DYNAMICS AND STELLAR CONTENT OF THE GIANT SOUTHERN STREAM IN M31. I.
KECK\(^1\) SPECTROSCOPY OF RED GIANT STARS

PURAGRA GUHATHAKURTA,\(^2\) R. MICHAEL RICH,\(^3\) DAVID B. REITZEL,\(^3\) MICHAEL C. COOPER,\(^4\)
KAROLINE M. GILBERT,\(^2\) STEVEN R. MAJEWSKI,\(^5\) JAMES C. OSTHEIMER,\(^5\) MARLA C. GEHA,\(^6,7\)
KATHRYN V. JOHNSTON,\(^8\) AND RICHARD J. PATTERSON\(^5\)

Received 2004 June 7; accepted 2005 November 9

ABSTRACT

This paper presents the first results from a large spectroscopic survey of red giant branch (RGB) stars in M31 using DEIMOS on the Keck 10 m telescope. A photometric prescreening method, based on the intermediate-width DDO51 band centered on the Mg\(^b\)/MgH absorption feature, was used to select spectroscopic targets. RGB candidates were targeted in a small section of M31’s giant southern tidal stream at a projected distance of 31 kpc from the galaxy’s center. We isolate a clean sample of 68 RGB stars by removing contaminants (foreground Milky Way dwarf stars and background galaxies) using a combination of spectroscopic, imaging, and photometric methods: radial velocity and the surface gravity—sensitive Na\(^i\) doublet are particularly useful in this regard. About 65% of the M31 stars are found to be members of the giant southern stream, while the rest appear to be members of the general spheroid population. The mean (solar) radial velocity of the stream in our field is $-458 \text{ km s}^{-1}$, blueshifted by $-158 \text{ km s}^{-1}$ relative to M31’s systemic velocity, in good agreement with recent velocity measurements at other points along the stream. The intrinsic velocity dispersion of the stream is found to be $15^{+8}_{-8} \text{ km s}^{-1}$ (90% confidence limit). A companion paper by Font and coworkers discusses possible orbits, implications of the coldness of the stream, and properties of the progenitor satellite galaxy. The kinematics, and possibly the metallicity distribution, of the general spheroid (i.e., nonstream) population in this region of M31 indicate that it is significantly different from samples drawn from other parts of the M31 spheroid; this is probably an indication of substructure in the bulge and halo. The stream appears to have a higher mean metallicity than the general spheroid, $[\text{Fe/H}] \sim -0.54$ versus $-0.74$, and a smaller metallicity spread. The relatively high metallicity of the stream implies that its progenitor must have been a luminous dwarf galaxy. The Ca \(II\) triplet line strengths of the M31 RGB stars are generally consistent with photometric estimates of their metallicity (derived by fitting RGB fiducials in the color-magnitude diagram). There is indirect evidence of a population of intermediate-age stars in the stream.

Key words: galaxies: abundances — galaxies: individual (M31) — galaxies: kinematics and dynamics

1. INTRODUCTION

The growth of galactic halos through the accretion of smaller stellar subsystems has been the subject of many studies over the last few decades (Searle & Zinn 1978; White & Rees 1978). Numerical simulations and semianalytical modeling of the accretion process have reached new levels of detail in recent years (e.g., Johnston et al. 1996; Johnston 1998; Helmi & White 1999; Helmi & de Zeeuw 2000; Bullock et al. 2001). The discovery of the Magellanic Stream (Mathewson et al. 1974) provided early observational evidence of an ongoing accretion/merger event in the Galaxy involving the Large and Small Magellanic Clouds. The fact that the Magellanic Stream is seen only in neutral hydrogen and not stars has led to some debate over whether ram pressure stripping or tidal forces are at play (Moore & Davis 1994; Putman et al. 1999; Maddison et al. 2002). The best example of an ongoing accretion event in the Milky Way is the Sagittarius dwarf satellite galaxy (Ibata et al. 1994) with its associated tidal debris (Majewski et al. 2003; Newberg et al. 2003). More recently, Sloan Digital Sky Survey (SDSS) stellar density maps have revealed that the low-luminosity, remote star cluster Palomar 5 is undergoing tidal disruption (Odenkirchen et al. 2001; Rockosi et al. 2002). SDSS and Two Micron All Sky Survey data have led to the discovery and characterization of the Monoceros Stream, an arclike structure at low Galactic latitude that is probably the result of an encounter with a dwarf galaxy (Yanny et al. 2003; Rocha-Pinto et al. 2003). The Monoceros Stream and other Milky Way structures like it are only just being identified; their large angular extent requires the use of wide-field surveys. Moreover, follow-up studies of these structures can prove to be difficult from our vantage point within the Galaxy’s disk.

In contrast to the Milky Way, its neighbor, the Andromeda spiral galaxy (M31), offers certain advantages for halo studies: its disk is highly inclined, we have a global external perspective of the galaxy, and yet it is close enough to allow us to characterize the properties of individual stars in detail. Ibata et al. (2001) and Ferguson et al. (2002) present star-count maps covering a
large area around M31 and find a giant stream extending to its south (hereafter referred to as the giant southern stream) along with several other signs of disturbance in the spheroid and outer disk. While the giant southern stream appears to be tidal debris from a merger event, the origin/nature of the other features is not clear. An investigation of M31’s innermost satellites by Choi et al. (2002) confirms the presence of tidally distorted outer isophotes in NGC 205 and reveals ongoing stripping in M32, with the amount of mass lost estimated to be of the same order as that seen in the giant southern stream. However, any association between the stream and M32 (or any other satellite) based on the above studies is largely circumstantial at this point.

Spectroscopy of large samples of individual stars in M31 is the only secure approach to establishing connections between specific streams/features and satellites and to global mapping of the kinematics and chemical composition of the halo. Reitzel & Guhathakurta (2002, hereafter RG02) carried out spectroscopy of ~100 candidate red giant branch (RGB) stars in an outer minor-axis spheroid field in M31 using the Keck 10 m telescope and Low Resolution Imaging Spectrograph (LRIS; Oke et al. 1995). Another 150 spectra of M31 RGB candidates in inner minor-axis and outer disk fields were analyzed by Guhathakurta & Reitzel (2002) and Reitzel et al. (2004). Recent papers by Ibata et al. (2004, 2005), McConnachie et al. (2004), and Kalirai et al. (2006a, 2006b) present radial velocities for several hundreds of RGB stars in M31. These spectroscopic studies taken together have shed light on a variety of topics, including the dynamics and metallicity distribution of the M31 bulge and halo, the relation between the spheroid and outer disk, evidence of faint debris trails in the spheroid, and the orbit of the giant southern stream. There are important questions that remain unanswered, however; e.g., what fraction of the spheroid is composed of identifiable streams?

In fall 2002 we started a spectroscopic survey of the M31 halo with the new Deep Imaging Multi-Object Spectrograph (DEIMOS; Faber et al. 2003) on Keck. This survey has yielded spectra for several hundred RGB candidates to date. The exposures are deep enough to yield information on both radial velocity and spectral absorption features; we consider the latter to be especially important for understanding the formation and evolution of the halo. The broad goals of the project are to characterize the dynamics, chemical abundance distribution, and structure/substructure of the M31 halo, with an emphasis on the global statistics and properties of debris trails from past mergers.

In this, the first paper from our DEIMOS survey, we focus our attention on the giant southern stream in M31. Details of the spectroscopic data set on which this paper is based are described in §2, including target selection, observations, data reduction and verification methods, survey efficiency, success rate of the target-selection procedure, and velocity measurement error. The division of our stellar sample into M31 stream RGB, M31 general spheroid RGB, and foreground Galactic dwarf populations is described in §3. The dynamics of the stream and spheroid are described in §4. The metallicity distributions of the stream and general spheroid are compared in §5 based on broadband photometry and spectral absorption features. The main points of the paper are summarized in §6. The companion paper by Font et al. (2006, hereafter Paper II) uses these and other data to obtain constraints on the orbit of the stream and the nature of the possible progenitor.

9 We use the term “spheroid” to refer to M31’s bulge and/or halo (Guhathakurta et al. 2005).

2. DATA

2.1. Spectroscopic Target Selection

2.1.1. Photometric Prescreening of M31 Red Giants

Candidate RGB stars in M31’s giant southern stream were selected from the field a3 photometry/astrometry catalog of Ostheimer (2002). The catalog is based on Kitt Peak National Observatory (KPNO) 4 m telescope and MOSAIC camera images in the Washington system $M$ and $T_2$ bands, as well as the intermediate-width DDO51 band. Sources were identified and photometered using the DAOPHOT II and ALLFRAME software packages (Stetson 1992, 1994). The photometric transformation relations of Majewski et al. (2000),

$$V = -0.006 + M - 0.200(M - T_2),$$

$$I = T_2,$$

were used to obtain magnitudes on the Johnson-Cousins system (the $T_2$ band is essentially identical to the $I$ band). The Ostheimer (2002) survey obtained KPNO MOSAIC data in 10 fields around the M31 halo, six of them on the southeastern minor axis extending out to $\pm 10^\circ$ from the galaxy center. The observations were carried out before the discovery of the giant southern stream, and it is pure chance that field a3 happens to intersect the stream.

The location of the $35^\prime \times 35^\prime$ field a3 relative to the stream is shown in Figure 1 (inset). The field center is located $\xi = +74^\circ 8$ (east) and $\eta = -127^\circ 5$ (south) with respect to M31’s center. Only a portion of the field, the southwestern or lower right half, is on the stream; the northeastern edge of the stream runs more or less diagonally across field a3 from northwest to southeast. It should be noted that the surface density of luminous M31 RGB stars is quite low in this remote field, even in the onstream portion, and there is substantial foreground and background contamination (see §2.4).

The DDO51 filter has a passband of width $\Delta \lambda \approx 100$ Å centered at $\lambda \sim 5150$ Å, designed to include the surface gravity–sensitive Mg I triplet and MgH stellar absorption features (Majewski et al. 2000). The features are strong in dwarf stars but weak in RGB stars. Following Palma et al. (2003), each object is assigned a probability of being an RGB star, $P_{\text{giant}}$, based on the degree of overlap of its photometric error ellipse with the (pre-determined) locus of dwarf stars in the $(M - DDO51)$ versus $(M - T_2)$ color–color diagram. The $P_{\text{giant}}$ parameter is effective in guiding the selection of RGB stars and in greatly reducing dwarf contamination in the sample, but it is not a perfect discriminant at the relatively faint magnitudes of our survey (as shown in §3.3.3); metallicity variations cause RGB stars to have a relatively broad distribution in two-color space, and the few more metal-rich ones that happen to intersect the dwarf locus are assigned a low $P_{\text{giant}}$ value; moreover, RGB/dwarf stars in the tails of the photometric error distribution can scatter close to or far from the dwarf locus and would then be assigned a low or high $P_{\text{giant}}$ value, respectively. Since metallicity is a secondary parameter (after surface gravity) in determining the strength of the Mg h and MgH features, one might worry about DDO51 selection introducing a bias against metal-rich RGB stars. We examine this issue in §5.1.2 and show that such a bias is not important for our sample.

