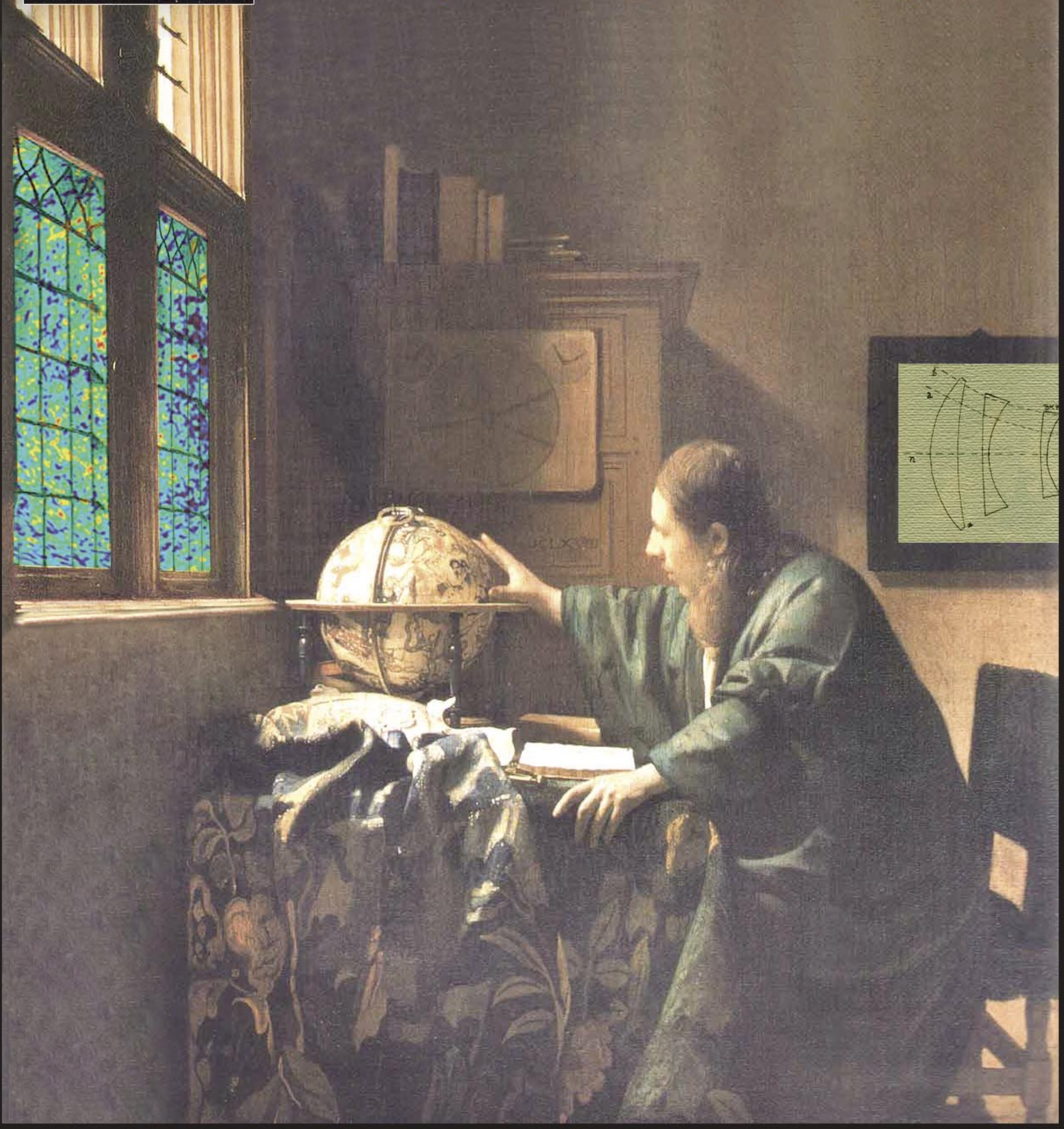
K.A.O.S.

Kilo-Aperture Optical Spectrograph



KAOS Purple Book

Editors:

Arjun Dey
Brian Boyle

Contributing Authors:

Sam Barden
Chris Blake
Joss Bland-Hawthorn
Brian Boyle
Terry Bridges
Michael Brown
Arjun Dey
Daniel Eisenstein
Karl Glazebrook
Francis Keenan
Eric Linder
Anna Moore
David Meyer
Jeremy Mould
Joan Najita
Knut Olsen
John Peacock
Hee-Jong Seo
Alex Szalay
Rosie Wyse

KAOS Science Team

Sam Barden	National Optical Astronomy Observatory
Chris Blake	University of New South Wales
Joss Bland-Hawthorn	Anglo-Australian Observatory
Brian Boyle	Anglo-Australian Observatory
Michael Brown	National Optical Astronomy Observatory
Matthew Colless	Mt. Stromlo Observatory
Warrick Couch	University of New South Wales
Len Cowie	Institute for Astronomy, University of Hawaii
Arjun Dey	National Optical Astronomy Observatory
Daniel Eisenstein	University of Arizona
Ken Freeman	Mt. Stromlo Observatory
Karl Glazebrook	The Johns Hopkins University
Buell Jannuzi	National Optical Astronomy Observatory
Francis Keenan	Queens University, Belfast
Rolf Kudritzki	Institute for Astronomy, University of Hawaii
Eric Linder	Lawrence Berkeley Laboratory
Anna Moore	Anglo-Australian Observatory
David Meyer	Northwestern University
Jeremy Mould	National Optical Astronomy Observatory
Joan Najita	National Optical Astronomy Observatory
Robert Nichol	Carnegie Mellon University
Knut Olsen	National Optical Astronomy Observatory
John Peacock	Royal Observatory Edinburgh
Thaisa Storchi-Bergmann	Universidade Federal do Rio Grande do Sul, Brasil
Steve Strom	National Optical Astronomy Observatory
Nick Suntzeff	National Optical Astronomy Observatory
Alex Szalay	The Johns Hopkins University
Rosie Wyse	The Johns Hopkins University

Executive Summary

What is the equation of state of dark energy? How does it vary with redshift? What formation mechanisms result in the observed structure of our Galaxy? Why is our Galaxy different in structure from other local group members? How does galaxy formation and evolution depend on the formation and evolution of large scale structure? How was the Universe reionized? What mechanisms govern the creation and dispersal of "metals" in the Universe?

The nature of these questions illustrate that observational astronomy has recently matured from a largely discovery oriented science, where astronomers discovered new phenomena in the Universe, to a physical science, where our goal is to understand the astrophysical processes that are responsible for the observed phenomena. We now realize that several central issues regarding the geometry, structure and evolution of the Universe and its observable contents are intrinsically complex, since many diverse astrophysical processes contribute to global observable trends. The result of this complexity is that the answers to many of the Big questions must be approached statistically, with samples large enough to permit us to determine from the observables what are the driving astrophysical processes and isolate the primary processes responsible for the global trends. Albeit complex, these answers are now within our grasp: the technological advances of the last decade in telescope construction, spectrograph design, detector technology, and computing power, place us at a unique moment in history where astrophysical questions can be addressed using a philosophically different approach.

A new facility capable of taking detailed spectroscopy of millions of objects in the Universe overcomes the statistical complexity of the Universe and allows us to answer these fundamental questions. In this document we present the scientific case for and technical feasibility of a wide-field, highly multiplexed multiobject fiber spectrograph for one of the Gemini telescopes: the Kilo-Aperture Optical Spectrograph (KAOS).

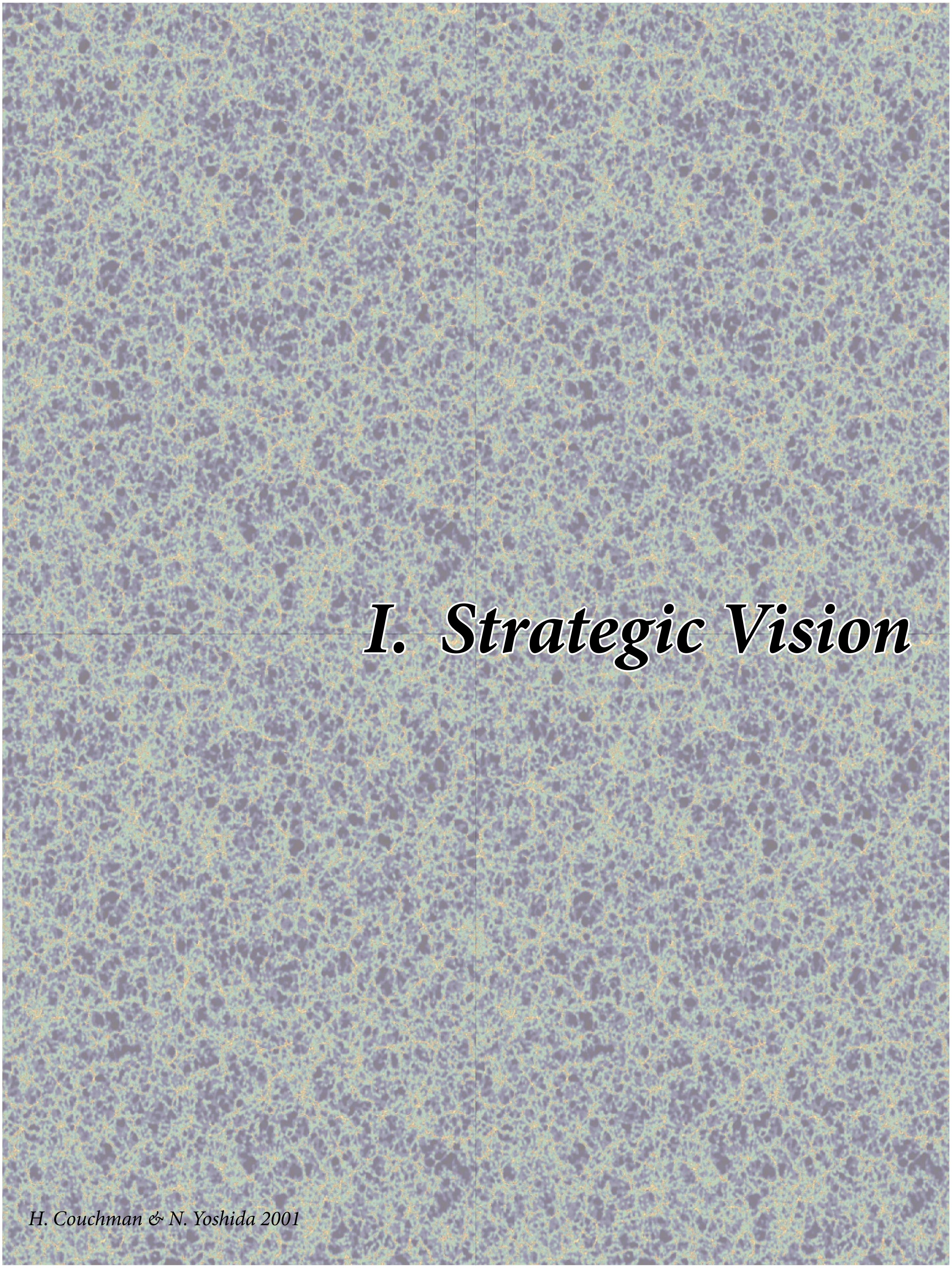
KAOS provides Gemini with the capability of simultaneously obtaining moderate to high-resolution (R=1,000-40,000) spectra of almost 5000 targets in a field of view of 1.5 degrees in diameter. This capability would enable Gemini to be uniquely capable, providing the Observatory with a 10-100 advantage over existing and planned multiobject spectrographs. KAOS will deliver of order 20,000 astronomical spectra per night!

The two primary science drivers for KAOS are (1) the determination of the equation of state of dark energy, and (2) the study of the origin of our Galaxy. The first is a project to use the acoustic oscillations in the baryons to directly measure the time variation in the equation of state of dark energy. This is best undertaken by measuring the acoustic oscillations in the linear regime, where one can obtain a direct metric measure of the angular diameter distance by comparison to the oscillations seen in the CMB at the last scattering surface. This requires very large galaxy redshift surveys at $z\sim1$ and $z\sim3$ which only become practical with the instrument described here. The second key project is aimed at studying the genesis of our Galaxy's old thin disk, thick disk and halo by mapping the kinematics (to < 3 km/s) and abundances (to 0.1 dex) of a million stars. These two projects, and several other scientific studies that motivate and are enabled by this new capability are described within these pages (Part II).

Part III of this document outlines the technical feasibility of KAOS. KAOS is mounted at the prime focus. It contains a 4-element corrector, an atmospheric dispersion correcting prism assembly, and an Echidna-style fiber optic focal plane. The ADC doubles as a wobble plate, which provides a fast guiding capability and the ability to beam switch. Sky subtraction with KAOS is done using the nod-and-shuffle observing mode, which obviates many of the traditional limitations of fiber spectrographs. Configuring the fibers is done through a novel approach that images the focal plane and allows configuration of the ~5000 fibers in less than 10 minutes. The fibers feed to an array of 12 spectrographs in the pier of the telescope. The Gemini telescopes were designed to have a replaceable top-end. KAOS is a prime-focus instrument that will require a new top-end structure which holds the corrector and the fiber assembly.

We are now entering a new era in multiwavelength astrophysical studies. A large variety of sensitive instruments both in space and on the ground will map the sky to unprecedented depths resulting in very high source densities of astrophysically important objects. In order to realize the full scientific potential of these missions, we will need a highly multiplexed spectroscopic capability even to follow-up representative samples of these sources. The current suite of 8m class telescope, adapted as they are for small field high spatial resolution studies, will undoubtedly carry out epochal explorations of the multiwavelength universe uncovered by these surveys, but will barely begin to realize the scientific potential of these missions. As we have learned over the years, the most interesting (and cosmologically relevant) astrophysics will undoubtedly lie at the limit of our surveys! KAOS equips the Gemini Observatory to stand alone in its ability to maximize the scientific yield during this new era. KAOS is not simply the next step in our exploration of the Universe and its complexity — it is demanded by it.

I. Strategic Vision



H. Couchman & N. Yoshida 2001

Chapter 1

KAOS and Gemini:

A Spectroscopic Facility for the New Millennium

At the start of the third millennium, astronomy has come of age as a physical science. Fuelled by major advances in technology, observational experiments can now aspire to produce spectroscopic data sets of the size and quality required to test the most sophisticated physical theories for our understanding of the Universe and its contents.

Astronomers now know that many key scientific questions regarding the structure and evolution of the Universe and its contents are intrinsically complex, with many diverse astrophysical processes contributing to the observable trends. Consequently, many of the Big Questions must be approached statistically, with spectroscopic samples large enough (10^5–10^6 objects) to disentangle the individual physical processes responsible for the trends seen in the data.

In this document we present the case for the Kilo-Aperture Optical Spectrograph (KAOS), an instrument for one of the state-of-the-art Gemini telescopes which will yield a two order of magnitude gain in information-gathering power over that provided by extant or planned facilities. KAOS will uniquely position the Gemini community to construct the large spectroscopic samples required to address many key questions in astronomy at the start of the next decade.

1.1 A Vision for Gemini

The Gemini Observatory holds a unique place on the international astronomical map. It is the only ground-based facility that is jointly owned by countries on four continents. As such, Gemini can tap into a vast international talent base and plan large collaborative projects for the future of a scale that would be unthinkable for the individual researcher, small research group or individual nations.

In the few years since First Light, the Gemini Observatory has served communities previously deprived of large telescope access. Gemini's success to date and, indeed its continued vision for the future, has been to focus on areas of astronomical parameter space unexplored by other 8m-class telescopes due to their technical limitations. However, over the next few years the astronomical landscape will continue to change rapidly. By the end of this decade, there will be 10 other operational telescopes with apertures in excess of 8-m, three more with apertures of 6.5-m, and a further two more >8-m telescopes under construction. There also will be competition from space: the James Webb Space Telescope (*JWST*), a 6-m successor to the Hubble Space Telescope, will be launched near the end of this decade, making obsolete much ground-based high angular resolution imaging in the near-infrared.

As well as competition, the development of new astronomical facilities also brings opportunity. Over the next few years, there will be a dramatic increase in the growth of dedicated mid-size facilities producing all-sky optical/infrared imaging surveys designed for spectroscopic exploitation by the 8-m telescopes. New space missions such as *XMM*, *GALEX*, *SIRTF*, *GAIA* and *NGSS* will further extend imaging databases of astronomical objects to unprecedented wavelength coverage, size and depth.

In this landscape, how should the Gemini Observatory position itself in order to excel?

The answer proposed in this document is for Gemini to *dominate* the field with a new wide-field, multiobject spectroscopic capability, such as the one provided by KAOS.

Clearly Gemini must also retain its unique capability for diffraction-limited imaging in the near-infrared. Development of the multi-conjugate adaptive optics (MCAO) system is consistent with maintaining Gemini's competitive edge in the area. Nevertheless, Gemini serves a broad community which also includes many world-leading groups in the area of wide-field survey science across the Gemini partnership. Many of these groups have played a key role in recent years in delivering high-impact science results from major surveys such as CFRS, 2dF and SDSS. These surveys have themselves been facilitated by the world-renowned strength in innovative instrumentation within member countries of the Gemini project.

The research strength across the partnership in survey science and instrumentation, coupled with the availability of next-generation optical/IR imaging surveys, provides Gemini with the opportunity to develop an important new and unique scientific direction, complementary to the MCAO initiative, in the area of highly multiplexed multi-object spectroscopy.

In this document, we sketch out the concept for KAOS, an instrument designed to provide Gemini with almost a hundred-fold advantage in multiplex-gain/field-of-view over any other facility, existing or planned (see Table~1.1 and Figure 1.1). As a direct result of this enhanced capability, the Gemini Observatory can play a key role in addressing a wide variety of astrophysical problems that cannot be even attempted by any other 8-m class facility. This document maps out the science cases for a number of these projects including two potential 'flagship programs'—the Dark Energy project, to investigate the form of the equation of state of the dark energy; and the Galaxy Genesis project, to provide a complete history of the formation of the thick disk and halo of our Galaxy. These two proposed studies strike at the heart of key questions in astrophysics fundamental to our understanding of the evolution of the Universe and the formation of our Galaxy, and are discussed in some detail in the following chapters.

Table 1.1: Existing & Planned Optical Spectrographs on Large Aperture Telescopes

Telescope	Aper. (m)	Spectrograph Name	Slit or Fiber	Resolution ($\lambda/\Delta\lambda$)	η[1]	FOV (sq deg.)	N_{Max}[2]	ε[3]	μ[4]
AAT	3.9	2DF	f	600-3000	0.05	3.14	400	1.0	1.0
Magellan	6.5	IMACS	s	2000-20000	0.3	0.20	600	25	1.1
MMT	6.5	Hectospec	f	1000	0.24	0.78	300	10	3.3
MMT	6.5	Hectochelle	f	32000	0.10	0.78	300	4.2	1.4
MMT	6.5	Binospec	s	1000(-3000)	0.40	0.067	<150	8.3	0.5
Gemini	8.1	GMOS	s	600-3700	0.40	0.0084	100	8.6	0.1
VLT	8.2	VIMOS	s	180-2520	0.30	0.062	<750	50	0.5
Subaru	8.3	FOCAS	s	250-2000	0.30	0.01	50	3.4	0.1
HET	9.2	LRS	s	550-1300	0.40	0.0035	13	1.5	0.05
Keck	10	LRIS	s	300-3000	0.35	0.010	20	2.3	0.15
Keck	10	DEIMOS	s	1700-4800	0.35	0.022	80	9.2	0.3
GTC	10.4	OSIRIS	s	500-2500	0.27	0.012	20	1.9	0.15
Gemini	**8.1**	**KAOS**	**f**	**1000-30000**	**0.3**	**1.77**	**4000**	**260**	**15**

[1] Optical throughput of spectrograph and telescope.

[2] Maximum number of objects which can be simultaneously observed.

[3] Survey efficiency relative to 2DF, defined as $(\text{Aperture})^2 N_{obj} \eta$

[4] Sky mapping efficiency relative to 2DF, defined as $(\text{Aperture})^2 (\text{FOV}) \eta$

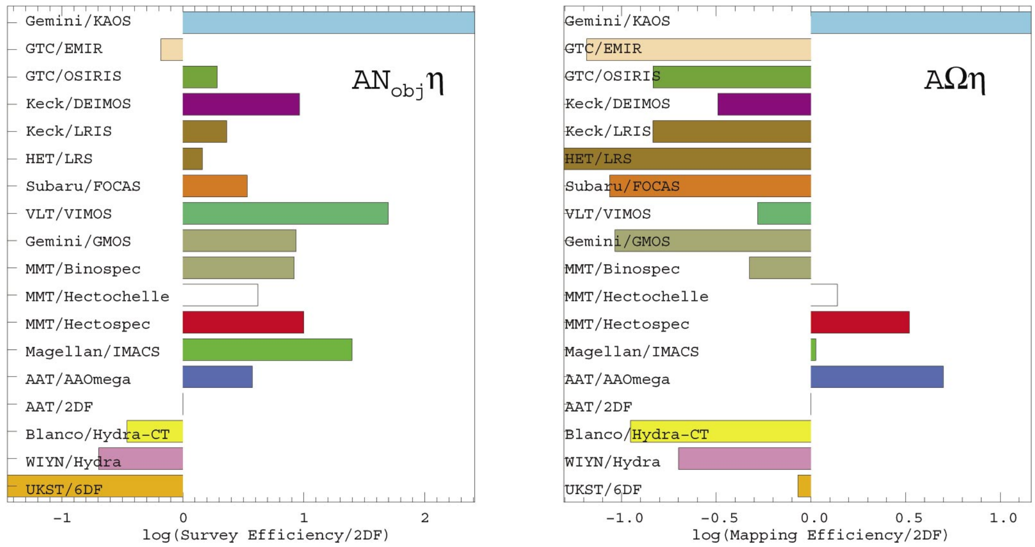


Figure 1.1: The survey efficiency (left) and mapping speed (right) of KAOS and other existing and planned optical multiobject spectrographs relative to 2DF.

1.2 Gemini and the 2010 Astronomical Landscape

JWST

The James Webb Space Telescope, the 6-m successor to the Hubble Space Telescope, is now in its design phase and is expected to be launched near the end of this decade. *JWST* will provide excellent diffraction-limited images with high Strehl (≈0.8), and will operate from L2, far from the thermal environment of the earth-moon system. It will therefore perform exquisitely at infrared and thermal IR wavelengths.

Currently, both Gemini telescopes are optimized to deliver excellent image quality and for observations in the thermal infrared. In the current decade, this is a strategically good choice, since it provides Gemini with a niche in which the Observatory can excel in the pre-*JWST* era.

However, what role does Gemini have in the *JWST* era? The current situation with an operational *HST* is a good comparison. *HST*, a 2.2-m telescope, has made obsolete most high resolution optical imaging on comparably sized telescopes, and forced most 4-m class telescopes into modes which excel at carrying out wide-field imaging and spectroscopic surveys (a regime in which *HST* cannot compete). *HST* has also had a strong synergy with ground-based 8-m class telescopes (and, to a lesser extent, with 4-m telescopes) which have provided spectroscopic observations that both support *HST* and have been complementary to it. For example, ground-based telescopes have been invaluable for acquiring spectroscopic redshifts for targets in the Hubble Deep Field; and both the VLT and Keck telescopes have recently implemented wide-field spectroscopic modes (VIRMOS and DEIMOS) which are carrying out wide field surveys of galaxy evolution and cosmological studies which are very complementary to the deep, narrow-field studies conducted by *HST*.

A similar approach based on the availability (towards the end of the decade) of the next generation of all-sky, deep (B<25 mag) imaging surveys will be extremely successful for Gemini. Indeed, the Gemini Strategic Plan (Mountain 1999) already proposes that the f/6 wide-field mode (with a maximum field of view of 45 arcminutes) be implemented on one or more of the Gemini telescopes, and that the mode be used for 'deep wide-field surveys' in observing campaigns lasting as long as a semester. As we shall see below, the KAOS proposal not only yields a more effective & unique capability than the proposed f/6 wide-field, but one that is also simpler and more flexible to operate as part of Gemini's instrument suite.

Other Space Missions

Other space missions may provide an important complementarity with the new directions implicit in the KAOS proposal for Gemini. There is increasing scientific interest in the investigation of galaxy formation through the detailed kinematical and chemical study of stellar populations in our own Milky Way galaxy (Freeman and Bland-Hawthorn 2002). The *GAIA* mission (launch date: 2012) is specifically designed to make a giant step forward in the study of stellar populations in our Galaxy; providing an extraordinarily detailed map of 1 billion stars in the Galaxy and beyond. This will include accurate proper motions for all stars down to V=20 mag and radial velocities/spectra for all stars to V=18 mag. The proposed Galaxy Genesis project with KAOS will provide a complementary deeper study of the stars within our own Milky Way Galaxy and other local Group members. In combination with *GAIA*, it will provide the most comprehensive and detailed picture of our Milky Way Galaxy ever produced.

In addition, by 2010, a host of NASA missions will be generating images of the sky at wavelengths from gamma rays to the far infrared. Some of these missions will provide all-sky surveys, whereas others will provide deep images of 'small' (~10-100 square degree sized) regions of the sky. For example, *GALEX* (UV) and *NGSS* (mid-IR) will conduct deep (<25 AB mag) imaging surveys over wide areas of sky (>100 square degrees) in passbands outside the traditional optical/near-IR regime. These new missions will result in detections of typically 10^3 to 10^4 objects per square degree. To fully realize the NASA investment, spectroscopic follow-up studies from the ground are critical.[1]

Other Ground-Based Survey Facilities

Notwithstanding *JWST*, a highly multiplexed survey capability on an 8-m class telescope will be essential to maximize the scientific return from planned ground- and space-based survey instruments. On the ground, several observatories are already implementing wide-field imaging cameras to carry out surveys specifically matched to the spectroscopic limit of 8-m class telescopes. Indeed, in the last decade and continuing into this decade, there has been a spectacular growth in wide-field ground-based imaging facilities. Almost all the 4-m class telescopes have implemented wide field cameras (see Table 1.2 for an incomplete list) and many new facilities are being planned or are in development. By the end of the current decade, even prior to LSST becoming operational, nearly the entire sky will have been mapped at least once using CCDs (Figure 1.2).

A key aspect of many of these new surveys is not only their depth but also their areal coverage. In both hemispheres, at least one multi-thousand square degree survey is planned in both the optical (North: PANSTARS/LSST, South: VST) and the infrared (North: UKIRT-WF, South: VISTA) yielding a database of the entire sky to the spectroscopic limit of 8m-telescopes by the end of the decade. For many key scientific applications — either from a perspective of statistics (e.g., cosmic variance) or rarity (e.g., z > 7 AGN) — the ability to carry out spectroscopy of faint targets over many hundreds of square degrees will be an important factor. The wide field of the Gemini/KAOS combination would be uniquely powerful in this regard; enabling groundbreaking studies on a scale that will not be possible with any other facility (or even all of them combined).

[1] Spectroscopy of large numbers of sources detected at other wavelengths is crucial to quantifying their contribution to the overall source population and extragalactic backgrounds and understanding their cosmological significance. In addition, the most unusual source, the one which often provides a critical cosmological clue, is invariably hidden in the vast number of sources at the catalog limit. For example, IRAS FSC10214+4724, was only discovered by spectroscopically surveying a very large sample of faint IRAS sources.

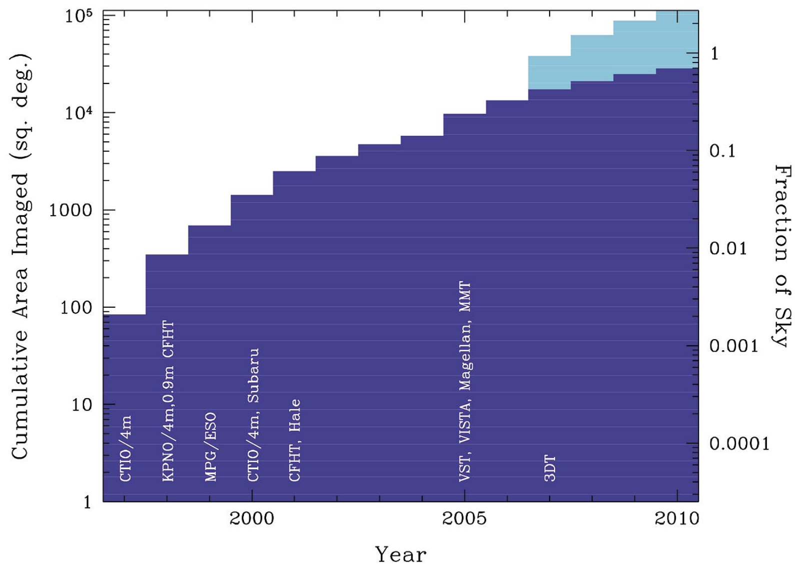


Figure 1.2: The equivalent area covered to R=26 AB mag (5-sigma, 2 arcsecond aperture) in a single filter assuming that 20% of the time is used for imaging. (For VISTA, VST and LSST, 60% is assumed). The entire sky will be surveyed once to this depth by 2010.

Table 1.2 Existing & Planned Optical Wide-Field Imaging Cameras

Telescope	Aper.(m)	Camera	FOV	Years of Operation	Area Covered by 2010[1] (Sq. Deg.)
Mayall	3.8	MOSAIC	36'x 36'	1998-2010	1240
KPNO	0.9	MOSAIC	59'x 59'	1998-2010	220
Blanco	3.9	BTC	4x15'x15'	1997-1999	250
Blanco	3.9	MOSAIC	35'x 35'	2000-2010	1050
CFHT	3.6	UH8K	29'x 29'	1998-2000	200
CFHT	3.6	CFHT12K	42'x 28'	2001-2005	480
CFHT	3.6	Megacam	1 sq. deg.	2006-2010	1480
Hale	5.0	LFC	23'x23'	2001-2010	720
Subaru	8.3	SUPRIME-CAM	34'x27'	2000-2010	3960
VST	2.5		60'x 60'	2005-2010	2302
VISTA	4.0	IRCAM	30'x 30'	2006-2010	5000
Magellan	6.5	IMACS	27'x 27'	2005-2010	1050
MMT	6.5	Megacam	24'x 24'	2005-2010	700
LSST	6.9		180' dia.	2007-2010	2x42000
PANSTARRS	2.5		3° dia.	2005-2010	28000
UKIRT	4.0	WFCAM	24'x24'	2004-2007	7000 (K<19)
					35 (K<21)

1. Equivalent area covered to R=26 AB (5σ, 2" aperture) in a single filter assuming 20% of time is utilized for imaging. In the case of the VLT Survey and LSST 8.4m imaging telescopes, it was assumed that only 60% of the time was used for imaging. The benchmark used was that the Mayall + MOSAIC reaches this depth per field in 2 hours (including overheads), and therefore covers roughly 1.4 sq. deg. per clear night.

In concert, wide-field surveys of the sky using both space-based and ground-based facilities will cover significant areas of the sky, yielding large numbers of sources requiring spectroscopic follow-up. Gemini, with KAOS, could position itself uniquely to exploit this astronomical landscape with unprecedented efficiency.

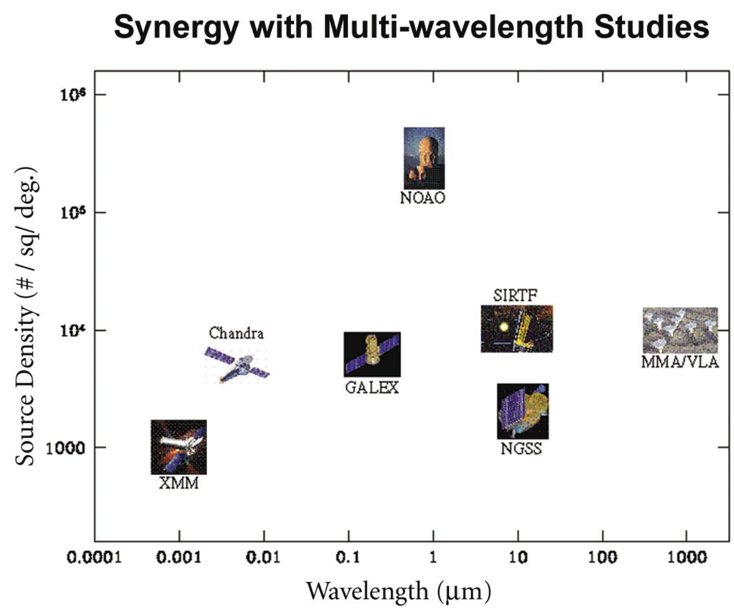


FIgure 1.3: Wide-field surveys from the X-ray to the radio now yield surface densities of objects exceeding 1000/□°

1.3 Gemini science operations

Gemini is highly focused on a cost-effective approach to science operations; scheduling its telescopes effectively to maximize the scientific output. KAOS would provide the final realization of this operational model by providing large scientific yield from a single instrument operated in a campaign mode. Such an operational mode would also enable partner countries with currently small shares of time to access a large data product from the Observatory which would enable scientific programs not otherwise feasible by members of these nations.

Equally, it is also important to recognize the key role that KAOS would have in the facilitation of smaller programs. KAOS must be viewed as a general-purpose user facility as accessible and relevant to the science needs of individual users/small groups with specific programs as to large Gemini-wide consortia behind mega-surveys. In this regard, the use/operation of KAOS is analogous to the operation of 2dF on the Anglo-Australian Telescope.[2]

The Gemini International Partnership is unique in spanning four continents. Moreover, groups in many of the Gemini partner countries represent the strongest in both the technical development and scientific exploitation of MOS studies. There can be no better group or partnership to undertake a project of this scale and vision.

[2] To date, over 30 different science programs have been scheduled using 2dF on the AAT, with less than half of the total time on 2dF having allocated to the major galaxy and QSO survey programs.

1.4 Complementarity to Gemini instrumentation

Gemini has a world-leading capability in near-infrared imaging. KAOS is a complementary not competing initiative, which seeks to add an important new scientific capability to Gemini in its second decade of operation. KAOS would form part of a pro-active approach to position Gemini and its community in pre-eminent position to take full advantage of the astronomical landscape at that time..

Indeed, the initial vision for Gemini was to have a removable top end to implement the f/6 wide-field capability at the Indeed, the initial vision for Gemini was to have a removable top end to implement the f/6 wide-field capability at the Cassegrain focus. Our goal has been to reinstate this vision—but with an expanded capability. To that end, the concept design for KAOS has specifically addressed the key issue of operability within the existing multi-instrument suite. As it transpires, KAOS provides for a much simpler changeover between instrumentation than the original f/6 top end, so that it may be viewed as *extending* the capabilities of Gemini's instrumentation suite, rather than replacing it.

1.5 New scheduling opportunities

By routinely providing over 20,000 spectra per night, the multiplex power of KAOS is capable of delivering major scientific gains in a number of key areas with short runs, best suited to Gemini queue-scheduled mode of operation. However, many of the 'killer applications' (Dark Energy, Galaxy Genesis) featured in the following sections will require significant allocations of time - many tens of nights - to achieve their scientific goals. Since KAOS is conceived as a facility instrument, these key projects need not monopolize the spectrograph, since many other projects targeting similar regions of sky could be executed simultaneously. For example, kinematic and chemical surveys of our Galactic halo may not have the star density to fill the spectrograph fibers, and could be executed simultaneously with, say, an extragalactic survey or an observation of a galaxy cluster simply by sharing the field. After all, all the spectrographs fed by the fibers do not need to be configured identically.

Despite the argument that Gemini's current time allocation system precludes large allocations of time, it is already becoming clear that significant awards of time to specific programs are not only possible, but also yield high profile science outcomes (e.g., Gemini Deep Deep Field with GMOS). Towards the end of this decade, KAOS may then provide Gemini with an opportunity to investigate different modes of time allocation, more focussed on long-term goals and those which facilitate greater opportunity for scientific collaboration between partner countries. In particular, several of the projects described in this document will provide coherent data sets, which, through public databases, will effectively enable much wider access to Gemini observations than is currently possible. Countries with only small shares on the telescope will be able to tap into databases of millions of spectra in addition to their individual investigators having access to the plethora of data generated by very small allocations of telescope time.

1.5 The Gemini Data Archive and the NVO

KAOS will generate Gemini data which will have a lasting legacy, since the datasets resulting from KAOS will be obtained and processed in a systematic way.[3] Gemini and KAOS will therefore be a uniquely valuable component of the future virtual observatory.

---

[3] In the *HST* Archive, the datasets with the highest scientific impact (measured by how often these datasets are accessed by the community) tend to be the ones resulting from large systematic surveys (i.e., the Key Projects, DD Surveys like the HDF, Treasury programs, and large Snapshot programs; M. Postman and M. Corbin, private communication).

A key part of the astronomical landscape in the coming decade will be the Virtual Observatory (VO). All Gemini partner countries are members of the International Virtual Observatory Alliance (IVOA) and significant intellectual property already resides within the Gemini community (USA: SDSS/IPAC/STSCI/NOAO, UK: WFAU, CASU, Canada: CADC) in the development of the Virtual Observatory and its relationship to astronomical datasets.

Much of the VO work is currently focused on underpinning grid technologies, with current initiatives also seeking to provide limited VO-tools over next 1 to 2 years. Much of the work is also focused on the development of the VO in relation to imaging datasets; understandable given the more immediate requirements of programs such as VISTA/VST.

Since data pipeline processing must be an integral part of KAOS, VO-compliance must be built into the project from day one. As such, there is a real opportunity for Gemini partners to be part of the vanguard for 'spectroscopic data' in the VO. This would include the definition of data standards and observing protocols to maximize observing efficiency and developing Web services to maximize the science return from the data. In doing so, Gemini partner countries will develop valuable intellectual property in this area, in turn leveraging engagement in other fundamental spectroscopic datasets of the future (e.g., *GAIA*).

1.6 Timescale

Gemini's current 2nd-generation instrumentation initiative (heralded by the Aspen workshop in June 2003) is extremely timely in the context of KAOS. Given the astronomical landscape, KAOS would be set to maximize its scientific output by the end of the current decade. If we use previous experience to guide us, the timescale for instrument inception to delivery for the Gemini program is approximately six to seven years—this being the gap between the first Gemini instrumentation workshop held in 1997 and the delivery of the resulting instruments to Gemini over the coming one-two years.

A similar time scale following the Aspen meeting leads to 2nd generation instrument delivery in 2009-10. This timescale is also consistent with the scale of the instrument.

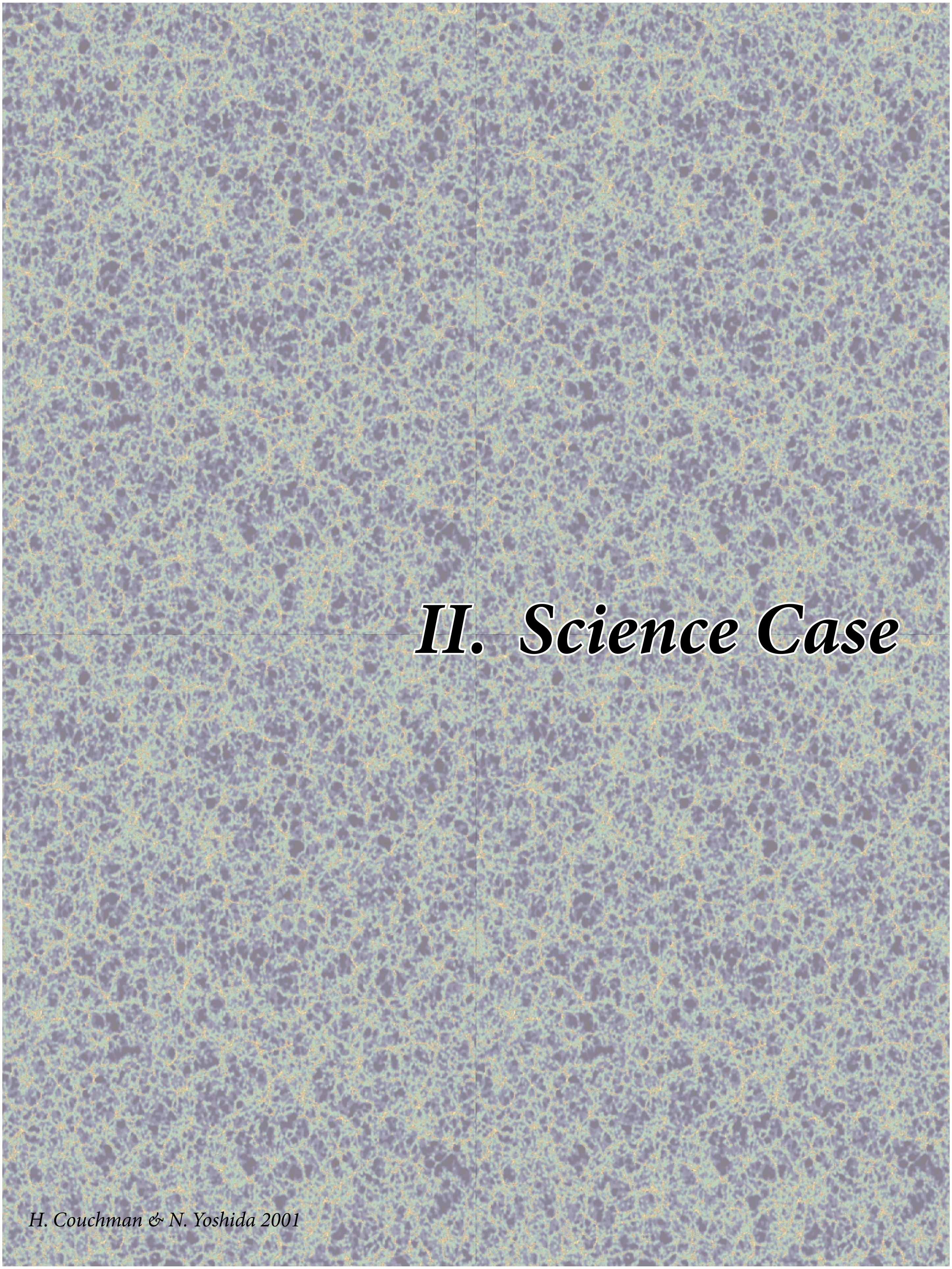
1.7 Cost

A ballpark estimate for the cost of the instrument lies in the range $20-30M. Such an amount would use up all, if not more, than the instrumentation funds available for 2nd-generation Gemini instrumentation initiative. It is therefore imperative that additional funds are identified if KAOS — even it is identified as a high-priority item from the Aspen process — to proceed. There are three possible sources of additional funds:

- Contributions from the member countries' national centers;
- Private monies;
- Contributions from non-member countries.

For example, the Anglo-Australian Observatory has already identified $1.25M in prototype positioner development costs as part of its in-kind contribution to the Australian Major National Research Facility Grant which could be used to offset against the total projects costs. Further contributions may come from private universities in exchange for guaranteed access to observing resources or data products. Finally, other observatories may see this capability as unique and may therefore be interested in either investing in it in return for access, or alternatively in trading access to complementary observing capabilities. Given the large number of 8-m class telescopes currently operational and limited funding for instrumentation, with KAOS Gemini could lead the way in establishing a more global system of astronomical resources.

II. Science Case



H. Couchman & N. Yoshida 2001

Chapter 1

Introduction

KAOS on Gemini provides the ability to conceive and carry out astrophysical investigations of an unprecedented scale. With each exposure, KAOS can generate nearly 5000 spectra of targets in a contiguous 1.5° field of view; this multiplexing ability will enable a new class of scientific studies that have been heretofore impossible, since they require either very large samples, or very large areal coverage, or both.

In the following chapters, we describe two key projects that are only enabled by KAOS: the *w(z)* Dark Energy Project (Chapter 2) and the Galaxy Genesis Project (Chapter 3). The goal of the Dark Energy Project is to provide unique constraints on the equation of state of the universe by using a direct metric measure – the scale of the acoustic oscillations in the linear regime. The Galaxy Genesis Project will be a first attempt at providing a detailed, nearly *complete*, map of the thick disk and halo of our galaxy with the goal of deciphering its formation history and directly testing competing theories. Chapter 4 provides a sampling of a variety of other science applications of KAOS. This is by no means a complete list, or even a completely representative one, but provides some idea of the impact of this instrument on a broad range of astrophysical problems.

KAOS is conceived as a facility instrument. Although the two main projects described here are telescope time intensive, they need not monopolize the telescope and instrument. Many other smaller proposed projects can be scheduled together with these projects; multiple scientific investigations can coexist on the same target exposure, constrained only by target density, spectrograph configurations and field selection.

Table 2.1.1 Summary of Selected KAOS Projects

Project	No. galaxies	Area (sq.deg.)	Spectral Resolution	KAOS Fields	Pointings per KAOS Field	Targets per KAOS Field	KAOS Nights
I. Dark Energy:							
1. z~3 Sample	6e5	150	1,000	85	1.5	7070	60
2. z~1 Sample	9e5	1000	1,000	566	1	1590	113
II. Galaxy Genesis:							
1. V=17 Sample	1e6	1666	20,000	943	1	600	400
III. Other Science:							
1. Local LSS	1.7e5	75	1,000	43	1	4000	22
2. Galaxy Evolution	5e5	100	2,000	57	1.8	8840	100
3. Growth of Structure	1e6	3x100	1,000	3x57	1	3330	200
4. AGN Physics at High-z	3.3e5	100	1,000	57	1	2000	25
5. IGM & Galaxy LSS	1500	10	5,000	6	1	250	4
6. M31 Halo/Bulge	2e5	25	10,000	50	1-4	4000	30
7. LMC Disk	1.4e5	100	20,000	45	1-2	100-6000	45
8. Milky Way Halo	1e5	400	5,000	225	1	440	112

Chapter 2

Dark Energy and Cosmic Sound:

A New Road to Cosmic Acceleration and the Equation of State from Giant High Redshift Surveys[4]

Abstract

KAOS will enable galaxy redshift surveys at the scale of a million galaxies at $z \approx 1$ and $z \approx 3$. From the clustering of galaxies, one can extract the scale of the acoustic oscillations as a cosmological standard ruler. This permits a precision measure of the Hubble parameter and angular diameter distance as a function of redshift, allowing one to study the acceleration of the universe at a level of detail comparable to and independent of future supernovae experiments.

2.1 Introduction

2.1.1 Cosmology in 2003

2003 was the year when the basic geometric parameters and key constituents of our Universe - Ω_b, Ω_m, Ω_λ, H_0 were determined to a few percent. This advance was made possible by a trio of observations: the *WMAP* CMB observations (Spergel et al. 2003); the 2dF Galaxy Redshift Survey galaxy clustering results (Percival et al. 2001); and the Supernovae cosmology results (Perlmutter et al. 1999; Riess et al. 1998). Each of these observations is critical: only the combination of *all* three datasets breaks the degeneracies between model parameters and renders a unique cosmology. The CMB can measure the matter density $\Omega_m h^2$ and the distance to z=1100 with great precision, but it cannot determine the present-day mix of components in the universe. Large-scale structure and supernovae studies are needed to distinguish the roles of matter and dark energy today.

In the last five years, we have detected and confirmed the acceleration of the universe, thereby inferring the existence of an unknown energy component of the universe that acts with *negative* pressure. This discovery poses a tremendous challenge to theoretical physics. The natural scale for the dark energy is far larger than what is observed, and there is no known physical principle that would make the value so small and yet be compatible with other cosmological measurements. This mystery has sparked investigation of a wide range of ideas.

One venerable candidate for the dark energy is the cosmological constant, but several more dynamic possibilities have been proposed. These include Quintessence (Ratra & Peebles 1998; Frieman et al. 1995; Caldwell et al. 1998), tracking scenarios (Wetterich 2002), k-essence (Armendariz-Picon et al. 2000), the 'Cardassian Expansion Scenario' (Freese & Lewis 2002), a unification of dark energy and dark matter known as a Chaplygin gas (Bilic et al. 2002), or even the alteration of gravity due to leakage into extra dimensions (Deffayet et al. 2002. All of these require exotic new physics and can be constrained by observations of the accelerating universe. Astronomical observations may therefore provide one of our few experimental handles on quantum gravity and string theory.

[4]Contributed by Daniel Eisenstein, Hee-jong Seo, Karl Glazebrook, Chris Blake & Eric Linder

The characterization of the acceleration of the universe and its anomalous source is therefore *the* major frontier of cosmology. Dark energy makes its presence known by altering distance measures in cosmology, including the evolution of the Hubble parameter with time, and the growth of structure. Precision measures of distance, Hubble parameter, or object counts across a range of redshifts can measure the equation of state. It is common to parameterize the effect of the dark energy on the acceleration of the universe depends on the ratio between its pressure P and energy density ρ. We write the equation of state as $w=P/\rho$, where w is in general a function of time (Steinhardt 1997; Turner and White 1997). The cosmological constant has $w=-1$, whereas other possibilities take on other values. CMB, galaxy surveys, and supernovae constraints together require $w<-0.7$ (Perlmutter et al. 1999; Percival et al. 2002; Spergel et al. 2003) but this is for a constant w. Observational cosmology is seeking methods to measure w as a function of time, but this is very challenging because the subtlety of the differences in observables that are created by variations in w.

2.1.2 Cosmology in 2010

Looking forward to the end of the decade, what can we expect the state of the cosmological art to be? The Planck satellite, due to be launched in 2007, will provide an order of magnitude improvement in CMB measurements over *WMAP*. The *SNAP* mission (recommended by the Turner report to the DOE) could launch as soon as 2008 and will provide a similar improvement in Supernova cosmology. New wide-field imaging surveys will be mapping the sky in multiple colors to perform weak-lensing mass reconstructions and cluster finding.

