*, in press (astro-ph/9908025)
 Johnston K.V., Spergel D.N., Hernquist L., 1995, *ApJ*, 451, 598
 Johnston K.V., Majewski S.R., Siegel M.H., Kunkel W.E., 1999, *AJ*, 118, 1719
 King I.R., 1966, *AJ*, 71, 64
 Klessen R., Kroupa P., 1998, *ApJ*, 498, 143
 Majewski S.R., Siegel M.H., Kunkel W.E., Reid I.N., Johnston K.V., Thompson I.B., Landolt A.U., Palma C., 1999, *AJ*, 118, 1709
 Mateo M., 1998, *ARA&A*, 36, 435
 Mateo M., Olszewski E.W., Morrison H.L., 1998, *ApJ*, 508, L55
 Velázquez H., White S.D.M., 1995, *MNRAS*, 275, 23L
 White S.D.M., 1983, *ApJ*, 274, 53
 Zaritsky D., White S.D.M., 1988, *MNRAS*, 235, 289
 Zhao H.S., 1998, *ApJ*, 500, L149

*If you can look into the seeds of time,
and say which grain will grow
and which will not,
speak then to me...*

W. Shakespeare: "Macbeth"

Chapter 6

The phase-space structure of a cluster halo: Insights into our Galactic halo*

ABSTRACT

We analyse high resolution simulations of the formation of a cluster in a Λ CDM cosmology. Our goal is to determine how much phase-space substructure is left over from the halos that merged to build up the cluster we observe at the present day. We study debris streams originating in such halos and find that their evolution can be explained simply in terms of the conservation of phase-space density. The properties of these streams are consistent with having a phase-mixing origin. After scaling the cluster to a galactic size halo, we find that the velocity ellipsoid in the equivalent of the "Solar vicinity" is formed by roughly a thousand dark matter streams with velocity dispersions smaller than a few tens of km s^{-1} . These results imply that the small scale structure present in the dark halo velocity ellipsoid will not strongly affect the signal expected by current and future experiments designed to determine the nature of dark matter in our immediate neighbourhood.

*Based on: Amina Helmi, Simon D.M. White and Volker Springel, in preparation

6.1 Introduction

Over the last twenty years, a paradigm has emerged for the formation of structure in the Universe, which may be summarised as follows.

- All structures observed today grew by gravitational amplification of very small initial density fluctuations thought to have been produced during the inflationary expansion of the Universe (Peebles 1974).
- The dominant mass component in the Universe is dark, and in the form of non-baryonic, weakly interacting particles. If at the time of decoupling from the general expansion of the Universe these particles had non-relativistic motions, the dark matter is termed “cold” (CDM). If, on the contrary, they were relativistic at decoupling, they represent “hot dark matter” (HDM). The natural candidate for HDM are massive neutrinos, whereas the CDM particles may be axions, neutralinos, or some variety of more massive particles left over from the early Universe.
- In the case of HDM, structure formation takes place top-down, starting from superclusters onto galaxy scales. The relativistic motions of the particles at decoupling are sufficient to erase the smallest scales fluctuations present right after the inflationary era. In the case of CDM, the first objects to collapse are galaxies, these then merge and form the larger scale structures we observe today. Thus structure formation occurs in a hierarchical fashion (White & Rees 1978).
- This hierarchical formation of structure is the scenario currently favoured, as a remarkable agreement is found when detailed comparisons of models and observations, ranging from galaxy scales onto clusters and superclusters scales are made.
- Galaxies are known to have two components: a baryonic (gas and stars) and a non-baryonic (or dark) component. Most of the mass in a galaxy is contributed by the dark component (90% or more) distributed in a halo leaving only a few percent for the visible mass which, in turn, is concentrated towards the centre of this dark halo. On large scales, the physics is dominated by the dark matter dynamical evolution. On scales of a kiloparsec or so, the dominant processes are star formation, gas cooling, feedback, etc. These processes are not yet fully understood, making the modelling of the formation of a galaxy a very difficult enterprise. Currently numerical simulations cannot address simultaneously both the gravitational dynamics affecting the dark halos and the behaviour of the collisional components because of the large dynamic range in the processes affecting each component at the same time, even though substantial progress has been made along these lines (e.g. Steinmetz & Navarro 1999, Pearce et al. 1999).
- Because of this difficulty, a phenomenological description of the evolution of the baryonic components has been put forward in the past few years (White & Frenk 1991). This prescription may either be combined with analytic approximations for the evolution of the dark matter component (Baugh, Cole & Frenk 1996; Kauffmann 1996) or with high-resolution simulations of the dark matter evolution in the Universe (Kauffmann et al. 1999). The prescription is based on a set of simple yet physically motivated equations describing the cooling of gas, transformation of gas in stars, etc.

An interesting question that may be formulated and answered in this general hierarchical scenario is what it predicts for a typical galaxy like our Milky Way. With this idea in mind,

several groups (Moore et al. 1999; Klypin et al. 1999) have performed high resolution simulations of a galactic size halo in CDM cosmologies. They found that the dark matter satellites outnumber the known population of satellite galaxies in the Local Group by a factor of ten. It is still unclear how to interpret this result, in particular in view of the uncertainties involved in transforming mass into light. Some attempts have been made to reconcile the disagreement, both by changing the initial power spectrum of density fluctuations (Kamionkowski & Liddle 2000) and by taking into account the effects of a reionising background which may inhibit star formation in the smallest mass halos (Bullock, Kravtsov & Weinberg 2000).

Broadly speaking, the hierarchical paradigm would predict that the Milky Way formed through mergers of smaller galaxies. These galaxies would contribute to the dark halo, the spheroid (bulge and stellar halo) and to the Galactic gas reservoir. It may be very difficult to determine the relative gas contribution of these progenitor galaxies, since gas “easily forgets” its site of origin. However, if the dynamical mixing timescales are sufficiently long (i.e. longer than the age of the Universe) it may be possible to “break-up” the Galactic spheroid (stars and may be even dark matter particles) into coherent structures in phase-space directly related to the galaxies that merged to form the Milky Way we observe today.

A first attempt at determining whether the merging history of the Milky Way may be imprinted in the phase-space structure of halo stars, and thus be recovered, was made by Helmi & White (1999; hereafter HW and Chapter 2). They studied the infall of satellites onto a fixed Galactic potential, and the evolution of their debris in phase-space. They found that after 10 Gyr stars having a common origin are distributed smoothly in space, and appear very clumped in velocity space, where they define streams with very small velocity dispersions. The expected number of such streams scales with the initial size r , velocity dispersion σ and orbital period P , of the disrupted object:

$$N_{\text{stream}} \sim 10 \left(\frac{r}{1 \text{ kpc}} \right)^2 \left(\frac{\sigma}{15 \text{ km s}^{-1}} \right) \left(\frac{P}{0.23 \text{ Gyr}} \right)^{-3}. \quad (6.1)$$

The total number of stars associated with the object is $N_* \propto r \sigma^2$ (from the virial theorem) while the volume V over which they are spread scales with the cube of the size of the orbit, and so approximately as P^3 . Hence the number of stars per stream in the Solar neighbourhood scales as $N_*/V N_{\text{stream}} \sim \sigma/r$; objects with large initial velocity dispersion and small initial size should produce the most easily detectable streams with little dependence on initial period. This implies that the Solar neighbourhood velocity ellipsoid should be formed by 300 – 500 kinematically coherent structures, originated in past merger and accretion events. Furthermore, very recently the first set of halo streams that could be directly linked to a disrupted satellite (Helmi et al. 1999) were detected in the Solar neighbourhood. The progenitor of the two streams found was probably similar to the dwarf galaxy Fornax, a present satellite of the Milky Way.

One of the weak points of the HW analysis which deserves further investigation, is the assumption of a fixed, smooth potential onto which galaxies are accreted. In the fully hierarchical regime, the concept of a background galaxy is meaningless, in the case of a spiral galaxy probably until $z \sim 2 - 3$. Moreover, the potential may be varying very violently, with large numbers of clumps orbiting the centre of what will become a virialised Galactic halo.

One of the goals of the present paper is to determine to what extent the previous analysis may have been correct or not, and also understand why this may be the case. We also wish to address here how the phase-space substructure present in the dark matter halo, which may be expected in view of the previous discussion, will affect on-going and future experiments which try to find out the nature of dark matter, and measure the mass of its constituent particles (e.g.

TABLE 6.1— Parameters of the simulations for the high-resolution region. The gravitational softening ϵ was kept fixed at the given values for redshift below $z = 9$ in physical coordinates, and above that redshift in comoving coordinates (Springel, 2000).

simulation	S1	S2	S3	S4
particle mass $m_p [h^{-1} M_\odot]$	6.87×10^9	1.36×10^9	2.38×10^8	4.68×10^7
initial redshift z_{start}	30	50	80	140
gravitational softening $\epsilon [h^{-1} \text{ kpc}]$	6.0	3.0	1.4	0.7
number of particles N_p	450088	1999978	12999878	66000725

HEAT collaboration: Barwick et al. 1998). We will tackle these problems by analysing a high-resolution simulation of the formation of a cluster in a Λ CDM cosmology scaled down to a galactic size halo (Springel, White & Tormen 2000).

6.2 Methodology

6.2.1 The simulations

The simulations we analyse here were carried out using a parallel tree-code (Springel, Yoshida & White 2000) on the Cray T3E at the Garching Computing Centre of the Max Planck Society. These simulations were generated by zooming in and resimulating with higher resolution a particular cluster and its surroundings formed in a cosmological simulation (as in Tormen, Bouchet & White 1997). The Λ CDM cosmological simulation has parameters $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.7$ and $\sigma_8 = 0.9$. The cluster selected is the second most massive cluster in the simulation and has a virial mass of $8.4 \times 10^{14} h^{-1} M_\odot$. The particles that end up in the final cluster of the cosmological simulation and in its immediate surroundings (defined by a comoving sphere of $70 h^{-1} \text{ Mpc}$ radius) were traced back to their Lagrangian region in the initial conditions for resimulation. The initial mass distribution between 21 and $70 h^{-1} \text{ Mpc}$ was represented by 3×10^6 particles. In the inner region, where the original simulation had 2.2×10^5 particles, new initial conditions were created for 4.5×10^5 , 2×10^6 , 1.3×10^7 and 6.6×10^7 particles, and small scale power was also added onto this volume. The original force softening was also decreased to obtain better spatial resolution. All four simulations were run from very high redshift until $z = 0$. A summary of their properties is given in Table 6.1.