The object-detection algorithm in DAOPHOT tends to reject sources that are extended relative to the point-spread function.
but some background field galaxies, especially compact ones, do slip through into the object catalog. The two-color method described above tends to assign high \( P_{\text{giant}} \) values to galaxies because they are dominated by the light of RGB stars and/or because their Mg \( b \)/MgH features are redshifted out of the DDO51 passband, making them appear featureless. For this reason the \( P_{\text{giant}} \) criterion was supplemented by the DAOPHOT-based morphological criteria \( \text{chi} \) and \( \text{sharp} \) to reject galaxies. Naturally, the compact galaxy rejection efficiency of DAOPHOT’s source-finding algorithm and \( \text{chi}/\text{sharp} \) measurements depend critically on the seeing and depth of the KPNO MOSAIC imaging data.

### 2.1.2. Slit Mask Design

Three DEIMOS multislit masks (1–3) were designed for field a3 using A. C. Phillips’ DSIMULATOR software.\(^\text{11} \)

| Parameter                  | Mask 1          | Mask 2          | Mask 3          |
|----------------------------|-----------------|-----------------|-----------------|
| Pointing center:           |                 |                 |                 |
| \( \alpha_{\text{J2000}} \) | 00 48 21.16     | 00 47 47.24     | 00 48 23.17     |
| \( \delta_{\text{J2000}} \) | +39 02 39.2     | +39 05 56.3     | +39 12 38.5     |
| Position angle (deg east of north) | 64.2          | 178.2           | 270.0           |
| No. of spectroscopic science targets | 85             | 80              | 83              |
| Breakdown of targets by lists 0/1/2/3/4 | 5/53/23/9/8 | 3/54/15/11/17/1 | 4/49/17/16/1 |
| Date of observations (UT)  | 2002 Aug 16    | 2002 Oct 11     | 2003 Oct 26     |
| Exposure time (s)           | 2 x 1800       | 3 x 1800        | 3 x 1200        |
| Number of \( Q \) = \-2/1/2/3/4 cases | 8/2/2/14/15/26 | 6/19/14/9/32 | 2/0/22/11/48 |

Notes: Units of right ascension are hours, minutes, and seconds, and units of declination are degrees, arcminutes, and arcseconds. The number of spectroscopic science targets selected from lists 1–3 for masks 1 and 2 and from lists 1–4 for mask 3 (in order of decreasing priority) is indicated (see Appendix A1); this number does not include alignment stars (list 0). The number of objects is also broken down according to \( Q \), a code indicating the reliability of the measured redshift and spectral quality (\( \frac{1}{2} \).3.2).

---

\(^{11}\) See http://www.ucolick.org/~phillips/deimos_ref/masks.html for details.
Keck II 10 m telescope and DEIMOS with its 1200 lines mm\(^{-1}\) spectral resolution is 3.44 km s\(^{-1}\). The 1\(\sigma\) triplet. While this is the characteristic width of stellar absorption lines, the precision with which the centroid of a line can be determined is typically a small fraction of its width. The centroiding accuracy, and hence the radial velocity measurement error (\(\S 2.5\)), depends on the S/N or, more specifically, on the significance of the cross-correlation peak (Tonry et al. 1996).

The mask 1 observation was carried out within 2 months of the initial commissioning of the DEIMOS spectrograph, and even the mask 2 observation was carried out during a period when improvements were continually being made to the instrument; the mask 3 observation was carried out during a period when improved rectification, flat-fielding, and wavelength-calibration purposes (each tick mark represents unity), and ordered by (\(V-I\)) color (indicated on the right side of the y axis). One of two instrument settings was used covering 6450–9150 or 7200–9900 \(\AA\). The atmospheric A-band correction is inadequate in the present reduction, so this region is excluded. The night-sky spectrum (bottom plot), a composite of the two wavelength settings, has been smoothed and shifted by +458 km s\(^{-1}\), the mean shift for stars in M31’s giant southern stream. The strongest TiO band heads, the Na\(d\) doublet, and the Ca\(ii\) triplet lines are marked. The TiO bands increase in strength with (\(V-I\)) color; the redder TiO band lies in the A-band gap, but the break in the spectrum is apparent for red stars. The Na line is surface gravity–sensitive and is thus a discriminator between M31 RGB stars and foreground Galactic dwarfs for (\(V-I\)) \(\approx\) 2 (triangles).

in the inter-CCD gap and to extend the blue side coverage to the TiO \(\lambda 7100\) and H\(\alpha\) \(\delta 6563\) features. The 1200 lines mm\(^{-1}\) grating yields a dispersion of 0.33 \(\AA\) pixel\(^{-1}\), and the spatial scale is 0.12 pixel\(^{-1}\). The 4\(\times\)2 array of 2K \(\times\) 4K CCDs span a spectral range of about \(\Delta \lambda \approx 2700 \AA\) and a total slit mask length of over 16\'. Thus, spectra from masks 1 and 2 (fall 2002) cover the range 7200–9900 \(\AA\), while mask 3 spectra (fall 2003) cover 6450–9150 \(\AA\). The width of the mask (in the dispersion direction) over which slitlets are distributed is about 2000 pixels, so the exact spectral coverage varies from slitlet to slitlet by up to a few hundred angstroms; moreover, some spectra are truncated by vignetting (see Fig. 3).

The 1\(\sigma\) slitlet width used on our masks subtends 4.8 pixels, given an anamorphic demagnification factor of 0.57 for the 1200 lines mm\(^{-1}\) grating at 8500 \(\AA\). The actual resolution is slightly better than this: for a typical seeing of 0.8 (FWHM), the spectral resolution is 3.8 pixel = 1.26 \(\AA\), which corresponds to 44 km s\(^{-1}\) or \(R \approx 7000\) at the Ca\(ii\) triplet. While this is the characteristic width of stellar absorption lines, the precision with which the centroid of a line can be determined is typically a small fraction of its width. The centroiding accuracy, and hence the radial velocity measurement error (\(\S 2.5\)), depends on the S/N or, more specifically, on the significance of the cross-correlation peak (Tonry & Davis 1979).

Prior to starting spectroscopic exposures on each slit mask, we used the mask-alignment procedure developed by the DEEP2 (phase 2 of the Deep Extragalactic Evolutionary Probe) team: guide stars on the TV guider camera were used for coarse alignment;
direct images were then obtained through the mask with the grating in zeroth order to fine-tune the alignment (both position angle and translation). The procedure converged after 2–3 iterations, with typical residuals of $\geq 0.1^\prime$ between the position of the alignment star and the center of its alignment box. This is indicative of the level of astrometric precision in the Ostheimer (2002) catalog from which our spectroscopic targets were drawn.

2.3. Data Reduction

2.3.1. Pipeline Processing and One-Dimensional Spectra

The three DEIMOS masks in field a3 were processed through the specl2d software pipeline (ver. 1.1.4) developed by the DEEP2 team at the University of California, Berkeley (UCB) for that survey. Briefly, the pipeline performed flat-fielding, rectification, slit function and fringing corrections, wavelength calibration, and sky subtraction on the two-dimensional spectra. The individual exposures of each spectrum were averaged, and one-dimensional spectra were extracted. Finally, the specl2d pipeline developed for the DEEP2 survey was used to cross-correlate the one-dimensional spectra against a variety of stellar and galactic templates to determine the redshift of each science target. A detailed description of the main data reduction/processing steps is given in Appendix A2.1.13

The one-dimensional spectra shown in Figure 3 represent the top one-third of our stellar sample in terms of S/N, typically $\sim 15 \AA^{-1}$ at the Ca II triplet for this subset. The spectra have been normalized to unit flux at 8500 Å; they are not flux calibrated, and the telluric A-band feature is still present (it will be removed is Poisson-limited subtraction of the bright night-sky emission our Fig. 3 to their Fig. 1). The most significant improvement in this section we assess the efficiency of our spectroscopic survey and the success rate of the photometric target selection procedure, i.e., the yield of M31 RGB stars. We do this by counting up the numbers of objects from all three masks in the different redshift quality ($Q$) code categories.

Starting with the failed $z$ measurements, the number of $Q = -2$ is 221, and the fraction of spectroscopic successes is higher. There are a few reasons for this. The fall 2002 mask designs were based on a preliminary optimal model for the spectrograph in DISIMULATOR; design imperfections resulted in some slitlets on masks 1 and 2 being affected by vignetting and inter-CCD gaps. The DEIMOS optical model was refined in time for the fall 2003 mask designs. The procedure for taking proper calibration data to achieve good wavelength calibration and flat-fielding was fine-tuned over the course of a few months by the DEEP2 and DEIMOS teams, so the calibration is less than ideal for the two early masks. The fall 2002 wavelength setting was not optimal in that there was a finite chance of one of the three Ca II triplet lines landing in the inter-CCD gap; the wavelength setting was improved for the fall 2003 observations. Finally, the observing conditions were slightly subpar for the two early masks, with occasional thin cirrus and/or worse-than-average seeing.

It is noteworthy that there is no significant difference in success rate between masks 1 and 2, even though the integration time is 50% longer for the latter. This suggests that factors other than total exposure time (e.g., the factors listed above) are responsible for determining the final data quality and success rate for the two early masks.

There are nine catastrophic failures ($Q = -2$) among the 221 unique RGB candidates targeted on the three masks. While these failures lower the overall efficiency of our survey, it is important to keep in mind that they are instrumental failures that in no way reflect the physical properties of the targeted objects or the success rate of the photometric screening procedure; i.e., it is as though these objects were not observed at all. For this reason, the percentages quoted below are measured relative to a denominator of 212 objects (221 – 9), those for which spectra were successfully obtained.

The number (and percentage) of $Q = 1$ and 2 cases is 38 (18%) and 46 (22%), respectively. Among the remaining 128 objects...
(60%) with definite redshifts, there are 44 galaxies (21%), mostly compact emission-line galaxies, spanning the redshift range $z = 0.1$–1.5. That leaves 84 stars (39%), of which 68 (32%) are M31 RGB stars and 16 (7.5%) are foreground Galactic dwarf stars, as we show in §3.3. The efficiency and success rate for the rest of our DEIMOS survey are expected to be higher than the fractions quoted here: most of the remaining slit masks (i.e., those not presented in this paper) were observed in fall 2003 or later and should therefore be comparable to mask 3.