Large galaxy redshift surveys at $z>0.5$ offer an independent, high-precision route to dark energy. The clustering of galaxies contains a preferred scale imprinted at the epoch of reionization that manifests itself as a series of acoustic

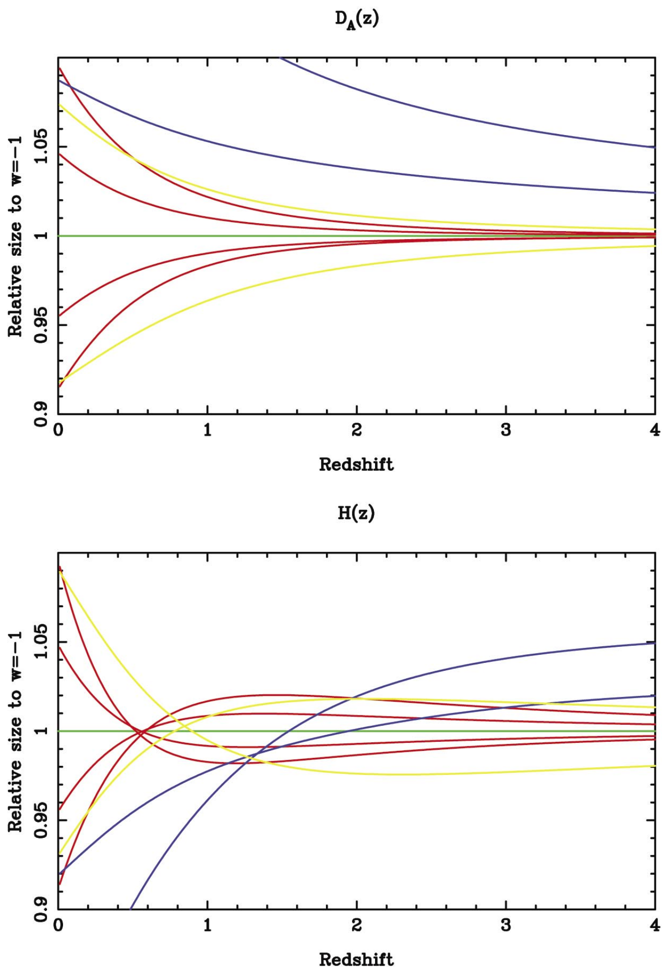


Figure 2.2.1: Comparison of distance-redshift and Hubble parameter relations between model cosmologies. All cosmologies are chosen to have $\Omega_m h^2$=0.14 and a constant angular diameter distance to z=1000; these constraints hold the CMB anisotropies nearly invariant. All models have $0.3<\Omega_m<0.45$ and $0.56<h<0.85$, within 2σ of current best values. The family of red curves are a set of constant w, i.e. "quintessence", models with $-0.8<w<1.2$. The Cardassian scenarios would be similar. The blue curves are for a Chaplygin gas with K parameters of 8 and 17 (Alam et al. 2003). The yellow curves show two braneworld models; here the vacuum energy is produced by the brane tension but the geometry is modified by the brane length scale l. The models here show two solutions with Ω_l=0.018 (Alam et al. 2003).

oscillations in the power spectrum. When combined with future CMB data, one can use the acoustic oscillations as a "cosmological standard ruler". By measuring the apparent size along and across the line of sight, one can measure the Hubble parameter $H(z)$ and the angular diameter distance $D_A(z)$ as a function of redshift. Such measurements would produce an accurate measurement of the equation of state at intermediate redshift and perhaps its time evolution.

We will show that the precision achievable from samples of roughly one million high-redshift galaxies with KAOS is comparable to that attainable from the *SNAP* satellite. Moreover, we stress that the methods are fully independent. Given the importance of dark energy and the spectre of undiagnosed systematic errors, it is crucial to pursue multiple paths towards to the precision measure of the dynamics of the universe.

Surveys with KAOS offer the unique chance to pursue the dark energy to $z\approx3$. While it is true that cosmological constant (i.e., $w=-1$) models predict that the dark energy has tiny effects at $z>1.5$, this must be tested. A number of the more dynamic models predict detectable effects at $z\approx3$. This is shown in Figure 2.2.1. Given the stakes and having been surprised by the simple existence of dark energy, we shouldn't fail to look for its effects even where our current prejudices would suggest a null result.

The required surveys must be large not only in number but in volume and sky area. Only an instrument with a very large $A\Omega$ can perform such a survey. The opportunity offered by KAOS is beyond the scope of any current spectrograph.

Regardless of what method is used to measure cosmological distances, one must remember that measuring even the time-average of $w(z)$ requires enormous precision. Dark energy is slightly dominant today but becomes subdominant to the "normal" dark matter at $z>0.5$. Moreover, w is extracted as the second derivative of the distance-redshift relation or the first derivative of $H(z)$. To measure w at few-percent precision and to have any chance to measure its evolution in time requires cosmological measurements at the few percent level.

2.2 Acoustic Oscillations as a Standard Ruler

The clustering of galaxies on large scales contains the fossil record of the growth of structure in the early universe. These can include signatures of the initial seeding of the perturbations, such as by an epoch of inflation, and of the processing of those fluctuations through the transition to matter domination and the epoch of recombination.

Here we will focus on the effect that is generically predicted to exist, namely the acoustic oscillations imprinted at the epoch of recombination (Peebles & Yu 1970; Holtzman 1989; Hu & Sugiyama 1996). Prior to recombination, the gas in the universe was locked to the photons of the cosmic microwave background, and the high pressure of this sea of photons caused the plasma to resist gravitational instability and instead to oscillate as a series of sound waves. After recombination, gas and light could separate, but the effects of the acoustic oscillations are imprinted in their spatial structure. We are familiar with this signature as the now-famous Doppler peaks in the anisotropies of the cosmic microwave background (Figure 2.2.2); however, the same structure is predicted to be present, as a weak sinusoidal modulation of the amplitude of fluctuations as a function of scale, in the late-time clustering of galaxies (see Figure 2.2.3). Detecting the matter acoustic oscillations at high significance would be a triumph of the Λ-CDM paradigm, with gravitational instability as the growth mechansim, adiabatic perturbations as input & CDM as the moderating component to the baryons.

The oscillatory pattern in the power spectrum has a characteristic scale, known as the 'sound horizon', which is the distance that a sound wave can travel between the Big Bang and the epoch of recombination. This scale depends only on properties of the early universe and is well measured by the CMB. *WMAP* presents a 3% measurement of the sound horizon (Spergel et al. 2003), and future CMB data should improve this to better than 1%. Hence, this scale is effectively a standard ruler (Eisenstein et al. 1998; Eisenstein 2003). As we will see, the scale of the oscillation can be measured to high precision in large redshift surveys.

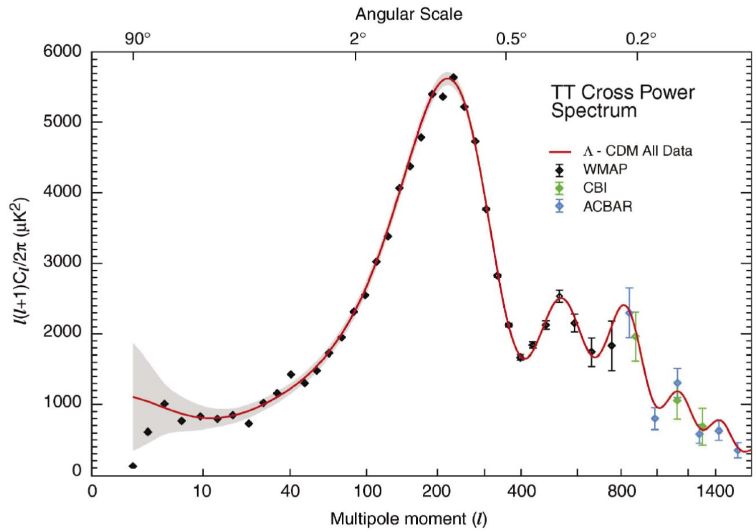


FIgure 2.2.2: The state of experimental data on the anisotropies of the CMB as of February 2003 (WMAP, Bennett et al. 2003). Acoustic peaks have been clearly detected, and cold dark matter cosmologies are a great fit.

An obstacle to the detection of this signal in the galaxy distribution is that the gravitational evolution of structure works to erase the primordial record in the clustering patterns on smaller scales. This occurs when perturbations on a given scale become of order unity in amplitude, leading to non-linear coupling between Fourier modes. Today, this veil of non-linearity extends to wavelengths of about $60h^{-1}$ Mpc, enough to wipe out all but the first of the acoustic oscillations (Meiksen et al. 1999). At higher redshift, the process is less advanced, and we can recover the primordial signals on smaller scale, including the full series of acoustic oscillations. For example, at z=3, we should be able to recover primordial information to roughly $12h^{-1}$ Mpc, which is even a factor of two smaller than what can be found in the anisotropies of the microwave background. Leaving aside acoustic oscillations for the moment, it is important to stress that direct measurements of the primordial fluctuations signatures on scales of 10 to 40 Mpc are *only* possible at high redshift and that the detailed nature of these fluctuations is essentially terra incognita. While signatures beyond the acoustic oscillations are more speculative, we might find other modulations or non-Gaussianities that could provide a window onto the inflationary generation of cosmic structure.

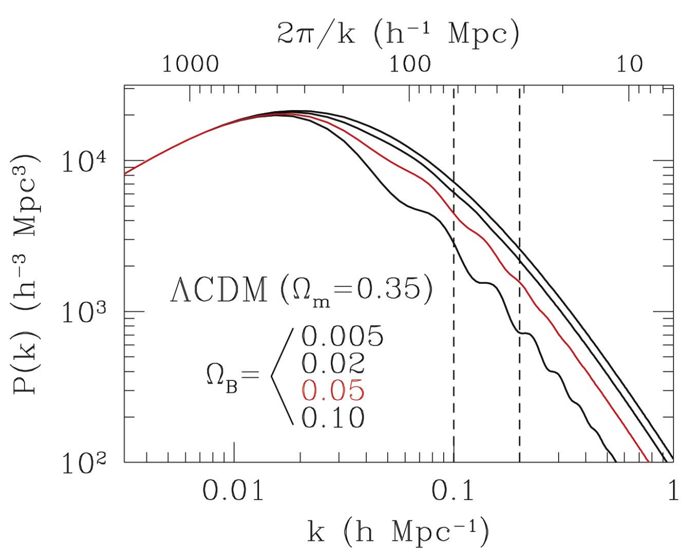


FIgure 2.2.3: Power spectra for four different cosmologies, with an increasing baryon fraction from top to bottom. Note the appearance of the acoustic oscillations as the baryon fraction increases. From Eisenstein et al. 1998.

The linearity of the density field on these large scales is a crucial advantage for using the acoustic oscillations as a standard ruler. Linear cosmological perturbation theory is well understood and offers a direct connection to the anisotropies of the CMB. Moreover, the preferred scale of the acoustic oscillations is much larger than any reasonable scale from galaxy formation physics. These make the acoustic oscillations a unique probe, free from the astrophysical & non-linear entanglements that plague, e.g., cluster based measurements.

Has this acoustic signal already been detected? Percival et al. (2001) claim a 2σ detection after integrating across the whole linear power spectrum in the 2dF Galaxy Redshift Survey. However, at low redshift we will always be limited by small sample volume, the dependence of the distance-redshift relation on just h, and the small extent of the linear regime. Observations at low redshift can only provide constraints on the first acoustic peak, which is limited for using oscillations as a precision measuring rod. The situation improves dramatically as we go to high redshift, where the matter overdensities naturally are smaller and the linear regime extends to smaller physical scales – a survey of order several times the local SDSS volume would likely reveal many peaks at very high statistical precision even by z=1.

It is worth comparing the distance measurements from the acoustic peaks to those inferred from the observations of type Ia supernovae (Riess et al. 1998; Perlmutter et al. 1999). The supernovae measure the luminosity distance as a function of redshift, which is equivalent in standard cosmologies to the angular diameter distance. The SN surveys cannot directly extract the Hubble parameter. While the distance-redshift relation is a less direct measure of $w(z)$ than the Hubble parameter, future SNe programs such as the *SNAP* satellite could achieve extremely good precision on distances at redshifts below 1.7. Of course, to realize their statistical potential, supernova cosmology programs must control their systematic errors at the 1% level.

2.3 Survey sizes and measurement precision

2.3.1 Empirical Goals

In a galaxy redshift survey, the fundamental observables are the angular separations and redshifts of the galaxies. We require a cosmological model to translate these angles and redshifts into physical distances. If we imagine observing a standard ruler at a particular redshift, then alterations in the dark energy model will distort the radial and transverse apparent size of the ruler. Since we know the true size, we can use the radial clustering signal to recover the Hubble parameter $H(z)$ and the transverse clustering signal to recover the distance $D_A(z)$. It is worth noting that this discussion differs from the familiar Alcock & Paczynski (1979) test in that we are working with a known length scale and therefore can extract $H(z)$ and $D_A(z)$ separately rather than merely their ratio.

Figure 2.2.4 shows the ratio of the radial and transverse distances between two cosmological models and a reference model as a function of redshift. The reference model is a cosmological constant model with Ω_m=0.3, Ω_w=0.7 and w=–1. The first variation model has w=–0.9. We pick Ω_m=0.329 and adjust the Hubble constant to hold the value of Ωh^2 fixed. This leaves the location and shape of the acoustic peaks in the CMB unchanged. A second model shows w=–0.8 and Ω_m=0.361.

Clearly, the effect of changing w in this manner makes only a small impact on cosmology. One must measure the distances to of order 1% at $z\approx1$ to distinguish these models. Given this requirement, we can proceed to calculate the necessary size of the redshift survey.

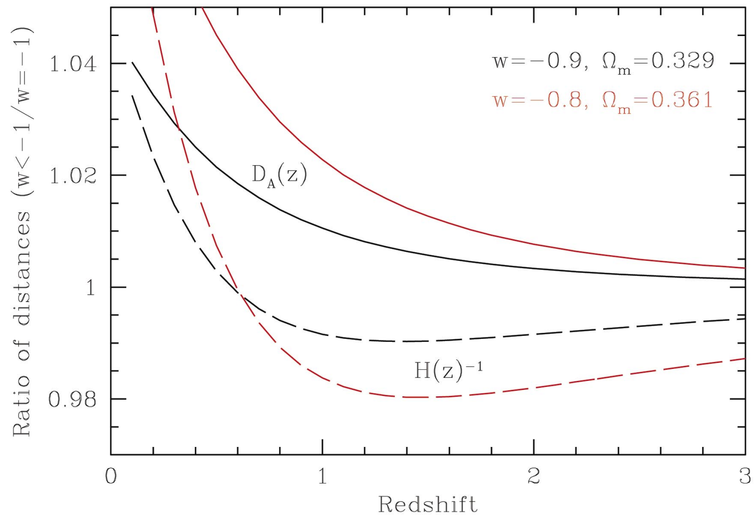


Figure 2.2.4: The length distortion of a rod as a function of redshift, supposing the true cosmology is $\Omega_m = 0.3$, $w = -1$ and the assumed cosmology is either $\Omega_m = 0.329$, $w = -0.9$ or $\Omega_m = 0.361$, $w = -0.8$, all at constant $\Omega_m h^2$. These two models are indistinguishable in the CMB but, as shown by the Figure, disagree in their low-z cosmography. The dashed and solid lines illustrate respectively the distortion if the rod is oriented radially (which depends on the ratios of $H(z)^{-1}$) and tangentially (which depends on the ratios of angular diameter distances). Thus it can be seen that the primary effect of assuming the incorrect cosmology is a re-scaling of distances away from their true values.}

2.3.2 Statistical Performance

The statistical errors on the power spectrum resulting from a redshift survey can be approximated as (Tegmark 1997a):

$$\frac{\sigma_P}{P} = 2\pi \sqrt{\frac{1}{Vk^2 \Delta k}} \left(\frac{1+nP}{nP} \right) \tag{2.1}$$

where V is the comoving volume of the survey, n is the comoving number density of galaxies in the survey, and P is the comoving power at the central wavenumber. This formula has a simple origin: the errors scale inversely with the square root of the number of Fourier modes measured, where the unit of volume in Fourier space is $2\pi^3/V$, and each mode is measured to order unity in the power with a penalty for shot noise. The shot noise penalty occurs when the white noise from the Poisson sampling of the density field exceeds the true clustering power. This happens when the product of the number density and power, nP, is less than unity. Note that this product is wavenumber dependent.

For the power spectrum measurement, if observational resources scale strictly with the number of survey objects (and not, e.g., with field of view), there is an optimal sampling density where $n = 1/P$ (Kaiser 1986). In other words, were our only goal to measure the power at a particular wavenumber, we would be most efficient by choosing the number density so that $nP = 1$. However, like all optimizations, the utility is a slow function of the controlling parameter near the maximum. In this case, $nP = 3$ or $nP = 1/3$ increases the errors by only 15%. Performance degrades more steeply as one moves further from the optimum.

In practice, we would recommend adopting higher sample densities than shown here:

1. For nP=1, the resulting survey measures each Fourier mode to a signal-to-noise ratio of 1, which makes for a very noisy picture. By using nP>1, we increase the fidelity of the map. This increases the scientific usefulness beyond the power spectrum statistic; for example, it would increase leverage on higher-order correlations and hence on non-Gaussianity.

2. nP>1 also means that one can divide the sample into several equal parts, e.g., by galaxy properties, and make comparisons that are not strongly limited by shot noise.

3. At $z\approx3$, the clustering of the matter remains linear to smaller scales than where the acoustic peaks are; since the power on these scales is smaller than the power near the acoustic peaks, there will be linear-regime science return from using a slightly higher n.

4. Finally, it is not the case that the observational resource requirements do scale simply with the number of survey objects. Field of view may be more expensive than additional spectroscopic fibers. For a fixed survey volume, one always wins by adding more survey objects; however, this benefit saturates for large numbers, and $nP = 5$ is within 20% of the $nP = \infty$ limit while being 80% larger than $nP = 1$. A counter argument is that higher number densities imply fainter objects and longer integration times. However, the number densities we seek can be found on the luminous tail of the luminosity function, where the source counts are quite steep.

What P should we use? We are interested in the higher acoustic peaks, which occur at $k\approx 0.2h$ Mpc^{-1}. The power at this wavenumber is about $2500\sigma_{8,g}{}^2 h^{-3}$ Mpc^3, where $\sigma_{8,g}$ is the *rms* overdensity of the galaxies in spheres of $8h^{-1}$ Mpc comoving radius. nP=1 would hence give $n = 4\times10^{-4}\,\sigma_{8,g}{}^{-2}h^3$ Mpc^{-3}. Note that this is considerably less than the density of L^* galaxies. Power is higher at smaller k, so smaller densities can be used when measuring larger scales.

2.3.3 Choice of galaxy targets

The sampling required is somewhat less than the densities of known spectroscopic galaxy populations at high redshift. Thus we have some freedom to pick our targets.

It is important to note that because we are seeking oscillatory features in the power spectrum, the fact that some types of galaxies might be highly biased in their clustering is not a problem. Oscillations in Fourier space correspond to preferred scales in real space. Bias will likely be scale-dependent and complicated on scales where hydrodynamics matter (i.e., 1 Mpc) or even where the structure formation becomes non-linear (i.e., few Mpc), but there is no reason for it to introduce a preferred scale on 100 Mpc scales (Coles 1993). Similarly, systematic errors can tilt the power spectrum but they are unlikely to produce oscillations. Non-linear redshift distortions are similarly smooth in power (Hamilton 1997). The acoustic oscillations in the power spectrum are very hard to obscure or distort; one needs only to reach the level of precision to detect them.

Table 2.1 - Dark Energy Fiducial Survey Parameters

Redshift Range	**Number of galaxies**	**Area (sq. deg.)**
$2.5<z<3.5$	6×10^5	150
$0.5<z<1.3$	9×10^5	1000

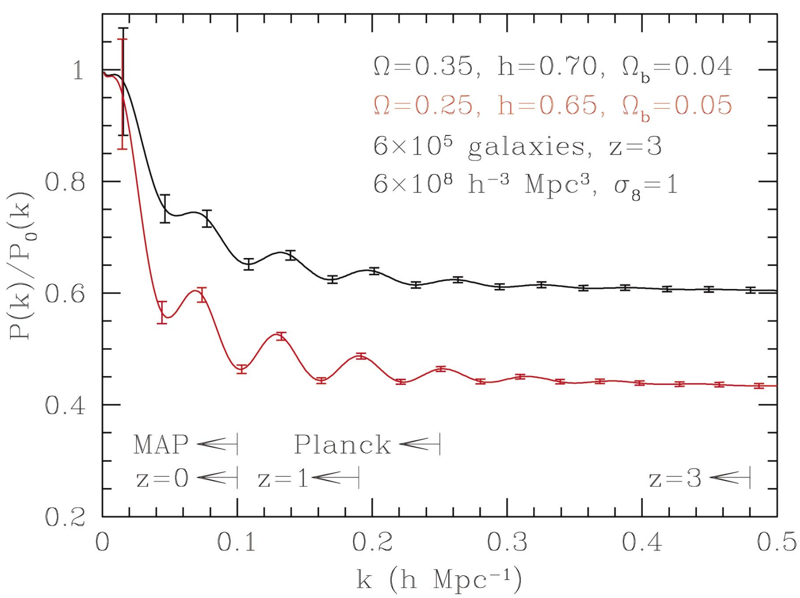


Figure 2.2.5: Simulated power spectrum for the z=3 survey. Two different cosmologies with different baryon fractions are shown; both are consistent with the baryon density inferred from big bang nucleosynthesis and WMAP. The power spectra have been divided by the power spectra from the same cosmology with zero baryons; this emphasizes the effects of the acoustic oscillations. Overlaid on the curves are the projected 1-σ errors from a survey at z=3 of 600,000 Lyman-break galaxies with a comoving density of $10^{-3}h^3$ Mpc^{-3}. If the redshift range of the survey were Δz=1, then the survey would cover 150 square degrees. Lines at the bottom show the extent of the linear regime as a function of redshift and the coverage of the two major CMB satellites (WMAP and Planck). Note that the z=3 survey probes the linear regime on a scale inaccessible in the CMB.

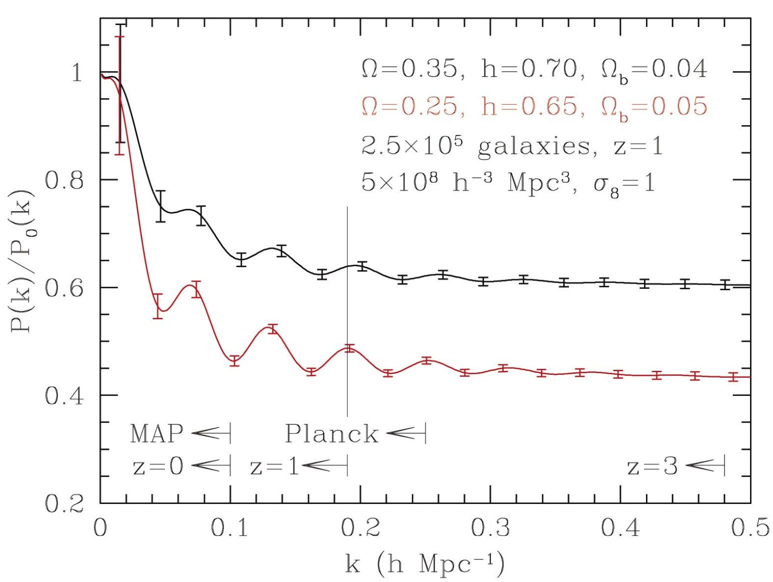


Figure 2.2.6: Similar to Figure 2.2.5, but the error bars refer to a survey at z=1 of a half million galaxies with a comoving density of $5\times10^{-4}h^3$ Mpc^{-3}. If the redshift range were Δz=0.2, then the survey would cover 1000 square degrees. We propose surveys extending from z=0.5 to z=1.3; the above figure includes only one quarter of the redshift range. The vertical line reminds the reader of the non-linear scale at z=1.

At $z\sim3$, the obvious set of galaxies to use is the Lyman-break galaxies. These have been measured to have $\sigma_{8,g} \approx 1$ (Steidel et al. 1996). We will seek a number density of $10^{-3}h^3$ Mpc^{-3}. A Ω_m=0.3 universe made flat by a cosmological constant has a comoving volume of $1120h^{-3}$ Mpc^3 per square arcminute between $z = 2.5$ and $z = 3.5$. This means that our survey will have a surface density of about 1 galaxy per square arcminute, which is very similar to the depth of Steidel et al. (1996). We adopt a fiducial survey size of 6×10^5 galaxies over 150 square degrees, the size being chosen to yield clear detections of the first three acoustic peaks (Figure 2.2.5).

At $z\sim1$, the choice of galaxy is less obvious, as one could use either giant ellipticals or luminous star-forming galaxies. Luminous early-type galaxies have the advantage that they have a high bias, probably yielding $\sigma_{8,g} > 1$, and strong 4000Å breaks, but measuring the redshift does require detecting this continuum break. Later-type galaxies may be less biased, but they have strong [OII]3727Å emission lines, which are easily identified because the line is a doublet. In either case, we have chosen a straw man of $\sigma_{8,g} = 1$ and $n=5\times10^{-4}h^3$ Mpc^{-3}. We use a slightly lower n because we are forced to $k<0.2h$ Mpc^{-1} by non-linear clustering at $z = 1$. From $z = 0.5$ to $z = 1.3$, there is a comoving volume of $540h^{-3}$ Mpc^3 per square arcminute, leading to a surface density of 0.26 galaxies per square arcminute. We adopt a fiducial survey size of 1000 square degrees, which corresponds to 9×10^5 galaxies, chosen to sample a similar volume to the SDSS luminous red galaxy sample. For presentation, we break the redshift range into four slices of $\Delta z = 0.2$ each; Figure 2.2.6 shows the forecasted power spectrum for one of the four slices.

To be clear, because of the difference in sky coverage and presumably exposure times, we are not imagining that these two surveys would share KAOS pointings. They might of course share pointings with other, better-matched, observational programs.

2.4 Measuring the Standard Ruler

Of course, we are particularly interested in how precision measurements of the power spectrum would translate into recovery of the apparent size of the sound horizon. We have done this in two different ways, first with an empirical treatment of simulated power spectra and then with a Fisher matrix calculation in a fairly general space of cold dark matter models.

2.4.1 A Robust, Empirical Approach

The empirical approach is drawn from Blake & Glazebrook (2003) and more details can be found there. In brief, we make a Monte Carlo realization of a survey power spectrum respecting the errors given above, divide the measured power spectrum by a smooth curve, and fit the remaining oscillations to a decaying sinusoidal function of wavenumber. The wavelength of the sinusoid is measured as a free parameter. An example of the fitting is shown in Figure 2.2.7. By considering many trial surveys, we construct the *rms* of this wavelength, thereby determining how well we can measure the standard ruler and in turn the distance to the redshift of the survey.

Figure 2.2.8 shows how the resulting precision depends on the KAOS survey volume and number density. Given that we need ~1% accuracy in the wavelength to obtain significant $w(z)$ constraints, it can be seen that surveys of order 10^6 galaxies covering of order 1000 deg^2 of sky are needed and that neither number nor area can be significantly shortchanged, in agreement with our estimates.

The advantage of the empirical approach is that it is essentially independent of the broadband shape of the power spectrum and hence of the details of the Λ-CDM model, scale-dependent galaxy bias, redshift distortions, etc. It focuses simply on the acoustic oscillation and recovers this as a distance scale that is imprinted by the CMB. While this model independence comes at a small price in precision, the results are similar to the model-dependent procedure to be described next, which justifies the acoustic oscillation as a robust feature of the power spectrum.

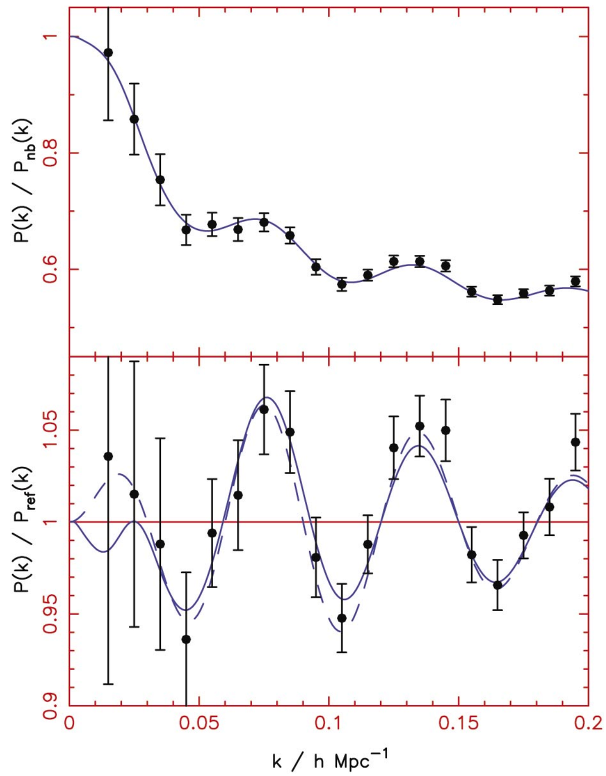


Figure 2.2.7: Power spectrum measurement for an example simulated survey of $N = 2\times10^6$ galaxies over a volume V = 1.2h^{-3} Gpc^3 at redshift z=1. The solid line in the top panel is the input model power spectrum divided by a reference zero-baryon power spectrum. The bottom panel shows this power spectrum with a smooth fit divided out. This reveals the oscillatory pattern. The dashed line in the lower panel is the best fit to the data of a simple decaying sinusoidal function. The wavelength of the oscillation can be fit to ~1% fractional accuracy.

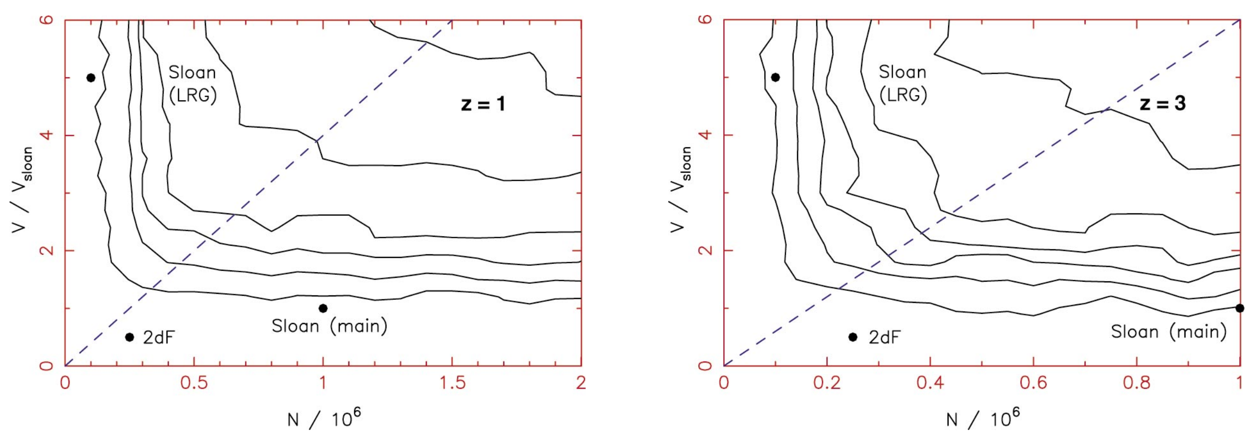


Figure 2.2.8: Fractional accuracy $\Delta\lambda/\lambda$ with which the wavelength of the baryonic oscillations in k-space can be measured at redshift z~1 (left) and z~3 (right), as a function of the number of galaxies N and the survey volume V. Linear bias factors b = 1 at z~1 and b = 3 at z~3 are assumed. Contours are shown corresponding to (beginning in the top right-hand corner) $\Delta\lambda/\lambda$ = 1.5%, 2%, 3%, 5%, 10% and 20%. The approximate positions of the 2dF and Sloan surveys (main and luminous red galaxy samples) are marked on the plot (although the contours are not appropriate for z~0 given the more restricted extent of the linear regime). We normalize the volumes to V_{sloan}=2×10^{8}h^{-3} Mpc^3. The diagonal dashed lines indicates the most efficient observational strategies, and corresponds to a surface density ≈ 2400 galaxies deg^{-2} (left) and ≈ 3400 galaxies deg^{-2} (right).

2.4.2 A Fisher Matrix Approach

We next analyze the performance of the KAOS surveys through the Fisher matrix formalism (e.g., Tegmark 1997b). This method computes the best possible error bars that a set of experiments can impose on a parameterized set of theories. A great advantage is that it makes it straightforward to incorporate more details of the Λ-CDM model and combine the constraints of the CMB and supernova datasets, either separately or jointly with KAOS. This approach will be published in full detail in Seo & Eisenstein (2003).

We use a very general space of cold dark matter models. Our parameters include the matter density ($\Omega_m h^2$), the baryon density ($\Omega_b h^2$), the optical depth to reionization (τ), the spectral tilt (n), and the tensor-to-scalar ratio (T/S). Our fiducial model is $\Omega_m h^2 = 0.35 \times 0.65^2$, $\Omega_b h^2 = 0.021$, $\tau = 0.05$, $n=1$, and $T/S=0$. We include an independent parameter for the angular diameter distance and Hubble parameter for each redshift bin that enters our experimental set (but only distance for the $z=1000$ CMB point). In particular, this means that distance-redshift relations cannot help data from the CMB or a survey at a particular redshift reconstruct the behavior at another redshift. This allows us to recover error bars on the distances and Hubble parameters at each redshift without reference to a dark energy model. Redshift distortions also produce an anisotropy in the power spectrum that is distinct from the cosmological distortions but could be confused with the latter. To ensure that broadband effects do not determine the redshift distortions, thereby breaking the degeneracy and giving an overly optimistic constraint on the cosmological distortions, we also include a separate parameter for the power spectrum amplitude, redshift distortion parameter β, and shot noise at each redshift. With these assumptions, nearly all of the ability to determine distances in the redshift surveys comes from detection of the acoustic oscillations.

Within this large parameter space, we combine Fisher matrices from multiple galaxy surveys and the *Planck* satellite. We then extract the covariance matrix of the parameters that are relevant to dark energy, namely $\Omega_m h^2$ and all of the distances and Hubble parameters. All other parameters are marginalized over and removed from the calculation. The result is a model-independent set of constraints on the distance-redshift relation and the Hubble-parameter-redshift relation.

When computing Fisher matrices for the redshift surveys, we have used a non-spherically averaged version of equation (2.1) so that the redshift and cosmological distortions can be fully assessed. This is one of the first fully 3-dimensional calculations of a Fisher matrix for a redshift survey and is likely the most general parameter space yet considered in this context. We neglect information from the non-linear regime, using a conservative assessment of the non-linear scale as $\sigma(R)=0.5$.

For CMB data, we use the predicted errors for the *Planck* satellite including polarization from Eisenstein et al. (1999). The *Planck* satellite is scheduled to launch in 2007, a similar time scale to building KAOS. With *Planck* data, the errors on $\Omega_b h^2$ are negligibly small and the errors on $\Omega_m h^2$ are comfortably smaller than needed to calibrate the sound horizon for these surveys. CMB data also provides a superb constraint on the angular diameter distance to $z=1000$.

For our fiducial KAOS surveys at $z\approx1$ and $z\approx3$, we present in Figure 2.9 the error bars ($1-\sigma$) on the Hubble parameter and angular diameter distances at each redshift bin. The points are covariant only through the uncertainty in the physical scale of the acoustic oscillations, which is minor in the case of *Planck*.

In addition to the KAOS redshift surveys, we include the projected results of the SDSS luminous red galaxy sample (Eisenstein et al. 2001). This sample will map $10^9\ h^{-3}$ Mpc^3 with a comoving number density of $10^{-4}\ h^3$ Mpc^{-3} at $z\approx 0.3$. Because this survey is well underway, we include it in all of our forecasts.

As one can see, the redshift surveys can produce precision measurements of these distances. Performance improves at higher redshift because the non-linear scale is smaller, allowing more acoustic peaks and Fourier modes to be measured, and because there is more volume at fixed sky area at $z=1.2$ than at $z=0.6$. At $z=3$, the constraints are particularly good, better than 2% on both quantities.

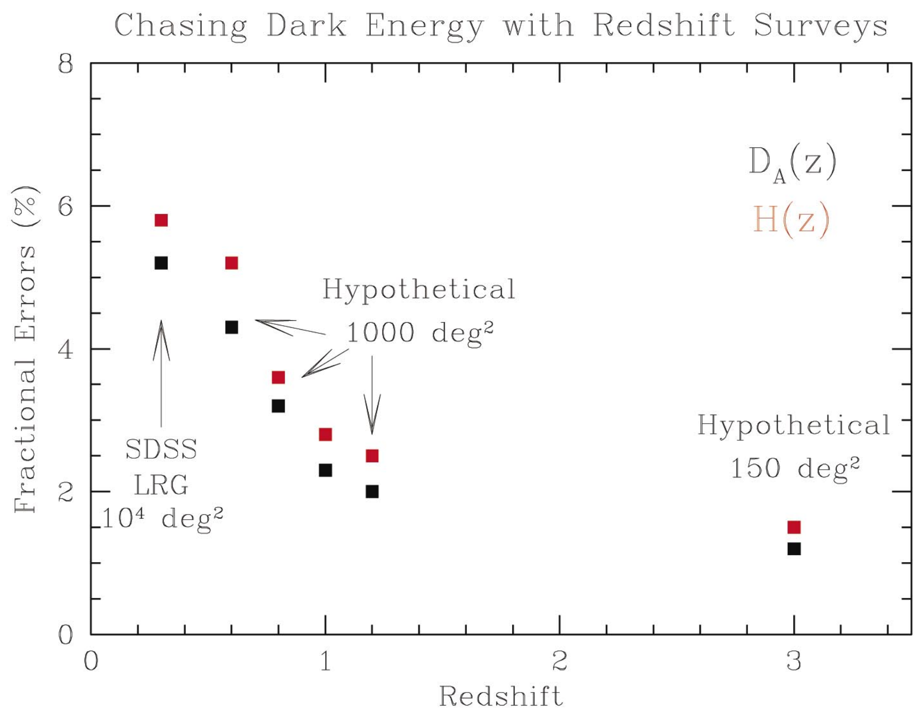


Figure 2.2.9: The 1σ errors on the Hubble parameter and the angular diameter distance for our hypothetical surveys. One survey is 1000 square degrees from z=0.5 to z=1.3, separated into 4 redshift bins, with a comoving density of $5\times10^{-4}h^3$ Mpc^{-3}. The other survey is a half million galaxies at z=3 with a comoving density of $10^{-34}h^3$ Mpc^{-3} (roughly 150 square degrees). Performance on the Hubble parameter is slightly worse than on the angular diameter distance because there are more tangential modes than radial modes.

2.5 Parameter Forecasts for Dark Energy

Armed with these forecasted measurements of the distance to and Hubble parameter at each redshift, we next consider the constraints on dark energy models. We consider models with four parameters – Ω_m, h, and $w(z) = w_0 + z\, w_1$ (with w constant at z>2) – and produce covariance matrices in this four-dimensional space.

Figures 2.2.10 and 2.2.11 display the constraints in the $w_0 - w_1$ plane, marginalizing over all other parameters. All ellipses include the forecasts for the Planck and the SDSS LRG sample. We show the results from the KAOS redshift surveys, the *SNAP* survey, and the two together.

For *SNAP* data, we assume a 1% measurement of the luminosity distance (which is trivially related to the angular diameter distance) in 15 redshift bins of 0.1 in width from 0.3 to 1.7. We also assume a 1% measure to z=0.05 from local supernovae, e.g., the supernova factory. Finally, we include an overall 5% uncertainty in the distances, correlated across all redshifts, because the supernovae do not measure the Hubble parameter to 1% accuracy. Our formulation differs in detail from the model of the *SNAP* team itself, but we believe it to be a close match in its quantitative results. Finally, we stress that this model for *SNAP* is very aggressive in its assumptions about systematic errors. Such errors are likely to be smooth residuals in redshift, just like the effects of dark energy that we are trying to isolate. Since the 16 Δz=0.1 bins are all statistically independent, we are essentially asserting better than 0.01 mag calibration between low redshift and high redshift. This is beyond the current state of the art and is essentially the design goal of the *SNAP* satellite.

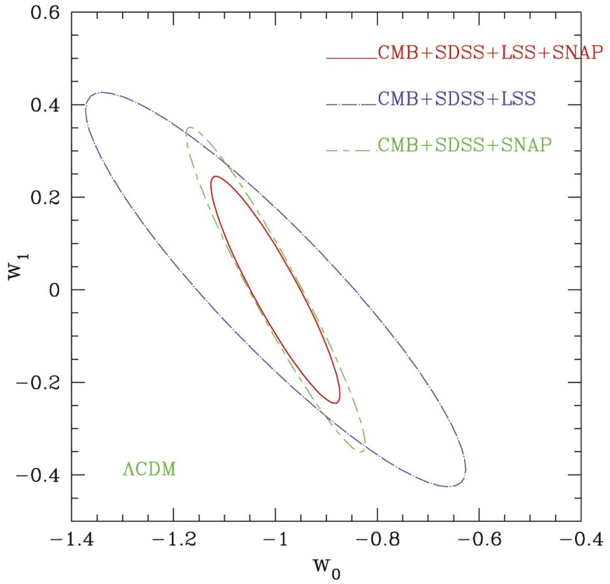


Figure 2.2.10: Constraints (1σ) in the $w_0 - w_1$ plane from a combination of the surveys in Figure 2.9, Planck CMB data, and the SNAP dataset. All other cosmological parameters have been marginalized over. The fiducial model is Ω_m=0.35, Ω_b=0.05, h=0.65 with a cosmological constant.

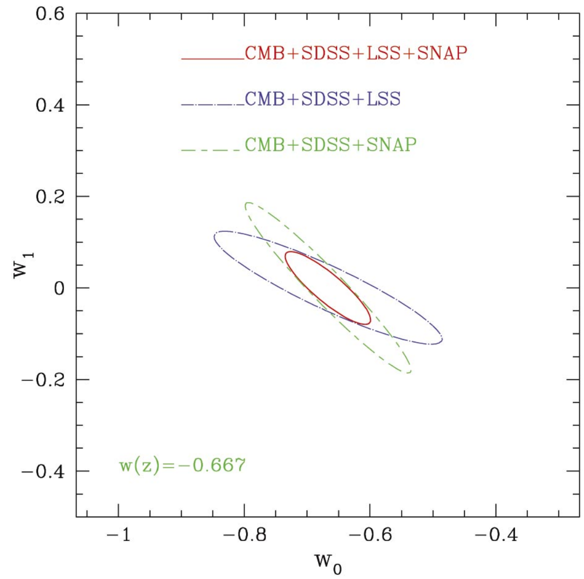


Figure 2.2.11: Similar to Figure 2.10, but the fiducial model is dark energy with w=–2/3. Because the dark energy is more important at high redshift, its parameters are easier to measure and the galaxy surveys, which perform better at high redshift, compete more evenly with SNAP.

Figure 2.2.10 shows the result when the fiducial model is a cosmological constant (i.e., we are perturbing around a w=–1 model). One sees that *SNAP* performs somewhat better than the KAOS surveys. Taking the performance on w_1 as the metric (as this is likely to be the frontier in five years), KAOS produces a 1σ error of 0.28, *SNAP* produces 0.23, and together they produce 0.16.

Figure 2.2.11 shows the results of a more interesting fiducial model which has an equation of state w=–2/3. *SNAP* and KAOS are now more evenly matched. In w_1, KAOS produces 0.08, *SNAP* produces 0.12, and together they produce 0.05.

The explanation for these trends is that *SNAP* could achieve a higher precision (obviously if the systematics are as good as predicted), especially at z<1, in the shape of the distance-redshift relation, but it suffers from an overall 5% uncertainty in the distance scale. Meanwhile, KAOS can measure to higher redshift (z=3 versus z=1.7) and can measure $H(z)$. The result is fairly evenly matched performance, with the exact balance depending on the details of the dark energy. Dark energy models with more positive w (or positive w_1 parameters) have more dynamical importance at higher redshift and can be better constrained by either experiment, with KAOS picking up more ground because of its $z \sim 3$ component. Interestingly, the overlap in redshift between KAOS and *SNAP* helps leverage the supernovae measurements: KAOS calibrates the distance scale, while *SNAP* provides the best measurements of the shape of the distance-redshift relation.

Both Figures show a clear covariance between the estimates of w_0 and w_1, in the sense that positive w_1's can be permitted if w_0 is more negative. This occurs because both KAOS and *SNAP* are making their measurements at z>0. If instead of asking for the value of the equation of state at z=0 we ask for the value at $z \approx 0.7$, we would find much less covariance between the intercept and slope of our linear equation of state model. Indeed, one can measure $w(0.7)$ to impressive accuracy, roughly 3%, while simultaneously estimating w_1. In detail, *SNAP* does better on this $w(0.7)$ measurement than KAOS.

An independent theoretical analysis more concentrated on the dark energy properties is even more optimistic about the value of the baryon oscillation method. Their well understood nature and linearity remove many of the systematic uncertainties that present obstacles to the use of other large scale structure methods. Since they are sensitive to the bare Hubble parameter, not just its integral, they possess a different combination of the dark energy equation of state parameters from other cosmological probes.

The physically critical clue of time variation in the equation of state can be treated naturally by the parametrization $w(z)=w_0+w_a\ z/(1+z)$, which serves as an excellent fit to slow roll quintessenece dynamics and is well behaved at high redshift (Linder 2003). This also allows for merging of data sets from different epochs in the expansion history, such as supernovae, baryon oscillations, and CMB information, to strengthen constraints on the time variation $w'=dw/d\ln(1+z) \approx w_a/2$. While baryon oscillations alone cannot break degeneracies sufficiently to obtain good estimates, they do show excellent complementarity with precision distance measurements from SNAP supernovae. As demonstrated above, the baryon oscillations aid in constraints as much as Planck CMB data, offering an important and independent crosscheck. Moreover, Fig. 2.2.12 shows that for a dark energy model with time varying equation of state, the oscillation data can yield further valuable benefits: they sharpen the limits on w' and w_0 by 40-50%! Note that spectroscopic data is key for these results: the Alcock-Paczynski method that uses only the ratio of the transverse to radial distance measurements has no such sensitivity (Linder astro-ph/0212301).

Perhaps the greatest promise of the baryon oscillation method is its freedom from the entanglement of astrophysical uncertainties, providing both robust crosschecks and strong synergy with other cosmological probes. But if we are willing to consider not only the matter power spectrum wiggle location, but also the evolution of the power spectrum amplitude with redshift, then further possibilities open up. So as to retain a firm hold on our understanding of the physics, suppose we remain in the linear region of the power spectrum, where the KAOS observations would focus. Then the linear growth factor of matter density fluctuations can also provide sensitivity to dark energy properties.

If KAOS measures the linear growth factor to 5% (i.e. the power spectrum amplitude to 10%) at redshifts z=1 and 3 then their evolution places complementary constraints on the equation of state. The addition of growth factor data leads to dramatic improvement in the case of time varying equation of state. In Fig. 2.2.13 one sees this enables determination of $w'=w_a/2$ to 0.05, or a 6 sigma detection of time variation! Even beyond this, the distinct dynamics of the growth factor allows breaking of degeneracies between models that have nearly identical CMB power spectra, and to test the framework of general relativity, e.g. distinguishing higher dimension braneworld models that mimic dark energy models in the distance measure (Jenkins & Linder, in preparation).

Thus a deep, wide field spectroscopic survey of a million galaxies at redshifts around z=1 and 3 offers both the promise of independent checks on the cosmology and the hope of strong synergy with other next generation probes to answer fundamental questions about our universe.

2.5.1 Dependence on Survey Parameters

We have experimented with varying the assumed volumes and number densities of the redshift surveys. We find that the dependence on number density is as designed. Decreasing the number densities while increasing the survey volume proportionally does slightly improve performance, as it would if nP were roughly 3. Similarly, increasing the number densities at fixed survey volume aids performance by the expected amount. The optimal choice of number density will depend on the details of the source counts, clustering bias, and exposure times, but the results are behaving as expected and we expect that we are within a factor of 2 of the best compromise. **In particular the volumes or galaxy numbers surveyed cannot be substantially reduced because the results will be compromised by cosmic variance or shot noise.**

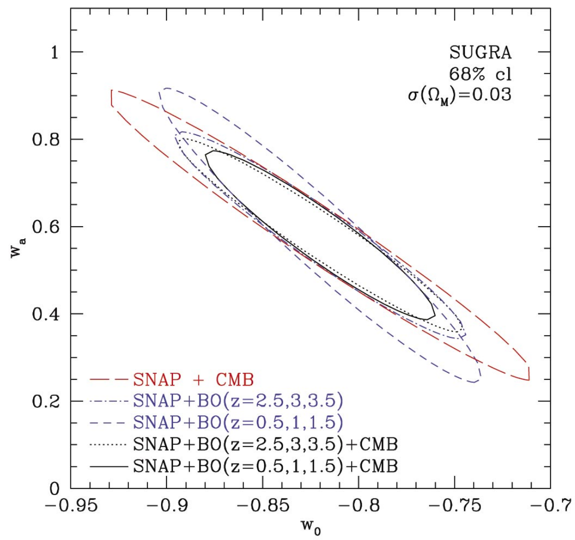


Figure 2.2.12: Joint probability contours (68% confidence level) for estimation of the time variation wa and the present value w0 of the dark energy equation of state, assuming a supergravity inspired dark energy model. The baryon oscillation and CMB data are more complementary with each other as well as the supernova data and there is a pronounced improvement when all three constraints are combined. [from Linder astro-ph/0304001].

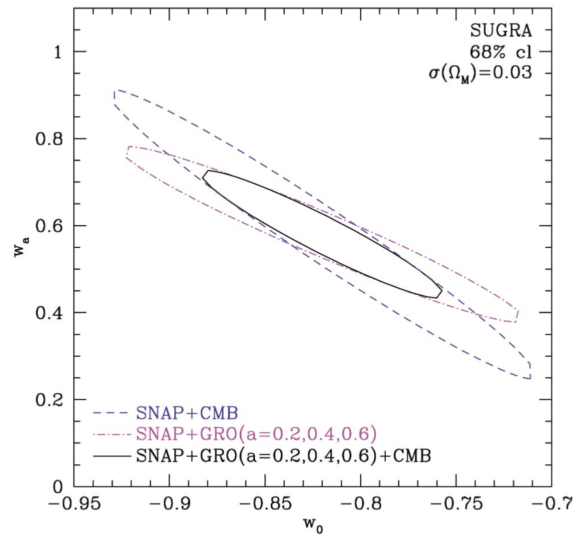


Figure 2.2.13: Parameter estimations (68% confidence level) of the present equation of state and its time variation. For the case of dark energy with time varying equation of state (here the SUGRA model with w0=–0.82 and wa=0.58), measurment of the linear growth factor at different redshifts adds valuable constraints on time variation, achieving 6 sigma detection [from Jenkins & Linder, in preparation].

As regards the survey size, increasing the volume at fixed number density causes the performance to scale nearly as the square root of the size. In detail, the improvement falls just short of this scaling, indicating that external information from the CMB is imperfect at a detectable but subdominant level. For the cosmological constant fiducial model and omitting *SNAP*, increasing both KAOS surveys by a factor of 5 improves errors on w_1 by a factor of two. Increasing only the $z{\approx}1$ survey is a 60% improvement; increasing only the $z{\approx}3$ survey is a 40% improvement. With *SNAP*, increasing both KAOS surveys by a factor of 5 improves errors on w_1 by 60%, an accuracy of $\Delta w_1 = 0.1$. Increasing only the $z{\approx}1$ (3) survey improves errors by 30% (35%). The w=–2/3 fiducial model would favor the z=3 survey more.

2.6 Discussion

Our calculations demonstrate that large redshift surveys such as could be performed by the KAOS instrument can provide accurate measurements of the dynamics of the late-time universe by using the acoustic peaks in the galaxy power spectra as a standard ruler.

The level of precision is competitive with that of the *SNAP* satellite, and the systematic errors are completely unrelated. We therefore regard this as a great opportunity to do these important cosmic acceleration measurements in a second and independent manner. By measuring angular diameter distance as opposed to luminosity distance we obtain an extremely important cross-validation of the supernovae results. That the surveys described here give comparable performance to *SNAP* validates the choice of survey sizes. Were they a factor of 4 smaller in $A\Omega$, they would be significantly off of the expected state of the art.

Weak lensing provides another possible precision route to cosmology. Full-sky surveys such as PANSTARRS and LSST could generate mass-selected cluster catalogs with sufficient numbers to place superb constraints on dark energy, essentially as a volume measurement at $z<1$. However, this does require excellent (<1%) control over the mass threshold of the cluster catalog, and it remains unproven that this challenge can be met. Moreover, the LSST is unlikely to be completed this decade, and so the deepest planned weak lensing maps could likely be considered as the instrument generation after KAOS.