6.2.2 The cosmology

In a Λ CDM cosmology the scale factor of the Universe in the matter dominated era evolves as

$$\frac{a(t)}{a_0} = \left(\frac{\Omega_0}{\Omega_\Lambda} \right)^{1/3} \left[\sinh \left(\frac{3\sqrt{\Omega_\Lambda} H_0 t}{2} \right) \right]^{2/3}, \quad (6.2)$$

where Ω_0 (Ω_Λ) is the present-day matter (vacuum) density of the Universe in units of the critical density today, H_0 is the Hubble constant which we take to be $70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and a_0 is the scale factor today (generally set to unity). We may invert the previous equation to determine the age of the Universe at redshift z

$$t = \frac{2}{3H_0\sqrt{\Omega_\Lambda}} \ln \left[\sqrt{\frac{\Omega_\Lambda}{\Omega_0} \frac{1}{(1+z)^3}} + \sqrt{1 + \frac{\Omega_\Lambda}{\Omega_0} \frac{1}{(1+z)^3}} \right]. \quad (6.3)$$

The Hubble parameter H is also a function of time

$$H(t) = H_0 \sqrt{\Omega_\Lambda + \Omega_0(1+z)^3}. \quad (6.4)$$

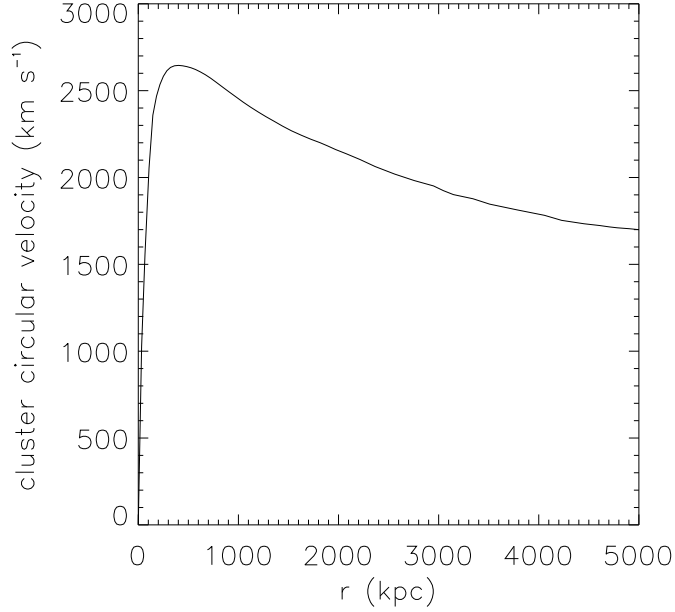


FIGURE 6.1— The circular velocity $V_c(r) = \sqrt{GM(r)/r}$ of the cluster for our simulation S3.

The position of a particle may be expressed in physical coordinates in terms of its value in the comoving frame as

$$\mathbf{r} = a(t) \mathbf{x}, \quad (6.5)$$

and the velocity in terms of the peculiar velocity $\mathbf{u}$ as

$$\mathbf{v} = H(t) \mathbf{r} + \mathbf{u}. \quad (6.6)$$

6.2.3 Scaling to a Milky Way halo

The simulation described above is currently the only one with such high resolution and is thus (almost) perfect to study the phase-space evolution of debris from disrupted halos that end up forming the cluster dark halo at the present time. Because we are not only interested in phase-space evolution but we would also like to know particular properties of a galactic halo, such as the number of streams one might expect in the Solar neighbourhood, we scale the cluster simulation to a galactic size halo.

In Figure 6.1 we show the circular velocity profile of the cluster at $z = 0$. We determine the cluster centre by successively refining a mesh located on the cluster, and determining the cell containing the largest number of particles. This process is repeated until the largest number of particles in a given (now small size) cell is sufficiently small to determine by simple counts which particle has the largest number of neighbours. Its position is defined as the cluster centre. This determination is robust against changes in the mesh shape and size, and we estimate the error in the final position of the centre of the cluster to be of the order of 2 kpc for our simulation S3. The circular velocity is then determined after spherically averaging the mass distribution around the centre of the cluster and is derived from $V_c(r) = \sqrt{GM(r)/r}$.

We scale the cluster to represent a “Milky Way” halo by scaling the circular velocity so that

at its maximum it is equal to 220 km s^{-1} . The scaling factor

$$\frac{V_c^{\text{cl}}}{V_c^{\text{MW}}} = \frac{r_{\text{vir}}^{\text{cl}}}{r_{\text{vir}}^{\text{MW}}} \sim 12. \quad (6.7)$$

The virial radius of our Milky Way dark matter halo is about 260 kpc. The validity of this scaling relies both on theoretical and numerical results (Lacey & Cole 1993; Moore et al. 1999; see also Jing & Suto 2000). This should be a good approximation to the dark matter evolution of a typical galactic halo, except possibly for the fact that assembly is shifted towards later times.

6.3 Results

6.3.1 Phase-space evolution of halo debris

To study the phase-space evolution of disrupted halos during the formation of the cluster, we proceed by identifying halos at high-redshift which are directly accreted onto the most massive halo at the time of identification.

We identify halos at each output time using a Friends-of-Friends (FOF) algorithm, which links particles within spheres of radius 0.2 in units of the mean interparticle separation. In this way we can construct a fairly detailed merger history of the formation of the cluster. At each redshift we identify the most massive halo, which may be considered to be the cluster main progenitor at that redshift. We say that a halo identified at redshift z will be directly accreted onto the cluster main progenitor at z' (the redshift of the next simulation output) if at least one of its particles has become part of the cluster main progenitor at z' .

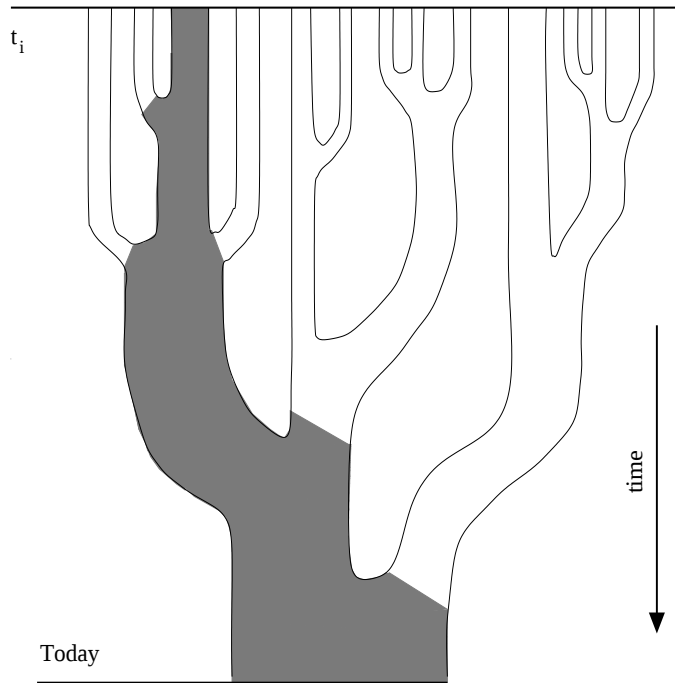


FIGURE 6.2— A schematic representation of the merger history of the cluster. The dark trunk at any given time represents the cluster main progenitor, and the branches linked to this trunk correspond to the halos we analyse in the text (based on Lacey & Cole, 1993).

This is like following the (thickest) trunk of the merger tree, which would correspond to the cluster main progenitor, and studying what happens to halos which fall along the tree branches

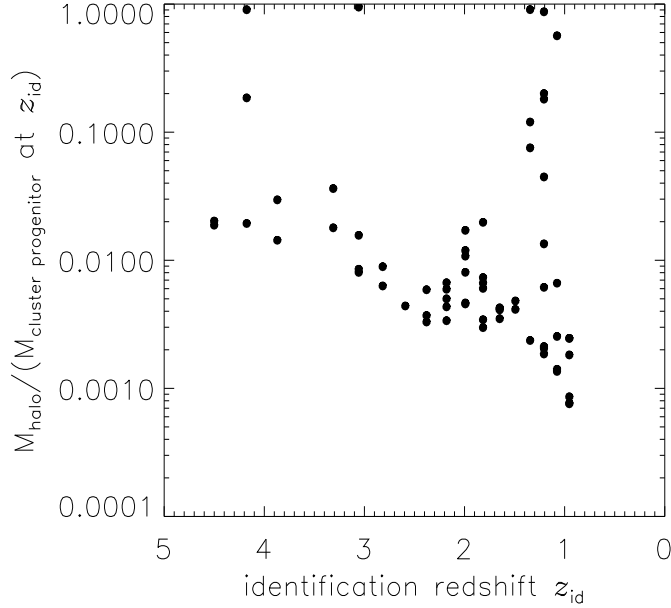


FIGURE 6.3— We show here the mass fraction of the halos directly accreted with respect to the mass of the cluster main progenitor at the time of identification. Note that in some cases this mass ratio approaches one, meaning that the cluster is experiencing a major merger. These cases may be considered to be dramatically different from the static infall of satellites discussed by HW.

at the given output time. This idea is illustrated in Figure 6.2. We have selected halos in this way since a direct merging onto the cluster main progenitor resembles more closely the HW static case: i.e. a halo falling directly onto a given potential. We should keep in mind that the results derived for the evolution of halos directly merging with the cluster are likely to be valid for all halos that end up forming the cluster, since halo–halo mergings inside the cluster virial radius are not very common (Springel 2000).

For the intermediate resolution simulation S3 we have identified 66 halos with at least 1000 particles, which fall onto the cluster main progenitor in the redshift range $z = 4.5 - 0.84$. The halos have a large spread in sizes as shown in Figure 6.3.

To determine the phase-space evolution of streams leftover from disrupted halos, we proceed by defining, for each halo identified before, a “reference” particle. We follow this particle in time to study the properties of the velocity ellipsoid and density in its immediate neighbourhood. This is equivalent to following the evolution of the properties of the stream in which this particle resides.

The choice of the reference particle is only subject to the condition that it should be located initially in a region of the halo with a large density. At the time of identification of the halo, we measure the space density and velocity ellipsoid in a cubic volume of 1 kpc on a side (83 pc for the Milky Way halo) around the reference particle. Then, for each output, we derive a local density and velocity ellipsoid from the halo particles in the same orbital phase as the reference centre. When the density has decreased by more than a factor of 3 in a ten times larger volume (i.e. 10 kpc, or 830 pc in scaled units), we proceed to follow the reference centre in this larger volume. This particular time, which we call t_i , can be interpreted as the “time of the formation of the stream”, that is, when the dynamics begin to be dictated by the cluster field rather than

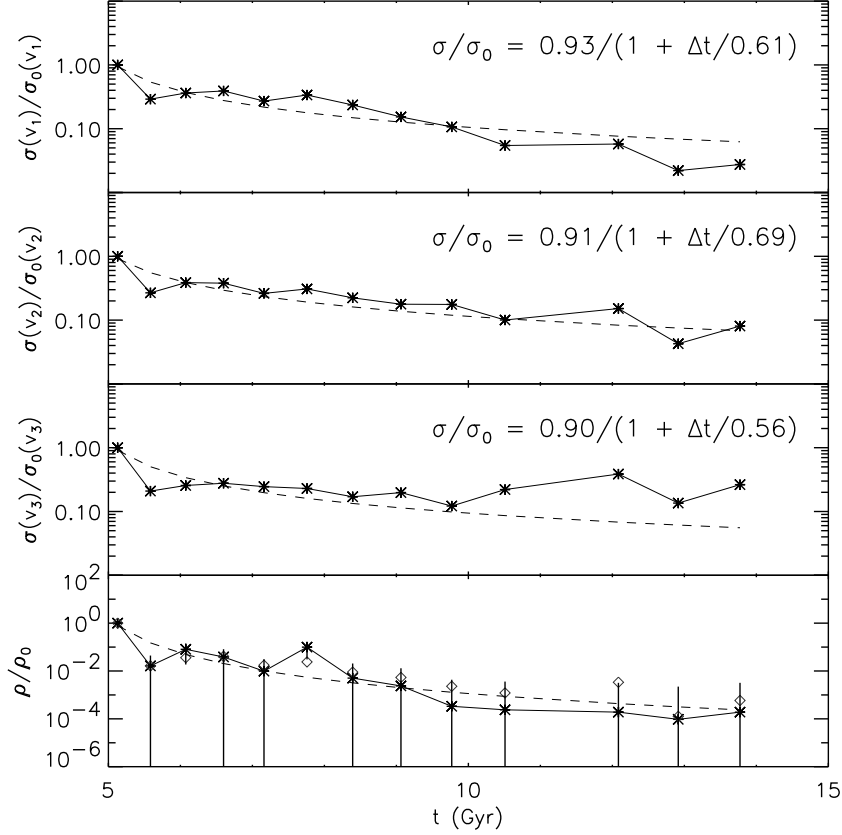


FIGURE 6.4— The phase-space behaviour of a debris stream in our S3 simulation. The top three panels show the time evolution of the velocity dispersions in the principal axes frame starting from the time the halo falls onto the main progenitor of the cluster at redshift 1.2. In the bottom panel we plot the evolution of the density along the stream. All quantities are derived from particles located inside a cubic volume of 10 kpc on a side and with the same orbital phase as the halo reference particle. The dashed curves correspond to the fits indicated in each of the panels. For comparison with the behaviour in a static potential see Figure 6.5.

by that of the halo. The enlargement in the size of the volume allows us to study, for a fraction of our halos, the properties of streams until the present day.

