

DWARF GALAXIES IN THE VIRGO-ACS SURVEY

Klein Onderzoek report

PATRICK (E.G.P.) BOS

Supervisors: Reynier F. Peletier and Edwin A. Valentijn

Kapteyn Astronomical Institute, University of Groningen

June 12, 2008



Heigh Ho...

Image: <http://kevin.p.monaghan.googlepages.com/SevenDwarfs.jpg/SevenDwarfs-full.jpg>

Contents

1	Introduction	3
1.1	Dataset	4
1.2	Expectations from theory	4
2	Data reduction	7
2.1	Dataset retrieval	7
2.2	Astro-WISE Ingestion	8
2.3	GalPhot model subtraction	9
2.3.1	GalPhot residual ingestion	10
2.4	SExtractor	12
3	Dataset clean-up and calibration	15
3.1	First bulk: false detections and stars	15
3.2	Association	16
3.3	Comparison to SDSS	17
4	Analysis	22
4.1	Background separation	22
4.2	Close-up inspection	25
4.3	Parameters	27
4.3.1	Distribution	30
4.3.2	Parameter-parameter relations	32
4.3.3	Luminosity function	32
5	Discussion	34
6	Conclusions	38
7	Acknowledgements	39
A	Research description (Dutch)	45
B	Derivations	46
B.1	Average surface magnitude within the effective radius: $\langle\mu\rangle_e$	46
B.2	From μ_0 and Sérsic n to $\langle\mu\rangle_e$	47
B.3	Dependence of $\langle\mu\rangle_e - \mu_0$ to Sérsic n	48
C	Software setup	49
C.1	Astro-WISE	49
C.1.1	Kapteyn version	49
C.1.2	Custom checkout	49
C.2	GalPhot in IRAF	50
D	Astro-WISE	51
D.1	Philosophy and personal contributions	51
D.2	Python scripts	51
E	Source images	52

1 Introduction

As a lot of authors would agree (Sabatini et al., 2003b; Brosch, 2002; Sabatini et al., 2003a), the Virgo Cluster offers a great opportunity for studying several kinds of large populations of galaxies. Not only is it the nearest cluster in our environment, it is also the dominant structure in the Local Supercluster containing a large concentration of galaxies. It is located between 15 and 20 Mpc away from us and spans about 140 square degrees on the northern sky. Its crossing time of about $0.1H_0^{-1} = 0.98h^{-1}\text{Gyr}$ and mass of $1.2 \times 10^{15}M_\odot$ make it a place in which high accelerations will already have caused a lot of interaction, possibly influencing the evolution of the cluster members. It is in short a very dynamic and certainly as of yet unrelaxed cluster in which a lot of interesting astrophysics might be observed.

The first massive survey of the cluster (which was part of the Las Campanas survey of galaxies) contains 2096 catalogued galaxies (Binggeli et al., 1985), culminating in the Virgo Cluster Catalog (VCC), over a thousand of which were identified as cluster members. This survey was incomplete though for magnitudes $M_B \gtrsim -14$, because of the use of photographic plates. Over the years the objects in this survey have been studied extensively. Data from several other surveys were used to extend our knowledge of the Virgo Cluster and its claimed members (some of which were excluded as members over time).

One of the things that makes the Virgo cluster interesting is that it offers us the possibility of studying large dwarf galaxy populations. Among the VCC galaxies about a thousand were classified as dwarfs, but a lot more were to be expected when you take the incompleteness into account; if the Schechter luminosity function (Schechter, 1976) is correct for lower luminosities than those of the identified galaxies then thousands more in the low luminosity range should be present, but beyond our detection range. Later studies indeed showed that clusters like Virgo, Coma and Fornax (all relatively close clusters) seemed to contain quite a lot of dwarf galaxies, with luminosity function (LF) faint-end slope values of around $\alpha \approx -1.8$ whereas on the basis of VCC alone values around -1.3 were predicted. Outside these clusters though (i.e. in so-called *field* galaxy populations) the relative dwarf count seems to be much lower, with faint-end slopes of -1.2 derived from the Sloan Digital Sky Survey (Blanton et al., 2001) and Two degree Field survey (2dF) (Cross et al., 2001).

These values indicate a large environmental dependence on the evolution of dwarf galaxy populations. In fact this is one of the illustrations of the dire need for research into these objects. The Cold Dark Matter (CDM) paradigm of hierarchical structure formation predicts steep faint-end slopes in all environments, which is inconsistent with these observations. This is where theories on environmentally induced formation processes become inevitable. Several of these theories have indeed been proposed in recent years; notably the squelching model (Tully et al., 2002) (representing a 'nature' approach to formation), in which dwarf galaxy counts are dependent on whether their environmental structure is formed before or after reionization, and the harassment model (Moore et al., 1999) (a 'nurture' approach), where dwarf galaxy formation is caused by tidal interactions between infalling galaxies. Both these theories seem to be able to explain at least the Virgo Cluster's abundance (or should we say the field's scarcity?) of dwarf galaxies, but certainly more data is needed on the presence of dwarf galaxies in different environments to be able to fully reconstruct the history of our universe.

For the latter effort we have tried to contribute our part during this research. Our interest was in finding low luminosity Virgo dwarfs, so as to partially complete the incompleteness inherent in the VCC. For these dwarfs we would then investigate their structural properties

as well as their distribution, both as function of environment. This could then perhaps lead us to conclude more about some of the herefore mentioned models.

1.1 Dataset

For this project we have used two hundred Hubble Space Telescope (HST) images. The images were obtained using the Wide Field Channel (WFC) on the Advanced Camera for Surveys (ACS), an instrument that is in use since March 7, 2002. They were made under HST proposal number 9401 for the ACS Virgo Cluster Survey (ACSVCS), a project run by Patrick Côté of Dominion Astrophysical Observatory (Côté et al., 2004). They are images of 100 early-type, low-redshift Virgo cluster galaxies taken in two bandpasses: SDSS g and z (see table 1), which are green and near-infrared filters. The g-band data have a total exposure time of 750 seconds (combined from two 375 second exposures for dithering) and a zeropoint of 26.068 magnitudes (Jordán et al., 2004) whereas the z-band data have a combined exposure time of 1120 seconds and have a zeropoint of 24.862 magnitudes (Jordán et al., 2004). The WFC has a 202 by 202 arcsecond field of view and a plate scale of about 0.05 arcseconds/pixel (Jordán et al., 2004).

Description	ACS WFC filter name	Central wavelength (Å)	Width (Å)
SDSS g	F475W	4760	1458
SDSS z	F850LP	9445	1229

Table 1: ACS WFC filters used for the project data (Pavlovsky, 2006, table 5.1)

In the end what we have is 100 g-band and 100 z-band images covering a total area of about 0.315 degrees squared. Compared to the total area of the Virgo Cluster which is about 140 degrees squared according to the VCC (Binggeli et al., 1985) this is not a lot. As we will see later it was not easy to find dwarf galaxies in this dataset, mainly due to an enormous amount of background sources. In any case, new dwarf populations have not been searched for by the ACSVCS project yet, because their main foci were on 1) globular cluster populations, 2) determining accurate distances through surface brightness fluctuations and 3) isophotal analysis of the targetted galaxies.

1.2 Expectations from theory

Before we begin our investigations we need to know what to expect from these images. Dwarf galaxies are defined as galaxies with a magnitude of $M_B > -18$ and have typical total sizes of kiloparsecs, while their half-light radii (the radius within which half of the object's total light is contained) are on the order of 20 - 200 pc¹. What does this imply for their observable features in the Virgo Cluster and in particular in our images of the Virgo Cluster?

Virgo is on average located at a distance of $d \sim 17$ Mpc (Tonry et al., 2001), corresponding to a distance modulus of

$$m - M = 5 \log_{10}(d/10\text{pc}) \quad (1)$$

$$m - M = 5 \log_{10}(17 \times 10^6/10) \quad (2)$$

¹General range determined by looking at some values for known Local Group dwarfs from the NASA/IPAC Extragalactic Database (NED). URL: <http://nedwww.ipac.caltech.edu/>

$$m - M \approx 31.2 \quad (3)$$

This means we will be looking for objects with an apparent magnitude of $m_B \gtrsim 13.2$ (of course we will need to convert this B band magnitude to the g and z bands used in our research, which we will do later on). We will be measuring half-light radii in arcseconds. We calculate the expected range using:

$$\tan(\alpha) = l_{\text{physical}}/d \Rightarrow \alpha = \arctan(l_{\text{physical}}/d) \quad (4)$$

This gives us a range of 0.24 - 2.4 arcseconds in size, corresponding to 4.7 - 47 pixels in the ACS images. The FWHM of the ACS is ~ 2.1 pixels, so especially the smaller objects might be hard to properly detect and model, but it should be possible as they are resolved. We can expect significant errors at the smaller end of the distribution.

With the basic characteristics determined we knew what to look for. Before we started though it was essential that we determined the number of them we expected to find, lest we started on a wild goose chase. To determine this we used the Schechter luminosity function (Schechter, 1976), which has become a standard theoretical tool in this field. Luminosity functions tell you how many galaxies you might find in a certain luminosity range per area on the sky. It is of course a statistical relation and will therefore by no means exactly determine the amount of sources we will find, but it is the best thing we have got, so we will go with it. It is given by equations 5, 6 and 7 below, where δN is the number of galaxies in a range of either luminosities δL or magnitudes δM , based on the distribution functions $\phi(L)$ or $\phi(M)$ (Sandage et al., 1985; Schechter, 1976).

$$\delta N = \phi(L)\delta L = \phi(M)\delta M \quad (5)$$

$$\phi(L) = \frac{N_0}{L_*} \left(\frac{L}{L_*} \right)^\alpha e^{-L/L_*} \quad (6)$$

$$\phi(M) \propto \left(10^{0.4(M_*-M)} \right)^{\alpha+1} e^{-10^{0.4(M_*-M)}} \quad (7)$$

M_* and L_* are the distribution's "typical" magnitude and luminosity and N_0 is a normalizing number of galaxies. We shall now only further consider equation 7; in equations 8, 9 and 10 we show that for faint objects the function is significantly simplified.

$$\lim_{M \rightarrow \infty} \phi(M) \propto \lim_{M \rightarrow \infty} \left(10^{0.4(M_*-M)} \right)^{\alpha+1} \lim_{M \rightarrow \infty} \left(e^{-10^{0.4(M_*-M)}} \right) \quad (8)$$

$$\lim_{M \rightarrow \infty} \left(-10^{0.4(M_*-M)} \right) = 0 \Rightarrow \lim_{M \rightarrow \infty} \left(e^{-10^{0.4(M_*-M)}} \right) = 1 \quad (9)$$

$$\Rightarrow \phi(M) \propto \left(10^{0.4(M_*-M)} \right)^{\alpha+1} \quad (\text{approximately}) \quad (10)$$

Now it is easily seen why α is called the faint-end slope value. We used an approximate value of $\alpha = -1.5$ to estimate the number of dwarfs we might find in the data. This value underestimates the number of galaxies if we are to take recent findings of -1.8 (Sabatini et al., 2003b, where they included VCC data) or even lower (Phillipps et al., 1998) for the Virgo Cluster seriously, but lacking any definite theoretically backed numbers, we decided it best not to become overconfident. The luminosity function then becomes:

$$\phi(M) \propto 10^{-0.2(M_*-M)} = 10^{-0.2M_*} 10^{0.2M} \propto 10^{0.2M} \quad (11)$$

Now to find the number of dwarfs in the range we are searching in we need one more number: a somewhat definite number of galaxies in the Virgo Cluster within a certain range. We need this number if we are to quantitatively use the Schechter function, because up to now we still only have a proportionality (in effect it will determine both N_0 and M_* , though we will not explicitly calculate them). Because the VCC becomes incomplete for sources fainter than -13 magnitudes (according to Sandage et al. (1985, figure 9), taking into account the difference in distance moduli between their paper and this report), we will take two bins: one from -14 to -15 and another from -13 to -14. The resulting normalization factors from the two bins will in the end be averaged. In the VCC catalog from Binggeli et al. (1985)² we found the number of galaxies from -14 to -15 to be 149 and from -13 to -14 to be 200 (only including what they thought to be certain cluster members). The area of the VCC survey was ~ 140 square degrees (Binggeli et al., 1985), so we have $N_{-15 \rightarrow -14} \sim 149/140 \sim 1.1 \text{ deg}^{-2}$ and $N_{-14 \rightarrow -13} \sim 200/140 \sim 1.4 \text{ deg}^{-2}$.