It is instructive to examine the foreground/background contamination rates in the present study in the context of the earlier RG02 survey. The surface density of M31 stars at the $R = 31$ kpc location of our field a3 (including the contribution of the giant southern stream) should be roughly similar to that in RG02’s $R = 19$ kpc minor-axis field; the surface density of foreground Galactic stars and background field galaxies is also expected to be the same in the two fields. Milky Way dwarf stars constituted an estimated 57% of RG02’s sample of 80 stars, whereas they represent only 16 out of the 84 stars (19%) in the present study, or 6 out of 62 stars (10%) if one excludes the “filler” targets (lists 3 and 4) for which the $P_{\text{giant}}$ criterion was relaxed or dropped. The suppression of foreground contaminants in our study is attributable to DDO51-based preselection of spectroscopic targets (§2.1). Spectroscopically confirmed galaxies comprise about a fifth of our sample of 212 targets, but the true galaxy contamination rate could easily be higher by a factor of 2, since some (unknown) fraction of the $Q = 2$ cases (and possibly $Q = 1$ cases) are probably background galaxies. The efficiency with which the morphological criteria $\chi$ and sharp reject compact galaxies is very sensitive to the seeing and depth of the KPNO MOSAIC image; there is a fair bit of variation in seeing/depth across the different fields in the Ostheimer (2002) survey. The RG02 study had a lower galaxy contamination rate (estimated to $\sim 10\%$) thanks to their use of four-band photometry to prescreen against galaxies and better seeing.

2.5. Velocity Measurement Error

The 27 duplicate measurements across the three masks include 13 cases for which both members of the pair have secure $z$ determinations ($Q = 3$ or 4); eight stars, three emission-line galaxies, and two absorption-line galaxies. The rms radial velocity difference between pairs of measurements for the eight stars is $21 \pm 5$ km s$^{-1}$. The $\pm 1\sigma$ uncertainty in the rms estimate was derived from 1000 sets of eight Monte Carlo drawings from a Gaussian distribution. Assuming that the measurement uncertainty is the same for each member of the pair, the radial velocity error for an individual measurement is $\sqrt{2}$ times smaller than the rms of the difference, or $15 \pm 3.5$ km s$^{-1}$. The velocity error is probably smaller for mask 3 than for the other two masks, but this is ignored in the above estimate. We will use this velocity error estimate in §4.1.3 to constrain the coldness of the stream.

The overlap areas between masks, where we have duplicate observations, are located near the ends of the masks (Fig. 1), and the data quality appears to be suboptimal in these regions (§2.4). This provides a plausible explanation for the fact that 5 of the 16 duplicate star measurements (31%) are “$z$-by-hand” cases, whereas only 13% of all secure $z$ determinations are in this category. The velocity measurement error for the typical star in our sample is likely to be somewhat smaller than the above estimate of 15 km s$^{-1}$.

3. SORTING OUT THE STELLAR SAMPLE

In this section we describe how the 84 confirmed stars in our sample are sorted into three groups: (1) RGB stars in M31’s giant southern stream, (2) RGB stars belonging to M31’s general spheroid population, and (3) foreground Galactic dwarf stars. The surface density of luminous RGB stars is relatively low at the location of our field a3, in the outer spheroid of M31 at a projected distance of $R = 31$ kpc from the center of the galaxy, close to its minor axis. Despite the $\approx 3\times$ overdensity due to the presence of the stream (§3.2) and the use of DDO51 photometry to screen out Milky Way dwarf stars (§2.1), our spectroscopic sample contains a nonnegligible fraction of foreground dwarf star (and background field galaxy) contaminants. We demonstrate below that a few key pieces of photometric and spectroscopic information can be used to distinguish between stream and general spheroid populations and to eliminate all contaminants from our field a3 sample without any significant loss of M31 RGB stars. The radial velocity distribution of stars is a logical starting point for this analysis.

3.1. Line-of-Sight Velocity Distribution

Figure 4 shows radial velocity histograms for stars in our field a3 spectroscopic sample, first slit mask by slit mask and then for the combined sample of 84 unique objects. Three features of the distribution are worthy of note: (1) a strong and narrow peak that dominates each of the four histograms and presumably corresponds to M31’s giant southern stream, (2) a low-level broad component that is shifted toward less negative velocities than the main peak ($-500$ km s$^{-1} \leq v < -200$ km s$^{-1}$), and (3) a weak concentration of stars seen as a separate group in the range $-150$ km s$^{-1} \leq v < 0$ km s$^{-1}$. A two-Gaussian maximum-likelihood fit to the first two components is shown as a solid line in Figure 4d. We show below that the second component represents...
M31’s general spheroid population (§ 4.2), while the third consists of Galactic dwarf stars along the line of sight but well in front of M31, of course (§ 3.3).

Mask 3 has the highest success rate of the three masks for reasons discussed in § 2.4: their velocity histograms contain 26, 27, and 39 stars, respectively, even though the number of spectroscopic targets on the three masks is comparable. The foreground dwarf star fraction is highest in the mask 3 velocity histogram because the $P_{\text{giant}}$ criterion was dropped entirely while selecting lists 3 and 4 filler spectroscopic targets (see § 2.1.2).

3.2. Contrast of Stream against Smooth Spheroid

The two-Gaussian fit indicates that the ratio of giant southern stream to general spheroid stars is about 45:23; i.e., the stream comprises $65^{\pm12}\%$ (90% confidence limit from the maximum-likelihood analysis) of the M31 population in field a3. The actual ratio of stream to general spheroid stars may be somewhat higher than this: the $I < 22.5$ limiting magnitude, and possibly the DDO51 criterion, used in the spectroscopic target selection process tends to bias the sample against the most metal-rich RGB stars, and these stars constitute a larger fraction of the stream than the general spheroid population (§ 5).

It is important to note that we can only distinguish between stream and general spheroid stars on a statistical basis. The radial velocity range over which the stream dominates, $v < -410 \text{ km s}^{-1}$, contains 47 stars, of which 45 are estimated to be members of the stream while the remaining two are members of the spheroid. The 21 M31 RGB stars outside this velocity range are all likely to be members of the spheroid.

The surface density of the stream appears to be roughly constant along its length over a wide range (see the star count map in Ferguson et al. [2002]). Since the spheroid density falls monotonically with increasing distance from the galaxy center, the contrast of the stream is expected to become progressively stronger. This is evident from the velocity histograms presented in the Ibata et al. (2004) study (compare fields 1, 2, 6, and 8 in their Fig. 1); in fact, the contrast in their innermost field (field 8) is so low that the stream is not readily discernible as a distinct population. It should be noted that the Ibata et al. fields were chosen to run along the highest surface density part of stream, whereas our field a3 sample is drawn from the edge of the stream (by chance), so the stream-to—general spheroid ratio in their fields should be slightly higher than in ours at the same radial distance from M31.

3.3. Rejecting Foreground Galactic Dwarf Stars

This section discusses foreground Milky Way dwarf star contaminants in our sample. Galaxies with high-quality spectra ($Q = 3$ or 4) are easily identified on the basis of redshift and spectral characteristics and removed from the sample, but it is a little more difficult to screen out foreground dwarf stars. Five pieces of information can be used to distinguish M31 RGB stars from foreground Galactic dwarfs: (1) radial velocity, (2) location within the $(V - I, I)$ CMD, (3) $P_{\text{giant}}$ parameter derived from the $(M - \text{DDO51})$ versus $(M - T_2)$ two-color diagram, (4) strength of the Na i $\lambda 8190$ doublet in red (cool) stars, and (5) comparison of photometric versus spectroscopic metallicity estimates. Each of these criteria is discussed in turn below. In general, no single criterion of the five is by itself a perfect discriminant between RGB and dwarf stars, but the combination is very effective. A likelihood-based approach to combining the above diagnostics will be discussed in a future paper (K. M. Gilbert et al. 2006, in preparation).

3.3.1. Radial Velocity

The radial velocity histogram of the combined stellar sample (Fig. 4d) has a distinct gap separating a group of 68 stars with $v < -200 \text{ km s}^{-1}$ from a group of 16 stars with $v > -150 \text{ km s}^{-1}$ (shaded histogram). The former group of stars includes the prominent peak associated with the giant southern stream and is roughly centered on M31’s systemic velocity of $-300 \text{ km s}^{-1}$; they are therefore designated “candidate M31 RGB stars.” The latter group of stars occupies the radial velocity range predicted by the Institute for Advanced Study Galaxy (IASG) star-count model (RG02; Bahcall & Soneira 1984; Ratnatunga & Bahcall 1985); they are therefore designated “candidate Galactic dwarf stars.” We now test the validity of these designations by comparing the properties of these two groups of stars.

3.3.2. Color-Magnitude Diagram

The CMD locations of the two subgroups of stars are consistent with the above designations. Candidate M31 RGB stars have a distribution that is nicely bracketed by the model RGB tracks (Fig. 2d), with the majority lying below the RGB tip (the few exceptions will be discussed in § 5.1.3). By contrast, very few candidate Galactic dwarfs lie below the RGB tip (Fig. 2c); this is not surprising because the density of Milky Way (thin and thick disk) stars is on the decline at these faint apparent magnitudes, and the probability of including a foreground dwarf star is further diminished by the onset of M31’s RGB (see Figs. 2a and 2b).

3.3.3. DDO51-based Selection

A feature of our DEIMOS spectroscopic survey that sets it apart from previous studies of M31 is DDO51-based prescreening of RGB stars (§ 2.1). Figure 5 shows a plot of $P_{\text{giant}}$ versus radial velocity for the 84 stars in our sample. As expected,
spreads have much lower resolution than theirs but should be fine for an assessment of the relative line strengths of RGG versus dwarf stars. The error \( \sigma_{\text{EW}(\text{Na})} \) is assumed to scale inversely with S/N and is estimated empirically from the eight stars with duplicate measurements. Red stars display a bimodal distribution of Na i line strengths as expected. It is very reassuring that our velocity-based subsamples track this bimodality. The high-S/N cases with \((V-I)_0 > 1.8\) in Figure 6 best demonstrate this: candidate M31 RGB stars (circles) lie at or below the \( \text{EW}(\text{Na}) = 1.8 \text{ Å} \) threshold (dashed horizontal line), while candidate Galactic dwarfs (crosses) lie at or above it.