In Figure 6.4 we show the evolution of the principal axes of the velocity ellipsoid and the local space density in the neighbourhood of the reference particle, for a halo identified at redshift 1.2. This halo had an initial mass of $1.65 \times 10^{13} M_{\odot}$ (18% of the mass of the cluster progenitor at the identification time) and directly merged with the cluster main progenitor at that redshift. Note that the halo is completely disrupted immediately and the local density decreases dramatically. The velocity dispersions along the principal axes are also found to decrease very rapidly. The dashed curve in each of the three top panels represents a fit to the velocity dispersion of the form: $\sigma/\sigma_0 = A/(1 + \Delta t/t_0)$, where $\Delta t = t - t_i$, and t_i is the “time of the formation of the stream”. In the bottom panel, the solid curve corresponds to the density, determined from number counts at each output. The dashed curve is the product of the fits to the velocity dispersions described above. As expected from the conservation of phase-space density, this dashed curve traces very well the behaviour of the local space density of the stream.

We are thus finding that the phase-space behaviour of halo debris appears to be fairly similar to the static case. A halo with a given initial velocity dispersion gives rise to streams with velocity dispersions ten times smaller after 10 Gyr of evolution in the cluster potential. This is the same type of scaling HW had found for the case of infall of satellites onto a fixed static po-

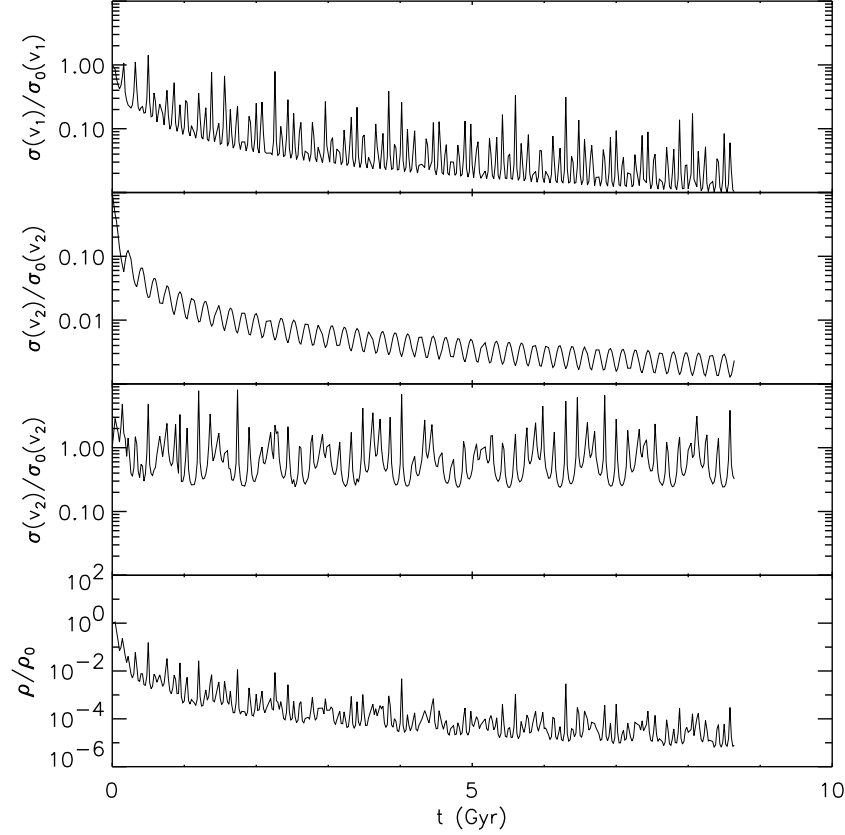


FIGURE 6.5— The phase-space behaviour of a debris stream in a static NFW potential consistent with the present day properties of the S3 cluster. The top three panels show the time evolution of the velocity dispersions in the principal axes frame. In the bottom panel we plot the evolution of the density along the stream.

tential, even though in the regime that we are probing here the potential changes very violently. We can make this comparison more direct in the following way.

As for many halos in CDM simulations, the cluster halo at the final time can be well fit by a Navarro, Frenk & White (NFW) profile (Navarro, Frenk & White 1996):

$$\rho(r) = \rho_0 \frac{\delta_c r_s^3}{r(r + r_s)^2}, \quad (6.8)$$

where $\rho_0 = 3H_0^2/(8\pi G)$ and δ_c is a function of the concentration of the halo $c = r_{\text{vir}}/r_s$:

$$\delta_c = \frac{200}{3} \frac{c^3}{\log(1 + c) - c/(1 + c)}. \quad (6.9)$$

The potential associated with this density can be obtained by integrating Poisson's equation, and is found to be:

$$\Phi(r) = -\Phi_0 \frac{r_s}{r} \log\left(1 + \frac{r}{r_s}\right). \quad (6.10)$$

Here $\Phi_0 = 3/2 H_0^2 \delta_c r_s$. In this (spherically averaged) potential we can integrate the orbit of the reference particle backwards in time, starting from its current position and velocity until the time of infall of the halo (or the “time of formation of the stream”). At this time, we may approximate the phase-space density around this particle by a multivariate Gaussian distribution.

This is possible because the multivariate Gaussian is determined from the properties of particles in a volume much smaller than the size of the halo.

We can follow the evolution of streams produced in the spherical and static NFW potential of the cluster using the approach developed by HW. The basic idea here is to map the initial system onto action-angle space, then follow the much simpler evolution in this space, and finally transform back *locally* onto observable coordinates (all these being linear transformations; for details see HW). In Figure 6.5 we show the evolution of the velocity ellipsoid and space density along a stream that moves in this spherical static NFW potential. When comparing with the actual behaviour of the stream, shown in Figure 6.4, we find that the true potential is not spherical, since none of the principal axes velocity dispersions remains constant. But because the final orbit is so much confined to the inner halo, the decrease in the density in the spherical case is of the same order of magnitude as that observed in the simulations. This is to say that the particle (or the stream) has made many more revolutions in the spherical halo than in the simulations, so as to produce the same decrease in the space density, or equivalently the same number of streams.

We may understand our results in view of the conservation of phase-space density which holds even when the potential varies violently with time, and depends only on the assumption that the system is collisionless, which clearly is a reasonably good approximation to the formation and evolution of a cluster (Springel 2000).

This characteristic behaviour is observed for all halos which are directly accreted onto the main cluster progenitor and which are completely disrupted at some point while orbiting the cluster. This also appears to be the case for the evolution of streams originating in the second most massive halo of the simulation that merges with the cluster main progenitor at redshift 1.34. However, because the halo rapidly sinks to the centre by dynamical friction, the determination of the orbital phases of the particles is more uncertain. Streams are thus harder to follow, as the phase-mixing timescales become extremely short. On the other hand, some of the halos identified manage to keep a bound core, and thus their phase-space evolution is not so dramatic: they do lose a fair amount of mass because of the cluster tidal field but their velocity ellipsoids do not vary much with time. This is particularly so when the reference particle is located right at the centre of the satellite halo.

6.3.2 The Solar neighbourhood

In Figure 6.6 we plot the velocities of particles inside a cubic volume of 2 kpc on a side, located at 8 kpc from the Milky Way centre (all in units of the scaled halo). Note that we did not determine the principal planes of the simulated dark halo, so that the only criterion we have for determining the “Sun’s” position in the halo is its distance to the “Galactic” centre.

In Figure 6.7 we show the velocities of the particles in the same volume as before, but now we identify with different colours particles that belonged to the same halo at redshift 2.38. At this redshift, which corresponds to 11 Gyr ago, our FOF algorithm finds 460 halos with at least 1000 particles for S3. In the box shown in this figure, there are 248 particles. Fifteen different halos contribute to this volume, but only seven with more than one particle. We generally expect particles of the same colour to be clustered in streams. Some such substructures are observable, but it is also apparent that the number of particles inside this box may be too small to populate each expected stream, thus making it difficult to distinguish streams by simple eye inspection.

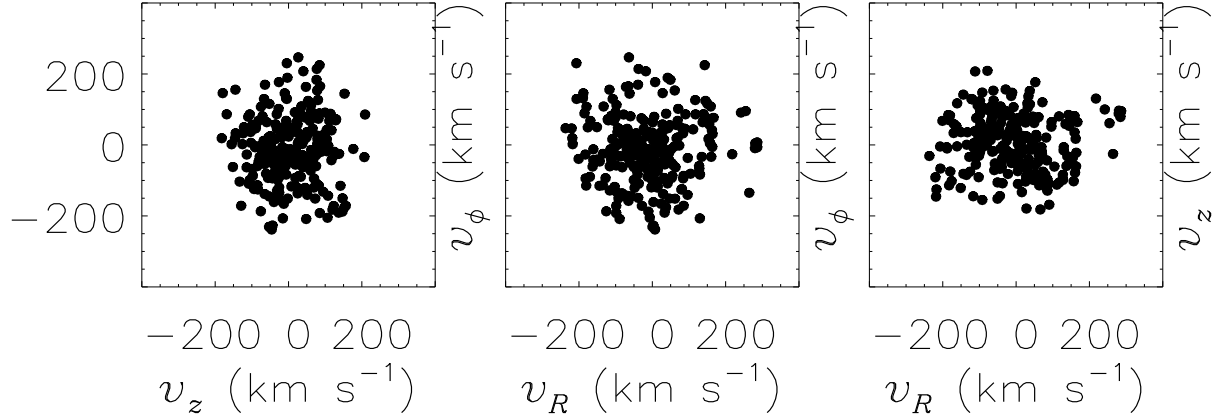


FIGURE 6.6— Principal axes projections of the velocity of particles located in a box of 2 kpc on a side at the Solar radius, where all quantities have been scaled to the “Milky Way halo” as described in the text. The distribution looks slightly clumped.