Calculating the faint end distribution from this can be done by first integrating equation 11 over the wanted range of M (here we only calculate explicitly for the -15 to -14 range):

$$N_{M_1 \rightarrow M_2} = \int_{M_1}^{M_2} \phi(M) dM \propto \int_{M_1}^{M_2} 10^{0.2M} dM = \frac{10^{0.2M_2} - 10^{0.2M_1}}{\ln 10^{0.2}} \propto 10^{0.2M_2} - 10^{0.2M_1} \quad (12)$$

Finally, to calculate the actual expected number of galaxies per square degree in a magnitude bin we divide the proportionality out by using the derived value of $N_{-15 \rightarrow -14}$.

$$\frac{N_{M_1 \rightarrow M_2}}{N_{-15 \rightarrow -14}} = \frac{N_{M_1 \rightarrow M_2}}{1.1 \text{ deg}^{-2}} = \frac{10^{0.2M_2} - 10^{0.2M_1}}{10^{0.2*(-14)} - 10^{0.2*(-15)}} = \frac{10^{0.2M_2} - 10^{0.2M_1}}{5.85 \times 10^{-4}} \quad (13)$$

$$N_{M_1 \rightarrow M_2} = \frac{1.1 \text{ deg}^{-2}}{5.85 \times 10^{-4}} (10^{0.2M_2} - 10^{0.2M_1}) = (10^{0.2M_2} - 10^{0.2M_1}) \times 1820 \text{ deg}^{-2} \quad (14)$$

Given the limited coverage of our dataset at 0.309 deg^2 , the total number for a magnitude bin in our survey will be:

$$N_{M_1 \rightarrow M_2, \text{total}} = (10^{0.2M_2} - 10^{0.2M_1}) \times 562 \quad (15)$$

We can do the same calculation for the -14 to -13 range which gives us a normalization factor of 476. Averaging this with 562 the final equation becomes:

$$N_{M_1 \rightarrow M_2, \text{total, final}} = (10^{0.2M_2} - 10^{0.2M_1}) \times 519 \quad (16)$$

In figure 1 the resulting expected distribution is shown. As a reminder: this is both an underestimation, an approximation for the faint-end and merely a statistical measure and therefore by no means conclusive. In practice though these numbers could turn out even lower, especially when you take into account that for the fainter magnitudes, all kinds of measurement and modelling trouble will come into play, decreasing completeness and accuracy. Moreover, the images were no random samples of the Virgo cluster but were specifically targetted at galaxies, thus further disturbing the predictive power of our Schechter function. If we were to accurately predict the number of galaxies we might better have used some two-point correlation function, giving us the probability of finding certain galaxies within the vicinity of the targetted galaxies. This seemed a bit overzealous though, so we did not. As will be shown in section 4, it has indeed proven quite difficult to find any dwarfs in our dataset.

²A digital version can be found at the Simbad Astronomical Database.
URL: <http://webviz.u-strasbg.fr/viz-bin/VizieR?-source=J/AJ/90/1681>.

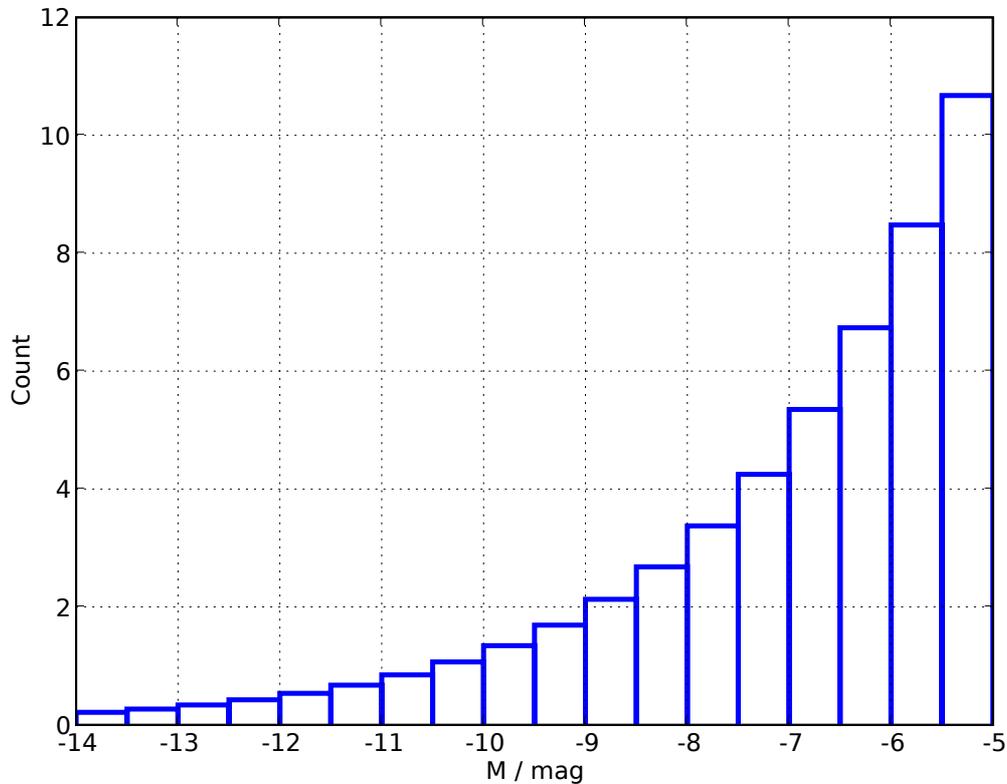


Figure 1: Faint-end distribution of galaxies we expect to find in the ACSVCS data.

2 Data reduction

2.1 Dataset retrieval

We downloaded the data from the HST archive³. The exact search parameters that will give you all the ACSVCS images we have used are:

```
Imagers: ACS
Exp Time: 1120,750
Proposal ID: 9401
Maximum records: 500
```

You then just **Mark all** and **Submit marked data for retrieval from STDADS**. You can then choose to let them upload your data via **sftp** to your server of choice, but for this you first need to register at their site⁴. You can also just use username *anonymous* and your email address for the password, because the ACSVCS data are public. When we tried getting all the data with **sftp** it all went a bit too slow so we registered and put our data on a *Stage disk* on their server so we could get the files ourselves using **ftp**.

³URL: <http://archive.stsci.edu/hst/search.php>

⁴URL: <http://archive.stsci.edu/registration/>

After some puzzling (there is enough data at the HST archive to drown in; thanks to John McFarland for helping us out here) we decided to use the *drizzled* images. These are fully calibrated and reduced (flat fields, bad pixels, cosmic rays, etc.), geometrically corrected and dither-combined images using the `drizzle` routine in IRAF. This routine thus basically creates ready to use reduced frames. One thing to be careful of is that while in the FITS headers of the data files the exposure time is still set to the original amount of seconds (750 and 1120 respectively), the drizzled images are in electrons (or ADU, since the gain was set to one) per second.

2.2 Astro-WISE Ingestion

Our next step introduces us to Astro-WISE (Astronomical Wide-field Imaging System for Europe). From a recent brochure:

The Astronomical Wide-field Imaging System for Europe, Astro-WISE, is an environment consisting of hardware and software which has been developed to be able to scientifically exploit the ever increasing avalanche of observational data produced by science experiments. Astro-WISE started out as a system geared towards astronomy, but is now also being used for projects outside astronomy. Astro-WISE is an all-in-one information system: it allows scientists to archive raw data, to process and calibrate data, perform post-calibration scientific analysis and archive all results in a distributed environment. The system architecture links together all these steps of the data flow. The complete linking of all input and output in the data flow, including software code used, for arbitrary data volumes has only been feasible thanks to a novel paradigm devised by the creators of Astro-WISE, named "target processing". The Astro-WISE information system started with one particular astronomical optical wide field imager, Omega-CAM. Subsequently, the system was expanded to host arbitrary optical wide field imagers, radio data and any other kind of digitized images.

For this Klein Onderzoek we have used Astro-WISE to store and keep track of our data and process it using astronomical tools that are incorporated into the system. The first step though is *ingesting* your data into the system, which in our case meant committing the drizzled ACSVCS images to the Astro-WISE database.

Astro-WISE is an object oriented system which uses several classes to distinguish data e.g. on the basis of either an everyday astronomical description or - in the case of end-products - steps in the reduction process. Data classes based on common astronomical terms are e.g. `BiasFrames` and `FlatFrames`, the purpose of which is immediately clear. For the science images there is a series of classes describing steps in the reduction process: from `RawScienceFrames` (straight from your telescope), through `ReducedScienceFrames` (on which bias and flat reduction have been applied) up to `RegriddedFrames` (astrometrically correct images) and `CoaddedRegriddedFrames` (frames from different chips put together into one final image).

All these classes can on the one hand make storing and finding data very clear when used consistently. For the instrument we have used though this is not the case. The general procedure in Astro-WISE is that you commit your raw data together with some reduction frames (bias frames, flat-fields). Astro-WISE will then reduce your data for you and within

a few processing cycles at the distributed processing unit (200 processing nodes at a Linux cluster in the *Rekencentrum* in Groningen that take the load off your own workstation) you have got your ScienceFrames. Our data was already reduced at STScI (HST headquarters) though, so the ingestion starts beyond this reduction point. We did not need any reduction frames, but could begin right away by ingesting the data as ReducedScienceFrames. This is in fact a bit confusing at first, because the drizzled ACS frames are actually coadded frames; they are composed of images from two chips.

Long story short: it is important to get acquainted with naming conventions in Astro-WISE for your instrument specifically. When you have got this sorted out actually ingesting the images is easy with a bit of help from the Astro-WISE team. The procedure, where we first ingest into our personal MyDB to ensure everything goes well and migrate to a public database afterwards, is as follows.

1. Set up your Astro-WISE environment (see appendix C).
2. Empty your MyDB (this was necessary, because at that time it was not supported to migrate only part of a database to another). In a terminal issue the following command:

```
env project=AWPBOS awe $AWEPIPE/Toolbox/dbdeleteall.py
```

Where AWPBOS is the name of the personal MyDB we have used.

