WFMOS Stellar Archaeology and Galaxy Genesis: Draft OCDD

Mike Irwin (IoA), Geraint Lewis (Sydney Uni), Mark Wilkinson (IoA), Vasily Belokurov (IoA), Kim Venn (UVic), Alan McConnachie (UVic), Beatriz Barbuy (IAG-USP), Luciana Pompeia(IAG-USP), Chris Evans (IfA), Dante Minniti (U.Catolica, Chile) ....

July 2006

1 Main Science Surveys

Survey	Rlim (AB)	SD deg^{-2}	Total Area	No. Stars	R	Wavelengths	Nights
DE	22.7	1000	2000	$2x10^6$	500	5600-7500Å	153
z = 0.5-1.3					2500	7500-10000Å	
DE	24.5	2000	300	$6x10^5$	500	3900-5200Å	136
z = 2.3-3.3							
GA-LR	21.0	1000	500	$0.5x10^6$	2000	3900-10000Å	140
GA (new)	20.0	1000	2000	$3x10^6$	5000	8400-8900Å	100
Low res							
GA-HR	17.0	500	3000	$1.5x10^6$	40000	3900-10000Å	490
GA (new)	17.0	500	2000	$1x10^6$	20000	4800-6800Å	200
High res							

Original proposed surveys and new suggested replacement surveys

2 Dark Energy ....

...... and the answer is .... w ≈ -1.

Contents

3 Stellar Archaeology and Galaxy Genesis

3.1 Science case

3.1.1 Overview

The origin and evolution of galaxies like the Milky Way and M31 remain among the key unanswered questions in astrophysics. The galaxies we see today in and around the Local Group are representatives of the general field population of the Universe and have been evolving for the majority of cosmic time. As our nearest neighbour systems they can be studied in far more detail than their distant counterparts and hence provide our best hope for understanding star formation and prototypical galaxy evolution over the lifetime of the Universe (Freeman & Bland-Hawthorn 2002). Although significant observational progress has been made recently we are still a long way from understanding galaxy genesis. To unravel this formative epoch, detailed large area spatial, kinematic and chemical surveys are required, providing the link between the local near-field cosmology and predictions from the high-redshift Universe.

ΛCDM cosmologies demonstrate the ubiquity of hierarchichal merging as the main driver in galaxy formation and evolution, and the discovery of the tidally disrupting Sagittarius dwarf (Ibata et al. 1994) provided the first compelling nearby evidence of this. However, the detailed process by which large galaxies such as the Milky Way arrive at their current state is still largely speculative (*e.g.* Abadi et al. 2003, 2006; Bullock and Johnston 2005; Font et al. 2006) despite recent observational and theoretical progress.

Current large scale spectroscopic surveys such as RAVE (Steinmetz, 2003) and the M31 kinematic surveys on Keck (Ibata et al. 2005, Guhathakurta et al. 2005) highlight the potential of multi-object spectroscopy to probe the structure and properties of the Galaxy and M31 at modest resolution, while FLAMES on the VLT provides a compelling argument for the multiplexed use of mid- to high-resolution spectroscopy particularly in understanding nearby dwarf galaxies (Tolstoy et al. 2004). However, FLAMES, although producing exquisite spectra of hundreds of objects, is still at least an order of magnitude shy of the requirements needed for analysis of the Galaxy and M31.

All of these efforts are necessarily limited in either coverage or depth. While future proposed Galactic surveys such as ARGOS with AAOmega (Lewis et al. 2005) and the ESA cornerstone space mission GAIA (Perryman et al. 2001) will go some way to alleviate this problem for the Galaxy, they will still leave unexplored a huge volume of parameter space that holds the key to full understanding of Galactic evolution.

The diversity of current and planned surveys illustrates the scientific importance and community interest in this field and WFMOS presents a unique opportunity for a major advance in our understanding of the origin, chemical evolution and mass assembly of nearby galaxies. The key advantages of WFMOS over existing 8m-class facilities are the greater multiplexing and wavelength coverage together with the much larger field-of-view. This provides enormous leverage in Galactic studies by enabling a range of structural probes in a multi-parameter space sensitive to both dynamical and chemical properties.

Galaxies are dynamically and chemically inhomogeneous and it is only by using the higher dimensionality afforded by combining kinematic and chemical measures that a full understanding of their constituent parts will arise. The fossil record of past events is recorded in the spatial distribution, kinematics and chemical fingerprint of the individual stars. Whereas a low resolution (LR) survey enables the structural decomposition of the Galaxy through the identification of kinematically and metallicity-related groups of stars, a targetted high resolution (HR) survey enables the detailed evolutionary history of these components to be determined. This combination will provide a unique insight into galaxy assembly.

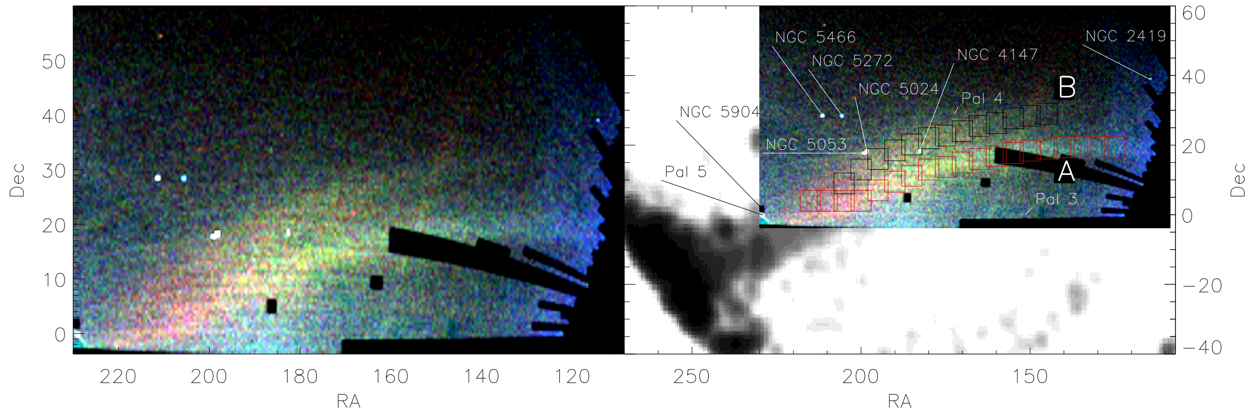

Figure 1: Field of streams figure from Belokurov *et al.* 2006 showing SDSS stars having $g - r < 0.4$ for a region around the North Galactic Cap. Colours denote depth along the line-of-sight with red being the most distant objects at d≈50 kpc and blue the nearest at d≈10 kpc. The right-hand panel shows a composite including the 2MASS M-giant distribution from Majewski *et al.* (2003). Several streams are readily visible.

Figure 3.1.1 shows an example of the current state-of-play in tracing substructure in the Galactic halo from large area photometric surveys. Several giant streams (up to ≈60° long) from disrupting globular clusters and satellite dwarf galaxies are now known to criss-cross the halo providing probes of both the nature of dark matter and the gravitational potential of the Galaxy.

3.1.2 Overall design of GA surveys

Although the feasibility study provided an excellent starting point in highlighting the huge potential gains available for Galactic science, the ambitious nature of the programme led to a rather imbalanced set of proposed surveys. After independently reassessing the science goals and the technical requirements needed to meet them, we feel that the approach adopted in the feasibility study is not necessarily the best strategy. In our opinion we can achieve, or even improve on, the original science goals, with a more modest amount of telescope time.

The science drivers for GA with WFMOS still naturally split into two components:

- LR kinematic and general abundance analysis of Galactic structure;
- detailed HR abundance of constituent populations to trace temporal chemical development;

but the detailed strategy within each is somewhat different.

LR survey

The LR mode should focus upon obtaining accurate kinematics and metallicity measures for a large stellar sample. In this mode, the main Galactic survey science for WFMOS (15 < V < 20) is complementary to the GAIA ($V < 15 - 17$) and RAVE ($V < 12$) large area surveys. GAIA synergy (see section xx.yy) is important since for both full kinematic and HR abundance analyses, good distances and accurate proper motions are required. This suggests that in general matching to GAIA survey depths for the LR survey, *i.e.* to V≈20, is sensible. However, we also do not want to lose sight of the advantage of an 8m-class telescope for deeper surveys in selected regions.

Sufficient velocity accuracy is essential for identifying streams and substructure against the Galactic background population. As an example, the required velocity errors for 3-sigma detection of streams for an n-dimensional kinematic survey are:

$$\epsilon = \frac{\sigma}{3} \times \frac{1}{N^{1/n}}$$

where σ is the dispersion of the background population, and N is the number of streams (e.g Helmi & White 1999). For a survey limited to radial velocity information this implies that errors of ≈0.1 km/s are required to locate halo streams in the Solar neighbourhood (assuming σ = 100 km/s and N=500; Wilkinson et al. 2005). Adding proper motions, improves this dramatically implying required errors of about ≈4 km/s for Halo substructure. For disk streams the radial velocity error would need to be more like ≈2 km/s.

Furthermore, satellite galaxies have velocity dispersions in the range 6-12 km/s whilst clusters range from 1 km/s (open clusters) to 10-20 km/s (massive globulars). As the velocity dispersion of tidal tails of disrupting systems reflects that of the progenitor, detecting and analysing dissolving clusters therefore requires errors also of a few km/s. Additional complications arise from the smooth foreground population (*e.g.* Brown, Velazquez & Aguilar 2005) against which it is harder to pick out substructure. Here, LR abundance measures will help differentiate substructure, and again velocities errors of around 2 km/s are necessary.

An efficient mid-resolution (R=5000) spectroscopic survey, comparable to FLAMES in LR mode, could be used to obtain velocities to ≈2 km/s precision and proxy [Fe/H] measurements to 0.1 via the NIR CaT. Although this would probably not give more exotic alpha-element or r- and s-process element measures at the individual star level it would alleviate the need for an unfeasibly large HR sample. It would also directly allow the tracing of the dissolution and tidal debris from clusters, where km/s precision is needed, and would also enable detailed studies of Halo substructure.

Similar reasoning holds for M31 and M33 surveys. The current ongoing surveys sensibly make use of modest resolution spectrographs to derive both kinematic and population-averaged overall metallicity measures. It is now well known (*e.g.* Tolstoy et al. 2004) that individual velocity and [Fe/H] parameters are required to unravel the complex star formation histories in dwarf galaxies. The same will inevitably be true for their giant neighbours. A further possiblity, is that by combining spectra from similar kinematic components, modest R=5000 resolution spectra, but at high signal:to:noise, could be used to extract additional chemical information, such as alpha-element ratios, by being more creative with analysis tools such as ANNs specifically trained (*e.g.* using currently available FLAMES spectra) to extract this information.

HR survey

For the HR mode we feel there is little to gain by having R=40000 with limited wavelength coverage. Here we are proposing a much larger simultaneous wavelength coverage but at R=20000. Results from VLT FLAMES, Keck HIRES and MIKE Magellan, suggest that chemical tagging using this strategy is possible and by focussing on redder wavelengths (4800–6800Å) than those highlighted in the feasibility report, studying a million stars at HR is emminently achievable.