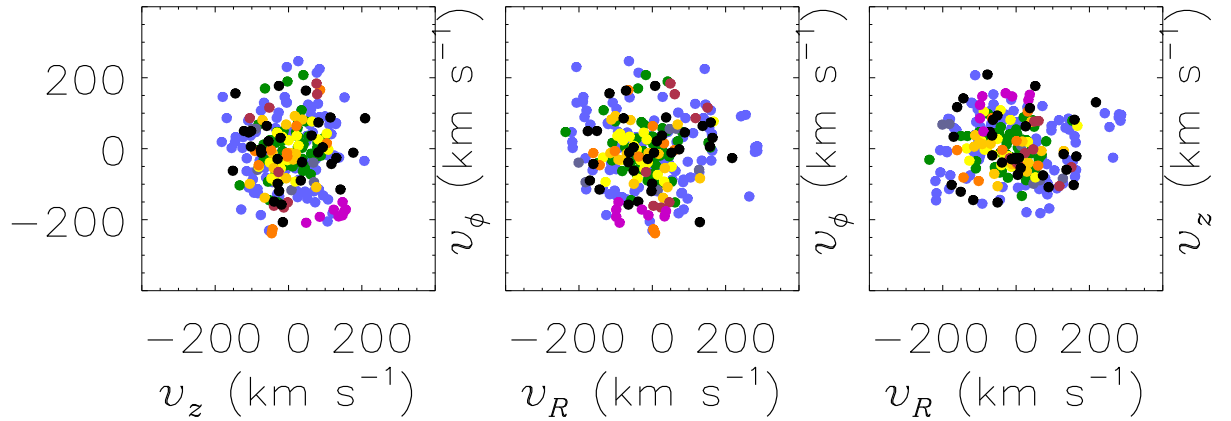


FIGURE 6.7— Same as Figure 6.6 but colours here indicate particles originating in the same halo 11 Gyr ago. Blue dots correspond to field particles (i.e. not belonging to a halo with at least 1000 particles 11 Gyr ago), and black dots to particles from halos which do not contribute substantially to this volume.

6.3.3 The filling factor

A relevant quantity is the number of streams we should expect to find in the Solar neighbourhood. This is important not only to characterize the kinematic distribution of halo stars, but in particular and probably more appropriate in this context, to understand the effect of velocity substructure on the different experiments which attempt to determine the nature of dark matter particles.

We compute the number of streams inside cubic volumes located in the inner $([-20, 20])^3$ kpc³ in the scaled Milky Way halo for our highest resolution simulation S4 as follows.

A stream is defined by particles with the same orbital phase and coming from the same halo at some initial (high) redshift. The orbital phase of a particle may be determined, in principle, by counting how many times it has crossed the planes $x = 0, y = 0, z = 0$. However, because of the limited number of outputs which are logarithmically spaced in the expansion factor, this determination does not yield completely correct answers. The logarithmic spacing of the outputs means, for example, that the time intervals range from $\Delta t = 0.14$ Gyr at $z = 4$, $\Delta t = 0.52$ Gyr at $z = 1$ to $\Delta t = 0.86$ Gyr at $z = 0$. It is therefore difficult to determine correctly whether two particles are at the same phase in an automated way, in particular towards low redshifts since the orbital timescales for stars in the Solar neighbourhood are of the order of 0.25 Gyr only. This is why we also keep track, not just of the crossing of the planes $x = 0, y = 0, z = 0$, but also of $v_x = 0, v_y = 0, v_z = 0$. We define the phase of a particle in the x -direction at time $t + \Delta t$ as:

$$\varphi_x(t + \Delta t) = \varphi_x(t) + \begin{cases} 1 & \text{if } v_x \text{ has changed sign but not } x, \\ 2 & \text{if } x \text{ and } v_x \text{ have changed sign,} \\ 3 & \text{if } x \text{ has changed sign but not } v_x, \end{cases} \quad (6.11)$$

and similarly for the y and z -directions. In this way we keep all the information about the orbital phase of a particle. This is like breaking up the orbit in 4 pieces: from the principal plane to a turning point, back from the turning point to the principal plane, to a second turning point and back to the principal plane. Clearly for the first few outputs such a scheme would break up true streams into components. We thus map back all this phase information by, at the final time t_0 , defining the final phase of a particle as

$$\phi_f = \text{int} \left[\frac{\varphi(t_0)}{4} \right] + 1 \quad (6.12)$$

Thus now we may say that particles with the same phase ϕ_f and coming from the same halo at some high redshift, which we fix here to be $z = 2.38$, are on the same stream. If we find L separate streams in a given box, each containing N_k particles ($k = 1, \dots, L$) then the mass-weighted mean number of particles per stream is

$$\langle N \rangle = \frac{\sum_k^L N_k^2}{\sum_k^L N_k}. \quad (6.13)$$

The mass-weighted number of streams, or the filling factor in the box is then

$$F = \frac{\sum_k^L N_k}{\langle N \rangle}. \quad (6.14)$$

In Figure 6.8 we plot the filling factor as a function of distance from the Galactic centre for each one of the boxes considered. The thick grey line is the median filling factor, derived from boxes

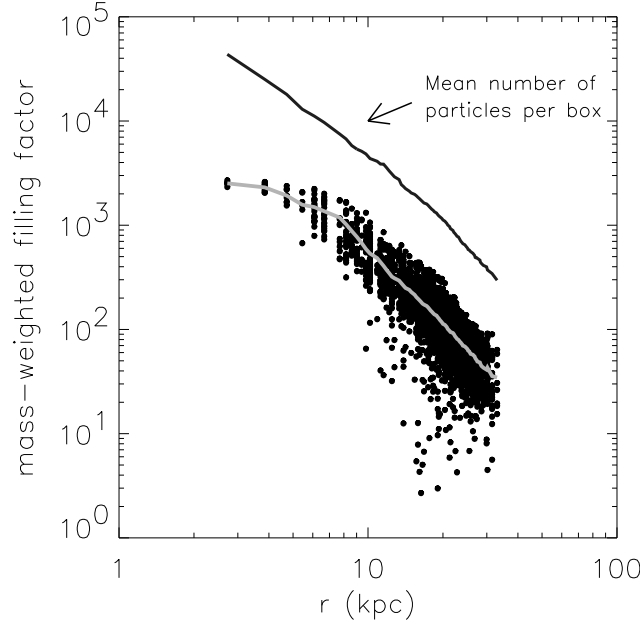


FIGURE 6.8— The mass-weighted number of streams in boxes of 2.72 kpc on a side for our simulation S4 as a function of distance from the centre of the Galactic halo. The thick grey line going through the data points corresponds to the median mass-weighted number of streams, derived for boxes at the same distance from the Galactic centre. The dark thick line corresponds to the average number of particles inside the same volumes. In such, relatively large, boxes, and only for the highest resolution simulation do we find enough particles to populate all streams with at least two particles. The average number of *dark matter* streams in the Solar neighbourhood is thus of the order of a thousand.

located at the same distance from the Galactic centre. At the position of the Sun we find of the order of a thousand streams.

Another interesting quantity which we can calculate for each box is the characteristic mass-weighted dispersion within a stream in units of the underlying 3-D velocity dispersion:

$$\sigma_{\text{stream}} = \frac{1}{\sqrt{N_D}\sigma} \sqrt{\sum_k^L \frac{\sum_{ij} |\mathbf{v}_i - \mathbf{v}_j|^2}{N_k - 1}} \quad (6.15)$$

where the ij run over all particle pairs in the k -th stream in a given box, N_k is the number of particles in this stream in this box, and σ is the 3-D velocity dispersion for all the individual particles in the box. This fraction is expected to be relatively small in view of our results of the phase-space evolution of halo debris. We plot this quantity in the left panel of Figure 6.9 as a function of distance from the Galactic centre for each one of the boxes considered. The velocity dispersions inside a stream are small, but still appear to be between 10% and 30% of the overall velocity dispersion inside the box. In the right panel of Figure 6.9, we give the characteristic mass-weighted velocity dispersion in km s^{-1} . The values of $20 - 30 \text{ km s}^{-1}$ are partly due to gradients inside the box for particles on the same stream. This effect is likely to be even stronger for the boxes located very close to the centre of the Galaxy, as observed. But we should also bear in mind that here the orbital timescales are so short that our scheme for determining the particles phases may not give the correct answer. A hint that this may be happening is the turn over at roughly 7 kpc in the mass-weighted filling factor, as shown in Figure 6.8.

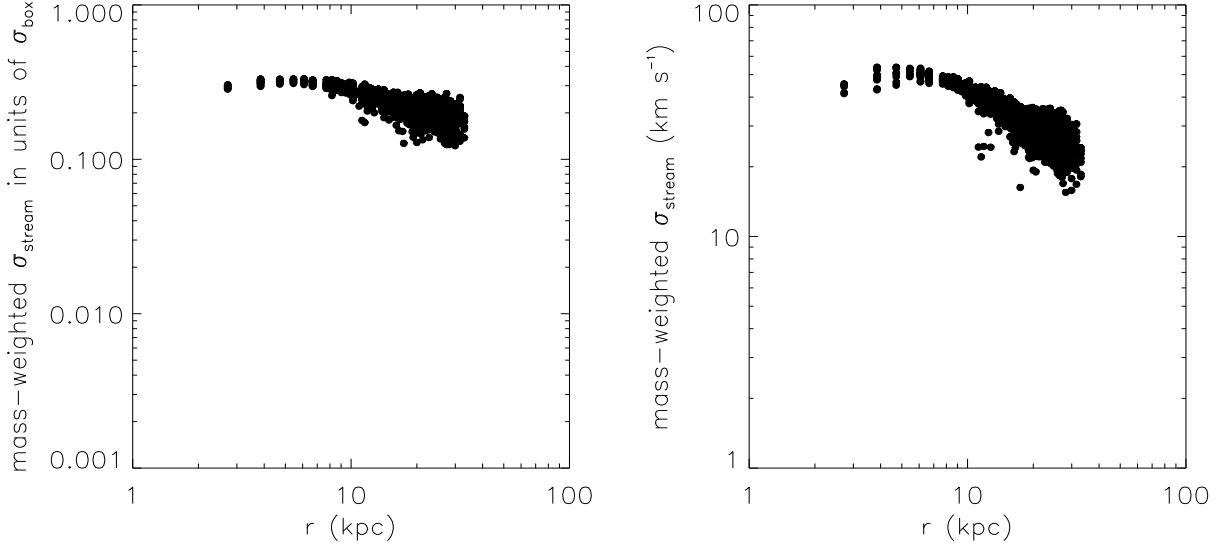


FIGURE 6.9— Mass-weighted velocity dispersions of streams inside cubic volumes of 2.72 kpc on a side, located in the inner $([-20, +20])^3 \text{ kpc}^3$ of the Galactic halo. In the left panel, the velocity dispersion is expressed in units of the 3D velocity dispersion in each particular cubic volume. In the right panel it is given in km s^{-1} .

If instead of taking boxes of 2.72 kpc on a side, we take boxes 8 times smaller (thus 1.36 kpc on a side), we will not be able to populate all streams, as the average number of particles inside such boxes is slightly less than a thousand. We thus only focus on boxes having at least ten times more particles than streams. The streams observed will be well-populated and thus enable us to obtain a fair estimate of the velocity dispersion of a stream. In Figure 6.10 we show the number of particles in boxes of 1.36 kpc on a side, containing a ten times smaller mass-weighted number of streams (left panel); in the central panel the mass-weighted velocity dispersion of streams in units of the 3D velocity dispersion in each particular volume is plotted. Finally the right panel corresponds to the mass-weighted velocity dispersions of streams in km s^{-1} .