3. Ingest the data into your MyDB, using the following command:

```
env project=AWPBOS awe $AWEPIPE/Toolbox/ingest/Ingest.py\  
-i *drz.fits -p redscience -commit
```

4. Check if the ingestion went as planned by using the online database viewer; image data as well as header (meta-) data at <http://dbview.astro-wise.org/>.
5. If all went well you can migrate the MyDB data to the ACS@HST database by issuing:

```
env project=AWPBOS awe $AWEPIPE/Toolbox/dbmigrateall.py ACS@HST
```

The data were from then on accessible from the ACS@HST database as ReducedScienceFrames.

2.3 GalPhot model subtraction

We are interested in as of yet uncharted dwarf galaxies in the survey. Because the target galaxies from the ACSVCS may hinder the photometry for these weak galaxies we decided to model these galaxies and subsequently subtract these models from the original images. We have used GalPhot for this purpose, a tool that fits ellipses with constant surface brightness to a preset number of isophotes in the original source. Because the ACSVCS target galaxies are early-type (and therefore elliptical) GalPhot is perfectly suited for this task of removing them from our images.

We wanted to use the residual images (original image minus GalPhot model) to find our sources, but since it was not yet supported to store complete residual images in Astro-WISE so as to use them for further analysis (i.e. inside Astro-WISE; this feature will soon be added

though, see appendix D.1) we had to run GalPhot using IRAF, i.e. outside of Astro-WISE (see appendix C). After this we could then reingest the residual images into Astro-WISE (see section 2.3.1).

In table 2 the reader will find the basic configuration parameters we have used in GalPhot. For most cases these parameters worked out fine, for others we have had to adjust them slightly. In general the `errshap` and `errcen` parameters could be set a bit higher to allow GalPhot a slightly larger error in fitting and the `iter2`, the number of harmonic fit iteration steps, which is not necessary for a good fit, could be cranked down a notch as well in case of need. In more specific cases, e.g. when there were sources that needed to be masked out to ensure a proper fit, we could indeed mask the source in IRAF with the `mark` command, but on several occasions we also just set a limit to the `rmax` parameter, to stop the fit from reaching the other source's influence. Each image also had a band in the middle which corresponds to the space between the two WFC chips; this could be masked out for all images and so we put this area in a default mask file for all the images. The `x` and `y` parameters (the galaxy center coordinates) were determined manually for each image.

Parameter	Value	Parameter	Value	Parameter	Value
<code>image</code>	<code>bla</code>	<code>rshap</code>	0.	<code>r2</code>	0.
<code>output</code>	<code>.res</code>	<code>rcen</code>	0.	<code>nsammax</code>	66
<code>galaxy</code>	<code>.pos.gal</code>	<code>errshap</code>	0.02	<code>fracmin</code>	0.6
<code>ngal</code>	1.	<code>errcen</code>	0.02	<code>cliplow</code>	0.
<code>deletio</code>	<code>.pos.del</code>	<code>iter1</code>	30	<code>cliphig</code>	0.
<code>intable</code>		<code>iter2</code>	5	<code>debug</code>	1
<code>outtable</code>	<code>.tab</code>	<code>hmax</code>	6	<code>extend</code>	yes
<code>rmin</code>	1.	<code>dposmax</code>	0.1	<code>outfill</code>	yes
<code>rmax</code>	0.	<code>dellmax</code>	0.1	<code>npolres</code>	4
<code>radfac</code>	1.1	<code>dangmax</code>	0.1	<code>mode</code>	ql
<code>linear</code>	no	<code>r1</code>	20		

Table 2: GalPhot basic configuration parameters for ACSVCS target galaxy subtraction

We produced residual images for all but four of the images. The galaxies we were not able to fit were VCC2095, VCC21, VCC778 and VCC798 (where VCC stands for the Virgo Cluster Catalog from Binggeli et al. (1985)). These respectively suffered from being very elongated, apparently containing two cores, having three relatively bright galaxies quite nearby and simply not fitting properly. This omission of 4% of the data will thus lead to a decrease in the expected number of galaxies by 4%. Some of the good residual images still needed a slight correction; GalPhot added some pixels valued -33000 to the images, which we removed by hand.

To give an illustration of the GalPhot residual image quality we included 'before' and 'after' shots of a few images in figure 2. A left-over, caused by the fact that no galaxy is perfectly elliptical, can be seen in all the images. We will deal with this in section 3.

2.3.1 GalPhot residual ingestion

We shortly comment on the way we ingested GalPhot residuals into Astro-WISE. Because Astro-WISE has specific routines (FITS file header translators to be exact, which translate

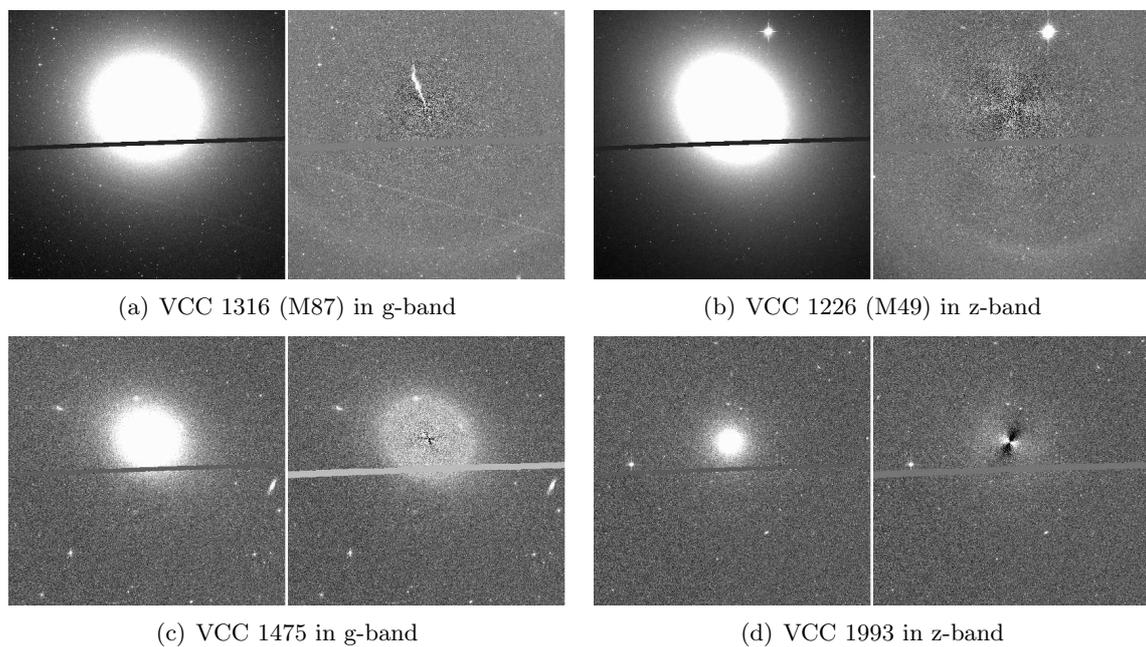


Figure 2: A comparison of original to GalPhot residual images. Left is before GalPhot model subtraction, right is after. SExtractor can partly deal with problems like too much background in figure 2(c) by locally determining the background value, which we do in this project (see section 2.4).

instrument specific header data to general Astro-WISE format data) for each instrument you cannot simply ingest your own custom images. In principle you could write your own GalPhot residual-image header translator, or better yet, enable Astro-WISE to automatically store GalPhot residual images in the database, which is something that is actually being developed now. As this feature was not available to us yet we used a trick by Gijs Verdoes Kleijn.

The trick came down to making Astro-WISE believe it was ingesting original ACS data. We did this by taking the drizzled files ingested earlier (see section 2.2) and replacing the original image data in the FITS files by the corresponding image data in the residual frames from GalPhot. We then simply changed the OBJECT header keys to read `VCCxxxx.residual` instead of the original `VCCxxxx`, to ensure others would be able to recognize them as not being original images. In the end we ingested these files into Astro-WISE.

2.4 SExtractor

One of the most difficult and thus time consuming parts of data reduction (in the wider sense of the term, i.e. as in reducing your image data to numbers) for a first time user is running Source Extractor (SExtractor or SE from here on) on your data. We used this program by Emmanuel Bertin to extract object data from our images. SE basically identifies clumps of pixels with values above a certain threshold as objects. It then derives photometry for these objects using various models, ranging from simply measuring the identified pixel values corresponding to the object or applying aperture photometry, up to fitting Kron and Petrosian models. All in all, it is able to deliver quite a lot.

This enormous amount of data does not come cheap though. SE hosts a wealth of configuration options⁵, the results of which you need to understand to be able to make sense of its output or even produce sensible output in the first place. Especially when dealing with faint extended galaxies like in our research it proved quite a task to optimize between keeping faint junk out and getting faint galaxies in.

The crux of the matter is deblending. The problem is that faint galaxies usually have patches of intensity below the SE detection threshold (which the user needs to set himself). To SExtractor this poses a problem, because it must determine if two adjacent intensity bumps on top of a bump below the threshold are actually two separate sources or if they belong together as one source. To do this it determines the fraction of the intensity the bumps would have were they to be regarded as separate to the intensity the bumps would have were they to be regarded as one object. If this fraction is low enough (i.e. the bump is small compared to its surroundings) they will be blended into one object, otherwise there will be two. See figure 3 (extracted from Bertin and Arnouts (1996)) for an illustration of the process.

Setting deblending parameters (detection threshold and deblending fraction, as explained above) for faint galaxies is hard because of the junk you introduce when you set the detection threshold too low. Noise fluctuations will soon be included as objects as well and frankly these are quite hard to distinguish from the really fainter galaxies. Furthermore the deblending fraction needs to be set quite low so you include all patches of faint galaxies in one object, but this can also lead SE to merge objects that are close to each other.

We eventually used the set of parameters in table 3. In this we also optimized background

⁵For an excellent overview of all the options and scenarios when using SExtractor see Holwerda (2005).