Our proposed high-resolution spectroscopic surveys (R=20000) are comparable to FLAMES/GIRAFFE in HR mode (where we already have built up a lot of experience) but with the considerable advantage of simultaneously covering all the primary wavelength region of interest (4800Å-6800Å) at a single visit. Combined with the 5- to 10-fold increase in target allocation this produces a 25- to 50-fold increase in survey efficiency enabling a range of GA projects that would be infeasible with current generation instruments. By taking advantage of the large simultaneous wavelength coverage and by adopting multi-component template and line profile fitting, the science requirements on

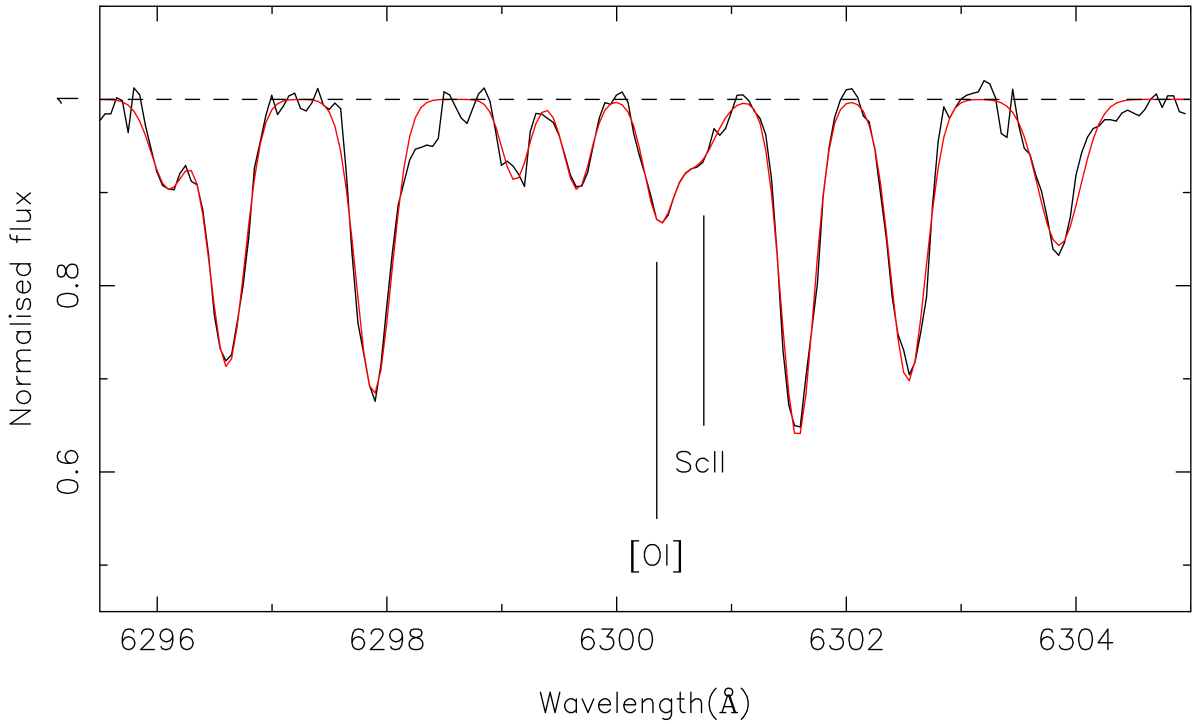

Figure 2: The region around the $[OI]_{6300}$ line from a 5400s VLT-FLAMES exposure at R=20000 of a metal-rich Bulge giant. An unconstrained multi-component Gaussian model fit of the most prominent lines is shown overlaid.

accurate stellar atmosphere modelling together with detailed abundance analysis for a broad range of atomic species can be comfortably met.

The advantage of contiguous wavelength coverage in the range 4800-6800Å is the ability to make use of a dozen FeII lines and more than a hundred FeI lines to derive accurate stellar atmosphere parameters: effective temperature; surface gravity; metallicity and microturbulent velocity; using excitation and ionization equilibrium and known curves-of-growth. The most important elements for chemical tagging all have measurable lines in this region.

3.1.3 GA requirements

The preceding considerations lead to the following overall survey requirements:

LR survey

The LR feasibility study case has velocity errors of order 10 km/s which we feel is far too coarse to detect tidal streams and accurately assign population membership. A resolution of R=2000 also precludes accurate EW estimation. In contrast R=5000 in the CaT region opens up the parameter space with reasonable s:n spectra (see section xx.yy), allowing velocity precision of a few km/s accuracy and [Fe/H] estimation to 0.1 dex. It also leads naturally to surveys better matched to GAIA.

HR survey

The HR observations are needed to measure the individual lines of a number of chemical species that are required for chemical tagging. A resolution of R=20000 is a compromise between resolving distinct lines and wavelength coverage, yet still allows the majority of lines of interest to be quantified. By working in the "visible" regime the problem of line blending is lessened but still affects some important lines. A good example is provided by the forbidden oxygen line $[OI]_{6300.3}$, shown in figure 3.1.3 from a 5400s VLT-FLAMES R=20000 resolution spectrum of a metal-rich Bulge giant. Here the [OI] line is half-blended with $ScII_{6300.7}$, but is still distinct enough to quantify using multi-component profile fitting as illustrated in the figure. This same star was observed at R=47000 with VLT-UVES. A direct comparison of derived equivalent widths suggests that, even in cases like this, EWs to better than 5mA precision can be obtained. The main limitation for metal-rich stars such as this is not the partially unresolved lines but the continuum placement.

A further advantage of a resolution of R=20000 is that by adopting a non-cross-dispersed spectrograph design it is feasible to measure extended spectral regions. This is impossible in échelle-mode instruments and can be used, for example, to tie in measurements of spectral indices (often all that is available for distant galaxies) with detailed line analysis. As an example, to measure indices of Mg2, Fe_{5270}, Fe_{5335} which are widely used requires fluxed spectra covering a wide spectral range. For example, to measure the Mg2 index at 5156-5197Å requires continuua defined at 4897-4958Å and 5300-5366Å therefore requiring continuous coverage from 4897Å to 5366Å. A large spectral coverage also makes it feasible to analyse the extended wings of the Balmer lines to derive more accurate effective temperatures.

We also note that absolute flux calibration is not necessary since it is straightforward to use extant photometry to provide the conversion to from relative to absolute fluxes and thence spectrophotometry.

These considerations allow a well-defined set of technical requirements to be drawn up that meet the GA science aspirations. The following table summarises them.

Summary of Galactic Program Requirements

To exploit kinematic synergy with GAIA implies:

LR survey to V=20 (distances, proper motion, radial velocity, [Fe/H])
HR survey to V=17 (uniqueness of abundance analysis for chemical tagging)

LR	(extent of halo, ancient streams, disk structure, Fe/H) stellar spectra LR: kinematics vels to ≈2-3 km/s and [Fe/H] R=5000 CaII triplet 500Å from 8350-8850Å Npix=1500, V < 20, 0.5hr exposures 2000 deg^2, 3 10^6 stars, 100 nights FWHM 3 pixels, pitch 8 pixels, 3000 fibres
HR	(chemical tagging of elements, IMF, SF history) R=20000, Npix=20000, 4800-6800Å Npix=20000, V < 17, 1hr exposures 2000 deg^2, 10^6 stars, 200 nights FWHM 3 pixels, pitch 8 pixels, 1000 fibres

3.2 Science Goals

Two of the key challenges facing modern astrophysics are our understanding of the fundamental nature and properties of dark matter, and of the inter-related complex baryonic processes that led to the formation of the Milky Way galaxy we see now.

The large-scale gravity-driven properties of dark matter guide the gaseous matter into proto-galactic environments, but the dominance of baryonic physics on small scales, through cooling, collapse and star formation, emphasises that understanding only the nature of dark matter provides a very limited view of galaxy formation and evolution.

However, both of these aspects can be simulataneously addressed with large scale WFMOS surveys of the stellar components of the Milky Way and companion galaxies. In addition to providing tests of cosmological models on low-mass scales, such surveys will also unravel the complex process of hierarchical formation wrapped within various Galactic structures. These studies directly compliment the dark energy aspects of this document, and taken as a whole these proposed studies with WFMOS will provide a new window on the three major energy constituents of the universe and their influence on the Cosmos we see around us.

To address the nature of dark matter and the baryonic physical processes within the the Galaxy we need to obtain the kinematic and chemical signatures of large populations of stars. Such a situation naturally breaks the surveys into two main parts; a low resolution survey to yield the kinematics and overall chemical abundance of stars over a large volume, and an overlapping higher resolution study to uncover the detailed patterns in stellar chemistry[1]. The synergy of these two main Galactic surveys will provide the most detailed picture of the processes that formed our Galaxy and of the nature of the component parts.

3.2.1 Addressing the big questions

Low Resolution (LR) Survey: This will result in a detailed kinematic and general abundance map of the Galaxy for which WFMOS depth and areal coverage are crucial. The major scientific goals for this survey are:

What is the nature of dark matter?

While our understanding of dark matter appears to be adequate on cosmological scales, its application is being challenged on the scale of individual galaxies. The WFMOS survey will focus on several crucial galactic-scale properties of dark matter. These are:

- **The mass and extent of the MW dark matter halo**

  The shape of the dark matter halo of the Galaxy is a consequence both of its formation history and the very nature of dark matter itself, and the Milky Way is the one massive spiral galaxy in which we can hope to map out the distribution of dark matter in fine detail. However, the majority of kinematic tracers currently available are tied to the Galactic disk, and the shape of the overall potential is very poorly constrained. Recent advances using tidal streams have met with some success (*e.g.* Fellhauer et al. 2006), although the paucity of data has ensured that the argument over the shape of the dark matter distribution in the Milky Way has continued to rage between oblate, spherical and prolate. In fact, the available data sets are tiny and will remain so until after GAIA - beyond 50 kpc there are currently only a few tens of tracers (dwarf galaxies, globular clusters, halo stars; *e.g.* Wilkinson & Evans 1999; Battaglia et al. 2005). GAIA will provide a few thousand objects out to 60kpc, but WFMOS could push the radial velocities much further out (to V=20 where there will still be proper motion constraints from GAIA); WFMOS will measure velocities of Carbon stars, AGB stars and TRGB stars to distances out to 250 kpc, with several tens of objects expected in a 1000 square degree survey - GAIA will provide proper motions for these stars out to 60-80kpc. These new tracers will make it possible to place significantly stronger constraints on the mass and extent of the Milky Way halo, as well enabling detailed study of the velocity distribution

[1] It should be noted that while GAIA will revolutionise this field, it lacks detailed abundances and accurate radial velocities to fully explore Galactic history. Hence WFMOS will fully compliment the GAIA endeavour and significantly enhance its scientific impact

of halo stars. Furthermore, WFMOS observations of the disk of the Milky Way will provide a detailed characterisation of the mass distribution close to the Galactic Plane, and it will be possible to investigate whether the Galaxy has a central cusp or core in its dark matter distribution. Since cusps are a generic prediction of galaxy formation in ΛCDM cosmologies, the existence (or lack thereof) of dark matter cores is a fundamental probe of our underlying cosmological ideas. Current data on the Galactic rotation curve and microlensing optical depth suggest that cusped models cannot reproduce the observations (Binney & Evans 2001). However, the detailed stellar kinematics provided by WFMOS will provide an essential probe of the dark halo profile in the central regions as they are implicitly dependent upon the form of the dark matter profile.

- **The lumpiness of dark matter from tidal stream properties**

  As globular clusters dissolve in the tidal fields of massive galaxies, stripped material forms into extended, cold stellar streams. Such cold streams are very sensitive to the heating, and dispersement, through interaction with ∼kpc substructure within the dark matter halo, while the hotter streams of dwarf galaxies are significantly more robust to such heating. Recently, extended tidal tails have been associated with globular clusters, including Pal5 and NGC5466, and more examples are coming to light in extensive imaging surveys (*e.g.* Belukorov *et al.* 2006; Grillmair & Dionatos 2006). As cold, young tails may morphologically resemble older, disrupting streams, it is vital to determine the heating of streams along their length, as this reveals the true nature of the stream and determines the interaction rate with any dark matter substructure. The WFMOS LR survey provides an ideal opportunity to kinematically map satellite tidal streams and hence constrain the dark matter substructure of the Milky Way halo.

- **The missing satellite problem - dark satellites**

  The dwarf spheroidal galaxies (dSphs) of the Local Group are valuable laboratories in which to test the properties of the dark matter, as they are extremely dark matter dominated in their inner regions (*e.g.* Kleyna et al. 2001, Wilkinson et al. 2004, Mateo et al. 1998) and appear to be the smallest stellar systems that contain dark matter in dynamically significant quantities. The number of low mass dark matter haloes in the vicinity of a Milky Way-type Galaxy is a strong function of the nature of the dark matter - cold dark matter (CDM) simulations typically over-predict the numbers of dSphs around the Milky Way by at least an order of magnitude (*e.g.* Moore et al. 1999) while warm dark matter models contain significantly fewer low-mass haloes (*e.g.* Bode, Ostriker & Turok 2001); this is known as the "satellite crisis" of CDM. The recent discovery of the extremely low luminosity Ursa Major dSph (Willman et al. 200, Kleyna et al. 2005), Boötes (Belokurov et al. 2006) and Canes Venatici (Zucker et al. 2006) satellites raises the interesting possibility that a number of dSphs containing very small stellar populations might have evaded discovery in previous surveys. However, photometric data alone are insufficient to establish definitively whether an object is a dSph or a star cluster - only kinematic measurements can achieve this. WFMOS will be the ideal instrument for the spectroscopic follow-up of new dSph candidates as wide-area surveys of stars to faint magnitudes (V=20-21) are required to provide velocity dispersions and abundance distributions. A candidate dSph with a low-luminosity ($M_V \sim -5$) and a large half-light radius ($\sim$ 200pc) would have a velocity dispersion of $< 1\mathrm{km\,s^{-1}}$ if it were a self-gravitating system containing no dark matter, while all dSphs observed to date have velocity dispersions in excess of $6\,\mathrm{km\,s^{-1}}$. The velocity dispersion thus provides the required strong discriminant between dSphs and extended star clusters, and the existence of a previously unknown population of dSphs would be a real victory for CDM hierarchical structure models. WFMOS will be able to provide efficient kinematic analysis of all newly identified candidate dSphs and thus to resolve the question of whether the so-called "satellite crisis" of CDM has been solved.

