Mon. Not. R. Astron. Soc. **269,** 957–974 (1994)

A dynamical study of the Sextans dwarf spheroidal galaxy

J. C. Hargreaves,[1] G. Gilmore,[1] M. J. Irwin[2] and D. Carter[2]
[1]*Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*
[2]*Royal Greenwich Observatory, Madingley Road, Cambridge CB3 0EZ*

Accepted 1994 March 18. Received 1994 February 1; in original form 1993 November 19

ABSTRACT

We have observed 26 giant stars in the Sextans dwarf spheroidal (dSph) galaxy and obtained high-quality spectra for 21 of these. We have multi-epoch spectra for nine stars from which two possible binary stars have been found. The velocity dispersion of the sample calculated by a maximum likelihood method is $7.0 \pm ^{1.3}_{1.0}$ km s^{-1} and the mean velocity is 224.4 ± 1.6 km s^{-1}. There is no sign of rotation about any axis within the errors of the measurements, the formally derived value around the minor axis being 0.4 km s^{-1} at 300 pc from the axis. It is possible that binary stars may be making a significant contribution to the velocity distribution or that Sextans may be being tidally disrupted. If further work eliminates both these possibilities then, assuming isotropy in the velocity dispersion and dynamical equilibrium, the core mass-to-light ratio is 124^{+85}_{-60} $M_\odot/L_\odot$ and the total mass-to-light ratio is 121^{+84}_{-58} $M_\odot/L_\odot$, implying the presence of large quantities of dark matter.

Key words: stars: giant – galaxies: fundamental parameters – galaxies: individual: Sextans – galaxies: kinematics and dynamics – Local Group – dark matter.

1 INTRODUCTION

This paper presents the results of the first multi-epoch observations and the most accurate velocity measurements for stars in the Sextans dwarf spheroidal (dSph) galaxy. These are used to calculate a velocity dispersion and mean velocity for Sextans.

The Sextans dSph galaxy is the most recently discovered (Irwin et al. 1990) and least studied of the eight of its kind known to be in orbit around the Milky Way Galaxy. In general these galaxies contain a mixture of old and intermediate population stars and show no evidence of gas or recent star formation. They have similar total luminosities to most globular clusters (10^5–10^7 $L_\odot$) but their core radii are about an order of magnitude larger for the same luminosity.

Study of the kinematics of stars in both globular clusters and dSph galaxies by application of various simple models has been used to produce measurements for the mass-to-light ratios of these systems. The answers for globular clusters are about 2 or 3 $M_\odot/L_\odot$, whereas those for the different dSph galaxies vary between 6 and 250 (Irwin & Hatzidimitriou, in preparation), and at face value imply that all the dSph galaxies are of similar mass but contain different proportions of luminous matter.

It is important to find the smallest scale at which large quantities of dark matter exist because it places constraints on what form that dark matter might take; for example cold (or dissipative) dark matter can cluster on arbitrarily short scales while relativistic matter (e.g. neutrinos) cannot. Therefore, since dSph galaxies are the next largest scale objects after globular clusters, they are also the next largest scale on which to look for dark matter. However, a mass-to-light ratio of 250 for a dSph galaxy does seem inconsistent with the general trend of mass-to-light ratio compared with size in the rest of the Universe; for example, the value for the Milky Way Galaxy is only about 30, measured out to 80 kpc, and that for the Local Group is closer to 100. This and the wide range of measurements for the mass-to-light ratios of the different dSph galaxies lead to a question over whether the stellar kinematics are quite such simple indicators of the masses of these galaxies as is assumed by the simple models employed to date.

The rest of this paper is divided into several sections. First, the observations at the William Herschel Telescope and the reduction procedure are described. Next, the errors on the observations are discussed. Then the velocity dispersion calculation is described, the mass-to-light ratio theory explained and a value for the mass-to-light ratio obtained.

Finally, other possible contributions to the velocity dispersion are discussed.

2 OBSERVATIONS, DATA REDUCTION AND ERROR ANALYSIS

2.1 Observations

The periods of observations were the nights of 1991 May 4–7, 1991 December 27–29, 1992 April 7–9 and 26–27. All the observations were made using the William Herschel Telescope in La Palma.

The spectral range observed was 8300–8750 Å, which is the region containing the prominent calcium triplet absorption lines. We used the red arm of ISIS, the R1200R grating, a slit width of 1 arcsec with the slit aligned with the parallactic angle and an EEV 1280 × 1180 CCD as detector. The CCD was windowed to 400 pixels, which is about 2 arcmin in the spatial direction. The resulting dispersion was about 0.35 Å per pixel, and the resolution measured from sky and arc lines was twice this. Arc lamp frames were taken before and after each stellar exposure using a CuNe lamp for the 1991 May run and CuAr and CuNe lamps for the others.

The stars observed were giant branch stars ranging in brightness from 17 to 19 mag in the *R* band. Observations were made of 47 Sextans candidates, of which 26 turned out to be members. Six of these stars were already known to be members from the AAT observations of Da Costa et al. (1991). Of the Sextans members, 24 were observed more than once and 22 were observed at more than one epoch. Good spectra were obtained for 21 stars with nine of these having reasonable spectra at more than one epoch. Two of these nine stars may be velocity variables.

Additionally, four bright radial velocity standard stars (RV stars) were observed with integration times of only 5 s, one or two on each night of each run. These spectra provided an estimate of the random and systematic errors for high signal-to-noise ratio, short-exposure spectra. The random part of this error gave an estimate of the minimum random error for the Sextans data, though it appears that the RV stars may have greater systematic error due to slit centring problems. The RV stars were also used as a check on the data reduction procedure because their actual velocities were already known, and as a base to obtain the absolute mean velocity of Sextans. The coordinates of all the Sextans and RV stars observed are shown in Table 1.

Table 1. Coordinates of the Sextans stars, and of the RV stars. The centre of the Sextans dSph galaxy is at $10^h10^m.5$, $-01°22'$.

Star	RA 1950	DEC 1950	
Sextans			
1	10 09 52.1	-01 07 35	
2	10 10 8.9	-01 30 36	*
3	10 09 47.6	-01 36 58	
4	10 10 34.6	-01 21 07	
5	10 10 8.6	-01 18 01	*
6	10 10 33.9	-01 23 00	*
7	10 11 17.9	-01 30 19	*
8	10 10 38.6	-01 28 09	*
9	10 11 24.3	-01 29 51	*
10	10 11 7.3	-01 18 47	*
11	10 09 53.2	-01 23 39	*
12	10 10 27.8	-01 36 30	*
13	10 10 39.5	-01 27 16	*
14	10 10 32.9	-01 13 08	*
15	10 10 35.3	-01 29 58	
16	10 10 45.7	-01 16 21	*
17	10 10 33.0	-01 30 00	*
22	10 09 32.3	-01 22 55	
23	10 11 33.2	-01 27 47	
32	10 11 21.9	-01 02 10	
37	10 08 39.0	-01 34 01	
38	10 09 14.3	-01 38 39	
44	10 10 32.0	-01 24 21	*
45	10 10 49.8	-01 07 35	
46	10 09 18.3	-01 39 53	
48	10 09 49.1	-01 32 11	*
RV Stars			
HD908061	10 29 35.3	28 36 34	
HD92588	10 40 7.4	-01 42 45	
HD132737	14 59 38.3	27 10 55	
HD107328	12 20 4.2	03 20 35	

Notes. The asterisks indicate those stars that have also been observed by Suntzeff et al.

2.2 Data reduction

Preliminary processing of the CCD frames to remove bias and cosmic ray events was done, mostly at the telescope, using FIGARO routines. Cosmic rays were identified as local events of amplitude greater than 4σ over local signal. Each was removed by interpolation only in the spatial direction, with independent interpolation in each CCD column. This process minimizes the creation of correlated noise on scales comparable with a wavelength resolution element. Regions of strong sky lines affected by cosmic rays were excluded from use later in the reduction procedure.

The spectra from the 1991 May run were analysed independently using FIGARO and IRAF routines. Once we were satisfied that the two procedures produced equally good results, the rest of the data were reduced with IRAF, it being the faster and more flexible system. The procedure used is described in detail below.

2.2.1 *Wavelength calibration*

The NOAO.TWODSPEC.LONGSLIT programs IDENTIFY, REIDENTIFY, FITCOORDS and TRANSFORM were used to perform wavelength calibration of the spectra by identification of the arc spectral lines, calculation of two-dimensional fits to the arc spectra and transformation of these fits to appropriate stellar spectra, respectively.

IDENTIFY found 16 lines for the 1991 May CuNe arcs and 12 lines for the other CuAr+CuNe arc lamp spectra;

REIDENTIFY made three-piece spline fits along the wavelength axis every 10 lines of the CCD. The rms residuals about the fits were about 0.01 pixels which corresponds to 0.12 km s^{-1}.

Arc line fits from one or more spectra can be combined by FITCOORDS into one two-dimensional Chebyshev polynomial fit. In our case we combined the wavelength fit for the arc spectra on either side of the stellar spectrum in question unless the telescope had been moved a considerable distance in between the two arc lamp exposures. For example, if the previous star had been an RV star then one arc would have been taken at this star's coordinates and the next at Sextans' coordinates so only one arc would be used for each of these two stars. As the stellar spectrum took up just 30 lines near the centre of the CCD, most of the image was occupied only by sky spectrum and it was possible to extract this alone. The order of the fit was chosen by cross-correlating extracted sky spectra from the images of several different stars, and that order which gave closest to zero shift for all the sky spectra examined was used for the whole of that run. In the wavelength direction the fits were of third order for the 1991 May and 1991 December runs and fourth order for the 1992 April data. In the spatial direction all the data were fitted best by third order. The rms residuals about these fits were about 0.04 pixels, corresponding to 0.48 km s^{-1}. The TRANSFORM program interpolated the fits produced by FITCOORDS on to the stellar spectra using spline fits. Thus wavelength-calibrated spectra were produced.

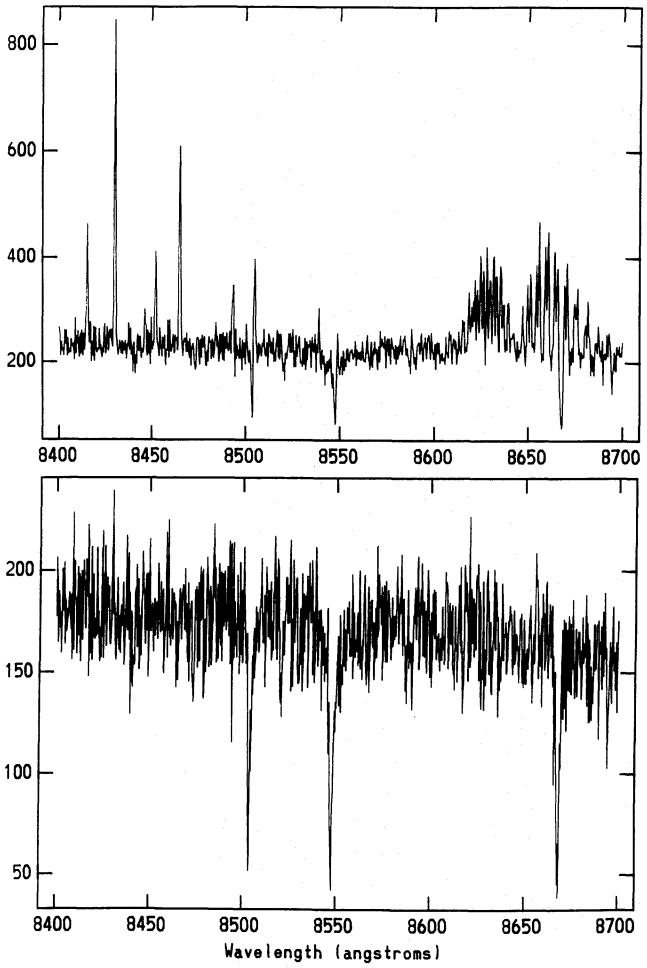

Figure 1. Comparison of a spectrum before and after sky subtraction. The lower spectrum has had the sky spectrum subtracted.

2.2.2 *Sky subtraction*

Sky subtraction was completed by the NOAO.TWODSPEC.LONGSLIT.BACKGROUND program. This calculated a second-order Chebyshev polynomial fit to each column of the image using an iterative procedure which excluded extreme pixels from the fit and interpolated across the 30 columns containing the stellar spectrum. This fitted background level was then subtracted from each column leaving a sky-subtracted image. Fig. 1 shows a typical spectrum from a Sextans member before and after sky subtraction.