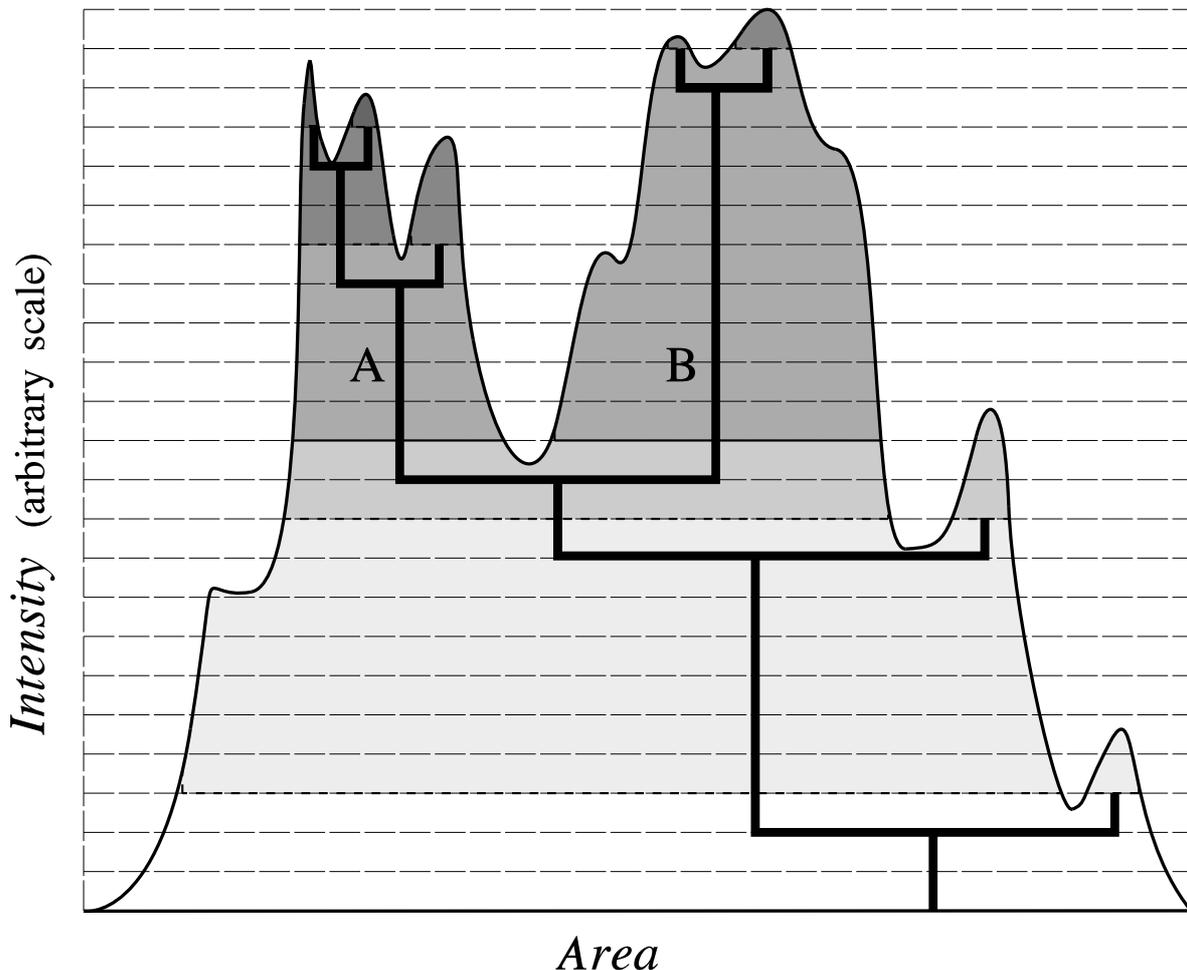


Figure 3: From Bertin and Arnouts (1996): A schematic diagram of the method used to deblend a composite object. The areal profile of the object (the smooth curve) can be described as a tree-structure (thick lines). The decision to regard a branch as a distinct object is determined according to its relative integrated intensity (tinted area). In the case above, the original object is split into two components A and B. Pixels lying below the separation threshold are assigned to their most credible "progenitor" afterwards.

modelling and filter parameters to get the optimal result for the type of galaxies we were looking for. Similar parameters were used in Benítez et al. (2004), corroborating our results. The resulting dataset included quite a lot of false detections, which in the end turned out to be a necessary evil. The exclusion of actual faint galaxies would have been too large otherwise.

We used SExtractor from within Astro-WISE⁶. This has the benefit that some of the

⁶This brings with it the usual Astro-WISE benefits, but also introduces a few restrictions compared to using SE outside of Astro-WISE. You cannot use all SE output parameters because there are no AW database tables for them (e.g. the ISO[0-7] parameters) and it is not possible to obtain more than one fixed-aperture-size magnitude (MAG_APER) per object (whereas outside of AW you can get as much as you want). These are things that could easily be fixed by respectively adding the proper database tables and adding to the source models in the database a hasMany relationship to a separate table of aperture magnitudes. In any case it should not be that using AW lays these kinds of easily solvable restrictions on your work as an astronomer.

Parameter	Value	Parameter	Value
MAG_ZEROPOINT	26.07 (g)	CLEAN_PARAM	1.2
MAG_ZEROPOINT	24.86 (z)	ANALYSIS_THRESH	1.5
BACK_FILTERSIZE	5	BACKPHOTO_TYPE	'LOCAL'
BACK_SIZE	128	BACKPHOTO_THICK	26
FILTER	'Y'	MASK_TYPE	'CORRECT'
FILTER_NAME	'gauss_2.0_5x5.conv'	PHOT_APERTURES	50
WEIGHT_TYPE	'NONE'	PHOT_AUTOPARAMS	[2.5,3.3]
INTERP_TYPE	'NONE'	PIXEL_SCALE	0.05
DETECT_MINAREA	10	GAIN	750 (g)
DETECT_THRESH	1.5	GAIN	1120 (z)
THRESH_TYPE	'RELATIVE'	PHOTFLUX_FRAC	0.5
DEBLEND_NTHRESH	16	STARNNW_NAME	'default.nnw'
DEBLEND_MINCONT	0.025	SEEING_FWHM	0.105 (arcsec)
CLEAN	'Y'		

Table 3: Source Extractor configuration parameters

configuration has already been done and a basic set of output parameters is defined. We added a few non-default output parameters to this set: MU_MAX, MU_THRESHOLD, FLUX_ISOCOR, FLUXERR_ISOCOR, MAG_ISOCOR, MAGERR_ISOCOR, ELONGATION, ELLIPTICITY, FLUX_AUTO, FLUXERR_AUTO, MAG_AUTO, MAGERR_AUTO, MAG_BEST, MAGERR_BEST, FLUX_BEST and FLUXERR_BEST. A description of these and other parameters can be found in Bertin and Arnouts (1996); Holwerda (2005).

3 Dataset clean-up and calibration

The source lists we eventually obtained consisted of 342652 sources in the g-band images and 630995 in the z-band images. Datasets on this order of magnitude need some proper filtering, which actually turned out to be the most time consuming effort of this project.⁷ A lot of false detections had to be filtered out and after that we needed to separate the actual Virgo cluster dwarf galaxies from contaminants like background sources and globular clusters⁸. Before we started doing anything serious with the SE parameters though, we needed to check their quality. We did this by comparing to SDSS data and using this comparison to properly calibrate our data. In this section we will describe these efforts.

3.1 First bulk: false detections and stars

The first step in the clean-up process was getting rid of a good part of false detections and stars to keep us from drowning in them. The first part of the former is actually easily disposed of by simply using a basic ACS/WFC characteristic, namely its PSF (point spread function) size, measured in FWHM (full width at half maximum) of a fitted gaussian. For our ACS/WFC data this was $\sim 0.105''$ which corresponds to ~ 2.1 pixels (Benítez et al., 2004). For simplicity's sake we used this value instead of measuring PSF sizes on all the images themselves. Using this measure we can logically exclude sources if they are smaller than this. Also sized based was our filter on number of pixels; below a certain number, sources are not likely to actually be real.

Next we can filter on blatantly false values of parameters. Some of the sources had *dummy valued* magnitudes of -9999, which indicate that the software ran upon an error while determining the magnitude.

A third filter we could readily apply is that on stars. We accomplished this by using the `CLASS_STAR` parameter, which is a measure of "star-ness" (measure of how much it looks like a star) as determined by a neural network, and the central surface brightness of the stars. Too starry or too centrally bright objects were excluded. The SE parameter closest to central surface brightness is `MU_MAX`, which is the maximum surface brightness of the object. Since the brightest objects in our set are stars which have their maximum value in the center we can indeed use this parameter to filter out stars. We have to be careful though in simply assuming this parameter to be the central surface brightness, as we will see in section 4.3.2.

Finally we made use of SExtractor's built-in magnitude error parameters. If these were too high, i.e. if SE was not able to determine magnitudes properly, the sources would be excluded as well. We list the actual filter criteria (general ones and band specific ones) in table 4. After applying these filters, we were left with 9491 sources in the g-band and 21757 sources in the z-band.

⁷This is probably the reason for the present day interest in data-mining algorithms. Doing all this by hand indeed gives strong feelings that there might be more efficient ways.

⁸The latter are actually one of the prime research targets for the ACSVCS project, which makes sense, as close-ups of large galaxies are bound to contain a lot of globular clusters as well.

General filter criteria	Band specific filter criteria
FLUX_RADIUS > 2.1	g-band
FWHM_IMAGE > 2.1	MAG_ISOCOR < 26
A > 2.1	MAG_ISOCOR > 12
B > 2.1	MU_MAX < 23
NPIX > 10	MU_MAX > 14.15
CLASS_STAR < 0.94	z-band
MAGERR_APER < 1	MAG_ISOCOR < 25
MAGERR_ISO < 1	MAG_ISOCOR > 11
MAGERR_ISOCOR < 1	MU_MAX < 22
MAGERR_AUTO < 1	MU_MAX > 13.35

Table 4: First dataset filters for getting rid of false detections and stars based on PSF size, impossible magnitudes, "star-ness" and errors.

3.2 Association

GalPhot is a good routine, but it cannot perfectly model all galaxies. Especially when modelling very bright galaxies which might have started saturating the CCD in the center⁹ and in cases where there are bright sources that cannot be properly masked (e.g. when it is too close to the center of the modelled galaxy) the residual will leave artifacts that SE might identify as sources. These artifacts will usually not be in both the g- and z-band residuals though as the GalPhot models will certainly not be the same in both bands.

We can thus filter out almost all of these false sources by building so-called associate lists from the g- and z-band source lists. This comes down to simply associating every source in one list with a source in the other list at the same coordinates, within a certain radius. The latter condition means that sources from the first list that have no counterpart within the radius in the other list are not included in the associate list. This means that the GalPhot artifacts, which are band specific, will only be included if by a small chance there happened to be an artifact or a false detection in the other band as well. Of course this chance depends on how large a radius is used within which association is allowed. We set this radius at 0.25", corresponding to 5 pixels. On visual inspection this turned out to include most ($\sim 95\%$) of the sources we wanted to include and indeed excluded a lot of false detections. Unfortunately this also excludes a bit of good data that has not been properly detected in both bands. This loss is minimal though, compared to the amount of clean-up it delivers; this procedure left us with 5838 objects with data in both bands, which means that $\sim 38.5\%$ of the g-band data have been excluded as false detections.

Filtering out objects like this leads to a lot of galaxies in the z-band, that are not detected in the g-band images, being left out. This is not a problem though, because the objects we are looking for are likely to be about equally luminous in both bands. Objects that are far more luminous in the z-band are more likely to be highly redshifted background galaxies and so we have thus excluded a lot of these in this process as well.

A more serious problem is that there might also be real objects in the g-band that are

⁹The ACSVCS project actually took this into account for the z-band by making another short exposure (90 seconds) of all the galaxies to be certain of having accurate central data. We have not used these images, as the shorter exposure times would not have added any better data of the weak dwarfs we are looking for.

not detected in the z-band, i.e. very blue but faint galaxies. These might include starforming dwarfs in the Virgo cluster. On visual inspection though, it seemed the bulk of g-band-only detections was of falsely deblended objects (multiple sources for one object), GalPhot residual detections and galaxies with multiple bright patches that were thus probable background sources (dwarfs do not possess that much structure). There were a few sources that were not associated because their detected centers were too far from each other in the two bands. Of these we did not find any possible dwarfs (mostly they were extended patchy sources as well, so probably background spirals), but of course it is possible that they are present in the excluded dataset (we did not check all images). This is a possible sacrifice we were prepared to make in order to get proper color information on the remaining sources. We from here on simply take this possible defect of very blue galaxies as a given.