If *SNAP* is not funded, then large redshift surveys will likely be the only way to measure the distance to $z>1$ at such high precision. KAOS will essentially pick up where concerted effort from ground-based supernova work leaves off. The $z\approx 1$ survey produces ~2% distance measures, which should safely exceed performance from ACS supernovae searches.

Supernova cosmology is of course more mature at present. Acoustic oscillations have not yet been convincingly detected in the galaxy power spectrum (although there are hints in Percival et al. 2001), in part because our largest surveys, 2dF and Sloan, must look at $z\sim 0$ where the sample volume is small and the fog of non-linear clustering is worst. However, there is reason to hope that with the progress on the SDSS luminous red galaxy survey and with the success of the *WMAP* mission, we may soon have a detection of the acoustic oscillations at low redshift. If so, then we will essentially be where supernovae cosmology was four years ago (Riess et al. 1998; Perlmutter et al. 1999). Extending further will require large surveys at higher redshift, such as would be enabled by KAOS.

Using the acoustic peaks for precision measurement of the dark energy requires more than just large samples. One must survey a large volume. Comoving number densities much in excess of $10^{-3}\,h^3\,\mathrm{Mpc}^{-3}$ are essentially of little help. Hence, for example, it is not enough to gather a million galaxies at $z\approx 1$; one must do so over ~1000 square degrees. This means that current multi-object spectroscopic instrumentation on 8-meter class telescopes is insufficient; one simply cannot cover enough sky. For example, this would be a minimum of 5000 pointings with IMACS, taking over 1000 clear dark nights.[5]

It is worth describing why spectroscopy is crucial to the endeavor. One might imagine that an imaging survey using photometric redshifts could make similar measurements of the acoustic peaks. Indeed, it is true that the first two oscillations do survive the projection into angular photometric redshift slices (e.g., see figure in Cooray et al. 2001). However:

1. Redshift accuracy at the level of $\sigma_z/(1+z) = 0.0015$ is required to extract the Hubble parameter, so an imaging survey is restricted to the distance-redshift relation.
2. Given a photometric redshift accuracy of even 4% in $(1+z)$, one must image 20 times the area to match the number of Fourier modes of the spectroscopic survey.
3. Higher harmonics of the acoustic oscillations are reduced in amplitude because of the thickness of the photometric-selected redshift slab. This further harms performance.
4. Control over systematic effects is greatly reduced. Spectroscopic surveys have redundancy for internally controlling their errors in photometric calibration or radial selection functions. Imaging surveys do not have this redundancy and are dependent on the external control of their calibrations (Tegmark et al. 1998).

Finally, we remind the reader that measuring the acoustic oscillations would be only one of the science products of these large redshift surveys. One is probing the linear regime with high precision on scales that are inaccessible at $z=0$. Primordial non-gaussianity or narrow features in the power spectrum (e.g., Broadhurst et al. 1990) can be recovered more easily than from the CMB or weak lensing surveys. One will of course use the redshifts, spectra, and maps to study the properties of galaxies at

[5]This estimate assumes that each 1/5 sq deg field needs to be covered only once with IMACS; an assumption that is almost certainly invalid given the limitations of simultaneously targeting objects over the entire field accessible to a slit spectrograph. A more realistic estimate would at least double the number of nights.

high redshift. This is another item of synergy between *SNAP* and KAOS, as *SNAP* seeks to provide 300 square degrees of high resolution, 9-band deep optical-infrared imaging that could provide SEDs and morphologies of the galaxies. Ground-based imaging surveys such as PANSTARRS or LSST would provide similar, albeit less capable, imaging data.

References

Alam U., Sahni V., Saini T.D., Starobinsky A.A., 2003, MNRAS, submitted, (astro-ph/0303009)
Alcock, C., & Paczynski, B., 1979,
Armendariz-Picon, C., Mukhanov, V. & Steinhardt, P.J., 2000, PRL, 85, 4438
Bennett, C. et al., 2003, ApJ, submitted (astro-ph/0302208)
Bilic N., Tupper G.B., Viollier R.D., 2002, Phys.Lett. B, 535 17
Blake, C.A. & Glazebrook, K., 2003, ApJ, submitted (astro-ph/0301632)
Broadhurst, T.J., Ellis, R.S., Koo, D.C., & Szalay, A.S. 1990,
Caldwell, R.R., Dave, R. & Steinhardt, P.J., 1998, PRL, 80, 1582
Caldwell R.R., 1999, Phys.Lett.B, 545, 23
Coles, P. 1993, MNRAS, 262, 1065
Cooray, A., Hu, W., Huterer, D., & Joffre, M., 2001, ApJ, 557, L7
Deffayet, C., Dvali, G., & Gabadadze, G., 2002, PR-D, 65, 044023
Eisenstein, D. J., Hu, W., & Tegmark, M., 1998, ApJl, 504, L57
Eisenstein, D. J., Hu, W., & Tegmark, M., 1999, ApJ, 518, 2
Eisenstein, D.J., et al., 2001, AJ, 122, 2267
Eisenstein, D.J., 2003, in *Wide-field Multi-Object Spectroscopy*, ASP Conference Series, ed. M. Brown & A. Dey
Freese K., & Lewis M., 2002, Phys.Lett., B, 540, 1
Frieman, J.A., Hill, C.T., Stebbins, A., & Waga, I., 1995, PRL, 75, 2077
Hamilton, A.J.S. 1997, in *Ringberg Workshop on Large-Scale Structure*, ed. Hamilton, D. (Kluwer Academic: Dordrecht)
Holtzmann, J.A. 1989, ApJs, 71,1
Hu, W., & Sugiyama, N. 1996, ApJ, 471, 542
Kaiser, N., 1986 MNRAS, 219, 785
Linder, E., 2003, PRL 90, 091301
Linder, E., astro-ph/0212301
Meiksin, A., White, M., & Peacock, J.A., 1999, MNRAS, 304, 851
Peebles, P.J.E. & Yu, J.T. 1970, ApJ, 162, 815
Percival, W.J., et al., 2001, MNRAS, 327, 1297
Percival, W.J.~et al. 2002, MNRAS, 337, 1068
Perlmutter, S., et al. 1999, ApJ, 517, 565
Perlmutter, S., Turner, M.S., & White, M. 1999, PRL, 83, 670
Ratra, B. & Peebles, P.J.E., 1988, PRD, 37, 3406
Riess, A.G., et al., 1998, AJ, 116, 1009
Spergel, D.N., et al., 2003, ApJ, submitted (astro-ph/0302209)
Steidel, C.C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K.L., 1996, ApJ, 462, L17
Steidel, C.C., Adelberger, K.L, Dickinson, M., Giavalisco, M., Pettini, M., & Kellogg, M., 1998, ApJ, 492, 428
Steinhardt, P.J., 1997, in *Critical Dialogues in Cosmology*, ed. N. Turok (World Scientific)
Tegmark, M. 1997, PRL, 79,3806
Tegmark, M., Taylor, A.N., & Heavens, A.F. 1997, ApJ, 480, 22
Tegmark, M., Hamilton, A.J.S., Strauss, M.A., Vogeley, M.S., & Szalay, A.S., 1998, ApJ, 499, 555
Turner, M.S., & White, M., 1997, PR-D, 56, 4439
Wetterich C., 2002, Space Science Reviews, 100, 195

Chapter 3

Galaxy Genesis:

Deciphering the Early History of the Milky Way[6]

Study of the resolved stellar populations of galaxies in the Local Group offers great scientific returns in our understanding of how galaxies form and evolve. The Local Group member galaxies (cf. Mateo 1998) include the three disk galaxies M31, M33 and the Milky Way, and numerous dwarf companions, both gas-rich and gas poor. What causes their similarities and also their diversity?

There are two main aspects to galaxy formation and evolution, namely the history of *mass assembly and re-arrangement* and the *history of star formation*. The old stellar populations play a particular role in deciphering these histories, since old stars usually retain a memory of certain aspects of their early life, such as the surface chemical abundances, and often orbital angular momentum and orbital energy.

The important questions concerning the *mass assembly*, apart from its rate, both past and present-day, include: what was the nature of the mass? — since collisionless dark matter, collisionless stars and collisional gas behave differently; what was the density distribution of any and all components? – since physical processes such as tidal stripping depend on relative densities, and dynamical friction timescale depends on mass ratios; what are the specific angular momenta and orbits? — since the coupling efficiencies of various processes depend on these.

The important questions concerning *star formation history* include: what was the rate of star formation and how did/does it vary as function of spatial location?; what was and is the stellar Initial Mass Function? – since the visibility of galaxies at high redshift, their contributions to background light in different passbands, their chemical enrichment, stellar feedback, the supernova rate, gas consumption rate, etc., depend on the IMF; what was/is the mode of star formation? what fraction formed in super star clusters?; what is the connection to the history of mass assembly of the various components.

Spiral galaxies are clearly diverse in their properties, for example in bulge-to-disk ratio, but theories should be able to produce the galaxy population in the Local Group rather naturally, without appeal to special conditions. Thus the Local Group members are 'typical' galaxies in their properties and for theory, but they are atypical for observation. From their resolved stellar populations we can obtain age distributions, chemical elemental abundance distributions, kinematics — all as a function of spatial location. The tracers that can be used include stars of a range of evolutionary stage and mass, planetary nebulae, star clusters, and gas through HII regions, 21cm emission and CO emission. The stellar properties in satellite companion galaxies provide important complementary constraints, for example, limiting the possible contribution of disrupted satellites to larger systems (cf. Unavane et al 1996).

[6]Contributed by Rosie Wyse

The study of Galaxy Genesis therefore requires observations of both our galaxy as well as its local group neighbors. In this chapter, we discuss the need for a KAOS survey of old stars in the Milky Way that provides unique constraints on the formation of our Galaxy. A discussion of KAOS surveys of the other local group members is included in the following chapter.

3.1 Background

The fossil record of galaxy formation and evolution is written in the distributions of age, chemical composition, kinematics, and spatial distributions of the stars, as well as the stellar initial mass function. These are the defining characteristics of a stellar population that we desire to study. We do not yet have a full characterization of the stellar populations in the Milky Way, throughout the main stellar components of disk, thick disk, bulge, and stellar halo. For example, even the star-formation history of the local region of the disk, the solar neighborhood, has only recently been determined to high precision using *Hipparcos* data (Hernandez, Valls-Gabaud & Gilmore 2000).

The well-known degenerate effects of age and metallicity on the positions of stars in color-magnitude diagrams (CMDs) frustrates any effort to learn the detailed star formation history from integrated spectra, or color-magnitude diagrams, alone. Chemical abundances from spectra are needed, and the more detailed the analysis, the greater the potential information to be gleaned. Different elements are synthesized in stars of different masses, and hence different evolutionary timescales; Wallerstein et al. (1997) provides a comprehensive discussion of our knowledge of sites and timescales of elemental production. Iron abundance by itself gives only an indication of membership in a specific stellar population.

If 'alpha'-elements (O, Si, Mg, etc. mainly produced in massive stars) can be measured, then the timescale of metal enrichment in a stellar population may be constrained (e.g., Wheeler, Sneden & Truran 1989; McWilliam 1997), and constraints placed on the massive-star IMF that pre-enriched the low-mass stars under study (e.g., Wyse & Gilmore 1992). The alpha elements, from Type II supernovae, are among the first metals to be produced and enrich on timescales of 10^6 yr, while iron is produced in Type I SNe that are thought to be related to white dwarfs, and contribute iron on timescales from around 10^8 yr, up to a Hubble time after the birth of the progenitors. The neutron-rich s-process elements, produced in the envelopes of AGB stars, can probe enrichment timescales of order 10^9 yr and may give interesting insights into the form of the primordial initial mass function. The r-process elements are produced in all supernovae, but are often especially enhanced in some of the most metal poor (ancient?) stars known (Truran 1981; McWilliam et al. 1995; Hill et al. 2002). The abundances in these stars of nuclear chronometers such as U, Th and Nd may provide important constraints on the ages of the oldest stars in the Galaxy (cf. Butcher 1987, Nature; Hill et al. 2002).

The kinematics, and kinematic gradients within a stellar population, also should be derived, to quantify relationships between the different stellar components, and, through the determination of correlations between kinematics and other parameters, elucidate the main physics determining Galaxy formation (slow dissipative collapse with angular momentum conservation or accretion by fragments?). The kinematics constrain the mass density profile of the Galaxy, and phase space structure constrains the disruption and accretion of smaller-scale substructure, e.g., globular clusters and companion galaxies.

As noted above, the characterization of the IMF in various environments is an important science goal. The high-mass IMF in old systems may be constrained by the values of the different elemental abundances in the long-lived stars they enriched; this is because the mass of alpha-elements synthesized and ejected in Type II supernovae is a function of progenitor mass, whereas the iron is fairly insensitive to progenitor mass. The low-mass IMF is derived by direct star counts, with intermediate masses requiring information about the star-formation history. The scatter in elemental abundance ratios of the vast majority of Galactic halo stars, those with [Fe/H] $\gtrsim -2.5$ dex is low, consistent with good mixing and the yields from an IMF-average Type II supernova, with a fixed IMF (Wyse 1997). However, for the most metal-poor stars, the scatter is large, consistent with seeing individual Type II supernovae (e.g., McWilliam 1997; McWilliam et al. 1995; Tsujimoto & Shigeyama 2003). Interestingly, the most metal-poor star known, at [Fe/H] = –5.3 dex (Christlieb et al. 2002), has only upper limits on neutron-capture elements, and has a pattern of elemental abundances that for the most part can be explained by a standard Type II supernova, but the elevated carbon and nitrogen abundances may point to a more exotic enrichment mechanism (e.g., Umeda & Nomoto 2003, astro-ph/0301315).

3.2 Where are the Old Stars in the Milky Way?

The stellar halo of the Milky Way is the classical 'Population II' old component; however, this contains only some few times 10^9 $M_\odot$ of stars (Carney, Latham & Laird 1990), with about 1% in globular clusters. The central bulge of the Milky Way is now generally considered to be a separate component from the halo (e.g., Wyse & Gilmore 1992), and contains around ten times as much stellar mass as does the stellar halo, or equivalently perhaps 20% of the disk stellar mass (e.g., Gerhard 2001). The bulge is also predominantly old, as old as globular clusters (Ortolani et al. 1993; Feltzing & Gilmore 1999), despite the bulge stars being significantly more metal-rich in the mean than are stars in the stellar halo. Old stars locally are predominantly in the thick disk; this stellar component has a vertical scale height some 3 to 4 times that of the thin disk, was discovered by Gilmore and Reid (1983) through star counts towards the SGP, and has been characterized as a distinct component of the Milky Way through studies of its chemical abundance distribution and kinematics (e.g., reviews of Gilmore, Wyse & Kuijken 1989 and Freeman & Bland-Hawthorn 2002). At the solar circle the thick disk contains 10%–20% of the stars in the thin disk (e.g., Chen et al. 2000). The old age is established through a combination of colors and chemical abundances, as shown in Figure 2.3.1 (cf. Gilmore, Wyse & Jones 1995). Recently Dalcanton & Bernstein (2002) have presented evidence that *all* late-type disk galaxies have a thick disk, with (surprisingly) uniform properties, supporting that the Milky Way is a typical galaxy.

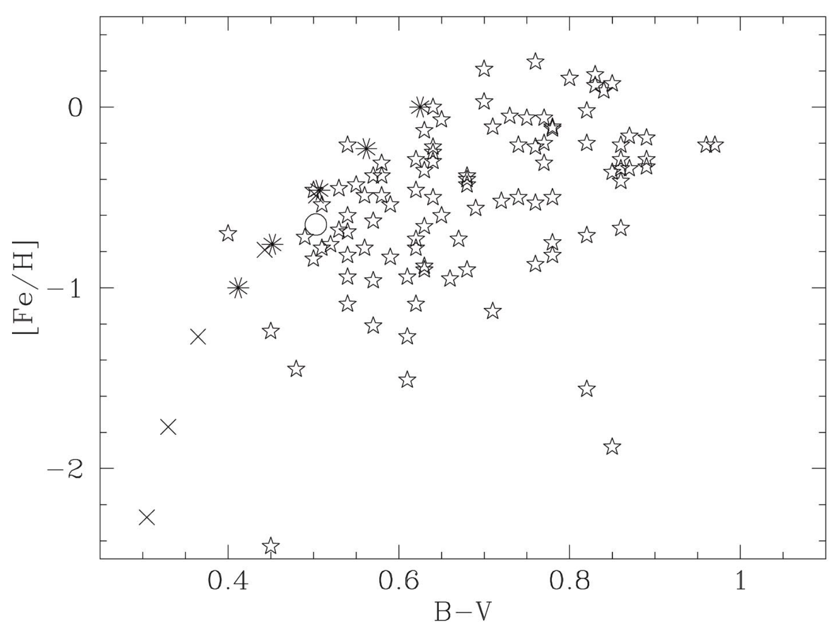


Figure 2.3.1: Scatter plot of iron abundance *vs B–V* color for thick disk F/G stars, selected in situ in the South Galactic Pole at 1-2kpc above the Galactic Plane (stars), together with the 14 Gyr turnoff colors (crosses) from VandenBerg & Bell (1985; *Y*=0.2) and 15 Gyr turnoff colors (asterisks) from VandenBerg (1985; *Y*=0.25). The open circle represents the turnoff color (de-reddened) and metallicity of 47 Tuc (Hesser et al.~1987). The vast majority of thick disk stars lie to the red of these turnoff points, indicating that few, if any, stars in this population are younger than this globular cluster. This figure is based on Fig. 6 of Gilmore, Wyse & Jones (1995).

Thus the thick disk presents an ideal opportunity to study stars formed during the early stages of evolution of the Milky Way.

3.3 The Thick Disk as a Probe of Early Times

There are several models for the formation of the thick disk, each with its own implications for the larger problems of the formation and evolution of disk galaxies. Discrimination among these models to decipher the fossil record of the Milky Way, as we demonstrate below, requires reliable kinematics, accurate to better than 10 km/s, and chemical elemental abundances, accurate to ~<0.3 dex, for samples of ~10^6 stars in several key lines-of-sight.

3.4 The Thick Disk is the Flattened Stellar Halo

At the time of its discovery, it was postulated that the thick disk was formed by local compression of the stellar halo by the gravitational potential of the thin disk (Gilmore & Reid 1983). This was soon disproven (Gilmore & Wyse 1985) by the determination of distinct metallicity distributions of (local) thick disk and stellar halo (see Figure 2.3.2 here).

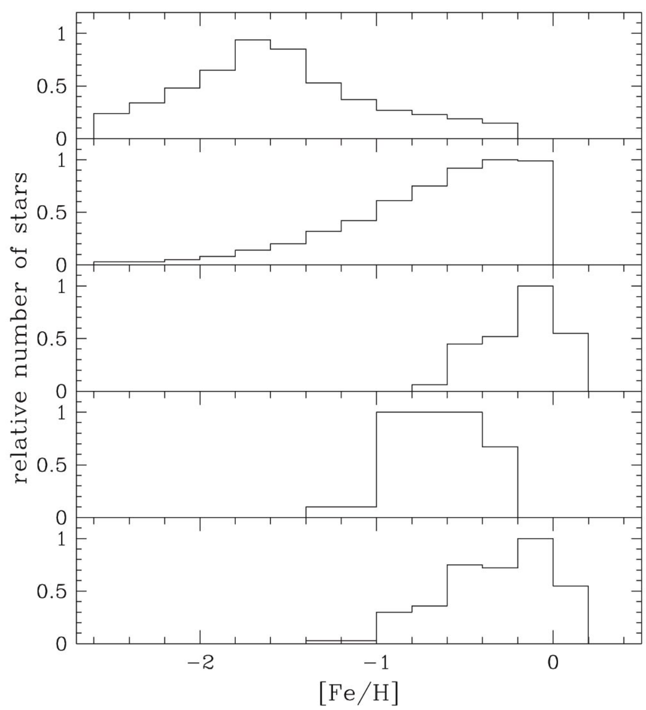


Figure 2.3.2: The metallicity distributions of representatives of the stellar populations of the Milky Way Galaxy. Where possible, a measure of the true iron abundance is plotted. The panels are (top to bottom) the local stellar halo (Carney et al. 1994, their kinematically-selected sample); the outer bulge K-giants (Ibata & Gilmore 1995), truncated at solar metallicity due to calibration limitations; the volume-complete local thin disk F/G stars (derived from the combination of the Gliese catalogue and in situ survey); the volume-complete local thick disk F/G stars (derived similarly); and lastly the 'solar cylinder', i.e. F/G stars integrated vertically from the disk plane to infinity. This figure is based on Figure 16 of Wyse & Gilmore (1995).

3.5 The Thick Disk is the Tail of the Thin Disk

The premise of this model is that the physical mechanisms by which the thick and thin disks form and evolve are the same. Evidence that the thick disk is kinematically distinct from the thick disk is shown in Figure 2.3.3, which shows cumulative vertical speed against rank metallicity (essentially vertical velocity dispersion as a function of metallicity, for the extant sample of two hundred or so local F/G stars from Edvardsson et al. (1993). However, this does not necessarily preclude similar physics (cf. Norris & Ryan 1991). There are two 'flavors' one should consider, one in which the disks evolve by thickening (or heating) and the other in which they evolve by settling (or cooling).

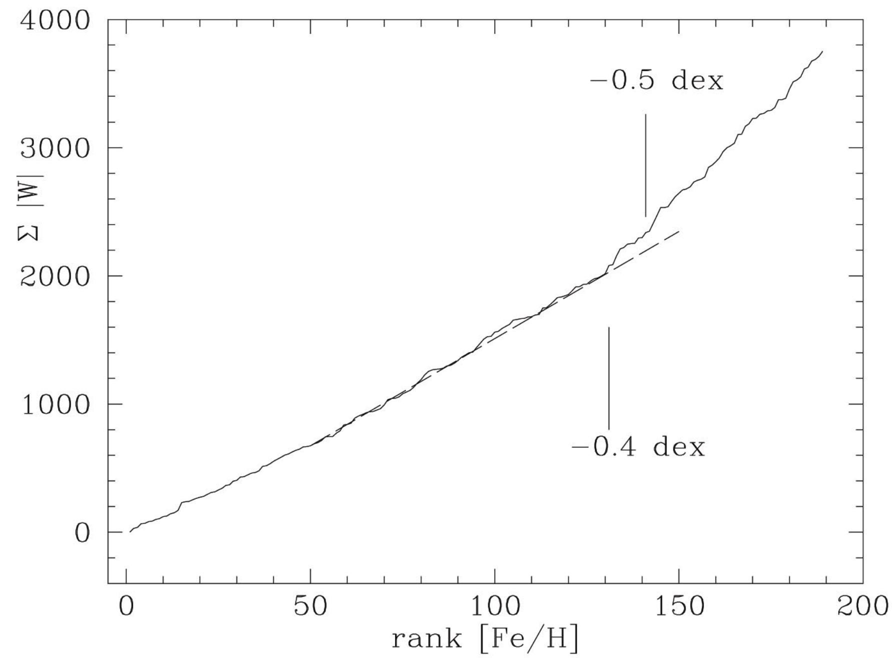


Figure 2.3.3: The rank number of a star in iron abundance versus the sum of the absolute value of the vertical velocity of the stars up to and including that rank, for the sample of Edvardsson et al. (1993) (taken from Wyse & Gilmore 1995). For a Gaussian velocity distribution, the slope of this plot is the value of the velocity dispersion. There is a clear change of slope at the metallicity at which the thick disk dominates over the thin disk, indicating distinct kinematics.

3.5.1 Heating

An increase in scale-height and velocity dispersion with stellar age within the thin disk is well-established (e.g., Wielen 1977), and can be understood to result from secular heating mechanisms in the disk, the most important being scattering of stars by local gravitational perturbations such as Giant Molecular Clouds or transient spiral arm segments (Spitzer & Schwartzschild 1957; Lacey 1991; Binney 2001). However, these mechanisms saturate at a vertical velocity dispersion of ~20 km/s, presumably not-coincidentally the actual value of the vertical velocity dispersion of the oldest thin disk stars (see Figure 2.3.3). Thus the value of the vertical velocity dispersion of the thick disk, some 40–45 km/s, is unexplained by this process. More exotic phenomena, such as close encounters with massive black holes in the dark halo, can provide the required high amplitude of random motions for a small fraction of the thin-disk stars (Ostriker & Lacey 1985; see also Sanchez-Salcedo 1999). However, in this case the thick disk should be a random sample of the thin disk, and have very a similar stellar population. The different metallicity distributions of thick disk and thin disk (Figure 2.3.2) then argue against this, as does the different age distributions of thick disk and thin disk, at least in the small local samples available.

3.5.2 Cooling

The thick disk could have formed as part of the (dissipational) settling of the proto-thin disk, with the scale-height set by the balance between cooling (and star formation) and gravity; the discontinuity in kinematics/scale-height between thick and thin disks could reflect the change in the cooling law as metallicity increases above ~1 dex and line radiation from metals becomes dominant over hydrogen line cooling (Gilmore & Wyse 1985; Burkert, Truran & Hensler 1993; Burkert & Yoshii 1996). One might then expect all (moderately metal-rich) disk galaxies to have a thick disk, which has recently gained support from the finding of ubiquitous thick disks by Dalcanton & Bernstein (2002). One would also expect a smooth vertical metallicity gradient within the thick disk (Burkert et al. 1992; their Figure 10). The available small samples do not support this (cf. Wyse & Gilmore 1995), but samples are small and errors on metallicity and distances are uncomfortably large.

3.6 The Thick Disk is Made of Shredded Satellites

Cold-dark-matter-dominated cosmologies posit that a large galaxy has undergone many mergers in its past. These many mergers with satellite galaxies could form a thick disk in which most of the stars now in the thick disk are former members of accreted satellites (e.g., Abadi, Navarro, Steinmetz & Eke, 2003). Figure 3.4 shows that even restricting the age range of the stars under consideration, both the old thick disk and the old thin disk are a complex mix of former satellites. The rather high mean metallicity of the thick disk stars observed so far then requires that the shredded satellites all be rather massive, given the observed correlation between mean galaxy metallicity and its velocity dispersion/luminosity, and our current understanding of that trend in terms of supernovae-driven winds. However, they should not be too massive so as not to over-produce the thick disk or destroy the thin disk!

Note that even in the scenario whereby a merger creates the thick disk by heating a pre-existing thin disk, discussed below, one would still expect that a signature of the shredded satellite would remain.

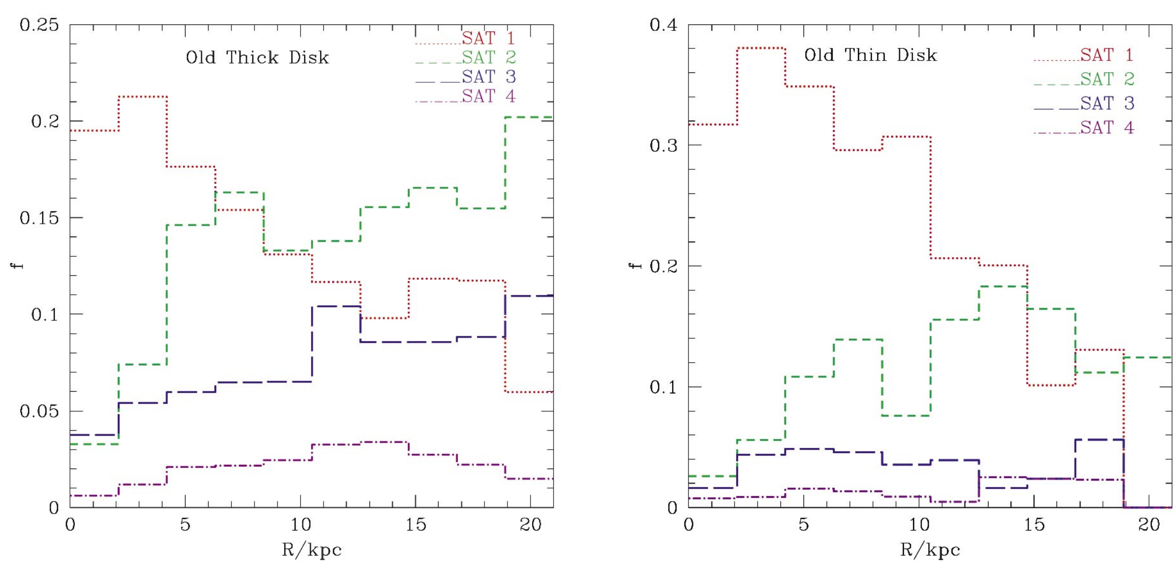


Figure 2.3.4: Taken from Abadi et al. (2002). (a) Left panel: The fraction of stars contributed by each of four large satellites (greater than 10^9 $M_{/}$) to the old (τ>10 Gyr) thick disk, as a function of radius. (b) Right panel: as the left panel, but for the old thin-disk component.

3.7 The Thick Disk is the Merger-Heated Early Thin Disk

A thick disk can be produced by merger-induced heating of a pre-existing thin disk (cf. Quinn & Goodman 1986). The measured vertical velocity dispersion of the local thick disk in the Milky Way can be provided for if a significant part of the orbital energy of a moderate-mass satellite galaxy is transformed into additional internal energy of the old stellar thin disk (Gilmore & Wyse 1985; Ostriker 1990; Majewski 1993). The effect of the accretion of a companion galaxy on the disk depends on many parameters such as those of the satellite's orbit (initial inclination to the disk plane, pericenter and apocenter distances, sense of angular momentum) and the satellite's density profile and total mass. Simulations of the merging process between a stellar disk and satellite have become increasingly sophisticated in recent years, including more physics such as allowing the excitation of the internal degrees of freedom of the dark halo, which lessens the heating effect on the disk (e.g., Huang & Carlberg 1997; Walker, Mihos & Hernquist 1996; Velazquez & White 1999). The extant simulations suggest that the accretion by the present-day stellar disk of a stellar satellite with mass some 20% of that of the disk can produce a thick disk similar to that observed in the Milky Way. However, gas has yet to be included in the simulations investigating disk heating, which is an important shortcoming, since gas if present (which is likely), would absorb and subsequently radiate away some of the orbital energy, again lessening the impact of the merger.

The lack of young or even intermediate-age stars in the (local) thick disk, combined with continuous star formation in the thin disk, then limits the last significant merger and associated heating of the thin disk to have occurred a long time ago (cf. Wyse 2001), at lookback times greater than around 12 Gyr, or redshifts of ~2 in 'concordance' Λ–dominated flat cosmology.

The shredded satellite(s) will also contribute to the 'non-thin-disk' stars, but likely with distinct kinematical and chemical signatures of the discrete stellar populations.

3.8 The Thick Disk is the Pre-Equilibrium Early Thin Disk

Should the cold-dark-matter paradigm not be correct, and instead of many mergers, large galaxies form by 'monolithic collapse', one can still form thick disks from dynamical heating of pre-existing thin disks. This would occur if the overall gravitational potential were non-spherical, and an initial thin disk formed after the gravitational collapse along the shortest axis, to relax to a new equilibrium after the gravitational collapse along the remaining axes (Jones & Wyse 1983). In this case there would be no contribution to the thick disk from shredded satellites, and again the thick disk would contain the oldest disk stars.

3.9 Observational Tests

How might one distinguish the various scenarios for thick disk formation noted above? It is clear that the age distribution, chemical elemental abundance distributions and phase space structures of the thick disk, all as a function of mean location within the Galaxy, are likely to be different in all cases.

What can be done with existing facilities? A first step to quantifying the phase-space structure of the Milky Way, through a comprehensive statistical study of the kinematics and metallicity distributions of stars in the interface between the thick disk and stellar halo, those stellar components for which mergers are most often implicated, was undertaken by Gilmore, Wyse, Norris & Freeman with the 2-degree-field fiber-fed spectrograph on the Anglo-Australian telescope, providing 400 spectra simultaneously. These spectra are used to obtain radial velocities and absorption line-strengths for samples of F/G main sequence stars at distances from the Sun of 5–10kpc (dependent on metallicity and magnitude), beyond significant contamination by the thin disk, down several key lines-of-sight.

The initial results (Gilmore, Wyse & Norris 2002) apparently detected a new kinematic component of the Milky Way Galaxy, plausibly the shredded remnant of the satellite whose merger with the Galaxy produced the canonical thick disk, by heating the pre-existing thin disk. This result was based on a sample of ~2,000 stars, to V=19, providing radial velocities to ~15 km/s and metallicities to ~0.3 dex. This tests the limit of the capabilities of the 2dF spectrograph.

Ideally, to decipher the history of the old thin disk, thick disk and stellar halo of the Milky Way Galaxy one needs:

1. kinematics accurate enough to detect cold streams, and hence to a couple of km/s;
2. good metallicities, hence [Fe/H] to 0.1 dex to be able to distinguish different star-forming entities; and
3. as good as possible elemental abundances, (relative) [α/Fe] to 0.2 dex to distinguish the signature of Type Ia supernovae in the star-forming region already identified by kinematics and metallicity.

Good elemental abundance estimates at present are limited to sample sizes of a few hundred in each of these three stellar components (Edvardsson et al. 1993; Nissen & Schuster 1997; Prochaska et al. 2000; Feltzing et al. 2003; Fuhrmann; Gratton et al. 2000).

However, due to the overlapping properties of the stellar components, and the need to define properties at the extremes of the populations, samples of many hundreds of thousands, or a million stars are required. Identification of substructure needs, at minimum, a 5σ deviation from a smooth distribution. The wings of the metallicity distribution of none of the thin disk, thick disk or stellar halo are well defined, but clearly all overlap in the extremes. Remembering that only 0.3% of a Gaussian-distributed sample is more than 3σ from the mean, very large samples are clearly required.

These goals set constraints on the spectral resolution, S/N and wavelength range. The fundamental stellar parameters of effective temperature and gravity should be constrained by the spectra themselves, although there will be several imaging surveys available by the time of the next generation Gemini instruments. Complementary constraints are set by the requirement of going faint enough to have an efficient haul of thick disk and halo stars.

3.9.1 Specifics: Spectral Regions and Resolution

The spectral region around the Ca triplet has many attractions, for example having little telluric absorption. Munari and co-workers, motivated by the proposed capabilities of the *GAIA* satellite, have discussed what can be achieved, specifically using the rather restricted wavelength range of only the 250Å from 8500Å–8750Å, initially focusing on radial velocities. Most relevant for the present discussion are the overview by Munari (1999), Munari & Tomasella (1999) and Munari & Castelli (2000) who analyze a library of synthetic spectra of cool stars (F-M) over wide ranges of metallicity and gravity (fixed microturbulence), with spectral resolution R = 20,000. They identify unblended lines of FeI, TiI, CrI, MgI, SiI, CaII and NI. R~20,000 is somewhat less than typical for elemental abundances, but e.g., Edvardsson et al. (1993) used some spectra with R~25,000 and found good agreement with their R>60,000 spectra. Armed with the synthetic library one should definitely be able to get at least 0.3 dex, based on the results of Carney et al. (1987) who have similarly defined grid of synthetic spectra and use χ^2 techniques to determine best fits.

TABLE 2.3.1: Optical windows for measuring iron and alpha elements

Wavelength Range (Å)	**Species**
8500-8750	CaII(3), NI(11), CrI(1), MgI(1), NiI(1), Ti(13), FeI(23), SiI(6)

(from Munari & Tomasella 1999; their Table 4) Note: the numbers in brackets refer to how many unblended transitions there are within the spectral window at R=20,000.

Note that the required velocity accuracies of <3 km/s can be achieved with modest S/N (around 10) and resolutions of more like 10,000 (see, e.g., Kleyna et al. 2002 — who quote this accuracy for R~7,500 and range of S/N; Carney et al. 1987 use R~30,000 and state velocities to 1 km/s for S/N~5). Canonically one can center a line to ~0.1 of a velocity pixel, given sufficient (modest) signal-to-noise ratio.

Getting metallicities to 0.1 dex is more of a challenge than the kinematics, and deriving the α-elements even more so. We need to have good estimates of gravity and stellar effective temperature; Boschi et al. (2002) show that various ratios of the unblended lines in this region provide diagnostics. As noted above, we should also have good optical and IR photometry from other facilities in any case.

It should be noted that the science of really deciphering the Milky Way early history cannot be achieved by just getting radial velocities good enough for a statistical assignment to 'thick disk' or to 'halo' (that is accuracy ~ 20 km/s) and 'metallicity' to 0.2-0.3 dex. This would not suffice to analyze kinematic substructure, nor would it determine star formation timescales or distinguish star-forming events.

The conclusion is that a spectral resolution of R~20,000 in the red is required, and wavelength coverage of several hundred angstroms around 8600Å.

3.9.2 FOV, Time and Sample Requirements

We need to target the old stars, those in the old thin disk, the thick disk and the halo. Star count models (e.g., Gilmore 1991) and simple considerations of distances (e.g., Table 3.2 below) point to an apparent magnitude limit of V=17 (or equivalently I~16.5). This gives a reasonable (relative) number of candidate thick disk and halo stars.

Adopting the 'standard' model galaxy from Gilmore (2000) we have that at high latitude, say at the SGP, in the two magnitude range 15<V<17 (restricting the range to minimize scattered light considerations) there are some 230 stars per square degree, of which some 140 are main sequence stars in the thin disk, around 45 are main sequence stars in the thick disk, around 10 are main sequence stars in the stellar halo, 15 are evolved (subgiant or red giant) stars in the thick disk and another 15 are subgiant or red giant stars in the stellar halo, with 5 halo horizontal branch stars (all per square degree). At lower latitudes, say at l=90, b=+30, these numbers increase to a total of 820 stars per square degree, composed of some 660 main sequence stars in the thin disk, around 80 main sequence stars in the thick disk, around 10 are main sequence stars in the stellar halo, 4 evolved (subgiant or RGB) stars in the thin disk, 43 subgiant or red giant stars in the thick disk and another 15 subgiant or red giant stars in the stellar halo, with 5 halo horizontal branch stars (all per square degree).

Multiplexing capability is clearly required; the target surface densities above imply that multi-object spectroscopy with around 1,000 objects/square degree is optimal. Fibers provide superior multiplexing over slitlets, and good sky subtraction, even at much fainter magnitudes than the current requirements, is possible with fibers with modern techniques such as 'nod and shuffle' (Glazebrook & Bland-Hawthorn 2001; see also Part III, Chapter 3). An 8-m telescope equipped with an efficient fiber-fed spectrograph (i.e., Gemini and KAOS) can obtain S/N ~ 25 at V=17, with R~20,000 in a 4 hr exposure.

TABLE 2.3.2: Log(distance, pc) as a function of apparent and absolute V mag.

Star	V	13.	14.	15.	16.	17.	18.	19.	20.
MPG	-2.0	(4.0)	4.2	4.4	4.6	4.8	(5.0)	5.2	5.4
	-1.5	3.9	4.1	4.3	4.5	4.7	4.9	5.1	5.3
MRG	-1.0	3.8	(4.0)	4.2	4.4	4.6	4.8	(5.0)	5.2
	-0.5	3.7	3.9	4.1	4.3	4.5	4.7	4.9	5.1
	0.0	3.6	3.8	(4.0)	4.2	4.4	4.6	4.8	(5.0)
CG/BHB	0.5	3.5	3.7	3.9	4.1	4.3	4.5	4.7	4.9
	1.0	3.4	3.6	3.8	(4.0)	4.2	4.4	4.6	4.8
	1.5	3.3	3.5	3.7	3.9	4.1	4.3	4.5	4.7
A	2.0	3.2	3.4	3.6	3.8	(4.0)	4.2	4.4	4.6
	2.5	3.1	3.3	3.5	3.7	3.9	4.1	4.3	4.5
	3.0	(3.0)	3.2	3.4	3.6	3.8	(4.0)	4.2	4.4
F	3.5	2.9	3.1	3.3	3.5	3.7	3.9	4.1	4.3
	4.0	2.8	(3.0)	3.2	3.4	3.6	3.8	(4.0)	4.2
	4.5	2.7	2.9	3.1	3.3	3.5	3.7	3.9	4.1
G	5.0	2.6	2.8	(3.0)	3.2	3.4	3.6	3.8	(4.0)
	5.5	2.5	2.7	2.9	3.1	3.3	3.5	3.7	3.9
	6.0	2.4	2.6	2.8	(3.0)	3.2	3.4	3.6	3.8
	6.5	2.3	2.5	2.7	2.9	3.1	3.3	3.5	3.7
K	7.0	2.2	2.4	2.6	2.8	(3.0)	3.2	3.4	3.6
	7.5	2.1	2.3	2.5	2.7	2.9	3.1	3.3	3.5

Note: MPG/MRG = metal poor/rich giant; CG/BHB = clump giant/blue horiz. Brackets help to delineate the 1 - 10 - 100 kpc transitions.

A survey sample size would need to be around 10^6 stars to define the transition between the thick disk and the halo in terms of kinematics and elemental abundance distributions and to identify and analyze phase space and chemical substructure. The strategy to adopt would be to target circles in the sky at constant latitude, then change latitude, do say b=30°, 45°, 60°, 90° and also half-circles at fixed longitude. This provides good coverage of kinematic cardinal directions. This survey strategy is optimal for establishing radial trends in all major Galactic components; note that our position at the Solar radius gives us an extra 8 kpc in radial coverage.

There are on average say 600 target stars/sq degree, to V=17, requiring sky coverage of 1666 square degrees for the required sample. At 4 hours per integration, clearly a FOV of a least a square degree is required.

To summarize, this project requires a wide-field optical spectrograph with a FOV of a least one square degree, and multiplexing capability of at least 1,000, and spectral resolution of R ~ 20,000. With KAOS on Gemini, this project would require of order 400 clear nights. Although this is a huge amount of observing time, most of the KAOS fibers are unfilled – hence many other high-latitude projects could be carried out simultaneously with this project.

References

Abadi, G., Navarro, J., Steinmetz, M. & Eke, V. 2002, ApJ submitted, astro-ph/0212282

Binney, J. 2001 in Galaxy Disks and Disk Galaxies, ASP Conf series vol 230, p63.

Boschi, F., Munari, U., Sordo, R. & Marrese, P.M. et al. 2002, astro-ph/0208319

Burkert, A. & Yoshii, Y. 1996, MNRAS, 282, 1349

Burkert, A., Truran, J. & Hensler, G. 1992, ApJ, 391, 651

Butcher, H. R. 1987, Nature, 328, 127

Carney, B., Laird, J., Latham, D. & Kurucz, R.L. 1987, AJ, 94, 1066

Carney, B., Latham, D. & Laird, J. 1990, AJ, 99, 572

Chen, B. et al. 2001, ApJ, 553, 184

Christlieb, N. et al. 2002, Nature, 419, 904

Dalcanton, J. & Bernstein, R. 2002, AJ, 124, 1328
Edvardsson, B., Andersen, J., Gustafsson, B., Lambert, D.L., Nissen, P.E. & Tomkin, J. 1993, A & A, 275, 101
Feltzing, S. & Gilmore, G. 2000, A & A, 355, 949
Feltzing S., Bensby, T., Lundstrˆm, I. 2003 A&A 397, L1
Freeman, K.C. & Bland-Hawthorn, J. 2002 ARA&A, 40, 487
Fuhrmann, K. 1998, A & A, 338, 16
Gerhard, O. 2001 in Galaxy Disks and Disk Galaxies, ASP Conf series vol 230, p21.
Gilmore, G. & Reid, I.N. 1983, MNRAS, 202, 1025
Gilmore, G. & Wyse, R.F.G. 1985, AJ, 90, 2015
Gilmore, G., 2000, in Allen's Astrophysical Quantities, 4th edition, edited by A.N. Cox (AIP Press, Springer New York).
Gilmore, G. 1981, MNRAS, 195, 183
Gilmore, G., Wyse, R.F.G. & Jones, J.B. 1995, AJ, 109, 1095
Gilmore, G., Wyse, R.F.G. & Kuijken, K. 1989, ARAA 27, 555
Gilmore, G. Wyse, R.F.G. & Norris, J. 2002 ApJL, 574, L39
Glazebrook, K. & Bland-Hawthorn, J. 2001, PASP, 113, 197
Gratton, R.G., Carretta, E., Matteucci, F. & Sneden, C. 2000, A&A, 358, 671
Hernandez, X., Valls-Gabaud, D. & Gilmore G., 2000 MNRAS, 316, 605
Hill, V. et al. 2002 A&A 387 560
Huang, S. & Carlberg, R., 1997, ApJ, 480, 503
Jones, B.J.T. & Wyse, R.F.G. 1983, A&A 120, 165
Kleyna J., Wilkinson, M. I., Evans, N. W., Gilmore, G. & Frayn, C. 2002, MNRAS, 330, 792
Lacey, C.G. 1991, in 'Dynamics of Disc Galaxies', ed B. Sundelius (Goteborg, Goteborg University), 257
Majewski, S.R. 1993, ARAA, 31, 575
Mateo, M. 1998, ARAA, 36, 435
McWilliam, A. 1997 ARAA, 35, 503
McWilliam, A., Preston, G., Sneden, C. & Searle, L., 1995 AJ 109 2757
Munari, U. and Castelli, F. 2000, A&AS, 141, 141
Munari, U. and Tomasella, L. 1999, A&AS, 137, 521
Munari, U., 1999, Baltic Astronomy, 8, 73
Nissen, P. & Schuster, W. 1997 A&A, 326, 751
Norris, J. & Ryan, S. 1991, ApJ, 380, 403
Ortolani, S. et al. 1995, Nature, 377, 701
Ostriker, J.P. & Lacey, C. 1985, MNRAS, 299, 633
Ostriker, J.P. 1990, in 'Evolution of the Universe of Galaxies', ASP Conf. Ser. Vol. 10, ed. R.G. Kron (San Francisco: ASP), 25
Proschaska, J.X., Naumov, S. O., Carney, B., McWilliam, A. & Wolfe, A. 2000, AJ, 120, 2513
Quinn, P. & Goodman, J. 1986, ApJ, 309, 472
Sanchez-Salcedo, F.J. 1999, MNRAS, 303, 755
Spitzer, L. & Schwartzschild, M. 1951, ApJ, 114, 385
Truran, J. 1981, A&A, 97, 391
Tsujimoto, T. & Shigeyama, T. 2003 ApJL, 584, L87
Umeda, H. & Nomoto, K. 2003, Nature in press, astro-ph/0301315
Unavane, M., Wyse, R.F.G. & Gilmore, G. 1996, MNRAS, 278, 727
Velazquez, V. & White, S.D.M. 1999, MNRAS, 304, 254
Walker, I., Mihos, J.C. & Hernquist, L. 1996, ApJ, 460, 121
Wallerstein, G. et al. 1997 Rev Mod Phys 69, 995
Wheeler, J.C., Sneden, C. & Truran, J. 1989, ARAA
Wielen, R. 1977, A&A, 60, 263
Wyse, R.F.G. & Gilmore, G. 1992 AJ 104 144
Wyse, R.F.G. & Gilmore, G. 1995, AJ, 110, 2771
Wyse, R.F.G. 1997, ApJL,490, L69
Wyse, R.F.G. 2001, in Galaxy Disks and Disk Galaxies, ASP Conf series vol 230, p71.

Chapter 4

Other Science Applications for KAOS

4.1 Studies of local large-scale structure[7]

Although most cosmological applications of KAOS would focus on high redshift, the instrument is also very well suited to one of the outstanding challenges that remains from the studies of local large-scale structure by 2dFGRS and SDSS. Samples of size >~10^5 galaxies suffice to map the distribution of L^* galaxies over regions large enough to be cosmologically representative (e.g., Colless et al. 2001). The low-order correlations of the distribution of light are measured with high accuracy, forming arguably the most important of the measurements used to break the degeneracies in cosmological parameters from the CMB (Percival et al. 2001; Spergel et al. 2003). The texture of large-scale structure is also well studied, with detailed delineation of the network of clusters, filaments and voids. The latter are especially interesting: regions with no L^* galaxies, of diameter typically 30 – 50 h^{-1} Mpc (see Figure 4.1).

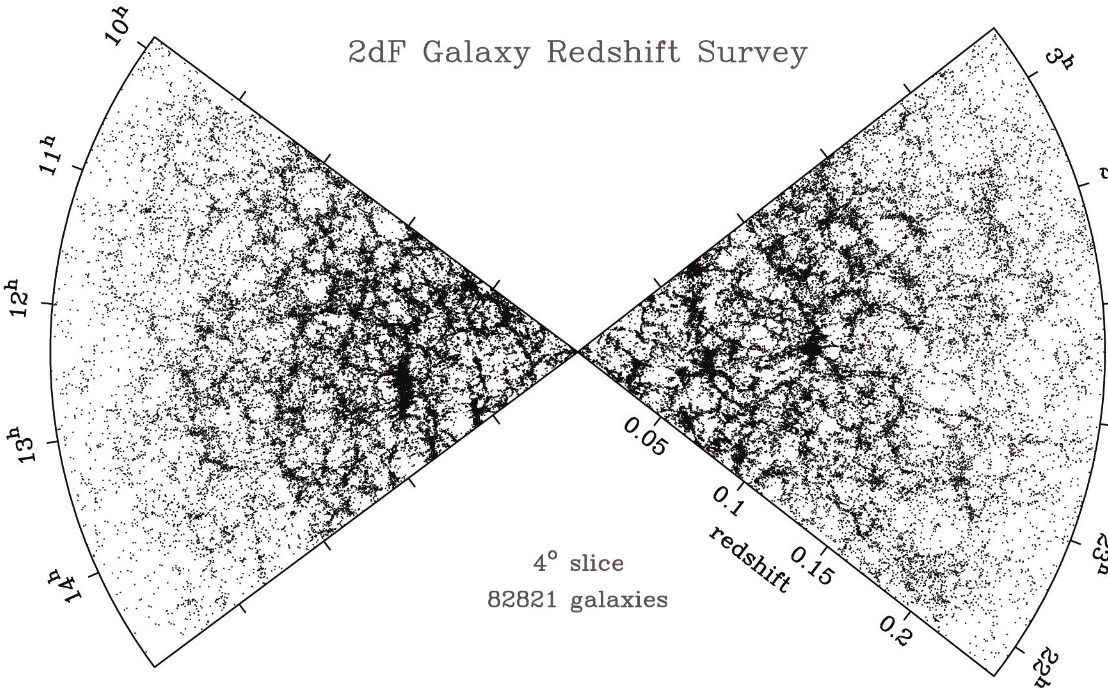


Figure 2.4.1: The distribution of galaxies in part of the 2dFGRS, drawn from a total of 213,703 galaxies: slices 4° thick, centered at declination –2.5° in the NGP and –27.5° in the SGP. This image reveals a wealth of detail, including linear supercluster features and voids of dimension 30 – 50 h^{-1} Mpc.