The most intense sky line is at 8430 Å and its residual after sky subtraction from a spectrum of high signal-to-noise ratio is about 3.5 per cent. The data were rebinned to 8460–8700 Å, which contains the calcium triplet but little else to prevent any unavoidable sky contamination affecting the cross-correlation. If too few lines of sky spectrum were used in the sky-subtraction calculation (this could happen if the two-dimensional images were trimmed too much in the spatial direction), the interpolation across the section containing the stellar spectrum was inadequate and poor results were obtained, though this produced no systematic velocity shift. As shown by Fig. 2, which is a comparison of template and sky spectra, there are sky emission lines close to each calcium absorption line. The low signal-to-noise ratio spectra were sufficiently poor that, even after sky subtraction, they produced very uncertain velocities. Even in some higher quality spectra one or two of the calcium absorption lines could be badly affected, producing a very non-Gaussian look to the cross-correlation peak. Methods of dealing with such effects are discussed later in the cross-correlation section.

2.2.3 *Cross-correlation*

The NOAO.TWODSPEC.APEXTRACT.APEDIT program was used to extract one-dimensional spectra optimally by summing the pixels in each column within a defined aperture at each wavelength. There is a fault in the cross-correlation program FXCOR which means it is necessary to bypass the parts of the code involved in rebinning, or meaningless results are obtained. Therefore both the template and stellar spectra should be binned identically before cross-correlation is performed. The NOAO.ONEDSPEC.DISPCOR program was used to do this, rebinning the stellar spectra and the template logarithmically to 1024 bins in the wavelength range 8460–8700 Å.

The template was created by the NOAO.ARTDATA.MK1DSPEC program and consisted of the three calcium absorption lines with a one-Gaussian width of 1 Å (full width at half-maximum of 1.18 Å). This artificial spectrum was found to be a far more reliable template than either a summation of Sextans spectra or the spectrum of an RV star, because only in this template was there no other information, apart from the absorption lines, for the stellar spectra to correlate with. For example, use of a template formed by summing Sextans spectra could result in systematic velocity shifts caused by beating in velocity space between residual sky lines in the

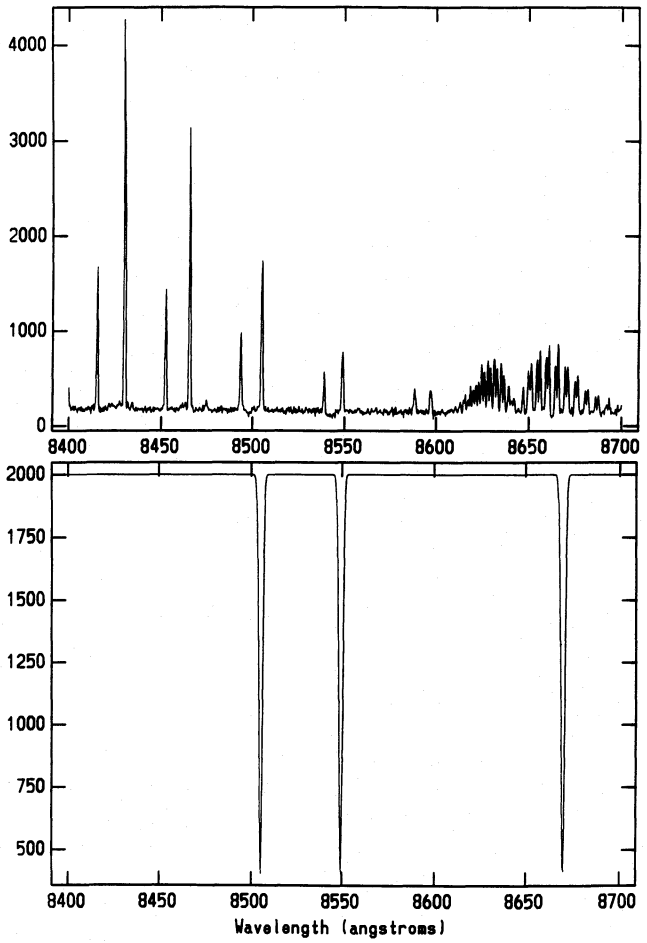

Figure 2. Comparison of sky spectrum and template spectrum. The sky spectrum is shown above the template. The existence of sky lines close to each absorption line in the template shows how crucial it was to have good sky subtraction. The middle template line has a weaker coincident sky line which explains why this line was generally the highest quality line in the reduced spectra.

template and in the star of interest. The accuracy of the template was tested by cross-correlating single lines in the template spectrum against their counterparts in the RV stars. Shifts of about 0.5 km s^{-1} between the different lines were observed.

All the stellar spectra were cross-correlated against the same template by the RV.FXCOR program. Initially this program subtracts any continuum from the spectra and then the correlation is performed. A Gaussian was fitted to the highest peak in the cross-correlation function and the centre of this Gaussian was taken to be the shift between the template and stellar spectra. This Gaussian fitting procedure performs much the same task as any filtering of the spectra, producing a smooth fit to a somewhat irregular correlation peak; it was not necessary to filter the spectra prior to the correlation as this would in no way increase the accuracy of the fit. FXCOR calculates the measured velocity shift and then the heliocentric velocity relative to the template, given sufficient information in the image headers. If the Gaussian fit was poor then the fits to individual lines were checked. If only one or two lines were contributing to the poor quality, the cross-correlation was done omitting the region containing those lines.

FXCOR produces a Tonry & Davis R value (Tonry & Davis 1979) for each correlation, and it was this parameter that was used to quantify the quality of the correlation. The noise in a stellar spectrum can distort the cross-correlation peak, causing the measured centre of the peak to be displaced from the actual velocity shift between the star and template. The R value is the ratio of the height of the cross-correlation peak at the velocity shift to the root mean square of the antisymmetric component. The resulting error in the velocity shift is related to the R value, the displacement of the correlation peak being proportional to $1/(1+R)$. Spectra producing sufficiently low-quality correlations were discounted by the means discussed in the next section.

It was sometimes possible to combine spectra of the same star with low R values to produce one spectrum of higher quality; the heliocentric shift between spectra taken at different times was calculated using the values output from the FXCOR program, the spectra were shifted by the correct amount using the IMAGES.IMSHIFT program, then both spectra were rebinned to a slightly smaller wavelength range (as was the template), and the spectra were added together by IMAGES.IMARITH and correlated in the same way as the other spectra.

Stellar spectra that on correlation with the template produced R values of less than about 5 were found to produce very different velocities for each of the three calcium absorption lines. For these very low signal-to-noise ratio spectra, differences of more than 10 km s^{-1} between lines were not uncommon. There was too much sky noise in these spectra for any meaningful result to be obtained, and although removal of lines may have increased the R value there was no objective way of using these spectra alone. Two spectra of the same star with this low an R value were not combined together, although meaningful results could be obtained from one of these spectra combined with another spectrum of higher quality.

2.3 Errors in the velocities

The error on the Sextans data was calculated using the stars for which there were repeat measurements of the velocity. An error distribution was created from the differences of the velocities obtained from individual observations for a star compared with the mean velocity of that star. The Gaussian 1σ width of this distribution was taken to be the error on an individual observation. The procedure for finding this value is described in more detail below. There were three sources of error in the results. These were random error caused by data of low signal-to-noise ratio, which is parametrized by the R value, systematic error caused by instrumental drift and imperfect centring in the slit, and real velocity variations. Further investigation of the errors was conducted by finding the error distribution for the RV stars. This distribution was less Gaussian in shape with a small central core and wide wings. These data are of much higher signal-to-noise ratio, so the width of the central core was seen as an estimate of the minimum random errors of our experiment. The wings to the distribution, which were not apparent in the Sextans error distribution, were probably caused by slit centring problems. This effect is greater in the RV star data because of the short

exposure times. Undetected real velocity variations of the stars caused by their being binary stars would have the effect of broadening the velocity distribution of the system. All three issues are discussed below.

2.3.1 *The error calculation for Sextans*

The reduction procedure described in the previous section produced the velocities and R values shown in Table 2. The R value was the best indicator of the quality of the spectra, better than the number of counts or air mass which were also inspected. According to theory described in the previous section, the inaccuracy in the position of the centre of the correlation peak rises sharply with decreasing R value. Therefore, if there is some value of R at which the accuracy is deemed to be acceptable, all the results that produced higher R values should be of even better accuracy. So a cut-off value for R whereby all spectra producing lower values were disregarded was obtained by the following method.

The difference between the velocity obtained for each observation and some mean value was required to give an idea of the accuracy of each measurement, and this could then be compared with the R value for each observation. Therefore an average velocity was calculated for each star. We shall refer to the straightforward average for each star as $\overline{V}$, and the deviation of each observed velocity from the mean for the star as ΔV. An alternative velocity is shown as $\overline{V}_{ex}$ in Table 2. Any velocity that differed from the mean (calculated excluding that velocity) by 2.5 or more times the standard

Table 2. The velocities and Tonry & Davis R values for the Sextans observations. (The columns are explained in the table footnotes and the text.)

Star	Date & No.	V_t km s^{-1}	R		$\overline{V}_{ex}$ km s^{-1}	ΔV_{ex} km s^{-1}	$\overline{V}_{7.5}$ km s^{-1}	$\Delta V_{7.5}$ km s^{-1}
1	M91 / 4 c	-0.4	6.4	*		-1.4	—	—
	M91 / 21	-0.1	12.0			-1.7		2.0
	M91 / 70	-3.0	26.9			-1.2		-0.9
	D91 / 3	-2.0	17.0			-0.2		0.1
	A92-1 / 3 c	3.7	5.8	*	-1.8	5.4	—	—
	A92-1 / 65	-2.7	21.7			-1.0	-2.1	-0.7
	A92-1 / 128	-1.9	10.7			-0.1		0.2
	A92-2 / 29	-2.2	8.5			-0.5		0.2
	comb /	-3.6	8.5	—	—	—		-0.5
2	M91 / 5	-0.5	9.1			-0.3		1.3
	M91 / 22	-1.8	10.9			-1.6		-0.0
	M91 / 71	0.4	10.5			0.6	-1.8	2.2
	D91 / 61	-9.8	11.4	*	-0.2	-9.6		-8.0
	A92-1 / 9	1.0	11.7			1.2		2.8
	A92-1 / 134	-0.2	9.4			0.1		1.7
	A92-2 / 3	-3.8	6.0	*		-3.6	—	—
3	M91 / 6 c	-4.3	7.5			-5.6	—	—
	M91 / 23	-2.4	9.1			-3.8		-2.2
	A92-1 / 13 c	9.2	7.0	*	1.4	7.8	—	—
	A92-1 / 67	-2.8	12.7			-4.2	-0.2	-2.6
	A92-1 / 136	0.7	13.3			-0.7		0.9
	A92-2 / 4	2.0	9.3			0.7		2.3
	comb /	1.4	9.8	—	—	—		1.6
4	M91 / 24	-10.9	9.5	*		-2.3		-1.6
	A92-1 / 20	-8.8	11.0		-8.6	-0.2	-9.3	0.5
	A92-1 / 138	-8.3	11.9			0.3		1.0
	A92-2 / 5	-32.8	3.4	*		-24.2	—	—
5	M91 / 25	-1.8	10.3			-0.5	-1.4	-0.5
	D91 / 15	-0.9	11.2		-1.4	0.45		0.5
	A92-2 / 6	-20.4	5.8	*		-19.0	—	—
6	M91 / 26	10.2	9.0		7.8	2.4	7.8	2.4
	M91 / 43	5.3	8.7			-2.4		-2.4
	A92-1 / 26	-5.5	12.4			0.8		0.8
	A92-1 / 69	-6.2	8.4		-6.3	0.08	-6.3	0.1
	A92-2 / 7	-7.2	9.1			-0.9		-0.9
7	M91 / 27	10.6	11.2	*		-2.2		-1.5
	M91 / 44	13.3	8.4			0.5	12.0	1.2
	D91 / 63	12.3	9.8		12.8	-0.5		0.3
	A92-2 / 8	4.9	5.6	*		-7.9	—	—
	A92-2 / 9	-37.7	1.8	*		-50.5	—	—
8	M91 / 46	11.8	8.2		2.5	9.3	11.8	
	A92-2 / 33	-6.7	7.4			-9.3	—	—
9	M91 / 73 c	0.5	6.8		-5.7	6.2	—	—
	A92-1 / 87 c	-11.8	4.1			-6.2	—	—
	comb /	-10.7	9.3		—	—	-10.7	
10	M91 / 74	-5.0	4.7		-7.5	2.5	—	—
	D91 / 17	-10.0	9.4			-2.5	-10.0	
11	M91 / 47	15.8	7.7		15.8		15.8	
12	M91 / 48	-13.9	4.2			3.0	—	—
	D91 / 33	-17.1	4.8		-16.9	-0.2	—	—
	A92-1 / 89	-19.8	3.6			-2.9	—	—
13	M91 / 49	22.8	3.7		20.9	1.9	—	—
	D91 / 5	19.0	4.4			-1.9	—	—
14	M91 / 50	-6.5	4.2		-6.1	-0.4	—	—
	D91 / 7	-5.7	11.2			0.4	-5.7	
15	M91 / 75	-22.7	1.8			-1.9	—	—
	M91 / 76	-44.8	1.8		-20.8	-24.0	—	—
	D91 / 9	-18.8	5.6			1.9	—	—
16	D91 / 19 c	2.1	5.9		-6.4	8.5	—	—
	D91 / 21 c	-14.8	4.1			-8.5	—	—
	comb /	4.6	7.7	—	—	—	4.6	
17	D91 / 23 c	-9.3	2.4		-0.0	9.2	—	—
	D91 / 25 c	9.2	6.8			-9.2	—	—
	comb /	-2.4	8.8	—	—	—	-2.4	—
22	D91 / 29	-5.4	9.8		-5.4		-5.4	—
23	D91 / 35	9.5	7.4		5.5	4.0	—	—
	A92-2 / 11a	1.4	9.92			-4.0	1.4	
32	D91 / 53	-7.3	12.0		-6.4	-1.0		-1.0
	A92-2 / 30	-5.4	12.7			1.0	-6.4	1.0
37	A92-1 / 7	0.3	11.4		-0.4	0.7		0.7
	A92-2 / 31	-1.1	11.9			-0.7	-0.4	-0.7
38	A92-1 / 11	-1.9	7.3		-3.9	2.0	—	—
	A92-2 / 32	-5.9	10.3			-2.0	-5.9	
44	A92-1 / 75	2.2	8.7		0.8	1.5		1.5
	A92-2 / 34	-0.7	8.2			-1.5	0.8	-1.5
45	A92-1 / 77	-12.4	4.8		-11.1	-1.2	—	—
	A92-2 / 35	-9.9	8.5			1.2	-9.9	—
46	A92-1 / 79 c	-0.8	5.5		-4.2	3.5	—	—
	A92-2 / 36 c	-7.7	7.1			-3.5	—	—
	comb /	2.2	7.7	—		—	2.2	—
48	A92-1 / 83	-12.6	3.4		-11.2	-1.4	—	—
	A92-2 / 37	-9.9	3.6			1.4	—	—