3.3 Comparison to SDSS

After these steps in cleaning up the dataset came the calibration of the magnitudes we were left with. Most of the objects at this point were indeed very likely to at least be physical objects, so we could start comparing these to known data. The Sloan Digital Sky Survey (SDSS) was the perfect candidate for this task¹⁰. In this survey, with coverage for about half of the northern sky, the part of the Virgo Cluster we are interested in (i.e. the part that matches with our ACSVCS data) is included fully (Lisker et al., 2008). Although it is less resolved than our images, its photometry is calibrated very accurately and thus it serves us well. The wavelength bands used are *ugriz*, so we can perfectly compare our g- and z-band magnitudes to theirs.

First we needed to couple the sources in our dataset to actual SDSS sources. We used the SDSS Imaging CrossID Upload for DR6 service¹¹ to obtain an SDSS dataset overlapping with our own data. We set *Search type* to *All nearby objects* and search radius to 2.2', which gives us a generous overlap for each image. We obtained right ascension and declination values (among other (structural) parameters determined by the ACSVCS project) from Ferrarese et al. (2006, table 4) to make an RA/DEC upload list. A few output parameters needed to be added to the default ones, so we edited the database query to read the following:

```
SELECT
  p.objID, p.ra, p.dec, dbo.fPhotoTypeN(p.type) as type, p.modelMag_g,
  p.modelMag_z, p.extinction_g, p.extinction_z, p.petroR50_g, p.petroR50_z,
  p.isoA_g, p.isoB_g, p.isoA_z, p.isoB_z, p.petroR90_g, p.petroR90_z,
  p.deVRad_g, p.expRad_g, p.deVAB_g, p.expAB_g, p.deVMag_g, p.expMag_g,
  p.deVRad_z, p.expRad_z, p.deVAB_z, p.expAB_z, p.deVMag_z, p.expMag_z
FROM #x x, #upload u, PhotoObjAll p
WHERE u.up_id = x.up_id and x.objID=p.objID
ORDER BY x.up_id
```

This way we ended up with a dataset of all SDSS sources covering the area of the ACSVCS data. Next up was matching these sources to our own. TOPCAT's Pair Match function was chosen for this task, using its Sky match routine to look for the best matching objects within a 5.0" radius. Unfortunately, this again excluded a great part of our dataset. We were left with

¹⁰See Adelman-McCarthy et al. (2008) for a description of the contents of the latest data release 6

¹¹URL: <http://cas.sdss.org/astrodr6/en/tools/crossid/upload.asp>

3050 matches to SDSS out of a total of 5838. These we thus used for comparison. Note though that after comparison we again used all 5838 sources, corrected using the factors determined for the 3050 matches. Most of the sources that are not matched are very faint and hence were probably simply not detected in SDSS. The HST obviously has less background noise, whereas SDSS data have to cope with sky noise. The brighter sources that were not in SDSS have slightly different model values so that they simply do not match perfectly in TOPCAT. In the end it does not matter that we do not take these very faintest sources into account in our calibration process, as these will have too much scatter to be of any use anyway.

Now we had our matched sources we needed to choose what to compare. We used both SExtractor's and the SDSS project's algorithms for determining the best magnitude. SDSS contains the 'best fit magnitude' parameter `modelMag` which is either `expMag` or `deVMag` (which are based respectively on an exponential and a De Vaucouleurs model) depending on which model produces the best fit to the data. In SExtractor there is a slightly different algorithm for the `MAG_BEST` parameter which chooses between `MAG_AUTO` and `MAG_ISOCOR`.¹² The former is determined using a Kron model (Kron, 1980) and the latter using a circular aperture including all the pixels that were identified as belonging to the source and centered on the source's determined coordinates. SE chooses to use `AUTO` in most cases (93% for our dataset) and `ISOCOR` when there is more than 10% influence from neighbouring sources.¹³

To add a bit to the above loss (the one leaving us with 3050 out of 5838), we had yet to exclude bad SDSS data, i.e. data with magnitudes of -9999. Our SDSS sample actually contained quite a few sources with -9999 values for a lot of the parameters (especially the modelled radii). For our purposes though only sources with -9999 values for `modelMag_g` and `modelMag_z` need to be excluded, as these will be used for comparison. This reduced our final set of sources to use in our SDSS comparison to 2970.

In figure 4 the differences in magnitude between our data and SDSS are plotted. A great deal of scatter is seen here for which we tried to find some explanations.

First we must note that although we used a best match algorithm in TOPCAT, this does not guarantee that the matched sources are actually the same (there might have been false detections in either our data or in SDSS) or that they have been analysed in the same way (central coordinates and thus modelled magnitudes might differ). To illustrate that indeed this plays a role we show in figure 5 that when we leave out sources that actually had a larger coordinate separation (coordinates in our data minus coordinates in SDSS) than 1", the scatter is reduced.

Second there is the fact that we are simply looking at quite low magnitudes at which SExtractor (and presumably SDSS as well) is reaching its limits of accuracy. As can be seen in Bertin and Arnouts (1996, figure 4), offsets between modelled and true magnitudes can go up quite rapidly for weaker sources, especially galaxies.

Third we may explain the differences by the fact that the models used are not the same; Kron (or in 7% of the cases circular aperture) in SE versus exponential or De Vaucouleurs in SDSS. Although the optimization algorithms make sure that the most accurate one is chosen, they will still differ because of different models, which again could give us a few fractions of magnitudes difference (within SDSS itself the two magnitudes differ on average by 0.11 magnitudes with a standard deviation of 0.34, and the SE magnitudes differ by -0.04

¹²See Bertin and Arnouts (1996, paragraph 6) for a detailed explanation of this SE parameter.

¹³Our choice for `MAG_BEST` is also backed up by Häussler et al. (2007) in which an extensive comparison of magnitude determination algorithms is done.

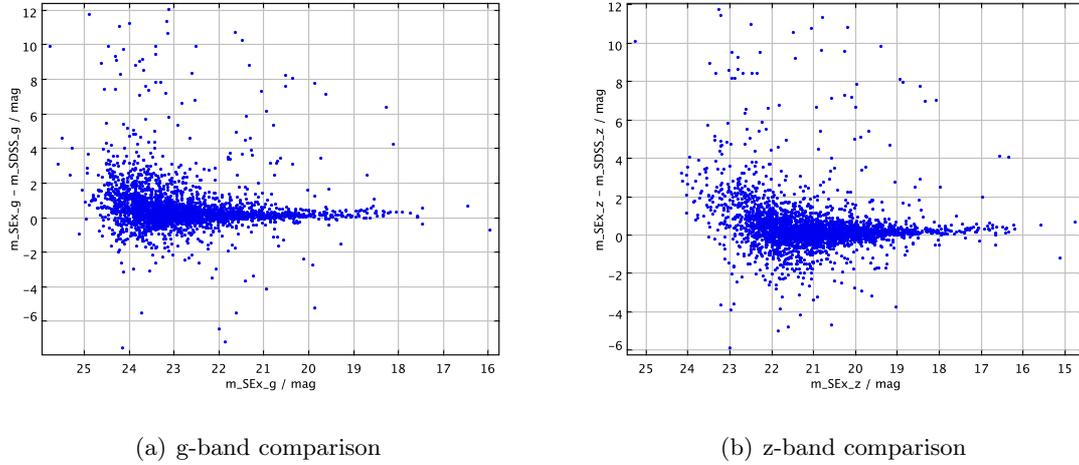


Figure 4: Difference between the magnitudes in our data and SDSS magnitudes (y-axis), plotted against our magnitudes (x-axis).

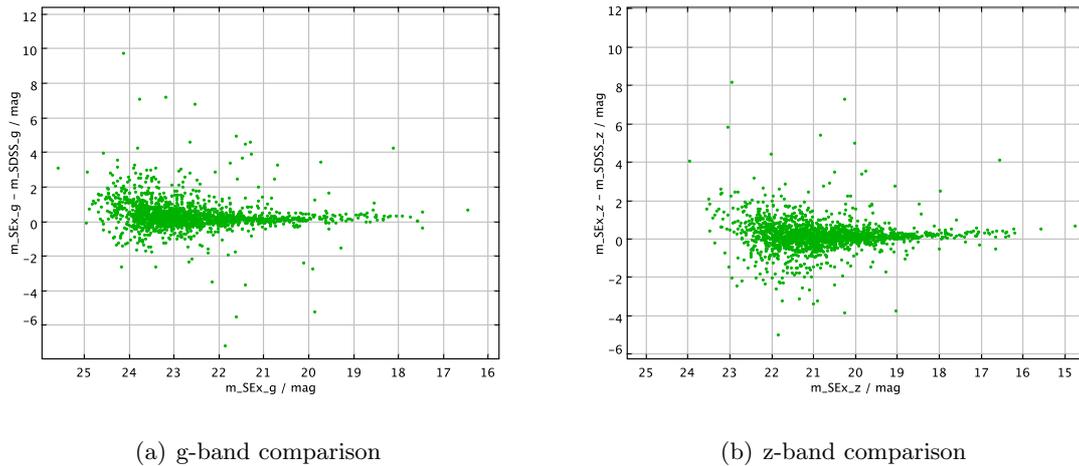


Figure 5: Difference between the magnitudes in our data and SDSS magnitudes (y-axis), plotted against our magnitudes (x-axis). Sources with a coordinate separation larger than 1'' are filtered out.

magnitudes with STD of 0.24, which illustrates that this suggested cause of scatter is highly likely).¹⁴

These scatter related arguments taken into account there still seems to be a general offset, especially visible for the brightest sources. This we have corrected for in the entire set by taking the 100 brightest galaxies and computing their median offset value. This gives us offsets (our data minus SDSS data) of 0.12 in g and 0.20 in z which we subtract from all the

¹⁴For a comparison between Sérsic (and thus exponential and De Vaucouleurs) and Kron profiles (among others) see Graham and Driver (2005).

magnitudes in our dataset. We have thus taken into account all we can for a dataset of this size, given the inherent modelling inaccuracies mentioned above.

The cause of these offsets, i.e. apart from the scatter, can be identified as twofold: first there is the fact that the transformation of ACS magnitudes to SDSS magnitudes requires a 0.02 magnitudes correction, as given in Pavlovsky (2006, table 3.2). This does not improve matters though as it leaves us with respectively 0.14 and 0.22 mag offsets. The other cause for the offsets can be found in the Kron aperture used for the `MAG_AUTO` parameter. This parameter as defined by SExtractor only determines about 90% of the total magnitude. To make matters worse the exact percentage is dependent on the degree of central concentration (e.g. characterized by the Sérsic n parameter). For an extensive discussion of these effects see Graham and Driver (2005), where it is explicitly assumed that the 'total magnitude' is determined by a Sérsic profile. For our purposes it suffices to say that an average offset of $(0.14 + 0.22)/2 = 0.18$ can indeed be explained by this effect. Taken into account that part of this offset and in particular the difference in offsets between both bands is caused by differences in scatter, we see that an offset on this order of magnitude comes from:

$$M_{\text{real}} - M_{\text{kron}} = 2.5 \log(L_{\text{kron}}/L_{\text{real}}) = 2.5 \log(0.8 * L/L) = 2.5 \log(0.8) = -0.24 \quad (17)$$

Here we take the value of L_{kron} to be 0.8 times the L_{real} because this is about the average for Sérsic n values between 0 and 2 (Graham and Driver, 2005, figure 10), which are the types of objects we are supposedly looking at. In any case, we clearly see that this is indeed a probable cause of the ~ 0.2 offset. To further ensure this is correct we also took a number of stars from our original unfiltered set and compared them to SDSS magnitudes. Stars have far steeper profiles than late-type and dwarf galaxies, so according to Graham and Driver (2005, figure 10) we would expect the light fraction $L_{\text{kron}}/L_{\text{real}}$ to drop even further to about 0.6. This would then give us an expected offset to SDSS of $2.5 \log(0.6) = -0.55$ mag. This is indeed what we found to be the case, so we can take this explanation to be pretty solid.