For the nearer satellites, GAIA will provide proper motion measurements of the brighter member stars. This will enable direct measurement of the velocity ellipsoids throughout the system and thereby help break the mass distribution-velocity dispersion anisotropy degeneracy.

- **The nature of dark matter "particles" from satellite profiles**

  The masses of dark haloes of dSphs can be used to constrain the properties of the dark matter, with the mass spectra of subhalos dependent upon its temperature. Furthermore, the density profile of the subhalos is also temperature dependent, with warm dark matter haloes are typically of lower concentration than those found in CDM simulations (Bode et al. 2001). Hence it is essential to obtain a robust estimate of the total masses and densities of the known dSphs. Significant observational effort over the past decade has produced large kinematic data sets in dSphs which are being used to constrain the mass distribution within the main body of each system - the shape of the density profile can be used to constrain the physical properties of the dark matter particles. However, to complete the picture, we must also determine whether the dark matter haloes of dSphs extend beyond the edge of their main stellar distributions, in a similar manner to the halo of the Milky Way and other massive galaxies. The outer parts of most, if not all, dSphs exhibit signs of tidal disturbance in their stellar distributions (*e.g.* Carina: Majewski et al. 2005). However, the stellar kinematics in the outer regions remain controversial (see *e.g.* Wilkinson et al. 2004, Muñoz et al. 2005 for the case of Ursa Minor). Deep, wide-area surveys of the outer regions of all dSphs, a task well-suited to WFMOS, are required to trace their stellar kinematics well beyond their nominal tidal radii. If the apparent disturbance of their outer stellar distributions is indeed due to the tidal effects of the Milky Way (which could be confirmed, for example, by the detection of photometric *and* kinematic tidal tails) then their total masses cannot be much larger than a few$\times 10^8 M_\odot$. On the other hand, if they are found not to exhibit evidence of tidal perturbation this would argue strongly that their masses are in excess of $10^9 M_\odot$, in which case the "satellite crisis" of CDM models may well be solved.

The mass and extent of the MW stellar populations

Our current understanding of the extent, structure and kinematics of the Milky Way is very limited. One of the primary goals of the WFMOS LR survey will be to dissect the Galaxy into its fundamental components, delineating structures based upon kinematics and metallicities. Without this deconstruction, understanding the detailed formation of the Milky Way is impossible. Such a survey will also reveal the presence of substructure, especially in the form of old, in-plane accretions. Furthermore, we only know the velocity distribution in the disk in the Solar neighbourhood, although a global understanding of disk kinematics is required to understand its origins. A combination of GAIA and WFMOS will extend kinematic knowledge out to the edge of the disk in the Galactic anti-centre direction; GAIA alone cannot do this as the radial velocities do not probe far enough out. The resulting dataset will, for the first time, provide answers to the major questions about the Milky Way, including; what features in the velocity distribution are resonances due to the bar (Dehnen and others have claimed some of the local features are bar-related - if they are resonances we should not see them at other radii)? How much substructure is in the disk? Can we see the effects of transient spiral arm passages through the disk? What is the variation of the velocity ellipsoid as a function of position in the disk? These will provide the important clues to the formation of the Galactic components.

The current generation of bulge surveys, for example the ARGOS survey using the AAOmega spectrograph, will provide first order maps of the kinematics of the bulge (velocity dispersion and mean velocity as a function of position through the bulge), as well as detailed metallicity distribution function. However, the much larger samples provided by WFMOS will make it possible to investigate in detail the second-order kinematic structure seen in simulations of the Galactic bulge (*e.g.* Figs. 10 and 16 of Fux 1997). These velocity features depend on the details of the stellar structure through the bulge. The large samples will also make it straightforward to detect the signatures of the resonances which may have produced the peanut-shaped bulge of the Milky Way (Quillen 2002) as well as other resonant features.

How unique is the Milky Way?

Is the Milky Way unique in its history and overall properties? To answer this question, any survey of the kinematic and chemical abundances of the Galaxy needs to be compared to similar studies of other spirals. In practice, this is impossible for all but the nearest spirals, M31 and M33. WFMOS offers the only realistic prospect of a panoramic spectroscopic survey of the stellar populations of our nearest large galaxies.

- **Large area M31/M33 velocity and abundance survey -** *e.g.* **outer halo M31**

  While M33 is distinctly different to the Milky Way, being 10× less massive and apparently undergoing a more quiescent formation, the similarity and differences between M31 and the Galaxy have long been debated. With 8-10m class telescopes, the ability to target the kinematics and chemistry of individual populations of stars, has shed some light on this, finding that both galaxies appear to possess similar overall structures, in the form of metal poor stellar halos and complex extended stellar disks. By the advent of WFMOS, it is envisaged that dedicated surveys by DEIMOS/GMOS-class instruments will have provided detailed kinematic and chemical maps of the bulge and disk structures in these galaxies, revealing details of the underlying stellar halo which will contaminate any study. Furthermore, current studies are already targetting significant substructures which are readily apparent in photometric maps of the outer disk of M31. As with the Milky Way, the halos of M31 and M33 hold great promise for unraveling the accretion history through the identification of extremely low contrast tidal streams in kinematic and chemical space. Again, such studies will also probe the underlying dark matter potential of these galaxies, providing further clues as to whether the Local Group dark matter distribution matches expectations from ΛCDM cosmological models. However, given their small field of view, it is impractical for DEIMOS/GMOS-class instruments to efficiently survey the sparse halo regions and it is clearly a region where the capabilities of WFMOS will provide a unique probe. Deep photometry (currently ≈80 sq deg of sky from the INT/WFC and CFHT/MagaCam) of the extended halos of both M31 and M33, is steadily accumulating, and down the SE minor axis of M31 already extends along the entire line-of-sight to the centre of M33, a distance of ≈220 kpc (15 degrees on the sky). These surveys are revealing an astonishing wealth of substructure, even in the outer halo. When complete coverage, ≈ 500 sq deg, is obtained in the next few years this unprecedented dataset will enable direct testing, via targetted efficient spectroscopic kinematic and metallicity followup surveys, of the predictions of ΛCDM cosmologies. Are fainter more diffuse tidal streams and relict debris dominant throughout the halo ? What is the shape and total mass of the dark matter halos of M31 and M33 ? Will this explain the puzzling apparently quiescent history of M33? WFMOS will be the only instrument that can tackle these important questions within a realistic time-frame.

High Resolution (HR) Surveys:

A major strength of WFMOS is the ability to exploit the natural synergy between targetting deeper surveys of stellar structure at low resolution with the more detailed picture afforded by the shallower higher resolution spectroscopy. For this type of work WFMOS/HR will provide a unique facility for undertaking the necessary detailed chemical tagging of large samples of stars, enabling deconstruction of the entire formation history of the Galaxy.

The primary question here is:

How has the Milky Way system evolved?

The formation and evolution of large scale structure channels the building material of galaxies, both dark matter and baryons, into high density environments. While this provides the birth site for galaxies, our understanding of 'What happens next?' is less clear. The combined LR and HR surveys with WFMOS will enable us to test the various scenarios that brought the Milky Way into existence. The fossil record of the chemical evolution of entire stellar populations and their subsequent dispersal around the galaxy is within reach and is opening up a new era of Galactic astronomy.

- **The evolutionary history of stellar components and subtructure**

  The HR surveys are crucial since they are required to fully unravel the formation history of Galactic populations. Chemical tagging through the accurate determination of chemical abundance ratios allows the identification of common formation sites for disparate populations and allows a detailed analysis of the physical processes governing star formation and evolution. Whereas the metallicity distribution of stars samples the chemical enrichment history of the galaxy, age estimates for individual stars provides information on the evolution with time of the abundances of a range of chemical elements. This impacts a wide area of astrophysics since the detailed progression of different species are directly affected by initial mass functions, star formation rates, infall and outflow of gas, nucleosynthesis pathways and stellar lifetimes.

- **The role of star clusters in structure formation**

  All stars form in clusters or associations. If we can reconstruct these using chemical tagging, that will be a valuable piece of information for star formation models used in cosmological simulations and will have relevance for feedback, galactic winds, etc. By reconstructing recently disrupted clusters we can determine current and recent cluster mass functions through combining kinematics and abundances. In addition, distances from GAIA combined with photometric information from wide field optical and near-infrared surveys hold out the tantalising prospect of age dating stars in groups and thereby linking coeval star clusters over the entire history of the Galaxy.

  Furthermore, Kroupa & Boily (2002) have emphasised the role that gas expulsion has on young cluster disruption and hence on the mass function of clusters that survive beyond star formation. This can be tested by looking for haloes of loosely bound or unbound stars near open clusters.

  [INLUDE SOME DISK CLUSTERS IN THE PROPOSED TARGETTED FIELDS]

- **Constraining nucleosynthetic pathways, sites, and stellar yields**

  Most elements still have uncertainties about the sites (*i.e.* on what type of object they form), the nucleosynthetic pathways from which they are produced (the different nuclear processes), and the yields from a given site (the amount of the element formed). The detailed study of abundance ratios of stellar populations of different environments impose very strong constraints on the different processes that produce elements in stars, improving our analytical power on the chemical evolution of the galaxies.

- **Signatures of ancient accretions**

The importance of combining kinematic and abundance information for tracing the members of disrupted systems has long been recognised. Stars now distributed far from their birthplace can be traced by tracking their location in a combined chemo-dynamic phase space (eg. Helmi *et al.* 2006). With detailed chemical abundance ratios, many of the ambiguities and degeneracies present in kinematic signatures can be resolved, enabling detailed allocation of various stellar components thereby allowing traceback to a common formation sites. Stars from common progenitors show distinct correlations between their orbital parameters, their abundance ratios and their evolution with age and can be directly used to probe evidence of the hierarchichal formation of the Milky Way.

- **The properties of extreme metal poor popIII stars**

  During the gaseous collapse of the Milky Way, the first generation of stars was left in a roughly spherical halo surrounding the newly formed disk, and this population of extremely metal poor stars are being uncovered today; Christlieb et al. (2003) report the identification of a number of ultra-low metallicity stars in the Galactic halo. These stars provide a direct link to the conditions in the protocloud from which our Galaxy was formed. The sparseness of the halo population, however, ensures the progress in the discovery of this population is very slow. Intriguingly, simulations of Galaxy formation suggest that the earliest pre-galactic stars will be concentrated in the central regions of the Galaxy at the present time. The photometric metallicity distribution function (MDF) in the Zoccali *et al.* (2003) Bulge field shows a peak at [Fe/H] ~ 0 but with a long tail extending to [Fe/H]~ -2, and a survey of the Galactic Bulge for very metal-poor stars should therefore yield a larger sample of such objects than the studies of the halo. Proposed surveys such as the AAOmega/ARGOS will begin this work, targetting selected Bulge regions to a limiting magnitude of I$\sim 16-17$, providing a complete spectroscopic MDF in several targetted Bulge windows.

  However, the vital key to the earliest epochs is the distribution, kinematics and chemical properties of the rarest, ultra-low metallicity stars. Here only a combined LR/HR large scale survey by WFMOS can obtain the spectra of sufficient numbers of stars to obtain a statistically useful sample. The LR survey necessarily acts as a first cut at locating extreme metal poor whereas the HR counterpart is required to guarantee the nature of the stars and to provide the detailed abundance ratios needed for atmospheric analysis.

  These extremely metal-poor stars are fossils of the first stellar formation and metal enrichment processes. As such they have triggered immense interest to try to answer the numerous questions relating to the initial mass function of the first stars and the subsequent mechanisms for early chemical enrichment in the Universe.