Notes. Date: M91, D91, A92-1, A92-2 are abbreviations for the 1991 May, 1991 December and 1992 April runs, A92-1 being the run at the start of April and A92-2 the one at the end.
No: this is the run number for the particular CCD frame on tape, so with the date acts as an identification number.
V_t: this is the heliocentrically corrected velocity with respect to the template.
R is the Tonry & Davis R value.
$\overline{V}_{ex}$: this is the average for each star of all the values of V_t excluding those that caused a change in the mean of more than 2.5 standard deviations.
Star 6 is probably a binary so two averages at the different observing times were calculated in order that the data from this star could still be used in the analysis of the errors.
$\overline{V}_{7.5}$: this is the average velocity for a star where data that produced a correlation with $R < 7.5$ are not included, and $\Delta V_{7.5} = (V_t - \overline{V}_{7.5})$.
The observations not included at the cut-off are marked by a dash.

deviation of the other velocities was left out of this calculation. Those observations excluded from the average are marked in the table by asterisks. The deviations of the individual velocities from this mean are shown by ΔV_{ex} in the table, and are more extreme for the more deviant velocities than the values of ΔV for the same observations. Star 6 had velocities measured on two different observing runs whose averages differed by 14 km s^{-1}, so this star was therefore assumed to be variable and was not included in the velocity dispersion calculations. It was also possible that star 8 was variable, although with only one observation at each of two observing runs and generally lower quality spectra this was not as likely, so this star was still included in the analysis. The relative probabilities of these stars being variable are considered later in the section discussing the relevance of binary stars. The values of ΔV_{ex} were grouped according to the appropriate value of R as shown in Table 3. When there were only two spectra of a star, both velocities were put in the bin corresponding to the lower R value. Each band of R had unit width, and the mean and standard deviation of the values of ΔV_{ex} in each band were calculated, leading to the values presented in Table 3. The standard deviations were then plotted against the R values for all the data, giving a representation of the accuracy variation with R. This plot is displayed in Fig. 3, and illustrates the way the inaccuracy rises as expected with decreasing R. For R greater than about 8 the standard deviation is below about 3, but it is substantially higher for lower R values. A suitable cut-off value (R_{cut}) was taken to be somewhere between 7 and 8. All the results were calculated for R_{cut} values of 7, 7.5 and 8, although there was little difference between these three when compared with the case when no R threshold was imposed on the data.

Table 3. ΔV_{ex} from Table 2 binned into values of R.

R	0.5–	1.5–	2.5–	3.5–	4.5–	5.5–	6.5–	7.5–	8.5–	9.5–	10.5–	11.5–	12.5+
ΔV_{ex}		-50.5	-24.2	3.0	2.5	-1.4	7.8	-5.6	-0.5	-2.3	-0.1	-1.7	-4.2
		-1.9	-1.4	-2.9	-2.5	5.4	9.3	0.1	-0.3	-0.5	-1.6	1.2	-0.7
		-24.0	1.4	1.9	-1.2	-3.6	-9.3	1.5	0.1	-0.5	0.6	0.8	-1.2
		9.2		-1.9	1.2	-19.0	2.0	-1.8	-3.8		-9.6	-1.0	-0.2
		-9.2		-0.4	3.5	-7.9	-2.0	0.5	0.7		-0.2	1.0	1.0
				0.4	-3.5	1.9	4.0		2.4		-0.7	0.3	
				8.5	-0.2		-4.0		-2.4		0.5		
				-8.5					-0.9		-2.2		
				6.2							0.7		
Mean		-15.3	-8.1	0.7	-0.0	-4.1	1.1	-1.0	-0.6	-1.1	-1.4	0.1	-1.5
St.Dev.		23.1	14.0	5.0	2.6	8.6	6.6	2.8	1.9	1.1	3.2	1.2	1.6

Notes. Each correlation produces a velocity and a Tonry & Davis R value. The velocities here are the same as ΔV_{ex} in Table 1.
Mean and St.Dev. are the mean value and standard deviation for each column.

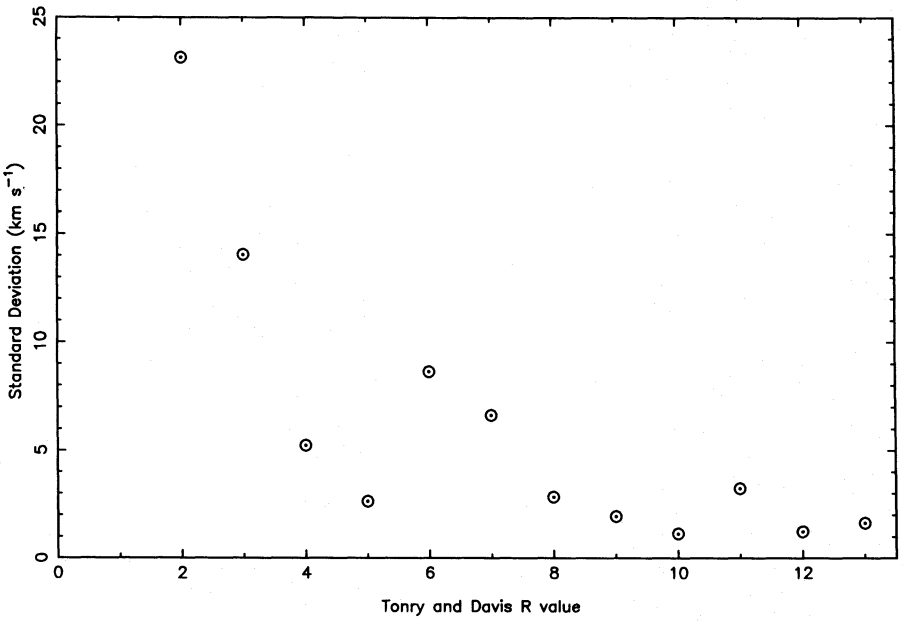

Figure 3. The standard deviations from Table 3 compared with the Tonry & Davis R value; the steep increase in the standard deviation with decreasing R at low R is as expected for an accuracy versus R diagram. The deduced cut-off value for R is between 7 and 8.

Having found a cut-off value of R it was then necessary to calculate the measuring error on each observation in the remaining sample. Up to this point all the observations had been taken in isolation. It was possible, however, to combine two spectra whose R values fell below R_{cut} to create a single spectrum with a sufficiently high R value for it to be used in the subsequent calculation. If the R values for the single observations were above the threshold R_{cut} then these observations were used separately, but if they fell below it and that of the combined spectrum fell above it then the combination was used instead. The results for an R_{cut} value of 7.5 are shown in Table 2 where the 'c's in the 'No.' column mark spectra that were combined to make the 'comb' spectra. Observations excluded at each value of R_{cut} are marked by dashes.

The standard deviations in Table 3 are basically the measuring errors for the observations at each value of R, so the standard deviation of the values of ΔV for the range of R greater than R_{cut} gives the measuring error of the sample used. This is equivalent to finding the width of a Gaussian fitted to the distribution of ΔV, so the assumption being made is that the errors were Gaussian, which is a point that will be returned to later. The calculation was made for each value of R_{cut} yielding a width σ_{err}^2, and the error on this width was ε_1, where $\varepsilon_1 = \sigma_{err}^2\sqrt{2/N}$. The values of σ_{err} obtained for R_{cut} values of 0, 7, 7.5 and 8 were 7.2 ± 0.6, $3.2 \pm {}^{0.3}_{0.4}$, 1.9 ± 0.2 and 1.9 ± 0.2 km s^{-1}, respectively. Kolmogorov–Smirnov (K–S) tests were performed on each fitted Gaussian to see if the errors were indeed consistent with Gaussian distributions. For R_{cut} values of 0, 7, 7.5 and 8 the probabilities of a Gaussian distribution were 0.006, 0.4, 0.7 and 0.7. So the errors are consistent with Gaussian errors except in the case where no threshold is imposed on R. In this case the errors would not be expected to be simply Gaussian because of the very much lower accuracy at the lower values of R making the distribution a summation of different Gaussians which would itself not be Gaussian in shape. The errors for R_{cut} of 7 have a lower probability of being Gaussian than those for the other two higher cut-offs, probably for the same reason. Fig. 4 shows the error distribution and fitted Gaussian for R_{cut} of 7.5.

2.3.2 *Comparison with the RV stars*

The RV stars were used to check the reduction procedure and to find a lower limit on the random errors. The velocities obtained for these stars are listed in Table 4. The reduction procedure and error calculation were identical to those used for the Sextans data so the results from the two sets could be compared. All the R values obtained for the correlations of these spectra were well above the threshold region, with an average value being around 17. The average velocities of the stars were within 1 km s^{-1} of the known velocities, verifying that the reduction procedure had worked well. The distribution of the differences from the mean velocities, ΔV, had a width of 2.8 km s^{-1} with a probability of being a Gaussian of 0.5, which is somewhat lower than the probability for the Sextans results for R_{cut} of 7.5 or more. Looking at the distribution of ΔV, extended wings can be seen in addition to a Gaussian-like shape close to the centre. This shape is probably due to the star not always being placed centrally in the slit. An experiment to test this was conducted during the 1992 April run whereby consecutive exposures of the same RV star were made with the star in the top half, middle and bottom half of the slit as it was viewed on the television screen. The 'top', 'middle', and 'bottom' readings in Table 4 show the velocities obtained. A difference of more than 6 km s^{-1} was obtained between the two sites of the slit, and this is sufficient to account for the wings to the distribution. This would be a systematic error, with no signal-to-noise ratio dependence. For the Sextans stars, however, the effect of positioning in the slit should be considerably less because of

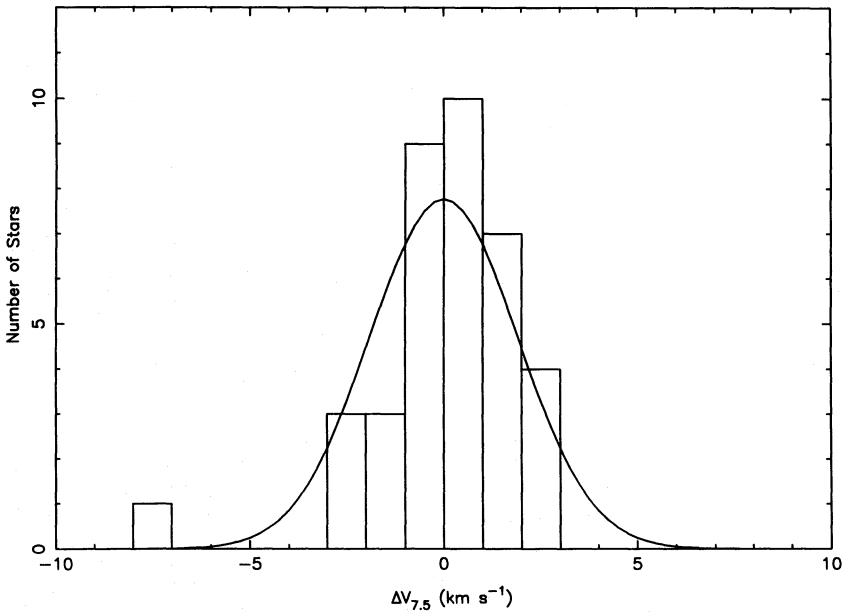

Figure 4. The Gaussian fit to the error distribution of the Sextans stars for R_{cut} of 7.5 with width 1.9 km s^{-1}. The K–S test of this fit produced a probability of 0.7.