6.4 Discussion

We have studied the phase-space evolution of different halos that merged to build up a cluster in a ΛCDM cosmology. Our analysis has shown that the debris streams originating in halos of different sizes and orbital characteristics all behave in a similar way: with velocity dispersions decreasing in time, and with local space density decreasing as $1/t^3$. The debris streams that we were able to follow until the present time, are consistent with having a phase-mixing origin. For halos that we could not follow for a very long time, because of their smaller initial number of particles or their shorter orbital timescales – which make mixing take place more quickly – we also find the same behaviour. On the small scales of the fine-grained distribution function we find that the dominant mixing mechanism is phase-mixing as opposed to chaotic mixing. The phase-space evolution appears to be very ordered and quite simple.

The simple behaviour of the debris streams can be understood in terms of the conservation of phase-space density. The system is collisionless and therefore, independently of how violent the potential changes in time, the streams will behave only guided by this conservation law. If chaos had been important, the behaviour of the streams would still have been dictated by Liouville's theorem, but the mixing would have taken place on a much shorter timescale, so

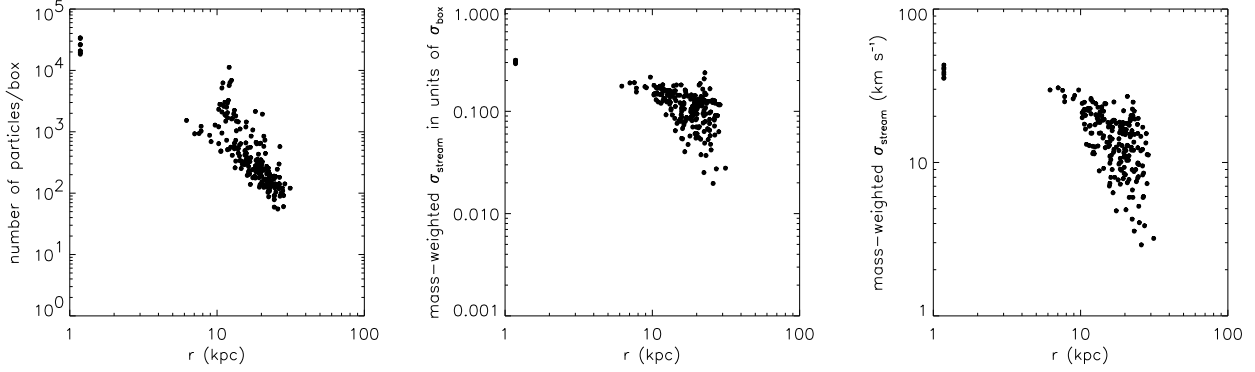


FIGURE 6.10— The number of particles in boxes of 1.36 kpc on a side, containing a ten times smaller mass-weighted number of streams (left panel). In the central panel the mass-weighted velocity dispersion of streams in units of the 3D velocity dispersion in each particular volume is plotted. Finally in the right panel, the mass-weighted velocity dispersion of streams in km s^{-1} is shown. The streams have a range of velocity dispersions that go from a few km s^{-1} to a few tens of km s^{-1} , and there is also an indication that there is a dependence on Galactocentric distance.

that all streams observed (independently of the number of particles available in the volume, i.e. of the resolution) would have been one-particle streams. The probably surprising result is that we have a finite number of massive streams, and as we explicitly show in Section 6.3.1, whose properties are not very different from those observed in static idealised potentials.

While violent relaxation takes place on the large scales, and is so effective as to produce a universal density profile for dark matter halos (Navarro, Frenk & White 1996), independent of their merger history and only weakly dependent on the cosmological initial conditions, on the small scales the halo retains memory of its origin and formation history, even after a Hubble time. It will be an interesting exercise to understand the link between the small (fine-grained) and the large (coarse-grained) scales.

Thus a cluster halo formed in a Λ CDM cosmology is not a smooth entity. Not only do dark matter halos contain a large number of dark satellites, dark matter halos have large amounts of substructure in the forms of streams. If we accept that we may scale a cluster halo down to a Galactic halo simply by the ratio of their maximum circular velocities (equivalent to a scaling their mass ratio to the cube), we find that the Galactic dark matter halo in the Solar neighbourhood appears to be clumped in roughly a thousand streams. These streams have their origin in the different halos that merged to form to our dark matter halo. The scaling used above is independent of the formation time of the halo, which thus means that one may expect a slightly larger number of streams since a galactic halo will have had more time to relax, and the streams to spread out in space. We estimate that this number cannot be more than a factor $(\Delta t_i^{\text{MW}})^3 / (\Delta t_i^{\text{cl}})^3 \sim 2.35$ larger, if the time elapsed since the formation of the Milky Way is $\Delta t_i^{\text{MW}} \sim 10.5 \text{ Gyr}$ ($z = 2$), and since the formation of the cluster is $\Delta t_i^{\text{cl}} \sim 7.9 \text{ Gyr}$ ($z = 1$).

Determining the characteristics of these roughly thousand streams is difficult from these high-resolution simulations, because of the number of particles and the small number of outputs available. In a box with a thousand particles, such as some of the ones shown in Figure 6.10, for example, we find on average just ten streams with more than one particle, and typically with only two particles! Even in boxes of 1.36 kpc on a side, the velocity gradient along the orbit of each stream may be of the order of $400 \text{ km s}^{-1} / 20 \text{ kpc} \times 1.36 \text{ kpc} \sim 27 \text{ km s}^{-1}$. The dispersion that we will measure will be of this order unless we reach higher resolution, so that boxes can be reduced in size and therefore the gradient can be suppressed to a minimum contribution. Using

our fitting formulae from Figure 6.4, we may estimate that a halo with an initial 1-D velocity dispersion of 70 km s^{-1} should contribute to the Solar neighbourhood streams with internal velocity dispersions of the order of 8 km s^{-1} , after 8 Gyr of evolution. If we now correct for the earlier assembly of a galactic halo, the expected 1-D velocity dispersion is about 6 km s^{-1} .

Our results indicate that dark matter experimentalists may quite safely assume that the distribution of dark matter particles is Gaussian. The large number of streams observed, none of which appears to be largely dominant in any of the volumes located around the Solar sphere (or circle), means that this is a very good approximation to reality. The important effect may only appear in the spatial distribution of the cores that may be left over even in an otherwise smooth sea of dark matter particles. In the analysis of the simulation S2 and S3, of small and intermediate resolution (at least in comparison to S4!) we did not find any prominent spatial substructures. This may change if we go to higher resolutions. An important property of these cores is that, if there at all, they have a negligible mass, and thus will not affect the thinness of the disk of the Milky Way, or other spiral galaxies.

How do these results affect our estimates derived for the stellar halo in Chapter 2 of this thesis? How many star streams should we observe in the Solar neighbourhood? The number of streams observed in the simulations corresponds to dark matter streams. Each dark matter halo will contribute by about a factor of ten more dark streams than stellar ones, since the baryons are clustered in the inner roughly one-tenth of their host halo. Thus, if we break the 2350 dark streams (now scaled to take into account the earlier assembly of a galactic halo) into ten contributors, each one of these will give rise to 23 stellar streams with orbital properties similar to those of the dark halo. However, because the baryons were initially clustered in the centre of their host halo, their orbital properties will not exactly follow those of their dark matter counterparts. A typical star in the Solar neighbourhood has a period of 0.25 Gyr, with an apocentre of about 11 kpc. A typical dark matter particle will have an apocentre of the order of 20 kpc, corresponding to an orbital timescale of the order of 0.3 Gyr. Because the time for dispersal scales as $1/t^3$, this implies that the number of star streams is actually a factor of 2 larger, i.e. each associated galaxy that merged to build up our Milky Way would have given rise to approximately 45 streams, assuming our scalings are roughly right. This estimate brings us close to the result derived in the beginning of this thesis: we should expect about five hundred stellar streams in the Solar neighbourhood.

Acknowledgments

A.H. wishes to thank the hospitality of MPA.

References

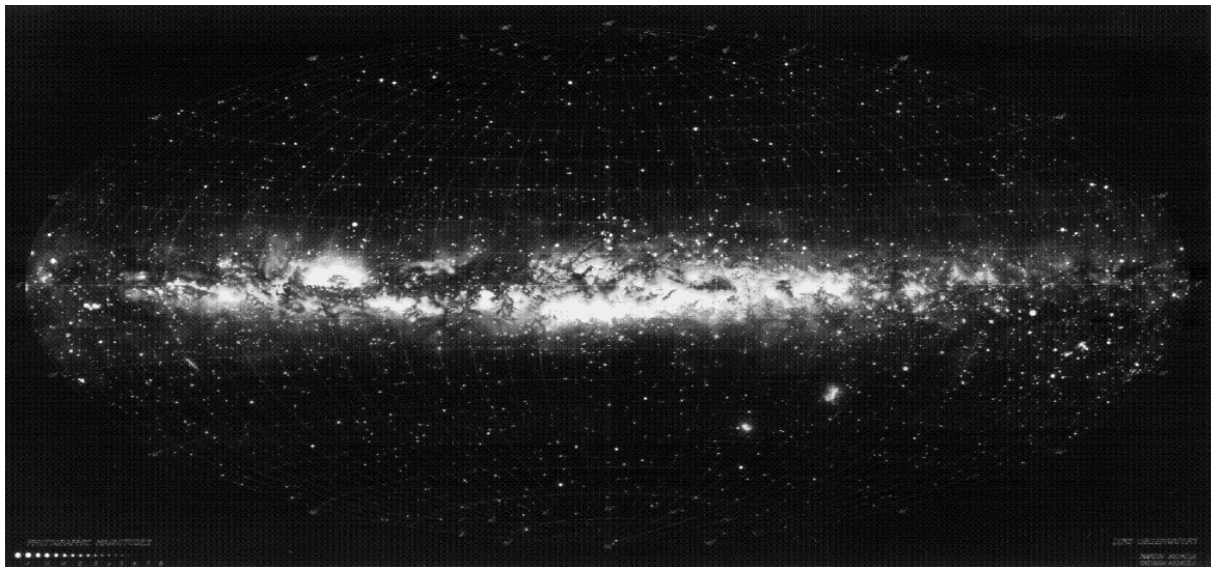
- Barwick S.W. et al. (HEAT collaboration), 1998, *ApJ*, 498, 779
- Baugh C., Cole S., Frenk C.S., 1996, *MNRAS*, 283, 1361
- Bullock J.S., Kravtsov A., Weinberg D.H., 2000, submitted to *ApJ* (astro-ph/0002214)
- Helmi A., White S.D.M., 1999, *MNRAS*, 307, 495 (Chapter 2)
- Helmi A., White S.D.M., de Zeeuw P.T., Zhao H.S., 1999, *Nature*, 402, 53 (Chapter 3)
- Jing Y.P., Suto Y., 2000, *ApJ*, 529, L69
- Kamionkowski M., Liddle A., 2000 (astro-ph/9911103)
- Kauffmann G., 1996, *MNRAS*, 281, 475
- Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, *MNRAS*, 303, 188
- Klypin A., Kravtsov A., Valenzuela O., Prada F., 1999, *ApJ*, 522, 82
- Lacey C., Cole S., 1993, *MNRAS*, 262, 627

- Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, *ApJ*, 524, L19
Navarro J.F., Frenk C.S., White S.D.M., 1996, *ApJ*, 462, 563
Pearce F.R., et al. (The Virgo Consortium), 1999, *ApJ*, 521, L99
Peebles, P.J.E., 1974, *ApJ*, 189, L51
Springel V., 2000, PhD Thesis, University of Munich
Springel V., White S.D.M., Tormen G., 2000, in preparation
Springel V., Yoshida N., White S.D.M., 2000, submitted to *New Astronomy* (astro-ph/0003162)
Tormen G., Bouchet F.R., White S.D.M., 1997, *MNRAS*, 286, 865
Steinmetz M., Navarro J.F., 1999, *ApJ*, 513, 555
White S.D.M., Rees M.J., 1978, *MNRAS*, 183, 341

Summary*

JUST like the Earth does not travel through space alone, but with eight other planets, our Sun does not live on its own but inhabits a conglomerate of a hundred billion stars (other suns) which we call our Galaxy, the Milky Way.