- **Linking the Bulge to the high Z universe**

  The Galactic bulge provides many clues to the nature of our galaxy. Laid down in the initial phases of the birth of the Milky Way, the bulge is related to spheroidal components of other galaxies and so establishing how it formed (collapse, accretion, disk instability) has an importance beyond merely improving our understanding of the history of the Milky Way. In addition, the interplay between the various processes involved will advance our knowledge of the stellar dynamics of galaxies more generally. The integrated spectra of Galactic Bulge fields and of Galactic globular clusters are directly comparable to those of distant ellipticals and the bulges of spirals as shown in Bica (1988). Spectral indices indicating abundances of alpha-elements such as Ca and Mg can be accurately calibrated from the HR spectra and allow a direct comparison between the Galactic Bulge and distant galaxies. Another link with the high-z universe comes from the measurement of Zn abundances. These are measurable in QSO absorption lines and probably reflect bulge formation at high-z that can be directly compared to the spatial and temporal distribution of Zn in Galactic Bulge giants covering a

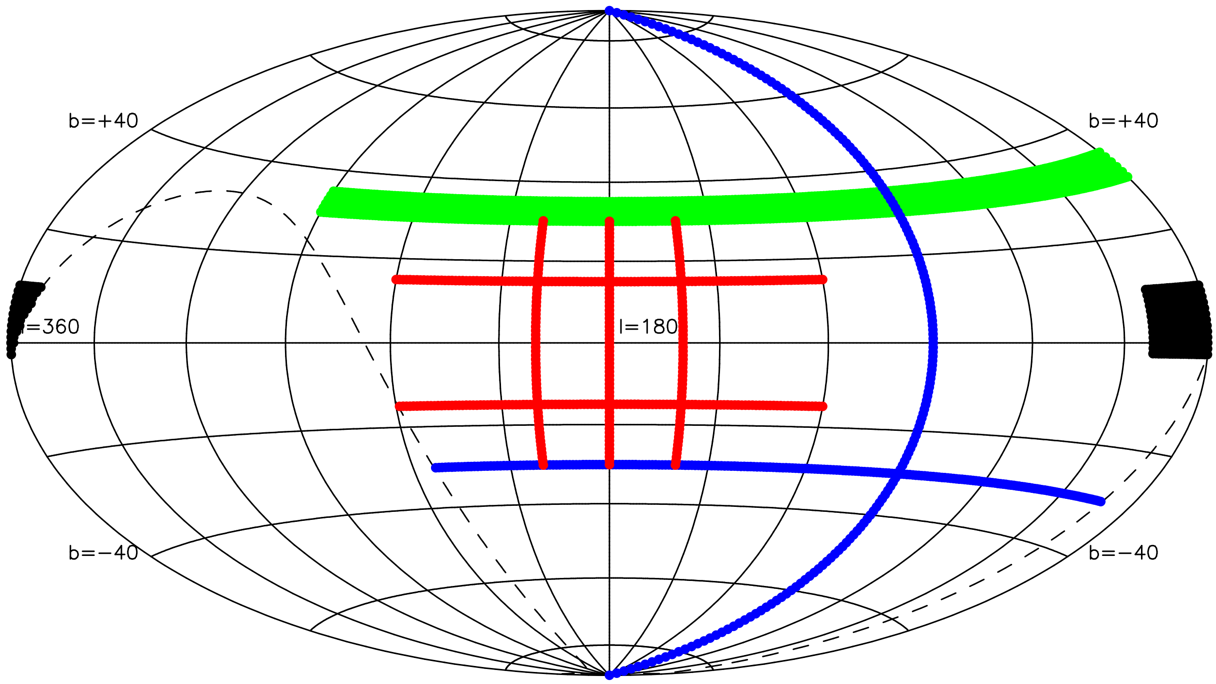

Figure 3: Aitoff projection showing sky coverage of LR and HR surveys - placeholder for now.

range of metallicities and ages. As discussed earlier, the presence of extremely metal-poor stars in the Galactic bulge also provides us with an opportunity to characterise the first stars in the Universe at a level of detail that is impossible for any external system.

[define main science aims and deliverables:

need 1, 2, 5 yr plans with interim deliverables - possibly in a subsection

targetted specifics, survey mode..... need time breakdown

disk and bulge simulations for survey sampling strategy to pick out streams and substructure in disk and outer disk]

3.3 Survey Design

While a Galactic archeology survey with WFMOS encompasses a broad spectrum of goals, it is essential that the survey is undertaken to provide maximum scientific return within realistic time-frames. Given the complexity of the various goals, no single, efficient survey strategy suffices. In the following sections we investigate the survey requirements needed to meet the science goals set out previously, based on current knowledge of the Galactic populations and some numberical simulations.

3.3.1 Streams in the Galactic Halo

The identification of halo streams in important for several of the key science goals of this survey, namely determining the mass and extent of the dark matter halo, as well as the identification of any putative dark substructures. Due to their low projected stellar density ($>$ 30mag/sq arc), it is extremely difficult to identify tidal streams in the Galactic halo within photometric surveys and

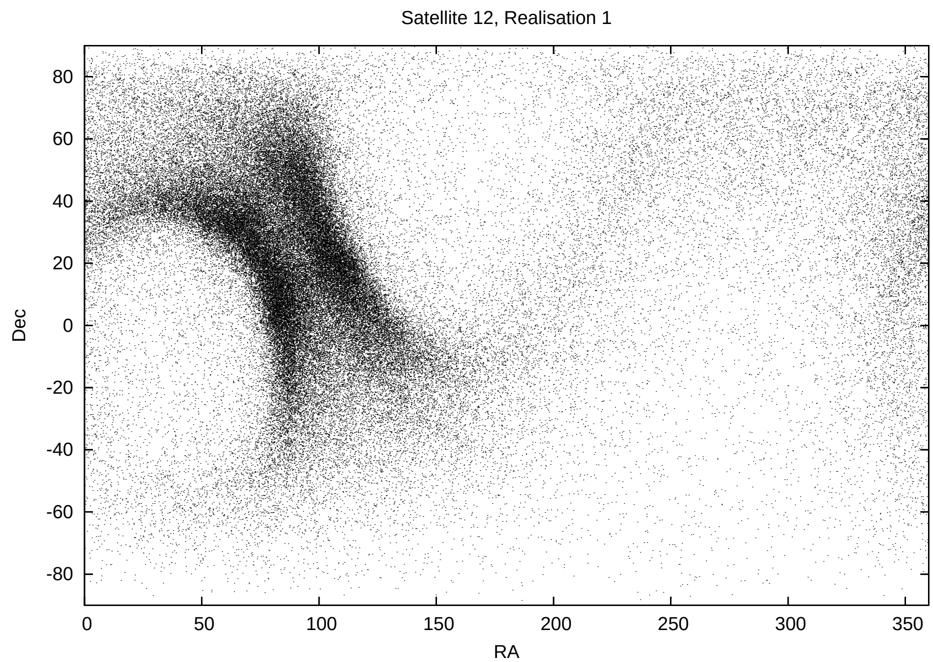

Figure 4: Distribution of debris from an inner halo satellite after 12 Gyr as a function of RA and Dec.

other clues, such as kinematic and chemical signatures, must be used to tease accreted systems from the Galactic populations. Due to their proximity, however, an individual stream covers a huge swathe of sky and hence optimizing a survey to maximize the detection and characterization of a population of streams is not a straight-forward task. In the following, N-body simulations (provided by A. Helmi) are used to construct mock skies such that differing observing strategies can be compared.

The simulations followed the demiseq of satellite galaxies orbiting in a static Milky Way potential for 12 Gyrs. A total of 32 satellites in the inner halo (apocentres $<\sim$ 20 kpc) were followed along with two satellites in the outer halo. All satellites were initially Plummer spheres containing 100,000 particles, the orbital planes were randomly oriented with respect to the plane of the Galaxy and the initial orbits were strongly radially anisotropic; a detailed description of the simulations is given in Helmi & de Zeeuw (2000). We construct an ensemble of debris trails by generating five realisations of each satellite's debris at 12 Gyr, obtained by first rotating in azimuthal angle and then observing the debris from the position of the Sun.

Figure 3.3.1 shows the distribution in RA and Dec of debris from a typical inner halo satellite at 12 Gyr. Orbital plane precession due to the non-spherical components of the Galactic gravitational potential means that the debris is distributed over much of the sky, although the surface density of material is low and highly non-uniform.

In order to establish the number of stars that would be observable in a given survey, we first estimate the fraction of satellite stars that are likely to be K-giants. Using the Besancon model of the Milky Way [2], 10-15 per cent of the halo population are K-giants. In the Besancon model, the halo is represented by a old, metal-poor stellar population with an initial mass function $dn/dm \propto m^{-0.5}$ (Haywood at el. 1997), which is a reasonable first-order approximation to the stellar populations expected in the progenitors of halo streams. In what follows, we conservatively assume that 10 per

[2]Robin et al. 2003; http://bison.obs-besancon.fr/modele/

cent of the stellar content of the satellite will be K-giants which would potentially be included in our survey.

The goal is to detect halo streams *and* to characterise those streams we detect; the characterisation of a stream would include a determination of the integrals of motion (*e.g.* energy E) of the satellite which gave rise to the stream, as well as an estimate of the internal velocity dispersion of the system based on the dispersion in the integrals of motion among stars in its debris trail. Relatively small numbers of stars suffice to *detect* the presence of substructure in velocity space, e.g. Helmi & de Zeeuw (2000) show that the two-point correlation function in velocity space can be used to detect clumping - the presence of substructure shows up as an excess of stars with small velocity differences. Similarly, streams with favourable energies and angular momenta can be detected against the Galactic foreground with small samples of three-dimensional velocities - Helmi et al. (1999) identified a debris clump containing 13 stars (out of a sample of 97). It is therefore assumed that for a robust detection of the phase-space overdensity associated with a particular satellite, approximately 100 member stars would be required. However, to characterise such a population requires significantly more data, and in what follows it is assumed that 1000 satellite stars is the minimum required for characterization.

Of course, all stellar debris in the halo will be observed through a foreground of Galactic stars. While multi-band photometric surveys (*e.g.* 2MASS) can be used to preferentially target K-type stars, the selection will not be completely efficient. Using the Besancon model, we find that at low Galactic latitude (b$\sim$ 0°) there are $\sim 1-2\times10^5$ stars per WFMOS field which satisfy V$<$ 20 and 2 $<$V-K$<$ 3 (this is only an approximate cut to detect halo K-giants, and targeting for the actual survey will use more than one colour (possibly supplemented by parallax information from *GAIA* if available) to perform the selection). For fields with b$>$ 30°, there are $\lesssim$ 4000 such stars per WFMOS pointing. We therefore propose that fields at b$>$ 30° should be the primary focus of the search for halo streams, since in these fields we can target all potential K-star candidates, and hence ensure maximal efficiency of the survey. Naturally, this will mean that low latitude, localised streams will not be detected. However, at low latitude the number of pointings to achieve complete targeting of K-giant candidates becomes prohibitively large - complete targeting is essential in order to ensure sufficient numbers of stars per stream will be observed.

Fig. 3.3.1 shows the result for an ensemble of surveys of 800 square degrees, to a limiting magnitude V= 20 and assuming that Galactic foreground is removed by brute-force observation of every K-giant candidate (*i.e.* b$>$ 30°). For each survey, one realisation of each satellite is selected and the number of stars falling within the survey region is counted. The total number of stars in the satellite is scaled to 10^6 and therefore 10^5 giants will be present. Satellites with only 10^5 stars (*i.e.* 10^4 K-giants) yield too low a surface density of giants to allow characterisation in a survey of 800 square degrees or less, although they can be detected. Results for three survey designs (all with total area 800 square degrees) are shown in Fig. 3.3.1: (a) a 5-degree-wide block (b=$[30°, 35°]$, l= 0° to 180°); (b) a cross made up of two 2-degree-wide strips (l=180°, b=$[-90°, 90°]$; l=$[0°, 180°]$, b=$[30°]$); (c) a 10-degree-wide block (b=$[30°, 40°]$, l= 0° to 90°). In each case, the histograms show the fraction of satellites which are characterised by the survey. What is clear from the figure is that a block design is strongly preferred over a cross design - this is because the surface density of the debris from an individual satellite is low over much of the sky and therefore "clipping" the edge of a debris region generally yields too few stars for full characterisation. Fig. 3.3.1 shows the results for a survey of 400 square degrees. In this case a block design is also preferred - the smaller total area of the survey means that only 10-15 per cent of streams will be characterised, compared to about 30 per cent for an 800 degree survey.

In the outer halo, streams are more spatially confined as the timescale for phase mixing are much longer and precession rates are much lower. A cross design clearly maximises the chances of clipping a stream - however, as discussed above, clipping a stream generally does not provide sufficient

Table 1: The results of our analysis for block surveys, listing the percentage of massive (10^7 stars) and low-mass (10^6 stars) streams which are characterised by block surveys. The 800 square degree survey also allows detection (100 stars per stream) of 50-100% of outer halo streams and $\sim 90\%$ of inner halo streams, although as noted above in-plane accretions are systematically missed.