Table 4. Velocities for the RV standard stars.

Star	Date	V_t km s^{-1}	$\overline{V}$ km s^{-1}	ΔV km s^{-1}	V_A km s^{-1}
HD908061	M91	36.7		0.5	
		37.6		1.4	
		36.0		-0.3	
		38.7		2.4	
	D91	30.6	36.2	-5.6	36.3
		42.4		6.2	
	A92	36.6		0.4	
		29.3		-7.0	
		36.7		0.5	
		37.8		1.5	
HD92588	M91	43.9		2.0	
	D91	38.5		-3.5	
		41.6	41.9	-0.3	42.8
	A92	42.3		0.3	
		43.4		1.5	
HD13273	M91	-19.9		4.0	
		-22.0		1.9	
		-26.7		-2.8	
	A92	-24.9	-23.9	-1.0	-24.1
		-24.1		-0.2	
		-24.7		-0.8	
	top	-21.4			
	middle	-25.2		-1.2	
	bottom	-27.6			
HD107328	M91	35.6	35.5	0.1	35.7
	A92	35.4		-0.1	

Notes. The notation here is the same as in Table 2; $\overline{V}$ is the mean velocity for each star and ΔV the difference of the velocity of each star from this mean. The values of V_A are the actual velocities of the radial velocity standard stars.

the motion of the star in the slit due to the guiding errors of the telescope during the exposures which are 180 times longer than those for the RV stars. So to get some idea of the minimum possible random error, which was signal-to-noise ratio dependent, a Gaussian was fitted to the centre of the distribution. The width was 1.3 km s^{-1} with a higher probability from the K–S test of 0.9. This is considerably less than the errors for the Sextans data, as would be expected for spectra of lower signal-to-noise ratio, so that, no matter how good the spectra obtained by similar observations are, the errors will not be less than 1.3 km s^{-1} for each observation. The effect of slit positioning in the Sextans data may not be negligible, but, as this would be expected to show up in the error distribution deviating from Gaussian in shape and no deviation of this sort is obvious, no contribution of this to the errors is included in the results in this paper. However, it is important to bear in mind that the effect of this systematic error would be to create wings to the velocity distribution of the galaxy, probably increasing the value calculated for the velocity dispersion. The error distribution for the RV stars, along with both Gaussian fits, is shown in Fig. 5, and all the results from the error calculations are displayed in Table 5.

3 RESULTS

3.1 The velocity dispersion calculation

The observations that had produced R values above the threshold were used to calculate a mean velocity for each star. The width of the distribution of these velocities defines the velocity dispersion (σ_{obs}) of Sextans. An unweighted Gaussian fit to the data and a weighted calculation were both made, producing slightly different results. All the velocity dispersion results, weighted and unweighted, for R_{cut} values of 0, 7, 7.5 and 8 are displayed in Table 5. The results mentioned in the text are those for an R_{cut} value of 7.5.

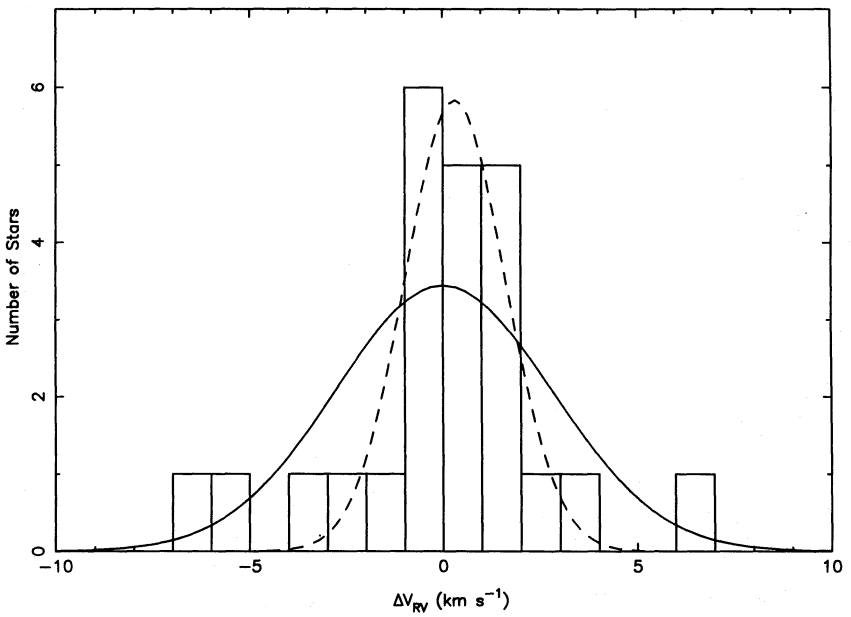

Figure 5. The Gaussian fit to the error distribution of the RV stars. The wider Gaussian is the fit to all the data and is of width 2.8 km s^{-1} with a probability of 0.5 for the fit. The narrower Gaussian is the fit for the distribution excluding the five most extreme velocities in the distribution. It is of width 1.3 km s^{-1} and the probability for the fit is 0.9. The relevance of the two different fits is explained in the text.

Table 5. Widths of the error distributions and the velocity dispersions calculated for different values of R_{cut}.

Distribution	$\bar{v}$ km s^{-1}	σ km s^{-1}	Error km s^{-1}	Prob	N
Sextans error calculation					
all R	0.1	7.2	± 0.6	0.006	75
$R \geq 7$	-0.03	3.3	$\pm^{0.3}_{0.5}$	0.4	44
$R \geq 7.5$	-0.0003	1.9	$\pm^{0.2}_{0.3}$	0.7	37
$R \geq 8$	-0.0003	1.9	$\pm^{0.2}_{0.3}$	0.7	37
RV stars error calculation					
all results	0	2.8	± 0.4	0.5	24
core	0.3	1.3	± 0.2	0.9	19
Sextans velocity dispersion					
Unweighted calculation (measuring errors included)					
all R	-3.8	8.7	$\pm^{1.2}_{1.3}$	0.7	25
$R \geq 7$	-1.7	6.7	$\pm^{1.0}_{1.1}$	0.9	21
$R \geq 7.5$	-1.1	7.3	$\pm^{1.0}_{1.2}$	0.8	21
$R \geq 8$	-2.5	6.4	$\pm^{1.0}_{1.2}$	0.8	18
Armandroff & Da Costa weighted calculation					
all R	-4.2	7.8	$\pm^{1.0}_{1.2}$	0.8	25
$R \geq 7$	-1.2	5.5	$\pm^{0.8}_{0.9}$	0.7	21
$R \geq 7.5$	-1.1	6.0	$\pm^{0.9}_{1.0}$	0.8	21
$R \geq 8$	-1.7	5.6	$\pm^{0.9}_{1.0}$	0.5	18
Maximum Lilkelihood calculation					
all R	-4.0	8.1	$\pm^{1.7}_{1.4}$	0.8	25
$R \geq 7$	-1.6	6.2	$\pm^{1.3}_{1.1}$	0.9	21
$R \geq 7.5$	-1.1	7.0	$\pm^{1.3}_{1.0}$	0.9	21
$R \geq 8$	-2.5	6.1	$\pm^{1.3}_{1.0}$	0.8	18

Notes. The value $\bar{v}$ is the average value calculated by the fit to the distribution.
The value of σ is the width of the distribution; so in the case of the error distribution it is the value of the error and in the case of the velocity dispersion calculation it is the velocity dispersion.
Error is the error in the value of σ calculated as described in the text.
N is the number of stars in the distribution.
The value 'Prob' is the probability obtained by a K–S test comparing a Gaussian distribution with the calculated width and average to the actual distribution of data.
The differences between the different calculations (weighted and unweighted) are described in the velocity dispersion section.

The velocity dispersion obtained by fitting a Gaussian to the unweighted distribution was $7.4 \pm^{1.1}_{1.3}$ km s^{-1} for an R_{cut} value of 7.5. The variance of σ^2_{obs} is $2\sigma^4_{obs}/N$, so this is the error quoted. This dispersion has not had the contribution due to measuring errors removed. The inclusion of these, as in equation (B9) of Appendix B, gives a corrected velocity dispersion of $7.3 \pm^{1.0}_{1.2}$ km s^{-1}.

From a theoretical point of view, the most accurate calculation should use the average velocity of each star weighted in some way to take account of its measuring error. The method used over the last few years for such a calculation was that of Armandroff & Da Costa (1986). The detail of this method is given in Appendix A. The result from this weighted calculation for an R_{cut} value of 7.5 was $6.0 \pm^{0.9}_{1.0}$ km s^{-1}. The K–S tests performed on all the results for the different R_{cut} values show consistency with Gaussian distributions.

We have used a different weighted calculation using a maximum likelihood estimator to find simultaneously the velocity dispersion and mean velocity ($\bar{v}$) of the sample. Fig. 6 shows the log likelihood function contoured in the $\bar{v}$, σ^2_{obs} space. The 1σ measuring errors are the first contour, where the log likelihood has fallen by 0.5 from its maximum. The theory is explained in Appendix B. Using this method the velocity dispersion for an R_{cut} value of 7.5 was $7.0 \pm^{1.3}_{1.0}$ km s^{-1}. For a sample where the velocity dispersion is considerably larger than the measuring errors, this result would be expected to be similar to the unweighted calculation, as explained in Appendix B. It can be seen, referring to Table 5, that this was indeed the case for our results.

We recommend use of the maximum likelihood estimator, with appropriate functional form for the intrinsic velocity dispersion, since not only does this give the 'best' estimate of σ_{obs} but also it enables reliable confidence limits to be readily determined. The Armandroff & Da Costa method gives too much weight to repeat measurements of the same stars, because the intrinsic dispersion of the sample is not allowed for. In cases where some stars are observed many more times than others (as is common) this will cause the true velocity dispersion to be seriously underestimated. This effect can clearly be seen in Table 5.

The value of the velocity dispersion adopted for later calculations is 7.0 km s^{-1}, which was that for the maximum likelihood calculation at an R_{cut} value of 7.5. The cut-off value of 7.5 was used because the measuring errors are low, Gaussian with a high probability, without losing too many stars from the sample. Fig. 7 shows the Gaussian with this width and the velocity distribution for an R_{cut} of 7.5.