Further on in this report we compare our data to data from Binggeli and Jerjen (1998) and Ferrarese et al. (2006). We therefore also calibrated these data in the same way as reported above. The correction factors in the g-band turned out to be 0.30 and 0.16 respectively (SDSS minus their data).

The value of 0.16 for the Ferrarese et al. (2006) data is on the same order of magnitude as the value for our data, which would make perfect sense given that they are using the same images. Their magnitudes though are integrated Sérsic intensity profiles, so that argument runs awry. A closer look at Ferrarese et al. (2006) tells us that the magnitudes are actually Sérsic total magnitudes *minus* the luminosity of a nucleus when present. This indeed should increase the magnitudes with respect to SDSS, as there are no nucleus subtractions made in SDSS. Unfortunately these nucleus magnitudes are on average 7 magnitudes fainter than the total magnitudes, so this will only give a difference of 0.002 mag; not what we are looking for either. Finally then, the only possibility is that the best fitting Sérsic model as used in Ferrarese et al. (2006) simply gives them a fainter magnitude than the exponential or De Vaucouleurs fits from SDSS. Unsatisfying as it may be, this is the only answer we can come up with at the moment.

The greater part of the offset from the Binggeli and Jerjen (1998) data is due to the different bands used in their paper; this gives a 0.21 mag correction factor, leaving us with 0.09 mags to explain. This offset is probably caused by slight differences in magnitude determination as well. The Binggeli and Jerjen (1998) magnitudes used are 'measured' as opposed to

modelled, although it is not completely clear from the paper in what way they are measured. Assuming they simply used some kind of aperture measurement in the images this could explain the discrepancy with SDSS, as SDSS contains modelled magnitudes.

4 Analysis

With our dataset cleaned up and calibrated we could analyse its contents. In figure 6 we show the distribution of magnitudes in our dataset after filtering and calibration. The distribution in z-band magnitudes is wider and brighter, which immediately indicates a large population of redshifted background galaxies. The first thing we need to do then is to separate these from our possible dwarfs. Another thing that may be said by looking at these histograms is that the limit for reasonable completeness is reached at about a g-band magnitude of -8 . This means that we can say with reasonable certainty that the distribution in figure 1 is only valid up to that point (g to B band conversion not taken into consideration, which as we will see below is a factor of 0.21). We would then be left with a total of some 15 possible detections.

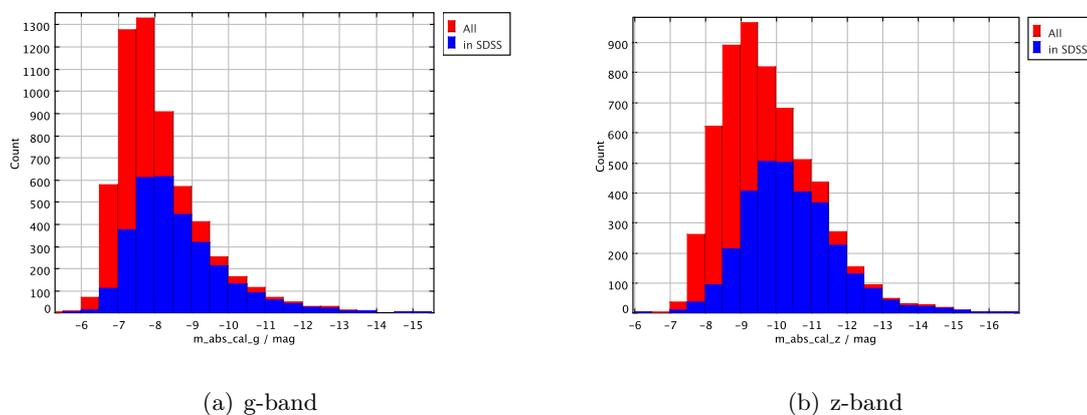


Figure 6: "Absolute" magnitude distribution for our dataset after cleaning and SDSS calibration. In red the entire dataset is shown, in blue only the parts that had a matching object in SDSS. Of course these are no true absolute magnitudes, as the majority of galaxies are in the background and hence have different distance moduli. Rather they are simply the apparent magnitudes minus the distance modulus to Virgo of 31.2.

4.1 Background separation

Our first measure for separating background from Virgo galaxies is a comparison to results from Graham and Guzmán (2003). In this paper data are collected from several other papers like Binggeli and Jerjen (1998); Stiavelli et al. (2001) to display relations between luminosity, surface brightness, effective radius and Sérsic index. We have also used these data, but added results from Ferrarese et al. (2006).

In this comparison B-band magnitudes are used, so we needed to calculate a conversion value from g to B. We obtained B-band magnitudes from van Zee et al. (2004, table 2) and g-band magnitudes for the same galaxies from SDSS. Corrected for extinction, the average conversion value (barring two outliers) turned out to be $B - g = 0.21$ (in this, extinction for SDSS data of the Virgo Cluster could be averaged at 0.1 magnitudes, which means we will also have to apply this extinction correction to our data).

The first comparison we are making is that of the surface brightness versus luminosity plot. To do this we need to determine a measure of surface brightness for our data. One

of the measures used in Graham and Guzmán (2003) is average surface magnitude within the effective radius (half-light radius): $\langle\mu\rangle_e$ (see appendix B for a derivation of the quantity from our SE data). Although they do not use this relation to prove one of the points in their paper¹⁵ it is fine for our purposes, which is comparing dwarfs to other dwarfs.

Using the definition of $\langle\mu\rangle_e$ from appendix B and applying corrections for respectively SDSS calibration, extinction and g- to B-band transformation we get:

$$\langle\mu\rangle_{e,g} = \text{MAG_BEST} + 2.5 \log(\pi \text{ FLUX_RADIUS}^2 0.05^2) + 0.75 + 2.5 \log(A/B) \quad (18)$$

$$\langle\mu\rangle_{e,B} = \langle\mu\rangle_{e,g} - 0.12 - 0.1 + 0.21 \quad (19)$$

$$\langle\mu\rangle_{e,B} = \text{MAG_BEST} + 2.5 \log(\pi \text{ FLUX_RADIUS}^2 0.05^2) + 2.5 \log(A/B) + 0.74 \quad (20)$$

Here we take the SE parameters from the g-band source lists. The data from Binggeli and Jerjen (1998) also need to be converted, as they list central surface brightnesses. Together with Sérsic n though, which is given, we can use these to calculate the values we need (see appendix B). We then plot these surface brightnesses against luminosity (absolute magnitude, also calibrated, converted and extinction corrected) in figure 7.

Using the relationship between $\langle\mu\rangle_e$ and M and approximately extrapolating from the Ferrarese et al. (2006) and Binggeli and Jerjen (1998) data, we can now exclude with a decent amount of certainty the group of sources that is outside this relationship as background galaxies. See figure 8 for an illustration of the cut we made using TOPCAT. This cut left us with 327 sources.

Now to further filter out background galaxies we have used a color filter. To determine a proper filtering range we have compared to color tables of Virgo dwarfs in Lisker et al. (2008). From their data we determined an average color of $g - z = 1.794 \pm 0.218$. However, this is for fairly bright galaxies so we need to take the slope in the color-magnitude relation into account. This can be averaged at about -0.05. Given the range used in their paper ($M \gtrsim -13$) and our range ($M \gtrsim -6$), we added five times this slope to our average and its absolute value to our deviation. We then arrive at a color range of $g - z = 1.5 \pm 0.5$ which should theoretically include all Virgo dwarfs. Because we want to use this filter to get out background galaxies we only use this measure to set an upper limit of $g - z < 2$. This will exclude highly redshifted background galaxies and prevent possible blue dwarfs from dropping out. After this we were left with 266 sources.

Using our estimates for size in paragraph 1.2, we can exclude 18 more, namely those with radii > 50 pixels.

Our final filter is based on profile flatness. Dwarf galaxies typically fit best to an exponential profile, i.e. to a Sérsic profile with $n \approx 1$ ($n = 1.2$ for a fit to the Draco Dwarf in Odenkirchen et al. (2001), $n = 0.6, 0.8$ and 0.7 for Carina, Sculptor and Fornax respectively in Walcher et al. (2003), a range of dwarfs listed in Jerjen et al. (2000, tables 2 and 3) with values $0.5 < n < 2$ and even simulations from Read and Gilmore (2005) show that dwarf spheroidals (a subclass of dwarf galaxies) are well fitted by a $n = 1$ profile). This is a flat profile, whereas larger galaxies usually fit to a profile with a much steeper central brightness. This means that in a dwarf galaxy, the difference between μ_0 and $\langle\mu\rangle_e$ will be smaller than for larger ones, so we can use this to separate potential dwarfs from other galaxies. Note

¹⁵Which is that in a central surface brightness versus luminosity relation dwarfs and large ellipticals seem to belong to the same continuum of objects, whereas in a $\langle\mu\rangle_e$ versus luminosity relation (which was used a lot before) there seems to be a clear dichotomy between these two types of objects.

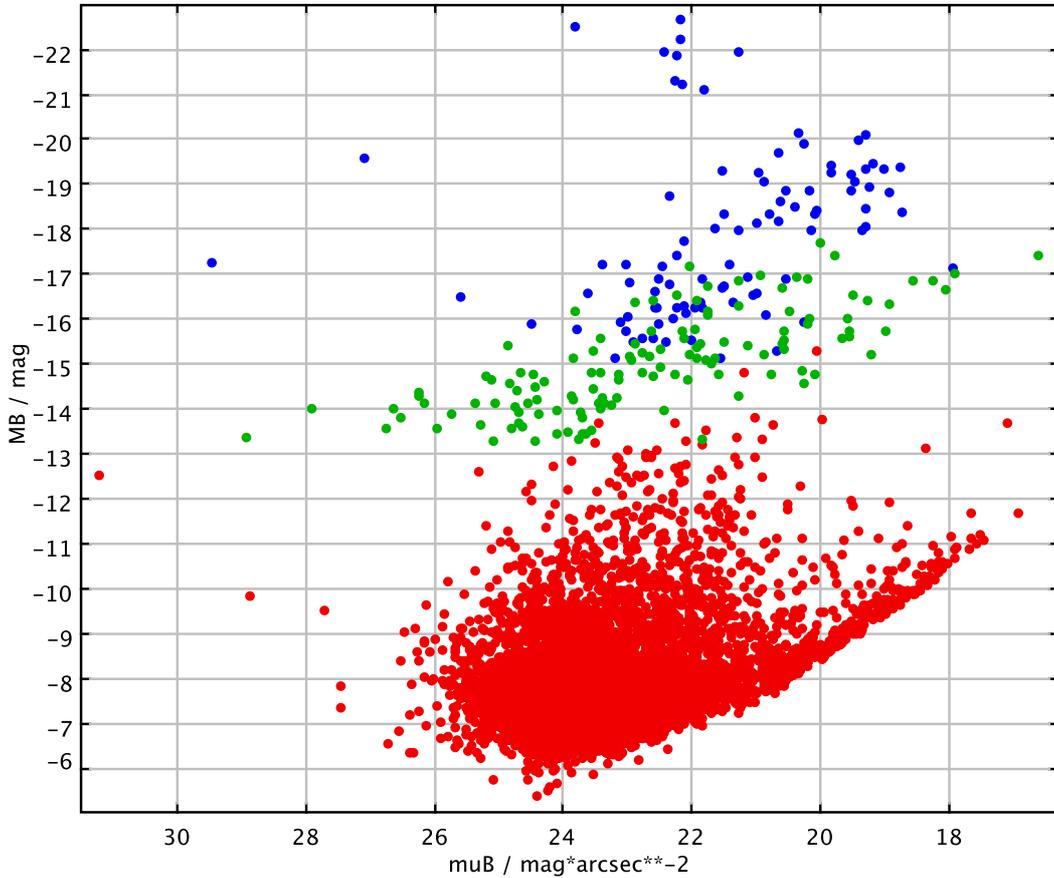


Figure 7: Average surface magnitude within the effective radius in the B-band $\langle \mu \rangle_{e,B}$ versus absolute magnitude in the B-band M_B . Data are from Ferrarese et al. (2006) (blue), Binggeli and Jerjen (1998) (green) and our own data (red). The upper group of blue galaxies are in fact no longer dwarfs, but other types from the ACSVCS.