Block of 800 square degrees (b$\sim 30°$)

$N = 10^7$		$N = 10^6$	
Inner	Outer	Inner	Outer
$50 - 60\%$	$50 - 100\%$	30%	50%

Block of 400 square degrees (b$\sim 30°$)

$N = 10^7$		$N = 10^6$	
Inner	Outer	Inner	Outer
60%	50%	10%	25%

targets to allow characterisation. The exception to this is the case of massive halo streams in which the surface density of targets is sufficiently high that the intersection with the cross survey yields enough members. Fig. 3.3.1 presents the results for a survey of such massive outer halo streams (10^6 K-stars per stream) and shows that a cross survey is indeed preferred. (The results are quantised due to the small number of satellites used – just two per survey). However, we note that the nature of these streams makes them the easiest to detect as they are localised on the sky. It is therefore to be expected that the majority of such streams will have been identified prior to WFMOS using all-sky surveys (see *e.g.* the recent work by Belokurov et al. 2006). We conclude that our survey should not be optimised for the detection of these streams, although we note that WFMOS will be the ideal instrument with which to obtain follow-up spectra in targeted regions of such streams in an efficient manner.

In summary, we conclude that an 800 square degree survey with the fields arranged in a 10-degree-wide block at high latitude (b$> 30°$) will characterise about 30 per cent of all halo streams. We note the this survey could be supplemented by a low-latitude block to observe streams which were missed by the main survey.

3.3.2 Targeting the Dwarfs

To solve the dark matter crisis, it is vital to uncover the true dark matter content of dwarf galaxies. Furthermore, the distribution of dark matter with dwarfs will reveal the presence of central cores or cusps, providing a measure of the fundamental nature of dark matter itself. Such measures require targeted surveys of the known dwarf population......

3.3.3 Targetting M31 & M33

Assuming an end-to-end efficiency of 15%, an 8hr integration for an I=21.5 star (probing 1 magnitude down the red giant branch) will result in a S/N$\approx$5/Å, sufficient to determine velocities to 5 km/s and [Fe/H] (from the CaT proxy) to an accuracy of 0.2 dex. A study with WFMOS would require a targetted survey covering a total of ???sq degrees (2 sq deg/ night) .... for a total of...... Efficient sampling is necessary and this will be driven by the photometric surveys, although we note that the strongest constraint on the shape of any dark matter halo is the gradient along stellar streams and any sparse tiling strategy has to be carefully thought through.

[THERE ARE SOME NUMBERS MISSING HERE eg. TARGET DENSITY IN THE HALO, turns out to be about 1000 /sq deg in the range 20.5 ¡ i ¡ 21.5 and covering a 1/2 mag wide plausible RGB locus ie. plenty SAMPLING STRATEGY, MORE ACCURATE ETC PREDICTION]

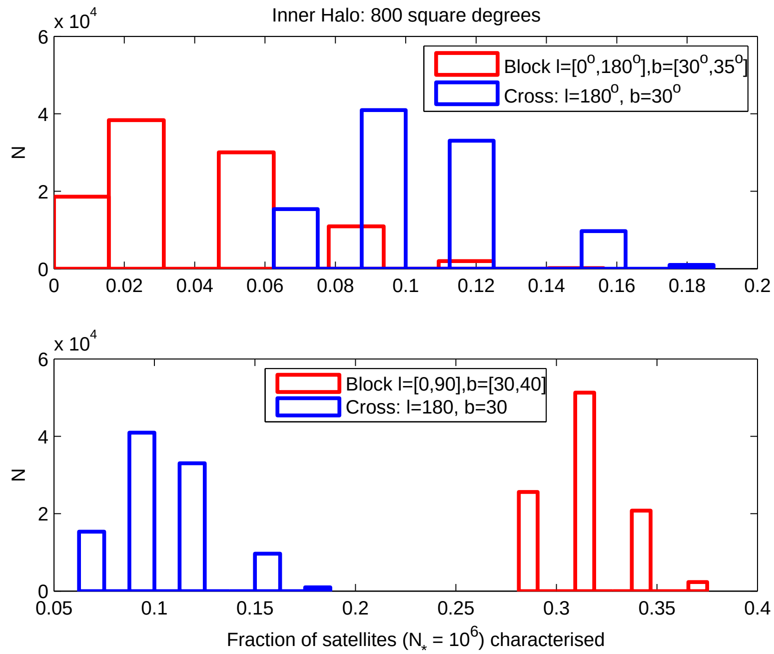

Figure 5: Histograms showing the fraction of inner halo streams with 10^6 stars which are characterised in a survey of 800 square degrees. Results for four different survey designs are shown.

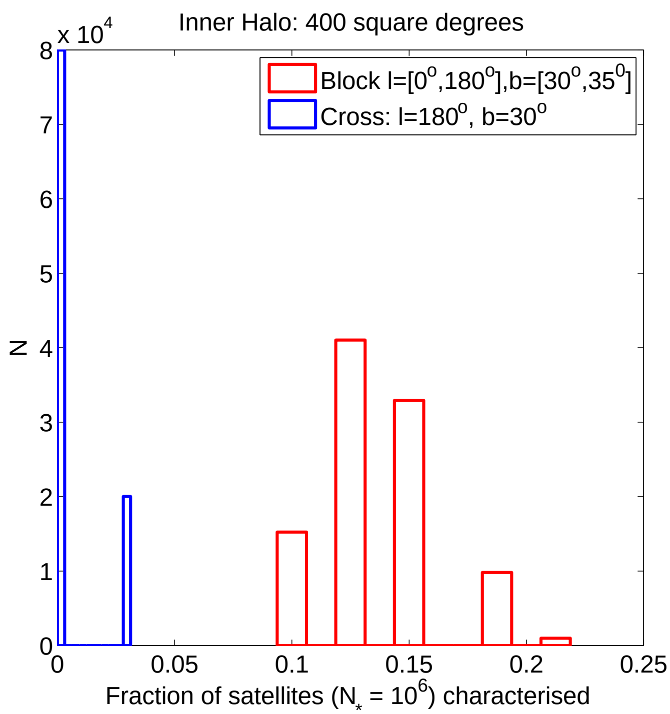

Figure 6: Histograms showing the fraction of inner halo streams with 10^6 stars which are characterised in a survey of 400 square degrees. Results for two different survey designs are shown.

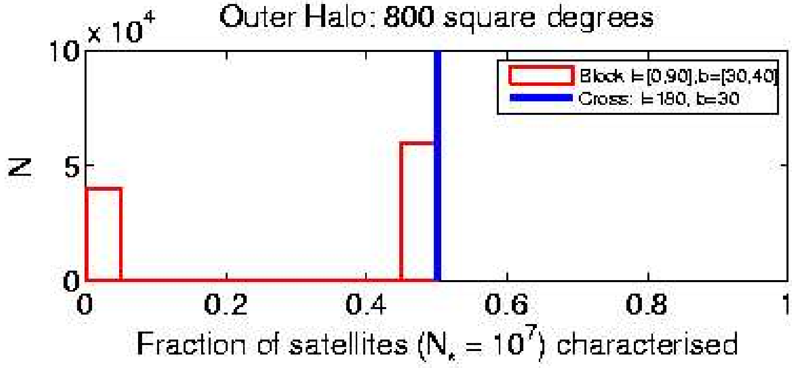

Figure 7: Histograms showing the fraction of massive outer halo streams (10^7 stars) which are characterised in a survey of 800 square degrees. Results for two different survey designs are shown.

3.3.4 The Galactic Bulge

The Galactic bulge provides many clues to the nature of our galaxy. Laid down in the initial phases of the birth of the Milky Way, it is expected that the bulge is home to a large population of the first stars that formed from the proto-galactic cloud. Furthermore, the present spatial and kinematic structure of the bulge, specficially the presence of the bar, has a long range influence on the stellar structure of the Milky Way. Unfortunately, most of the stars, gas and dust in the Milky Way are confined to the bulge and plane of the galaxy,with extinction and crowding making it difficult to unveil the inner structure. Previous surveys have concentrated in the clear "windows", where optical surveys can be carried out (MACHO, OGLE, etc), although success has been varied.

Large scale photometric surveys in the near-IR are will soon be peering through all the gas and dust towards the Galactic centre, with telescopes such as VISTA seeking to map the whole bulge systematically for several epochs in $JHKs$. Covering a total of 360 sq degs, this survey will identify a $\sim 5 \times 10^8$ point sources, with roughly a million variable stars providing distance estimators. While such a photometric dataset will provide a wealth of information on the bulge, including the 3-D structure of the bulge, the identification of the earliest stars and understanding the bulge's dynamical influence will require detailed spectroscopic follow-up.

It is currently unclear whether the inner Galaxy contains a single bar/bulge component formed via the buckling and thickening of a bar (Kormendy & Kennicutt 2004) or whether there is an additional, distinct spheroidal component which built up through accretion (Freeman & Hawthorn 2002). WFMOS can provide a definitive answer to this question by establishing the dynamical state of the bulge using a large sample of bulge stellar velocities and metallicities. By comparing the observed kinematics with the predictions of dynamical bulge models, (including both descriptive Schwarzschild orbit synthesis models, and fully evolutionary N-body models which track the formation of the bar/bulge via disk instabilities), it will be possiblew to understand the detailed properties of the stellar bar/bulge and learn which processes were dominant in the evolution of this region. Currently available kinematical data are totally inadequate for this purpose, and our survey will be the first survey large enough to probe the full kinematic structure of the inner Galaxy.

In order to estimate the number of stars necessary to distinguish different bulge models, kinematical predictions of an N-body stellar bar (Athanassoula 2005) and an axisymmetric bulge model (Dwek et al. 1995) can be compared, revealing that 600 bulge stars are needed to resolve the systematic

10 km/s differences in dispersion between the models over distances of about 1 kpc along a single sightline, allowing for foreground contamination. The limited size of previous surveys has prevented discussion of kinematic variations along the line of sight. As an example, Fig. 1 shows how the models differ in the variation of projected velocity dispersion at low Galactic longitude. Because the kinematics vary on scales of a few degrees, we require the good areal coverage of the region $|l| < 15$ as provided by our proposed sightlines. Further, although the location of the ends of the photometric bar are known from star counts, models of the resonant thickening of bars into peanut bulges predict the presence of resonances located within the bar (e.g. Quillen 2002); - our goal is to map out these resonances through their velocity signatures.

Clump giants will be used for much of the main bulge survey, because they are numerous in most stellar populations older than about 2 Gyr and have a well-defined relation between absolute magnitude and [Fe/H], allowing accurate estimates of their distances. In Baade's Window, about half of the stars with I¡16 are clump giants (Paczynski et al 1999). It is easy to measure their radial velocities and abundances, and verify their gravities using the Ca II 850 nm triplet. Clump stars are ideal for the dynamics, because we can measure their distances, using reddenings individually determined from the 862 nm diffuse interstellar band. The OGLE photometry shows that there are several thousand clump giants per WFMOS field, even in the outer bulge. Also included will be brighter giants, to measure the kinematics of other bulge populations, such as younger AGB stars and very old metal-poor stars in the bulge region. The younger stars are important tracers of any more recent evolution of the bulge. The metal-poor stars in the bulge are particularly interesting, because recent simulations suggest that the earliest pre-galactic stars will be concentrated now in the central regions of the Galaxy.

In terms of understanding its formation and chemical evolution, it is critical to understand the chemical properties of the Bulge. A WFMOS survey will (i) establish the shape of the detailed metallicity distribution function (MDF) in each field, for comparison with chemical evolution models, and (ii) quantify the contribution of the more metal-poor stars to the bulge and determine (in combination with kinematics) whether they are part of the bulge or part of the galactic stellar halo. Existing photometric MDFs do not have sufficient statistics for this purpose: we will dramatically improve this situation with our spectroscopic MDF. The bulge shows a strong vertical abundance gradient, and its overall MDF extends from [Fe/H] = +0.3 to -1.5 (e.g. Zoccali et al 2003). WFMOS will provide not only the mean metallicity but also the distribution of elemental abundances (most notably Fe-peak and alpha-elements) at different locations across the bulge, which will help us understand its chemical evolution. Observing an unbiased sample of giants in 6 fields distributed over the bulge, WFMOS will reveal how the MDF changes relative to the mean abundance gradient.