3.2 Comparison with previous results

The only other accurate measurements of the velocities of stars in the Sextans dSph galaxy are by Suntzeff et al. (1993). They measured velocities for 43 stars in Sextans during the period 1991 March 25–27 and obtained a velocity dispersion of 6.2 ± 0.9 km s^{-1}, analysing their results using the Armandroff & Da Costa method, and only using their most accurate 33 stars. We have 15 stars in common; for three of these our data were too poor to produce a reliable velocity and one is the variable star 6. Their observations were only 1 month before our first observations, and the velocity they obtained for star 6 is, within the errors, the same as the velocity we obtained during the 1991 May run, adding weight to our suggestion of its variability. Suntzeff et al. find their velocity measurement of star 5 (their star 17) statistically improbable for a Gaussian distribution, and it is interesting to note that our result for this star is quite reasonable in itself but the two velocities measured by the different groups are not within each other's errors. They have several observations of each star so were able to produce an individual error estimate for each star, and their measurements of the stars observed by both of us have errors between 2.4

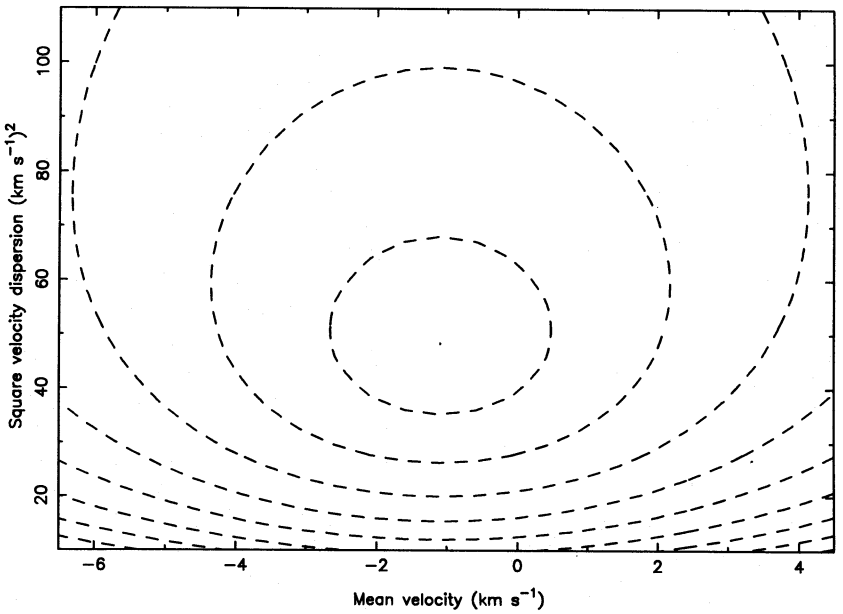

Figure 6. The log likelihood function described in Appendix B, for the data with an R_{cut} value of 7.5. The central dot is at the maximum of the function, and the contours are analogous to the 1σ to 8σ errors.

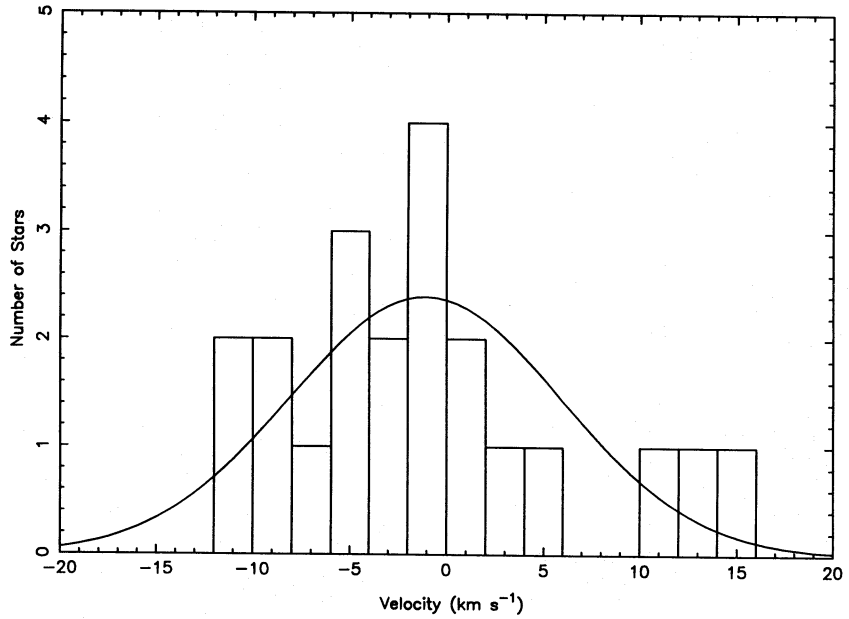

Figure 7. The velocity distribution of the Sextans dSph galaxy with R_{cut} of 7.5. There are 21 stars in the sample and the Gaussian shown is that produced by the error-weighted calculation. The velocity dispersion is 7.0 km s^{-1} and the average velocity -1.1 km s^{-1} with respect to the template.

and 7.4 km s^{-1}. Using the analysis described by Godwin & Lynden-Bell (1987) to compare our results with those of Suntzeff et al., and taking an average error estimate for the results of Suntzeff et al. of 5.2 km s^{-1}, between us we have underestimated the measuring errors by a factor of 1.3. With only two sets of data it is not possible to say what the real errors are for each group, though since our errors are much smaller it is highly likely that the underestimate of the errors of Suntzeff et al. comprises most of the discrepancy; they would have to increase their errors by a much smaller factor to make up the difference. If the discrepancy were all due to us, our errors would have to be as large as theirs. Fig. 8 shows our results plotted against those of Suntzeff et al. obtained for the 12 stars for which we both have velocities, with our own predicted errors plotted as the error bars.

3.3 The absolute mean velocity of Sextans

The other result obtained was that for the systemic velocity of the Sextans dSph galaxy. From the RV standard stars, the

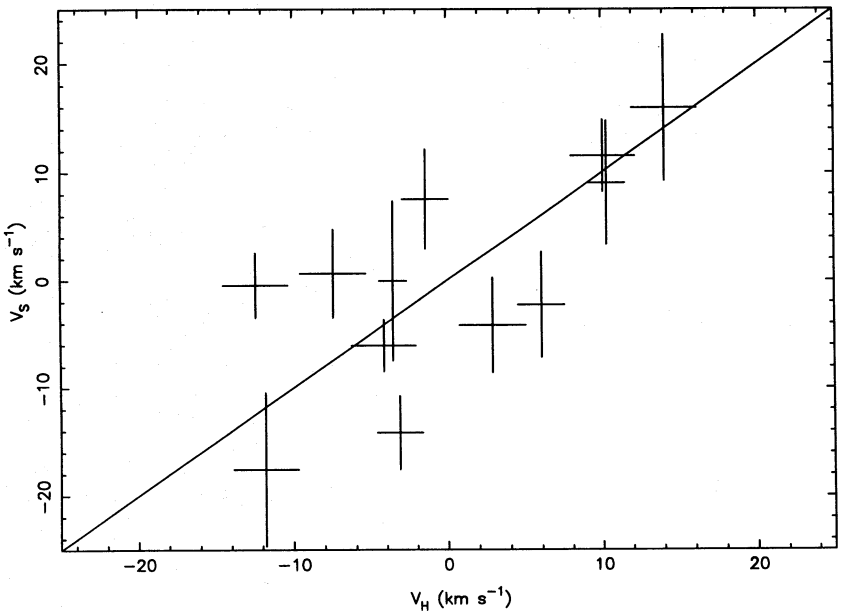

Figure 8. The velocities obtained from our data compared with those from Suntzeff et al. for the 11 stars for which we both have data. The V_S and V_H are the velocities from Suntzeff et al. and our velocities, respectively, and the error bars show the upper and lower limits of each star's velocity according to the errors calculated by the two groups. The rectangles described by the error bars should intersect the straight line drawn on the diagram if the two sets of data were consistent; the error underestimate is a factor of 1.3.

velocity of the template was found to be 223.3 km s^{-1}. The maximum likelihood method output values for the velocity dispersion and the mean velocity of the sample simultaneously. For a value of R_{cut} of 7.5 the mean velocity of the sample with respect to the template was -1.1 ± 1.6 km s^{-1}. The error here is derived from the log likelihood plot in the same way as the velocity dispersion (see Appendix B). This gave the velocity of Sextans as 224.4 ± 1.6 km s^{-1}, in line with the previous values of 230 ± 6 km s^{-1} (Da Costa et al. 1991), and 227.9 ± 1.8 km s^{-1} (Suntzeff et al. 1993).

4 ANALYSIS

4.1 Mass-to-light ratios

4.1.1 Theory

As discussed by Richstone & Tremaine (1986), the central mass-to-light ratio of a system may, dimensionally, always be written

$$\frac{\rho_0}{I_0} = \eta \frac{9\sigma_0^2}{2\pi G r_{hb} S_0}, \tag{1}$$

where $I_0 \approx S_0/2r_{hb}$, S_0 is the central surface brightness, r_{hb} is the half-brightness radius and σ_0 is the measured central velocity dispersion. The factor η is given by

$$\eta = \frac{4\pi}{9} r_{hb} \frac{[\int_0^\infty \rho(r)\,dr]^2}{\int_0^\infty M(r)\rho(r)/r\,dr}, \tag{2}$$

so is dependent on the model fitted to the density profile. It turns out, however, that, for all simple models producing realistic profiles, η is always close to 1. The equation for ρ_0/I_0 with η set to 1 is called King's method of core fitting; it was first derived by King in 1965 (see Richstone & Tremaine 1986) without use of a King model. This method is clearly independent of any model used to fit the density profile, all the variables being measurable quantities. This is not quite true for dSphs, however, because of the difficulty of defining the central brightness, and hence the half-brightness, without first fitting a model to the data. Therefore a King model has been fitted to the surface brightness profile (Irwin & Hatzidimitriou, in preparation). From this the half-brightness radius was calculated.

King models are lowered isothermal spheres with velocity distributions given by

$$f(v) = \begin{cases} k[e^{\beta(E-E_0)} - 1], & E < E_0, \\ 0, & E > E_0, \end{cases} \tag{3}$$

where $E = v^2 + \Phi(r)$, $\Phi(r)$ is the potential, and k, β, and E_0 are constants. These models are parametrized by $W_0 = \beta[e_0 - \Phi(0)]$ or by the concentration, $c = \log(r_t/r_c)$ where r_t is the tidal cut-off radius and r_c is the core radius which King originally defined empirically (King 1962). As σ_0 is the measured central velocity dispersion, we shall call $\sigma(r)$ the central velocity dispersion as a function of distance from the centre of the galaxy. The value of β changes with the different models, as is shown by the plot in Binney & Tremaine (1987) which relates $\sqrt{\beta\sigma^2(r)}$ to distance, r, from the centre of the galaxy. Rather confusingly, their σ^2 is our $1/\beta$, whereas (assuming isotropy of the velocity dispersion) our $\sigma^2(r)$ is equal to one third of the mean square velocity which, in Binney & Tremaine's notation, is $\bar{v}^2/3$.

The central density of a King model (King 1966) is

$$\rho_0 = \frac{9}{4\pi G \beta r_c^2}. \tag{4}$$

For $W_0 \geq 5$, β is close to σ_0^2 and r_c is close to r_{hb}, giving some feeling as to why this equation is equivalent to King's method of core fitting. In the same paper, King also developed an expression for the total mass of a system fitted by a King model. For different values of W_0 he calculated values for μ, where μ is the integral

$$\mu = \int_0^{R_t} \frac{\rho}{\rho_0} 4\pi R^2 \, dR. \tag{5}$$

Here $R = r/r_c$ and $R_t = r/r_c$. It therefore follows simply that

$$M_{tot} = \rho_0 \mu r_c^3 = \frac{9 r_c \mu}{4\pi G \beta}. \tag{6}$$

So the total mass-to-light ratio for a particular King model is given by

$$\frac{M_{tot}}{L_{tot}} = \frac{9 r_c \mu}{4\pi G \beta L_{tot}}. \tag{7}$$

This method (first used by Illingworth (1976) and hereafter called Illingworth's method) is more sensitive to the King model fitted to the density distribution than is the core fitting method (where the value of the η correction is 0.96–1.01 for W_0 from 2.5 to 11). It seems logical that this should be the case because to obtain a total mass must involve integration of the density over the whole galaxy, which must be a more model-sensitive procedure than using only core parameters. For the two methods it would therefore be expected that the core fitting procedure is the more accurate. It should also be emphasized that Illingworth's method is only applicable to a King model, whereas King's method of core fitting is applicable to all reasonable models.

When it comes to relating these equations to actual observations, there is another point which has to be considered. Both σ_0^2 and $1/\beta$ are values calculated through the centre of the galaxy, and, as can be seen from the plot in Binney & Tremaine previously referred to, the measured line-of-sight velocity dispersion varies with distance from the centre of the galaxy. Now, in the case of real observations, all the stars looked at cannot be at the centre of the galaxy, particularly in such a diffuse system as Sextans, so the average distance of the stars observed was taken as the radius for the velocity dispersion measurement, σ_{obs}; the corrections at this radius can be read off the plot for the appropriate King model. Therefore $1/\beta$ can just be read off, and σ_0^2 is equal to one third of the mean square velocity at the centre for the particular value of W_0.