### 3.3.5. Calcium Triplet Line Strength

The final criterion for distinguishing candidate M31 RGB stars from candidate foreground Galactic dwarfs is a comparison between photometric and spectroscopic metallicity estimates. The \([\text{Fe/H}]_{\text{phot}} \) estimate is derived from the position of the star in the CMD relative to model RGB tracks (see \S 5.1.1 below). The \([\text{Fe/H}]_{\text{spec}} \) estimate is derived from the strength of the Ca ii absorption-line triplet and empirical calibration relations based on RGB stars in Galactic globular clusters (\S 5.2.1). RG02 found that M31 RGB stars lie close to the \([\text{Fe/H}]_{\text{spec}} = [\text{Fe/H}]_{\text{phot}} \) line, whereas Galactic dwarfs have weaker Ca ii lines so that \([\text{Fe/H}]_{\text{spec}} < [\text{Fe/H}]_{\text{phot}} \) for the most part (see their Figs. 13 and 14). We show in Figure 12 and \S 5.2.3 that the Ca ii triplet line strengths measured in co-added spectra of M31 RGB stars and Galactic dwarfs in our field a3 sample follow these same trends.

### 3.3.6. Discussion: Toward a Clean and Complete Sample of M31 Red Giant Stars

Of the five diagnostics discussed above, radial velocity and Na i doublet line strength are the most powerful for foreground dwarf rejection. Based on the (admittedly arbitrary) color and \( \text{EW}(\text{Na}) \) thresholds in Figure 6 (dashed lines), there are 10 and 19 stars in the top right and bottom right sections that should be considered definite Galactic dwarfs and definite M31 RGB stars, respectively. All 29 stars were required to be well measured, \( \sigma_{\text{EW}(\text{Na})} \leq 1 \text{ Å} \), and to be separated by at least 1.5 \( \sigma_{\text{EW}(\text{Na})} \) from the horizontal line. The Na i doublet feature is easily visible for the 10 definite dwarfs in Figure 3 (triangles). In Figure 5 the definite dwarfs and definite RGB stars are marked as crosses and filled circles, respectively. It is very reassuring that these two subgroups maintain perfect separation in radial velocity; in other words, none of the candidate Galactic dwarfs turned out to be a definite M31 RGB star and vice versa.

It is worth checking whether any of the 16 candidate foreground Galactic dwarfs might possibly be M31 RGB stars. With this in mind, we examine their properties using Figures 2, 5, and 6, going through the subsample one star at a time trying to identify possible RGB interlopers:

1. Obviously, the 10 cool stars that lie well above the \( \text{EW}(\text{Na}) \) threshold are definite dwarfs. It is useful to check how they fare with respect to some of the other diagnostics. All 10 lie at or above M31’s RGB tip in the CMD. Three of them have \( P_{\text{giant}} \geq 0.7 \), two have \( P_{\text{giant}} \approx 0.4 \), and five have \( P_{\text{giant}} < 0.1 \), again emphasizing that the \( P_{\text{giant}} \) parameter is not a perfect discriminator between dwarfs and RGB stars.

2. The candidate cool dwarf at \((V-I)_0 = 1.98 \) appears to have a strong Na i doublet feature, \( \text{EW}(\text{Na}) = 3.19 \text{ Å} \), but its spectrum is so noisy that its offset above the threshold corresponds to only 1 \( \sigma_{\text{EW}(\text{Na})} \). In other words, the Na test suggests that it is a dwarf but is not definitive. The object lies well below M31’s RGB tip in the CMD. Its relatively low \( P_{\text{giant}} \) value of 0.35 tips the scale in favor of a Galactic dwarf classification.
3. Two candidate dwarfs lie close to the EW(Na) threshold, one 0.5 mag redder than the color cut and the other 0.2 mag bluer than it. Even though the Na diagnostic is inconclusive, they are both likely to be dwarfs based on two other factors: (1) relatively small $P_{\text{giant}}$ values, 0.41 and 0.09, respectively; and (2) CMD locations close to M31’s RGB tip.

4. The candidate dwarf with $(V-I)_0 = 1.46$ and EW(Na) = 0.65 Å is too hot for the Na test to be used. It is about 1 mag fainter than M31’s RGB tip and has a relatively high $P_{\text{giant}}$ value, 0.63. These factors suggest that the object may be an M31 RGB star.

5. Finally, two candidate dwarfs lie within $\approx 0.1$ mag of the color cut but on the blue side. Both have noisy spectra but appear to lie significantly below the threshold if our empirical scaling of $\sigma(\text{EW(Na)})$ is to be trusted. It is intriguing that the two have relatively large $P_{\text{giant}}$ values, 0.54 and 0.86, and both lie $\approx 1$ mag below the tip of M31’s RGB in the CMD. Like the previous candidate dwarf, these two should also be considered possible M31 RGB stars.

In summary, there are three stars among the 16 velocity-selected candidate dwarfs that could be M31 RGB stars. All three are relatively faint and do not have well-measured photometric and spectroscopic parameters as a result. None of the three present a strong/definite enough case to warrant reclassification at this stage. We return to a discussion of these objects in § 4.2.

We next use similar reasoning to test whether any of the 68 candidate M31 RGB stars might possibly be foreground Galactic dwarfs:

1. The 48 candidate M31 RGB stars with $v < -400$ km s$^{-1}$ can be ruled out as potential dwarfs on the basis of the IASG Galactic structure model (Bahcall & Soneira 1984; Ratnatunga & Bahcall 1985), which predicts that the foreground dwarf contamination rate should be negligible at these large negative radial velocities (see Fig. 5 of RG02). The giant southern stream population should be completely free of foreground dwarf contamination by virtue of its large blueshift.

2. The remaining 20 candidate M31 RGB stars in the radial velocity range $-400$ km s$^{-1} < v < -200$ km s$^{-1}$ all have $P_{\text{giant}} \geq 0.6$, so are unlikely to be dwarfs.

3. The Na diagnostic does not turn up any compelling evidence to suggest that there are dwarf interlopers among this sample of 20. Twelve are to the left of the color cut, and eight are to the right of it. None of the 20 are significantly above the EW(Na) threshold; six are significantly below it, three on each side of the color cut.

In summary, there is nothing to suggest that any of the candidate M31 RGB stars requires reclassification. Given the clean RGB versus dwarf star distinction provided by the radial velocity and Na i doublet diagnostics in Figures 5 and 6, coupled with all the earlier evidence, it is probably safe to drop the term “candidate” from the designations: unless otherwise mentioned, we refer to the $v < -200$ km s$^{-1}$ group as M31 RGB stars and the $v > -150$ km s$^{-1}$ group as Galactic dwarfs throughout the rest of this paper.

3.3.7. Comparison to Earlier Studies

Most previous spectroscopic studies of M31 stars have adopted a more limited approach to RGB versus dwarf star discrimination. Most did not have access to DDO51 photometry, which is so valuable for suppressing foreground dwarf contamination in our DEIMOS survey. RG02 used radial velocities, CMD information, and the [Fe/H]$_{\text{spec}}$ versus [Fe/H]$_{\text{phot}}$ comparison; the resulting RGB/dwarf star separation was not as clear as in this study. The Reitzel et al. (2004) and Ibata et al. (2005) studies used radial velocities alone, while Ibata et al. (2004) made no mention of foreground contamination; none of these last three studies made use of spectral contamination line strengths.

4. DYNAMICS

The dynamics of the giant southern stream and general spheroid (nonstream) populations in M31 are described in this section. A maximum-likelihood fit of the sum of two Gaussians is carried out on the radial velocity distribution of the 68 M31 RGB stars. This is used to characterize the mean velocity and velocity dispersion of the two components. Figure 7 shows $\Delta \chi^2 (\equiv \chi^2 - \chi_{\text{min}}^2)$ curves for these four quantities (four of the five parameters in the fit); the optimal value of each parameter and the uncertainty, in the form of 90% confidence limits, are indicated.

4.1. Giant Southern Stream

4.1.1. Mean Radial Velocity

The taller/narrower of the two Gaussians in the maximum-likelihood fit corresponds to the main peak in the stellar radial velocity distribution (Fig. 4d). Based on this fit, we determine that the giant southern stream has a mean heliocentric radial velocity of $-458 \pm 6$ km s$^{-1}$ (90% confidence limit) or a line-of-sight velocity of $-158$ km s$^{-1}$ with respect to M31 (see Fig. 7a). Our measurement is consistent with recent measurements by Lewis et al. (2004) and Ibata et al. (2004) at other points along the stream in the sense that our field a3 lies between their fields 2 and 6 in terms of both sky position and radial velocity. Paper II uses the field a3 mean velocity from this study, along with other velocity data (Lewis et al. 2004; Ibata et al. 2004) and line-of-sight distance estimates (McConnachie et al. 2003) to constrain the orbit of the stream and its progenitor satellite.

4.1.2. Velocity Gradients

We next explore possible trends in the mean velocity of M31’s giant southern stream as a function of sky position. Figure 8 shows the radial velocity distribution of M31 RGB stars as a function of projected position along and perpendicular to the stream (Fig. 8, top and bottom, respectively). The positions are defined relative to the center of field a3 with $\Delta r_1$ increasing from northwest to southeast and $\Delta r_1$, increasing from southwest to northeast; in other words, the projection of the stream is assumed to be at a position angle of $-45^\circ$ (Fig. 1). We conclude that there is no strong velocity gradient along either axis. The dashed lines in Figure 8 mark the nominal slopes, $dv/dr_1 = -0.5$ km s$^{-1}$ arcmin$^{-1}$ and $dv/dr_1 = +0.6$ km s$^{-1}$ arcmin$^{-1}$, but there is uncertainty about the membership of any given RGB star (stream vs. general spheroid; see § 3.2) and large Poisson errors. As discussed in Paper II, our measurement of the local velocity gradient along the stream in field a3 is consistent with the published radial velocities over a longer spatial baseline (Lewis et al. 2004; Ibata et al. 2004).

Any systematic trend in the stream’s mean velocity with position tends to broaden the peak in the velocity histogram. The radial velocity gradient along the direction parallel to the stream is $dv/dr_1 = -0.47$ km s$^{-1}$ arcmin$^{-1}$, so this translates to a spread of $\Delta v = \pm 4.7$ km s$^{-1}$ over the $\Delta r_1 \sim 20'$ spanned by our three masks along the length of the giant southern stream (Figs. 1 and 8). This spread is small compared to the width of the peak associated with the stream. Indeed, we have checked explicitly that correcting for the gradient has a negligible effect on the width of the best-fit Gaussian.
4.1.3. Intrinsic Velocity Dispersion

Figure 7b shows the likelihood function for the width of the Gaussian corresponding to the giant southern stream. The best-fit width is $\sigma_v^{\text{stream}} = 21 \pm 7$ km s$^{-1}$ (90% confidence limit). The 1 $\sigma$ velocity measurement error of 15 km s$^{-1}$ ($\S$ 2.5) is subtracted in quadrature from the measured velocity dispersion of the stream. Our best estimate of the intrinsic line-of-sight velocity dispersion of the stream is $\sigma_v^{\text{stream}}(\text{intrinsic}) \approx 15$ km s$^{-1}$, but, given the large uncertainty in the measured value, we conclude that $\sigma_v^{\text{stream}}(\text{intrinsic}) \lesssim 23$ km s$^{-1}$ (90% confidence limit). This is comparable to the velocity dispersions measured in the Milky Way’s Monoceros and Sagittarius Streams (Crane et al. 2003; Majewski et al. 2004). We consider the coldness of the stream and its implications in Paper II.