3.4 Data Analysis

Automated LR data analysis particularly in CaT region well advanced, but still worth applying information theory techniques in this wavelength regime to see if we can extract more information other than velocities and proxy [Fe/H] measurements (*e.g.* Munari et al. 2005).

HR data anlysis is seriously non-trivial. Tools to help automate abundance analysis are required; need to see if ANNs GAs and SoMs etc. are any use or, are we limited to feature extraction based on line lists (*e.g.* EWs) and then modelling/templating (cf. Kurucz model atmospheres as in Munari's extensive library) .... MORE

3.4.1 Stellar atmosphere modelling

General atmospheric parameters: temperature, overall metallicity, systemic velocity, microturbulent velocity, surface gravity

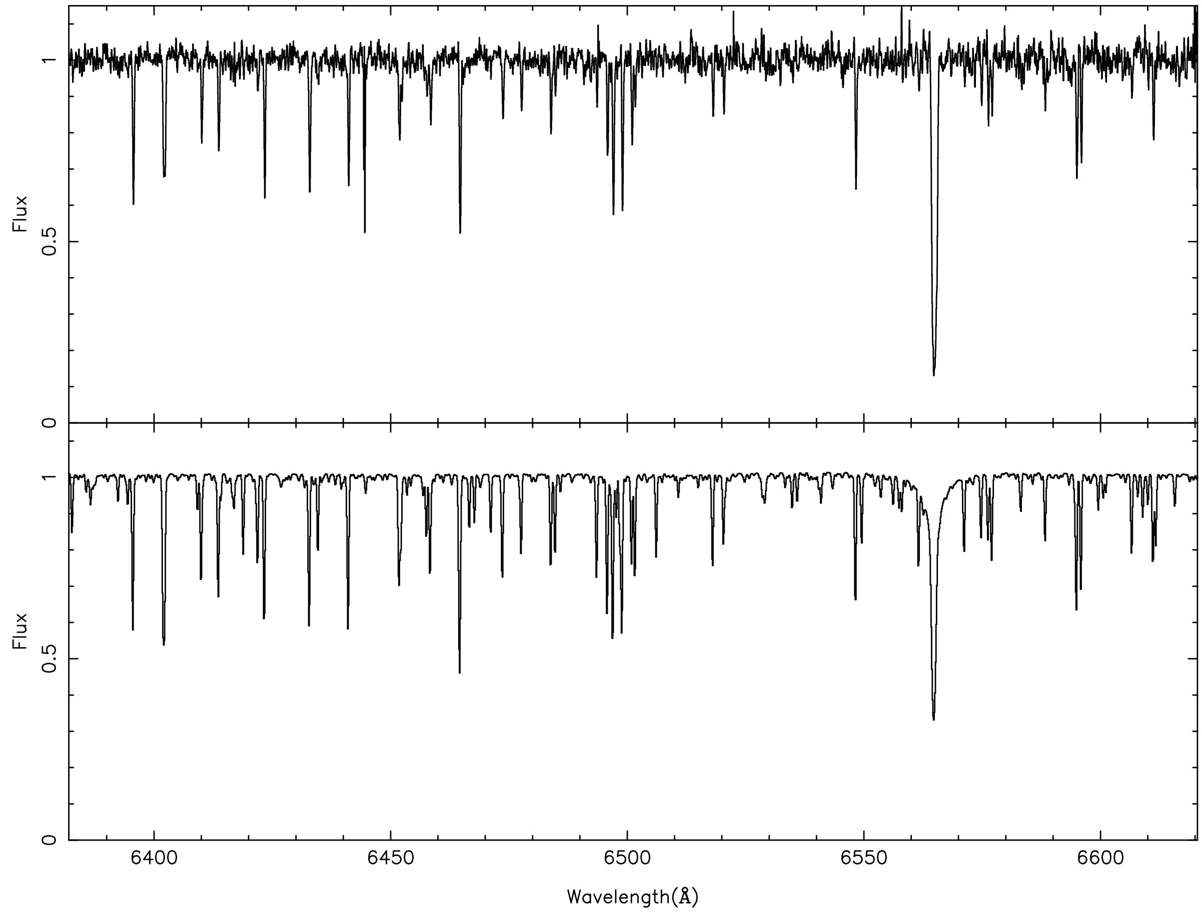

Figure 8: The best fitting continuum-subtracted model spectrum cf. to an HR VLT-FLAMES continuum-subtracted spectrum .... well it will be eventually .... this one is just a random ballpark model spectrum.

Acquiring the Munari et al. grid of UHR simulations .... 50,000 at R 20000

We have the 50,000 template model stellar spectra at R=20000 online at the IoA.

been investiating the on-line web form for generating model spectra

got a set of line lists used for HR analysis Hill, Barbuy, Bensby, Shetrone ......

template fitting using GAs ?

ANNs ?

SoMs ?

need an example of template model fitting - this one is just a place holder example model spectrum cf. HR data.

3.4.2 Chemical tagging

Main categorisation probably along the lines of the following:

light elements (*e.g.* C,N,O,Na,Mg,Al) tracers of deep mixing patterns, found only in globular cluster environments, gives limit to no. of dissolved GCs in a stellar pop - evidence for differences

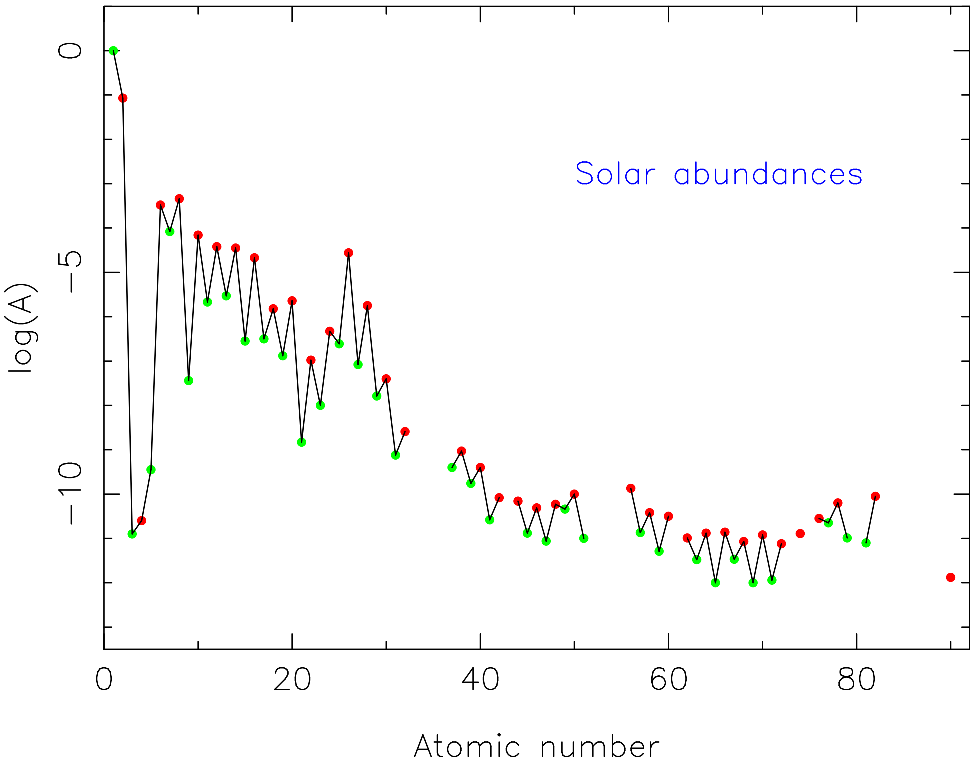

Figure 9: Example of chemical tagging for the solar spectrum

in these ratios (to Fe) between field pop II giants and GC giants (*e.g.* Pilachowski, Sneden and Kraft 1996; Shetrone 1996).

alpha elements (*e.g.* O,Mg,Si,Ca,Ti) occur predominantly in SNeII, abundance of different alpha elements sensitive to mass of SNeII progenitor (even Z) alpha abundance limits no. of SNeII explosions and hence constrains IMF /ejecta fractions with time

heavy elements (Z>30) (*e.g.* Y,Ba,Ce,Sm,Eu) mix of r- and s-process *i.e.* those beyond iron peak formed by repeated neutron capture

rapid capture high energy events such as in SNe explosions *e.g.* Eu thought to be produced almost entirely by r-process

slow capture probably in more quiescent situations *e.g.* from stellar winds during AGB phase typical elements are Ba and La *e.g.* Ba/Eu measure of contribution AGB (can have anomalous via self pollution but causes large Ba/Fe

iron peak elements (V,Cr,Mn,Co,Ni,Cu,Zn) mainly from SNe1a *i.e.* older binary systems so ratio of alpha to Fe gives leverage on SFH Zn is good tie in with high redshift studies - also elements heavier than about Zn are predicted NOT to exist in zero metallicity stars - or those enriched purely by very massive first stars (especially pair production SN) another sign to look for. Cu supposedly produced almost solely in SNe1a

Bear in mind that typical errors are ≈0.1 dex and the spread may only be 0.5 dex max. So this limits the likely detective work to ≈3-5 divisions in each of the above tracer groups and even here many of these characteristics may be correlated.

Excellent articles to get this kind of general information are McWilliam 1997 (ARAA) and Tolstoy et al. (2003)

We are trying to figure out a realistic parameter space for "popn typing" To be conservative maybe we should just consider 3 categories under,over and "normal" in each "bin".

This is for analysis of simulations in an attempt to figure out an optimal observing strategy and a realistic assessment of whether its worth going after the bulge.

need to stress importance of good scattered light properties and fluxing for line indices spectrophotometric calibration

3.4.3 Line lists

4800-6800A good regime *e.g.*

alpha-elements OI, MgI, SiI, CaI, TiI, TiII ....;
s-process elements Y, La, Zr, Ba ....;
r-process elements Eu ....;
light elements NaI, AlI, CI and $C2_{5135,5634}$ ....;
iron-peak elements ....;

FeI FeII OI NaI MgI CaI ScII TiI TiII CrI NiI Yii BaII EuII LaII ZnII, CuI, MnI

see Shetrone et al. 2003 for more details and line lists.

etc ....

[NEEDS TIDYING UP :-)]

3.5 Instrument requirements

Most lines of interest are weak *e.g.* EW $\approx$ 100mÅ. Ideally we want them on, or close to, the linear part of curve-of-growth but this is not always possible so we tend to work with lines in range 10mÅto 250mÅ

Turbulence + rotation velocities broadening of late-type giants typically $\approx$few km/s and hence lines are unresolved at R=20000 and line profile is generally dominated by spectrograph

Early-type stars have velocity dispersions much greater than this (eg. faster rotation velocities) and hence spectra are fully resolved.

Define resolution as FWHM = 2.35σ

e.g. $\sigma_v = 2$ km/s then FWHM = $4.7 km/s$ *i.e.* $\approx$ 0.1Åat 6000Å

and since flux = $I_p\sqrt{2\pi}\sigma = I_p\times$ FWHM $\times 1.07$, implies line saturation (*i.e.* $I_p = C$) when EW $\approx$ FWHM !

exception is CaT lines which are generally saturated all the time

Assume line is weak then continuum noise σ_n (/Å) dominates

$$EW = \frac{I_{total}}{C} \qquad \Delta EW = \frac{\sigma_n\sqrt{w}}{C} = \frac{\sqrt{w}}{s:n}$$

where C continuum level/Å, $s:n$ is signal:noise per Å and w is effective width integrating line over.