that this will exclude two very interesting (but probably not present in our data) subsets of the dwarf(-like) family: ultra-compact dwarfs and compact ellipticals. These are very rare systems in the range $0 < n < 11$ (Evstigneeva et al., 2007, 2008). Interesting as they may be, the chance of finding them is too slim to justify not using this filter for throwing out early-type background galaxies. In appendix B we show numerically that when we take limits to the Sérsic n parameter of $0 < n < 2$ this means that only sources for which $0 < \langle \mu \rangle_e - \mu_0 < 3$ are included. Since we have these parameters in both g- and z-band images, we can even apply this filter twice. This finally leaves us with 120 sources.

To resume, here are the steps we have taken to exclude background galaxies:

1. Extrapolate (within wide margins) $\langle \mu \rangle_{e,B}$ versus M_B relationship from Graham and Guzmán (2003).

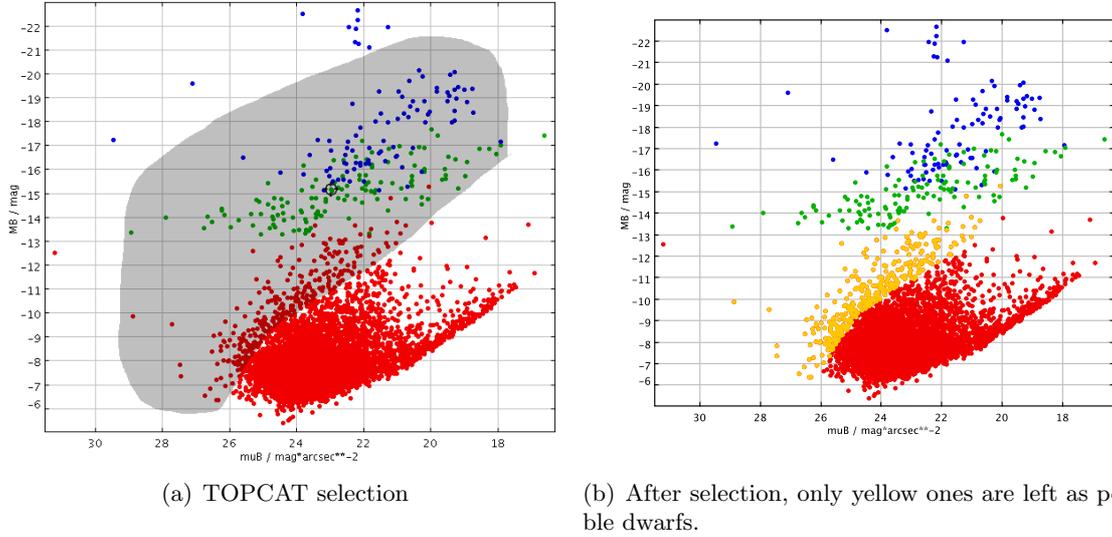


Figure 8: Same plot as figure 7, but with a first exclusion of background galaxies based on the relationship from Graham and Guzmán (2003).

2. Filter on color $g - z$, using data from Lisker et al. (2008); $g - z < 2$.
3. Filter on radii using our expectations from paragraph 1.2: $r_e < 50$ pixels (both bands).
4. Filter on profile flatness using relations from Graham and Driver (2005): $0 < \langle \mu \rangle_e - \mu_0 < 3$ (both bands).

4.2 Close-up inspection

Now that we have managed to extract a manageable selection of sources we can take a closer look at them, i.e. by eye. We have taken images of all 120 sources. For comparative reasons we have first taken all images with the same logarithmic color scale matched to the same pixel values. This color scale is displayed in figure 15. In appendix E all the source images are displayed (in grey frames).

We then started a rough classification by eye to sort out the still invasive background galaxies. We obtained the following groups:

- Background / certainly not dwarfs:
 - Spirals: 4, 5, 6, 7, 9, 10, 12, 13, 14, 15, 17, 19, 20, 21, 22, 23, 24, 25, 26, 32, 35, 36, 37, 38, 40, 42, 43, 45, 46, 51, 54, 56, 57, 58, 59, 60, 61, 66, 69, 74, 79, 81, 82, 83, 87, 91, 98, 106
 - Other (high central slopes, clearly redshifted, collisions, too much substructure, etc.): 2, 8, 18, 31, 34, 39, 47, 48, 49, 50, 53, 55, 70, 71, 72, 76, 78, 93, 96, 110, 113, 117
- Junk:
 - GalPhot: 3, 41, 44, 52

– Too close to the edge: 77, 102, 118, 119, 120

- Possible dwarfs: 1, 11, 16, 27, 28, 29, 30, 33, 62, 63, 64, 65, 67, 68, 73, 75, 80, 84, 85, 86, 88, 89, 90, 92, 94, 95, 97, 99, 100, 101, 103, 104, 105, 111, 112, 114, 115, 116

For some images the GalPhot process made the background negative (see for example sources 68, 86 and 88), so for these we had to customize the value range a little. We did this in such a way that the background has about the same color in all images so you can optimally compare both bands. We also applied this to the images of possible dwarfs to make sure they are dwarfs. The result of this custom ranging is in the second set of images (with red frames) in appendix E. On this basis we made the following final classification of the above "Possible dwarfs" category:

- Background / certainly not dwarfs:
 - Spirals: 1, 16, 27, 29, 63, 64, 67, 68, 86, 89, 101
 - Other: 28, 30, 80, 84, 85, 90, 92, 95, 99, 100, 104, 105, 107, 108, 112, 115, 116
- Too close to the edge: 88
- Possible dwarfs: 11, 33, 62, 65, 73, 75, 94, 97, 103, 109, 111, 113, 114

This then gave us a final selection of 13 dwarf candidates. We will shortly classify the individual dwarf candidates using their characteristics, justifying inclusion. Of course, as with any kind of qualitative classification, the classifications below are subject to discussion. We have used a Level 5 webpage¹⁶ on dwarf galaxy classification that in turn is based on van den Bergh (1959) (which we were not able to get our hands on) as our guide.

- 11 Slightly extended source with a flat surface brightness profile. No clear structural features apart from a few blue regions, which might also be background sources. Likely to be a dwarf spheroidal (dSph) or dwarf irregular (dIrr). Brighter in g than in z, indicating low redshift and thus Virgo cluster membership.
- 33 Same goes for this one, except for the blue regions. It also has a little less spherical symmetry and might thus better be classified as a dIrr.
- 62 Apart from the same general features as source 11, this source displays a higher surface brightness core. This would indicate its being a nucleated dwarf spheroidal (dSph,N).
- 65 Same story as source 11; dSph or dIrr, more likely to be a dIrr though.
- 73 Again, dSph or dIrr.
- 75 dSph or dIrr.
- 94 Same as source 62, though with a less bright nucleus; dSph,N.
- 97 This is probably also a dSph,N, though with a blue nucleus.
- 103 dSph/dIrr (depending on whether the southern blob is part of the galaxy (dIrr) or not (dSph)).

¹⁶URL: <http://nedwww.ipac.caltech.edu/level5/CLASSIFICATION/dwga.html>

109 dSph,N.

111 dIrr.

113 dIrr.

113 dIrr.

We see no dwarf ellipticals (dE) or dwarf spirals (dS). The lack of the former was to be expected, as about a third of the images were actually targetted at known dEs, thus biasing the distribution and ruining the Schechter function's predictive power. The latter are in the first place quite rare, especially in clusters where gravitational influences will readily destroy their structure, and secondly they might have simply been filtered out by eye in the above process. Also, dSphs and dIrrs, being at the faint/small end of the classification scheme, are basically what we have optimized our search parameters for in the first place, so this outcome should not be surprising at all.

A final test to make them eligible for dwarfhood in this report could have been a test of Virgo cluster membership, which could have been performed for those objects that are in SDSS or in NED (NASA/IPAC Extragalactic Database¹⁷). We could then have collected photometric redshifts for them from SDSS, which to be honest would not have been very accurate for faint galaxies in general and for dwarfs specifically as they are poorly calibrated in the photometric redshift algorithm (Csabai et al., 2003). We herefore chose not to do this. NED could have provided us with additional data like spectra to compute redshift. Unfortunately though, none of our objects were catalogued in NED. The above selection is then our final dwarf set. They are once again grouped in the final images (the ones with the green frames) in appendix E.

4.3 Parameters

Here we study the SE and derived parameters for the identified dwarfs. We list the most important SE parameters in table 5. In table 6 we list a number of derived parameters used in the plots elsewhere in this document and also the Astro-WISE SID/SLID pairs that identify the sources in the AW database.

We give a brief description of the columns in these tables. Table 5: ID is the image number, b is the color-band, α and δ are right ascension and declination, X and Y are image coordinates, m is apparent magnitude, f is flux, μ_0 is central surface magntiude, f_0 is the corresponding central flux, σ_m and σ_f are the errors in m and f respectively, r_e is the effective radius (the half-light radius), e is ellipticity ($e = 1 - b/a$ where a and b are semi-major and semi-minor axes respectively), θ is the position angle, f_{bg} is the local background flux and pix is the number of pixels SE identified as part of the source, before extracting photometry. Table 6: ID is image number again, $m_{x,c}$ are calibrated apparent magnitudes, $M_{x,c}$ are calibrated absolute magnitudes, M_B is the absolute B-band magnitude, $\langle\mu\rangle_{e,x}$ are the mean surface magnitudes within the effective radius, and SID/SLID $_x$ are the Astro-WISE source identifiers.