4.2. General Spheroid (Nonstream) Population: Evidence of Substructure

In this section we turn our attention to the kinematics of M31’s general spheroid (i.e., nonstream) population. The broad, low-level component in the combined stellar radial velocity histogram (Fig. 4d) spans the range $-500 \leq v < -200$ km s$^{-1}$ and extends to the right of the main peak. As we show in $\S$ 5.1.3., this component has a large spread in [Fe/H] comparable to that seen in other studies of the spheroid (even though there are differences in detail). The broad component is not merely the tail of the giant southern stream’s radial velocity distribution, judging from the differences in their stellar populations: the former appears to be more metal-poor, has a larger metallicity spread, and lacks stars above the RGB tip (see $\S$ 5.1.3 and Figs. 2 and 10).

The broader of the two components in the maximum-likelihood fit to the radial velocity distribution of M31 RGB stars is centered at a heliocentric velocity of $-333^{+33}_{-51}$ km s$^{-1}$ and has a Gaussian width of $65^{+32}_{-21}$ km s$^{-1}$ (90% confidence limit; see Figs. 7c and 7d). The broad spheroid component appears to be adequately fit by a Gaussian, but the large Poisson fluctuations (only $\approx 21$ stars belong to this component) make it impossible to tell whether its radial velocity distribution is drawn from a truly smooth and virialized underlying distribution or whether it contains substantial substructure.

The velocity distribution of the general spheroid population in field a3 appears to be anomalously narrow. Its width $\sigma_v^{\text{phil}}(\text{field a3}) = 65^{+32}_{-21}$ km s$^{-1}$ is significantly smaller than the width of $\sigma_v^{\text{phil}}(\text{other}) = 150$ km s$^{-1}$ measured for other M31 spheroid tracers: field RGB stars in an $R = 19$ kpc minor-axis field (RG02) and global samples of globular clusters and planetary nebulae (see Evans & Wilkinson 2000 and references therein). A plausible explanation of this is that our field a3 sample is dominated by substructure in the spheroid; e.g., it consists of one or two subclumps instead of a smooth, virialized distribution. This is the first of three lines of evidence presented in this paper pointing to the possible existence of substructure in the general spheroid population in this remote M31 field.

We next consider the three possible M31 RGB interlopers among the 16 stars with $v > -150$ km s$^{-1}$ (candidate Galactic dwarfs) that have been excluded from the dynamical analysis so far (see the list in $\S$ 3.3.6). Their radial velocities, $-21$, $-127$, and $-18$ km s$^{-1}$, place them right among the foreground Galactic dwarf population, but it is possible that they belong to M31 instead. For example, they would be within 3 $\sigma_v^{\text{phil}}$ of a broad distribution centered on M31’s systemic velocity of $-300$ km s$^{-1}$. Including all three as members of the M31 spheroid changes the
best-fit width to $\sigma_{\text{ph}}^{\text{best-fit}} = 116_{-22}^{+31} \text{ km s}^{-1}$, while including only the last pair of stars as spheroid members (probably more realistic) changes the width to $\sigma_{\text{ph}}^{\text{best-fit}} = 103_{-21}^{+28} \text{ km s}^{-1}$, where the quoted uncertainties are 90% confidence limits derived from new maximum-likelihood fits in each case. Our basic conclusion about the field a3 spheroid sample having an unusually narrow velocity distribution seems to be independent of whether or not the possible RGB interlopers are included in the sample.

The centroid of the radial velocity distribution of the spheroid population is offset by $-33 \text{ km s}^{-1}$ with respect to M31’s systemic velocity of $-300 \text{ km s}^{-1}$, and the offset is significant at the 90% confidence level based on the maximum-likelihood fit to the primary sample of 68 M31 RGB stars (possible interlopers excluded). Inclusion of the possible interlopers leads to a small shift in the best-fit spheroid mean velocity toward less negative values and therefore a reduction in the magnitude of the offset relative to systemic, along with a slight increase in the uncertainty. The net result of these changes is that the best-fit offset of $-19 \text{ km s}^{-1}$ is no longer statistically significant. If the observed offset is real/significant, it could be a second sign of substructure in M31’s spheroid. A third possible sign of substructure is the apparent difference between the metallicity distributions of the field a3 spheroid sample and RG02’s sample (see § 5.1.3).

An alternative explanation of the observed velocity offset (taken at face value) is that it is a result of global rotation of M31’s spheroid. The sense of spheroid rotation would then be the same as for M31’s disk, whose south-western half is known to be blue-shifted relative to systemic. Since our field a3 lies close to the minor axis, this would imply that we are seeing only a small component of the full spheroid rotation speed and would therefore place a lower limit of $(v_{\text{rot}}/\sigma_{\text{ph}})^{\text{best-fit}} \geq 0.2$. The M31 inner spheroid (bulge) appears to have an aspect ratio of $5:3$ (in projection) judging from the isopleths in the Ibata et al. (2001) star-count map. If the observed flattening is due to spheroid rotation, this would require $(v_{\text{rot}}/\sigma_{\text{ph}})^{\text{best-fit}} \sim 0.8$ (Binney & Tremaine 1987), which is consistent with our lower limit. If M31’s spheroid is confirmed to be rotating at this level, it would be in stark contrast to the Milky Way halo, which appears to have little or no (or even slight retrograde) rotation (Majewski 1992; Majewski et al. 1996; Popowski & Gould 1998).

In closing, we note that there are very few constraints on the global dynamics of M31’s stellar spheroid at the current time. Only two sight lines have been studied using field RGB stars—field a3 from this study, at a projected distance of $R = 31 \text{ kpc}$ slightly off the minor axis, and RG02’s 19 kpc minor-axis field—and both studies are hampered by small number statistics. Ibata et al. (2005) present radial velocity measurements from a large spectroscopic survey but only comment on M31’s disk and not its spheroid. Future papers from our survey will use all of the existing fields with DEIMOS spectra (about a dozen locations scattered around M31), and it should be possible to derive tighter constraints on the dynamics of the bulge and halo.

5. CHEMICAL ABUNDANCE (AND AGE)

The giant southern stream in M31 appears to be the remnant of a massive/luminous satellite. The chemical abundance and age distribution of the stream are best viewed in the context of the ensemble of former satellites that merged to form the galaxy’s halo. With this goal in mind, we compare the properties of luminous RGB stars in the stream and those in the general spheroid. The analysis of RGB star metallicities in this section is based mostly on their photometric properties with limited use of spectroscopic information; a full analysis of spectroscopic chemical abundances will be presented in a future paper. Some indirect age constraints are also obtained in this section.

5.1. Photometric Metallicity Estimates

5.1.1. Measurement Method and Errors

Photometric metallicity estimates are obtained for the 68 confirmed M31 RGB stars in our field a3 sample. As illustrated in Figure 2d, the position of each star in the $(V - I)_0$ versus $(V - I)_0$ CMD is compared to a set of model RGB fiducials (Girardi et al. 2000). The conversion of stellar photometry from the Washington system to the Johnson-Cousins system and star-by-star dereddening are described in § 2.1. The fiducials span a wide range of metallicities for $[\alpha/Fe] = 0$ (solar elemental abundance ratios) and an age of $t = 12.6 \text{ Gyr}$; they have been placed on the CMD using a true distance modulus of $(m - M)_0 = 24.47$, corresponding to an adopted distance to M31 of 783 kpc (Stanek & Garnavich 1998; Holland 1998). In order to estimate $[\text{Fe/H}]_{\text{phot}}$ for the stars, a Legendre polynomial of 6th order in $(V - I)_0$ and 10th order in $I_0$ is used to interpolate between the model RGB tracks.

If an M31 star is 2–6 Gyr old instead of our assumed age of 12.6 Gyr, its photometric metallicity estimate would need to be revised upward by about +0.3 dex (e.g., a $t = 6.3 \text{ Gyr}$ fiducial is shown as a dashed line in Fig. 2a). The overall uncertainty in the metallicities is about 0.3 dex, dominated by systematic errors such as age error/spread, residual differential reddening, $[\alpha/Fe]$ variations, and model inaccuracies; however, relative metallicity rankings can be achieved to somewhat greater accuracy than this for stars of comparable age. For stars located above the RGB tip in the CMD, the $[\text{Fe/H}]_{\text{phot}}$ estimates are based on linear extrapolation and are therefore very uncertain.

5.1.2. Selection Biases

Before studying the metallicity distribution of M31 stars, we investigate whether our RGB star sample is an unbiased sample. A couple of selection effects in particular are worth discussing.

It is well known that metal-line blanketing causes the RGB tip to get fainter with increasing metallicity. This causes the most metal-rich of the luminous RGB stars, those with $[\text{Fe/H}] > -1$, to be underrepresented in any magnitude-limited sample. RG02 characterized this bias in detail and corrected for it. We do not correct for the bias in this paper, since we are primarily interested in a differential stream versus general spheroid comparison rather than the absolute shape of the $[\text{Fe/H}]$ distribution. Moreover, the bias should be less pronounced in our field a3 sample (and the rest of the DEIMOS survey) than in RG02’s sample: they used a limiting $I$-band magnitude of 22.0, whereas our study used 22.0 for list 1 primary targets but relaxed it to 22.5 for lists 2–4 secondary and filler targets (§ 2.1.2).

Since metallicity is the second parameter (after surface gravity) in determining the strength of the Mg b/$\text{MgH}$ feature, our DDO51-based spectroscopic target-selection procedure is expected to introduce a bias against the most metal-rich RGB stars (§ 2.1.1). We investigate this by plotting $[\text{Fe/H}]_{\text{phot}}$ versus $P_{\text{giant}}$ in Figure 9. There is no strong or obvious trend for the primary and secondary spectroscopic targets ($P_{\text{giant}} > 0.5$) that form the bulk of the sample. The handful of lists 3–4 filler targets in our sample with $P_{\text{giant}} < 0.5$ appear to have the same $[\text{Fe/H}]_{\text{phot}}$ distribution as the rest of the sample—they are not particularly metal-rich. This test will ultimately be carried out with a larger sample of filler targets from the full DEIMOS survey. In summary, it appears that DDO51-based prescreening does not introduce any strong metallicity bias over that introduced by $I$-band selection.
The metallicity distribution of RGB stars in M31’s giant southern stream is compared to that of its general spheroid in Figure 10. The top panel shows radial velocity as a function of metallicity for 68 confirmed RGB stars. The portion below the dashed horizontal line is the velocity range occupied by the stream, \( v < -410 \text{ km s}^{-1} \). It appears that the 47 stars in this range are more skewed toward high metallicities than the 21 stars outside it. The thin solid histogram in the bottom panel is the \([\text{Fe/H}]_{\text{phot}}\) distribution of stars in the stream’s velocity range, while the dashed histogram represents the nonstream general spheroid population. We show below that the stream contains a few stars brighter than the RGB tip; \([\text{Fe/H}]_{\text{phot}}\) estimates for such stars are unreliable because they are based on a naive linear upward extrapolation of the RGB fiducials (§ 5.1.1). We therefore recompute the stream’s metallicity distribution, this time excluding the seven stars that lie above the tips of the model RGB tracks (thick histogram).