[7]Contributed by John Peacock

The 2dFGRS has studied in detail how the clustering of galaxies depends on luminosity (Norberg et al. 2001). The relative bias rises with luminosity, in a way that can be fitted by the formula $b/b^{*} = 0.85 + 0.15(L/L^{*})$. Thus, sub-$L^{*}$ galaxies cluster slightly more weakly, reflecting the tendency of the most luminous galaxies to inhabit high-density regions. Nevertheless, galaxies of luminosity down to $M_B \approx -18.5$ (roughly $L^{*}/3$) share much the same overall 'skeleton' of large-scale structure, as has been established for many years. The open area, however, is the distribution of galaxies at much lower luminosity. The 2dFGRS measures the luminosity function robustly down to $M_B \approx -13$, at which point the space density of galaxies is still increasing (Madgwick et al. 2001). However, the limiting redshifts for such galaxies are far too small at the 2dFGRS depth to allow meaningful study of their large-scale structure. What is really required is a redshift survey over part of the regions studied by 2dFGRS to a limit around 5 magnitudes deeper.

Voids and the spatial distribution of dwarf galaxies

Perhaps the most interesting question that would be opened up by such a survey would be the distribution of dwarf galaxies in the voids. Individual nearby voids have been noted since the first redshift surveys (e.g., Kirshner et al. 1981; Geller & Huchra 1989), and the long-standing puzzle has been the extent to which they were truly devoid of mass. The formation of voids is only partly understood in structure formation models. Cold Dark Matter (CDM) naturally produces a biased galaxy distribution, where luminous galaxies form preferentially in the high-density regions and are strongly clustered. The amount of mass expected in voids is strongly dependent on Ω_m: in a high-density $\Omega_m = 1$ universe, a high degree of bias is required, and the matter density would in practice exceed about half the mean value. In the preferred low-density $\Omega_m = 0.3$ universe, however, L^{*} galaxies are good tracers of the underlying mass distribution (Bahcall et al. 2000; Verde et al. 2002), so the density in voids is probably in the region of 20% of the mean.

Although CDM theories do a reasonable job of predicting the existence and approximate sizes of voids, they fail in a number of key respects (see Peebles 2001). First, CDM predicts that voids must potentially contain a high density of low-mass objects. This reflects the skewing of the conditional halo mass function towards low masses in voids: large numbers of $M{\sim}10^{9}\ M_{/}$ haloes should survive there, which have undergone merging into larger units in higher-density regions. In a sense, voids exhibit in a purer form the aspect of CDM that causes the 'substructure problem': incomplete destruction of merged sub-haloes means that a system like the local group is expected to contain thousands of low-mass dwarf companions (Moore et al. 2000). In the latter case, the situation is complicated by the stripping of the outer parts of the sub-haloes, and the need to understand how this affects the number of stars remaining in the core. In a void, the much lower degree of dynamical evolution means that this complication is evaded: in a sense, voids provide a 'time machine' by which we can inspect conditions analogous to those at high redshift, where L^{*} galaxies were dispersed into subgalactic chunks.

The difficult question is of course the baryonic content of these low-mass haloes. If we assume that voids in the galaxy distribution arise as a continuation of the morphology-density relation, then they should contain a homogeneous population of late-type, star-forming systems. Theory tends to back up this empirical extrapolation: for example, Benson et al. (2003) predict that star-formation rates per unit stellar mass should peak in voids, and decline with increasing distance away from the void into regions of higher galaxy density (see Fig. 4.2). However, in practice voids appear to be more strongly delineated compared to the surrounding LSS network than expected even in simulations with a uniform M/L (Cole et al. 1999). This has led many authors to propose ways to suppress galaxy formation in the lowest density regions (e.g., Rees 1985; Bode, Ostriker & Turok 2001). Since the theoretical situation is rather unclear, there is a clear motivation to measure empirically the distribution of dwarf galaxies (equivalently, how the very faint end of the galaxy luminosity function is modulated by environment). The results would be very important in removing current degrees of freedom in galaxy formation models.

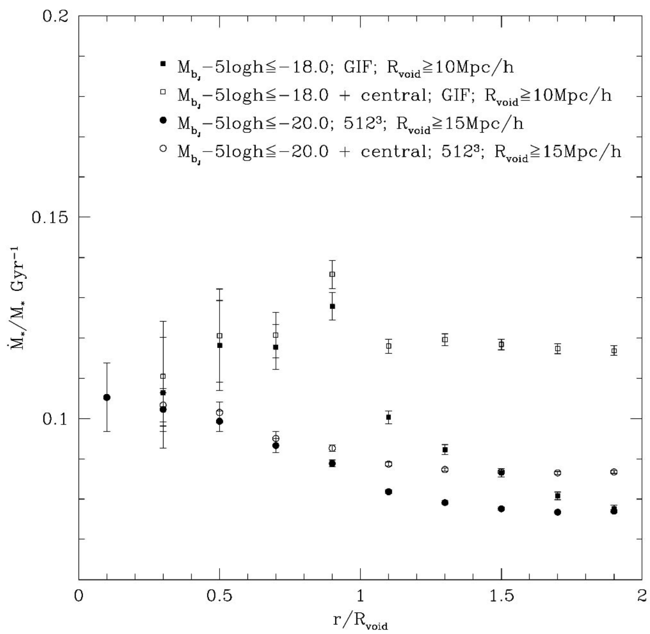


Figure 2.4.2: The specific star-formation rate (i.e., SFR per unit mass of stars) in model galaxies as a function of the radius to their nearest void centre. Figure taken from Benson et al. (2003). Filled symbols include all model galaxies in the specified luminosity range; open symbols include only those existing at the centre of their halo.

Application to KAOS

The parameters of KAOS are readily seen to be well suited to this problem. The typical depth of 2dFGRS is $z \approx 0.1$, at which distance a 30 h^{-1} Mpc void subtends 5°. This could be well covered by a 3×3 mosaic of KAOS fields. In order to study a full range of galaxy densities from one void to the next, through the intervening supercluster, a strip of ≈5°×15°, or of order 50 KAOS fields would be required. One would aim to reach the bottom of the luminosity function as presently defined, corresponding to $B \approx 25$ at this redshift (cf. the current 2dFGRS limit of $B \approx 19.5$). The number of targets is reasonably well matched to KAOS: the total galaxy density to these sort of depths is of order 0.2 h^3 Mpc^{-3}, so that 4000 targets would exist over a single KAOS field to a redshift of approximately 0.15. KAOS would therefore provide a definitive measurement of LSS to the bottom of the luminosity function, within the volume where it is presently well measured in L^* galaxies. This would be a substantial project, but modest in comparison to the time requirements of some of the high-redshift applications.

A key practical aspect will be the necessity for photometric pre-selection, since most galaxies at $B \approx 25$ will lie at redshifts much larger than those of interest. The COMBO-17 project (e.g., Wolf et al. 2002) demonstrates that multi-band CCD imaging can deliver the necessary separation of foreground and background. By the time KAOS is operational around 2010, it will be routine to obtain the necessary CCD imaging in order to allow the selection of the targets of interest.

References

Bahcall et al. 2000, ApJ, 541, 1
Benson et al., 2003, MNRAS, 340, 160
Bode, Ostriker & Turok 2001, ApJ, 556, 93
Colless M. et al., 2001, MNRAS, 328, 1039
Geller & Huchra, 1989, Sci, 246, 897
Grogin & Geller 2000, AJ, 119, 32
Jones & Bland-Hawthorn 2001, ApJ, 550, 593
Kirshner et al. 1981, ApJ, 248, L57
Madgwick D.S. et al., 2001, astro-ph/0107197
Norberg P. et al., 2001, MNRAS, 328, 64
Peebles 2001, ApJ, 557, 495
Percival W.J. et al., 2001, MNRAS, 327, 1297
Rees 1985, MNRAS, 213, 75
Spergel D. et al. 2003, astro-ph/0302209
Verde et al. 2002, MNRAS, 335, 432
Wolf C. et al., 2002, astro-ph/0208345

4.2 Formation and Evolution of Galaxies[8]

How do galaxies form and evolve? Primeval galaxies undergoing their very first episodes of star formation have remained elusive, either because these nascent systems are shrouded in dust or because galaxy formation is a slow hierarchical process wherein galaxies are assembled over time from small building blocks. The aging of galaxies has also remained a puzzle: so many factors contribute to the chemical, dynamical and morphological development of a galaxy, that the small observational data sets painstakingly obtained barely provide us with clues even to the most global issues (e.g., the evolution of the luminosity function) over a large range in lookback time.

Our investigation of the processes of galaxy formation and evolution is currently restricted to theoretical simulations and small observational data sets. Current simulations suggest that galaxy assembly is a hierarchical process, wherein mergers and interactions play a significant role in determining the present-day morphologies and stellar constituents of galaxies. At present, the observational data are ambiguous, and results suffer from selection effects and small number statistics. The pioneering studies during the last decade have been largely restricted to small pencil beam surveys or shallow surveys of the low-redshift galaxy population (e.g., CFRS - Lilly et al. 1996; LDSS/Autofib - Ellis et al. 1996; various Keck surveys - Cowie et al. 1996, Koo et al. 1996, Cohen et al. 1999). More recently, observations with *HST* and the Keck telescopes have demonstrated the existence of star-forming galaxies at redshifts beyond 3 (Steidel et al. 1996, 1999). These galaxies are believed to be the building blocks of the present-day galaxies, but their properties (masses, chemical composition, stellar content, ages and evolutionary histories) remain largely unknown.

In order to address the question of galaxy formation and evolution for the entire population, we would ideally want to trace the evolutionary history of galaxies (i.e., their star-forming history, chemical evolution, merging and morphological evolution) as a function of mass, redshift and environment. It is critical to understand and interpret the formation of galaxies in the context of structure formation and evolution. These astrophysical problems are inextricably linked, since the large-scale environment plays a crucial role both in the assembly of galaxies and in their evolution (through merging, exclusion, harassment, etc.).

Need for Large Area / Depth / Large Samples

To trace the evolutionary history of galaxies as a function of environment, we need to sample galaxies over the entire range of environments: from the lowest density regions (voids) to the rarest high-density environments (cores of rich clusters). Surveys over large volumes (~100 square degrees — see previous section) are needed to accomplish this.

In order to explore the early evolutionary history of galaxies ($z\sim$ 4-6), deep spectroscopic surveys (to ~25 AB mag) are essential. These depths are needed not only to probe the highest redshifts, but also to ensure that the samples are not restricted to the rarest, most luminous objects, and instead sample more typical objects as well (i.e., as much of the luminosity function as possible).

Galaxies in the present epoch exhibit a large range in physical properties (e.g., masses, chemical abundances, star forming histories, morphologies, gas and stellar content). Since the evolutionary processes responsible for these properties are numerous and complex (e.g., star formation, mergers and interactions, infall), an obserational program to unravel the formation histories of present-day galaxies inherently requires large samples (~10^6 galaxies). For example, in order to trace the evolutionary history of galaxies as a function of mass, redshift and environment, we would need at least 5 redshift bins ($1<z<6$), 4 mass bins (logarithmic intervals spanning 10^8 – 10^{12} $M_\odot$), 5 bins in mean

[8]Contributed by Arjun Dey and Joan Najita

stellar age or star-formation rate (0.1 – 10 Gyr or 0 – 10^3 $M_\odot$/yr), 4 bins in morphology (E/S0, S, Irr, multi-component), 4 bins in mean chemical abundance (0.002 – $2Z_\odot$), and 3 bins in environmental density (field, groups, clusters). With at least 100 galaxies per bin, this implies a total sample size of at least 500,000 galaxies. This is truly a lower limit since we have required that the bins of rare objects are also well populated. Populating such rare bins is critical in order to address questions such as the formation history of the most massive galaxies, or the evolutionary history of the most metal-poor galaxies.

A Representative KAOS Project

The surface densities on the sky of R<25 AB mag $z\approx3$ and I<25 AB mag $z\approx4$ galaxies are ≈ 4000 and 800 per square degree respectively. A survey of 5×10^5 galaxies over 100 square degrees requires a wide-field, highly multiplexed multi-object capability. We are only beginning to scratch the surface with current studies: a typical Keck+LRIS campaign results in ~ a dozen redshifts of R~ 24 – 25 galaxies per night. The low-resolution spectra that are obtained probe only a restricted class of objects (mostly low-extinction, star-forming systems) and are generally only sufficient for measuring redshifts. Detailed investigations of the physical properties of these objects and their evolution requires spectroscopic capabilities which are neither available nor planned.

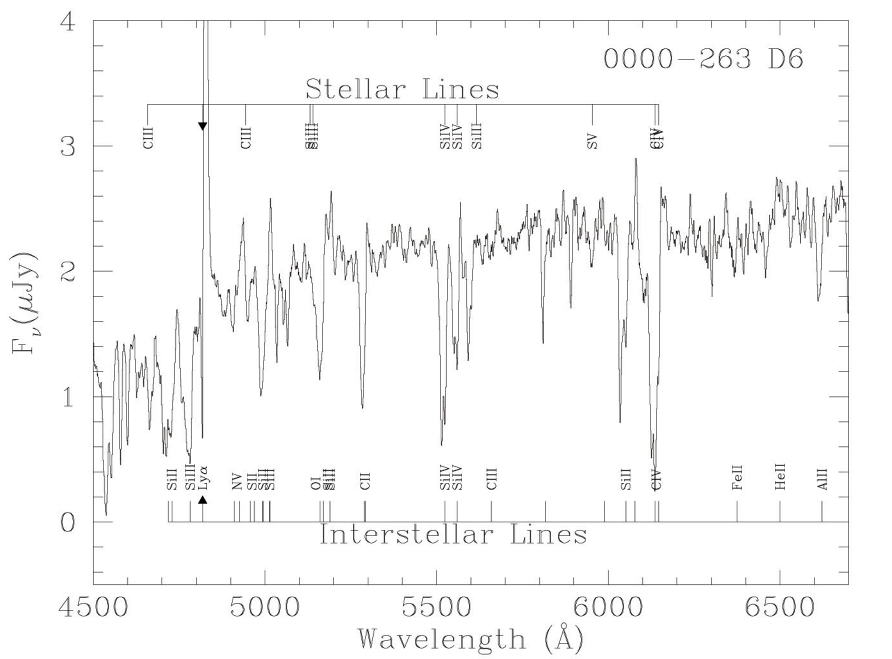


Figure 2.4.3: Spectra of distant galaxies at $\lambda/\Delta\lambda$>1000 will probe stellar content, chemical abundance, and stellar and gas kinematics. This 4.8-hour $\lambda/\Delta\lambda$=1200 spectrum of a very bright (R≈23) star-forming galaxy at a redshift z=2.96 obtained with the W. M. Keck Telescope (Dey et al. 1999) is dominated by absorption lines from the interstellar medium and stellar winds, but also shows weak photospheric absorption lines of hot stars.

For the detailed spectroscopic studies described here, we need a minimum resolution of $\lambda/\Delta\lambda\approx$ 2000 and signal-to-noise ratios per resolution element of ~20 (e.g., figure 4.3). The redshift range targeted by this survey requires spectroscopy at both optical and infrared wavelengths. In 0.7" seeing with DEIMOS on Keck, this requires an exposure time of ≈1 night per 80-object mask; hence it will take 6250 clear nights to obtain spectra of 5×10^5 galaxies, or a minimum of 17 clear observing years with Keck and DEIMOS dedicated to this one project. In comparison, KAOS would execute this project in only ~100 clear nights. A large spectroscopic study of this nature would yield a very large slit survey of 'blank' sky, resulting in a high potential for serendipitous discoveries (e.g., of extremely high redshift objects not directly targeted by the study). This study would be a by-product of the surveys required for the *w(z)* Dark Energy project described in Chapter 2.

4.3 The Growth of Structure[9]

One of the most important problems in astrophysics is understanding how structures in our Universe have evolved from the tiny spatial fluctuations observed in the cosmic microwave background to the clumpy and highly non-Gaussian matter distributions that we observe today. Thus far, our understanding of structure formation, lacking comprehensive observational data, has been largely guided by theoretical studies. Numerical simulations of structure formation have now attained a high degree of sophistication, and demonstrate that clear differences between competing cosmological structure formation models may be observable at large lookback times (z~1-5; Figure 4.5). Confronting these models with observations of the matter distribution at a range of lookback times remains one of the most robust methods of discriminating between present competing models and constraining future theories of structure formation. This requires obtaining redshifts for galaxies over very large areas of the sky (to probe the characteristic scales on which structures exist) to very faint limiting magnitudes (to probe the distant universe). Redshift surveys would not only probe the overall topology of the matter distribution, but also the detailed dynamical history of galaxies in a variety of environments from the sparse voids to the densest clusters.

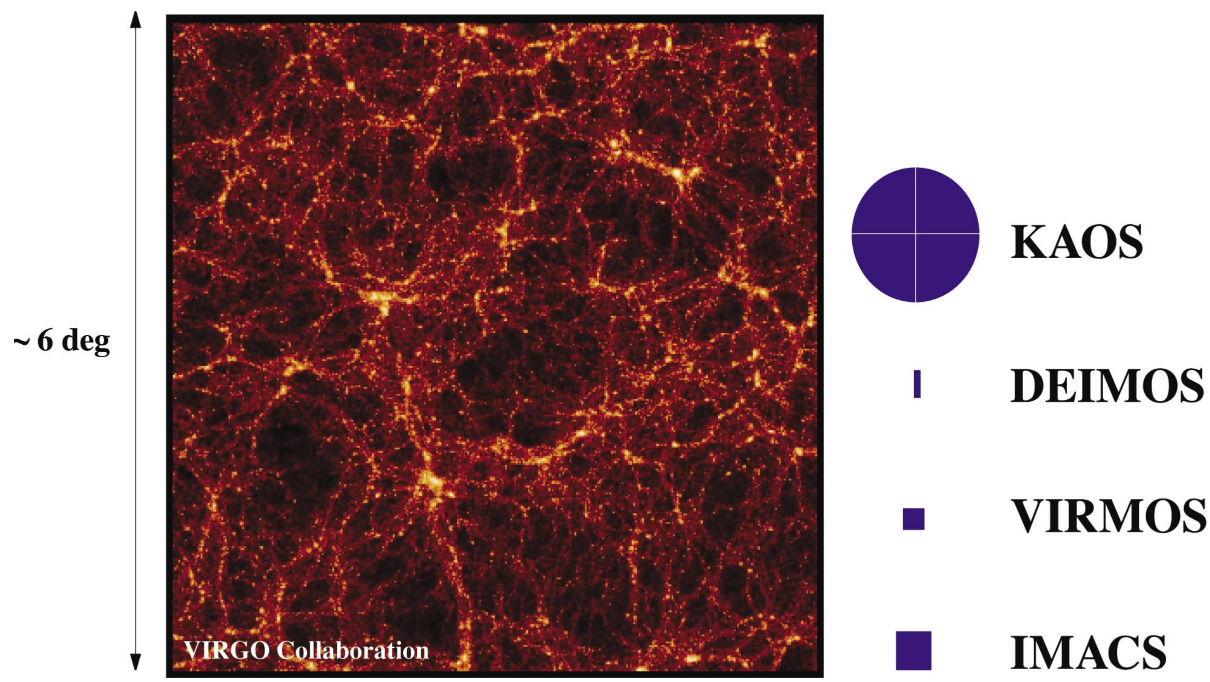


Figure 2.4.4: Numerical simulations of the dark matter distribution predict structures on scales of clumps and filaments (arcminutes) to large voids and walls (several degrees). Characterizing this structure from observations of galaxy clustering requires a wide-field, highly multiplexed spectroscopic capability.
(Picture courtesy J. Coelberg; Jenkins et al. 1998)

[9]Contributed by Arjun Dey, Joan Najita and George Jacoby

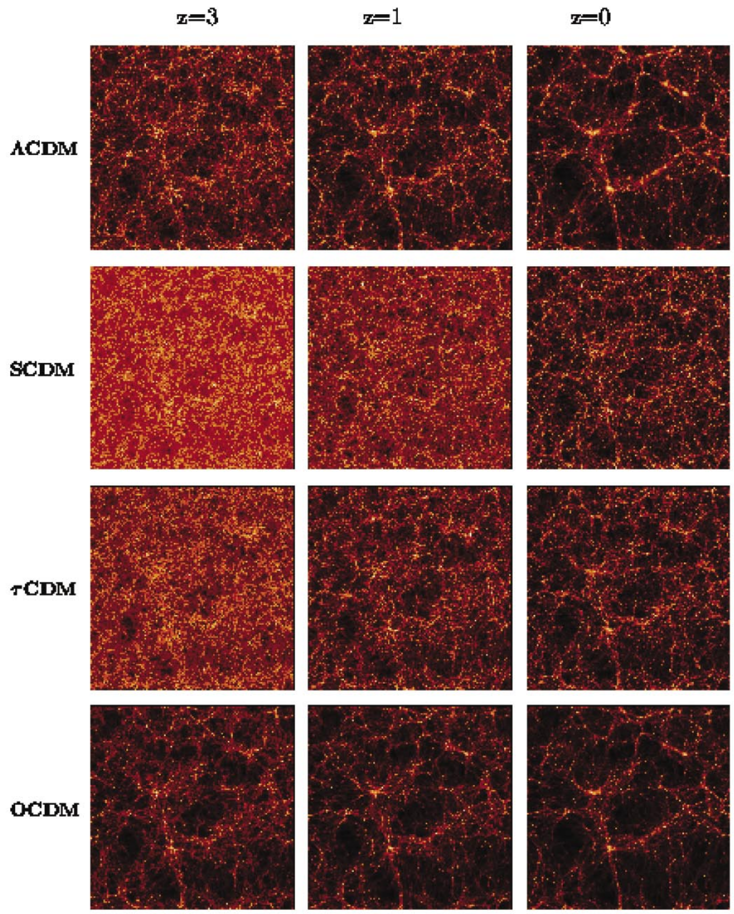


Figure 2.4.5: In numerical simulations of the growth of large scale structure, differences between competing models are especially obvious at large lookback times. In these simulations from the VIRGO Consortium (e.g., White 1997), the boxes ($240h^{-1}$ Mpc on a side, corresponding to ~6° at z=1) show the dark matter distributions at different epochs and predict discernible differences at z>1. (Picture courtesy Joerg Colberg; Jenkins et al. 1998 ApJ 499,20)

Need for Large Area / Depth / Large Samples:

Current surveys suggest that there is significant structure on ~$100h^{-1}$ Mpc scales (corresponding to ~6° at z=1), comparable to the largest scales thus far investigated (e.g., Landy et al. 1996, Doroshkevich et al. 1996). The mass distribution may be comprised of large filamentary or sheet-like structures which are inferred from pencil-beam

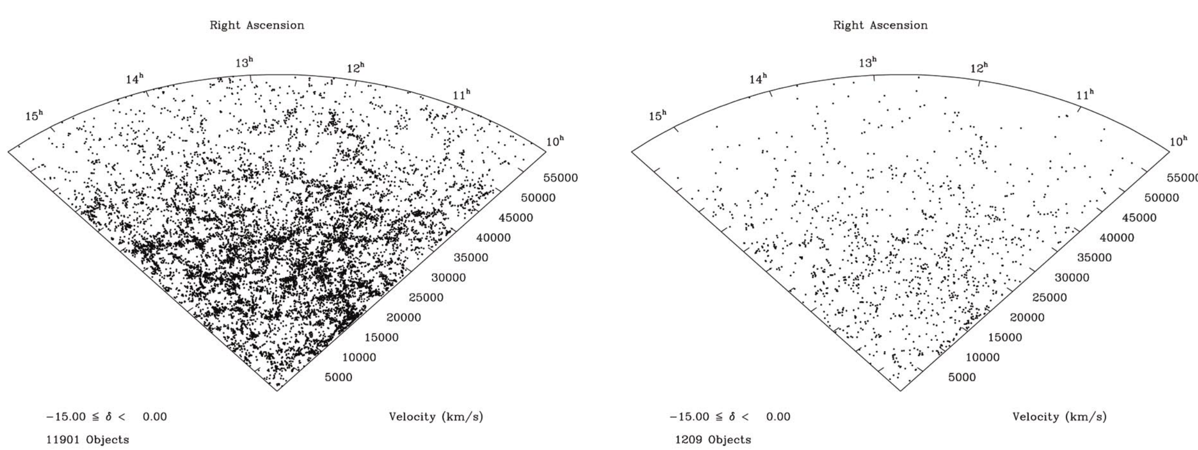


Figure 2.4.6: The need for dense sampling: when all the galaxies in the Las Campanas Redshift Survey (Left) are sampled at 10% (Right), the strongest features disappear. (Figure courtesy M. Postman).

surveys — little is known about the true large scale angular distribution at each epoch due to the paucity of redshift information over large angular scales for all but the very nearest galaxies (z~<0.2; e.g., Davis et al. 1982, Shechtman et al. 1996, Ellis et al. 1996, Geller et al. 1997, da Costa et al. 1998). This implies the need for spectroscopic surveys over several times this characteristic scale (i.e., >10°x10°) in order to sample the large volumes that contain these structures (e.g., Figures 2.4.4 and 2.4.5).

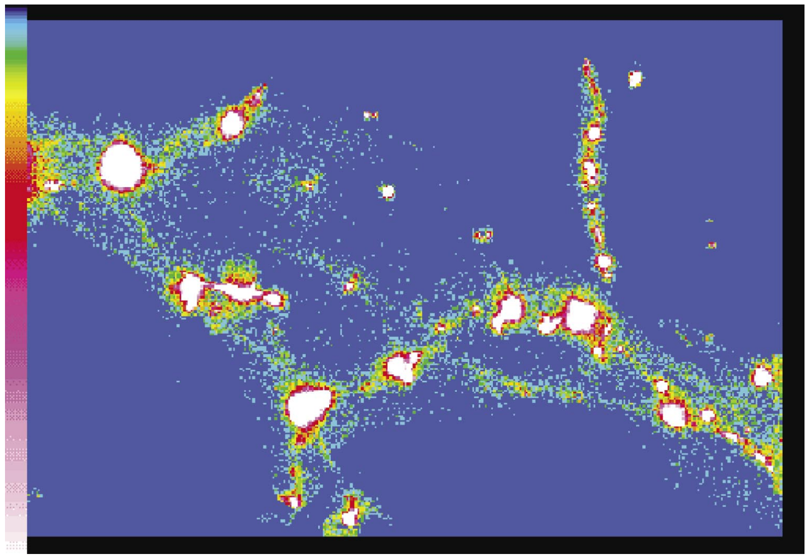


Figure 2.4.7: Simulation by M. Pierre (1998) of an *XMM* observation of an X-ray emitting filament located between two z=0.5 clusters (not shown). By measuring the velocities of galaxies in the different regions, KAOS will locate the filaments and clumps in velocity space and measure the total mass in the virialized regions. The region shown is ≈1.25°x1.82°, and is well matched to the FOV of KAOS.

A deep survey (R~25 AB mag) is needed to densely sample the high redshift populations and characterize the three-dimensional matter distribution to redshifts beyond 4. Dense sampling is needed to detect and highlight compact structures: most existing surveys of the sky are non-contiguous and sparsely sample the galaxy distribution. Sparse sampling is insensitive to strong clustering (Figure 4.6) and results in significant uncertainties in the derived structure statistics due to both sampling and 'cosmic' errors (e.g., Szapudi & Colomb 1996, Szapudi & Szalay 1996). Since the simulations show structures on scales ranging from tens of kpc to ~100 Mpc, the observations must be sensitive enough to sample the velocity fields on small and large scales, as well as comprehensive enough to have sufficient numbers of objects on all scales in each redshift interval to provide robust statistical measures of the spatial structures.

A Representative KAOS Project

To study the growth of structure from z≈5 to z≈1, we would map out galaxy redshifts over multiple ~10°x10° fields. Dividing the redshift range crudely into 10 bins of Δz=0.3, and requiring ~1,000 galaxies per square degree per Δz bin for an accurate estimate of a clustering statistic, we would need a total sample size of about 10^6 galaxies per field. To R~25 AB mag, the surface density of galaxies is about 10^5 per square degree, the majority of which lie at z>1. This program will reveal the evolution of the matter distribution as traced by the galaxies as a function of redshift, and provide a relatively unbiased discovery tool for high redshift clusters and groups. An investigation of the velocity fields of galaxies in regions imaged at X-ray wavelengths (e.g., Figure 4.7) will isolate the hot X-ray emitting gas in redshift space, determine the fractional baryonic mass in galaxies versus that in hot gas, and provide a direct measure of the ratio of baryonic to total mass in filaments and clusters. Measuring redshifts of ~10^6 galaxies at a resolution $\lambda/\Delta\lambda \approx 1000$ requires of order 200 clear nights with KAOS, whose field of view is well matched to the angular scale of structures predicted from simulations (Figure 2.4.5). A project of this scale would require about 12,500 nights with Keck and DEIMOS, and about 4,000 nights with Magellan and IMACS per 100 square degree field.

A Kinematic and Chemical Study of Intracluster Stars

In addition to mapping out the evolution of the large-scale topology of the matter distribution, a project of this scale would simultaneously yield detailed dynamical information on clusters of galaxies over a large range in redshift. Since the densest clusters identify sites of the strongest initial perturbations and the earliest ones to collapse, interactions play an important role in shaping the evolution of galaxies in clusters and perhaps in the chemical and stellar enrichment of the inter-cluster medium as a whole. The dynamical analysis of member galaxies in clusters at high redshift, when combined with detailed kinematic and chemical studies of galaxies and intracluster stars and gas in clusters at the present epoch, can yield a comprehensive understanding of the chemical enrichment of the Universe in these over-dense regions. In particular, the intracluster stellar component in nearby clusters contains a fossil record of the past dynamical history imprinted in the kinematics of the intra-cluster stellar population. A chemical and kinematic study of the intra-cluster planetary nebulae (a good velocity tracer population that is relatively easy to identify from the ground) in these clusters, would reveal the entire tidal history of many cluster galaxies over the evolutionary history of the clusters! For example, the areal density of intracluster planetaries in Virgo (with line fluxes F([OIII]λ5007)>10^{-17} erg s^{-1} cm^{-2}) is ≈ 1000 per square degree, and the cluster spans an area of ~ 100 square degrees. Obtaining moderate resolution, moderate signal-to-noise ratio spectra of ≈10^5 planetaries requires ~20 nights with KAOS.

4.4 AGN Physics and High Redshift[10]

The aim of this project is to understand the physics of AGN evolution from the end of the epoch of re-ionization (z=6.5) to the present day (z=0). This will be achieved by a conducting a comprehensive survey for over 30000 AGN with which to carry out detailed studies of their demographics and physical properties (space density, clustering, environments and black hole masses). This will complement the increasingly detailed dynamical studies of local AGN with similar luminosities that will be carried out over the coming decade with space-based facilities and IFU capabilities on ground-based telescopes.

Science Summary:

A large sample of faint AGN at z<6.5 will enable unprecedentedly detailed scientific studies to be made of:

1. Hierarchical models of QSO formation from z=6.5 to the present day.
2. The evolution of black holes in AGN from z=6.5.
3. The AGN contribution to the UV ionizing background.

[10]Contributed by Brian Boyle

Current Position:

Our knowledge of the statistical properties of the AGN population has been greatly enhanced by the two major QSO surveys of recent years: the 2QZ (Croom et al. 2001) and the SDSS (Richards et al. 2001). In many respects the surveys provide complementary information; the deeper (g<21) 2QZ survey probes ~1 mag below the break in the QSO luminosity function (L*) at z<2.5, whereas the shallower SDSS survey extends to much higher redshift, z<6.5, courtesy of its superior CCD-based photometric selection techniques.

The statistical precision provided by samples of over 20000 AGN has yielded new insights into many aspects of the QSO phenomenon, the QSO luminosity function (Boyle et al. 2000) QSO clustering (Croom et al. 2001), the z>5 Universe (Fan et al. 2001), limits on cosmological parameters (Hoyle et al. 2002) new populations of AGN (Brotherton et al. 1999, Hall et al. 2002) the physics of QSO emission lines (Croom et al. 2002) and QSO black hole masses (Corbett et al. 2003).

However, even surveys such as these still provide us with an incomplete view of the QSO phenomenon — missing a vital area of observational parameter essential for building up a full understanding of AGN physics and evolution. Both the SDSS and 2QZ are limited to magnitude (g<21) which at best probe only 0.5 – 1 mag below L* at low redshift (z<2) and only sample the very brightest end of the AGN luminosity function at z>4.

To achieve the a comparable range in luminosity coverage of the z~2 AGN luminosity function (the peak epoch of QSO activity) compared to the existing z=0 galaxy LF requires samples extending to B<25. Equivalent depths (I~24) for z>5 would reach ~ 2 mags below the break at these redshifts providing, for the first time a full picture of the high redshift AGN demographics.

Unfortunately, due to the relatively low surface density of QSOs (<0.1 QSOs/sq arcmin even at g~25) and the small fields-of-view available to MOS devices on 8m-class telescopes required to identify AGN at faint magnitudes, the faintest AGN studies conducted to date have been limited to B<22 and comprise less than 100 QSOs in total.

KAOS contribution

With its uniquely wide field-of-view, KAOS is the only instrument, either constructed or planned, capable of conducting a major new spectroscopic survey of AGN at the depths required to provide a full statistical picture of the AGN phenomenon over the full range in redshift between the end of the epoch of re-ionization and the present day. Although the shape and form of the AGN LF is unknown outside the L,z range probed by current surveys (B<22, I<21), extrapolation of the current AGN LF evolutionary models for z<2.5 obtained from the 2QZ (Boyle et al. 2001) and z>3 from the SDSS (Fan et al. 2002) yield the following AGN space densities at typical depths attainable with Gemini using nod-and-shuffle techniques:

Mag Limit	Redshift Range	Space Density of QSOs	Space Density of Other AGN
B<25	0<z<2.5	300 deg^{-2}	300 deg^{-2}
R<24.5	2.5<z<4	30 deg^{-2}	300 deg^{-2}
I<24	4<z<6.5	3 deg^{-2}	50 deg^{-2}

To ensure completeness, complete color-selected samples of AGN will also contain significant contamination from galactic stars. Indeed, for the key redshift range 2.5<z<3.5 (i.e. the epoch of peak QSO activity) this contamination can be as high as a factor 10 of greater. With its extremely high multiplex capability, KAOS is unique in being able to access simultaneously all AGN candidates (1500 per field) based on a color selection designed to yield a complete sample of the full range in redshift from 0<z<6.5.

Proposed KAOS experiment

To remove the effects of cosmic variance from any studies of QSO clustering, the survey will need to be conducted over an area whose minimum spatial dimension corresponds to a comoving length of ~ 200/h Mpc. At z>0.5 this corresponds to a maximum angular size of 10°. A survey covering 100 deg^2 would therefore produce 30000 z<2.5 AGN, 3000 AGN with 2.5<z<4 and 300 AGN in the range 4<z<6.5. Based on these numbers the AGN power spectrum would be determined to better than 5 per cent precision at z<2.5, and an estimate of the scale length of the AGN correlation length determined to better than 20% for all Δz=1 between z=0 and z=6.5. At resolutions of Δz=0.25, Δm=0.5 mag, errors on the estimates of the QSO LF would be better than 5% for all redshifts z<4. At the very highest redshifts, errors of the z~6 LF would be typically 10%. Note that the proposed survey would probe a factor of more than 10 further down the AGN LF at all redshifts than any previous survey. Indeed 90% of the AGN identified in this probe a region of (L,z) space unexplored by any previous survey.

The survey would provide information on high redshift AGN at similar luminosities to those low redshift AGN that will be studied with increasingly detailed dynamic/kinematical observations over the coming decade. Coupled with surveys such as SDSS and 2QZ, it will provide an unprecedentedly wide baseline in luminosity and redshift over which to:

- test the luminosity-dependence of AGN clustering predicted by recent hierarchical models of QSO formation (Kauffmann & Haehnelt 2002).

- determine the redshift-dependence of AGN clustering from 0<z<6.5 independent of luminosity. The redshift evolution of AGN bias at a fixed luminosity is a fundamental input parameter to models of QSO formation.

- extend our knowledge of the z>4 AGN LF below L*. This will provide discrimination between models of AGN evolution (density *vs.* luminosity) at these redshifts and providing an accurate (convergent) estimate of the ionizing background due to AGN in the early Universe.

- use the mean spectral properties of AGN (see Corbett et al. 2003) to derive black hole mass estimates over a wide range of luminosities and redshifts. Direct comparison of black hole mass estimates obtained from a KAOS survey for AGN at high redshift may be directly compared to black hole mass estimates in low redshift AGN of similar luminosity obtained directly from dynamical studies, to infer the evolution of fuelling rates/ black holes masses within AGN.

- coupled with the Dark Energy project, it will provide a unique opportunity to investigate the evolution of AGN environment over the redshift range 1<z<3.

The proposed survey relies on the availability of deep B<25, I<24 digital data covering the full range of the optical/IR passbands U->K for optimal AGN selection. It is assumed that, by the time KAOS is completed, that this data will exist (at least over 100 deg^2) accessible from both hemispheres (i.e., PANSTARRS/UKIDSS combination in the Northern Hemisphere, VST/VISTA in the Southern Hemisphere).

The survey can also be viewed as providing important complementary information to proposed deep surveys in other wavebands (X-ray: *Chandra*, radio: eVLA, UV: *GALEX*), many of whose source detection will be AGN.

The proposed survey will require approximately 70 pointing to cover the full 100 square degrees. Over each 1.5deg^2 KAOS field-of-view there will be 2,000 AGN candidates; ideally matched to KAOS's 4,000 fibers to enable a 'sky-object' paired fiber configuration, so that nod-and-shuffle can be implemented with 100% time on target, i.e., with no efficiency loss. Based on previous experience (2QZ), a minimum signal-to-noise ratio SNR ~ 5 per 4Å resolution element will provide 95% complete identification rate amongst an AGN sample. For the magnitude limits specified above, this SNR is achieved in 10,000 secs exposure; assuming median seeing of 0.8 arcsec (*R* band), system efficiencies of 0.1, 0.24 and 0.31 at 440nm, 670nm and 870nm respectively. In total, this survey could be completed in 25 dark nights.

The combination of high target density yet large coverage means that this survey would be impractical on any other existing facility. It would take over 10 times as long to complete using Magellan/IMACs (field of view 6 times smaller than KAOS, integration times longer by factor 2), 20 times as long with either VLT/FLAMES (f.o.v. 6 times smaller target density 3 x higher than offered by FLAMES) or MMT/Hectospec (f.o.v. 3 times smaller than KAOS, integration times 2 times longer, survey target density 3 x higher than offered by Hectospec). Furthermore, at the proposed survey depths (2 per of sky), the use of nod-and-shuffle is crucial. Currently, the lack of nod-and-shuffle on these other facilities would make the proposed survey impossible, no matter how much time was devoted to it.

4.5 The Relationship Between the IGM and Galaxies at High Redshift[11]

Introduction

One of the long-standing problems in investigations of the intergalactic medium (IGM) is the nature of the relationship between the gaseous and luminous constituents of the universe. By mapping the distribution of galaxies and of the IGM over the same volume, we can examine how these two constituents affect each other during the process of the formation and evolution of galaxies. Observations of this nature can provide statistical constraints on the feedback (radiation, winds, metal enrichment) from star formation in galaxies to the surrounding IGM.

To be specific, moderate resolution (*R*=1000 to 5000) spectroscopy of 120 quasars with redshifts between 2 and 4 drawn from a contiguous 10 square degree region of the sky will provide a mapping on scales of 10 to 80 Mpc of the properties of the IGM. When combined with extensive galaxy redshift surveys in the same field we will be able to investigate the relationship between the IGM and luminous matter in the same volume of the high redshift Universe. While these observations will not resolve individual subcomponents of the Lyman-alpha forest, the variation in the column density of absorption by neutral hydrogen will be traceable from z=1.6 to the redshifts of the available background quasars. In addition, metal absorption lines are sensitive tracers of the gas associated with the ISM of galaxies along the sight-lines to the background quasars. Even galaxies too faint to observe spectroscopically in the galaxy redshift survey would still be detectable by the absorption the gas in their interstellar media would cause in the quasar spectra. Quasar absorption line systems with significant absorption by metals (like the higher column density systems produced by gas directly associated with the ISM of galaxies), will produce identifiable absorption by Mg II (observable with KAOS over the range $0.3<z<2.9$) and C IV (observable with KAOS over the range 1.3 to 6.1).Combined with similar studies at lower redshifts, we will be able to trace the evolution of the IGM, stellar populations, and their impact on each other from redshifts of $z\approx4$ to the present epoch.

[11]Contributed by Buell Jannuzi

The IGM in 2003

The combination of the UV spectrographs on the Hubble Space Telescope and Echelle spectrographs on large ground-based telescopes have enabled us to study the properties of the IGM from redshift 6 to 0 (e.g., Kim et al. 2002; Kim et al. 1997; Weymann et al. 1998). Such data, when combined with the interpretive insight provided by modern cosmological simulations (e.g., Dave et al. 1999; Cen et al. 1998), have provided new insights into what drives the evolution of the observed properties of the IGM. Measurements of the fluctuations in the IGM have been used to map the underlying dark matter distribution (e.g., Croft, R. et al. 2002) and help constrain the most accurate determination of cosmological parameters (*WMAP*; Bennet et al. 2003, Spergel et al. 2003).

There have been extensive efforts to study the relationship between the IGM and the luminous matter (as traced by individual, groups, and clusters of galaxies). However, these are generally frustrated by the difficulty of obtaining detailed information about the distribution of the IGM and of galaxies in the same cosmological volume. Most studies have used *HST* spectroscopy of a small number of lines-of-sight toward quasars (generally less than 20) and incomplete redshift surveys of the brighter galaxies in modest fields (less than 15' in field-of-view) centered on the quasars (e.g., Chen et al. 2001; Chen et al. 1998; Le Brun, Bergeron, Boise, 1999). The small fields-of-view and low redshifts, $z<0.8$, of the quasar absorption line data make it difficult to identify the larger scale structures with which both the IGM and individual galaxies might be associated. This has complicated the interpretation of the observational results, yielding conflicting interpretations of effectively the same observational results. A few studies have used extensive redshift surveys over larger areas, but have been limited by the available UV spectroscopic data (e.g., Morris et al. 1993; Grogin and Geller 1998).

Another approach has been to obtain spectra of multiple quasars in a single large field in order to improve the knowledge of the distribution of gas in the IGM in the surveyed volume (e.g., vanden Berk et al. 1999; Impey, Petry, and Flint 1999). While yielding interesting constraints on the correlation lengths and/or sizes of structures traced in the IGM, these studies have still provided less than 10 probes of the IGM across scales of several degrees (i.e., only ~10 sight-lines over a region of more than 25 square degrees).

Similar investigations at high redshift are only just beginning. Working with the world's largest telescope, researchers are now obtaining many redshifts (i.e., hundreds) of Lyman-break galaxies in the fields of high redshift quasars. This has allowed the investigation of the interaction between galaxies with strong winds (driven by the forming stars they contain) and the surrounding IGM (Adelberger et al. 2003). Such studies are still limited to a relatively small number (<10) of isolated lines-of-sight (i.e., a single quasar in a survey region of 10s of arcminutes) and sampling a limited redshift range (centered around $z=3$, where the Lyman-break technique used to feed the galaxy redshift survey in this study was optimized).

Modern cosmological simulations can provide detailed predictions for the relative distribution of gas in the IGM including modeling the metal enrichment of the gas (e.g., Cen and Bryan 2001; Croft et al. 2002). This enables the prospect of being able to interpret the richer data set that would be provided by a program using KAOS to obtain a redshift survey and spectra of AGN in the same high redshift volume.

A KAOS Survey of the IGM

To investigate in greater detail the relationship between the IGM and other structures requires a three-dimensional map of large-scale structures as traced by galaxies, coupled with moderate resolution, high signal-to-noise ratio spectroscopy of many QSOs behind the same survey volume. A survey covering roughly 10 square degrees for the complementary galaxy redshift survey and including of order 100 QSO sight-lines probing the same volume, would map structures between redshift 2 and 4 on angular scales of 1 to 80 Mpc (H_0=70 km/s/Mpc, Ω_m=0.3, Ω_λ = 0.7). Such a survey would provide more than an order of magnitude improvement in the density of sampling of the structure of the

IGM at these epochs, providing unprecedented constraints for the models of the IGM and its relationship to galaxies at this epoch of the Universe.

High spectral resolution (R~40,000) spectroscopy of the majority of the background QSOs, while desirable, would be prohibitively expensive in 8-m class telescope time even with the multiplexing advantage of KAOS. However, low-to-moderate resolution (R=1000 to 5000) spectroscopy, while still requiring many long exposures, would be feasible for quasars as faint R~22. This would make possible obtaining spectra of 30 to 100 high redshift ($2<z<4$) AGN over the survey area. Such spectra, while not capable of resolving the individual sub-components of the Lyman-alpha forest, could provide maps of the variation in the opacity caused by fluctuations in the column density of neutral hydrogen along each line-of-sight. These can be compared to the distribution of galaxies in the complimentary redshift survey. To probe variations in the column densities as low as 10^{14} cm^{-2} in neutral hydrogen will require a SNR per spectral resolution element (5.23Å for R=1000 mode of KAOS) of at least 10. Interpretation of the spectra will be limited by our ability to determine or model the intrinsic AGN continuum emission, observed as modified by the IGM in our data. The SDSS is currently in the process of developing techniques identifying significant variations in individual QSO spectra when compared to a low-resolution template, based on their low-to-modest-resolution spectra of high redshift ($z>3$) AGN (Burles 2003), but this will be the major constraint in interpreting the spectra. The spectra will provide information about the distribution of $N(H)$ over the redshift range 1.6 out to the redshifts of the background AGN.

Simultaneously, we will gather information about the distribution of metal-line absorbing gas in systems with $N(H)>10^{13.5-14.0}$ cm^{-2} over a redshift range 0.28 to 6.1 (or the redshift of the background AGN). Over this range of column densities, the Mg II λλ2796,2803 doublet should be observable in the optical window for intervening gas at redshifts $0.28<z<2.9$ and the CIVλλ1548,1550 doublet will allow investigation over the redshift range 1.3 to 6.1.

The surface density of quasars with $R<22$ mag is approximately 150 per square degree (Boyle 2003), yielding about 250 QSOs per KAOS field. Most of these quasars will not have high enough redshifts to be of use in studying the distribution of $N(H)$ (as traced by the absorption by the Lyman-alpha forest) at high redshifts ($1.6<z<6$), but there should be between 5 to 25 such objects per KAOS field, 30 to 150 in the entire survey area. The low surface density of QSOs suggests that this project could be done in conjunction with the proposed redshift survey for the $w(z)$ Dark Energy project.

Total exposure times with KAOS are significant, but not prohibitive. For example, to obtain an R=1000 spectrum at 5000Å with SNR of 10 per 5.2Å spectral resolution element for an R=22 mag AGN (power-law spectrum) 4.5 hours per field, or roughly 30 hours for the entire survey. Covering larger areas than 10 square degrees (to mitigate the effects of cosmic variance on the conclusions) is therefore not out of the question. To increase the spectral resolution to R=5000 would require an increase to a total integration time of 20 hours per KAOS field. This would move the project into one requiring a large time commitment (100+ hours to complete the proposed survey), but might be possible as a major project for the instrument.

No extant facility can provide the data as efficiently with KAOS. Since the typical separations between target quasars will be ~5-20 arcminutes, current wide-field spectrographs like DEIMOS would be able to observe only 1 to 3 QSOs at a time, requiring ~80 pointings for every one of KAOS. The proposed survey requires the large field of view of KAOS and the light gathering capability of Gemini or larger telescopes.

References

Adelberger, K. L., Steidel, C. C., Shapley, A. E., and Pettini, M. 2003, ApJ, 584, 45
Boyle, B. 2003
Cen, R., and Bryan, G. L. 2001, ApJ, 81, 546L
Cen, R., Phelps, S., Miralda-Escude, J., Ostriker, J. P., 1998, ApJ, 496, 577

Croft, R., Weinberg, D. H., Bolte, M., Burles, S., Hernquist, L., Katz, N., Kirkman, D., and Tytler, D. 2002, ApJ, 581, 20
Croft, R., Hernquist, L., Springel, V., Westover, M., and White, M. 2002, ApJ, 580, 634
Dave, R., Hernquist, L., Katz, N., Weinberg, D. H., 1998, ApJ, 511, 521
Impey, C. D., Petry, C. E., Flint, K. P. 1999, ApJ, 524, 536
Kim, T.-S., Carswell, R. F., Cristiani, S., D'Odorico, S., and Giallongo, E., 2002, MNRAS, 335, 555
Kim, T.-S., Hu, E. M., Cowie, L. L., and Songaila, A., 1997, AJ, 114, 1
Morris, S., Weymann, R. J., Dressler, A., McCarthy, P. J., Smith, B. A., Terrile, R. J., Giovanelli, R., and Irwin, M., 1993, ApJ, 419, 524
Grogin, N. A., and Geller, M. J. 1998, ApJ, 505, 506
vanden Berk et al. 1998, ApJS, 122, 355
Weymann, et al. 1998, ApJ, 506, 1

4.6 Stellar Populations with KAOS[12]

What is the Nature of the Halo/Bulge of M31?

Cosmic variance demands that we extend our study of resolved stellar populations to other galaxies in the Local Group. In particular, the large disk galaxies (M31, M33) provide important constraints on disk galaxy formation, combined with those from the Milky Way (see Chapter 3). Indeed despite the bulge/halo population of M31 being the defining Population II (Baade 1958), we now know it to have a strong metal-poor component, with a metallicity distribution more like the thick disk of the Milky Way than that of the 'Population II' stellar halo.

The field stellar population of M31 has been studied through colour-magnitude diagrams (both ground-based and from the Hubble Space Telescope) of its evolved stars and spectroscopy of bright red giant branch stars. All studies (Durrell et al. 2001, Holland et al. 1996, Rich et al. 1996, Rich et al 2001, Ferguson 2001), following the pioneering work of Mould & Kristian (1986), find that the dominant field population probed down the minor axis has a mean metallicity[13] of around –0.6 dex, from projected distances of ~ 5 kpc out to ~ 20 kpc (the 'bulge' minor-axis effective radius is ~ 1.5 kpc, significantly larger than that of the Milky way; Walterbos & Kennicutt 1988). The metallicity distribution is asymmetric, and can be fit by the superposition of two populations, metal-poor and metal-rich, with peaks similar to the globular clusters, at ~ –1.5 dex and ~ –0.6 dex. The bulk of the stars, even out at 20 kpc, is in the metal-rich population. This raises the issue of which stars in the Milky Way should have been identified as 'Pop II', by comparison with the field halo of M31 — perhaps the members of the Milky Way thick disk, whose mean metallicity is comparable to that of the dominant population in M31's 'halo' (Wyse & Gilmore 1988). Indeed, perhaps the 'halo' in M31, which is rather flattened with an axial ratio of ~ 0.6, is actually a thick disk (cf. Wyse & Gilmore 1988). This could have been formed during a significant merger.

But what is the merging history of M31 — What merged, and when? Is the merging history similar or very different to that inferred for the Milky Way?

---

[12]Contributed by Rosie Wyse

[13]Most of these metallicities are based on the colour of the red giant branch and are subject to calibration uncertainties including the elemental abundance mix.

Intriguing evidence for a more recent (minor?) merger has been found through wide-area star counts of the evolved population in the 'halo' of M31 using the Wide Field Camera on the Isaac Newton Telescope (Ibata et al. 2001; Ferguson et al. 2002), which revealed large-scale (tens of arcmin, or many kiloparsecs) substructures through significant overdensities, most easily explained as a remnant star stream from tidal interactions. Further, assigning metallicity variations to variations in the color of the red giant branch (i.e., assuming a uniform old age), there are also large-scale chemical inhomogeneities that do not necessarily correlate with coordinate space inhomogeneities (Ferguson et al. 2002).