Putting equations (1) and (7) into units of $M_\odot/L_\odot$, with r_c in pc and the velocity dispersions in km s^{-1}, the factor $9(4\pi G)^{-1}$ becomes 166.5, so

$$\frac{\rho_0}{I_0} = \eta \frac{333 \sigma_0^2}{r_{hb} S_0}, \tag{8}$$

and

$$\frac{M_{tot}}{L_{tot}} = \frac{166.5 r_c \mu}{\beta L_{tot}}. \tag{9}$$

4.1.2 *Results*

As already explained, the observed velocity dispersion would be expected to decrease with distance from the centre of the dSph galaxy. Table 6 shows how the velocity dispersion varies with radius for our results. Beyond 400 pc it does appear to drop off considerably, down from 8 to 3 km s^{-1}. The values of 'No.' in the table give the number of stars within each range of radius. The dispersions and errors here were calculated as before, using a maximum likelihood estimator. Within the errors this drop-off of velocity dispersion with radius is consistent with the range of possible King models. Irwin & Hatzidimitriou (in preparation) have found that the best King model fit to Sextans has a concentration, $c = 0.98$, which is equivalent to $W_0 = 4.8$. This model has the parameters $\eta = 0.98$ and $\mu = 11.4$. Fig. 9 shows this model fitted to the photometric data. The brightness profile was obtained from the density in stars per unit area in the following way. It was assumed that surface brightness was proportional to surface density; in such a diffuse system there is unlikely to be mass segregation so this is a reasonable assumption. Therefore all that was required was the constant of proportionality. This was obtained by integrating the fitted profile of the number density distribution (in this case the King model) such that the integral was equal to the total magnitude of the dSph galaxy.

Irwin & Hatzidimitriou also calculated the following parameters for the dSph galaxy:

$r_c = 312 \pm 41$ pc;
$r_{hb} = 285 \pm 37$ pc;
$r_t = 3003 \pm 995$ pc;
$M_V = -9.0 \pm 0.5$;
$L_{tot,V} = (3.4 \pm {}^{2.0}_{1.3}) \times 10^5 \, L_\odot$;
$S_{0,V} = 0.5 \pm {}^{0.3}_{0.2} \, L_\odot \, pc^{-2}$.

All the distances quoted here are geometric mean distances. The average distance from the centre of the galaxy of our observations was 310.0 pc, which is very close to a core radius, leading to $\sigma_0 = \sigma_{obs}/0.93$ and $1/\beta = \sigma_{obs}^2/0.84^2$; throughout the following calculations $\sigma_{obs} = 7.0^{+1.3}_{-1.0}$ km s^{-1}. Calculation of the core mass-to-light ratio using equation (8) gives $\rho_0/I_{0,V} = 124^{+85}_{-60}$. Similarly the total mass-to-light ratio calculation of equation (9) produces $M_{tot}/L_{tot,V} = 121^{+84}_{-58}$. The errors quoted here include those due to the half-brightness radius, the luminosity and the velocity dispersion, with the luminosity error contributing more than half the total error. Not only is there considerable uncertainty in the measurement of the total luminosity and central surface brightness of Sextans, but the values obtained by different people are not consistent. The problem is that on the survey

Table 6. Variation of velocity dispersion with radius.

Radius pc	No. in bin	Velocity Dispersion. km s^{-1}
28-168	5	$6.7 \pm {}^{1.9}_{2.7}$
194-243	5	$8.5 \pm {}^{2.4}_{3.4}$
314-377	5	$8.2 \pm {}^{2.3}_{3.2}$
404-612	6	$2.6 \pm {}^{0.7}_{0.9}$

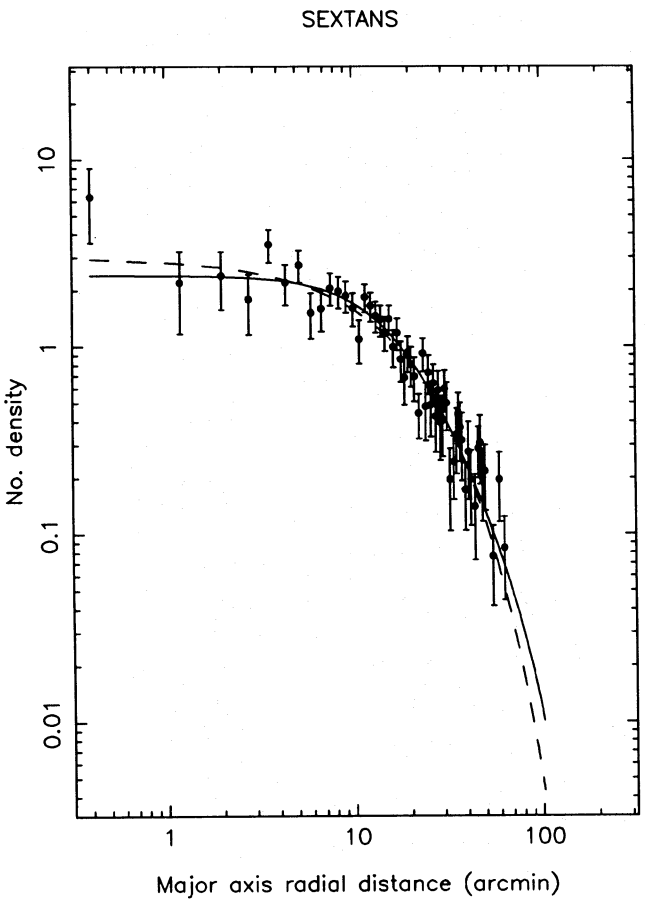

Figure 9. The best King model fitted to the Sextans dSph galaxy (by Irwin & Hatzidimitriou in preparation, from the APM results) with a concentration of 0.98. 1 arcmin is equivalent to 24 pc at the distance of Sextans (83 kpc).

plates there are only about twice the number of Sextans members as field stars, so, depending on what the background luminosity is calculated to be, there is large variation in the luminosity obtained for Sextans. This is particularly true close to the centre of the galaxy where, because of the small area, there are correspondingly fewer stars, leading to larger errors. For example, Caldwell et al. (1992) have also calculated the luminosity parameters for Sextans. The results were $M_V = -10.0$, $L_{\mathrm{tot},V} = 8.2 \times 10^5$ $L_\odot$, and $S_{0,V} = 1.5$ $L_\odot$ pc^{-2}. Using equations (8) and (9) as before, the mass-to-light ratios obtained are $\rho_0/I_{0,V} = 41$ and $M_{\mathrm{tot}}/L_{\mathrm{tot},V} = 50$.

There is also an error, not included so far, caused by uncertainty in the King model fitted to the data. As an example, taking a probable lowest value for the concentration of $c = 0.8$, and Irwin & Hatzidimitriou's measurements for luminosity and core radius, all the model parameters are tweaked to produce $\rho_0/I_{0,V} = 134^{+92}_{-51}$ and $M_{\mathrm{tot}}/L_{\mathrm{tot},V} = 103^{+67}_{-51}$. At this value of the concentration, the value of μ is quite variable depending on the exact concentration, and the value of β varies considerably with a small change in average radius, leading to lower reliability in Illingworth's method compared with core fitting.

So, taking all these things into consideration, treatment of Sextans as a simple lowered isothermal sphere produces a mass-to-light ratio of greater than 40 at the 1σ level. For the mass-of-light ratio to be 3, and with Irwin & Hatzidimitriou's parameters for Sextans, σ_{obs} would need to be about 1 km s^{-1}. This is well outside the 99.9 per cent confidence level value of 4.4 km s^{-1} from the maximum likelihood calculation. Such a small dispersion would only just be detectable because the dispersion caused by our errors alone is 1.5 km s^{-1}.

The mass-to-light ratios obtained by Suntzeff et al. were $\rho_0/I_{0,V} = 30^{+20}_{-13}$ and $M_{\mathrm{tot}}/L_{\mathrm{tot},V} = 54^{+44}_{-24}$. Most of the discrepancy between these results and those derived from our data was due to the differences in the values of $S_{0,V}$ and M_V used. The values used by Suntzeff et al. were 1.3 $L_\odot$ pc^{-2} and -9.4, respectively.

4.2 Other possible explanations of the velocity dispersion

4.2.1 Anisotropy of the velocity dispersion

The models used to calculate the mass-to-light ratio assumed isotropy in the velocity dispersion. That is,

$$\sigma^2_{\mathrm{total}} = \sigma^2_{\mathrm{los}} + \sigma^2_\theta + \sigma^2_\phi = 3\sigma^2_{\mathrm{los}}, \qquad (10)$$

where σ_{los} is the line-of-sight velocity dispersion, and σ_θ and σ_ϕ are the dispersions that would be seen along the other two perpendicular directions. So, the maximum effect that anisotropy could have on the mass-to-light ratio is a factor of 3 (the calculated mass-to-light ratio would be a factor of 3 bigger than the true answer) if all the dispersion were actually along the line of sight. It is impossible to measure this in any way since we can only measure the velocity dispersion in one direction, so some effect from anisotropy is always a possibility. This alone, however, cannot account for the huge discrepancy between the mass-to-light ratio we have calculated for Sextans and the values measured for larger and smaller stellar systems.

4.2.2 Rotation of Sextans

Rotation of the dSph galaxy would artificially increase the observed velocity dispersion as long as that rotation was not in the plane of the sky. Therefore plots were made of the velocities of the stars with respect to their distances from the major and minor axes of Sextans as seen on the sky. The plot for the rotation around the minor axis of the galaxy is shown in Fig. 10. In the centre of a galaxy such as Sextans, the rotation curve would be expected to be linear and, as we have only observed close to the centre, linear least-squares fits were made to the rotation plots. There is little sign of any rotation in either case. For example, at 300 pc, the formally derived value of the rotation around the minor axis is 0.4 ± 0.5 km s^{-1} while that around the major axis is 1.1 ± 1.0 km s^{-1}. Other randomly chosen axes give similar answers, and, as can be seen by the plots, the linear fits are extremely poor. The detection of rotation is hampered by the fact that all our stars are close to the centre of the dSph galaxy, making fitting of a realistic rotation curve impossible. All we can say with any degree of certainty is that the rotation is too small for us to detect, the values quoted being upper limits.

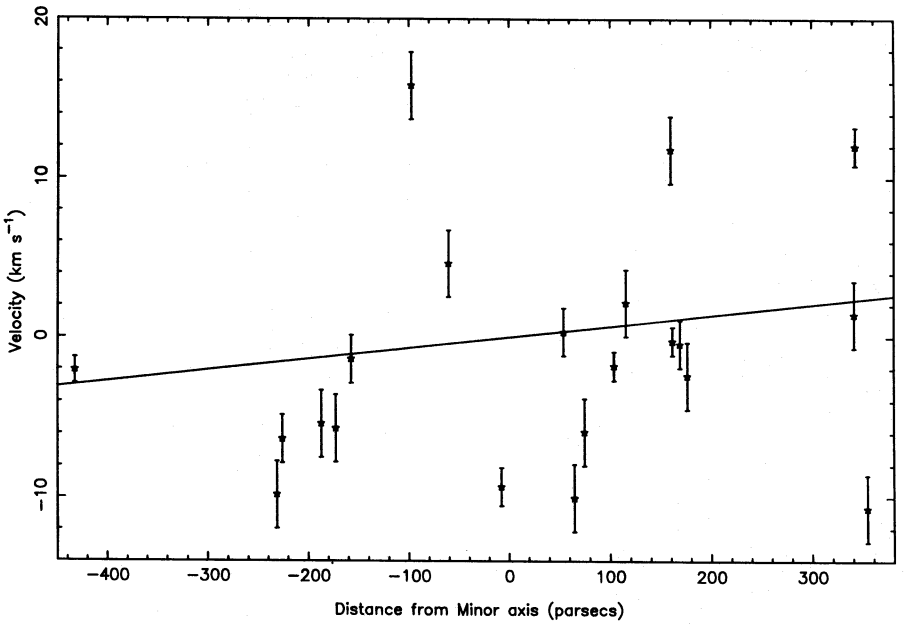

Figure 10. Test for rotation around the minor axis of the Sextans dSph galaxy! The best-fitting line to the data shows a rotation of 0.4 km s^{-1} at 300 pc. No significant rotation is detectable.

Table 7. Coordinates of non-members discovered.