The stream appears to be more metal-rich on average than the general spheroid, \( \langle [\text{Fe/H}] \rangle = -0.51 \text{ versus } -0.74 \), and appears to have a smaller metallicity spread, 0.25 versus 0.40 dex. A two-sided Kolmogorov-Smirnov test indicates that there is only a 5% probability that the two \([\text{Fe/H}]\) distributions are drawn from the same parent distribution—in other words, the apparent difference between the stream and general spheroid metallicity distributions is a 2 \( \sigma \) effect. Another difference between the stream and general spheroid populations is evident from the CMD (Fig. 2d): there are nine M31 stars that lie near or above the RGB tip, and all nine are in the stream’s radial velocity range (pentagons). A couple of these stars in particular are more than 0.5 mag above M31’s RGB tip and yet must be M31 members given their large negative radial velocities. These are best explained in terms of an intermediate-age asymptotic giant branch (AGB) population in the stream.

Two pieces of information suggest that the true mean metallicity of the stream may be even higher, and the difference in mean metallicity between the stream and general spheroid larger/ more significant, than observed. First, if a sizable fraction of the RGB stars in the stream are indeed of intermediate age, then their photometric [Fe/H] estimates will need to be corrected upward by about +0.3 dex. Second, any bias resulting from our photometric selection procedure will tend to deplete/truncate the high end of the [Fe/H] distribution (§ 5.1.2), the end at which the observed difference between the stream and general spheroid distributions appears to be the greatest.

Even if the observed mean metallicity of the stream is taken at face value, it points to a relatively high luminosity for the progenitor satellite galaxy of the stream. An empirical correlation between metallicity and luminosity has been noted for dwarf satellite galaxies in the Local Group (Mateo 1998; Grebel & Guhathakurta 1999; Dekel & Woo 2003), which would indicate an absolute B-band magnitude of \( M_B \approx -17 \) for the stream’s progenitor or a luminosity of \( L_B \approx 10^8 L_{\odot} \). This is consistent with the lower limits on progenitor luminosity derived from the stream’s width and internal velocity dispersion in Paper II.

The mean [Fe/H] of \(-0.7 \text{ dex observed for the general spheroid population in field a3 and its total range of 1.5 dex are broadly consistent with published studies of the metallicity distribution in other fields around the M31 spheroid (Mould &
Kristian 1986; Durrell et al. 1994, 2001, 2004; Rich et al. 1996; Holland et al. 1996; Reitzel et al. 1998; Sarajedini & van Duyne 2001; Bellazzini et al. 2003). However, several factors make it difficult for us to carry out detailed comparisons or to draw firm conclusions about real metallicity variations from field to field. First, our field a3 general spheroid [Fe/H] distribution is based on only 21 RGB stars and therefore suffers from large Poisson fluctuations. Second, while our study, like these others, is based on photometry of RGB stars, there are significant differences in terms of data analysis techniques and associated systematic errors, sample definition, contamination issues, and selection biases. The RG02 study may be comparable to ours, since both are based on spectroscopy of RGB stars. We note that RG02 found the spheroid to be significantly more metal poor, by $\geq 0.5$ dex in the mean, with a tail in the [Fe/H] distribution extending down to $\approx -2$ dex that is simply not seen in the present study (see their Fig. 17c). This may be yet another sign of substructure in M31’s spheroid, something that needs to be confirmed using larger, more uniformly selected, and directly comparable samples.

5.2. Spectroscopic Constraints on Metallicity

The discussion of the chemical abundance distribution in M31’s spheroid has so far been based purely on photometric metallicity estimates. In this section we present a brief analysis of the Ca ii line strength in the spectra of RGB stars. The EW(Ca) measurement has a relatively large uncertainty associated with it compared to the random error in [Fe/H]$_{\text{phot}}$, for example, so we prefer not to translate it into an estimate of the spectroscopic metallicity [Fe/H]$_{\text{spec}}$. Instead, the Ca ii line strength is used as a point of comparison: the observed EW(Ca) is compared to the value predicted from the photometric metallicity estimate. We demonstrate that the two are in good agreement.

5.2.1. Predicted Calcium Line Strength

The photometric properties of each star are used to predict the strength of its Ca ii triplet. First, the photometric metallicity estimate [Fe/H]$_{\text{phot}}$ (§ 5.1.1) is taken to be the same as the metallicity on the spectroscopic scale defined by Carretta & Gratton (1997), [Fe/H]$_{\text{CG97}}$. Next, we use a well-established empirical calibration relation, based on luminous RGB stars in Milky Way globular clusters, to derive [Fe/H]$_{\text{CG97}}$ from the Ca ii triplet (Rutledge et al. 1997a, 1997b):

$$[\text{Fe}/\text{H}]_{\text{CG97}} = -2.66 + 0.42(\Sigma\text{Ca} - 0.64(V_{\text{HB}} - V)), \quad (2)$$

where $\Sigma\text{Ca}$ is the weighted sum of the EWs, in angstroms, of the three lines comprising the Ca ii triplet,

$$\Sigma\text{Ca} \equiv 0.5\text{EW}(8498 \, \text{Å}) + 1.0\text{EW}(8542 \, \text{Å}) + 0.6\text{EW}(8662 \, \text{Å}), \quad (3)$$

and the luminosity-based correction for the effect of surface gravity is made relative to the apparent magnitude of M31’s horizontal branch, $V_{\text{HB}} = 25.17$ (Holland et al. 1996). Inverting equation (2), the predicted Ca ii line strength is defined to be

$$\Sigma\text{Ca}_{\text{pred}} = 6.33 + 2.38[\text{Fe}/\text{H}]_{\text{phot}} + 0.64(V_{\text{HB}} - V). \quad (4)$$

Since this relation is based on Galactic globular cluster RGB stars, we are making the implicit assumption that M31 RGB stars are comparably old ($t \geq 10$ Gyr) and, more importantly, $\alpha$-enhanced to the same degree ($\alpha/\text{Fe} = +0.3$ dex). As discussed in § 5.2.2, RGB stars are grouped according to the $\Sigma\text{Ca}_{\text{pred}}$ parameter for the purpose of co-adding spectra.

5.2.2. Stellar Absorption Features and Co-Added Spectra

The normalized one-dimensional spectra shown in Figure 3 are arranged in order of the stars’ $(V-I)_0$ color becoming redder upward. The onset of the TiO band seems to occur at $(V-I)_0 \approx 1.8$, which corresponds to $T_{\text{eff}} = 4000$ K (Alonso et al. 1999). As might be expected, there is a good correlation between the observed broadband color and the strength of the TiO bands for stars that are redder (cooler) than this. Other features, such as the Ca ii triplet and occasionally the Na i doublet, are also visible. It should be noted, however, that the one-dimensional spectra shown in Figure 3 represent the best third of our sample. Unfortunately, the spectral S/N is not high enough to support a detailed star-by-star abundance analysis of all RGB stars in our sample.

We therefore decided to co-add the 92 spectra (including eight duplicate observations) in groups of about a dozen to improve the S/N. The spectra were grouped according to the expected Ca ii triplet absorption line strengths of the stars, $\Sigma\text{Ca}_{\text{pred}}$. Two different grouping schemes were tried. In the first scheme, no distinction was made between stream and general spheroid populations, and the entire set of RGB stars were simply divided into five bins by line strength. In the second scheme, the 47 RGB stars in the stream’s radial velocity range were divided into four bins and the 21 general spheroid RGB stars outside this range into two bins. In both schemes, the 16 Galactic dwarf stars were placed in a bin by themselves. All spectra were shifted to zero velocity (rest frame) and then combined with inverse-variance weighting. The resulting co-added spectra were then smoothed with a weighted boxcar of 5 pixels (1.7 Å) in width, comparable to the instrumental resolution.

Figure 11 shows the six co-added spectra using the first grouping scheme. The lowest co-added spectrum is that of Milky Way dwarfs, and the next five are M31 RGB star co-added spectra with predicted line strength increasing upward. The top spectrum (thick line) is a model RGB spectrum from Schiavon & Barbuy (1999) computed using $T_{\text{eff}} = 4000$ K, $\log g = 1.5$, and $[\text{Fe}/\text{H}] = -0.3$. It is used merely to illustrate/identify usable absorption features in the far red region of the spectrum; e.g., Ca, Fe, Mg, Ti, and V lines for RGB stars, and these plus the surface gravity–sensitive Na i doublet for dwarfs. A detailed quantitative treatment of absorption lines is postponed until a future paper; only the Ca ii line strengths are discussed briefly below. We have therefore not made any attempt to fine-tune the match between model and data in terms of line strength or spectral resolution.

5.2.3. Measured Calcium Line Strength and Some Sanity Checks

The EWs of the Ca ii lines were measured in all co-added spectra using an 18 Å window for each line. The continuum level (close to unity, since these are normalized spectra) was measured on either side of each Ca ii line using multiple nonoverlapping windows of the same width while avoiding regions of the spectrum known to contain other strong lines (see the RGB model spectrum in Fig. 11). For each spectrum the weighted sum of the EWs of the three Ca ii lines, $\Sigma\text{Ca}_{\text{meas}}$, was computed according to equation (3).

Figure 12 compares the measured versus predicted combined Ca ii EWs for the seven co-added spectra from the second grouping scheme defined above. It is reassuring to see that most of the M31 RGB co-additions, both stream and general spheroid stars (Fig. 12, pentagons and squares, respectively), lie close to the one-to-one line. Taken at face value, this would seem to suggest
that these stars resemble those in Galactic globular clusters (from which the $[\text{Fe}/\text{H}]_{\mathrm{spec}}$ calibration relations were derived) in that they are old and have $[\alpha/\text{Fe}] \approx +0.3$. If confirmed, this would imply that RGB stars in M31’s spheroid have higher $[\alpha/\text{Fe}]$ than their counterparts in Local Group dwarf spheroidal satellites (Shetrone et al. 2001). It is also possible that there are differences in both age and $[\alpha/\text{Fe}]$ between M31 and Galactic globular cluster RGB stars but that the effects of these two parameters have conspired to cancel each other out in the $\Sigma \text{Ca}_{\mathrm{mean}}$ versus $\Sigma \text{Ca}_{\mathrm{pred}}$ diagram.