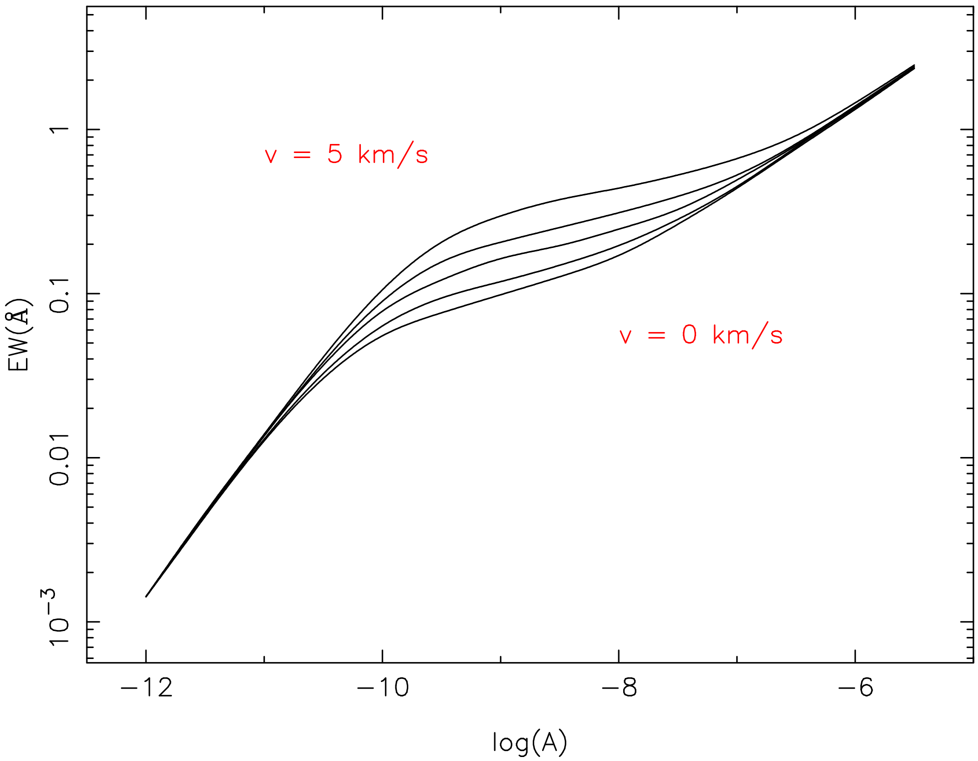

Figure 10: The curve-of-growth of typical weak lines highlighting the critical dependence of the abundance determination on knowledge of the microturbulent velocity.

For a fitted Gaussian profile $w = \sqrt{4\pi}\sigma \approx 1.5$ FWHM hence

$$\Delta EW = \frac{\sqrt{1.5FWHM}}{s:n} = \frac{\sqrt{1.5\lambda/R}}{s:n}$$

primary abundance error driver from measuring point-of-view.

e.g. if dominated by spectrograph R=20000 at 6000Åwith s:n 100/Å→ $\Delta EW = 7$mÅ

issues like r/o noise .... blending .... what is signal for V=17 at R=20000

Integration times - signal:noise - tradeoffs -v- science cf. GAIA

LR survey - survey requirements [Fe/H] to 0.1 dex and velocities to 2-3 km/s

HR survey -

3.5.1 Data reduction requirements

Standard data processing can build on existing tools for basic instrumental signature removal, book-keeping, spectral extraction and wavelength calibration (*e.g.* 2dF, 6dF, WYFFOS, FLAMES....).

Issues:

data formats and FITS structure and HDU content

calibration: flats, arcs, standards

spectra tracing and extraction

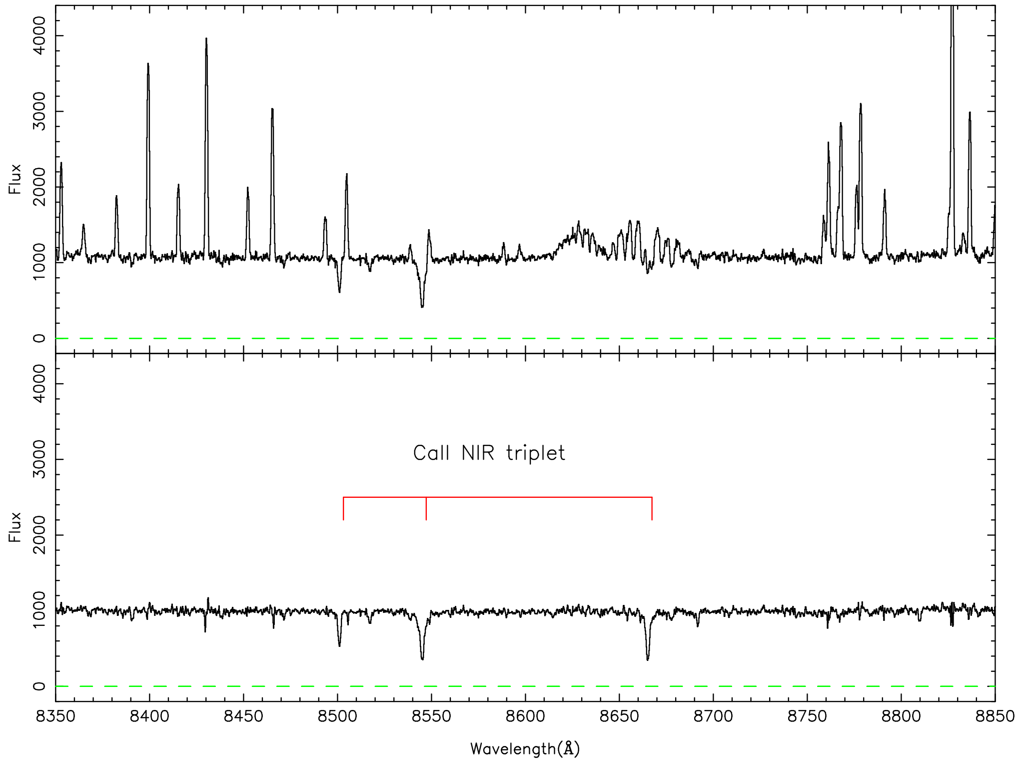

Figure 11: Example of automated LR sky subtraction from FLAMES benchmark test data.

wavelength calibration and verification

bookkeeping: spectra to target information

sky subtraction:

galaxies: redshifts, spectral typing

stars: velocities, abundances, atmosphere modelling (temperature, gravity, microturbulence)

→ automation challenging and crucial

- data reduction strategy

- benchmark data for testing data reduction strategy use VLT FLAMES GIRAFFE data

Nod and shuffle is a disaster area from an efficiency point-of-view since need up to 4x longer integration to achieve same signal:to:noise. (half time on target, twice as much sky variance) Issues here are, can more sophisticated sky subtraction strategies coupled with a more stable fibre-fed spectrograph design eliminate the need ? Tools for more sophisticated sky subtraction *e.g.* PCA, auto-ass ANNs, optimal scaling and so on need trialling on existing data. FLAMES data is probably a good benchmark for tests ?

For a stable fibre-fed spectrograph there are alternative strategies that are worth pursuing and developing further (*e.g.* Tolstoy et al. 2004 with FLAMES data; and Wild and Hewitt 2004 on SDSS data). We are investigating and will report on alternative sky subtraction strategies, building on and extending the above model fitting and PCA decomposition approaches, by making use of Genetic Algorithms and Artifical Neural Networks. Recovering up to a factor 4 in telescope time for a given survey is sufficient motive.

Modelling sky using dedicated sky fibres works well - show FLAMES CaII triplet example ....

- automatic continuum estimation

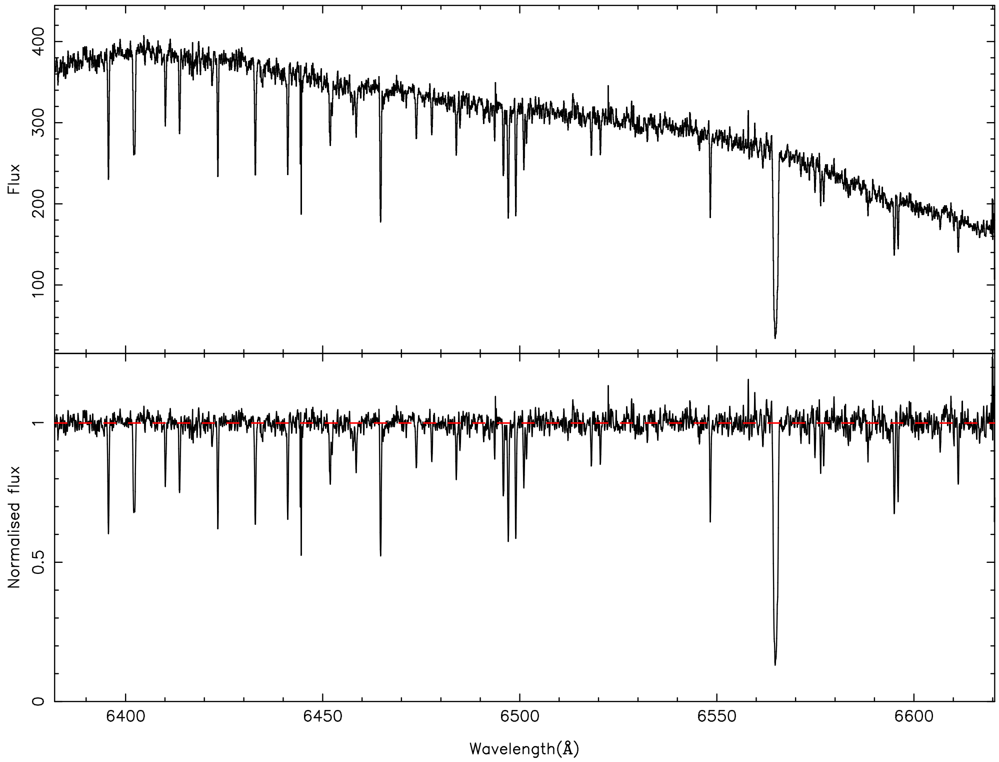

Figure 12: Example of automated continuum estimation and normalistation using FLAMES benchmark HR test data.

3.5.2 Acquisition requirements

light from targets should go down the fibres .... d'er!

3.5.3 Calibration requirements

flux calibration of spectra to take advantage of no cross-disperser can then measure continuum indices to tie in with other measures and inparticular extragalactic spectra (inc. high Z)

scattered light needs quantifying since need good estimate of continuum level for abundance measures

can use extant photometry for absolute calibration

3.5.4 Data products

****Edinburgh team**** Nick Cross was interested in this

3.6 Observing Scenarios

3.6.1 Fibering efficiency

Assume close-packed hexagonal fov for each POSM of area A, and N POSMs for the focal plane area $\pi D^2/4$. Define the fill factor of POSM f to allow for edge mounting supports, glue etc.. $f \approx 0.9$

Let the density of sources on the sky be ρ, then for a random distribution of sources (a good approximation for most Galactic science), the expected number per POSM fov is given by a Poisson distribution with $a = \rho.A$ and hence the probability of zero sources in a POSM fov is

$$P(n = 0) = e^{-\rho A}$$

Typical stellar number densities are given in the table below. So to give a concrete worst case example: consider a 1.5 degree diameter fov with 2700 POSMs/fibers packed optimally. This would imply the available (fixed) area A per POSM is

$$A = f.\frac{\pi/4 \times 1.5^2}{2700}$$

i.e. $\approx 1/1528$ sq deg. or equivalently an effective fibre density per sq deg of 1528. Using the table below, for a latitude $> 60°$ from the Plane the average total density of sources to V= 20 is $\approx$1500 per sq deg, hence $P(n = 0) = e^{-1500/1528} = 0.36$. Even to V= 21 the numbers are $\approx$ 2200 per sq deg and hence $P(n = 0) = e^{-2200/1528} = 0.24$.

For 30° from Plane the total density of source to V= 20 is $\approx$5600 per sq deg, hence $P(n = 0) = e^{-5600/1528} = 0.03$ which is fine.

Repeat this for HR but need to know fibre/POSM choice limitations if any

[NEED TO UPDATE THIS WITH LATEST NUMBERS WHEN THEY HAVE STABILISED]

3.6.2 Surface density of Galactic targets

$\|b\| =$	20°	30°	60°	90°	
V = 17	2290	1318	468	355	HR limit ?
18	4074	2239	741	550	
19	7079	3631	1122	832	
20	11482	5623	1698	1230	LR limit ?
21	17378	8128	2455	1778	
22	23988	11220	3467	2399	

Estimated stellar density per square degree in the V-band

LR survey limit V$\approx$20 HR survey limit V$\approx$17

available survey area $\approx$2 sq deg per pointing

3.7 Observing preparation support

3.7.1 Observation preparation

Presumably generic to both DE, GA and VA?

Each observation will be executed via the eqivalent of phase II software, *e.g.* the Gemini Observing Tool or ESO's P2PP. The simplest path for accomplishing this is to modify the Gemini observing tool for WFMOS, although obviously there may be interface issues with the Subaru hardware.

Prior to inclusion in the phase II software, each field will require a POSM configuration file in the database of survey fields. The software to enable this will be different to past multi-fibre instruments such as 2dF or VLT-FLAMES, but the concept remains the same – a large number of sources are provided as possible targets, and the configuration software is used to assign each POSM. If potential targets are given a weighting factor when selected (by colour or brightness) then in the case of crowded regions the higher priority source can be assigned. Note that even though the number of objects is significantly larger than in 2dF/AAOmega, because of the POSM approach the configuration software represents much less of a challange in terms of optimisation.