Kinematics and spectroscopic metallicities are of obvious importance in sorting out what is going on; inhomogeneities in both persist longer than in coordinate space. A wide-area, large multi-plexed spectrograph is the obvious instrument.

What would be the requirements for an efficient survey of the bulge/halo of M31? The Wide Field Camera survey is presently of some 25 sq. degrees, covering a broad (elliptical) annulus extending in the minor axis from ~ 10 kpc to ~ 35 kpc (at the distance of M31, 1.5° is ~ 10 kpc). Consistent with the star counts from Pritchet & van den Bergh (1994) in the innermost fields there are around 50 red giant stars within 2 mags of RGB tip (I=20.5) per sq. arcmin. Out at ~ 30 kpc one still expects around several of these bright giants per sq. arcmin. At a minimum, one requires spectroscopy good enough to measure metallicities to ~ 0.3 dex, and radial velocities to ~ 10 km/s, and thus discriminate between the crude 'disk', 'bulge', 'halo' stellar components and identifying (unresolved kinematically) cold streams.

This is achievable with R ~ 10,000 spectroscopy around the Ca triplet cf. Reitzel & Guhathakurta (2002) KECK study of ~ 100 similar target stars (4hr exposures using 1200line/mm grating and LRIS, again Ca triplet region, 0.62Å/pixel). The target surface density of several per square arcmin, combined with a total survey area of tens of square degrees, argues for several thousand fibres per square degree, and a FOV of around a square degree.

The capability of KAOS as proposed here, with 4000 fibers over a 1.7 sq. degree FOV, (less than one fibre per sq. arc min) would provide well sampled metallicity and radial velocity distributions over the WFC survey area in about 50 pointings. Adequate S/N for the chemical abundances will require ~ 4 hr exposures at I ~ 22.5, and with 50 pointings the survey would need ~ 200 hours, ignoring overheads.

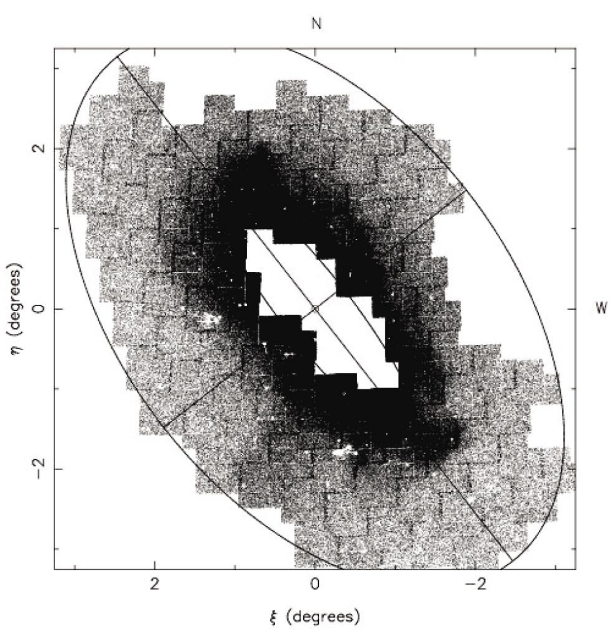


Figure 2.4.8: Star counts of red RGB stars in M31 from the WFC survey (Ferguson et al. 2002).

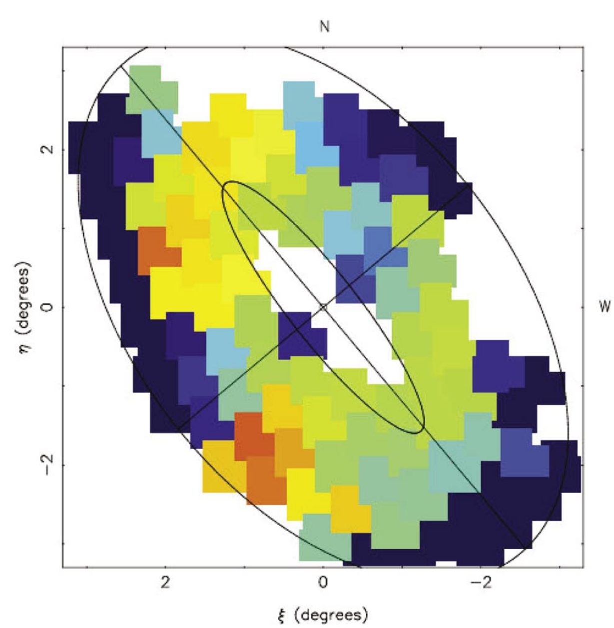


Figure 2.4.9: Map of inferred metallicity from colour of the RGB stars in M31 from the WFC survey; the range is ~ 0.5 dex (Ferguson et al. 2002).

Twenty nights is a reasonable allocation; note that the imaging survey showed the need for full area coverage to find the more interesting bits, since even if no structure in coordinate space some remains in kinematics and in metallicities (perhaps reflecting older mergers? — the proposed spectroscopic survey will tell!).

A new instrument is required; the proposed science cannot be accomplished in a reasonable time with other instruments.

What is the Nature of the Halo/Bulge of M33?

The early work of Mould & Kristian (1986) established that the field stars of M33 some ~ 7 kpc projected distance from the centre of that galaxy, along the minor axis and thus expected to have little contribution from the disk, have a mean metallicity of only ~ -2 dex, with a small spread. There is a kinematic 'halo' as traced by the globular clusters (Schommer et al. 1991). Analysis of the CMDs resulting from deep imaging with the Hubble Space telescope (Sarajedini et al. 2000) has shown that some of these 'halo' clusters have a red horizontal branch despite low metallicity (~ -1.5 dex), perhaps indicating a younger age (~ 7 Gyr), with others probably as old as the classical Galactic halo globular clusters. The field surrounding the globular clusters studied with HST are disk-dominated and show a complex star formation history (Sarajedini et al. 2000).

In contrast with M31, the WFC survey of the upper RGB stars found no inhomogeneities in terms of surface number densities (Ferguson et al 2003). But what of chemical or kinematic inhomogeneties? Again, a spectroscopic survey will reveal all!

Stellar Populations in Andromeda Dwarf Spheroidal Galaxies

Our understanding of the evolution of the low surface brightness, low luminosity satellite galaxies of the Milky Way (the dwarf spheroidals) has changed substantially in the last few years. The theoretical expectation that these galaxies were so fragile that they could sustain only one, short-lived epoch of star formation (e.g., Dekel & Silk 1986) was not fulfilled by the morphology of the CMD of the evolved stars (Smecker-Hane et al. 1994) or by the deep HST imaging data reaching below the oldest turnoff of these systems (Hernandez et al. 2000). The inferred extended star formation histories are also not immediately compatible with the proposals to truncate sharply star formation in low-mass systems to prevent Cold Dark Matter cosmologies producing a surfeit of dwarf galaxies (Bullock et al. 2000).

Some of the dwarf spheroidal companions to M31 have also been shown to contain luminous AGB stars, through (small field-of-view) HST imaging (DaCosta et al. 2000). Further, again like the Milky Way companions, the internal velocity dispersions, measured by radial velocities of a few luminous T-RGB stars using the Keck telescope, with required accuracy of ~ 1 km/s, are sufficiently high that dark matter is inferred to dominate the dynamics (Cote et al. 1999). Mapping the kinematics across the face using the AGB stars is possible in short order with KAOS.

References

A. Dekel & J. Silk, 1986, ApJ, 303, 39
T.A. Smecker-Hane, P. Stetson, J. Hesser & M. Lehnert, 1994, AJ, 108, 507
X. Hernandez, G. Gilmore & D. Valls-Gabaud, 2000, MNRAS, 317, 831
J. Bullock, A. Kravtsov & D. Weinberg, 2000, ApJ, 539, 517
M. Unavane, R.F.G. Wyse & G. Gilmore, 1996, MNRAS, 278, 727
P.R. Durrell, W.E. Harris & C. Pritchet, 2001, AJ, 121, 2557
G. Da Costa, T. Armandroff, N. Caldwell & P. Seitzer, 2000, AJ, 119, 705
P. Côté, M. Mateo, E. Olszewski & K. Cook, 1999, ApJ, 526, 147
S. Holland, G. Fahlman & H. Richer, 1996, AJ, 112, 1035

R.M. Rich, K. Mighell, W. Freedman & J. Neill, 1996, AJ, 111, 768
D. Reitzel & P. Guhathakurta, 2002, AJ, 24, 234
R.M. Rich, C. Corsi, M. Bellazzini, L. Frederici, C. Cacciari & F. Fusi Pecci in 'Extragalactic Star Clusters', IAU Symposium 207, eds Grebel, Geisler & Minniti, 2001, (astro-ph/0106015)
A.M.N. Ferguson, 2001, in 'New Quests in Stellar Astrophysics', eds M. Chavez et al. (astro-ph/01082330)
R.A. Walterbos & R.C. Kennicutt, 1988, A&A, 198, 61
R.F.G. Wyse & G. Gilmore, 1988, AJ, 95, 1404
J. Mould & J. Kristian, 1986, ApJ, 305, 591
W. Baade, 1944, ApJ, 100, 137
W. Baade, 1958, in 'Stellar Populations', ed D.J.K. O'Connell (North Holland, Amsterdam) p 303
C. Pritchet & S. van den Bergh, 1987, ApJ , 316, 517
A.M.N. Ferguson & R. Johnson, 2001, ApJL in press (astro-ph/0108116)
A. Sarajedini & J. van Duyne, 2001, AJ in press (astro-ph/0107344)
C. Pritchet & S. van den Bergh, 1994, AJ , 107, 1730
R.A. Schommer, C. Christian, N. Caldwell, G. Bothun & J. Huchra, 1991, AJ, 101, 973
A. Sarajedini, D. Geisler, R. Schommer & P. Harding, 2000, AJ, 120, 2437

4.7 The Structure of the LMC's Disk[14]

Motivation

The Galaxy Genesis project described in this Purple Book seeks to unravel the formation history of the Milky Way disk by identifying the remnants of the disrupted early generations of star clusters that populate the thick disk. Carrying out a similar project in the disk of the Large Magellanic Cloud (LMC) will in many ways be easier, and would allow us to study disk formation in a galaxy with mass and age close to that of the primordial galaxy building blocks. Here,

Figure 2.4.10: The KAOS field of view projected against the inner regions of the Large Magellanic Cloud. Courtesy of Sam Barden.

[14]Contributed by Knut Olsen

we propose to survey approximately 10^5 LMC red giants with KAOS at high resolution (R~20,000), so as to identify disrupted ancient star clusters through their unique chemical signatures. By observing near the Na D lines at ~5890Å, our single-order echelle spectra will include lines of Ba, Ca, Fe, Ti, and Si in addition to Na; Na and Ba are particularly interesting, since they have been found to vary between globular clusters (Sneden et al. 2000). This survey will demonstrate the concept of chemical tagging (Freeman & Bland-Hawthorn 2002) at a much lower cost in telescope time than the Galactic survey. At the very least, the results of this survey will provide a massive database through which we will study mean stellar abundance trends and dispersions for a clear understanding of dwarf galaxy chemical evolution.

Chemical evolution in the LMC: the current state of the art

By contrasting the abundance patterns in the LMC with those in the Milky Way, we may highlight the physical processes that affect galactic chemical evolution. However, only with the availability of high-resolution spectrographs on 8-m class telescopes in the southern hemisphere has it become possible to study the chemical evolution of the LMC in detail. Smith et al. (2002) used Gemini South and the Phoenix high-resolution infrared spectrograph to measure oxygen and heavy element abundances in 12 LMC red giants; Figure 4.11 shows their results for O. First, Smith et al. found that in the LMC, the [O/Fe] ratio begins to drop at a lower [Fe/H] than in the Milky Way. This observation is consistent with the LMC having experienced a strong early burst of star formation (Gilmore & Wyse 1991), while the Milky Way formed stars at roughly constant rate. Second, the LMC reaches lower [O/Fe] ratios at high [Fe/H] than does the Milky Way. This observation could be explained by a higher ratio of SN Ia to SN II in the LMC than in the Milky Way.

The Smith et al. results demonstrate how the comparison of abundance patterns in the LMC and Milky Way strengthens our general understanding of chemical evolution. With a much larger LMC abundance dataset, we could explore the source of the evident intrinsic scatter in the abundance patterns. Although our proposed KAOS project will not include the measurement of O, Ti and Si behave in a similar fashion to O (Edvardsson et al. 1993). The Smith et al. results thus provide strong motivation for the KAOS proposal.

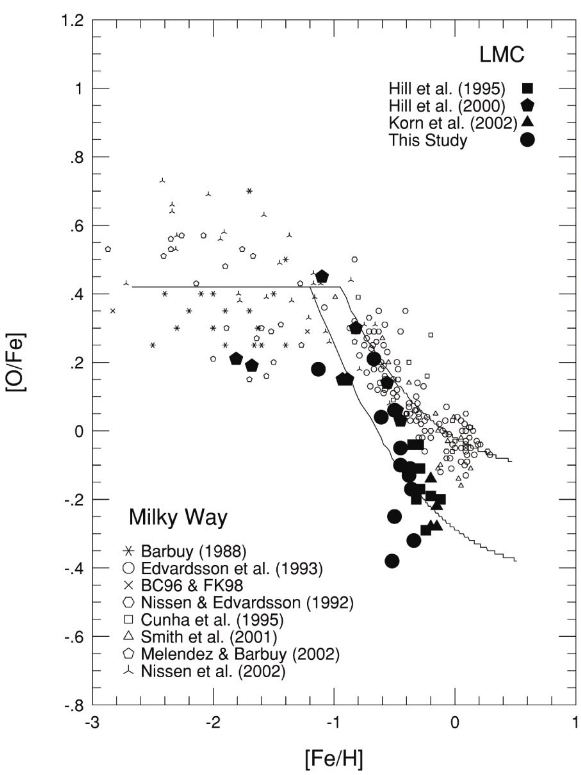


Figure 2.4.11: [O/Fe] vs. [Fe/H] for LMC and Milky Way stars, adapted from Smith et al. (2002), along with their simple chemical evolution models.

The globular clusters of the LMC: tracers of early star formation

The globular clusters (GCs) of the Milky Way halo are widely considered to represent the oldest known stellar population in the Galaxy, with the disk population being a few Gyr younger. The LMC has 13 known old GCs, which are remarkably similar to those of the Milky Way. Fig. 3 compares the *HST* color-magnitude diagram of the LMC cluster NGC 2019 (Olsen et al. 1998) with the MW GC M5. The extremely good match implies that NGC 2019 and M5 have the same age to within 1 Gyr and nearly identical abundances. Indeed, *all* of the LMC's old GCs studied to date have ages, abundances, and integrated luminosities which could have been drawn from the same parent population as the Milky Way GCs (Olsen et al. 1998, Johnson et al. 1999). The GC systems do, however, have one major difference: the MW GCs have halo kinematics, while the LMC GCs revolve with the HI disk (Schommer et al. 1992). Thus, the LMC contains the *oldest* disk that we know of in the Local Group; understanding the formation of this disk is potentially extremely exciting. There is no plausible way to form a disk GC system from the accretion of fragments (e.g., Searle & Zinn 1978) without also forming a stellar halo (Abadi et al. 2002). While many have looked (cf. Olszewski et al. 1996), there is not yet any evidence for a halo in the LMC. In the absence of a halo, a compelling explanation is that galaxies such as the LMC are indeed the 'building blocks' out of which the halos of larger galaxies formed (e.g., Côté et al. 1998). Kinematics of a large set of metal-poor LMC stars would establish whether the LMC indeed lacks a stellar halo.

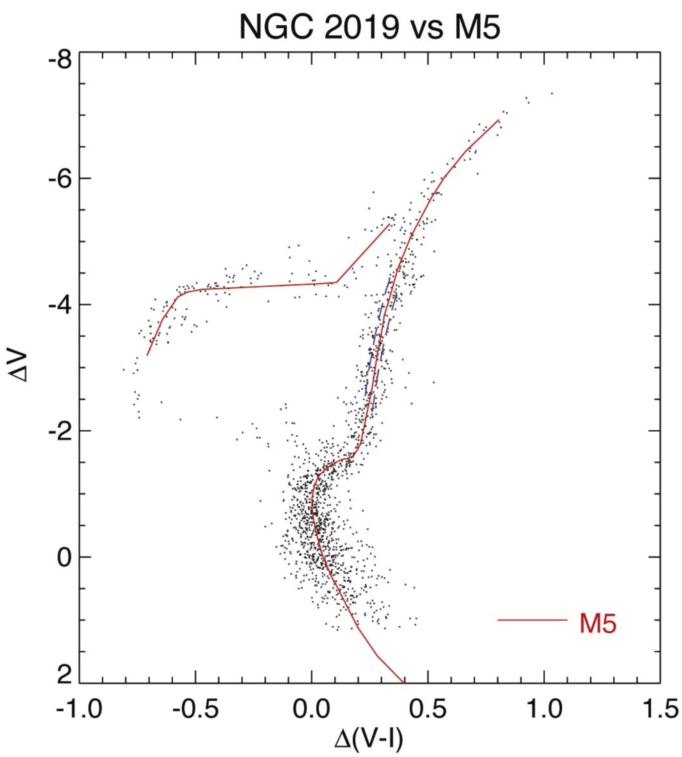


Figure 2.4.12: NGC 2019, an LMC globular cluster similar to the old Milky Way globular cluster M5. This VI color-magnitude diagram was produced from HST WFPC2 images of NGC 2019 (Olsen et al.1998); it has been shifted to line up with M5's fiducial sequence at the main sequence turnoff. The excellent matches of the red giant branch colors and slopes indicates that NGC 2019 and M5 have nearly identical abundances and ages. The dashed line to the right of the RGB indicates the color that NGC 2019's red giants would have if the cluster were 2 Gyr younger; the line to the left shows the position of red giants that are older by 2 Gyr.

Selecting an old LMC population

Unlike the Milky Way, the star formation history of the LMC has been punctuated by bursts. The LMC experienced an initial burst of star formation coinciding with the epoch of globular cluster formation. This was followed by an approximately eight-Gyr-long period of lower star formation rate, and a subsequent renewed star formation burst beginning approximately four Gyr ago and continuing to the present (e.g. Geha et al. 1998, Holtzman et al. 1999, Olsen 1999).

The star formation history is such that a random sample of LMC red giants with *V≤20* will contain ~30% stars formed at early times. Moreover, these stars are easily selected by their metallicity, since the initial burst produced rapid enrichment to [Fe/H] ~ -1 (Gilmore & Wyse 1991; see also Dopita et al. 1997), as seen in the cluster metallicity distribution shown in Figure 2.4.13 (Pagel & Tautvaisiene 1998).

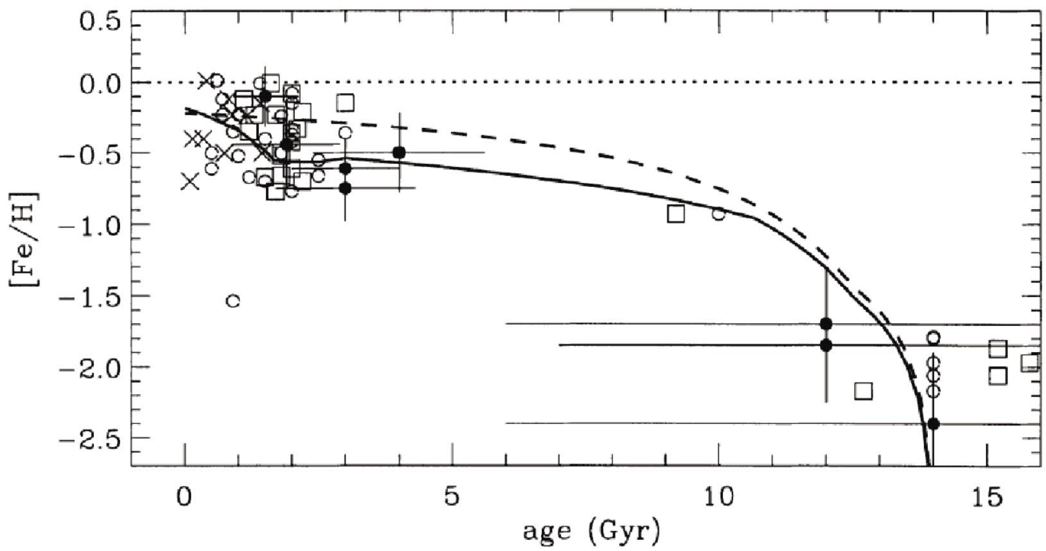


Figure 2.4.13: The LMC chemical evolution model and age-metallicity relation of LMC clusters, adapted from Pagel & Tautvaisiene (1998).

The Experiment

The main goal of this experiment is to identify through abundances and kinematics the population of disrupted low-mass star clusters that accompanied the formation of the LMC's existing globular clusters. How many clusters do we expect to identify, and how many objects will be available to KAOS? If the globular clusters are the surviving remnants of an LMC cluster population with a power-law mass spectrum similar to that of young LMC clusters (Elmegreen & Efremov 1997), then we expect the initial burst to have produced roughly 2000 star clusters with masses in excess of 10^3 $M_{\odot}$. In order to identify ~10 chemically unique stars per parent cluster, we will thus need to survey 60000 LMC stars, assuming 30% of the stars are ancient. Using the 9-million star color-magnitude diagram of the MACHO project (Alcock et al. 2000) as a guide, there are ~300000 LMC red giants with *V<18* distributed over an area in excess of 100 square degrees. To sample stars over the entire LMC disk, we will thus require ~45 KAOS pointings. Within the inner ~3 degrees of the LMC, the object density will exceed the number of fibers available, while at a 5 degree radius the density will drop to ~1000 objects per pointing. The high object density in the central regions led us to suspect that crowding might be problematic. However, as demonstrated by observations taken by the SuperMACHO project (Stubbs et al. 2003) with the CTIO 4-m of the LMC Bar, crowding is not a problem at *V=18*. Analytical considerations of crowding (Olsen et al. 2003) suggest that accurate photometry at *V=18* in the LMC is possible even in regions with surface brightness σ_V=20 mag arcsec^{-2}, while the LMC has central surface brightness σ_B=21.5 mag arcsec^{-2} (Bothun & Thompson 1988). The 45 pointings will allow KAOS to obtain spectra of ~140,000 red giants, or ~20 stars per disrupted cluster; the capabilities of KAOS are thus an excellent match to this project. Assuming 1 night of integration per pointing to achieve S/N=100 with R=20000 at 5800Å, this project will require ~45 nights of telescope time.

The LMC-SMC-Milky Way interacting system

The LMC has clearly interacted with its close companion SMC and the Milky Way, the most spectacular evidence coming from the Magellanic Stream (Mathewson et al. 1974) and its leading arm (Putman et al. 1998). Understanding this interaction is important for the analyses of the kinematics of both the LMC and Milky Way projects described in this Book, since the Clouds may be responsible for the Galactic warp (Weinberg 1998). While the SMC has borne the brunt of the damage from the interaction (e.g., Caldwell & Coulson 1986), in the LMC the effects are more subtle. van der Marel & Cioni (2001) found that the LMC is intrinsically elliptical with an asymmetric stellar density profile, which they ascribe to tidal interaction with the Milky Way. The kinematics of LMC carbon stars (Kunkel et al. 1997, Graff et al. 2000, van der Marel et al. 2002) most clearly show the rotation of the LMC disk, but also show intriguing evidence for kinematic disturbances. One of these disturbances corresponds in location to the region that Olsen & Salyk (2002) found from photometric analysis contains a warp. The analysis of LMC kinematics is particularly important, since the irregularities might be revealing structure that could be responsible for the microlensing rate seen in the direction of the LMC (Graff et al. 2000). The experiment proposed here will provide accurate kinematics for 10^5 LMC stars, a sample that is two orders of magnitude larger than all other samples combined. Thus, kinematic features that are invisible in today's data will be readily apparent in the KAOS dataset.

Abadi, M.G., Navarro, J.F., Steinmetz, M., Eke, V.R., astro-ph/0212282
Alcock, C., et al. 2000, AJ, 119, 2194
Bothun, G.D., Thompson, I.B. 1988, AJ, 96, 877
Caldwell, J.A.R., Coulson, I.M. 1986, MNRAS, 218, 223
Côté, P., Marzke, R.O., West, M.J. 1998, ApJ, 501, 554
Dopita, M.A., et al. 1997, ApJ, 474, 188
Edvardsson, B., Andersen, J., Gustafsson, B., Lambert, D.L., Nissen, P.E., & Tomkin, J. 1993, A&A, 275, 101
Elmegreen, B.G., Efremov, Y.N. 1997, ApJ, 480, 235
Freeman, K.C., Bland-Hawthorn, J. 2002, ARAA, 40, 487
Geha, M., et al. 1998, AJ, 115, 1045
Gilmore, G., & Wyse, R.F.G. 1991, ApJ, 367, L55
Holtzman, J., et al. 1999, AJ, 118, 2262
Johnson, J.A., Bolte, M., Stetson, P.B., Hesser, J.E., Somerville, R.S. 1999, ApJ, 527, 199
Kunkel, W.E., Demers, S., Irwin, M.J., Albert, L. 1997, ApJ, 488, L129
Mathewson, D.S., Cleary, M.N., Murray, J.D. 1974, ApJ, 190, 291
Olsen, K.A.G., Hodge, P.W., Mateo, M., Olszewski, E.W., Schommer, R.A., Suntzeff, N.B., Walker, A.R. 1998, MNRAS, 300, 665
Olsen, K.A.G. 1999, AJ, 117, 2244
Olsen, K.A.G., Salyk, C. 2002, AJ, 124, 2045
Olsen, K.A.G., Blum, R.D., Rigaut, F. 2003, AJ, submitted
Olszewski, E.W., Suntzeff, N.B., Mateo, M. 1996, ARAA, 34, 511
Pagel, B.E.J., Tautvaisiene, G. 1998, MNRAS, 299, 535
Putman, M.E., et al. 1998, Nature, 394, 752
Schommer, R.A., Suntzeff N.B., Olszewski E.W., Harris H.C., 1992, AJ, 103, 447
Searle, L., Zinn, R. 1978, ApJ, 225, 357
Smith, V.V., Hinkle, K.H., Cunha, K., Plez, B., Lambert, D.L., Pilachowski, C.A., Barbuy, B., Melendez, J., Bala*Chandra*n, S., Bessell, M.S., Geisler, D.P., Hesser, J.E., Winge, C. 2002, AJ, 124, 3241
Sneden, C., Pilachowski, C. A., Kraft, R. P. 2000, AJ, 120, 1351
Stubbs, C.W. et al. 2003, 2002, BAAS, 201, #7807
van der Marel, R.P., Alves, D.R., Hardy, E., Suntzeff, N.B. 2002, AJ, 124, 2639
van der Marel, R.P., Cioni, M.-R.L. 2001, AJ, 122, 1807
Weinberg, M.D. 1998, MNRAS, 299, 499

4.8 A Milky Way Halo Survey at the Main Sequence[15]

The current view (Freeman & Bland-Hawthorn 2002) is that the Galactic halo formed at least partly through the accretion of small metal-poor satellite galaxies that underwent some independent chemical evolution before being accreted by the Milky Way (Searle & Zinn 1978, Freeman 1987). Although we do still see such accretion events taking place now in the apparent tidal disruption of the Saggittarius dwarf (Ibata et al. 1995), most of them must have occurred long ago. We can expect to see dynamically unmixed residues or fossils of at least some of these accretion events (Helmi & White 1999). Of all the galactic components, the stellar halo offers the best opportunity for probing the details of its formation. There is a real possibility of identifying groups of halo stars that originate from common progenitor satellites (Eggen 1977, Harding et al. 2001, Majewski et al. 2000).

This genetic identification can sort a large sample of stars by kinematics and n-dimensional chemistry, where n is limited by resolution.

What do we expect to see kinematically ? Is it a normal distribution with $\sigma \approx 100$ km/sec ? Is it a distribution like the globular clusters. That is illustrated in Figure 2.4.14, based on the known kinematics and individual velocity dispersions of the cluster system.

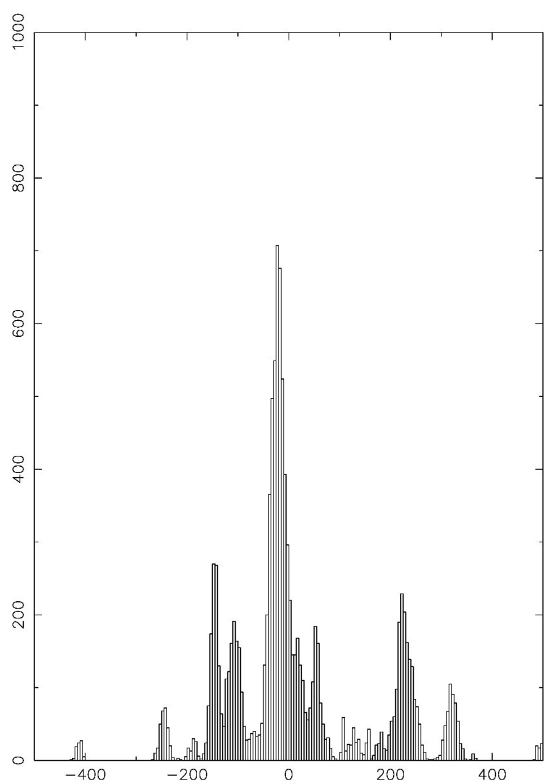


Figure 2.4.14: Distribution of velocities of stars in globular clusters. A possible small sample of halo kinematics. The abscissa is radial velocity in km/s.

Adding the kinematic distribution of stars in dwarf spheroidals would be a closer realization of what we expect to see when we sample the halo at 3 km/sec radial velocity resolution. It would be very valuable to pursue adequate simulations to tell us what we may see in a survey of halo genetics, but in a hand-waving way we could speculate that in a sample of 100,000 objects we might see ~200 kinematically and chemically distinct families of ~500 members each.

[15]Contributed by Jeremy Mould

[16]http://cfa-www.harvard.edu/cfa/oir/MMT/MMTI/hectochelle.html

A sample with $19 < V < 21$ mag would span the Milky Way and require ~225 KAOS fields to be observed for 4 hours each at $R = 5000$ at the calcium triplet. The resolution should be sufficient that the detailed chemical signature should also show its unique character. The density of halo stars at this magnitude is 250 per square degree. Of existing facilities, it is Hectochelle[16] on the MMT which is best matched to the parameters of this survey. Nevertheless, this project would still take nearly 700 nights with Hectospec! KAOS on Gemini would be able to execute this survey about 3.4 times faster than Hectochelle.

4.9 Population III and the Dwarf Spheroidals[17]

There is renewed theoretical interest in a flat IMF proto-population. Suppose, for a moment, that it is as sparse as one ω Cen in the whole Milky Way. The surface density would be less than 5 dwarfs per square deg. And for power law index s = 1 (as opposed to 2.35) that would be less than 1 dwarf per square deg. Such stars might begin to appear in the ultra-low metallicity tail of a halo genetics survey such as that described above. Their kinematics and detailed chemistry should be extraordinarily interesting (Figure 4.15).

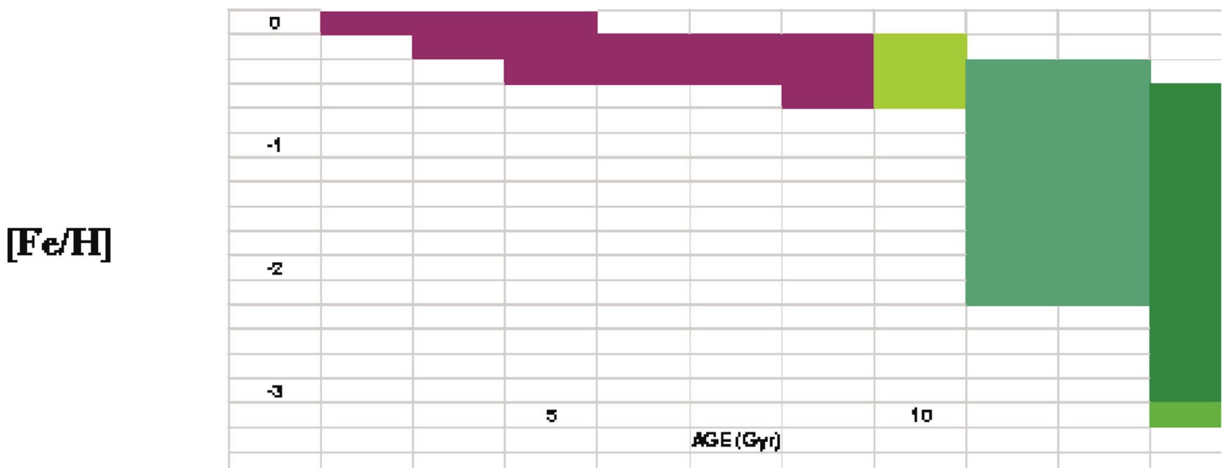


Figure 2.4.15: The age metallicity relation of the Milky Way by Freeman and Bland-Hawthorn (2002). The ultra-low metallicity tail of the metallicity distribution in a halo survey samples star formation in the first billion or so years of the Galaxy.

In the spirit, however, of searching for the Holy Grail in traditional sites, rather than randomly, we should consider a complete survey of the Milky Way's dwarf spheroidal galaxies as an advantageous strategy. The KAOS spectrograph is well matched to the size of the Draco or Ursa Minor dwarf galaxy. A complete sample of stars to the horizontal branch in the "seven dwarfs" would range from 1000 to 100,000 stars from Draco to Fornax. In addition to exploring the low metallicity tail of chemical evolution, one would gain a sample which would strongly constrain dynamical models of dwarf galaxies. Current models with *M/L*~100 for Draco regard dwarf spheroidals as dark matter laboratories with a sprinkling of baryons. Simulations are needed of structural models with different core radii for the dark matter distribution, to test the power of kilo-aperture spectroscopy to discriminate between such models. Models which incorporate environmental effects on these galaxies, such as tidal disruption, are also required. Existing facilities will be able to tackle the small dwarfs in the sample, such as Draco, but not the larger ones in reasonable observing times.

[17]Contributed by Jeremy Mould

References

Eggen, O. 1977, ApJ, 215, 812
Freeman, K. 1987, ARAA, 25, 603
Freeman, K. and Bland-Hawthorn, J. 2002, ARAA, 40, 487
Harding, P., Morrison, H., Olszewski, E., Arabadjis, J. & Mateo, M. 2001, AJ 122, 1397
Helmi, A. & White, S. 1999, MNRAS, 307, 495
Ibata, R., Gilmore, G., Irwin, M. 1995, MNRAS, 277, 781
Majewski, S. et al 2000, AJ, 120, 2550
Searle, L. & Zinn, R. 1978, ApJ, 225, 357

4.10 KAOS and the Interstellar Medium[18]

KAOS offers the unique opportunity of probing the interstellar medium of our galaxy over both small and large scales by studying the lines-of-sight to Galactic stars using high signal-to-noise ratio, high-resolution spectroscopy. Over a spectral window of 3800 Å to 1 micron, the following key interstellar absorption lines can be targeted:

Atomic Lines:
- Ca II H and K at 3934 Å and 3968 Å
- Na I D doublet at 5890 Å and 5896 Å
- K I at 7698 Å

Molecular Lines:
- CN 3874 Å,
- CH+ 4232 Å, and CH 4300 Å
- C2 and CN bands longward of 7000 Å

Diffuse Interstellar Bands:
- Strongest are at 4430 Å, 5780 Å, 6284 Å
- Over 100 others mostly > 5000 Å

For absorption line studies of the ISM, a minimum spectral resolution of 20,000 is required. This would be comparable with previous Copernicus and current FUSE UV ISM absorption line observations. Even higher resolution would be desirable since interstellar absorption line profiles have structure down to velocity scales of 1 km/sec and below. At a resolution of 20,000 at the Na I D lines, a spectrum with a S/N ratio of 100 would result in 5σ detections of 15 mÅ interstellar features. The WIYN Hydra MOS can reach this sensitivity on a V=11 star in a 100 minute net exposure in good seeing. In order to fully utilize the 4000 fibers of a KAOS, it would be hoped that this sensitivity could be improved by at least 2 magnitudes.

4. Likely targets for multi-object ISM absorption line studies would be open star clusters rich in early-type stars. Clusters such as NGC 663, M36, M35, and NGC 2244 typically have 80 to 100 stars with $9 < V < 11$ in a 1° field. The brightest rich cases, h and Chi Per, have about 100 stars with $8 < V < 10$ in a 1° field. Such clusters could easily yield hundreds if not a thousand targets up to V=13 mag over the 1.5° KAOS field.

Most optical Galactic ISM absorption line studies to date have focused on bright ($V < 8$) background stars. There are two basic reasons for extending this work toward much fainter stars: (1) a much wider expanse of the Galaxy can be explored (the halo, other spiral arms, etc.), and (2) the physical conditions of dustier (translucent and molecular) clouds can be probed. The dense multi-object capability of KAOS makes it possible to transform such expanded studies from core samples of single sightlines to wide-field structural probes of distant and/or dense interstellar clouds. Although this

[18]Contributed by Dave Meyer

optical absorption-line approach does not provide the complete spatial coverage of an H I 21 cm emission survey, its use of many extremely narrow (< 0.00001 arc sec) beams will be sensitive to structures that would be washed out in the much broader radio beam. Some specific potential ISM absorption projects with KAOS could include:

A Search for Mini-Clouds

Many have theorized that the Galaxy could be filled with a substantial population of very small (AUs) clouds. Some have suggested that these theoretical clouds could be dense enough to carry a significant fraction of the Galactic baryonic dark matter. Due to their very small size, such clouds would be extremely difficult to detect through H I 21 cm emission observations. Also, the statistics from repeat optical and UV absorption-line observations are currently far too low to place meaningful limits. However, in one three-night run with KAOS, it should be possible to obtain high quality interstellar Na I spectra toward a total of at least 15,000 stars from a selection of open clusters. This number of sightlines with high-quality Na I data would surpass by at least a factor of 5 the current holdings of such data. By repeating these observations one to two years later, it would be possible to look for any transient Na I variations (as compared temporally and with the other members of the cluster) that might be indicative of the passage through the line of sight of an intervening mini-cloud. Depending on the number and distance of the observed stars, even null results could place serious constraints on the mini-cloud hypothesis.

A Deep Molecular Absorption Survey

Almost all interstellar molecules (over 100 species at last count) have been identified via their radio and microwave transitions in dark molecular clouds. KAOS will have the capability to probe 100s of heavily extincted ($E(B-V) > 1.5$) stars at a time in clusters/associations embedded/behind dense clouds. Only a handful of such stars have been observed to date. The KAOS observations will afford by far the most sensitive search yet (either individually or by creating a net spectrum for the dense gas absorption) for new interstellar molecular species and/or transitions at optical wavelengths. In addition, the CN, CH, CH+, and C2 lines will provide an unprecedented view of the small-scale structure in the molecular gas of these clouds to compare with that of the atomic Na I, Ca II, and K I gas.

3-D Interstellar Absorption Imaging

One advantage in studying interstellar gas through optical (or UV) absorption rather than radio (or microwave) emission is that one knows that the optical absorption must be due to gas in front of the targeted background star. By blasting a targeted field with 4000 Na I and Ca II observations toward 4000 stars of various measurable types and thus, inferred distances, it should be possible to construct reasonably detailed 3D maps of the gas distribution in these fields. Potential applications include better determinations of the distances to high-velocity clouds in the Galactic Halo as well as the distances to hard-to-type objects which nevertheless have measurable interstellar absorption.

4.11 Studies of High Velocity Clouds with KAOS[19]

High velocity clouds (HVCs) consist of interstellar gas moving at velocities incompatible with simple models of differential Galactic rotation. Even though they were discovered almost 40 years ago, relatively little is still known about their physical

[19]Contributed by Francis Keenan

properties. However the determination of HVC properties is important, to assess their input to the mass and energy of the Galaxy, and help in our understanding of processes that distribute and ionize gas in the halo of our own and other galaxies. The (mostly) unknown distances to the HVCs remains the key problem in determining their nature and origin. For example, their distances D affect estimates of their linear diameters (as D^{+1}), masses (D^{+2}) and densities(D^{-1}). If the HVCs are relatively nearby they may be condensing gas from a Galactic fountain, or if very distant could be extragalactic gas being accreted by the galaxy.

Most optical absorption line studies of HVCs have employed hot stars (OB-type) as background probes, and have obtained spectra at high resolution (R>60,000). However our recent work has demonstrated that optical HVC studies can be performed using later-type objects and lower spectral resolutions, of around R≈30,000 (e.g., see Figure 2.4.16). We therefore believe that KAOS will be able to make a significant contribution to HVC research programs, particularly in 2 areas.

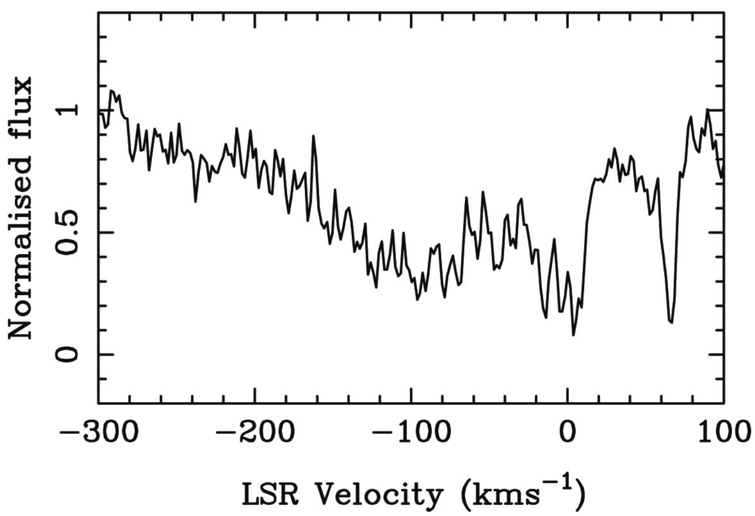
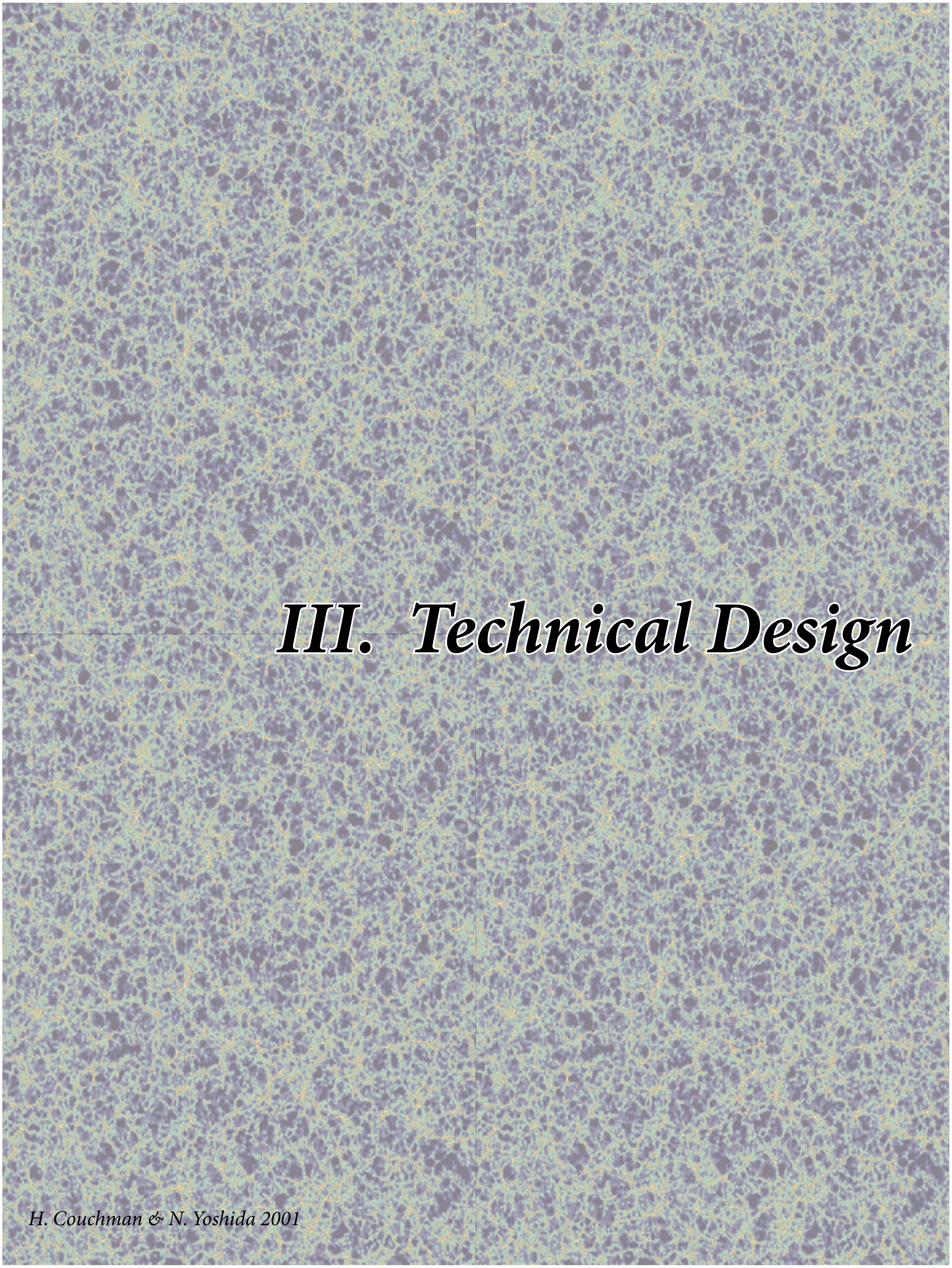


Figure 2.4.16: The CaII K 3933Å line profile towards a late F-type star, obtained with the WHT at a spectral resolution of R = 30,000. Strong HVC absorption is visible at V_{LSR} ≈ 70 km/s in the wing of the broad CaII K stellar feature, while local interstellar gas is also observed around V_{LSR} ≈ 0.

Firstly, KAOS may be employed to obtain moderate resolution (R = 30,000) spectra for a very large sample of stars of varying spectral type and magnitude (and hence distance), towards a small region of sky with known strong HVC emission in HI 21 cm. Detections or null detections of HVC absorption towards the stars would provide estimates of the HVC distance. A late F-type star (such as that shown in Figure 4.16) with V = 21 lies at a distance of around 25 kpc, and hence even completely null detections of HVC absorption in a KAOS sample would provide very important information – namely, that the HVC must be extragalactic in origin.

Another important parameter for HVC studies is their degree of 'clumpiness', i.e., the size distribution of the individual 'cloudlets' that form an HVC. This quantity plays a vital constraint to various HVC formation mechanisms, and also to numerical simulations to assess if star formation via HVC cloudlet collisions is a viable scenario. For such modeling, the HVC cloudlet dimensions are required, as opposed to simple angular size, and hence it is important to study HVCs of known distance, such as Complex A at D = 9 kpc. KAOS observations of a large sample of stars behind such an HVC would allow its spatial structure to be mapped on sub-arcmin angular scales (compared to the >arcmin spatial resolution available from aperture-synthesis HI 21 cm studies), corresponding to ~pc scales.

If an 'integral-field' mode is available, KAOS will be able to observe extended sources behind HVCs, such as active galaxies. For example, NGC 3783 (surface brightness of B ≈ 20.0 arcsec^{-2}) lies behind HVC 187, and integral field observations would provide details of the HVC structure on arcsec angular scales, corresponding to sub-pc dimensions.



III. Technical Design

H. Couchman & N. Yoshida 2001

Introduction

This segment of the KAOS Purple Book describes the instrument concept and its feasibility on Gemini (Chapter 1), the echidna fiber positioner (Chapter 2), and the efficacy of the nod-and-shuffle observing technique for sky subtraction on fiber spectrographs. The expertise for realizing this instrument concept and implementing the observing techniques lies entirely within the Gemini partnership, which has pioneered much of the technological innovations described (i.e., large optics, active control, wobble plates, fiber positioner technology, and nod-and-shuffle observing mode).

Chapter 1

KAOS Instrument Concept[19]

The science objectives described in Part II require a wide field of view with a relatively high density of spectroscopic apertures. This combination is well achieved with a fiber-fed instrument allowing a configurable focal surface. The wide range of spectral resolving powers can be accommodated by a spectrograph implemented on an optical bench where it can be easily reconfigured.

We present here a concept that achieves the primary objectives of the science goals by implementing a wide-field, prime-focus, fiber-fed; multi-object spectrograph for the Gemini telescopes. The following summarizes the general science requirements as instrument specifications for the concept studied.

- 1 to 2 degree field of view
- High density of targets per exposure (~1 per square arc-minute)
- High efficiency
- Spectral window of 3900 Å to 1.5μm
    - Optical channel of 3900 Å to 11000 Å
    - Non-thermal IR channel of 1.1 to 1.5 μm (future upgrade)
- ~1 arc-second apertures
- Spectral resolving power range of 1,000 to ~30,000
- Ability to do Nod & Shuffle observing

This chapter provides a discussion on the instrument concept. The fiber positioner, Nod & Shuffle observing, and the data pipeline requirements that will link the vast amount of KAOS data with the virtual observatory, are described in subsequent chapters.

KAOS

KAOS (Kilo-Aperture Optical Spectrograph) was conceived to address the science requirements described in Part II. It is a prime-focus, fiber-fed, multi-object spectrograph designed for use on the Gemini 8-meter telescopes. The specific implementation for Gemini grew from a concept study originally made for a 20 arc-minute FOV, wide-field spectrograph for the GSMT 30-meter telescope concept (see *http://www.aura-nio.noao.edu/book/index.html*).

The specifics for a Gemini instrument were derived at the workshop on the Next Generation Wide-Field Multi-Object Spectroscopy workshop held in Tucson in October of 2001 (*http://www.noao.edu/meetings/wfmos*). See ASP Conf Ser vol 280 for information on that workshop and preliminary papers describing the KAOS instrument and Echidna fiber positioner.

---

[19] *Contributed by Samuel Barden, Rick Robles, Ming Liang, James Robinson, Charles Harmer and Arjun Dey*

The Gemini telescope was considered to be an ideal platform for two reasons:

1. It serves a multi-national community that has a large interest in wide-field spectroscopic science.
2. The original Gemini concept included a wide-field mode that has not yet been implemented.

In order to refine the concept for Gemini, it is important to understand the various constraints of the existing design.

Gemini Design Constraints

The Gemini telescope design and performance give the following constraints for the implementation of KAOS:

- 8.1 meter, f/1.7 primary mirror;
- Tight space constraint at prime focus with only ~1.5 meters available space behind the uncorrected prime focus image;
- The need to counter wind shake with a fast tip-tilt system working at ~1 arc-second amplitudes at a rate of 2 to 4 Hz;
- Availability of a pier lab for locating the spectrographs, requiring ~60 meter long fibers.