The strongest lined (metal-rich) stream RGB stars lie slightly above (or to the left of) the one-to-one line in Figure 12. If these happen to be intermediate-age stars—i.e., in the $\approx 2–6$ Gyr range instead of the 12.6 Gyr age that is assumed in fitting model RGB fiducials—their photometric metallicity estimates would be biased low by $-0.3$ dex ($\approx$ 5.1). Thus, the $\Sigma \text{Ca}_{\mathrm{pred}}$ values would be biased high by 0.7 Å ($\approx$ 4) and would, for the most part, explain the observed offset from the one-to-one line. This assumes that intermediate-age RGB stars follow the same $\Sigma \text{Ca} \rightarrow [\text{Fe}/\text{H}]_{\mathrm{CG97}}$ calibration relation as old RGB stars (eq. [2]), which is yet to be verified.

Eight of the nine potential intermediate-age AGB stars located above the RGB tip in the CMD are in the first bin for spectral co-adds of stream stars in Figure 12; i.e., they have among the lowest $\Sigma \text{Ca}_{\mathrm{pred}}$ values and are expected to have relatively weak Ca ii lines. This bin also contains four metal-poor RGB stars located well below the RGB tip. The co-added spectrum for this group of stars appears to have a “normal” Ca ii triplet strength, in that it lies close to the $\Sigma \text{Ca}_{\mathrm{mean}} = \Sigma \text{Ca}_{\mathrm{pred}}$ line, but it is not clear how to interpret this. If the stars above or near the RGB tip are indeed intermediate-age AGB stars, (1) $[\text{Fe}/\text{H}]_{\mathrm{phot}}$ estimates are bound to be inaccurate, as they are based on an arbitrary extrapolation of the RGB fiducials ($\approx$ 5.1); and (2) the empirical $\Sigma \text{Ca} \rightarrow [\text{Fe}/\text{H}]_{\mathrm{CG97}}$ calibration relation (eq. [2]), which is based on RGB stars, is likely to be off for AGB stars. It is conceivable that these two errors are somehow canceling each other out; a detailed investigation of these issues is beyond the scope of this paper.

Judging from their location on the CMD and the agreement between $\Sigma \text{Ca}_{\mathrm{pred}}$ and $\Sigma \text{Ca}_{\mathrm{mean}}$, it appears that about two-thirds of the stars in our M31 RGB sample are old. This includes the two bins containing general spheroid members and bins 2 and 3 from the stream population in Figure 12. If the nine stream stars located above the RGB tip in the CMD (most are in stream bin 1) and the dozen strong-lined stream RGB stars located above (or to the left of) the $\Sigma \text{Ca}_{\mathrm{mean}} = \Sigma \text{Ca}_{\mathrm{pred}}$ line (stream bin 4) all turn out to have ages $t \approx 8$ Gyr, as suspected, it would imply that about $30\%$ of the overall M31 RGB population in field a3 is of intermediate age.

In a recent deep Hubble Space Telescope study of M31 main-sequence turnoff stars in an inner spheroid field ($R = 11$ kpc), Brown et al. (2003) found a surprisingly high fraction of intermediate-age stars, $\approx 30\%$. It has been suggested that the orbit of the giant southern stream might wrap around and intersect the Brown et al. field and that this might be responsible for the high intermediate-age fraction seen there. If this explanation was correct, and taking our estimate of the field a3 intermediate-age fraction literally, the wrapped stream population would have

---

**Fig. 11.**—Montage of co-added DEIMOS spectra (thin lines) showing the region around the Ca ii triplet, normalized and shifted to rest-frame wavelength and smoothed with a 1.7 Å weighted boxcar, with each spectrum vertically offset by one unit relative to the previous one. The bottom spectrum is a co-addition of 16 foreground Galactic dwarf stars. The next five spectra from the bottom are co-additions of about a dozen M31 RGB stars each, grouped and ordered by predicted Ca ii line strength (increasing upward) as estimated from the CMD-based photometric metallicity and luminosity ($\approx 5.2$). The thick line at the top is a model red giant spectrum with $T_{\text{eff}} = 4000$ K, $\log g = 1.5$, and $[\text{Fe}/\text{H}] = -0.3$ from Schiavon & Barbuy (1999). A few prominent spectral features of RGB stars are identified along with the Na i 5890 doublet, which is strong in dwarfs. Because we have corrected the Galactic dwarf and M31 RGB spectra to zero velocity, the 8228 Å telluric feature is Doppler shifted to $\approx 8231$ and $\approx 8240$ Å, respectively.

**Fig. 12.**—Measured equivalent width of the Ca ii triplet lines in co-added spectra plotted as a function of the predicted EW. The EW $\Sigma \text{Ca}$ is defined to be the weighted sum of the EWs of the three lines ($\approx 5.2$). The $\Sigma \text{Ca}_{\mathrm{pred}}$ value for each star is derived from its CMD/RGB fiducial-based metallicity estimate $[\text{Fe}/\text{H}]_{\mathrm{phot}}$ and stellar luminosity, as described in the text. Stars are placed in groups of a dozen or more according to $\Sigma \text{Ca}_{\mathrm{phot}}$ and their spectra co-added. M31 RGB stars in the giant southern stream (pentagons) and general spheroid (squares) show reasonable agreement between predicted and measured EWs, with the exception of the strongest lined (most metal-rich) stream stars for which $\text{Ca}_{\mathrm{meas}} > \text{Ca}_{\mathrm{pred}}$. The $\text{Ca}_{\mathrm{pred}}$ calculation makes no sense for foreground Galactic dwarf stars, so it is not surprising that they fall well off the one-to-one relation (crosses).
to outnumber the smooth spheroid population 2:1 in the inner $R = 11$ kpc field (i.e., the stream’s surface brightness would have to be twice that of the smooth spheroid) just as it does in our remote field a3. No such wraparound portion of the stream is obvious in the Ferguson et al. (2002) star-count map.

In contrast to the RGB stars, the foreground Galactic dwarf stars lie well below the $\Sigma \text{Ca}_{\text{cont}} = \Sigma \text{Ca}_{\text{pred}}$ line in Figure 12. The same effect was noted by RG02. This is yet another feature that can be used to distinguish between foreground dwarf star contaminants and M31 RGB stars (§ 3.3.5). It is evident from Figure 11 that the co-added dwarf star spectrum is significantly less noisy than the M31 RGB star co-additions. This is because the former are $\Delta I \approx 1$ mag brighter on average (compare Figs. 2c and 2d), and there are 16 stars in the former co-addition versus 12 or 13 in the latter. The error in $\Sigma \text{Ca}_{\text{pred}}$ for the dwarfs is estimated to be 50% of that of the M31 RGB stars. Thus, the offset of the dwarfs from the one-to-one line is highly significant.

6. SUMMARY

The following are the main points of this paper:

1. We are using the DEIMOS spectrograph on the Keck II 10 m telescope to carry out a moderate-resolution ($R \approx 7000$) spectroscopic survey of a large sample of RGB stars in the outer bulge and halo of M31. This is the first paper from that survey and describes data from three DEIMOS slit masks in field a3, located on the giant southern debris stream discovered by Ibata et al. (2001) at a projected distance of 31 kpc from the center of M31. The field a3 data presented here (≥200 spectroscopic targets) represent only a few percent of the DEIMOS survey data obtained to date.

2. Spectroscopic targets are selected using intermediate-band DDO51 and Washington $M$ and $T_2$ photometry by discriminating between RGB stars and foreground Galactic dwarfs on the basis of surface gravity. The method has proved to increase the yield of bona fide M31 RGB stars in our sample.

3. A sample of 68 definite M31 RGB stars is isolated. Careful attention is paid to the removal of sample contaminants, both background galaxies and especially foreground Milky Way dwarf stars. The latter are identified using a combination of data and methods: radial velocity, broadband color-magnitude information, DDO51 photometry, Ca ii triplet line strength, and Na i doublet line strength, with the first and last of these being the best discriminants between M31 RGB and Milky Way dwarf stars.

4. About two-thirds of the M31 RGB stars in our field appear to be members of its giant southern stream, while the rest belong to the general spheroid population.

5. The mean heliocentric radial velocity of the stream in field a3 is $-458$ km s$^{-1}$, which translates to $-158$ km s$^{-1}$ with respect to the systemic velocity of M31. The stream has a relatively low internal line-of-sight velocity dispersion, $15^{+18}_{-15}$ km s$^{-1}$ (90% confidence limit from a maximum-likelihood analysis). The interpretation of these and other data on the kinematics and three-dimensional structure of the stream, in the context of possible orbits and progenitor properties, is presented in Paper II.

6. The rms metallicity spread of M31’s giant southern stream is 0.25 dex, and its mean metallicity is $\langle [\text{Fe/H}] \rangle = -0.51$, possibly higher if one corrects for selection bias against the highest metallicity RGB stars. This is indicative of a fairly luminous progenitor satellite galaxy. The photometric and spectroscopic metallicity estimates are in good agreement with each other for the majority of RGB stars in our sample.

7. The most metal-rich RGB stars in the stream have $[\text{Fe/H}]_{\text{spec}} > [\text{Fe/H}]_{\text{phot}}$ (anomalously strong Ca ii lines). The stream also contains a few stars that lie above the RGB tip in the CMD. Both findings suggest that the stream contains a non-negligible fraction of intermediate-age stars.

8. The general spheroid population in field a3 has a mean metallicity of $\langle [\text{Fe/H}] \rangle = -0.74$ and thus appears to be more metal-poor than the stream on average. The spheroid component has a broad metallicity distribution spanning about 1.5 dex.

9. There is a hint of spheroid substructure in M31 based on the radial velocity (and possibly metallicity) distribution of the general spheroid RGB population in this field.

We are grateful to Sandy Faber and the DEIMOS team for building an outstanding instrument and for extensive help and guidance during its first observing season. We also thank Alison Coil, Drew Phillips, and Greg Wirth for observing field a3 masks 2 and 3 on our behalf, Drew Phillips for help with slit mask designs, Ricardo Schiavon for providing a model RGB spectrum and expert advice on spectral features, Jeff Lewis and Matt Radovan in the Lick instrument shops for their careful and timely fabrication of the masks, and the DEEP2 team for allowing us use of the spec1d/zspec software. The spec1d data-reduction pipeline for DEIMOS was developed at the University of California, Berkeley with support from NSF grant AST 00-71048. P. G. acknowledges support from NSF grant AST 03-07966 and a Special Research Grant from the University of California, Santa Cruz. R. M. R. and D. B. R.’s contributions were supported by NSF grant AST 03-07931. S. R. M., R. J. P., and R. J. P. acknowledge funding by NSF grants AST 03-07842 and AST 03-07851, NASA/JPL contract 1228235, the David and Lucile Packard Foundation, and the F. H. Levinson Fund of the Peninsula Community Foundation. M. G. is supported by NASA through Hubble Fellowship grant HF-01159.01-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. K. Y. J.’s contribution was supported through NASA grant NAG5-9064 and NSF CAREER award AST 01-33617.