The Milky Way is visible on a clear night as a band of silvery light across the sky, from which it derives its name. This light originates in millions of stars that are too distant to be visible with the naked eye. Even though it looks as though we are outside this milky path, we are in fact situated in a spiral arm close to the edge of this disk. The “milky way” is what one observes looking edge on into the disk.



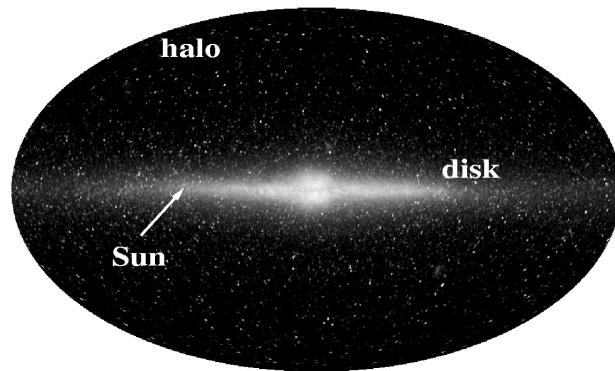
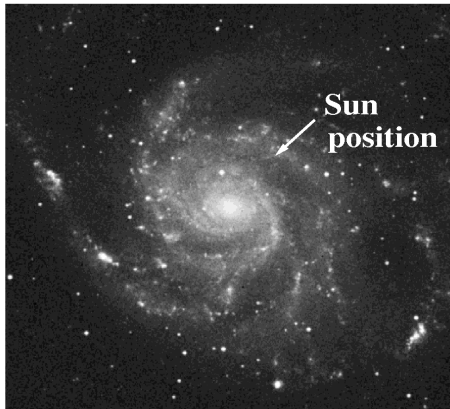
Panoramic view of the sky showing the Milky Way. Individual stars are shown as white dots. The “Milky Way” clouds, actually the combined light of dim, unresolved stars in the densely populated galactic plane, are interrupted by dramatic dark dust lanes. (Lund Observatory, Sweden)

There are hundreds of millions of galaxies like our own throughout the Universe. Not all of them have a spiral pattern, some are elliptical (similar to a rugby ball) and some are very irregular. Galaxies come in different colours and sizes, and a fundamental question astrophysicists try to answer is how did such systems form? What produces this large variety of shapes and colours? How do galaxies evolve?

Going back in time by observing the distant Universe

Let us for a moment return to the Earth and its immediate neighbourhood. Our Sun is about 150,000,000 kilometers away from us. Because light travels through space at 300,000 kilometers per second, it takes eight minutes for the light of the Sun to reach the Earth. One may rephrase this by saying that when we look at the Sun we do not see it as it is now, but as how it was eight minutes ago. One can thus imagine that if you put the Sun very far from the Earth, its light will

*This is a semi-popular summary. It provides a general introduction and an outline of the topics addressed in this thesis. For a more technical summary please refer to Chapter 1.



The left figure shows the galaxy M101, also a spiral system which probably would resemble our Milky Way seen face on. The corresponding location of the Sun is indicated with an arrow. The picture on the right was taken by the COBE satellite and shows the plane of our Galaxy in infrared light. The thin disk of our home spiral galaxy is clearly visible, as are some stars belonging to the halo.

take a very long time to reach us. For example, the closest star to our Sun is about 4 light-years away: its light travels for four years before it reaches the Earth! Similarly, if a galaxy is very distant we will observe it not as it is now, but as it was billions of years ago! This means that by looking at very distant objects, and progressively coming closer and closer, we actually have a sequence in time of how galaxies form and evolve to become what we see in our local Universe. Thus the history of the formation of galactic systems is actually written in the skies.

This idea has been extensively explored observationally. For example, from very deep observations of small regions of the sky taken by the Hubble Space Telescope, we have learned that the more distant galaxies are smaller and that they are often found to be interacting with other galaxies.

Can we understand this observational result? Yes, indeed. Over the last twenty years astrophysicists have been putting together a paradigm of how galaxies form starting from very small matter fluctuations in the early Universe (as early as when the Universe was only three minutes of age!). This standard scenario proposes that small galaxies are the first to form. They then merge given rise to larger and larger systems, like for example our own Milky Way.

This thesis: The Galactic halo

Another way of addressing the problem of how galaxies form is by taking our Milky Way as a landmark. If the Milky Way is a typical galaxy (as we believe) then we can use what we know about it to test the model described above, determine where this model fails and why, or rule it out completely. After all no theory can ever be proven right, all one can try is to prove it wrong (that is the theorist's nightmare and the observer's dream!).

The advantage of using the Milky Way to tackle this problem is that the stars in our Galaxy are close enough that we can actually distinguish them one from the other and determine their individual properties, like how far they are, how quickly they move, in what direction and so on. This is yet unachievable for other galaxies, except may be for a few of the Milky Way satellites. If stars retain memory of where they came from, we may then be able to tell if our Galaxy was built hierarchically, that is, from smaller bits and pieces.

Surrounding the Milky Way is a diffuse roughly spherical cloud of stars known as the Galactic halo. The disk of the Galaxy, in which the Sun orbits, is thought to have formed from the cooling of gas. Its stars have always been part of it: they were born in the disk and that is where

they will also die. But stars in the halo are generally very old and metal-poor. Because the very early Universe was free of heavy elements (these were only later produced inside stars), metal-poor stars may be considered to be more pristine. These are the stars thought to have been formed in galaxies that merged to form our Milky Way. This is why tackling the question of how the Galaxy formed may be equivalent to addressing how its halo formed.

What are the signatures of the Milky Way merging history?

This is the main issue addressed in Chapter 2 of this thesis. By means of computer simulations and analytic calculations (yes! the good old “pen and piece of paper” science!) we found that galaxies destroyed during the Milky Way infancy should still be detectable as ghostly sheets of stars moving on near-parallel orbits. Thus debris from a dwarf galaxy destroyed by a collision with our own Galaxy can survive as relic streams of stars wrapping around the Milky Way. We predict in Chapter 2 that, as a result of the hierarchical build-up of the Milky Way, halo stars in the vicinity of the Sun should be distributed in a few hundred streams, stars in the same stream moving with practically the same speed through space.

Finding fossils: the relic streams

Just like paleontologists hunting for fossils to reconstruct the evolution of species on the Earth, in Chapter 3 we looked for “relic streams” amongst nearby halo stars to understand how our own Galaxy formed and evolved into what we see today.

Between 1989 and 1993 the European Space Agency’s HIPPARCOS satellite measured the positions, distances and motions of 120,000 stars. The resulting catalogue, published in 1997, provides a unique opportunity to search for stellar streams. By sifting carefully through this catalogue, we were able to make a nearly complete census of old giant stars within a few thousand light years of the Sun. Earlier observations by other astronomers had used mountain-top telescopes to show that these stars contain less than a thirtieth as much iron and magnesium as the Sun, a sure sign that they are among the oldest stars in the Milky Way. These observations also measured how fast each star moves away from the Sun. HIPPARCOS, on the other hand, measured changes in the apparent direction of each star, thus establishing how fast it moves across the line-of-sight. Combining these data we determined the direction of motion and the speed of a few hundred nearby giant stars.

Analysing these observations, we were struck by the fact that while most stars move in apparently random directions, two small sets of stars seem to be moving together. One set of nine stars pursue near-parallel orbits which cross the Milky Way’s disk at high speed from North to South; a second set of three stars pursue orbits which cross the disk at the same speed and angle, but from South to North. In both cases the direction and speed of the motion is well separated from those of the other nearby giants. Comparing with our previous computer simulations, we concluded that each set of stars is part of a debris stream, and that both streams come from a single dwarf galaxy which was torn apart by the Milky Way’s gravity during or very soon after the Milky Way’s own formation. The characteristics implied for this dwarf make it very similar to other small galaxies which still survive in the outskirts of the Milky Way system. The stars in each set are distributed all over the sky, even though they move in parallel, showing that the Sun is actually sitting inside the two streams of debris. These ghostly streams are the first fossil evidence showing directly that our own Galaxy was put together from smaller building blocks.

Can we find more fossils?

With a larger sample of old stars we expect to find many more such streams, fossils of cannibalised galaxies. We need, however, to know the motions of a about five thousand halo stars very precisely to be able to constrain the evolutionary history of our Milky Way. This will be possible with future satellite missions, like the American FAME or the European GAIA.

In Chapter 4 we show that these future satellites will measure the motions of stars so accurately that we will finally be able to confirm or rule out the hierarchical model for the formation of our Galaxy. The predicted several hundred streams of stars should be easily detected if they are really there! We can do even better and, for example with GAIA we could identify 50% of all the different galaxies that we think merged to build up the Milky Way.

A living fossil: the Sagittarius dwarf galaxy

Mergers are not a thing of the past. For evidence of the contrary, one only needs to look at the closest satellite galaxy of the Milky Way: the Sagittarius dwarf. Sagittarius was only discovered in 1994, and shows clear signs of being torn apart by the Milky Way tidal forces just now. It is very elongated and not very bright, and it will probably have disappeared the next time it comes close to the centre of the Milky Way. As it is torn apart, its stars form streams which can be visible as a band of light across the sky. The curious and somewhat puzzling property of Sagittarius is that it has managed to survive for many billions of years, even though it comes extremely close to the Galactic centre. In Chapter 5 we describe models which help understand why it has not been fully destroyed yet. We find that Sagittarius was probably very different in the past, probably many times more massive than what it is today.