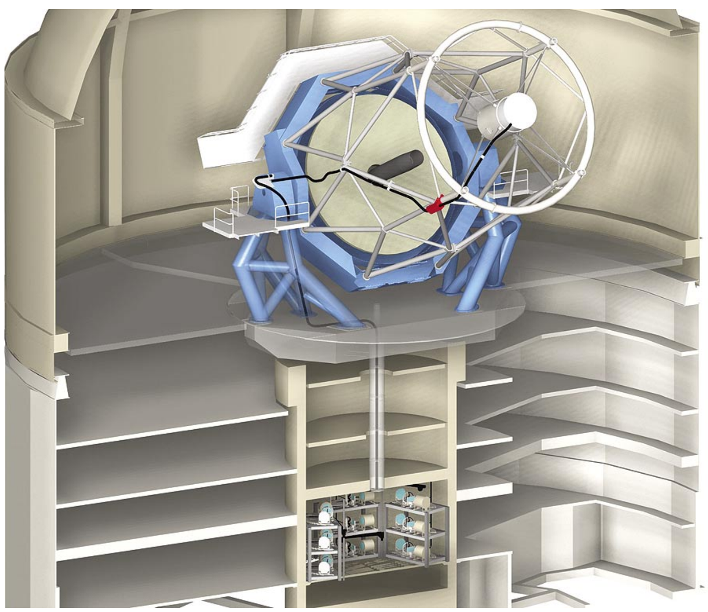


Figure 3.1.1: Display of a solid model rendering of the KAOS concept on Gemini. The fiber positioner is located at the prime focus of the telescope. The fibers run down the truss, into the telescope mount, and drop down to the lower floor where the spectrographs reside.

Prime Focus Corrector Concept

Producing a 1 to 2 degree field of view on a classically built 8-meter class telescope requires a large corrector. An initial look at a 2 degree diameter field of view showed that good images could be achieved across the full field, but that the front corrector element would be over 1.5 meters in diameter with a mass for the glass in the full corrector approaching 1500 kg! This approach was considered to be too risky and costly, and it was decided to scale back the size of the focal surface to a 1.5 degree diameter field of view.

A variety of concepts were explored including a design that had an atmospheric dispersion compensator (ADC) as part of the powered optics similar to the 2dF wide field corrector. The design selected for further conceptualization, however, has plane-parallel dispersion prisms. This approach was preferred due to its simplification and minimization of the risk in the fabrication of the large corrector elements. It is shown in the Figure 3.1.2. It is a four-element corrector made from Fused Silica and BK7. The ADC prisms are composed of LLF6 and BK7 prism pairs.

Wobble Plate

The Gemini telescopes are prone to vibration by the wind causing the images to move by about an arc-second in amplitude at a frequency of a couple of Hertz. The f/16 secondary has a tip-tilt capability that compensates for the wind buffeting. For a prime focus, a transmissive plate will need to be wobbled in order to stabilize the image against the wind. This is similar to techniques used to stabilize images in binoculars. The CFHT telescope has recently implemented a wobble plate in the MegaPrime unit (see *http://www.cfht.hawaii.edu/Instruments/Imaging /Megacam/overview.html*).

The nominal concept described here uses one or both of the ADC prism sets as the wobble plate. This approach was selected in order to minimize the number of elements in the corrector and, more importantly, the total mass of the prime focus assembly. Image quality across the field is preserved for image displacements of several arc-seconds. Figure 3.1.3 shows how the images degrade as the wobble plate (one set of ADC prisms) is tilted at increasing angle. It may be possible to utilize one of the ADC prisms for the rapid windshake compensation and the other for Nod&Shuffle observing where the image is moved on and off the fiber on a timescale of about every 30 seconds. Beam switching between fibers may be possible with image displacements of up to 10 or possibly 20 arc-seconds.

Figure 3.1.2 indicates the ADC prisms as the wobble plates. A solid model rendering of the wobble plate mechanism is shown here in Figure 3.1.4. Columns of piezo motors would provide the tip-tilt motion.

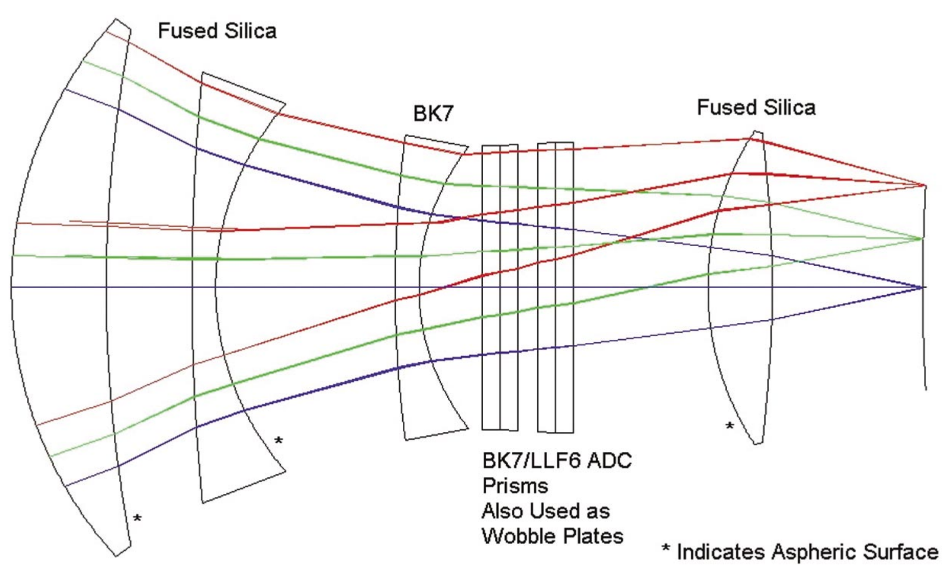


Figure 3.1.2: Zemax layout for the chosen, nominal corrector concept. Output is f/2.4 with 19187 mm focal length. Total length of system is 2305 mm. Image surface is 1.5 deg (515 mm diameter) with curved radius of 4382 mm.

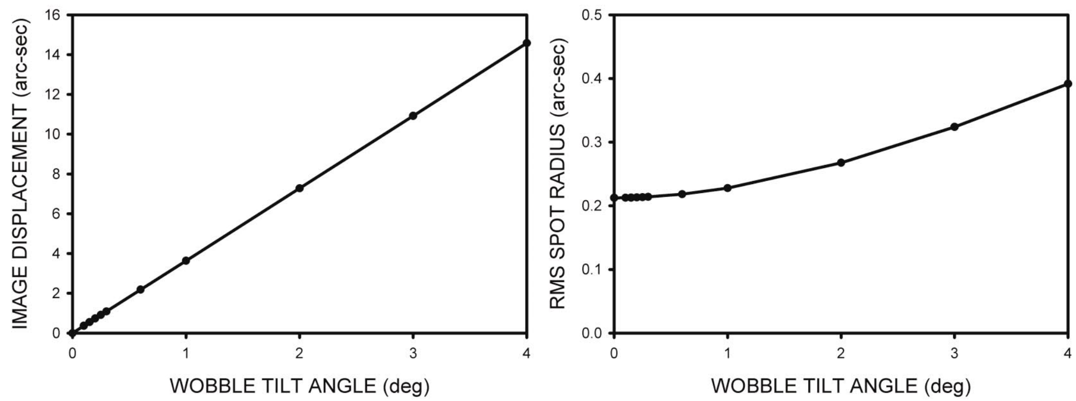


Figure 3.1.3: Image displacement and quality as a function of tilt angle for the wobble plate. Only one of the two ADC prisms was tilted for this evaluation.

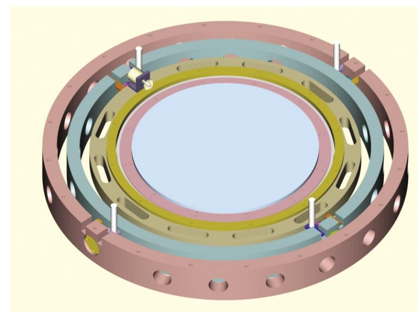


Figure 3.1.4: Solid model view of ADC/Wobble Plate rotation and tip-tilt assembly. View for one of the two assemblies.

Image Performance

The following plots give spot diagrams for the images produced by the nominal wide field corrector concept. The focal length of the system is 19187 mm (f/2.4) giving a plate scale of 93 microns per arc-second. The rms radii of spots across the full field when the telescope is pointed at the Zenith are all under 40 microns diameter (0.43 arc-seconds on the sky) integrated across the 3800 Å to 1 µm spectral window. At a Zenith angle of 70 degrees and a wobble displacement of about 1 arc-second, the images are still good with rms diameters of less than 60 µm (0.65 arc-seconds) across the full spectral window. The field has a diameter of 1.5 degrees and is curved with a 4.3 meter radius. The exit pupil is concentric with the curved field.

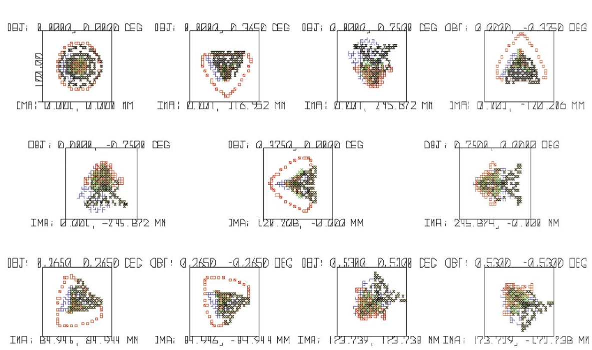


Figure 3.1.5: Spot diagram for Zenith, no wobble. Wavelength range is 0.38-1.0 micron. Box is 100 microns (1.07") wude,

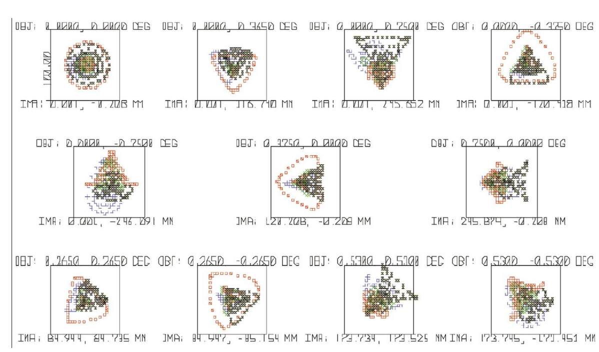


Figure 3.1.6: Spot diagram for Zenith, 0.3 degree wobble tilt.

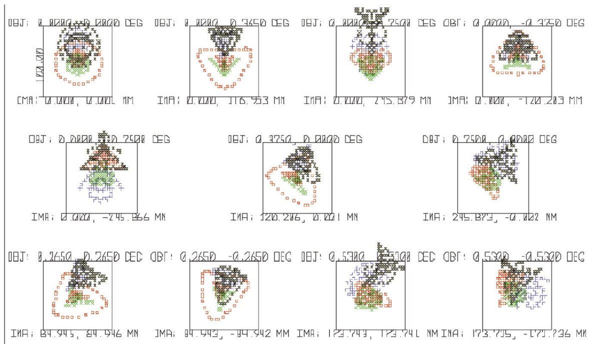


Figure 3.1.7: Spot diagram for 70 degree Zenith, no wobble. ADC prisms are compensating atmospheric dispersion. Box is 100 microns (1.07") wide.

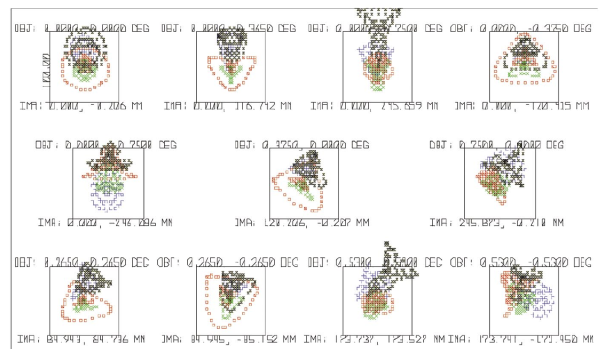


Figure 3.1.8: Spot diagram for 70 degree Zenith, 0.3 degree wobble tilt. ADC prisms are compensating atmospheric dispersions.

Design Issues

It was necessary to introduce aspheric surfaces onto three of the elements in order to get good image quality. These surfaces were forced to be conic in order to simplify their fabrication and testing. The concave surfaces of the front two elements contain 2 of the aspheric surfaces. The third surface resides on the leading convex surface of the final element. The asterisks in Figure 3.1.2 indicate these aspheric surfaces.

The front element has a total clear aperture diameter of 1.29 meters. Although the size of this element will introduce risk, it is within the current state-of-the-art being considered for other astronomical facilities (e.g., LSST telescope with a 1.4 meter front element, also containing aspheric surfaces. For further details on the corrector concept for LSST see *http://www.noao.edu/lsst/ScottEllisfinalreport.pdf*). The major challenge will be obtaining a piece of fused silica with adequate homogeneity in the index of refraction.

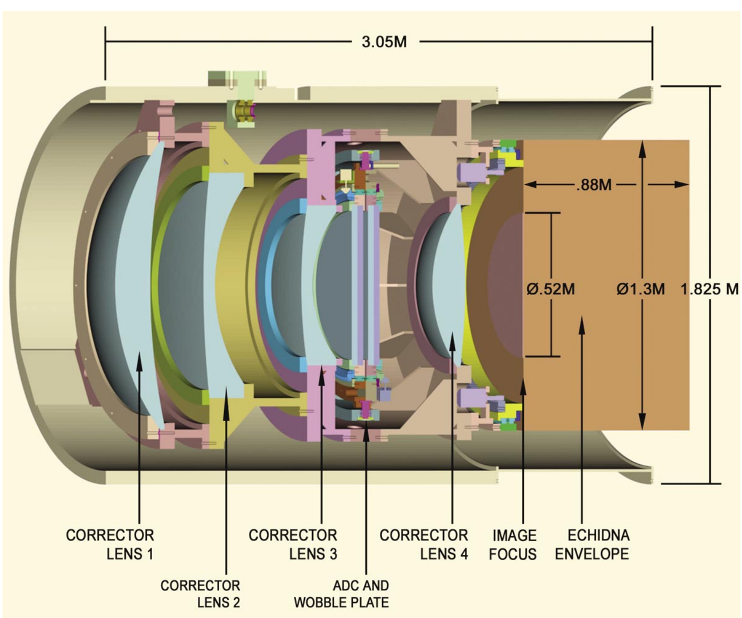


Figure 3.1.9: Solid model of KAOS prime focus wide field corrector assembly. The various components are indicated.

Preliminary optical tolerances were examined. If one assumes that the optics are mounted and pre-aligned in sub-assemblies, then individual lens decenters of 100 μm are acceptable for all components with the front element giving the worse case image degradation of 3%. Tilt and decenter tolerances are all reasonable.

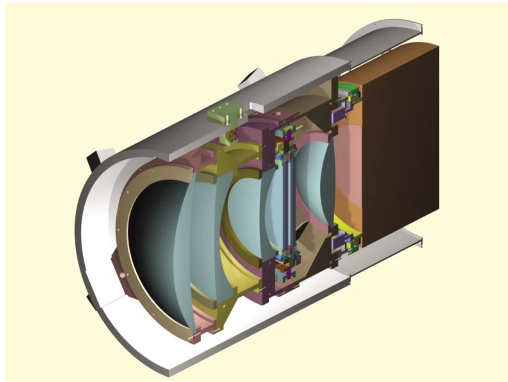


Figure 3.1.10: Solid model of KAOS prime focus wide field corrector. See Figure 3.1.9 for labeling components.

Preliminary FEA on the support of the lenses appears to give reasonable results. It may be desirable to draw a slight vacuum between the first and second elements in order to help support the mass of the front element. The pressure differential required is quite low. Alternatively, an axial support could be placed at the center of the front element and supported with a spider vane assembly at the expense of some slight additional vignetting of light in the outer regions of the field.

Figures 3.1.9 and 3.1.10 show solid model renderings of a notional support structure and housing for the wide field corrector and related components. The space envelope fits within that available at the top end of the Gemini telescopes. The volume of space allocated for the fiber positioner is a cylinder 1.3 meters in diameter and 0.88 meters in length. The overall diameter of the housing is about 1.9 meters, resulting in a central shadow of 6%.

Fiber Positioner

The nominal instrumental concept described here has a 4000 to 5000 fiber positioner located at the prime focus image. A brief description of the positioner is provided here. Echidna is described in more detail in chapter 2. Information can also be obtained at *http://www.aao.gov.au/local/www/echidna.*

The space envelope of the Gemini prime focus environment is inadequate to allow the implementation of large slit-fed spectrographs. The desire for high dispersion also eliminates the ability to utilize small, slit-fed instruments as well. Fiber optics are therefore required to transfer the target light down to remotely located spectrographs. The Echidna approach ideally satisfies the need for wide field coverage with a relatively high density of configurable apertures within a very constrained space envelope. Of course, a Cassegrain implementation of the wide field would somewhat mitigate the space envelope problems, but as will be discussed later, there are significant advantages of a prime focus versus Cassegrain instrument in terms of performance, cost, and support.

The nominal fiber spacing for Echidna is 7 mm with hexagonal packing for the fibers. For the nominal wide field corrector, described previously, a total of 4811 fibers can be located within the 1.5 degree diameter field of view. It may be possible to shrink the interfiber spacing to 6 and possibly 5 mm, allowing a significantly larger number of fibers to be located within the focal area. However, given the current scientific requirements along with the tradeoff in complexity, the 7 mm spacing appears to be quite adequate and will be assumed for further discussion of the concept.

The Echidna fibers are moved by tip-tilt of the fiber. Although this causes the input beam to be at a variable angle with respect to the fiber axis, the prime focus focal ratio of f/2.4 at most is degraded into an f/2.03 beam for fiber spines that are 200 mm in length with a 7 mm displacement (2.0 degree tilt). Longer fiber spines would minimize this effect further. The spectrograph collimator would be designed to accept this additional cone of light.

Figure 3.1.11: shows how the fibers are moved to acquire targets. Each fiber has can move within a circle of 1.25 arc-minute radius.

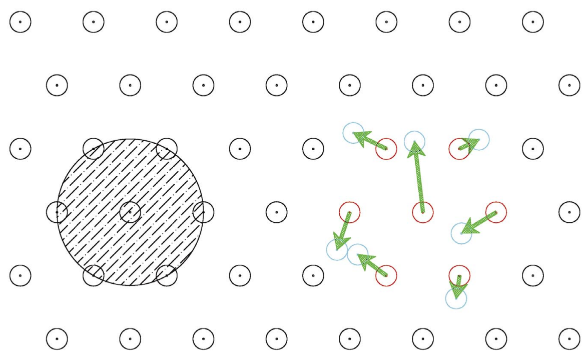


Figure 3.1.11: Schematic showing range of fiber motions and how they are used to acquire targets. Each fiber can reach its neighbor's home position giving a 2.5 arc-minute diameter range for each fiber.

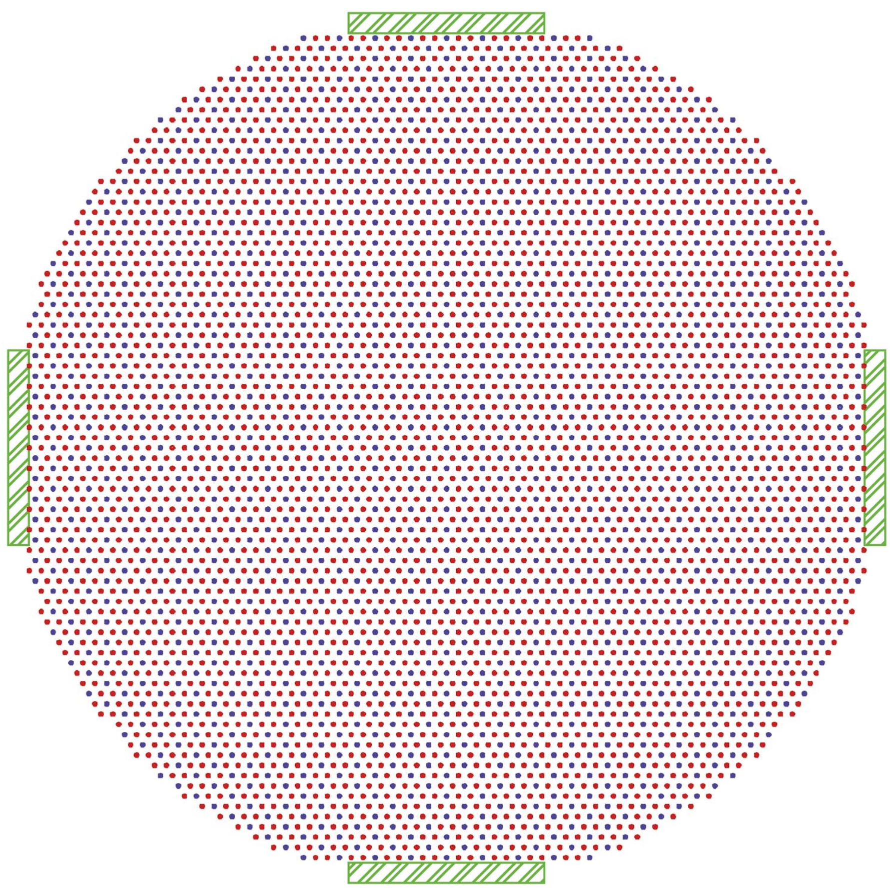
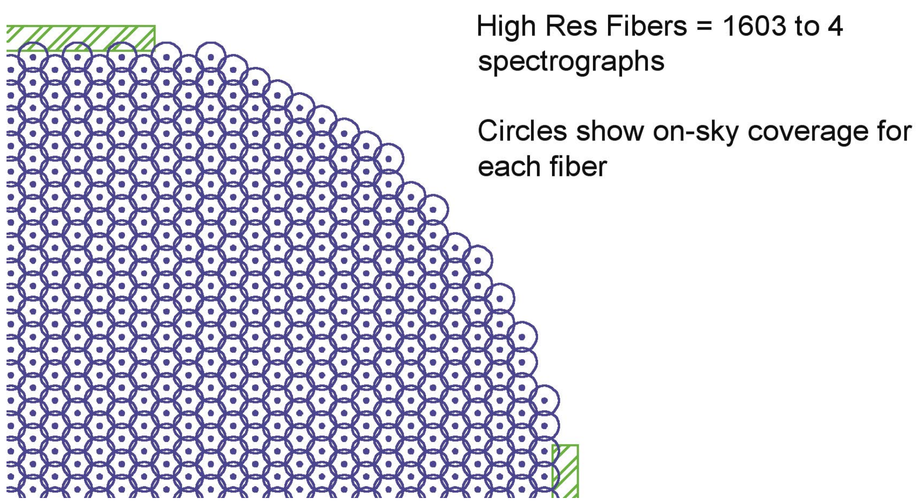


Figure 3.1.12: KAOS fiber arragement. 4811 total fibers. 1603 (blue) feed into high-resolution spectrographs, 3208 (red) feed low-resolution spectrographs. Four guide/wavefront sensor probe regions (1.5 by 15 arc-minute in area) are shown in green.

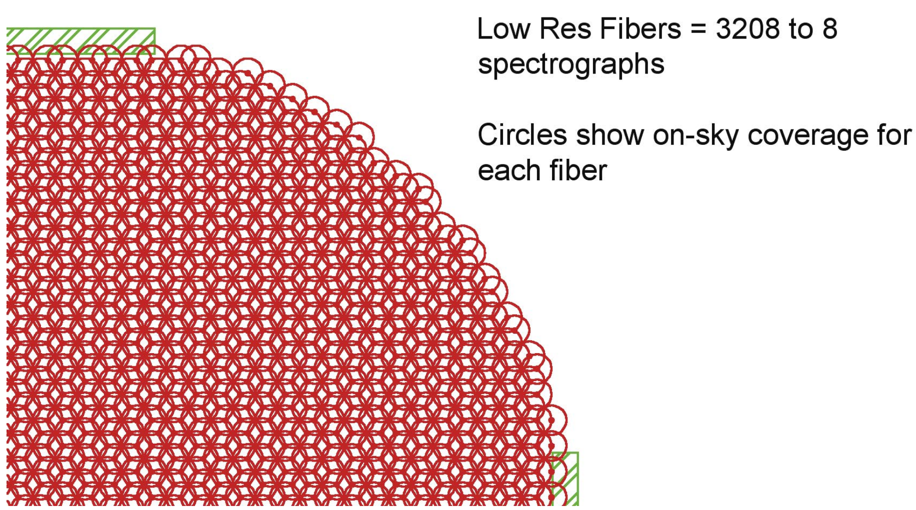
In order to maximize the multiplex capability of the instrument by allowing the possibility of observing targets at high resolution while simultaneously observing targets at low dispersion, 1/3rd (1603) of the fibers are allocated to a bank of high resolution spectrographs while the remaining 2/3rd (3208) fibers feed into a set of low dispersion instruments. This split gives a relatively good match of fiber density to target density on the sky for the galactic and extragalactic applications where each region of the field is accessible by at least 1 high-resolution fiber and by at least 2 low-dispersion fibers. The layout is shown in Figure 3.1.12, showing the distribution of fibers for each fiber type. Regions at the four quadrants of the field are reserved for acquisition and wavefront sensors.

Configuration times for the full set of fibers are expected to be under 5 minutes, assuming that the imaging system required for fiber configuration can simultaneously image the full set of fibers. Installing four imagers around the periphery of the prime focus assembly may do this. Each imager will view a quadrant of the fibers by looking at the primary mirror and through the corrector lenses.

Figures 3.1.13 and 3.1.14 display the high- and low-resolution fibers and their range of coverage in more detail.



Figures 3.1.13: View of quadrant of high-resolution fibers and their range of coverage.



Figures 3.1.14: View of quadrant of low-resolution fibers and their range of coverage.

Acquisition, Wavefront Sensing, and Calibration Unit

As mentioned above, four regions at the edge of the field are allocated for acquisition and wavefront sensing. The 1.5 by 15 arc-minute regions are more than adequate for viewing 15th magnitude and brighter stars within each region by translatable probes for this purpose. The four regions are desired in order to do tip-tilt measurements within each quadrant so that atmospheric tip-tilt can be distinguished from telescope induced tip-tilt errors. The four regions are indicated in Figure 3.1.12, shown previously. Two additional types of probes are desired for deployment in each region in addition to the tip-tilt sensors: these are an acquisition viewer and a wavefront sensor. The acquisition viewer would be a simple imager for initial target acquisition. Refined field alignment could then carried out by a distributed set of fiber probes within the science field of view if not by these exterior probes. The wavefront sensors are required for control of the active support for the primary mirror. It is assumed that these probes can be utilized at the edge of the field. If not, then it will be required that such a probe be capable of insertion into the center of the field in between observations.

Flat field and wavelength calibration can be done with the use of a translatable screen and bank of illumination lamps within the corrector assembly. Although the ideal location for such a screen is at the exit pupil of the telescope, implementation at that location is difficult, if not impossible. A good compromise location is just behind the set of ADC prisms where there is adequate space for insertion of a screen. The bank of lamps (quartz and hollow cathode tubes) can be located around the circumference of the final corrector element. Proper analysis of the illumination of the lamps will be required to ensure that the light seen by the fibers is uniform across the full field of view, particularly for the flat field (quartz) lamps. It may also be possible to utilize the night sky emission lines to help scale the relative intensity of flat field exposures for each fiber. An alternative location for a screen could be on the primary mirror cover.

Instrument Rotator

The alt-az mounting of the Gemini telescopes forces the implementation of an image rotator. The concept envisions a mechanical rotation of the fiber positioner and related hardware located in the instrument envelope of the prime focus assembly. The corrector elements, ADC, and calibration system do not rotate. The fiber cable must have a service loop that allows full motion of the rotator.

Other Instruments at Prime Focus

The space envelope for the fiber positioner can also be utilized for other instrument concepts that are not specifically considered in this document. Among these are: a wide field optical imager; other fiber feeds such as a monolithic fiber array IFU or a set of smaller, deployable fiber IFU's; and possibly a small, narrow-field, slit-fed, very low-resolution spectrograph(s).

Top End Assembly

The original Gemini concept included the ability to change top ends, particularly with a wide-field secondary. Although only one top end currently exists for each Gemini telescope, the as-built design of the telescopes includes this ability for relatively easy exchange of top end assemblies. In principle, the top end can be quickly and easily disengaged, and installed with excellent mechanical alignment to within 100 microns.

The KAOS implementation will take advantage of this design feature and will utilize that interface for both the mechanical structure of the top end and for the location of a fiber cable interconnect.

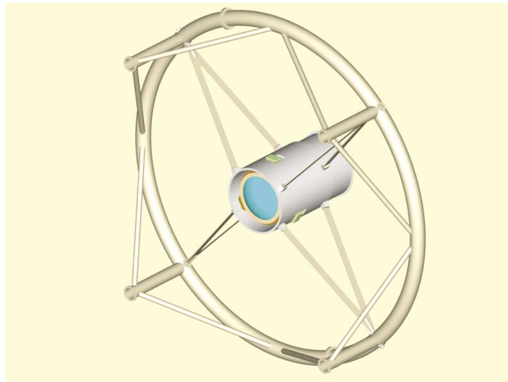


Figure 3.1.15: Solid model rendition of KAOS top end assembly.

Steel versus Composite Materials

The total mass of the top end is a critical concern for the KAOS concept given the significant mass of the wide field corrector. If the top end structure is made from steel, the total mass of the top end is estimated to be 10,100 kg or about 1.6 times that of the currently existing f/16 top end assembly. In order to balance the telescope for this KAOS assembly, an additional mass totaling 18,600 kg would be needed behind the primary mirror in addition to the currently existing mass of the ISS and mounted instruments!

If, on the other hand, the top end ring structure were to be made from a composite Carbon Graphite material, the mass of that component would decrease by a factor of three reducing the total mass of the KAOS top end to 6500 kg. This is nearly the same as that of the currently existing top end and would only require an additional counterbalance mass of 1200 kg behind the primary mirror. The following table summarizes these mass estimates.

Surely the composite material will be preferred to steel for the construction of the KAOS top end assembly. An additional benefit of a composite structure is the ability to tune the structural behavior of the assembly through engineering analysis. This could yield a structure that might help damp out windshake of the telescope. There are two significant drawbacks to a composite top end that must be considered: a cost of nearly three times that for a steel assembly and a sensitivity of the material to temperature and humidity. Assuming that the humidity sensitivity can be adequately minimized, the cost of the composite structure is more than offset by the advantage of its lower mass compared to that of steel.

Table 3.1.1: Comparison of mass for f/16 versus steel and composite KAOS top ends.

Assembly	F/16 Top End (kg)	Steel KAOS Top End (kg)	Composite KAOS Top End (kg)
Ring	5180	5180	1800
Secondary	705	-	-
Secondary Support	221	-	-
Corrector/Echidna	-	3900	3900
Fiber Cable	-	729	729
Vanes	151	300	300
Total	**6257**	**10109**	**6529**
Counter Balance Needed	*0*	*18600*	*1200*

Preliminary engineering analysis of the top end structure shows that the lowest natural frequency is around 10 to 12 Hz and that operation of the wobble plate will not impact the telescope.

Fiber Optic Cable

The fiber optic cable is the heart of the instrument and considerable care in the design and implementation of the cable must be made to ensure optimal performance of all fibers within the cable.

It is estimated that the fiber length will be 60 meters. Multiplied by 5000 such fibers, a total of 300 km (186 miles) worth of fiber will be utilized in KAOS! This is twice the distance between Tucson and Phoenix.

Fiber Optic

The nominal concept for KAOS utilizes two sizes of fiber. A small diameter fiber installed within the Echidna positioner, and a larger diameter fiber coupled to the small fibers at the top end structural interface and feeding into the spectrographs located in the pier lab of the observatory. The Echidna fibers, matched to ~1 arc-second on the sky, are 100 microns in diameter. This choice maximizes optimal signal to noise performance of sky limited targets imaged onto a circular aperture with respect to the median (0.6 to 0.7 arc-second) image quality anticipated. The f/2.4 input focal ratio will enter directly into the fiber without relay optics. Due to the angular deviation of the fibers as they are moved to different locations requires the fibers to be capable of accepting a beam of f/2 or faster. This implies a numerical aperture (N.A.) of at least 0.25 or larger. This N.A. is near the limit for glass clad fibers. Plastic clad fibers are also available with numerical apertures ranging from 0.3 to 0.66. An example of such fibers with a hard polymer cladding produced by Polymicro Technologies is given at *http://www.polymicro.com/fbr4.htm.*

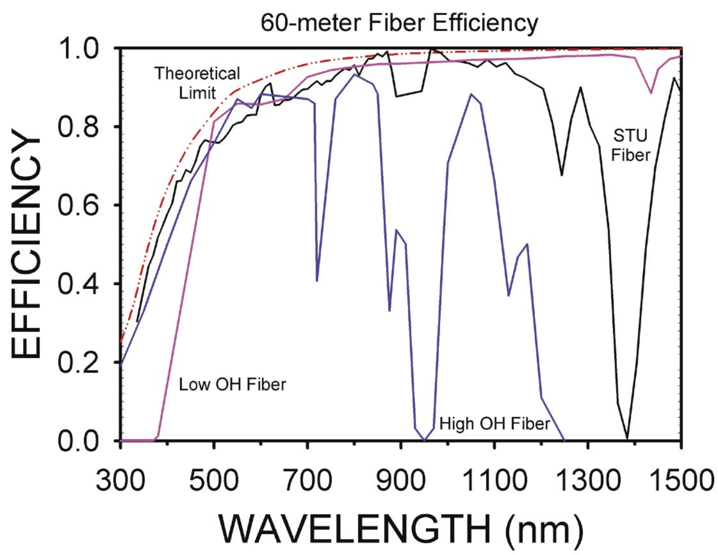


Figure 3.1.16: Expected internal transmission for three types of glass fibers and the maximum theoretical limit. The STU fiber is that preferred for good performances across the CCD response window. The Low OH fiber would be preferred for a near-IR implementation.

These fibers will follow from the Echidna positioner to the mechanical interface of the top end structure to the telescope. At that point there will be a fiber interconnect junction. This junction will serve two purposes. The first is to provide a convenient point for disconnecting the top end assembly without the need to remove the full 60 meter length of the fiber cable. This is required to minimize installation and removal effort of the KAOS assembly. The second purpose is to relay the light from f/2 to f/4 for propagation through the remainder of the fiber cabling. It is near this focal ratio that excellent image scrambling is achieved while focal ratio degradation is still kept at a minimum (see Barden, Elston, Armandroff, and Pryor, 1993, ASP Conf Ser 37, pp223-234 and Watson and Terry, 1995, Proc SPIE 2476, pp10-19). Glass clad fibers can be utilized for this portion of the fiber cable (see *http://www.polymicro.com/fbr1.htm* for examples of usable fibers from Polymicro Technologies). The transfer from f/2 to f/4 doubles the core diameter for this fiber to 200 microns.

It would be desirable to have the fiber drawn from the Heraeus STU glass for optimal transmission from the blue to the near-infrared. Figure 3.1.16 shows the expected transmission for the three types of typical silica core materials available.

It is important to note that the cladding of the fiber should be at least a factor of ten times larger than the longest transmitted wavelength to prevent leakage from the fiber. The fiber dimensions should therefore be 100 micron core surrounded by at least a 15 micron cladding.

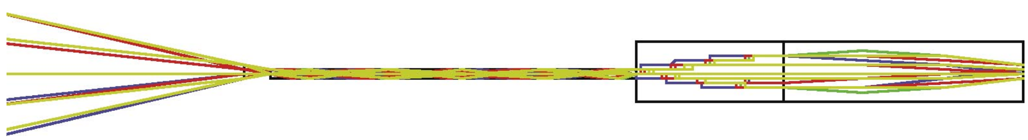


Figure 3.1.17: Zemax layout of input fiber and GRIN lens relay. The f/2.4 light from the corrector enters the 100 micron fiber on the left. The output of the second GRIN lens would feed into the 200 micron fiber.

Fiber Relay

The transfer relay in the fiber interconnect junction could be done with a pair of GRIN lenses, one cemented to the output of the small fiber, the other cemented to the input of the larger fiber. The lenses could then be optically gelled together in the connector for low loss coupling. An air gap could also be used with high performance AR coatings applied to the surfaces of the GRIN lenses. This might maximize cleanliness of the cable ends and minimize support issues related to ensuring the presence of adequate optical coupling material, but will likely result is slightly higher coupling loss. The use of micro-lenses might also be preferable to GRIN lenses in order to achieve a higher level of coupling efficiency by minimizing the chromatic aberrations inherent in GRIN lenses.

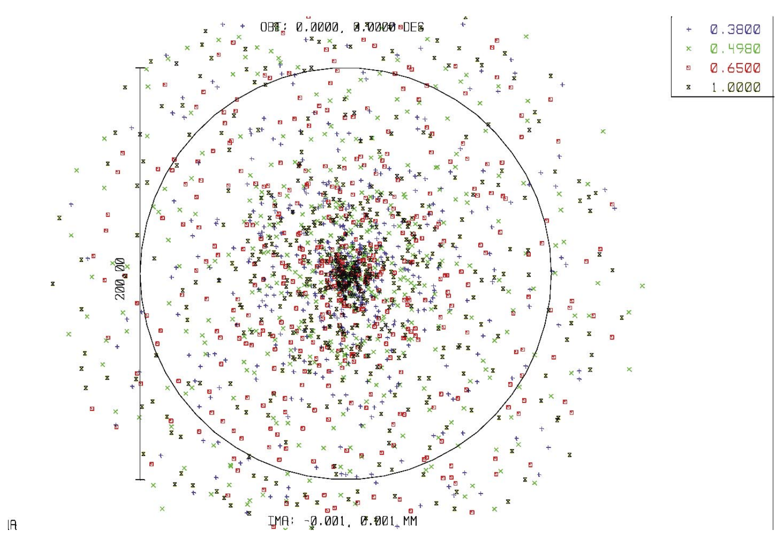


Figure 3.1.18: Spot diagram showing illumination of the 200 micron fiber by the GRIN lens relay.

Figures 3.1.18 and 3.1.19 display the spot diagram and encircled energy plot for the output of a preliminary design for the GRIN lens relay. This particular design gives a geometric coupling efficiency of only 80% between a 100 and 200 micron fiber. Additional analysis indicates that this coupling loss can be further minimized through the use of alternative GRIN or micro-lens designs. It is hoped that such coupling loss would be no larger than 5 to 10% rather than the 20% level shown here. It should be noted that the FMOS/Echidna instrument for Subaru will utilize a similar interfiber relay that the University of Durham is fabricating (http://www.naoj.org/staff/akiyama/FMOS).

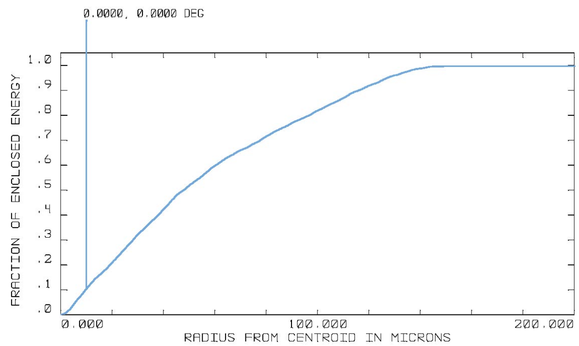


Figure 3.1.19: Encircled energy plot for the GRIN lens relay. 80% EE in capured by the 200 micron fiber.

This fiber relay junction will also be the location where other fiber instruments for the KAOS top end assembly would interface to the fibers feeding the pier spectrographs. For example, it may be desirable to fabricate a 69x69 fiber monolithic IFU array or a small set of deployable IFU's (e.g., 48 10x10 arrays) for use in place of the Echidna positioner. The output of those devices would plug into the KAOS fibers at this relay interface.

Fiber Cable Structure

Each fiber should be sheathed in its own protective jacket. Typically, this jacket is made of Teflon or some other low friction material. The fibers are then typically grouped inside a protective, flexible, metal conduit.

In order to minimize the potential for cable induced stresses on the fibers that might increase the fiber focal ratio degradation, it will be necessary to incorporate thermal expansion joints along the length of the cable (approximately every 10 meters) and appropriate routing of the fibers through constrained flexible channels for flexing about the elevation axis. Both of these concepts have been developed by the HectoSpec instrument (see *http://cfawww.harvard.edu/mmti/hectospec/hectospec.pdf*) to minimize cable induced stress.

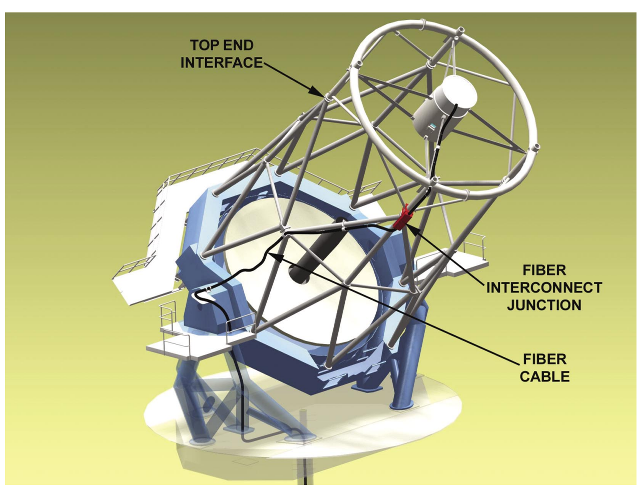


Figure 3.1.20: Solid model showing the path for one of the three fiber cable runs.

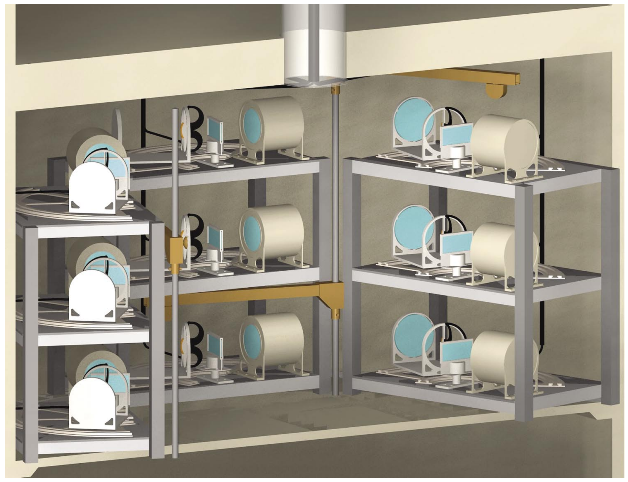


Figure 3.1.21: Fiber routing from the ceiling to the spectrographs in the pier lab. The 12 spectrographs are stacked in groups of three around the circumference of the room. Three of the four stacks are shown here.

Fiber Cable Routing

The intention is for the long set of fibers leading into the pier to remain permanently on the Gemini telescope. This allows the cable to be continuous and to utilize a pathway that minimizes the total length while keeping the fibers out of the way of the Cassegrain mounted instruments. Figure 3.1.20 shows the proposed routing for the fiber cables.

There are four main trusses that interface to the top end assembly. The fibers could be divided into four groups (1200 fibers in each), each one traversing down a separate truss. However, it should be noted that Gemini intends to utilize one of those routes for the optical train of a laser launch facility. That particular truss could be avoided by the fibers. This may be a better match to the concept of dividing the fibers into thirds, of which 1/3rd of the fibers feed into high resolution spectrographs while the remaining 2/3rd feed into low dispersion instruments.

The volume of space occupied by the fiber cables is non-negligible. The total set of 5000 fibers can squeeze into a circular cross-section of 100 to 125 mm diameter. However, it is likely that the cable must allow more “breathing” space for the fibers. The fibers could be grouped into subsets of 400 fibers each. These subsets would be housed into flexible tubes of 50 mm diameter. Four of these cables would route down the side of each of the three truss tubes.

Near the elevation axis of the telescope, the fiber cabling would pass through access ports into the tubular steel structure of the telescope mount. From there they are routed inside this well-protected environment to a point near the center of the Azimuth axis. The cables exit and descend down the center of the pier to the spectrograph lab located on the ground floor. This routing allows the fibers to completely avoid the area occupied by the Cassegrain mounted instruments. It provides protection to the fibers from Cassegrain instrument support activities and minimizes the path length and service loops required for telescope motion. A small loop will be required near the point of the elevation axis. Rotation about the Azimuth will be taken up by a gentle twisting of the fiber cable down the full length of the pier structure from the top of the pier to the ceiling of the spectrograph room. Figure 3.1.1 shows the fibers dropping through the central axis of the pier.

At the top of the spectrograph room, the fibers will branch out to feed the set of spectrographs (see Figure 3.1.21). The division of the fibers into twelve sub-cables will help aid in this distribution.

Modal Noise

The modes of light propagation in a fiber can shift as the fiber is flexed about. This is easily observed when illuminating a fiber with a laser and the fiber is touched. The speckle pattern of the output beam can be seen to shift and change pattern.

In the majority of astronomical applications, this aspect of light transmission through fibers can be ignored. However, when the light is highly dispersed, as in a high-resolution spectrograph, the shifting modes can change the way the light is illuminating the detector pixels. A difference in this pattern between the flat field, sky spectrum, and object spectrum will result in imperfect calibration of the spectrum. The noise of the spectrum is increased above that expected from normal photon statistics. If the wavelength bandwidth is broad enough, the modes (or speckles) are blurred enough to become negligible and explain why typical observations are not impacted by the modal noise. Baudrand and Walker 2001 (PASP, 113, pp851-858) discuss this and show a simple cure that involves vibrating the fiber so that the modes are continually moved around and blurred out. With the vibration, the noise in the resulting spectrum becomes photon noise dominated.

Modal noise in the KAOS instrument may be an issue for high signal to noise applications, particularly in the high-resolution regime. It may also be a limitation for the observation of extremely faint targets in which the sky must be exposed to very high signal to noise levels. The potential impact of modal noise in KAOS should be further explored. However, it can be minimized through the vibration of the fiber cables. It may be desirable to implement a vibration mechanism onto the fiber cables at some location in the pier room. The vibration should induce a physical motion of the fiber with amplitude of about 1 mm and a rate of several Hertz. The location of the vibrating mechanism should be such as not to induce vibration in the spectrographs themselves.

Spectrographs

Detecting nearly 5000 spectra with significant spectral coverage and a wide range of spectral dispersions can prove to be a very difficult task for a single spectrograph. Such a spectrograph is likely to be high risk and costly. It is assumed that the KAOS spectrographs will be an ensemble of multiple, identically made instruments, each viewing a subset of the fibers.

If the fibers are divided into two groups (1/3rd for high resolution and 2/3rd for low resolution applications), then two types of spectrographs could be constructed. For simplicity, the nominal concept described here addresses one generic spectrograph design for both applications. The division into two groups could still yield cost savings in that only 1/3rd of the spectrographs need high dispersion gratings.

Beam Diameter

The minimum monochromatic, on-axis beam diameter of the spectrograph is determined by the highest resolving power required for the science, the entrance aperture, and the highest angle of diffraction permitted by the gratings. For KAOS, the entrance aperture is approximately 1 arc-second. The highest resolving power desired is on the order of R=20,000 to 40,000. If one assumes a maximum diffraction angle of 63 degrees (R2 Echelle) for the R=20,000 case, then the minimum beam diameter is 200 mm (8 inches). An R4 Echelle (76 degree diffraction angle) could be utilized to achieve R=40,000.

It is possible to use an image slicer to reduce the beam diameter by narrowing down the entrance aperture. However implementing slicers on 5000 (or even just 1600) fibers would be a daunting and very expensive task. It would also require more detector pixels to cover the full spectrum. It is therefore assumed that KAOS will not utilize image slicing.

Therefore, it is desirable to have a beam diameter of at least 200 mm for the highest resolution spectrographs. The low-resolution instruments could have smaller beam diameters, but the case will be made for them to also have beam diameters of at least 200 mm.

Another factor that influences the choice of beam diameter is the physical size of the detector and related field angles for the camera. A small beamed instrument feeding a very large detector can result in a very complex camera design and minimize any cost advantage for such a small beamed instrument compared to that of a larger beamed instrument better matched to the detector size. Any further conceptual design study for an instrument like KAOS should involve a trade study between beam diameter and detector size.

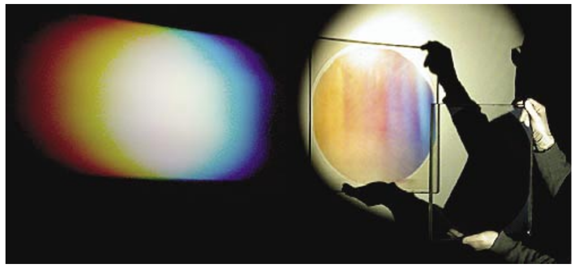


Figure 3.1.22: Photograph of light being diffracted from a 380 mm diameter VPH grating fabricated by the CSL group (image courtesy of the CSL group).

For the current study, it was assumed that an internal focus, Schmidt-like camera would be more cost effective than the multiple production of a complex, all-transmissive camera with an external focus. This decision also feeds into selection of the beam diameter for the spectrograph. The detector size will directly impact the central obstruction loss of the spectrograph with such an internal focus camera. It is desirable to have the beam diameter be at least a factor of three times larger than the diameter of the detector. Otherwise, the central obstruction loss becomes quite significant. This is particularly true for fiber-fed spectrographs in which any central shadow of the telescope is washed out by the focal ratio scrambling of the fiber, which is very likely the case for the 60 meter long fibers of KAOS.

It is therefore assumed that the KAOS spectrographs will have internally focused cameras and beam diameters of at least 200 mm. This makes an excellent match to 4K by 4K, 15 micron detector formats. The 200 mm beam diameter is about a factor of three times larger than the physical size of the detector array.

It may be desirable to increase the beam diameter up to at least 300 mm in size. This gains in a variety of areas. The highest resolving power of the instrument is increased by a factor of 1.5 (R=30,000 becomes possible with an R2 Echelle). It would also be possible to fit a 4K by 6K detector format and decrease the amount of lost light due to the central obstruction of the detector. A factor of 1.5 more fibers could be fed into such an instrument or a factor of 1.5 gain in spectral coverage be attained. For cost reasons, the 6K dimension would best be utilized to increase the number of fibers feeding the spectrograph reducing the total number of instruments required. However, the availability of a range of 300 mm sized gratings may be slightly more problematic than for a 200 mm beam diameter. Fortunately, there are some Echelle grating choices available from the Richardson Grating Laboratory (a 31.6 l/mm R3 grating and two R2 gratings with 87 and 110 l/mm, see *http://www.gratinglab.com/products/table4.asp*). A wide range of VPH gratings of this size (up to ~400 mm in diameter) can be obtained from a couple of vendors: Wasatch Photonics (see *http://www.wasatchphotonics.com*) and a spin off company from the Centre Spatial de Liege (CSL) in Belgium *http://www.ulg.ac.be/cslulg/Activities/AdvancedOptics/Advanced_main.html* (see Figure 3.1.22 for an image of one of their large gratings).

A 200 mm beam diameter was assumed for the nominal concept study described here.

Camera Focal Length

The pixel size and required resolution sampling set the camera focal length. The preliminary study for a spectrograph assumed a sampling of 3.3 pixels per resolution element across the 1 arc-second fiber (a plate scale of 0.3 arc-seconds per pixel). That results in a camera focal ratio of ~f/1.3.