Star	RA 1950	DEC 1950
21	10 11 19.8	-01 08 09
24	10 08 56.5	-01 18 27
25	10 11 33.6	-01 30 00
26	10 09 22.6	-01 44 27
27	10 08 54.3	-01 39 31
28	10 09 33.8	-01 04 00
29	10 09 31.5	-01 11 37
30	10 09 16.5	-01 16 18
31	10 11 37.4	-01 04 58
33	10 11 8.4	-01 02 17
34	10 09 21.4	-01 38 01
35	10 09 26.3	-01 36 45
36	10 09 16.1	-01 15 17
39	10 08 59.6	-01 35 23
40	10 10 15.4	-01 09 06
41	10 09 17.5	-01 28 49
42	10 11 53.1	-01 14 18
43	10 09 48.2	-01 44 34
44	10 10 32.0	-01 24 21
47	10 11 45.6	-01 25 05
50	10 09 26.2	-01 13 27

4.2.3 Binaries

The presence of binaries in the sample of observed stars would increase the observed velocity dispersion so that it is no longer a true indication of the mass of the galaxy. Suntzeff et al. (1993), made a calculation for Sextans using a Monte Carlo method. They obtained velocity dispersions of close to 6 km s^{-1} for a binary fraction of 0.25, assuming an intrinsic velocity dispersion of 2.1 km s^{-1}, which is equivalent to a mass-to-light ratio of 2.5 according to the parameters they used. Use of Irwin & Hatzidimitriou's parameters makes little difference to this result. The calculation involved several simplifications, such as uniformity of mass ratios and inclinations and a flat period distribution, which are not necessarily valid, but it is still useful for comparison with the observations.

We need to calculate how many binaries there were in our sample. The star that was not included in the calculations seemed to be obviously a binary but it is necessary to be more quantitative about this. There were five good-quality observations of star 6. Observation of such a large difference in velocity over the space of 1 yr is an 8σ result according to our calculated measuring errors and therefore extremely unlikely by chance. The other possible binary, star 8, was still included in the sample because of its fewer lower quality observations. The observation of the velocities of this star is a 3σ result, if we have estimated our velocity precision correctly. If there were two binary stars out of the nine for which there were multi-epoch observations, there should be a fraction of 0.22 in the sample used to calculate the velocity dispersion. If only one out of those nine was a binary then the binary fraction should be 0.11. This is not the true picture, however, because due to our observing criteria we could have detected only a fraction of the binary stars. To quantify the possible binary fraction accurately requires continued velocity monitoring. Such effects are important since a binary fraction of 0.1 still requires a true mass-to-light ratio of 50, whereas all extra mass above that found in globular clusters can be explained by a fraction of 0.25.

4.2.4 Tidal interaction with the Galaxy

The other alternative is that the dSph galaxies are being tidally disrupted by the Milky Way Galaxy, so that the

assumption of dynamical equilibrium underlying equations (8) and (9) is invalid. Structure observed in the Small Magellanic Cloud, along with the existence of the Magellanic Stream, is evidence that tidal disruption is currently occurring in this dwarf companion to the galaxy, and suggestive evidence for the existence of phase-space structure in the outer Galaxy, the reality of which would strongly support such a model of merging galaxies, continues to arise (Arnold & Gilmore 1992). Indeed, the concept of the small galaxies close to the Milky Way undergoing tidal disruption and merger is fundamental to standard CDM cosmologies.

A completely disrupted dSph galaxy of freely expanding stars would be expected to disperse in the time it takes for the galaxy to orbit our Galaxy once or twice, so that it would be unlikely for a significant proportion of the nine known dSphs to be undergoing tidal disruption now. It is possible, however, that the dSph could still be visible as a collection of stars having been tidally disrupted some time ago, as shown by Kuhn (1993) who has performed N-body calculations demonstrating that, in the case of strong velocity anisotropy, the time for an unbound dSph galaxy to disperse may be an order of magnitude larger than a free expansion argument would suggest. If Sextans is a tidally disrupted dSph galaxy, the velocity dispersion that we have been so diligently measuring would have nothing to do with the actual mass-to-light ratio of the galaxy.

The distance d_t that a dSph galaxy of mass M_{dSph} (in an orbit of ellipticity 0.5) would have to be from the Galaxy, mass M_G, for the force from the dSph on a star at distance r_t from the centre of the dSph to be balanced by the force from the Galaxy, assuming a Keplerian Galaxy potential, is

$$D_t = r_t \left(\frac{3.5 M_G}{M_{dSph}} \right)^{1/3} . \qquad (11)$$

This simple calculation can be made for all the dSph galaxies to gain an idea of whether they may be tidally disrupted. The distance D_t obtained is hereafter referred to as the tidal distance. Using Irwin & Hatzidimitriou's values for the luminosity and tidal radius for each dSph galaxy, a Galaxy mass of 10^{12} $M_\odot$ and a supposed mass-to-light ratio of 3, all the dSph galaxies except Fornax, LeoI and LeoII could be suffering tidal interaction with the Galaxy at the moment. Sextans particularly with such a large tidal radius is well inside its tidal distance of 370 kpc at 84 kpc, and Sculptor, Draco and Ursa Minor are more than 20 kpc inside their tidal distances.

Of course there is a question here over the form of the Galactic potential. Tidal distances are slightly smaller for alternative galaxy potentials like isothermal spheres, and it is probably true that fewer of the dSphs are tidally disrupted than is suggested by this calculation. Nevertheless, it is interesting to note that, with the exception of Sculptor, the dSph galaxies that are most likely to be undergoing disruption according to the preceding calculation are those with the largest apparent mass-to-light ratios.

If the same calculation is performed taking the masses derived from the velocity dispersion as M_{dSph}, Sextans and Sculptor are inside their tidal distances with all the other galaxies being 30 kpc or more outside theirs.

Therefore there are two consistent alternatives: the dSph galaxies may have small mass-to-light ratios, implying that several of the galaxies are being tidally disrupted or that some other effect such as binaries is causing the velocity dispersion not to be a true representation of the masses; alternatively, the galaxies have higher masses and fewer of them are being tidally disrupted.

4.2.5 *Dark matter*

Should all the ideas in the previous sections fail to account for the high mass-to-light ratios, the alternative is that dSph

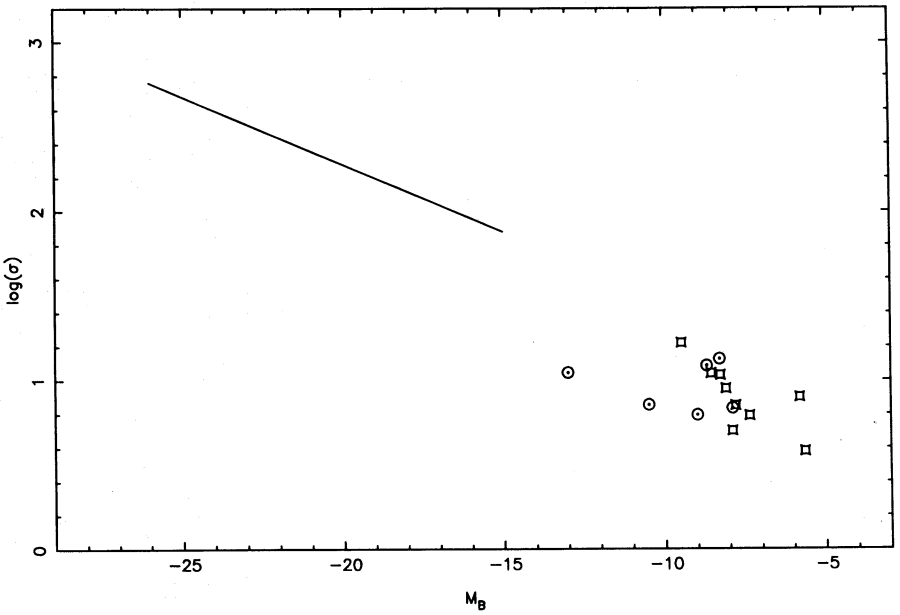

Figure 11. Velocity dispersion versus absolute magnitude for various systems. The straight line shows the relation for elliptical galaxies and the bulges of spirals. The squares are globular clusters, and the circles are Galactic dSph galaxies. The velocity dispersions of dSph galaxies are similar to those of globular clusters; it is the low central surface brightness of the dSphs which results in their comparatively high mass-to-light ratios (see Figs 12 and 13).

galaxies contain large amounts of dark matter, the brighter galaxies containing a lower proportion than the fainter ones. Each dSph galaxy would require a core dark matter density of about 0.07 $M_\odot$ pc^{-3}. The implication is a range in total luminosity of a factor of 80 compared with a range in total mass of a factor of 7. In some senses it is simpler to view the problem as one of missing brightness rather than missing mass. This is demonstrated by Figs 11, 12 and 13 which show the absolute magnitudes of globular clusters and dwarf spheroidals compared with their central velocity dispersions, central surface brightnesses and mass-to-light ratios, respectively. In each figure, the line marked is the empirical relation for these values for elliptical galaxies and the bulges of spiral galaxies. Also marked on the central surface brightness versus magnitude diagram are some dwarf elliptical galaxies from the Virgo and Fornax clusters. The velocity dispersions of globular clusters and dSph galaxies are similar; the discrepancy comes with the low central surface brightnesses of the dSph galaxies and their correspondingly high mass-to-light ratios. So, the velocity dispersions of the dSph galaxies follow a similar trend to other objects in the Universe; it is the low surface brightness that is anomalous, resulting in equally anomalous mass-to-light ratios.

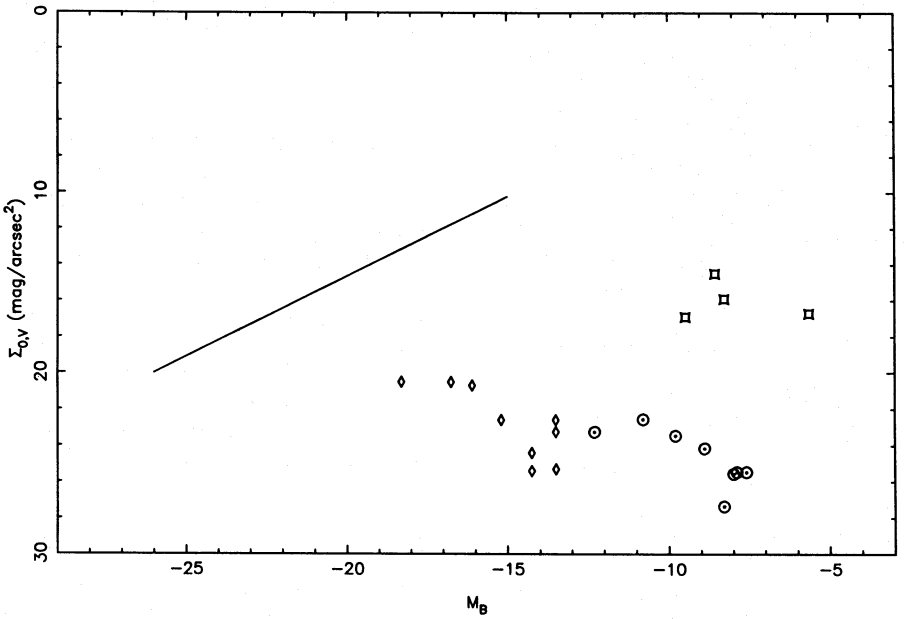

Figure 12. Central surface brightness versus absolute magnitude for various systems; see caption to Fig. 11 for details. Diamond symbols represent dwarf elliptical galaxies from the Virgo and Fornax clusters.

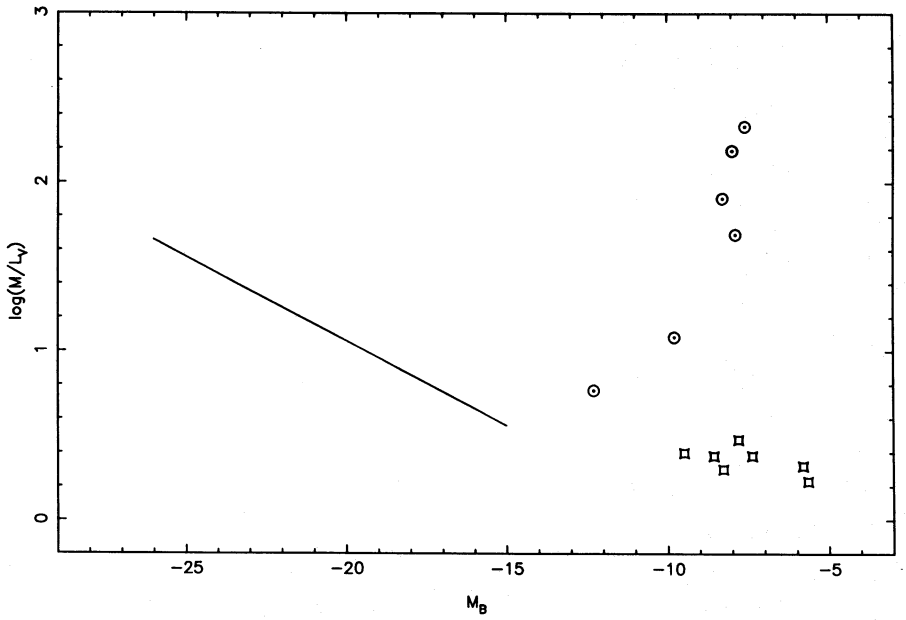

Figure 13. Mass-to-light ratio versus absolute magnitude for various systems; see caption to Fig. 11 for details.