The dark-matter halo of our Galaxy

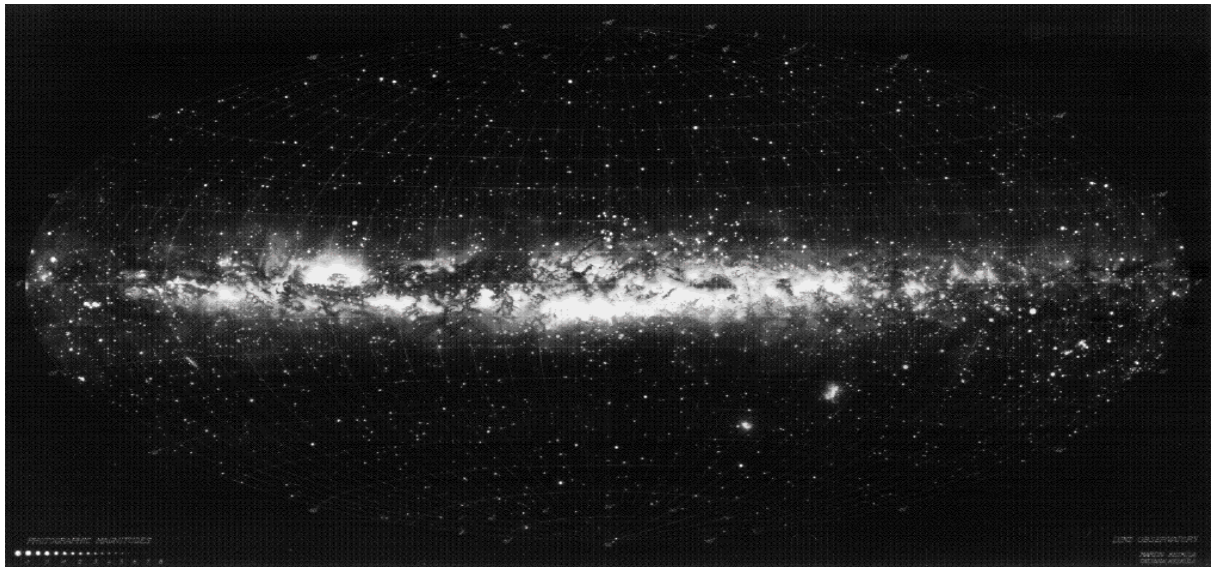
The Milky Way has both a “visible” stellar and a “dark” halo, which extends much farther out than the disk or the stellar halo we have been discussing so far. We derive its existence from the motions of the Milky Way satellite galaxies. These satellites have orbits that take them very far from our Galaxy, but still remain bound to it (i.e. they always come back). By measuring the mass in stars and gas in our Galaxy, one finds that this “visible” mass is about a factor of ten smaller than what is required to keep the satellite galaxies orbiting around the Milky Way. Because we cannot see the mass that is responsible for this gravitational pull, we call it “dark matter”. Dark matter is also found to be present in other galaxies, in aggregates of galaxies which are called clusters, and in aggregates of clusters of galaxies. The dark matter is present everywhere in the Universe, and like in our Galaxy, it is about ten times more abundant than the gas and stars we can observe directly. One of the key problems in astrophysics and particle physics is to determine what kind of matter makes up 90% of the mass budget in the Universe.

In Chapter 6 we address what should be the properties of the dark halo of the Milky Way, if the Galaxy was indeed built hierarchically. A key ingredient of the hierarchical paradigm is that it assumes that the dark matter is formed by some kind of elementary particles that interact very weakly, and are thus difficult to detect. Understanding the properties of our dark halo is particularly important because it is impractical to build experiments to detect every possible elementary particle candidate. Moreover, it is also important to understand if a signal is found, what is its origin and whether these experiments may have finally detected and determined the nature of the evading and dark component of the Universe.

Nederlandse samenvatting

ZOALS de Aarde niet alleen, maar samen met acht andere planeten door het heelal beweegt, zo leeft ook onze Zon niet alleen, maar woont zij samen met honderd miljard andere sterren (zonnen) in ons sterrenstelsel, de Melkweg genoemd.

Wanneer men op heldere avonden de hemel bekijkt, ontdekt men meteen de Melkweg: een zwak glanzende wolkachtige lichtband, vandaar zijn naam. Dit licht is afkomstig van miljoenen sterren, die te ver weg staan om met het blote oog te zien. Al hoewel het lijkt alsof wij buiten de "lichtstrook" zitten, behoren we tot een spiraalarm aan de rand van de melkwegschijf. De "melkweg" is in werkelijkheid een band van sterlicht, die men ziet, wanneer men een dunne schijf op zijn kant bekijkt.

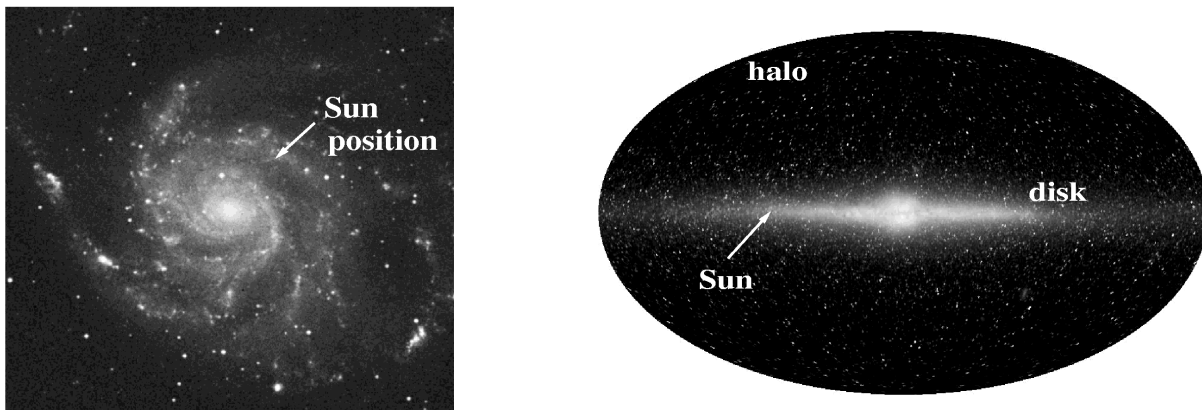


Panorama van de Melkweg aan de hemel. Sterren zijn te zien als witte puntjes. De melkwegwolken, eigenlijk het gecombineerde licht van duizenden sterren in de dunne schijf, worden onderbroken door banen van donker stof (Lund Observatory, Zweden).

Er zijn honderden miljoenen sterrenstelsels net als onze Melkweg verspreid door het heelal. Niet allemaal hebben zij spiraalarmen, sommige zijn elliptisch (ze lijken op een rugby-bal) en anderen zijn heel onregelmatig. Sterrenstelsels komen voor in verschillende kleuren en maten. Een belangrijke vraag in de sterrenkunde is hoe sterrenstelsels zijn ontstaan. Waarom zijn er verschillende soorten stelsels, en waarom hebben ze hun specifieke vorm?

Terug in de tijd door te kijken naar het diepe heelal

Laten wij even terugkeren naar de omgeving van de Aarde. Onze Zon ligt 150,000,000 kilometer ver van ons weg. Omdat de snelheid van het licht 300,000 kilometer per seconde is, duurt het voor het licht acht minuten om van de Zon naar de Aarde te reizen. Met andere woorden: het waargenomen beeld van de Zon is dat van acht minuten geleden. Zou de Zon op grotere afstand van de Aarde liggen, dan zou haar licht er langer over moeten reizen. Bijvoorbeeld, de dichtstbijzijnde ster is ongeveer 4 lichtjaar van ons verwijderd: haar licht moet vier jaar door de ruimte reizen voordat het op de Aarde schijnt! Voor sterrenstelsels, die nog veel verder



In het linker figuur is het sterrenstelsel M101 te zien, ook een spiraalstelsel dat op de Melkweg zou kunnen lijken gezien van boven. De positie van de Zon is aangegeven met een pijl. De rechter foto werd gemaakt door de COBE satelliet, en laat de melkwegschijf zien in het infrarood. Ook zijn een paar halo-sterren te zien buiten de schijf.

weg staan, kan deze tijd oplopen tot enkele miljarden jaren! Verweg gelegen sterrenstelsels worden dus waargenomen zoals ze er miljarden jaren geleden uitzagen, vlak na hun ontstaan. Dit betekent dat als we waarnemingen nemen van ver weggelegen stelsels en stap voor stap steeds dichterbij komen, we een tijdreeks krijgen van hoe sterrenstelsels zijn ontstaan en hoe ze zich ontwikkelden, tot wat we nu zien in onze eigen omgeving. Dus de geschiedenis van sterrenstelsels staat aan de hemel geschreven.

Dit idee wordt heel veel en vaak gebruikt door sterrenkundigen. Bijvoorbeeld, door de hele diepe waarnemingen met de Hubble Space Telescope van een klein gedeelte van de hemel weten we nu dat de meest verweg gelegen sterrenstelsels veel kleiner en onregelmatiger zijn, en dat ze elkaar vaak sterk beïnvloeden, zoals botsende sterrenstelsels.

Kunnen wij deze observationele ontdekking begrijpen? Jazeker. In de laatste twintig jaar hebben sterrenkundigen een theorie ontwikkeld, waarin sterrenstelsels ontstaan door zwaartekrachtsversterking van de oorspronkelijk in het heelal aanwezige massafluctuaties. In het vroege heelal zijn eerst de kleine sterrenstelseltjes gevormd, die botsen vervolgens met andere kleine sterrenstelsels en smelten zo samen tot grotere stelsels, zoals onze eigen Melkweg.

Dit proefschrift: De galactische halo

Een andere manier om te leren hoe sterrenstelsels zijn ontstaan is door het bestuderen van ons eigen Melkweg. Indien de Melkweg een typisch sterrenstelsel is (en dat denken we wel) kunnen we gebruik maken van de gedetailleerde informatie die we over de Melkweg hebben om te testen of het "hiërarchische" model klopt, wat er misschien fout aan is en waarom. Men kan niet bewijzen of een theorie, danwel een model helemaal klopt, men kan alleen proberen om de theorie op de proef te stellen door nieuwe gegevens te vinden die de modellen lek kunnen schieten (het is de nachtmerrie van de theoretische sterrenkundige en de droom van de observationele sterrenkundige!).

Het voordeel van het bestuderen van de Melkweg ligt in het feit dat de sterren in ons stelsel zo dichtbij zijn dat hun individuele kenmerken waargenomen kunnen worden, zoals hoe ver weg ze staan, hoe snel zij bewegen en in welke richting, enzovoort. Dit is nog onmogelijk voor andere sterrenstelsels, met de mogelijke uitzondering van een paar satellietstelsels van de Melkweg. Indien sterren onthouden waar ze vandaan komen, dan kunnen wij die informatie gebruiken om te bepalen of onze Melkweg door een hiërarchisch proces is ontstaan of niet.

De Melkweg heeft een dunbevolkte, bijna bolvormige halo om zich heen: de galactische

halo. De dunne Melkwegschijf, waar onze Zon heel regelmatig rondom het centrum beweegt, is ontstaan door afkoeling van gas. De sterren die zich in de schijf bevinden zijn in de schijf geboren en zullen hierin ook dood gaan. De sterren, die tot de halo behoren, zijn meestal heel oud en zeer metaalarm. Omdat er in het vroege heelal geen zware elementen bestonden (deze zijn later gevormd in het binnenste van sterren) kunnen we zeggen dat zeer metaalarm sterren lijken meer oorspronkelijke sterren. Het zijn deze sterren, waarvan men denkt dat ze zijn ontstaan in sterrenstelsels, die samengesmolten zijn tot de Melkweg. Daarom is de vraag van hoe ons melkwegstelsel is gevormd gelijkwaardig aan de vraag hoe de galactische halo is ontstaan.

Wat zijn de kenmerken van de samensmeltingsgeschiedenis van de Melkweg?

Dat is de centrale vraag van hoofdstuk 2 dit proefschrift. Met behulp van computersimulaties en analytische berekeningen (ja! de klassieke “potlood en papier” wetenschap!) hebben we ontdekt dat sterrenstelsels, die in botsing gekomen zijn met onze vroege Melkweg terug te vinden zijn als groepen sterren. Deze groepen sterren bewegen in onderling evenwijdige banen, en daarbij als het ware in hun beweging de herinnering aan hun gemeenschappelijke oorsprong bewaren. In dat hoofdstuk voorspellen we dat tengevolge van de hiërarchisch vorming van onze Melkweg, de halo-sterren in de Zonsomgeving zijn verdeeld in zo’n paar honderd groepjes, en dat sterren die tot dezelfde groep behoren met bijna dezelfde snelheid door het heelal bewegen.