There is the additional impact due to the tilt of the fibers when positioned at the edge of their accessible range effectively resulting in illumination at the input end in an f/2 cone rather than f/2.4. That impact requires that the camera either be sped up to f/1 or that the spectral resolving element be 4.4 pixels rather than 3.3. *We note that this aspect was ignored for this preliminary conceptual design exploration and that it should be accounted for in any future design studies. The remaining discussion of the spectrograph concept ignores this affect.*

Number of Fibers per Spectrograph

It is assumed that Nod & Shuffle observing (see chapter 3 for a discussion of the benefits of Nod & Shuffle observing for KAOS) will be required for KAOS. This requires that the CCD detectors have sufficient pixels to allow the spectra to be shuffled between two positions. Full chip shuffling requires a factor of three times as many pixels as get illuminated by the spectrograph. Short shuffling between spectra requires only a factor of two more pixels, but demands sufficient space between the spectra for the shuffled image.

Given the internal focus of the camera, it is desirable to keep the size of the detector to a minimum. Hence the short shuffling is assumed. If the spectra illuminate a swath of pixels with a FWHM of 3.3 pixels, then the spectra should be separated by a factor of at least 2.5 to 3 times that width. This yields coverage of 1 input spectrum for every 10 pixels on the detector, or 400 fibers per 4K detector.

Number of Spectrographs

With 400 fibers per 4K detector, a total of 12 spectrographs are required to cover the full complement of 4811 fibers in KAOS. If a 300 mm beam design is chosen instead with 6K of pixels available along the slit, then the number of spectrographs is reduced to 8.

With the division of fibers into high spectral resolution and low-resolution applications, four of the 200 mm diameter spectrographs are needed to cover the full complement of high dispersion fibers while eight are needed for the low dispersion use. Those numbers can drop to 2 or 3 spectrographs with 300 mm beam diameters for the high dispersion case and 6 spectrographs for the low dispersion application.

The high dispersion spectrographs may not necessarily need to be operated in Nod & Shuffle mode due to the brighter magnitudes of the objects most likely to be observed. Therefore, the number of spectrographs could be reduced by another factor of two for the high-resolution case. However, it may be desirable to retain Nod & Shuffle capability in those spectrographs, as that would allow appropriate sky subtraction on faint targets when the moon is bright. For the intention here, it is assumed that Nod & Shuffle mode is preserved for these spectrographs.

Nominal Spectrograph Concept

The design criteria for the nominal spectrograph concept presented here assumed the following parameters (please note that the impact of tilted fibers in the focal plane were ignored):

- 1 arc-second effective apertures (150 micron fibers)
- ~f/4 collimator (f/3.87 resulted in the design)
- 200 mm beam diameter
- 0.3 arc-second per pixel plate scale on the detector
- 15 micron pixels
- f/1.29 camera
- Camera focal length = 257.8 mm
- Demagnification ratio of 3.0

The resulting design has the following characteristics

- 200 mm beam diameter
- All spherical surfaces
- Reflective Collimator
- Quasi-Houghton Maksutov Camera
- 5 element Corrector

- Internal Focus
- 4K by 4K CCD format
- Use of volume phase holographic (VPH) gratings
- Articulated fiber/collimator arm
- 2 pixel RMS spot diameters
- Parfocal camera across full wavelength band.

Figure 3.1.23 displays a Zemax ray trace of the concept for all of the configurations modeled giving resolving powers that ranged from R=2000 to R=20000. A solid model rendition of the concept is shown in Figure 3.1.24.

Configurations for R=2000, 5000, 10000, and 20000 were modeled centered at a wavelength of 5000 Å. At 8500 Å, configurations were modeled to give R=2000 and 20000.

Table 3.1.2 shows a nominal set of gratings for KAOS for observations centered at 500 nm in the blue, 650 nm, and either 850 or 800 nm in the red. Other, intermediate gratings will also be desired, particularly for the higher dispersion applications. This set of gratings was used to generate the performance characteristics of KAOS.

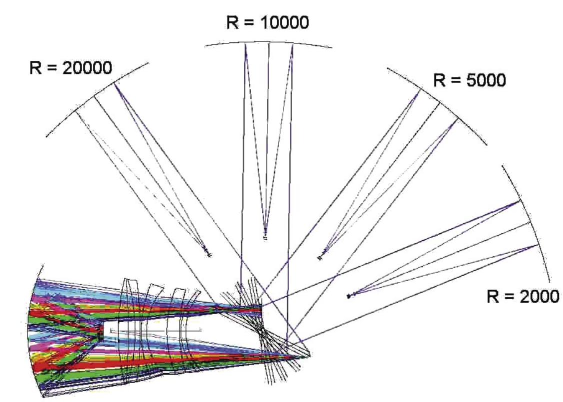


Figure 3.1.23: Zemax raytrace of one of the KAOS spectrographs.

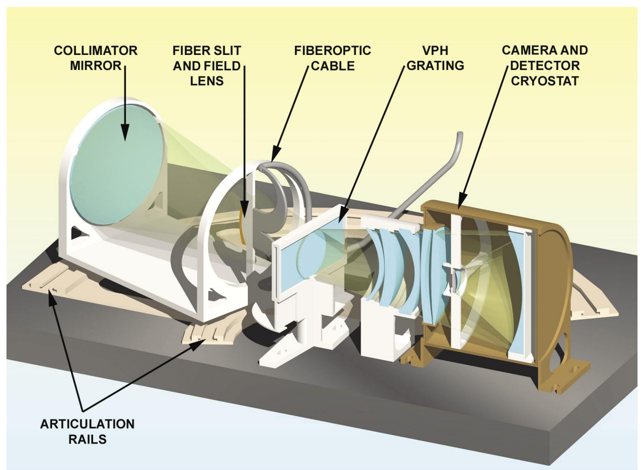


Figure 3.1.24: Solid model rendering of one of the KAOS spectrographs.

Layout of Spectrograph in Pier

The dozen spectrographs will reside in the pier lab underneath the Gemini telescope. Four stacks of three spectrographs each will fill the room, but leave sufficient space for handling equipment to be maneuvered by the support personnel. Figure 3.1.21 (shown previously) displays a cutaway view of the pier lab with three of the four spectrograph stacks displayed.

A portable instrument lift can be used to service each stack. A hoist for relaying the equipment from the lift onto the table can reside at each stack and service each of the three spectrographs in each stack.

The fibers will branch out from the center of the ceiling to each stack and then to each individual spectrograph. The cryostats for each spectrograph contain the final element of the camera corrector and the spherical mirror along with the detector array. They are kept cold by a system of recirculated liquid Nitrogen (LN2). A closed cycle cooler maintains the reservoir of LN2 used to service the cryostats. The LN2 reservoir and electronic components will likely have to reside immediately outside, but adjacent to the pier room.

This layout assumes that all of the spectrographs are identical.

Table 3.1.2: KAOS grating configurations.

λ_c	R	λ cov	disp	λ_{res}	ν	α	d	dn	η_{peak}	η_{ends}
Å	$\lambda/\Delta\lambda$	*Å*	*Å/pix*	*Å*	*l/mm*	*deg*	*μm*			
5000	1100	3380-6630	1.39	4.64	433	6.2	15	0.02	0.83	.35 - .78
5000	2000	3500-6480	0.77	2.56	770	11.1	6.5	0.04	0.89	.37 - .71
5000	5500	4450-5500	0.27	0.91	1915	28.6	4.5	0.06	0.93	.55 - .64
5000	9600	4680-5255	0.16	0.52	2802	44.5	4	0.07		
5000	19300	4835-5120	0.08	0.26	3564	63.0	9	0.07	0.87	.15 - .38
6500	1100	4500-9000	0.55	1.84	330	6.2	34	0.01	0.90	.32 - .55
6500	2000	4500-8300	0.98	3.28	600	11.1	8	0.04	0.89	.40 - .70
6500	5500	5750-7150	0.36	1.19	1473	28.6	6	0.06	0.92	.48 - .62
6500	9800	6080-6880	0.20	0.66	2156	44.5	5	0.07	0.79	.45 - .46
6500	19300	6280-6650	0.10	0.34	2740	63.0	11	0.07	0.87	.25 - .50
8500	1100	5620-11250	2.37	7.91	255	6.2	26	0.02	0.83	.30 - .78
8000	2000	5500-10350	1.23	4.08	482	11.1	10	0.04	0.90	.36 - .70
8500	5500	7550-9350	0.47	1.56	1126	28.6	7.5	0.06	0.93	.56 - .65
8500	9800	7950-8990	0.26	0.87	1648	44.5	6.5	0.07		
8500	19500	8210-8730	0.13	0.44	2096	63.0	14	0.07	0.85	.29 - .43

The R=1100 cases underfill the CCD detector by almost a factor of two yielding greater than 1 octave of coverage. The efficiencies are for the grating only. The peak column gives the efficiency at the center of the wavelength coverage, the ends column gives the efficiency for both ends of the spectrum covered by the CCD.

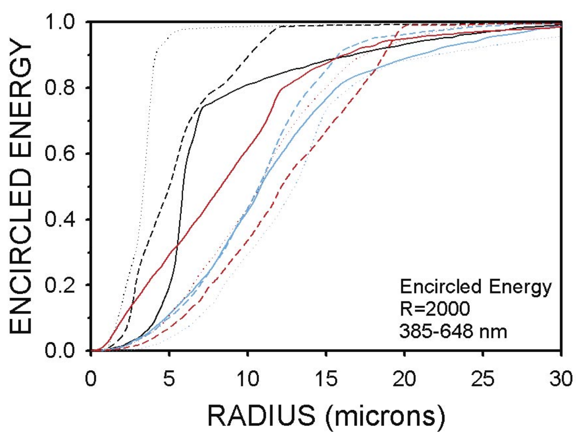


Figure 3.1.25: Encircled energy plot for blue R=2000 configuration. Two pixels equal 15 micron radius.

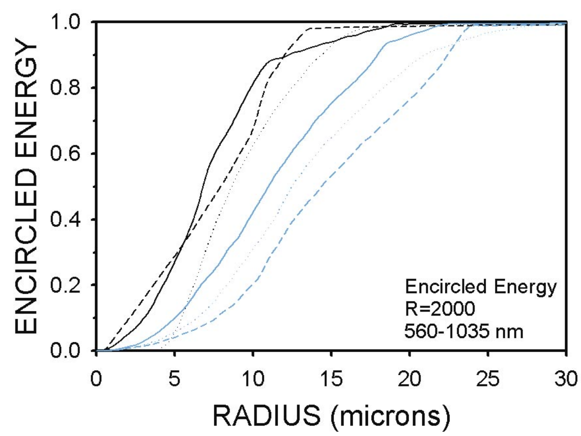


Figure 3.1.26: Encircled energy plot for red R=2000 configuration. Two pixels equal 15 micron radius.

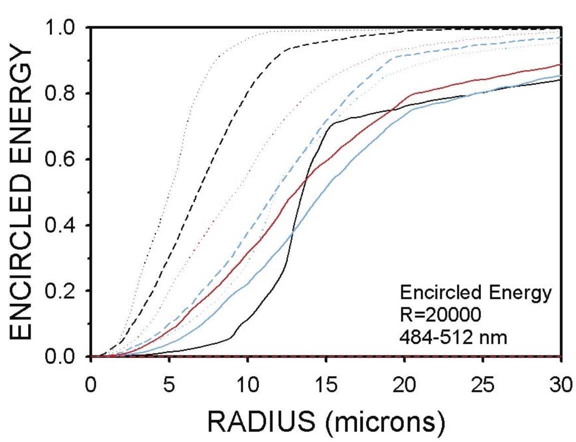


Figure 3.1.27: Encircled energy plot for blue R=20000 configuration. Two pixels equal 15 micron radius.

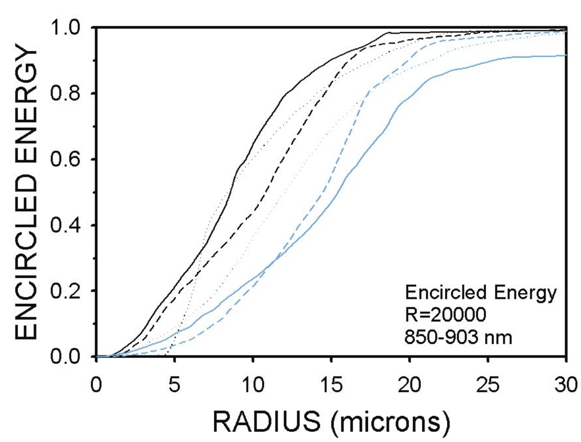


Figure 3.1.28: Encircled energy plot for red R=20000 configuration. Two pixels equal 15 micron radius.

Figures 3.1.25 through 3.1.28 display the ray traced encircled energy for a variety of the configurations across the detector (center and ends of spectrum and slit). In general, 50% encircled energy is well under a 2 pixel diameter.

Predicted Performance for KAOS

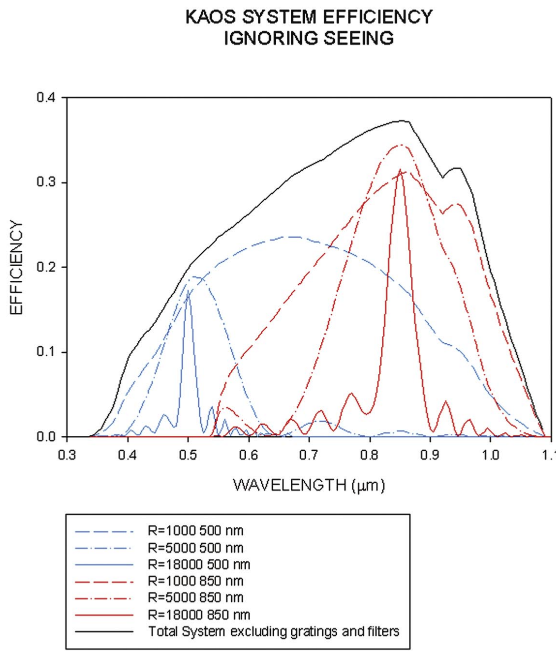


Figure 3.1.29: Total system efficiency for KAOS inclusive of telescope, fibers, and detector. Six different grating configurations are displayed along with the efficiency exclusive of the gratings

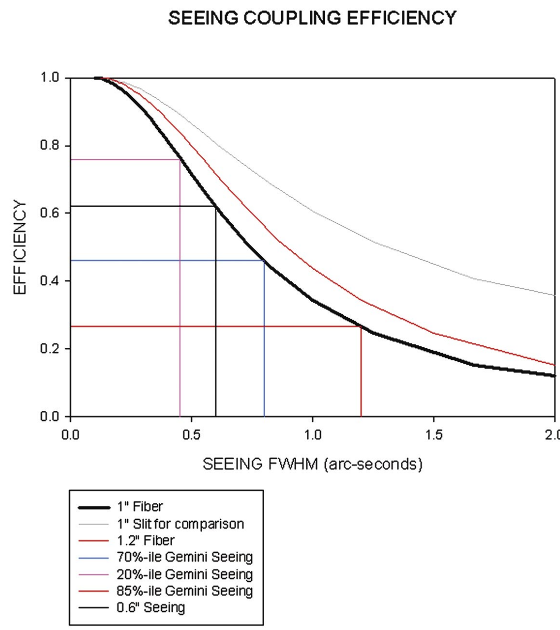


Figure 3.1.30: Aperture coupling efficiency for a 1" fiber as a function of seeing FWHM. A 1.2" fiber and slit are also shown for comparison. 20%, 70%, and 85% seeing cases are shown along with that for 0.6" seeing FWHM.

Figure 3.1.29 shows the predicted system effiency for KAOS as a function of fraction of detected photons that are incident on the telescope primary mirror. This includes all known sources of losses, except for the seeing loss on the circular aperture. Data from Figure 3.1.30 (showing the aperture efficiency as a function of image quality) should be utilized to scale Figure 3.1.29 as a function of input seeing.

The following set of figures (3.1.31 through 3.1.34) shows the breakdown of assumptions for each component in the KAOS system. The majority of components were appropriately modeled for efficiency as a function of wavelength. The anti-reflection coatings, however, were assumed to be flat over wavelength at a level of 1% reflective loss per air/glass surface. This assumption was used due to the wide variety of options available for AR coatings. The most likely favored coatings for the majority of surfaces are the combination of MgF2 and SolGel. Such coatings can have excellent broadband performance with the peak being well under the 1% level. It is assumed that such coatings on different elements would have their peak shifted across wavelength in order to optimize the broad band nature of the system rather than focus on a specific peak wavelength. That assumption led to the flat 1% average performance. Future studies should fold in a more realistic AR coating behavior.

The gratings were all assumed to be volume-phase holographic (VPH) gratings. The efficiencies were computed using a piecewise approximate, rigorous coupled wave analysis. The program used was Gsolver (*http://www.gsolver.com/*), which works well for VPH gratings. The grating parameters assumed in the calculations are listed in Table 3.1.2.

The resultant curves were utilized in an exposure time calculator that has been implemented on the Web (*http://www.noao.edu/gateway/spectime/gemkaos.html*). Figures 3.1.35 and 3.1.36 show the S/N versus magnitude for the various blue and red configurations respectively.

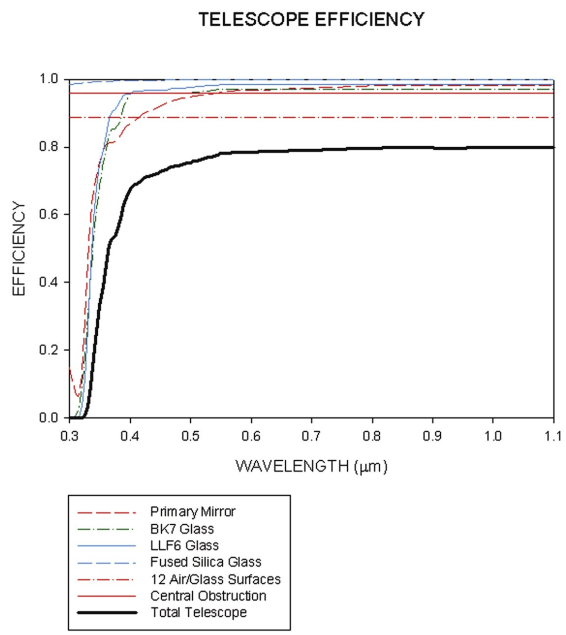


Figure 3.1.31: Efficiency of components in the telescope and wide field corrector.

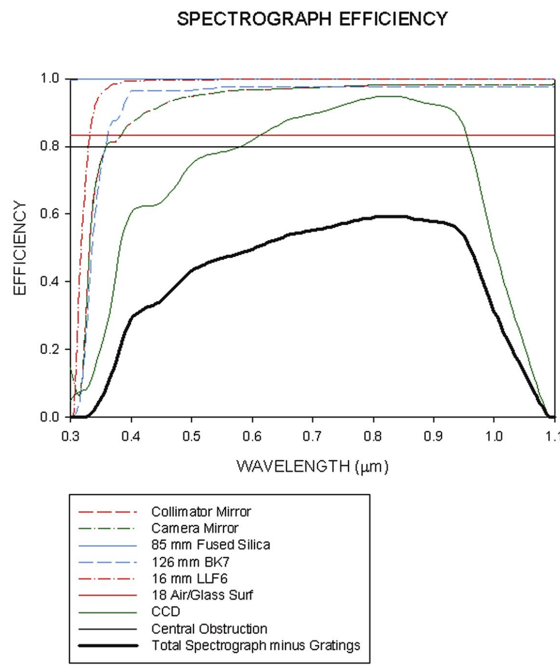


Figure 3.1.32: Efficiency of components in the spectrograph.

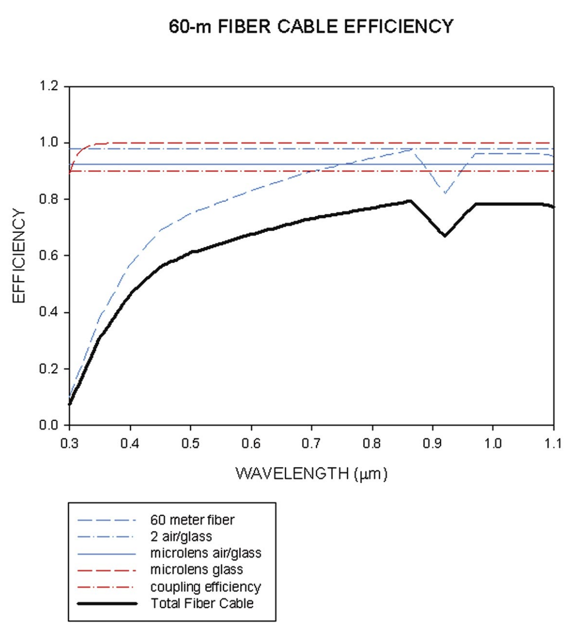


Figure 3.1.33: Efficiency of the 60 meter fiber cable.

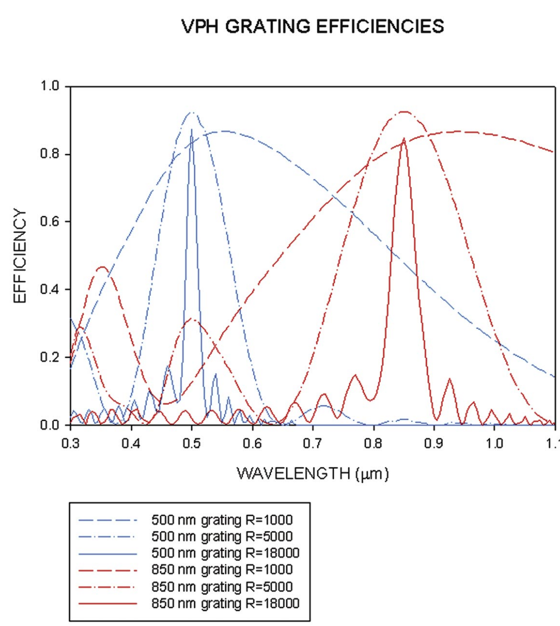


Figure 3.1.34: The computed efficiency for the six VPH gratings.

Other Spectrograph Concepts

There are many alternative approaches in spectrograph design that could and possibly should be explored during a full conceptual design study for KAOS. Among these are the following:

- Can a DEIMOS type spectrograph work for the low dispersion fibers?
  - DEIMOS has an 8K by 8K detector format. If it is assumed that such a format could replace four of the KAOS spectrographs, then a total of two DEIMOS style spectrographs would be required to image the full set of 3200 low resolution fibers and allow Nod & Shuffle observing. This requires the use of two parallel fiber slits per spectrograph. A cost/performance comparison would be required.
  - The DEIMOS concept will likely not work well for the high-resolution applications. An alternative design would be required for that set of fibers.
- Would other "transmissive" solutions work for KAOS?
  - An investigation into a transmissive collimator is underway. That will allow the fiber slit to be pulled out of the spectrograph beam, simplifying the fiber slit assembly since it will not obstruct any light.
  - The camera optics may become complex and expensive. The DEIMOS alternative might be the best approach in this manner: larger detector, fewer total number of spectrographs.
- Use of classical Echelle gratings for the high dispersion application.
  - The availability and performance characteristics of VPH gratings working at 63 and 76 degree diffraction angles are not well understood. An effort is underway to evaluate the performance of these gratings in this regime. These are first order gratings, so would require a set of several gratings to achieve adequate wavelength coverage at this resolution. Classical Echelle gratings may provide a better solution given that one grating provides performance over the full spectral window in different grating orders. However, the free spectral range of the Echelle may require the use of relatively inefficient order separation filters that are not needed for the VPH grating equivalents. A performance trade study is required. Figure 3.1.37 shows the comparison of a theoretical VPH grating to the real efficiency of an 87 l/mm, 63 degree (R2) Echelle grating.
  - A preliminary look at the current concept shows that Echelle gratings could be used with the current collimator and camera arrangement.

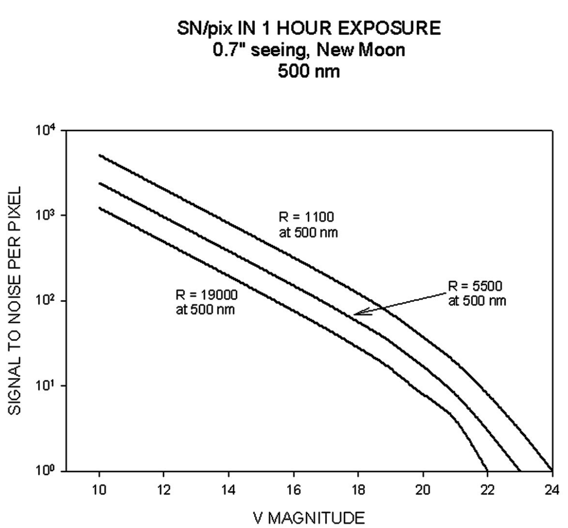


Figure 3.1.35: Predicted S/N versus V magnitude at 500 nm.

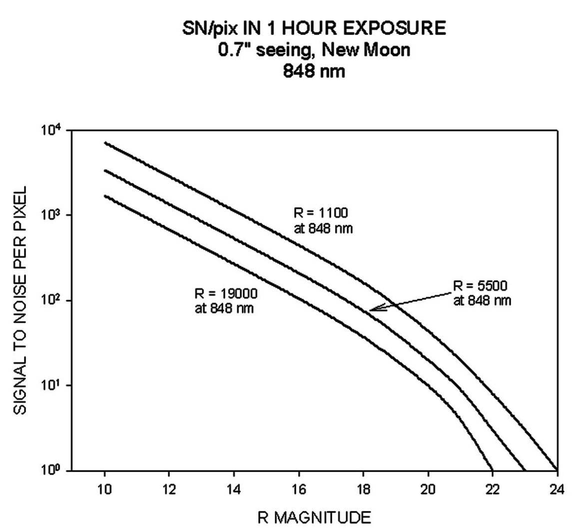


Figure 3.1.36: Predicted S/N versus R magnitude at 850 nm.

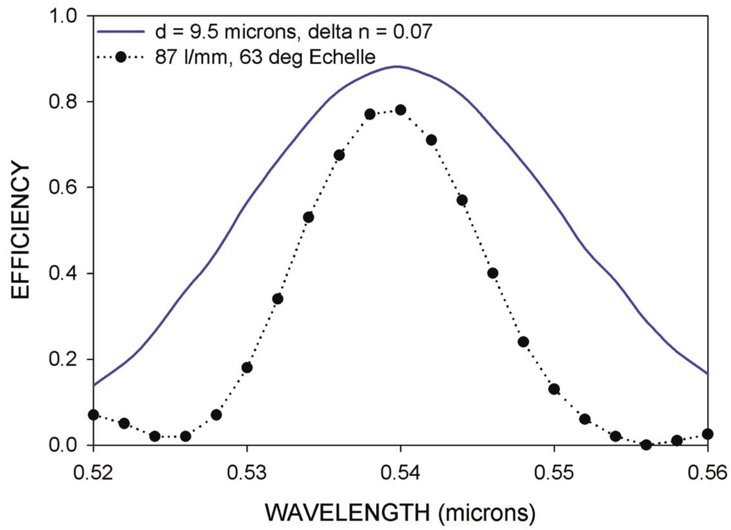


Fig 3.1.37: Comparison of the efficiency of a VPH grating (solid line) with that of an R2 Echelle grating (dotted line).

F/6 vs Prime Focus

Why is KAOS a prime focus implementation rather than the original Gemini concept for a wide field with an f/6 secondary?

The original concept of a wide field mode for the Gemini telescopes involved the implementation of a secondary mirror with a focal ratio of f/6 that would produce a 45 arc-minute field of view at the Cassegrain location. The intention was to place a multi-object spectrograph at Cassegrain, either slit or fiber fed. The final design and construction of the Gemini telescopes retained the f/6 concept as a future upgrade. It is this reason that the top end assembly was designed for relatively easy removal and installation.

There are numerous reasons why the f/6 design is not ideal for KAOS.

- The field of view is limited to 45 arc-minutes, considerably smaller than the desired 1 to 2 degree field of view for KAOS. This seriously reduces the potential multiplex advantage of KAOS compared to other existing facilities.
- The f/6 secondary requires a 2.4 meter diameter mirror with very tight positional tolerances.
- The central obstruction of the 2.4 meter secondary is significantly larger than that of the prime focus assembly (9% rather than 6%).
- A new ISS is required as the present f/16 instrument support system will not work in an f/6 beam.
- A wide field corrector and ADC are still required.
- A wobble plate will be required for wind shake since the 2.4 meter secondary can not be operated as a tip-tilt element for such correction.
- A new baffle is required as the current one is designed for the f/16 secondary.
- The f/16 instruments will not work with the f/6 secondary.
- The Echidna fiber positioner concept is not suitable for an f/6 beam due to the angular tilts of the fibers resulting in a higher degradation of the focal ratio. An alternative fiber positioning scheme would be required.
- A lower density of fibers will result with the f/6 option.
- The f/6 changeover would require at least a factor of two more effort due to the need to exchange the full set of Cassegrain instruments and ISS units in addition to the top end changeover required for either option. (See Table 3.1.3)

Table 3.1.3: Comparison of f/6 and prime focus installation tasks.

F/6 Secondary Wide Field	KAOS Prime Focus Wide Field
1. Move telescope to horizon and engage locking pins.	1. Move telescope to horizon and engage locking pins.
2. Disconnect services.	2. Disconnect services
3. Position top end storage cart to telescope.	3. Position top end storage cart to telescope.
4. Release motorized top end latches	4. Release motorized top end latches
5. Lower f/16 top end to basement, place on storage fixture.	5. Lower f/16 top end to basement, place on storage fixture.
6. Remove f/6 top end from storage fixture and raise to telescope level.	6. Remove KAOS top end from storage fixture and raise to telescope level.
7. Latch top end to telescope.	7. Latch top end to telescope.
8. Connect services.	8. Connect services.
9. Move telescope to Zenith and engage locking pins.	9. Connect fibers.
10. Remove all f/16 Cass instruments.	10. Install counterbalance.
11. Remove f/16 ISS.	11. Telescope is ready.
12. Install Cass wide field corrector.	
13. Install Cass ISS.	
14. Install f/6 instruments.	
15. Telescope is ready.	
Time Required to Change: ~16 hours	**Time Required to Change: ~8 hours**

The need to change out all of the Cassegrain hardware is estimated to double the time required to implement the change. Reverting back to the f/16 mode is effectively just the opposite of this procedure.

Advantages of the f/6 would include the following:

+ Image quality across the field of view could be excellent. However one could argue that the prime focus would achieve comparable image quality if the field of view were decreased to that of the f/6 option.
+ The fiber length would be decreased from 60 meters to 30, resulting in a slight gain in sensitivity.
+ A larger space envelope is available for wide field instrumentation. Slit-fed instruments would be possible.
+ The alternative fiber positioning scheme might be better suited for the observation of centrally concentrated target clusters, such as galaxy clusters or globular clusters.

Figure 3.1.38 shows a side by side solid model rendering for both prime focus and f/6 options. The wide field correctors for both implementations are displayed in Figures 3.1.39 and 3.1.40. Table 3.1.3 lists the tasks required to reconfigure the telescope for each mode of operation.

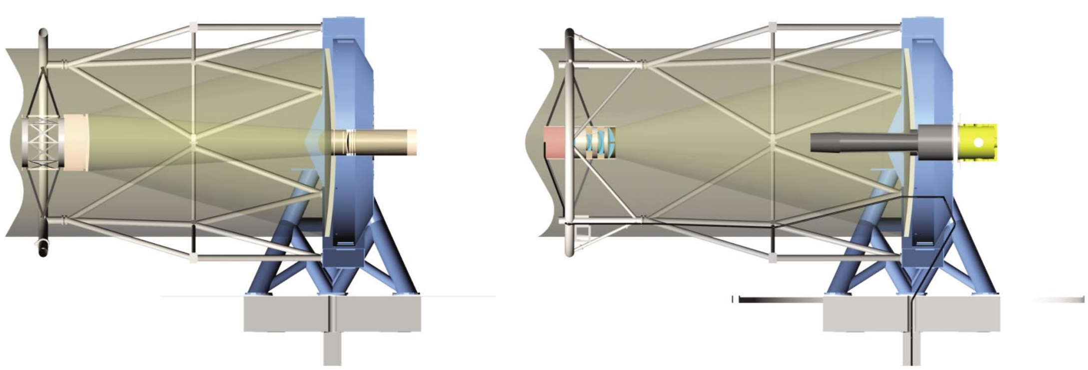


Figure 3.1.38: The original Gemini f/6 secondary option for a wide field (displayed on the left) compared to the proposed KAOS prime focus implementations (shown on the right)

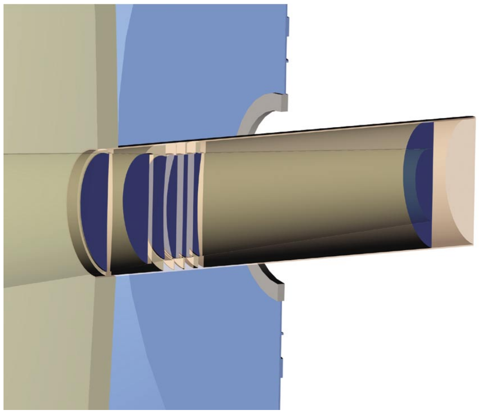


Figure 3.1.39: Gemini f/6 Cassegrain wide field corrector with ADC prisms.

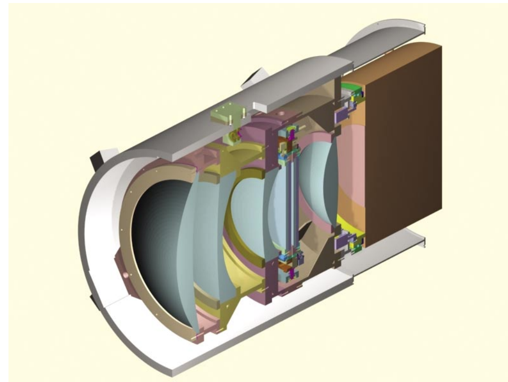


Figure 3.1.40: KAOS prime focus wide field corrector for comparison.

Remaining Issues

Residual issues regarding the KAOS instrumental concept are the following:

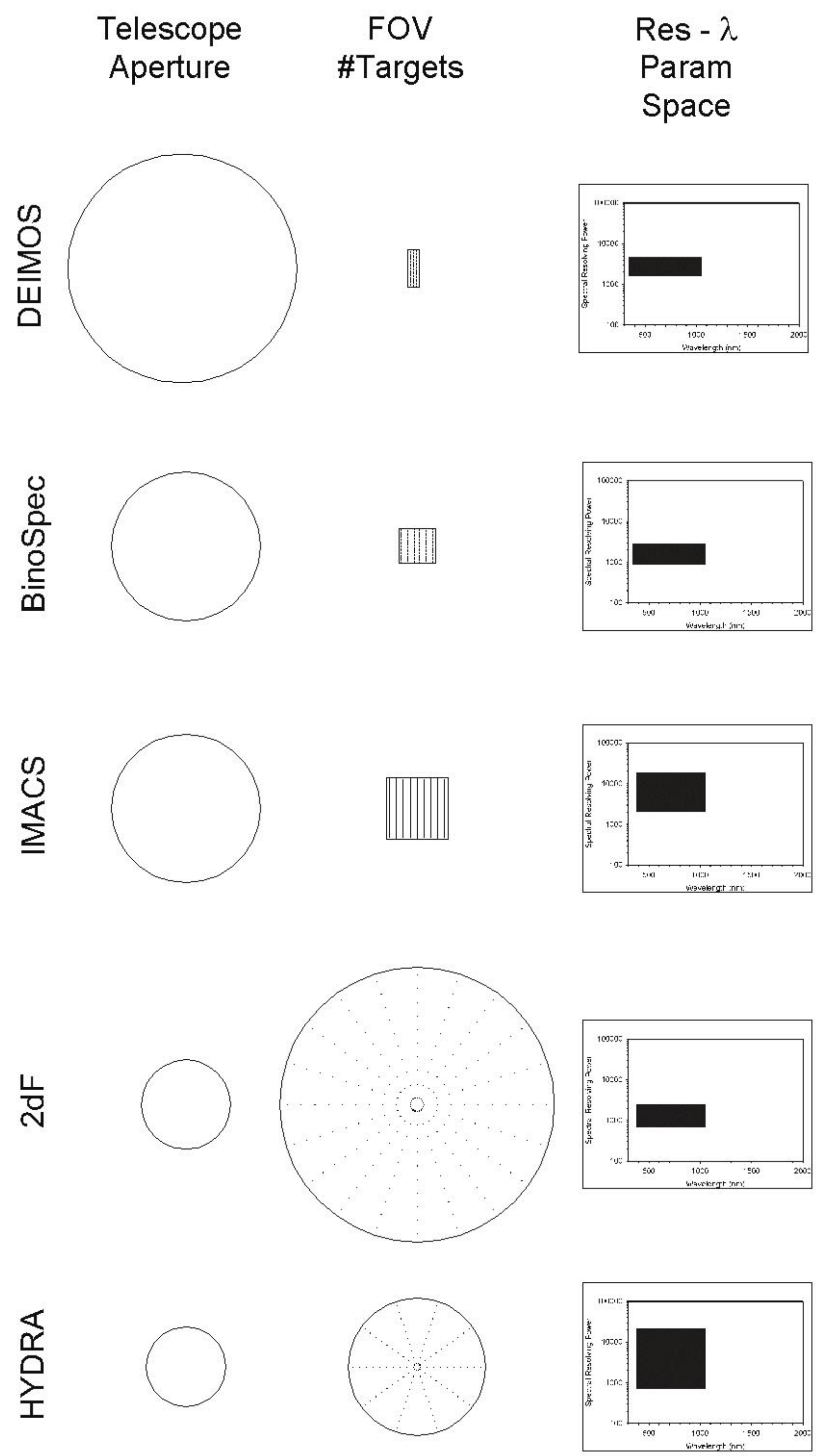
- Understanding the impact of Echidna spine tilt on spectrograph performance and design.
- Exploration of alternative spectrograph concepts, in particular a transmissive or off-axis collimator that allows the fibers to be outside the beam of the instrument. *(Preliminary alternative collimator concepts are currently ongoing.)*
- Design study of 300 mm beam spectrograph.
- Exploration of a single spectrograph concept similar to DEIMOS.
- Implementation of ruled R2 and R4 Echelle gratings into the spectrograph design concept. *(Quick look shows current optical concept will work with Echelles, but full evaluation is still required.)*
- Optimize fiber relay optics.

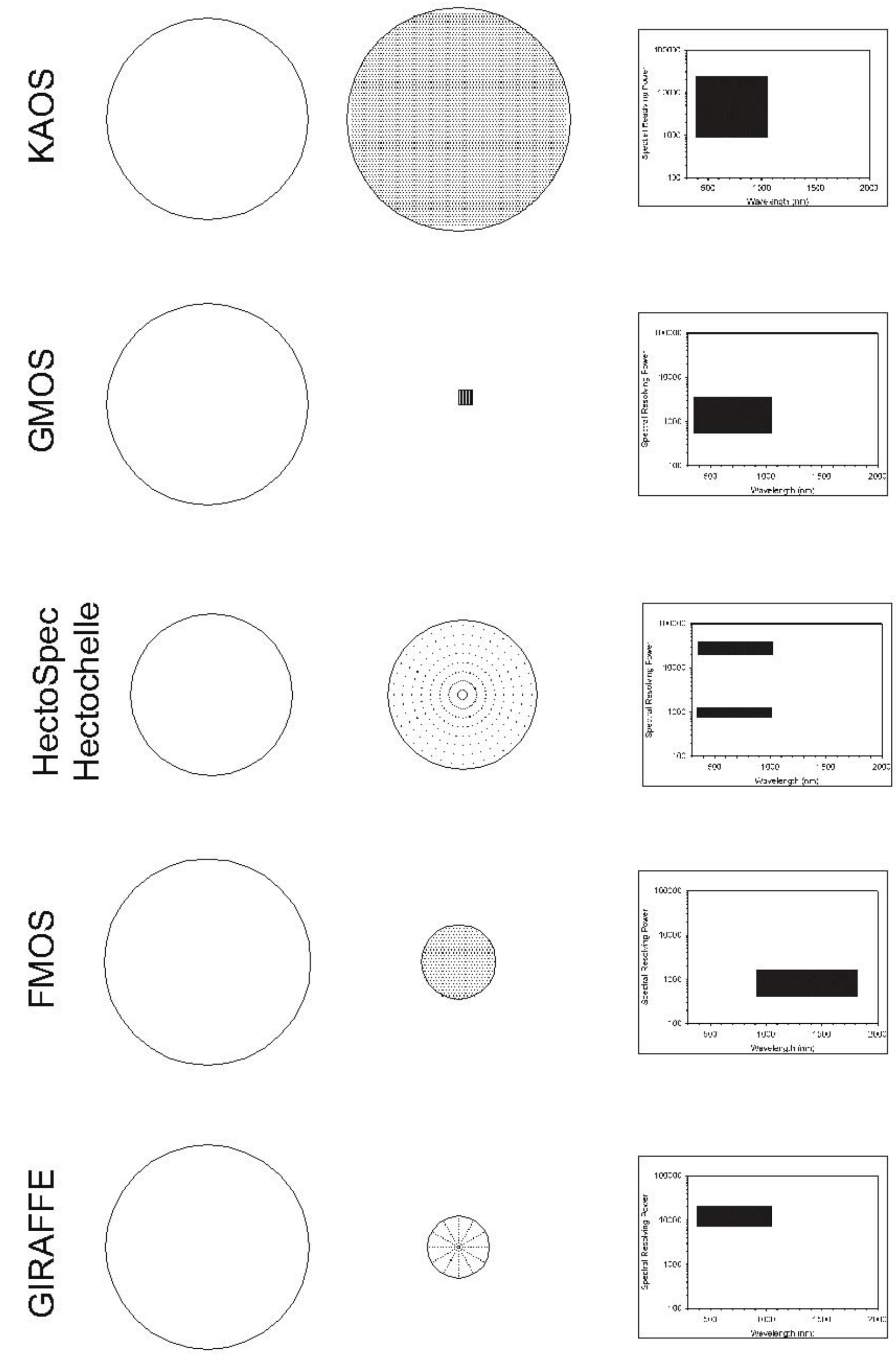
Concluding Comments

The instrument concept for KAOS is realizable with no apparent technical showstoppers. The strengths of this instrument are easily apparent for large survey projects as described in Chapter 2. Figure 3.1.41 shows a comparison of KAOS with other wide-field spectrographs on large telescopes. The relative aperture of the telescope is represented along with the relative field of view and density of objects that are observable by the instrument in one observation. The resolution vs wavelength regime covered by each instrument is represented in the plot to the right of the field of view.

Figure 3.1.42 shows the KAOS field of view and density of targets overlaid on an image of the Andromeda galaxy. The full moon is also shown to provide additional scale.

Figure 3.1.41: Comparison of MOS systems showing telescope aperture, field of view, density of targets, and resolution-wavelength parameter space (R(log): 100-100,000; wavelength: 300-2000 nm). Facilities are shown ranked by telescope aperture.




KAOS
GMOS
HectoSpec
Hectochelle
FMOS
GIRAFFE
Spectral Resolving Power
Wavelength (nm)

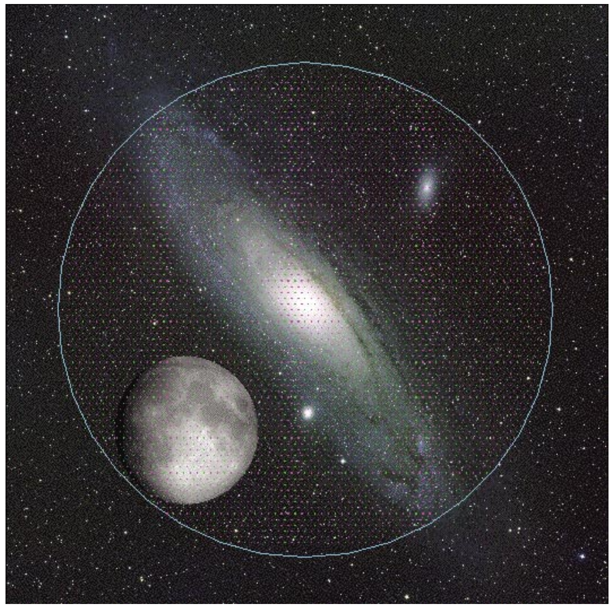


Figure 3.1.42: KAOS field of view in comparison to the Andromeda galaxy and the moon. Photo credit: REU program, N.A. Sharp/NOAO/AURA/NSF

Chapter 2

Echidna-5000: The KAOS Fiber Positioner[20]

1. Introduction

The Echidna positioner is a 400 fiber positioner for the Fiber Multi-Object Spectrograph (FMOS) to be located at the prime focus of the Subaru telescope (Masahiko et al 2002). The positioner represents a radical change from "traditional" pick and place fiber positioners in that each fiber is allocated its own robotic positioner resulting in simultaneous movement of all 400 fibers (Gillingham et al 2002). This feature dramatically reduces configuration times to the extent that no redundancy is required.

The FMOS-Echidna project at the Anglo-Australian Observatory is entering the manufacturing phase. A 2-module prototype is assembled and has been in operation for 12 months. The prototype instrument was recently tested at an environmental facility where, amongst other tests, the positioning performance of the spines was monitored as a function of temperature and humidity. The simulated worst case conditions at Mauna Kea did not affect the positioning performance of the spines.

This report outlines a concept for a 5000 fiber positioner, from now on referred to as "Echidna-5000", as part of the KAOS instrument. The design follows from the outline presented in Brown and Dey 2002. The design philosophy for Echidna-5000 is simple: to try to stick as closely as possible to the spinefilled, linear module system on which the FMOS-Echidna positioner is based. This seriously reduces the development time for the positioner and in addition minimises risk associated with novel development.

A comparison of the required specifications for FMOS-Echidna and Echidna-5000 is presented. Critical areas of difference between the systems are highlighted such as the longer spine and the substantially longer modules required for Echidna-5000. Each area is discussed in turn and accompanied by prototyping results where appropriate. It is concluded Echidna-5000 can be based on the linear module system with only moderate changes to the original FMOS-Echidna design.

1.1. A comparison of FMOS-Echidna and Echidna-5000

A low level comparison of the two systems is shown in Table 1. Given the primary mirrors of the Subaru and Gemini telescopes are virtually identical it is not surprising to see similarity between the systems, scale notwithstanding.

The major differences between the 2 systems, with consequences where appropriate, are as follows:

- The slower input beam of the Echidna-5000 system demands a reduction in the maximum spine tilt. To keep the same spine patrol area this results in an increase in the spine length from 160 mm to 200 mm, measured from the centre of the pivot ball to the fiber tip.

[20]*Contributed by Anna Moore and the AAO Echidna Team*

Table 1. A low level comparison of FMOS-Echidna and Echidna-5000 fiber positioners

	FMOS-Echidna	Echidna-5000
FOV angular diameter	0.5°	1.5°
Plate scale	83 µm	93 µm
FOV physical diameter	147 mm	515 mm
Input beam f/ratio	f/2.1	f/2.4
Radius of curvature	0	4382 mm
Telecentricity	~telecentric	~telecentric
Number of spines in FOV	400 (maximum 424)	4811 (maximum ~4850)
Fiber core size	100 µm	100 µm
Fiber positioning requirement	10µm or less	10µm or less
Field configuration time	10 min or less	5 min or less
Maximum spine tilt	2.6°	2.0°
Spine pitch	*7.2 mm*	*7.0 mm*
Minimum fiber to fiber spacing	*~6”*	*~6”*
Fiber position measurement device	*Focal plane imager*	*Spine tip imager*

- The diameter of the field of view is ~3.4 larger in Echidna-5000 compared to FMOS-Echidna. If one assumes a single linear module design, to compensate for increased deflection the module must be made substantially thicker. This results in a significant reduction in the height of the piezoelectric actuator lying at the heart of each fiber positioning robotic device.

- Each actuator has 4 electrodes requiring connection to the external electronics. In FMOS-Echidna this is achieved, somewhat elegantly, by inserting a multi-layered PCB over the module actuators during assembly and soldering each of the electrodes to the corresponding connection on the PCB. Keeping this design would be extremely advantageous but it must be shown there is available area on the Echidna-5000 PCB to route the required electrode traces while leaving sufficient space between individual traces to prevent voltage breakdown.

- The fiber measurement system for FMOS-Echidna is not suitable for Echidna-5000. The proposed replacement system comprises up to 4 cameras mounted on the prime focus unit, each imaging a section of the field via the primary mirror. This represents a major change in fiber position measurement technique.

Each of the above items was investigated and the results are discussed below. It is concluded all issues associated with the above items can be successfully answered.

1.2. The Echidna-5000 spine (“spine-200”)

The Echidna-5000 spine shown in Figure 3.2.1, from now on referred to as “spine200”, is virtually identical to the original. The 3 changes are as follows:

- A 40 mm increase in the length of the carbon fiber tube, to limit the maximum spine tilt to 2°
- An increase in the length of the Tungsten alloy counterbalance to properly balance the longer spine
- A 1 mm decrease in the diameter of the pivot ball to reduce the distance between the magnets hence aiding module assembly.

A 200 mm long spine was assembled and is shown alongside 6 FMOS-Echidna spines in Figure 3.2.2. Due to time restraint this prototype spine has a counterweight of equal length to the FMOS-Echidna spine so is not balanced about the pivot ball. An unbalanced spine will position perfectly well if it is operated in the vertical orientation, such that the length is parallel to the gravity vector.

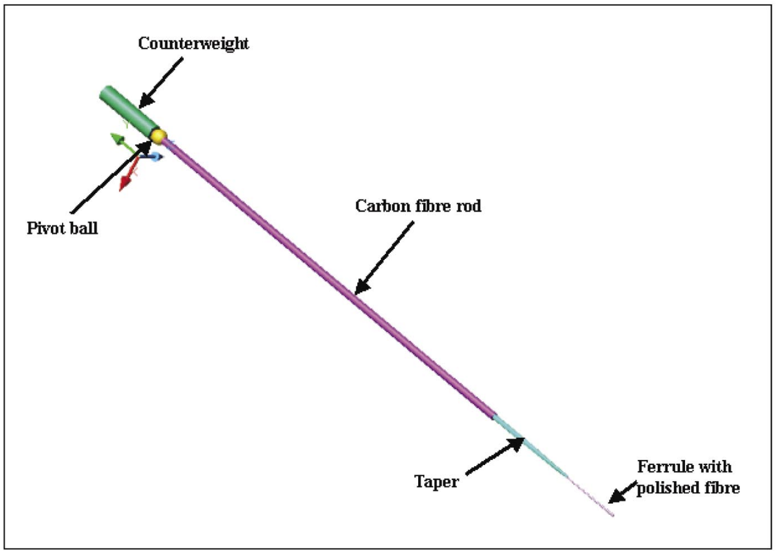


Figure 3.2.1: Spine-200 for the KAOS fiber positioner. The distance from the centre of the pivot ball to the fiber tip is 200 mm, resulting in a reduced maximum angular tilt of the spine compared to the FMOS-Echidna spine.