APPENDIX A

TECHNICAL DETAILS

A1. SLIT MASK DESIGN

A. C. Phillips’s DSIMULATOR software was used to design our DEIMOS masks (details can be found at the DEIMOS Web site). The DSIMULATOR software took as input multiple lists of spectroscopic targets: lists 1, 2, 3, etc., in order of decreasing priority. In addition, the targets in each list were assigned weights in direct proportion to their $P_{\text{grant}}$ values (§ 2.1.1). The mask pointing center and position angle were chosen manually (see thin rectangles in Fig. 1, inset, and Table 1). For each slit mask design the software started with the highest priority list (list 1) and automatically filled in the $\approx 16' \times 4'$ mask area to the extent

---

14 See http://www.ucolick.org/~phillips/deimos...ref/masks.html.
possible. It maximized the sum of weights over all selected targets subject to the following constraints: (1) a minimum length of 6" for each slitlet, (2) a minimum distance of 25" between the target and the ends of the slitlet, (3) a 0.5" gap in the spatial direction between the ends of adjacent slitlets to avoid spectral overlap, and (4) the avoidance of inter-CCD gaps and vignetted regions whose locations on the slit mask are predicted on the basis of an optical model of the spectrograph. Next, the software filled in available spaces on the slit mask with targets drawn from the next input list in the priority sequence (list 2), and so on for lists 3 and 4.

The mask 1 and 2 designs were each based on three input lists with the following selection criteria:

1. Objects with 20 < I < 22, \( P_{\text{giant}} > 0.5 \), \( \chi < 1.3 \), and \(-0.3 < \text{sharp} < +0.3 \).
2. Same as list 1 but with the I magnitude range expanded to \( I < 20 \) (above the RGB tip) and 22 < \( I < 22.5 \) (fainter RGB).
3. Objects with 20 < \( I < 22.5 \), a less stringent DDO51 constraint (\( P_{\text{giant}} > 0.25 \)), and slightly less stringent \( \chi/\text{sharp} \) morphology cuts.

After the initial round of observations in fall 2002, we decided to make some changes to the target selection procedure for our survey in order to fill the slit masks more efficiently. The design for mask 3 (and the other fall 2003 and later masks from our DEIMOS survey that are not presented here) was based on four input lists. The criteria for lists 1 and 2 were the same as those listed above except that the \( P_{\text{giant}} \) requirement was dropped altogether. The list 3 criteria were the same as those listed above except that the \( P_{\text{giant}} \) requirement was dropped altogether. The list 4 criteria were the same as for the modified list 3 but with further relaxation of the morphology cuts \( \chi \) and \( \text{sharp} \). The number of spectroscopic targets on masks 1, 2, and 3 is 85 (53/23/9), 80 (54/15/11), and 83 (49/17/16/1), respectively, where the numbers in parentheses indicate the breakdown by DSIMULATOR input list number.

The slitlet width was set to 1". In addition to these science targets, 3–5 bright stars per mask were selected manually for the purposes of slit mask alignment using the following (list 0) criteria: \( I < 20 \), the same morphology cuts as lists 1 and 2, and no \( P_{\text{giant}} \) requirement. Each alignment star was assigned a 4.0" × 4.0" box while avoiding overlap with the science target slitlets. After this step DSIMULATOR was used to maximize the lengths of all slitlets in the spatial direction while maintaining a 0.5" interslit separation. A few guide stars were also selected for each mask using the same list 0 criteria as for alignment stars; they are useful for coarse mask alignment and guiding off the TV guider camera, but no spectra were obtained for them.

A2. DATA REDUCTION

A2.1. Pipeline Processing

The DEIMOS masks were processed using the software pipeline spec2d developed by the UCB members of the DEEP2 team. The pipeline performed a number of functions. Flat-field exposures were used to rectify the curved spectra in the raw spectrogram into rectangular arrays by applying shifts in the spatial direction. Next, a one-dimensional slit function correction and two-dimensional flat-field and fringing corrections were applied to each slitlet. Using the DEIMOS optical model as a starting point, a two-dimensional wavelength solution was determined from the arc lamp exposures with residuals of order 0.01 Å. Each slitlet was then two-dimensionally sky-subtracted exposure by exposure using a B-spline model for the sky. The individual exposures of the slitlet were then averaged with cosmic-ray rejection and inverse-variance weighting. Finally, one-dimensional spectra were extracted for all science targets based on the optimal scheme of Horne (1986) and converted to plain text format using standard IRAF\textsuperscript{15} tasks.

The extracted one-dimensional spectra were processed through the spec1d pipeline developed for the DEEP2 survey at UCB (an adaptation of the corresponding SDSS pipeline). The pipeline cross-correlated the spectrum of each science target against a series of stellar templates spanning a range of spectral types and emission- and absorption-line galaxy templates to determine the redshift. The science and template spectra were continuum-subtracted and the science spectrum interpolated to the resolution of the template: \( \Delta \log \lambda = 2 \times 10^{-5} \) pixel\(^{-1} \). The cross-correlation was computed in pixel space (i.e., real space as opposed to Fourier space) with the relative line strengths and line widths held fixed. The software shifted and scaled the template to find the best fit in reduced-\( \chi^2 \) space. The galaxy templates used in the fitting procedure are linear combinations of the emission- and absorption-line templates, whereas the various stellar templates were used individually in the fit. The 10 best solutions for the redshift \( z \) of each object were reported, arranged in order of increasing reduced-\( \chi^2 \).

A2.2. Quality Assessment

The visual inspection software zspec, developed by D. Madgwick et al. at UCB for the DEEP2 survey, was used to view the sky-subtracted two-dimensional and one-dimensional spectra of each slitlet/science target. The extraction window used by the spec2d pipeline—i.e., the range of “rows” in the two-dimensional spectrum that are collapsed to form the one-dimensional spectrum—is indicated by markers along the spatial axis of zspec’s two-dimensional spectrum display. In rare cases this window appeared to be too narrow or wide or displaced from the target’s stellar continuum (or emission lines); we manually set the extraction window in these cases, reextracted the one-dimensional spectrum, and processed it through the spec1d pipeline.

The 10 best redshift choices from spec1d are listed by zspec. Selecting one of the \( z \) choices causes the corresponding template (appropriately redshifted) to be displayed overlaid on the science target’s one-dimensional spectrum. The positions of prominent absorption or emission lines in the template in question (e.g., Ca ii triplet lines for a stellar template and [O ii], [O iii], H\( \beta \), and other lines for an emission-line galaxy template) are marked on the one-dimensional and two-dimensional spectra. The night-sky spectrum, or more precisely, the variance versus wavelength, is plotted alongside the target’s one-dimensional spectrum; it proved to be very useful in deciding which spectral features are reliable and which are not. Different degrees of smoothing were tried on the target’s one-dimensional spectrum to enhance S/N at the cost of spectral resolution. This afforded a better view of its spectral features and allowed us to assess the reality of marginal/weak ones.

For about half the targets from masks 1 and 2 and two-thirds of those from mask 3 it was easy to pick out the correct \( z \) value from the choices provided by spec1d. For most of the rest, a careful inspection of the spectrum and/or redoing the fit on a segment of

\textsuperscript{15} IRAF is distributed by the National Optical Astronomy Observatory, which is operated by AURA, Inc., under cooperative agreement with the National Science Foundation.
the spectrum reveals the correct choice of redshift. On rare occasions, none of the 10 choices was accurate, but it was obvious from the one-dimensional spectrum what the correct z value was. The redshift was marked manually in these cases. It was typically based on the Ca ii triplet, often a subset of the three lines (the reddest line lies amid a cluster of near-sky lines and is sometimes affected by them), the TiO 27100 band (for red stars), and/or the Na i doublet (dwarf stars). Only 19 out of the 248 science spectra, or 17 out of the 221 unique science targets (8%), are in this “z-by-hand” category.

Following the classification scheme used in the DEEP2 survey (Coil et al. 2004), each science target was assigned a quality code Q to indicate the reliability of the measured redshift and the overall quality of its spectrum. The above cases with well-measured z values, including the z-by-hand cases, were placed in one of two categories: (1) Q = 4 for redshifts based on two or more robust spectral features, and (2) Q = 3 for those based on one robust feature and one or more marginal ones or on a few marginal features. The Q = 4 redshift measurements are expected to be “rock solid” with something like 99% confidence, whereas the Q = 3 redshift measurements are expected to have ≥90% reliability. No distinction is made between Q = 3 and Q = 4 cases elsewhere in this paper; both are treated as secure redshifts.

Objects for which the redshift measurement failed fall into three categories designated: (1) Q = −2 to indicate a catastrophic failure in the data reduction for instrumental reasons such as severe vignetting, the spectrum landing in the inter-CCD gaps, near the periphery of the CCD array, or on a bad column, poor sky subtraction, and scattered light problems; (2) Q = 1 for slitlets with a barely visible spectral continuum and/or very low S/N; and (3) Q = 2 for cases in which the S/N is marginal to adequate but there is not enough information for a reliable z determination. Judging from the DEEP2 survey, some of our Q = 2 cases are probably distant red galaxies for which the limited spectral coverage and low S/N prevented us from making a reliable redshift measurement. The low S/N in some of the Q = 1 cases in our sample, especially the handful of bright ones (Fig. 2c), could be the result of early mask design/fabrication problems. Such instrumental failures rightfully belong in the Q = −2 category, but we have not attempted to reclassify any of these cases, as there is no direct way to confirm this hypothesis.

REFERENCES

Alonso, A., Alcaino, L., & Martínez-Roger, C. 1999, A&AS, 140, 261
Bahcall, J. N., & Soneira, R. M. 1984, ApJS, 55, 67
Bellazzini, M., Cacciari, C., Federici, L., Fusi Pecci, F., & Rich, R. M. 2003, A&A, 405, 807
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
Brown, T. M., Ferguson, H. C., Smith, E., Kimble, R. A., Sweigart, A. V., Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton: Princeton Univ. Press)
Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. 2000, A&AS, 141, 371
Kulick, A., Arribas, S., & Martínez-Roger, C. 1999, A&AS, 140, 261