Fossielen ontdekken: de restanten van een dwergstelsel

Zoals paleontologen naar fossielen zoeken om de evolutie van levende soorten op Aarde te kunnen bepalen, zo zoeken wij in hoofdstuk 3 naar de specifieke eigenschappen van de beweging van nabije, oude sterren om daaruit te kunnen afleiden hoe ooit ons Melkwegstelsel is gevormd, en hoe het zich ontwikkelde tot het stelsel dat we nu kennen.

Tussen 1989 en 1993 heeft de Europese HIPPARCOS satelliet heel precies de plaats en de beweging gemeten van 120,000 sterren in de nabije omgeving de Zon. Dit was een unieke kans om sterrengroepjes op te zoeken. De catalogus verschaftte ons de precieze posities en bewegingen langs de hemel, die uit andere bestaande catalogi voor dezelfde sterren konden worden aangevuld met de bewegingen in de gezichtslijn, dus loodrecht op de beweging langs de hemel. Op deze wijze slaagden we erin om voor een groot aantal heldere, metaalarme reuzensterren in de buurt van de Zon de drie-dimensionale ruimtelijke beweging te bepalen. Deze sterren bestaan net als de Zon voornamelijk uit waterstof en helium, maar hebben minder dan een dertigste van het ijzergehalte van de Zon. Aangezien het ijzergehalte een goede maat is voor het tijdstip waarop een ster gevormd is, behoren zulke zeer metaalarme sterren dus tot de oudste sterren van ons Melkwegstelsel.

De analyse van deze waarnemingen liet zien dat van de uitgekozen zeer oude sterren veruit de meeste, zoals verwacht, een willekeurige beweging ten opzichte van de Zon hebben, maar dat sommige echter een gezamenlijke afwijkende beweging hebben. In het bijzonder vonden we twee groepjes van oude sterren: één van drie sterren die uit de Melkwegschijf omhoog beweegt, en één van negen sterren die met vrijwel dezelfde snelheid uit de schijf naar beneden beweegt. Vergelijking met computermodellen bracht ons tot de conclusie dat de twee groepjes een gemeenschappelijke oorsprong hebben en deel uitmaken van het restant van een satelliet-melkwegstelsel, dat al heel lang geleden uiteen is gevallen. De eigenschappen, afgeleid voor dat stelsel, maken het tot het evenbeeld van de dwergmelkwegstelsels die ook nu nog op grote afstand om ons Melkwegstelsel bewegen. De nu ontdekte stergroepjes lijken dus het eerste fossiele bewijs te zijn dat onze eigen Melkweg eveneens is samengesteld uit de restanten van vele en kleinere bouwstenen.

Kunnen we nog andere fossielen vinden?

Met een grotere catalogus van oude sterren verwachten we veel meer zulke stergroepjes te ontdekken, fossielen van samengesmolten sterrenstelsels. Om te bepalen hoe onze Melkweg zich heeft ontwikkeld hebben we dan rond vijf duizend halo-sterren nodig met zeer precies gemeten bewegingen. Dit zal mogelijk worden met nieuwe satellieten, zoals de Amerikaanse FAME of de Europese GAIA.

In hoofdstuk 4 laten we zien dat de nieuwe satellieten zo nauwkeurig zullen zijn, dat we eindelijk zullen kunnen zeggen of de hiërarchische theorie wel of niet klopt voor onze Melkweg. De voorspelde honderden sterrengroepjes moeten erg duidelijk te vinden zijn, als ze bestaan! Met GAIA zal het ook mogelijk zijn om 50% van alle sterrenstelsels die samengesmolten zijn tot de Melkweg, terug te vinden.

Een levend fossiel: het Sagittarius dwergstelsel

De samensmelting van stelsels is niet iets van het verleden. Men hoeft alleen maar te kijken naar het meest dichtstbijzijnde satellietstelsel van de Melkweg: het Sagittarius dwergstelsel. Sagittarius werd pas ontdekt in 1994, en is sterk vervormd door de getijdenkrachten van de Melkweg. Het is zeer uitgerekt en niet meer zo helder. Het is zo goed als zeker dat dit stelsel over een paar miljard jaar volledig door het Melkwegstelsel zal worden opgeslokt. De sterren die niet langer deel maken van Sagittarius zijn te zien als een lichtstrook aan de hemel. Wat opvallend is aan Sagittarius is dat het de sterke zwaartekracht van onze Melkweg heeft overleefd, ondanks het feit dat al miljarden jaren lang het stelsel keer op keer heel dicht bij het galactische centrum komt. In hoofdstuk 5 hebben we modellen ontwikkeld die kunnen verklaren waarom het stelsel nog niet helemaal verdwenen is. We denken dat Sagittarius in het verleden heel anders was; waarschijnlijk had het een tienmaal grotere massa dan nu.

De galactische donkere halo

De Melkweg heeft een zichtbare stellaire en een donkere halo, die veel groter is dan de schijf of de stellaire halo, waar we het tot nu toe over gehad hebben. De snelheden van de satellietstelsels van de Melkweg zijn te groot om veroorzaakt te kunnen worden door de zichtbare massa. Deze is slechts tien procent van de massa die vereist is om de bewegingen te kunnen verklaren. Omdat de rest van de massa niet zichtbaar is wordt het “donkere materie” genoemd. Donkere materie is ook te vinden in andere sterrenstelsels, in groepen sterrenstelsels die clusters worden genoemd, en in groepen van clusters. De donkere materie is te vinden op elke plek in het heelal, en schattingen tonen aan dat er misschien wel tienmaal zoveel donkere materie als zichtbare materie is. Waaruit deze donkere materie bestaat is een van de grootste vraagstukken in de hedendaagse sterren- en natuurkunde.

In hoofdstuk 6 bepalen we wat de karakteristieken van de donkere halo van onze Melkweg zijn, of de Melkweg werkelijk op een hiërarchische manier is ontstaan. Een belangrijk ingrediënt van de hiërarchische modellen is dat de donkere materie bestaat uit elementaire deeltjes die bijna geen wisselwerking met elkaar hebben, en dus moeilijk te meten zijn. Het begrijpen van de karakteristieken van onze donkere halo is zeer belangrijk, omdat het onpraktisch is experimenten te ontwikkelen voor iedere mogelijke kandidaat van de elementaire deeltjes. Men moet vantevoren ongeveer weten wat men zoekt. Indien men een signaal ontdekt, is het belangrijk om zijn oorsprong te bepalen en dan ook of zulke experimenten eindelijk de ware aard van de onzichtbare massa in het heelal kunnen achterhalen.

Curriculum Vitae

I was born in Bahía Blanca (Argentina) on October 6th 1970. After finishing my high school education, I began to study Astronomy at La Plata University in 1989. In 1994 I graduated with a grade average of 9.91 (out of a maximum of 10), and was selected the Best Student of La Plata University of that year. My diploma thesis was on the subject “Integrability and chaos in Kaluza-Klein theories”, which I carried out under the supervision of Prof. Héctor Vucetich. I then became a fellow of the Argentine Research Council and started a PhD project on “Chaos in the early Universe”. Between 1993 and 1995 I was a teaching assistant in the course “Elements of Theoretical Astrophysics”, at La Plata University.

After being a summer student at the Instituto de Astrofísica de Canarias, and taking part of the 5th Vatican Summer School, I was awarded an Amelia Earhart fellowship, which gave me the final push to move to the Netherlands, in the beginning of 1996. A year later I was granted a fellowship from Leiden University to work towards the PhD degree, on a subject of my choice, my preferences lying in the fields of galaxy formation and cosmology. After being introduced to Prof. Simon White, the subject of my PhD thesis became clear: The formation of the Galaxy in the context of hierarchical cosmologies.

During the past three years, I have visited the Max Planck Institut für Astrophysik in Garching (Germany) on many occasions. I have had the opportunity to present my work at several international astronomical meetings such as in Rome, Canberra, Aspen, Ringberg Castle (Germany), Liège, Asiago and Heidelberg. I have also spent short periods of time as a visitor at Mount Stromlo Observatory and the Institute for Advanced Study in Princeton, and I have given colloquia at the Astronomy Department of the University of Virginia (Charlottesville), the Astronomy Department of the University of Córdoba (Argentina) and the Kapteyn Instituut in Groningen.

In May 1997 I was awarded for the second time the Zonta International Amelia Earhart Fellowship.

Acknowledgments

Being close to the end of my PhD, I look back and find so many friendly faces that have accompanied me, in some way or another through this journey. My first officemates Renate, Nicole and Steve, have been of great support during a difficult start in a different country. Maria Rosa, Laura and Eleni, your friendship is one of the most valuable things I have from this period. I also want to mention here Nicolas and my patient officemate Philip.

Mijn nederlandse familie, in het bijzonder de van Oostveen's en de Degenhart's, wil ik heel graag bedanken voor de gastvrijheid die er voor zorgde dat Mariano en ik ons snel thuis voelden.

La distancia pone a prueba la amistad, pero también nos muestra quiénes nos acompañarán toda nuestra vida. Gracias Santiago, Ale, Mariela, Karina, Anita y Martha por tanto cariño, por estar cerca aún estando lejos.

At the Sterrewacht, I am grateful to Janet, Marja, Pedro K., and Vicente. My gratitude is also for our big brother for helping me give shape to my PhD, and for always keeping an eye on me.

I would like to acknowledge the hospitality of MPA during my frequent visits, and my gratitude especially goes to Maria and Cornelia, and to the computer support group for always being ready to help. My MPA friends made me feel at home so many times. I am most grateful to "el Santo", for leading me in my first steps through science with the highest standards, for teaching me almost everything I know (or ought to know!), for the enthusiasm and for helping me make some of my dreams come true.

I appreciate Martin Heemskerk's (et al.) help in translating the summary into Dutch. For my many trips to conferences and working visits I received financial support from LKBF, NWO, EARA and the TMR European Network. The printing of the colour figures was subsidised by LUF (Leids Universiteits Fonds). I would also like to acknowledge Zonta International for the Amelia Earhart fellowships, and George C. for believing in me.

I would not be an astronomer close to finishing a PhD if it were not for my parents. Gracias mami, por tantos sacrificios y tanto apoyo. Gracias papi, por enseñarme el camino de la ciencia, cuando desde chiquita me llevabas al instituto, y traías revistas Nature a casa, por sentarte conmigo y explicarme trigonometría, y química, y tantas otras cosas. Gracias a los dos, por haberme dado la libertad de elegir mi camino, y por la confianza que siempre tuvieron en mí. Los quiero mucho!

Finally, I want to thank Mariano, for making such a beautiful cover. I have always felt it would not look proper to thank my spouse in a paper, so this is my only chance to thank you for the many times that you were there to help me out, like the time of the luminosity function... Mi amor: gracias por tu paciencia, por tu cariño, por estar siempre listo para ayudarme, por alentarme constantemente, por creer en mí... Amor, este camino no hubiera sido posible sin vos, ni tampoco hubiera tenido sentido. Porque vos le das sentido a mi ser.

'Would you tell me, please, which way I ought to go from here?'
'That depends a good deal on where you want to get to', said
the Cat. 'I don't much care where...', said Alice. 'Then it
doesn't matter which way you go', said the Cat. 'So long as I
get somewhere', Alice added as an explanation. 'Oh, you're
sure to do that', said the Cat, 'If you only walk long enough'.

L. Carroll: "Alice's Adventures in Wonderland"