Further issues that will be addressed include mechanisms for target selection and the possible need for subsiduary imaging surveys to supplement those that will be available. Homogeneity and statistical balance of sample selection are also key issues that need resolving, as is interplay with other large scale surveys such as GAIA that will be happening concurrently.

3.7.2 Source catalogues

targets - area, sampling, magnitude and colour ranges

Source catalogues - 2MASS is excellent for selecting relatively clean all-sky samples of K- and M-giants out to ≈100 Kpc (though do we want to attempt much more than RVs for later M-type stars ?). SDSS and Segue extension can be used for general selection and by then equivalent VST optical surveys of the Southern sky will exist, complemented by deeper NIR imaging from VISTA and WFCAM. Is this sufficient given there are other all-sky optical catalogues to choose from already, and Pan-Starrs and Skymapper in the wings ?

How much do we want to selected various sub-types ?

Proper motion of fid's = use UCAC ?

astrometric error budget

3.7.3 Planning the observations - ETC

[need realistic estimate of overall QE - possibly rule out some targets *e.g.* M31/M33]

ETC exposure time calculator - the FLAMES VLT one is a good starting point but we need answer to overall throughput efficiency *e.g.* the feasibility study quotes 15% at one point (cf. 10% FLAMES HR, 5% LR at CaT, and 2% UVES - but also cf. Keck DEIMOS 30% and 10m diam).

Exposure times from VLT FLAMES GIRAFFE ETC

Estimated signal:noise LR for 1800s integration at 0.29Å per pixel, resolution 4 pixels *i.e.* 1.2Å, efficiency at central wavelength 4.5%; second part of table ditto but assuming efficiency 15%

Disadvantages M31 project re: Keck and DEIMOS 10m -v- 8m; overall efficiency 30% -v- 15%?; leads to factor of ≈3 extra exposure, also target density in inner regions too high cf. 2 arcmin POSM to really benefit from WFMOS.

Further disadvantage is that Keck studies will arleady have been up and running for nearly 10 years by the time WFMOS gets on sky BUT probably not in the outer M31 regions such as the extended "halo". [HOWEVER, TO REACH V=23 AND GET ENOUGH S:N TO BE ABLE TO MEASURE INDIVIDUAL ABUNDANCES REQUIRES INTEGRATING AROUND 8 HOURS PER FIELD]

V	K giant s:n/pixel	8500Å s:n/Å	G-dwarf s:n/pixel	s:n/Å
17	39.4	73.2	29.7	55.2
20	5.8	10.7	3.8	7.0
23	0.4	0.7	0.3	0.6
17	71.9	133.7	54.2	100.8
20	10.6	19.5	6.9	12.8
23	0.7	1.4	0.5	1.0

Estimated signal:noise HR 5340Å– 5619Å for 3600s integration at 0.07Å per pixel, resolution 4 pixels *i.e.* 0.29Å, efficiency at central wavelength 10%. Need to worry about readout noise *e.g.* 4.2e- per pixel for FLAMES max signal in 3600s is 23e- per pixel. Effect of 15% efficiency is shown approx in rows below.

V	K giant s:n/pixel	5500Å s:n/Å	G-dwarf s:n/pixel	s:n/Å
17	35.1	132.7	35.4	133.8
20	5.8	21.9	5.9	22.3
17	43.0	162.6	43.4	163.9
20	7.1	26.8	7.2	27.3

VLT FLAMES efficiencies for reference

Grating 50% Detector 80% to 50% Camera 80% Fibre 60% Corrector 90% Collimator 95% Filter 85% *i.e.* overall ≈ 10% at best.

Can do better with detectors particularly in red (*e.g.* deep depletion devices) and with holographic gratings so might expect ≈15% overall throughput. After all FLAMES will be 10-year old technology by then !

3.7.4 Fibre allocation and configuration time

Bearing in mind POSM concept implies targets are typically spaced by 2 arcmin with up to 3000 max over 1.5 degree fov. General things to ponder are: target density, sampling stratgey, depth, spectral resolution and wavelength coverage, signal:to:noise, integration times to achieve these, multiplexing *i.e.* simultaneous "red" and "blue" coverage using dichroics (*e.g.* CaT on "red" arm and other interesting regions on "blue"). Acreage of silicon *e.g.* 8kx10k ? deep depletion for good red response, and also require good blue response ? How to populate pixels viz no. of spectra -v- wavelength coverage -v- resolution.

Scattered light and pixel pitch (3 pixels/FWHM)

Fill factor, arrangement of slits sub-slits.

Is target crowding a problem *e.g.* Bulge

if 1 arcsec fibres need overall astrom error budget including targets

3.8 End-to-end observing cycle

3.8.1 Observation configuration

The use case analysis in Section **xx** highlights some of the issues that will have to be tackled for planning and execution of observations. The benefit of two well-defined, large-scale GA surveys with complimentary objectives enhances flexibility of observations, specifically with regard to image quality, sky transparency (cloud cover), and airmass. With prepared observations in the Gemini Observation Database the priorities for each night can be set, both with regard to the status of survey (*e.g.* relative status of LR survey compared to HR, the need to complete a spatial region etc) and the conditions. In simple terms this means observing the faint HR fields in the best conditions, and the bright LR fields in the worst – moreover, the fast reconfiguration time of the POSM means that the strategy can be adapted quickly at the telescope. Particularly in the LR survey, the large surface density of targets (see **table in 3.7.2**) will allow each prepared observation to have a well-defined magnitude range. This will ensure consistency of the final data quality for the spectra within each field, and will simplify decisions regarding execution at the telescope.

Following observation the observed targets can be integrated into the survey database, with a QC flag following pipeline reductions to indicate if the field requires reobservation at a later time. As mentioned elsewhere in this document, the survey database will also require cross-reference with the GAIA archives, conceivably with the prepared observations for the GA surveys drawn from GAIA data and other releases as they occur.

3.8.2 Operational model

A general overview of the operational model is given by the observation execution case analysis in Section **xx**. Daytime calibration files such as biases (and darks if required) will be sufficient, although wavelength calibrations and fibre flatfields may be required for each pointing, depending on the stability of the instrument with reconfiguration. For LR mode we can directly use the sky lines either as a primary wavelenght calibrator or as an extended check on the wavelength calibration. **Accurate wavelength calibration is crucial for successful sky subtraction.**

The two GA surveys are defined in Section **xx**, we now justify our estimated completion times. Assume an average science observation time per night of 8 hours. The POSM reconfiguration time is inconsequential compared to the time for telescope slewing, guide-star acquisition and active control via the WFS. Here we assume an overhead of 15 mins. The Phase II requirements for a new field in GMOS is 20 mins, with the overhead 15 mins for reobservations of a previously acquired field, so our assumption is plausible. Indeed, given that Phase II estimates are generally conservative, we see this as a worst-case estimate of the real timings.

The target densities from Table **table in 3.7.2** demonstrate sufficient sources in our proposed survey areas for at least one WFMOS pointing.

The LR survey requires 0.5 hr exposures of each field. So in an 8 hour night this (conservatively) would yield observations of 10 fields of 0.5 hr + 0.25 hr overheads. Thus, wiith $\sim$3,000 targets per field this gives 30,000 stars per night, giving a total of 100 nights would observe 3×10^6 stars.

Similar arguments can be employed for the HR survey, in which each pointing will require 2x0.5 hr exposures. In 'average' nights 6 fields will be observed, giving over 10^6 stars from 200 nights.

We have rounded the number of fields per night downwards, if arcs are required for each configuration this will not impact on our estimates. Note that median seeing on Mauna Kea is typically $0.2''$ better than that at Paranal, and our exposure time estimates were taken from the FLAMES ETC, with the default $0.8''$ seeing at the VLT. If the initial seeing estimate at the start of the night is $>0.8''$, (within reason) the brighter fields can be selected for that night, with relatively little impact on the science objectives.

A similar strategy will also enable us to take advantage of nights with thin cirrus conditions.

[Figure 34 from the feasibility study suggests that you could do the brighter fields in the thin cloud situation. So assuming it's (usuably) clear 75-80rather than the 1.7 that they use in the original report??]

3.9 End-to-end observing cycle

3.10 Competition

RAVE south V<12 250,000? stars R=8500 8420-8780Å

GAIA all-sky V<15 single pass targetting <17 final combined R=10000 8450-8750Å

[GAIA RVS only potentially achieves reasonable spectra for V=17 by coadding 100 separate scan stripe images at differing orientations (with different spectra overlapping in each case) prior to extracting the spectra and this probably only by mission end (*e.g.* 2020).]

Keck DEIMOS M31, M33 and dpshs vels and ¡[Fe/H]¿

ARGOS is a proposed AAOmega 160,000 star survey with similar resolution to RAVE (10000) but reaching some 5 magnitudes deeper. Souther LR survey of MW Bulge and targetted disk interfaces V < 18 R=10000 5160-5460Åand 8440-8900Å

Sloan plus SEGUE extension survey closer to the Galactic Plane of 200,000 stars has 5 times less resolution than ARGOS so is more limited in scope R=2000

Magellan MIKE southern MW satellites

AAO team - science focus and requirements differ

3.11 GAIA performance

GAIA is designed with the following overall performance goals:

Parallax errors:

V=10	4 μas	10%	at 25 kpc
V=15	10 μas	10%	at 10 kpc
V=20	200 μas	10%	at 500pc

Proper motion errors:

$$\mu\text{(mas/yr)} = \text{v(km/s)} / 4.74\text{*d(kpc)}$$

V=10	4 μas/yr	0.1 km/s	at 5 kpc

V=15	8 μas/yr	2 km/s	at 50 kpc	*e.g.* LMC giants
V=20	150 μas/yr	7 km/s	10 kpc	*e.g.* disk and inner halo dwarfs

Radial velocities:

V=10	1 km/s
V=15	2 km/s
V=17	10 km/s

3.12 Serendipity and Butterfly Statistics

May be useful but needs refactoring ....

We can investigate the problems in surveying for discrete phenomena and estimate the chances of discovering new examples, or classes, of object by borrowing a statistical analogy from biological systems (*e.g.* Thujll & Shipman 1988). In studies of the diversity of a population of butterflies, Fisher (1943) proposed a simple relationship linking the sample size, N and the number of species present, S, such that

$$S = A \times ln(1 + N/A) \tag{1}$$

where A the diversity index, parameterises the tendency of the population to produce separate sub-species. Although, the formula is necessarily a simplification of a complex classification process, Fisher *et al.* argued that for any procedural sub-classification scheme based on combining distinguishing factors multiplicatively, a logarithmic series provides a good practical approximation. In particular, if a similar survey has already been done, the above formula enables the diversity parameter A to be derived and subsequently used in predicting how many further observations are likely to be needed in order to, say, find one new "species".

Naturally, simply looking at more objects is not the only way to find new species or find more examples of rare objects as the following points emphasise:

- increasing N exponentially usually $\Rightarrow$ more sky coverage; this can be a problem at bright magnitudes.
- alternatives are to survey in new parameter spaces, particularly outside the wavelength range normally sampled.
- survey at different epochs, *i.e.* time series sampling, which can give orthogonal information to conventional spectroscopy or broad-band wavelength coverage.
- survey to fainter levels and potentially sample a different spatial volume, *e.g.* larger look-back time for a cosmological sample.
- higher accuracy measurements enable 'normal' object locii to be better defined, hence the parameter phase space volume for unusual objects is increased and they are easier to delineate.

3.13 References

[THIS NEEDS UPDATING]

Battaglia G., et al., 2005, MNRAS, 364, 433
Bica E., 1988, A&A, 195, 76

Binney J.J., Evans N.W., 2001, MNRAS, 327, L27
Bode P., Ostriker J.P., Turok N., 2001, ApJ, 556, 93
Christlieb N., 2003, Reviews of Modern Astronomy, 16, 191
Fellhauer M., et al., 2006, MNRAS, submitted
Fux R., 1997, A&A, 327, 983
Kleyna et al. 2001, ApJ, 564, L115
Kleyna J.T., et al., 2005, ApJ, 630, L141
Majewski S. R., et al., 2005, AJ, 130, 2677
Mateo et al., 1998, AJ, 116, 2315
Moore B., et al., 1999, ApJ, 524, L19
Muñoz R. R., et al., 2005, ApJ, 631, L137
Quillen A. C., 2002, AJ, 124, 722
Wilkinson M.I., Evans, N.W., 1999, MNRAS, 310, 645
Wilkinson M.I., et al., 2004, ApJ, 611, L21
Willman B., et al., 2005, 2005, ApJ, 626, L85
Zoccali M., et al, 2003, A&A, 399, 931