To investigate the effect on performance of reducing the pivot ball diameter, combined with the extra weight of the carbon fiber tube, the spine-200 was placed on a spare FMOS-Echidna actuator and tested. The actuator was one of the spares shown in Figure 3.2.2. Except for the contact angle between the 3-hemisphere mount and the spine-200 pivot ball increasing from 42° to ~50°, the driving mechanisms are identical. Since the prototype spine-200 is unbalanced the test was performed at 0° zenith angle.

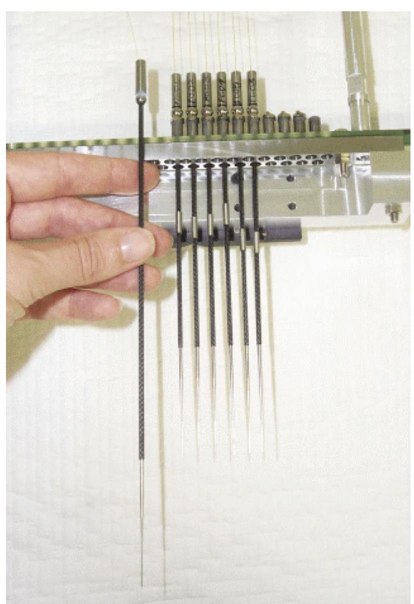


Figure 3.2.2: The prototype spine-200 for KAOS shown alongside 6 FMOS-Echidna spines

The performance of a newly assembled spine can be measured very quickly by performing a calibration test. In this test the spine is moved a known set of steps in each of the 4 possible directions of movement: +u, -u, +v and -v. After each movement the position of the fiber is recorded. The cycle is performed several times for each of the user-defined number of steps.

The calibration tests are identical for both spines and comprise the following:

- The position of the back-illuminated fiber is measured
- The drive signal, a 10-sawtooth waveform of 100V amplitude, is applied across one pair of actuator electrodes so moving the spine 10 steps in the +u direction.
- The position of the back-illuminated fiber is measured
- Procedure a-c is repeated for the -u, +v and -v directions
- The number of steps is reduced by 1 and procedure a-d is repeated until the number of steps =0
- Procedure a-e is performed 7 times in total

The calibration results for spine-200 and an FMOS-Echidna spine, tested on identical mounts and under the same conditions, are shown in Figure 3.2.3 and Figure 3.2.4 respectively. The driving signal has a sawtooth profile of 100 V amplitude. This corresponds to 62.5% of the maximum voltage applied across any pair of electrodes in FMOS-Echidna. The amplitude is relatively insignificant for initial assessment of spine performance.

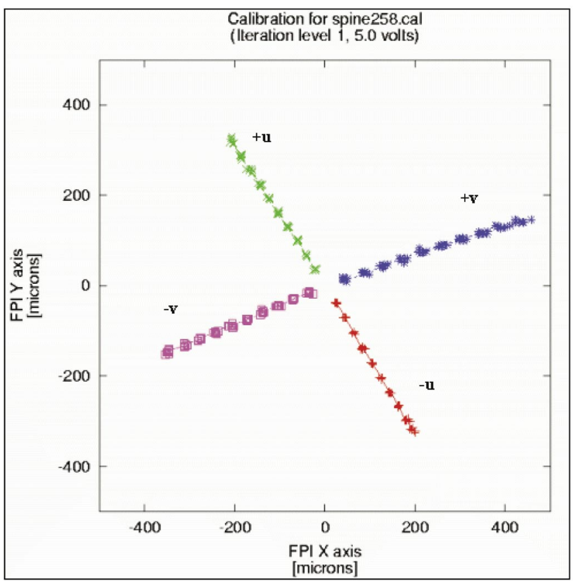


Figure 3.2.3: Calibration results for the spine-200 prototype positioned by and Echidna actuator.

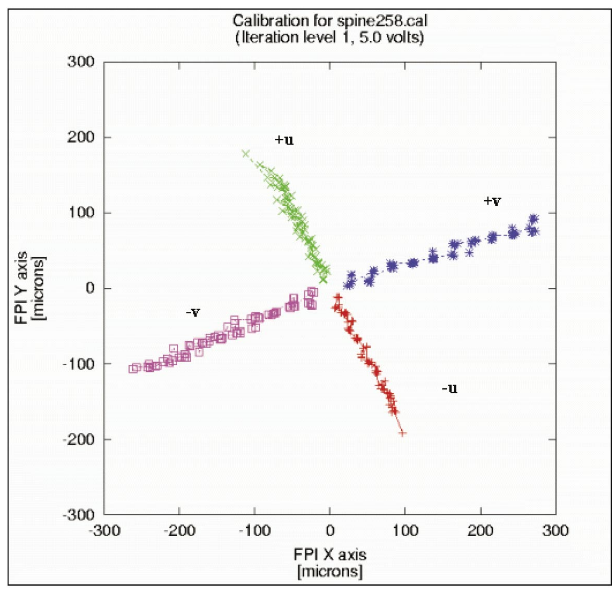


Figure 3.2.4: Calibration results for the FMOS-Echidna prototype spine and actuator.

The calibration results shown in Figure 3.2.4, corresponding to the FMOS-Ecliidna spine, are considered extremely good. The directions of movement are well defined with reasonable step sizes (~20-30 µm) and repeatability along each direction. Such a spine would have a very good positioning capability, assuming the calibration were constant over the spine patrol area.

Though the results are preliminary, the calibration results for spine-200 are superlative. The step size (40 µm) is, as expected for a longer spine, slightly larger than the original spine. The directions are again well defined with excellent repeatability evident in the spine step sizes. Though it is true that spine step size repeatability increases with step size magnitude, it is likely the increase in both the mass of the spine and contact angle between the 5 mm diameter pivot ball and actuator mount have improved the already perfectly adequate calibration results.

Spine deflection The FMOS-Echidna spine is designed with minimal tip deflection in mind. The carbon fiber tube has an excellent stiffness to weight ratio that aids in the minimisation of the spine tip deflection. Being a composite material the tube does not have as uniform a cross-section as a similarly sized extruded metal tube. This results in slightly different deflection values for the 400+ Echidna science and guide spines depending on the orientation of the spine in the instrument. So long as the deflection difference over a lhr observation period is smaller than ~1/10th of a fiber diameter the performance is not degraded. This requirement sets a maximum limit on the spine tip deflection of ~100 µm given current quality assurance procedures for the carbon fiber tube. The measured deflection of the 160 mm FMOS-Echidna spine is ~50 µm.

The theoretical deflection of spine-200 is approximately $(200/160)^3$ or ~2 times the deflection of the FMOS-Echidna spine. This results in a value of ~100 µm, the maximum limit acceptable. It is possible to use a larger diameter fiber tube and reduce the deflection, however, in the interest of minimising module deflection it is advantageous to keep the 2.3 mm diameter tube used in FMOS-Echidna. Further investigation into the maximum allowable deflection accompanied by measured deflections from several prototype spine-200 are required.

Multiple fibers per spine In principal, Echidna-type spines can be fitted with more than one science fiber (of ~100 µm core diameter). The guide spines, for example, have 7 fibers cemented to the spine tip albeit of a smaller core diameter. Though any suggested change to the number of fibers contained within a spine should be prototyped and position tested, the proposition raises no immediate concerns when considering target acquisition.

However, there is a definite concern implementing multiple fibers per spine for the purpose of cross-beam switching. Echidna-type spines are not constrained in rotation about the principal axis parallel to the spine length. In other words, when a spine is driven the tip can rotate in the focal plane. If multiple fibers per spine were suggested as a method for performing cross-beam switching this rotation could render the procedure useless.

Guide fiber rotation testing, performed as part of the FMOS-Echidna final design, has suggested the rotation is slowly varying - a worst case of ~50° after ~100 field configurations was noted. This rotation rate is acceptable for guide spines where only the orientation of the bundle on the sky is required to calculate the telescope adjustments necessary for fine pointing. The rotation rate is unacceptable for 5000 pairs of spines requiring continued alignment to successfully perform cross-beam switching.

1.3. Actuator

The driving mechanism behind the Echidna positioner, referred to here as the actuator, comprises a piezoelectric tube with 4 deposited electrodes on the outside surface. The tube is, at one end, cemented into the module and at the other cemented to a rare earth magnet. The spine pivot ball, a ball bearing located at the centre of mass of the spine, rests on 3 hemispheres cemented to the surface of the magnet. An FMOS-Echidna actuator is shown on the left-hand side of Figure 3.2.5. A full description of the driving mechanism can be found in Moore et al 2002.

The proposed actuator for Echidna-5000 is shown in Figure 3.2.6. The piezoelectric tube is noticeably shorter and a little thinner than its FMOS-Echidna counterpart though almost identical in every other way. The drastic reduction in actuator height produces a corresponding reduction in the velocity of the spine tip. This is a deliberate design change as an appreciable spine tip velocity during configuration is not essential for Echidna-5000. This results from a major difference in the proposed method of determining the fiber positions in the focal plane. Any reduction in the overall actuator height can be converted to module thickness. With an actuator height of 10 mm, combined with a maximum required spine tilt of 2°, one can use modules of ~25 mm thickness. It is then possible to span the KAOS 515 mm

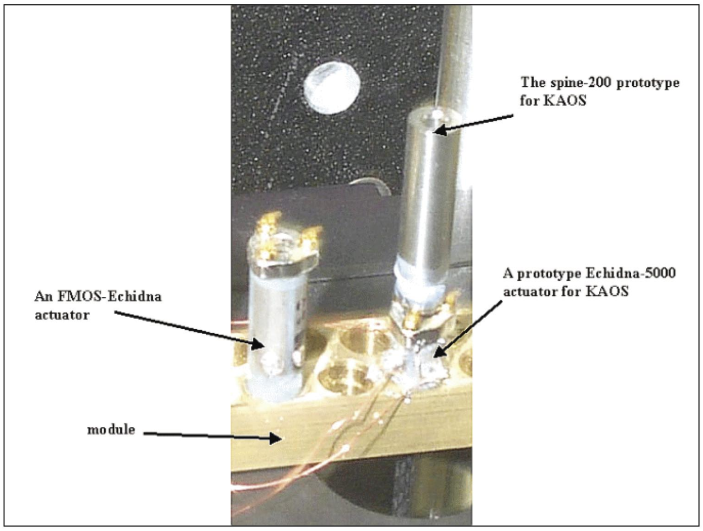


Figure 3.2.5: A comparison of the proposed actuator for Echidna-5000 and the larger version developed for FMOS-Echidna.

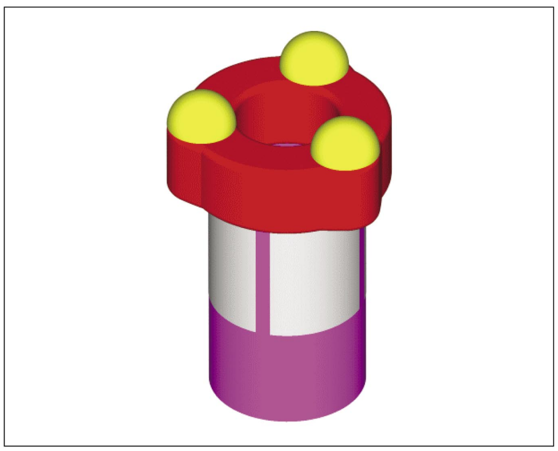


Figure 3.2.6: A rendered drawing of the KAOS actuator, the spine driving mechanism.

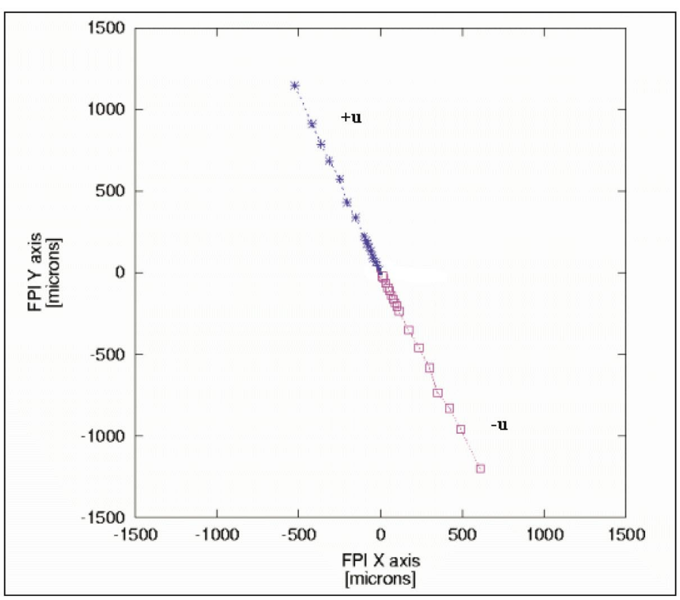


Figure 3.2.7: Spine-200 calibration performance along the X-axis with shortened actuator.

field of view with FMOS-Echidna style continuous linear modules without module deflection degrading the optical performance over a ~1hr observation.

Actuator performance A makeshift prototype actuator was made from a 5.5 mm outside diameter original and is shown on the right-hand side in Figure 3.2.5. The newly assembled spine-200 was positioned on the actuator and tested using a sawtooth voltage of amplitude 160V and 70 Hz trigger frequency. The voltage and frequency correspond to the high voltage level in FMOS-Echidna. Though the step size was small, the spine moved repeatably in all four directions with calibration results along the u-axis shown in Figure 3.2.7.

The maximum travel along the u-axis is approximately 1 mm, corresponding to 100 sawtooth steps. This gives an average step size of ~10 µm compared to ~40 µm for the same spine on the longer FMOS-Echidna actuator driven by the same signal. For Echidna-5000 this step size would be used for the fine positioning.

Without substantial change to the FMOS-Echidna electronics design one can increase the amplifier output voltage from 160V to 220V peak to peak. A more substantial increase can be obtained by modifying the drive signal so the voltage across the piezoelectric actuator slowly decreases to -220V before rapidly increasing to 220V (whereas the current drive signal starts from 0V). The cumulative effect would be to increase the step size to ~28 µm, a more than satisfactory value given this is translated to a spine tip velocity of ~2 mm/s using a 70 Hz trigger frequency.

Further testing of spine-200 and the shortened actuator will be performed in the near future.

1.4. Module

An Echidna-5000 module is shown in Figure 3.2.8. The module is 24mm wide, approximately 600 mm long and 30 mm thick. A suggested material is hardened steel, the same as an FMOS-Echidna module, though the use of carbon fiber or some other high stiffness to weight ratio material is not ruled out. Both top and bottom surfaces of the module are slightly curved (R~4200 mm) along the module length. The module cross-section can be rectilinear as the curvature of the field over such a small distance produces a negligible effect. There are 4 rows of 71 stepped holes, each bored perpendicularly to the module surface. Into each hole is cemented a piezoelectric actuator. For clarity the module is shown without the printed circuit board (PCB) present. Two spines are shown mounted on actuators in the module along with both fiducials. A section through the module is shown in Figure 3.2.9.

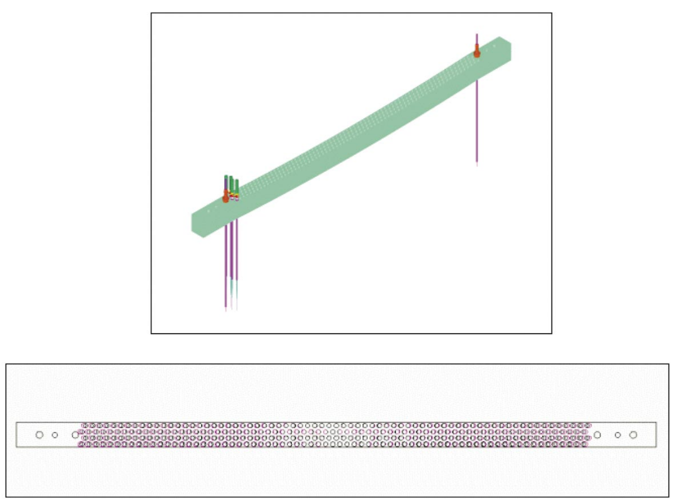


Figure 3.2.8: An Echidna-5000 module with 71 spines/row. The above drawing shows the module with 2 spines and both fiducials in place. The lower drawing is a plan view of the module.

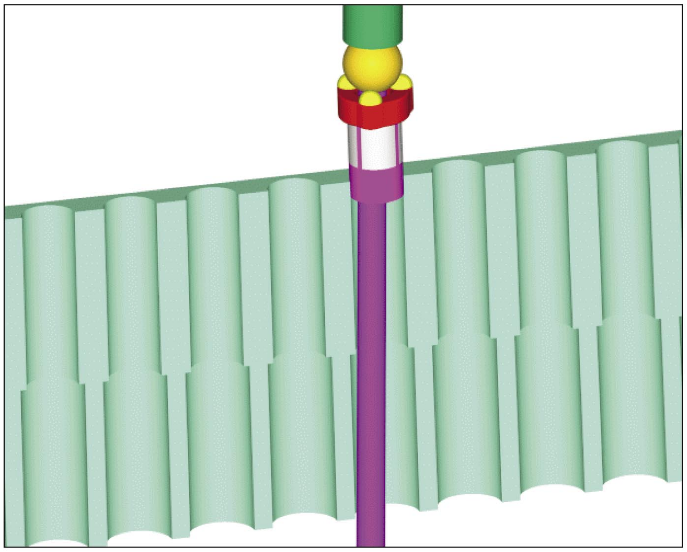


Figure 3.2.9: A spine positioned in a sectioned drawing of the module. The spine is tilted at the maximum angle of 2deg.

To dramatically reduce the number of non-allocatable spines, two versions of the module populate the field of view. This is discussed in section 2.

Fiducials As part of the fiber position measurement system every module has two fixed tubes, or fiducials, located at each end. Each fiducial has a fiber cemented at the tip. During a field configuration the system measures the position of every backlit spine relative to the fixed fiducials until the entire field is configured. So long as the locations of the fiducials on the sky are known one can position the configured spines onto the target objects. For this reason the fiducials are much stiffer than the spines and undergo a substantially reduced deflection.

Guide spines For every FMOS-Echidna module the 2 outermost positions are reserved for guide spines. Each guide spine has a 7 fiber bundle, with each fiber identical and smaller than the science fiber for sampling purposes. Out of the 12 modules populating the field in FMOS-Echidna the central 7 have active guide spines arranged in 2 rows along the edge. Guide spines in outer modules are located far enough outside the field for imaging quality to be unacceptable. This system can be directly implemented in Echidna-5000, resulting in *4* guide spines per module. In total *34* guide spines can access the field of view.

1.5. PCB

A multi-layered PCB is positioned above the module and serves to route all electrical traces to the module actuators. Each layer of the PCB has an array of holes large enough to allow insertion of a piezoelectric tube. The available space between any two holes, combined with the number of avenues between rows of spines, governs the number of traces carried by the layer. The maximum number of layers is ~32 with current manufacturing techniques.

There are 284 actuators, each with 4 electrodes, in an Echidna-5000 module. Two of the 4 electrodes are connected to a common ground. Given the location of identical switchboards at either end of the module the maximum distance a single trace has to travel is half a module length. In total, 284 isolated traces must be routed from either end of the PCB, not including ground connections.

This can be achieved, in theory, by implementing 4 rows per module combined with reducing the actuator diameter to 4.4 mm. With a material gap of 2.5 mm between every hole in the PCB one can route 4 traces along each of the 3 avenues available, supporting 12 traces per layer. Therefore a total of 284 traces can be routed using 24 layers. A total of 2-3 layers should be added for providing ground connections, still keeping the total number of layers below 32.

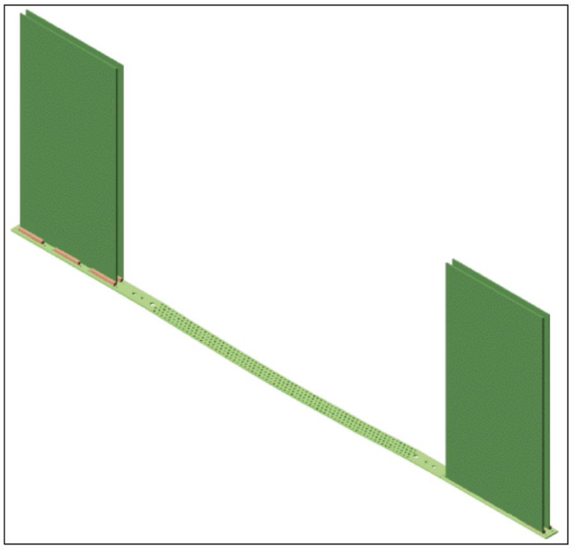


Figure 3.2.10: The module printed circuit board (PCB).

Figure 3.2.10 shows a rendered drawing of the Echidna-5000 PCB with switchboards in place. The proposed length of the PCB shown in Figure 3.2.10 is ~1m. This is certainly outside the realms of local PCB manufacture and would warrant further investigation. It is possible the PCB can be divided into two sections if overall length causes concern.

2. Echidna-5000 positioner assembly

A concept design for the Echidna-5000 is shown in Figure 3.2.11. The assembly consists of 20 modules positioned side by side, each module supported by 2 guide rails. Each guide rail can be stepped to achieve the required curvature of the spines in the direction perpendicular to the module length. The guide rails form part of the main structural plate shown in Figure 3.2.11. The plate supports the modules, switchboards and a 4-legged frame surrounding the

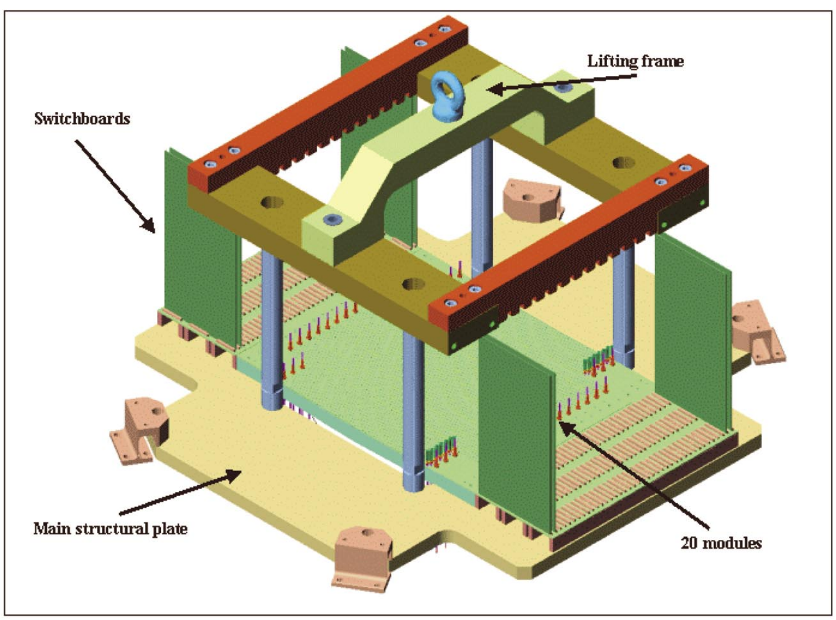


Figure 3.2.11: The Echidna-5000 assembly, partially filled with spines.

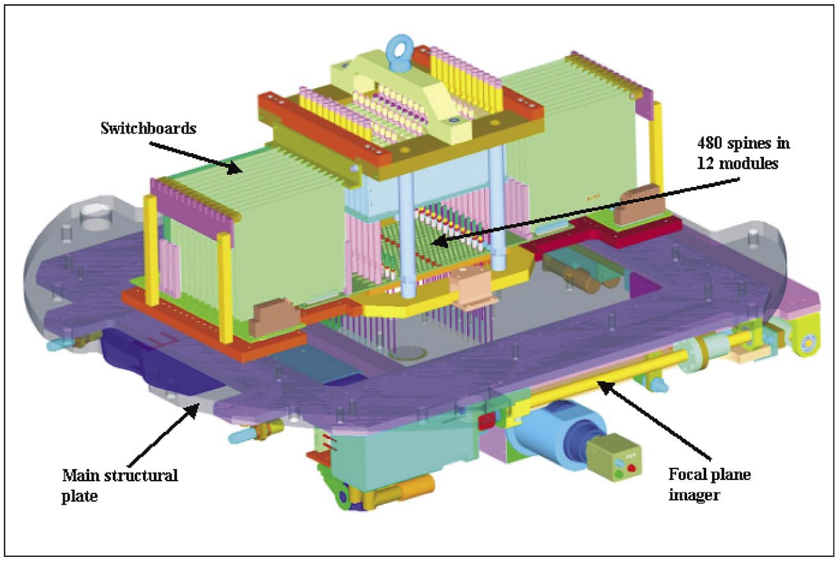


Figure 3.2.12: A rendered drawing of the FMOS-Echidna assembly with focal plane imager located underneath.

spines. The frame provides a convenient lifting point for removal of Echidna-5000 from the prime focus unit, as well as supporting the switchboards and module frames (not shown in Figure 3.2.11).

There are 2 module designs utilised. The 14 central modules, described in section 1.4, have 71 spines per module and a fiducial separation of ~510 mm. The 3 outer modules on each side of the field have a reduced number of 55 spines per row with a fiducial separation of ~400 mm. This reduces the number of non-allocatable spines from ~1020 to ~530 while keeping significant similarity between the modules to aid in assembly and allow module interchangeability.

The design is strongly influenced by the FMOS-Echidna positioner assembly shown in Figure 3.2.12. It is envisaged the method of fiber routing from the module shown in Figure 3.2.12, where each module has a frame providing mounting points for every spine-exiting fiber, will be directly scaled and applied to Echidna-5000.

The available space for Echidna-5000 in the prime focus unit is a cylindrical drum of diameter 1.3m and height 0.8m. A section of this drum is shown with the assembly in Figure 3.2.13. The boxes located alongside the positioner represent available space for electronics and the guide system. There is space available above the assembly for fiber routing purposes and below for location of the wavefront sensor and acquisition units.

It should be noted that focus adjustment of Echidna-5000 must be provided.

Division of the field The spine coverage across the field of view is shown Figure 3.2.14. With the two module design implemented the number of spines lying outside the field is ~484 or 10% of the total number (5296). This compares with a loss of 18% of the total number (5680) using identical modules of 71 spines per row, a saving of 380 spines.

The locations of the module fiducials are shown relative to the field in Figure 3.2.14. The number of fiducials located outside the field causes concern in that the corresponding image quality viewed by the fiber position measurement system may be inadequate. This situation does not arise in FMOS-Echidna given the measurement system directly views the fiber output. Further modeling of the image quality, combined with simulations of the measurement system, are required. A total of 34 guide spines, 17 along each side, can access the field.

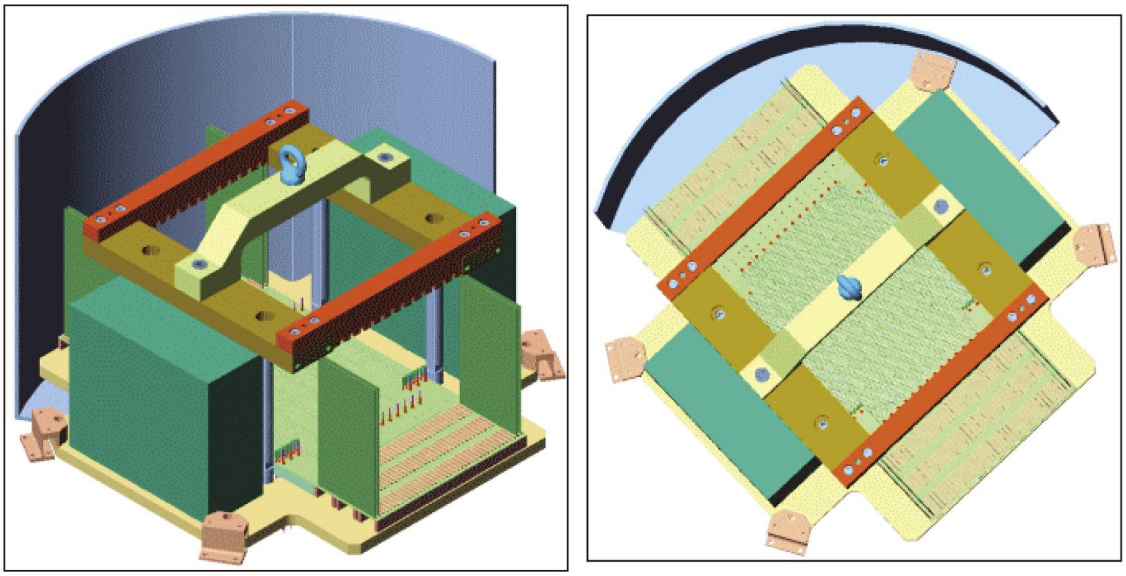


Figure 3.2.13: The Echidna-5000 assembly with a section of the cylin≠drical housing evident. The right-hand side drawing is a plan view of the assembly.

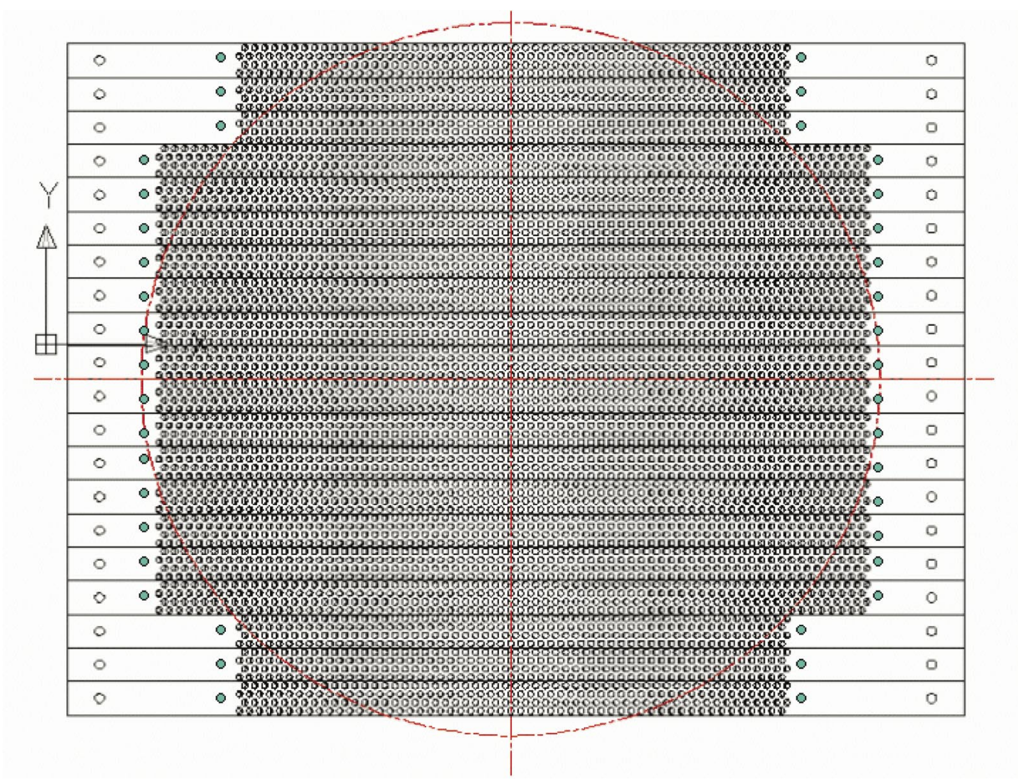


Figure 3.2.14: The KAOS field of view (circle) populated with 20 modules. The 2 outer rows of holes are for guide pin insertion; the central holes are for actuator insertion; the solid circles represent fiducial location.

Mass estimate The mass of the FMOS-Echidna positioner assembly is a mere 37kg. Though it is premature to give an estimate for the mass of the Echidna-5000 at this stage a direct scaling of a factor of 9 in the overall volume increases the mass to ~330kg. This is likely to be small compared to the mass of the corrector and is, in itself, an overestimate (the mass of the spines and actuators, for example, increase with r^2 rather than r^3).

2.1. Determining fiber positions: the "one-shot" STI

It is discussed in (Brown and Dey 2002) that direct scaling of the focal plane imager, used for fiber position measurement in FMOS-Echidna, is not viable for Echidna-5000. An alternative imaging method, referred to as the spine tip imager (STI), is introduced whereby a small telescope placed, for example, through the center of the primary mirror could be used to determine the positions of all 5000 backlit spines with very little loss in time. It is assumed that the STI will record images of fixed fiducial spines and that all spine measurements are relative to these fiducials. This avoids the need for particularly high stability of the mounting of the STI camera.

An improvement on the suggested STI design is to mount a telescope on the outside of the prime focus unit viewing the reflection of the backlit fibers from the primary mirror. This has the advantage of providing non-aberrated images across the field as the optical path includes reflection from the primary mirror.

In its simplest form, one can assume a CCD 2k x 2k x 15 µm (30.7 mm on a side) fed by a suitable lens. A reduction of about 17:1 is needed to view the entire 515 mm diameter field, resulting in a detector scale of 2.64"/pixel. With a required centroiding accuracy of ~0.1" (around 1/30th pixel), the STI can be deliberately set out of focus to spread the images across a few pixels. Centroiding accuracy of 1/15th-1/30th pixel is obtained with current AAO multi-fiber instrumentation with suitably sampled images. A magnification of 17:1 requires a telescope of ~0.8m focal length.

If 1/30th pixel centroiding accuracy is in practice too difficult to obtain in this application, it is possible to mount several telescopes on the outside of the prime focus unit, each imaging a fraction of the field. If one implements 4 telescopes, for example, each would image a quarter or ~0.75° section of the field. The detector scale, calculated using the above parameters, would reduce to 1.32"/pixel. The centroiding accuracy would in turn reduce to ~1/13th of a pixel. The focal length of the telescope would increase to 1.6 m.

A possible objection to the STI is that spines placed very close together would not be directly resolved. However, in the Echidna positioner, we have developed a method to resolve ambiguity amongst neighbouring spines by time coding their back illumination. This trick can be used to distinguish such spines.

Analysis and prototyping of an STI has recently begun at the Anglo-Australian Observatory as part of the MOMFOS project. This instrument, for the prime focus of the Giant Segmented Mirror Telescope (GSMT), includes an Echidnastyle positioner.

2.2. Configuration time

The configuration time is dependent upon the factors listed in Table 2. Using information from FMOS-Echidna and spine-200 prototyping on the shortened actuator, one can roughly assign the values shown in Table 2. The estimate assumes the worst case scenario of requiring an integrating camera rather than a CMOS image sensor or video camera (if a fast frame rate camera is implemented over an integrating variety the configuration time will be significantly reduced).

The corresponding configuration time is ~300 s or ~5 minutes. It should be strongly noted, however, that prototyping of the STI and drive electronics is required before confidence is placed in this value.

Table 2. Factors affecting the configuration time of Echidna-5000.

Spine tip velocity	2mm/s
Number of attempts to achieve target position	5
Maximum number of spines moved simultaneously	2500
Number of encoding bits to distinguish neighbouring spines	3
Detector integration time	15
Frame download time	1 s
Centroid processing time	3 s

2.3. Manufacturing implications

Current methods for achieving the required high precision assembly in FMOS-Echidna can be directly implemented in Echidna-5000. The jigs used in the actuator cementing process, for example, can be scaled with only medium level modifications. The spine assembly is identical, though ways of assembling components en masse would be required.

The latter point hints at probably the most significant design change of Echdna-5000 compared to FMOS-Echidna. While a 400 spine assembly can be manufactured, to a large extent, by hand this is not applicable to a 5000+ spine assembly. For example, there are 4 electrodes per actuator requiring electrical connection to a PCB in Echidna-5000. This equates to 20,000 solder joints without spare modules! A mandatory part of the preliminary design would be to develop methods of mass component assembly.

2.4. Conclusions

A concept for a 5000 fiber positioner assembly for the KAOS instrument is presented. It is shown the positioner assembly can be directly modelled on the FMOS-Echidna instrument due for commissioning on the Subaru telescope in mid-2005. This reduces the risk and development time associated with novel instrumentation.

An alternative fiber position measurement system is proposed, more suited for configuration of 5000 spines. Initial estimates of field configuration times are ~5 min. The system is currently under development at the Anglo-Australian Observatory as part of the MOMFOS design study for the 30-m Giant segmented mirror telescope (GSMT).

References

Kimura, M. et al, Proc. SPIE, 4008, (in press)
Gillingham, P. G. et al, Proc. SPIE, 4008, (in press)
Moore, A.M., Gillingham, P.R., Griesbach, J., Akiyama, M., 2002, Proc. SPIE, 4841, (in press)
Brown, M., J., I., Dey, A., ASP, 280, 2002

Chapter 3

Nod and Shuffle[21]

Nod & Shuffle (N&S hereafter) is a technique that in principle allows the best possible sky subtraction for multi-object spectroscopy. It is a key capability for KAOS, if it is to fulfil the key science requirement for spectroscopy at *R*~24-25 where the object signal is ~1% of the sky signal transmitted down the fibre. N&S involves nodding (beam-switching) the telescope quickly between objects and nearby sky, while simultaneously shuffling object and sky spectra to separate regions of the CCD. The advantages are that object and sky are both observed through the same slits/fibres and CCD pixels, while the rapid beam-switching timescales (~1 minute) allow sky variations to be accounted for; see Glazebrook & Bland-Hawthorn (2001, PASP, 113, 197) for further details. N&S was initially developed for use with *multi-slit* instruments on the AAT (LDSS++/Taurus++) and has recently been implemented with spectacular success on Gemini with GMOS (Abraham et al. NOAO newsletter, see also Figure 3.3.1).

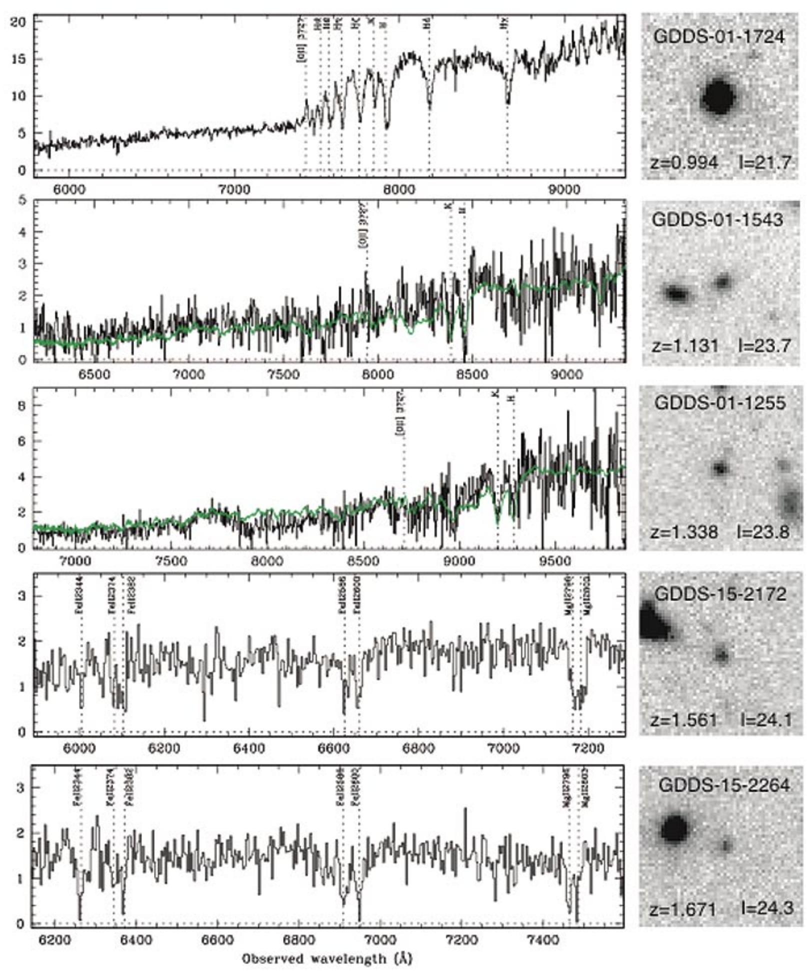


Figure 3.3.1: Montage of 100ks Gemini GMOS Nod & Shuffled spectra from the GDDS. Ground-based I-band images (in ~0.8" seeing) are shown at right. Objects shown span a redshift range of 0.994<z<1.671 and a magnitude range of 21.7<I<24.3 mag (Reproduced from Abraham et al.)

[21]*Contributed by Terry Bridges, Karl Glazebrook and Brian Boyle*

For the KAOS case, it is important to establish the efficacy of N&S for a *multi-fibre* system. Early results using the 2dF system on the AAT have shown that N&S works in principle (Glazebrook 1999, AAO Newsletter #90). However, it was only at the end of 2002 that long exposures of faint objects in dark conditions were obtained with 2dF N&S.

Two N&S modes were tested with 2dF. In **band-shuffling** mode, alternate retractor blocks (each retractor block contains 10 fibres, and is ~50 CCD pixels) are disabled and the parked fibre positions are blocked with a wooden mask. This allows shuffling by ~50 pixels (10 fibres) into these created gaps, with 100 fibres available for shuffling. In **micro-shuffling** mode, every second fibre is disabled, again by parking those fibres and masking the parked positions. This again leaves 100 fibres on the CCD, with 10-pixel spacing between the spectra.

N&S data were obtained on two nights, 28 August 2002, and 6 November 2002. In August both micro and band-shuffle modes were tested at low spectral resolution (4Å /pix) with wide spectral coverage 4700-9600Å. Fibres were placed at the positions of 40 deep radio sources and 10 sky positions. Three band-shuffled exposures of 600 sec total (ie. 300 sec on objects), with a nod time of 60 sec were obtained. This was followed by three micro-shuffled exposures of 420 sec total with a 5 pixel shift.

The observed RMS of the sky-subtracted sky-dedicated fibres were compared with their expected RMS (based on pure Poisson statistics) over the range 6000-8000Å. The data demonstrated excellent sky-subtraction performance, with Poisson-limited sky-subtraction with for both band-shuffled and micro-shuffled data.

The ability to recover faint object spectra using N&S was tested on observations of faint candidates the Subaru-XMM deep field in November. Data were taken in micro-shuffling mode; 95 fibres were placed on objects and 5 allocated to sky. A total of 12x1800 sec exposures were obtained, or 3 hours of on-source time. The objects were observed at low resolution (4Å/pixel) with a central wavelength of 6730Å. Unfortunately, conditions were poor during the night, with seeing ranging from 2-3" and thin cirrus present for much of the time. The results, however, are excellent. Again, qualitatively, N&S gives excellent sky subtraction, with much smaller residuals than normal 2dF reduction. Quantitatively, in almost every case (i.e. for different sky fibres, and for different methods of N&S data reduction), the sky-subtraction is Poisson-limited. This is true for both the individual 1800 sec exposures, and for the combined 3-hour dataset. Figure 3.3.2 shows some example spectra taken from the Subaru data, with a comparison of N&S reduced spectra and conventionally-reduced spectra. It is clear from this that systematic sky residual is removed by N&S, vastly improving the detection of real features, and hence the identification of faint objects.

Conclusions

N&S with the 2dF multi-fibre spectrograph yields sky-subtraction that is Poisson-limited, for both band- and micro-shuffled data. The quality of the sky-subtracted spectra is excellent, even in the far-red where there are many strong night sky lines that vary on short timescales. The removal of systematic sky residuals vastly enhances the identification of sources at or, indeed beyond, the conventionally accepted limit for fibre spectrographs.

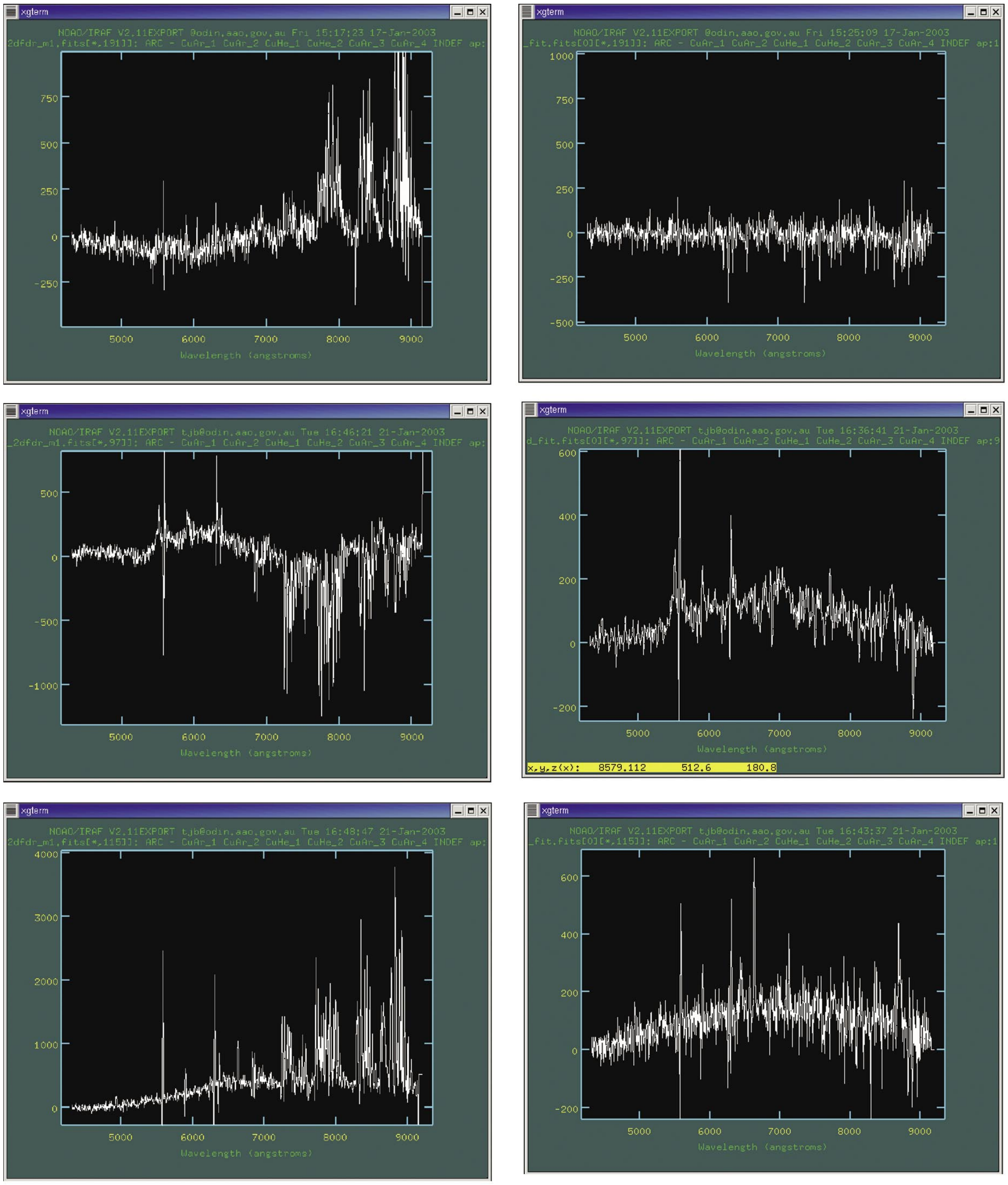


Figure 3.3.2: Representative spectra taken from the 2dF Subaru-XMM deep field observations in November 2002. Spectra on the left-hand side are for conventional (ie. non-N&S) reductions, while right-hand side are N&S reduced. (1a, top): Sky spectrum; (1b, middle): R=21 QSO at z~3.5; (1c, bottom): R=21.25 emission-line galaxy at z~0.3

Chapter 4

Software Architecture and System Engineering[22]

It is extremely important to understand the systems engineering aspects of the data systems. This is where the final results all come together. The system must formally define:

1. a sceince requirements document,
2. a detailed data model,
3. a detailed pipeline workflow,
4. a coherent calibration procedure,
5. a full data management system.

The definition of these components must be carried out with thinking about the possible science use cases beforehand. As an example, think of the impact of fiber collisions: we cannot place two fibers closer than a certain distance. For objects closer than this distance we must record, which object of the pair has been chosen, and why: the system as a whole has to capture all details of the target selection and fiber assignment procedure.

The data system must also contain all input data that impacted the selection of targets, including all photometric and astrometric catalogs, remarks and notes about serendipitous targets, etc.

Finally, it must be able to automatically track the provenance of every bit of data that comes out of the system. This relates to the formal definition of the data model, to the formal definition of the pipeline workflow, and in particular how these interact with the data management system as a whole. An example of this is that detailed logs must be written at every step of the processing, which must be stored in a database. It must be easy to get a high-level view of the log status, like have any errors occurred? If errors happen, they must trigger appropriate alarms.

Even several years after data was taken we must be able to go back to the logs and recover the smallest details of the data acquisition and the reduction process.

Links to the Virtual Observatory

While many of the standards for the Virtual Observatory are under development today, there are several aspects that one can already take into consideration. The most important aspect is that we will need to store and generate as detailed metadata as possible, as the system is developed and ran. Having rich metadata is crucial for an automated understanding of the content and the quality of the scientific end products, and in the end will also affect how the actual workflow systems are designed, implemented and ran.

We need to set up email exploders, which will contain all the project communication, and will become part of the eventual project archive. All design documents will be transformed to electronic form and become part of the project metadata.

[22] *Contributed by Alex Szalay & Michael Brown*

The data archive will adhere to the emerging VO standards. The templates for archives that will come out of the VO will make this job rather easy, if the rest of the system is well architected.

KAOS & Gemini: Data Reduction Pipeline

KAOS datasets of 10^5—10^6 spectra require a robust data reduction pipeline with minimal human intervention and VO-compliance. This requires a well-defined data reduction procedure and a data model for the Gemini-KAOS system. This includes information describing the target and fiber selection, fiber and guider positions and details of the wavelength and flux calibration. Even without VO-compliance, such information is required for science with large spectroscopic datasets if the uncertainties are to be dominated by shot-noise.

Data reduction pipelines capable of handling 10^5—10^6 spectra have been in regular operation for the 2dF spectrograph and SDSS. The GMOS and AAΩ (AAT) instruments will be using nod & shuffle sky subtraction for normal science operations about 4 years before KAOS first-light. It will therefore be possible to minimize the KAOS pipeline development costs by drawing upon the experience, algorithms and software of these instruments and surveys. The pipeline development costs and time should therefore be significantly less than those of SDSS (2 FTEs for 4 years), where much of the software was developed from scratch. There will also be a small pool of astronomers/programmers with experience in developing large multiobject spectroscopic pipelines and the hiring of experienced staff should further shorten the development time of the software.

A possible implementation of the KAOS data reduction pipeline, based on the experience of 2dF and SDSS, involves two software packages. The 2D package performs basic reductions including flat-fielding, wavelength calibrations and combining multiple exposures of targets. The 1D package performs a preliminary analysis including identification of object types and velocity measurements along with uncertainties. Neither package requires data processing that is unique to KAOS and, if necessary, it will be possible to modify an existing package to process KAOS data.

Ideally, the software packages should be portable and have open source code. Portability is required so astronomers can download and run the software at their home institutions. In addition, it is not possible to predict the dominant computer architecture within astronomy for the lifetime of the KAOS instrument. Open source software allows peer review of the KAOS data reduction pipeline. Astronomers would also be able to modify the software for particular science programs and implement or suggest upgrades to the data reduction pipeline. While this model of software development has not been widely used in astronomy, examples, including SExtractor (source detection and photometry), have been very successful. Coding the data reduction packages in a commonly used non-proprietary computer language aids portability and the aims of open source code.