5 CONCLUSION

The internal central velocity dispersion of the Sextans dSph galaxy is $7.0^{+1.3}_{-1.0}$ km s^{-1} measured from 21 giant stars. This leads to mass-to-light ratios of 124^{+85}_{-60} and 121^{+84}_{-58} in solar units, using core fitting and Illingworth's method, respectively. By comparison, apparently purely stellar systems, such as globular clusters and the stellar Galactic disc, have mass-to-light ratios of about 3. Thus the observed internal velocity dispersion of the Sextans dSph galaxy is several times larger than the value of about 1 km s^{-1} which is expected if the galaxy is a self-gravitating stable system whose gravitational potential is dominated by the mass in visible stars.

Of the possible explanations of this discrepancy, there are two that we think are more likely than the others. The first is as yet undetected binaries in the observed stellar tracers. We have found two possible binary stars in the sample of nine stars for which we have good multi-epoch observations, one of these being a more positive identification than the other. Since a binary fraction of 0.25 is sufficient to explain the excess mass-to-light ratio over about 3, further observations are required to rule out this possibility. The second possibility is failure of the dynamical assumptions due to current tidal disruption. There is quite a high likelihood that Sextans is being tidally disrupted, since simple calculation shows it to be the most likely of the dSph galaxies to be in this state. It may be relevant that the three dSph galaxies that are a priori the most likely to be undergoing tidal disruption are also the three with the highest apparent mass-to-light ratios. Other possible explanations include a substantial dark matter density in this galaxy or a serious underestimate of the measuring errors. We think the latter possibility is unlikely. Other effects which may contribute include velocity anisotropy and rotation in the dSph, although they are unlikely to provide a major part of the answer.

ACKNOWLEDGMENTS

The spectroscopy reported here was obtained with the William Herschel Telescope operated on the island of La Palma by the Royal Greenwich Observatory in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias.

REFERENCES

Armandroff T. E., Da Costa G. S., 1986, AJ, 92, 777
Arnold R., Gilmore G., 1992, MNRAS, 257, 225
Binney J. J., Tremaine S., 1987, Galactic Dynamics. Princeton Univ. Press, Princeton, p. 236
Caldwell N., Armandroff T. E., Seitzer P., Da Costa G. S., 1992, AJ, 103, 840
Da Costa G. S., Hatzidimitriou D., Irwin M. J., McMahon R. G., 1991, MNRAS, 249, 473
Eadie W. T., Drijard D., Jame F. E., Roos M., Sadoulet B., 1982, Statistical Methods in Experimental Physics. North-Holland Publishing Company, Amsterdam
Godwin P. J., Lynden-Bell D., 1987, MNRAS, 229, 7P
Illingworth G. D., 1976, ApJ, 204, 90
Irwin M. J., Bunclark P. S., Bridgeland M. T., McMahon R. G., 1990, MNRAS, 244, 16P
King I., 1962, AJ, 67, 471
King I., 1966, AJ, 71, 64
Kuhn J. R., 1993, ApJ, 409, L13
Richstone D. O., Tremaine S., 1986, AJ, 92, 72
Suntzeff N. B., Mateo M., Terndrup D. M., Olszewski E. W., Geisler D., Weller W., 1993, ApJ, 418, 208
Tonry J., Davis M., 1979, AJ, 84, 1511

APPENDIX A: THE ARMANDROFF AND DA COSTA METHOD

The mean velocities, equivalent to the values of $\overline{V}_{7.5}$ in Table 2, are written as v_i in the following equations. The error on a single observation is $\sigma_{\rm err}$, so the variance of the average squared velocity of a star is $\sigma^2_{\rm err}/N_i$, where N_i is the number of observations of star i used to calculate the velocity. The weight on star i is w_i, and is the inverse of this variance. The average velocity of the galaxy is then defined as

$$\overline{v} = \frac{\sum w_i v_i}{\sum w_i}, \tag{A1}$$

and the velocity dispersion is defined by

$$\sigma^2_{\rm obs} = \left[\frac{\sum w_i (v_i - \overline{v})^2}{\sum w_i} - \frac{N}{\sum w_i}\right] \frac{N}{N-1}. \tag{A2}$$

Here the first term in the brackets is simply the square of the observed velocity dispersion while the second term is the contribution to this dispersion from the measuring errors, which must be subtracted to produce $\sigma_{\rm obs}$, the intrinsic velocity dispersion.

The error on the square of the velocity dispersion is quoted as $\epsilon = \sqrt{\epsilon_1^2 + \epsilon_2^2}$ where $\epsilon_1 = \sigma^2\sqrt{2/N}$ and $\epsilon_2 = (2/N)\sqrt{\sum_{i=1}^{N} \varepsilon_i^2 \epsilon^2(\varepsilon_i)}$. Here ε_i is the error in velocity i and $\epsilon(\varepsilon_i)$ is the uncertainty in this error. The values of ϵ_2 were all very small, $\epsilon(\varepsilon_i)$ being of the order of 0.5 km s^{-1}.

To illustrate the overweighting this method gives, take the following extreme example. Suppose the true velocity dispersion in a system is 10 km s^{-1} and the measuring errors 1 km s^{-1}. Further, suppose we have only observed two stars, the first 10 times, the second once. If these two stars are a fair sample then their unweighted dispersion about the mean will be 10.0 km s^{-1}. To one decimal place this will also be the maximum likelihood estimate of the dispersion. The Armandroff & Da Costa method, however, would estimate a dispersion of 6.0 km s^{-1} mainly due to the mean velocity estimate being severely biased toward the star with most measurements. The simple weighting of the velocity dispersion summation then compounds the problem.

APPENDIX B: THE MAXIMUM LIKELIHOOD METHOD

Consider a series of velocity measurements v_i, $i = 1, 2, \ldots, N$, with associated errors σ_i, $i = 1, 2, \ldots, N$. We wish to estimate the systemic velocity, $\overline{v}$, and the velocity dispersion, $\sigma_{\rm obs}$, of the system. Assume that both the measuring errors, σ_i, and the intrinsic system velocities are Gaussian distributed (it is straightforward to generalize to other intrinsic forms). Then the probability of observing velocity v_i is given by

$$P(v_i)=\frac{1}{\sqrt{2\pi(\sigma_i^2+\sigma_{\rm obs}^2)}}\exp\left[\frac{-(v_i-\overline{v})^2}{2(\sigma_i^2+\sigma_{\rm obs}^2)}\right], \tag{B1}$$

and the likelihood of observing the series of observations $\{v_i, i=1, N\}$ is given by

$$L(v_1, v_2, \ldots, v_N)=\prod_i P(v_i). \tag{B2}$$

Using the log likelihood function for convenience we see that

$$\ln[L(\overline{v}, \sigma_{\rm obs}^2)]=\sum_i -\frac{(v_i-\overline{v})^2}{2(\sigma_i^2+\sigma_{\rm obs}^2)} -\frac{1}{2}\sum_i \ln(\sigma_i^2+\sigma_{\rm obs}^2)+\text{constant}. \tag{B3}$$

$\ln(L)$ is a function of $\overline{v}$ and $\sigma_{\rm obs}^2$ only, and is to be maximized with respect to these variables. Numerically, this is easily accomplished by constructing a two-dimensional grid of suitably fine mesh and simply 'contouring' the resulting array. By making use of the Central Limit Theorem we can argue that the error ellipses for $\overline{v}$ and $\sigma_{\rm obs}^2$ are then simply given by steps down in $\ln(L)$ from the peak where (peak $-1/2$) corresponds to the 1σ error or 68 per cent confidence limit, (peak $-4\times1/2$) to the 2σ error or 95 per cent confidence limit, and so on.

To proceed analytically we differentiate equation (B3) with respect to $\overline{v}$ and $\sigma_{\rm obs}^2$ to yield

$$\frac{\partial \ln(L)}{\partial \overline{v}}=\sum_i\left(\frac{v_i-\overline{v}}{\sigma_i^2+\sigma_{\rm obs}^2}\right)=0, \tag{B4}$$

and

$$\frac{\partial \ln(L)}{\partial \sigma_{\rm obs}^2}=\sum_i\left[\frac{(v_i-\overline{v})^2}{2(\sigma_i^2+\sigma_{\rm obs}^2)^2}-\frac{1}{2(\sigma_i^2+\sigma_{\rm obs}^2)}\right]=0. \tag{B5}$$

Re-arrangement of equations (B4) and (B5) yields an iterative scheme (cf. Gauss–Seidel) for finding the unknowns. This gives

$$\hat{\overline{v}}=\frac{\sum_i w_i v_i}{\sum_i w_i}, \tag{B6}$$

and

$$\hat{\sigma}_{\rm obs}^2=\frac{\sum[(v_i-\overline{v})^2-\sigma_i^2]w_i^2}{\sum_i w_i^2}, \tag{B7}$$

where $w_i=1/(\sigma_i^2+\sigma_{\rm obs}^2)$ is updated each cycle. If $\sigma_{\rm obs}^2 \gg \sigma_i^2$ then

$$\hat{\overline{v}}=\frac{1}{N}\sum_i v_i, \tag{B8}$$

and

$$\hat{\sigma}_{\rm obs}^2=\frac{1}{N}\sum_i[(v_i-\overline{v})^2-\sigma_i^2]. \tag{B9}$$

Since $\overline{v}$ is also estimated simultaneously, a factor $N/(N-1)$ is required for equation (B7) (and equation B9) to make $\sigma_{\rm obs}^2$ an unbiased estimator. Note that both the functionality of the 'weighting' factor and its form differ from that used by Armandroff & Da Costa (1986).

To estimate the analytic errors for $\overline{v}$ and $\sigma_{\rm obs}^2$ we first construct the information matrix (e.g. Eadie et al. 1982),

$$\mathbf{I}=\begin{bmatrix}\dfrac{\partial^2 \ln(L)}{\partial \overline{v}^2} & \dfrac{\partial^2 \ln(L)}{\partial(\sigma_v^2)\overline{v}}\\[2ex] \dfrac{\partial^2 \ln(L)}{\partial(\sigma_v^2)\overline{v}} & \dfrac{\partial^2 \ln(L)}{\partial(\sigma_v^2)^2}\end{bmatrix}. \tag{B10}$$

The parameter covariance matrix is then given by $\mathbf{I}^{-1}$. The off-diagonal term is given by

$$\frac{\partial^2 \ln(L)}{\partial(\sigma_v^2)\overline{v}}=\sum_i -\frac{(v_i-\overline{v})}{(\sigma_i^2+\sigma_{\rm obs}^2)^2}\approx 0, \tag{B11}$$

which is small relative to the diagonal terms (cf. equation B4) since

$$\frac{\partial^2 \ln(L)}{\partial \overline{v}^2}=\sum_i -\frac{1}{\sigma_i^2+\sigma_{\rm obs}^2}, \tag{B12}$$

and

$$\frac{\partial^2 \ln(L)}{\partial(\sigma_v^2)^2}=\sum_i\left[-\frac{(v_i-\overline{v})^2}{(\sigma_i^2+\sigma_{\rm obs}^2)^3}+\frac{1}{2(\sigma_i^2+\sigma_{\rm obs}^2)^2}\right]. \tag{B13}$$

Therefore, to a good approximation,

$$\text{var}\{\overline{v}\}=\left(\sum_i \frac{1}{\sigma_i^2+\sigma_{\rm obs}^2}\right)^{-1}\approx\frac{\sigma^2}{N}, \tag{B14}$$

and

$$\text{var}\{\sigma_{\rm obs}^2\}=\left\{\sum_i\left[\frac{(v_i-\overline{v})^2}{(\sigma_i^2+\sigma_{\rm obs}^2)^3}-\frac{1}{2(\sigma_i^2+\sigma_{\rm obs}^2)^2}\right]\right\}^{-1}\approx\frac{2\sigma^4}{N}, \tag{B15}$$

where $\sigma^2=\langle\sigma_i^2\rangle+\sigma_{\rm obs}^2$.