¹⁷URL: <http://nedwww.ipac.caltech.edu/>

ID	b	α	δ	X	Y	m	f	μ_0	f_0	σ_m	σ_f	r_e	e	θ	f_{bg}	pix
011	g	186.95169	13.075591	1528	206	18.7734	829.1249	19.8184	0.7917	0.0032	2.4807	46.8534	0.163	83.43	-0.0258	19342
011	z	186.95174	13.075596	1527	203	18.2285	449.3574	19.2318	0.4459	0.0043	1.7850	42.1124	0.137	-80.69	-0.0330	15386
033	g	187.69302	16.789017	3673	214	19.9683	275.8572	21.7316	0.1359	0.0087	2.2205	44.0733	0.159	0.44	-0.0318	4725
033	z	187.69301	16.789043	3674	213	19.3454	160.6420	21.0208	0.0858	0.0100	1.4778	35.7286	0.189	-13.52	0.0004	4080
062	g	186.68283	9.572543	503	1049	20.7599	133.0517	21.4980	0.1685	0.0093	1.1408	25.1402	0.268	-40.25	-0.0035	2658
062	z	186.68282	9.572548	504	1049	19.9243	94.2483	20.3600	0.1577	0.0086	0.7439	23.4345	0.275	-45.08	-0.0043	2396
065	g	191.06894	12.963397	292	1500	20.8200	125.8876	22.2631	0.0833	0.0193	2.2401	25.5181	0.297	37.42	0.0030	3967
065	z	191.06897	12.963407	291	1499	20.2400	70.4713	21.6142	0.0497	0.0212	1.3739	24.1878	0.189	40.74	-0.0038	3097
073	g	189.13203	11.441881	2802	3710	21.1964	89.0100	21.9531	0.1108	0.0178	1.4564	25.2708	0.582	-83.92	0.0008	1662
073	z	189.13203	11.441875	2802	3710	20.8704	39.4314	21.6004	0.0503	0.0248	0.8998	21.4301	0.345	-79.78	-0.0000	908
075	g	187.71855	16.751768	1619	2679	21.2259	86.6263	21.8960	0.1168	0.0205	1.6319	24.2064	0.510	-51.57	0.0011	2168
075	z	187.71857	16.751774	1619	2680	20.3236	65.2449	21.1819	0.0740	0.0106	0.6356	24.0777	0.528	-38.83	-0.0005	1691
094	g	188.87730	12.207076	1160	2795	21.6154	60.5098	21.9363	0.1126	0.0217	1.2082	20.6689	0.537	-24.30	0.0003	1244
094	z	188.87727	12.207104	1161	2793	20.5477	53.0797	20.4221	0.1490	0.0114	0.5553	17.7011	0.479	-24.96	-0.0015	1630
097	g	190.46481	9.422692	3418	1893	21.8024	50.9372	22.1544	0.0921	0.0200	0.9399	20.8703	0.646	5.72	-0.0015	895
097	z	190.46480	9.422675	3416	1893	21.1670	30.0049	21.6952	0.0461	0.0220	0.6087	20.1007	0.632	10.20	-0.0010	700
103	g	186.28790	12.873319	2629	780	22.2620	33.3588	22.2004	0.0883	0.0411	1.2612	17.4559	0.633	29.13	0.0014	441
103	z	186.28793	12.873364	2625	779	21.1634	30.1037	21.4183	0.0595	0.0387	1.0728	15.6081	0.406	23.99	0.0000	557
109	g	186.01205	11.197648	545	2888	22.5562	25.4411	22.0202	0.1042	0.0297	0.6951	15.7786	0.632	-8.62	-0.0035	472
109	z	186.01207	11.197608	544	2891	21.4642	22.8197	20.9735	0.0896	0.0209	0.4386	13.2797	0.536	-12.85	-0.0042	613
111	g	186.39774	12.827759	2773	358	22.6655	23.0045	22.5227	0.0656	0.0415	0.8782	16.4800	0.569	58.63	0.0018	312
111	z	186.39777	12.827776	2775	360	21.7375	17.7418	21.6635	0.0475	0.0358	0.5844	15.2566	0.367	50.12	-0.0096	329
113	g	188.52689	12.031106	378	1512	22.7194	21.8897	22.5835	0.0620	0.0367	0.7399	15.5424	0.688	70.75	-0.0182	344
113	z	188.52687	12.031132	380	1511	21.2585	27.5791	21.4346	0.0586	0.0204	0.5175	15.0826	0.504	80.31	-0.0177	825
114	g	190.71669	10.668942	573	1167	22.8395	19.5973	22.4751	0.0685	0.0348	0.6277	13.0304	0.783	66.58	0.0001	380
114	z	190.71663	10.668938	571	1164	21.8911	15.4009	21.8332	0.0406	0.0336	0.4761	14.8261	0.642	66.69	-0.0006	367

Table 5: SExtractor parameters for our 13 dwarf galaxies.

ID	$m_{g,c}$	$m_{z,c}$	$g - z$	$M_{g,c}$	$M_{z,c}$	M_B	$\langle \mu \rangle_{e,B}$	$\mu_{0,B}$	$\langle \mu \rangle_{e,g}$	$\langle \mu \rangle_{e,z}$	SID/SLIDg	SID/SLIDz
011	18.6534	18.0285	0.6249	-12.5466	-13.1715	-12.4366	22.7978	20.0284	22.5878	21.6989	0/292251	1/290661
033	19.8483	19.1454	0.7029	-11.3517	-12.0546	-11.2417	23.8548	21.9416	23.6448	22.5252	0/291651	0/292161
062	20.6399	19.7243	0.9156	-10.5601	-11.4757	-10.4501	23.5781	21.7080	23.3681	22.3112	494/290501	464/292151
065	20.7000	20.0400	0.6601	-10.5000	-11.1600	-10.3900	23.7143	22.4731	23.5043	22.5733	578/291941	497/291891
073	21.0764	20.6704	0.4060	-10.1236	-10.5296	-10.0136	24.6337	22.1631	24.4237	23.1730	3227/291211	7348/291241
075	21.1059	20.1236	0.9822	-10.0941	-11.0764	-9.9841	24.3980	22.1060	24.1880	23.2341	1428/291651	4916/292161
094	21.4954	20.3477	1.1478	-9.7046	-10.8523	-9.5946	24.5052	22.1463	24.2952	22.6831	1959/291441	4644/291221
097	21.6824	20.9670	0.7154	-9.5176	-10.2330	-9.4076	25.0065	22.3644	24.7965	23.9548	1072/291321	1327/292211
103	22.1420	20.9634	1.1785	-9.0580	-10.2366	-8.9480	25.0380	22.4104	24.8280	22.8841	217/291191	391/291551
109	22.4362	21.2642	1.1719	-8.7638	-9.9358	-8.6538	25.1093	22.2302	24.8993	23.1005	2315/290691	3615/291471
111	22.5455	21.5375	1.0080	-8.6545	-9.6625	-8.5445	25.1416	22.7327	24.9316	23.3384	124/291871	55/292661
113	22.5994	21.0585	1.5408	-8.6006	-10.1415	-8.4906	25.4199	22.7935	25.2099	23.1009	880/290841	746/291791
114	22.7195	21.6911	1.0284	-8.4805	-9.5089	-8.3705	25.5492	22.6851	25.3392	24.0490	591/291851	539/291631

Table 6: Derived parameters and Astro-WISE identifiers for our 13 dwarf galaxies.

4.3.1 Distribution

Let us now first take a look at the distribution of our dwarf galaxies on the sky to see if patterns emerge.

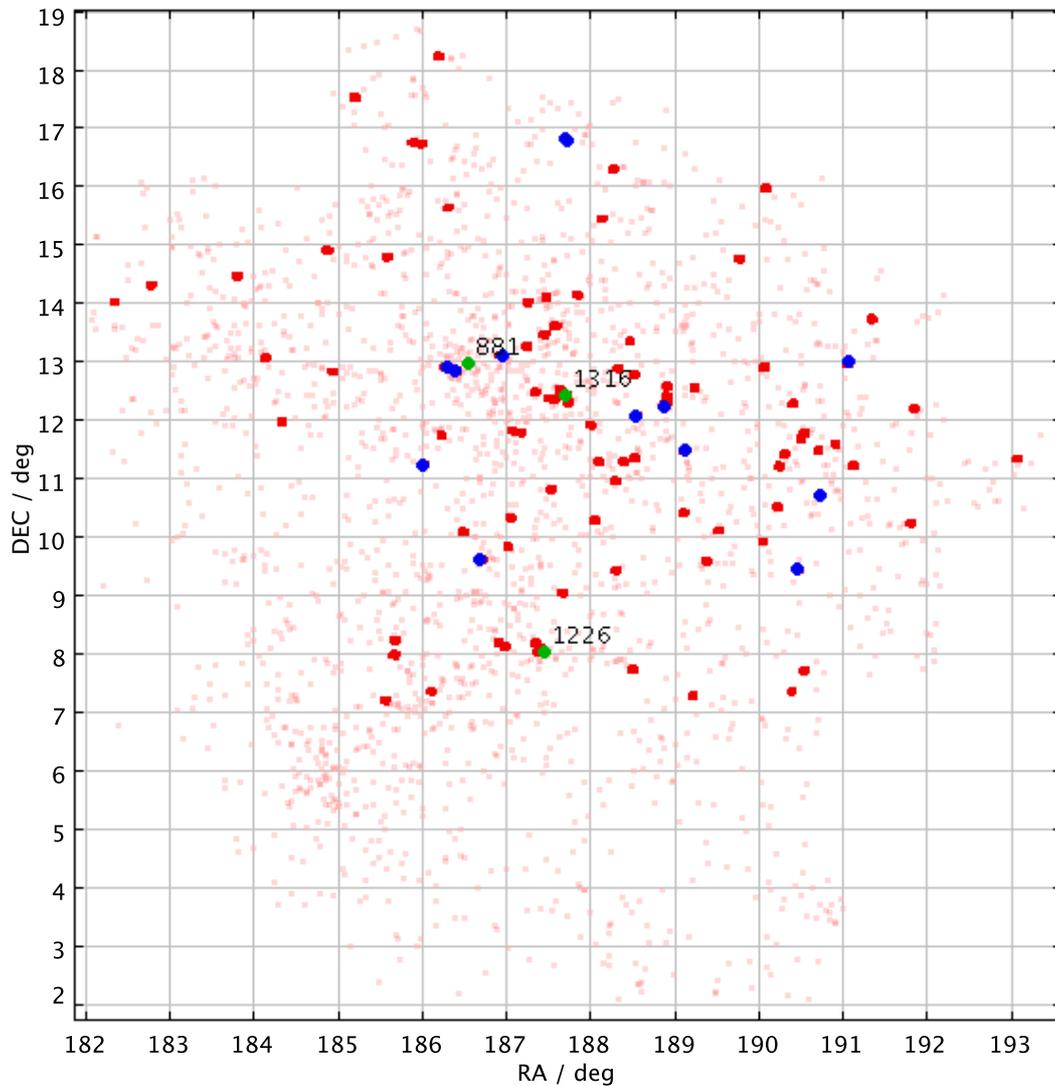


Figure 9: An overview on the sky of our sources. The faint pink dots are all the Virgo Cluster Catalog galaxies, the red dots are all sources in our dataset (the one containing 5838 sources), in blue are the locations of the dwarfs we have found and the green dots with the numbers above them are the three dominating Virgo cluster galaxies. 1316 is M87 (the one with the jet), 881 is M86 (the blueshifted, and thus probably infalling one) and 1226 is M49 (brightest Virgo cluster member).

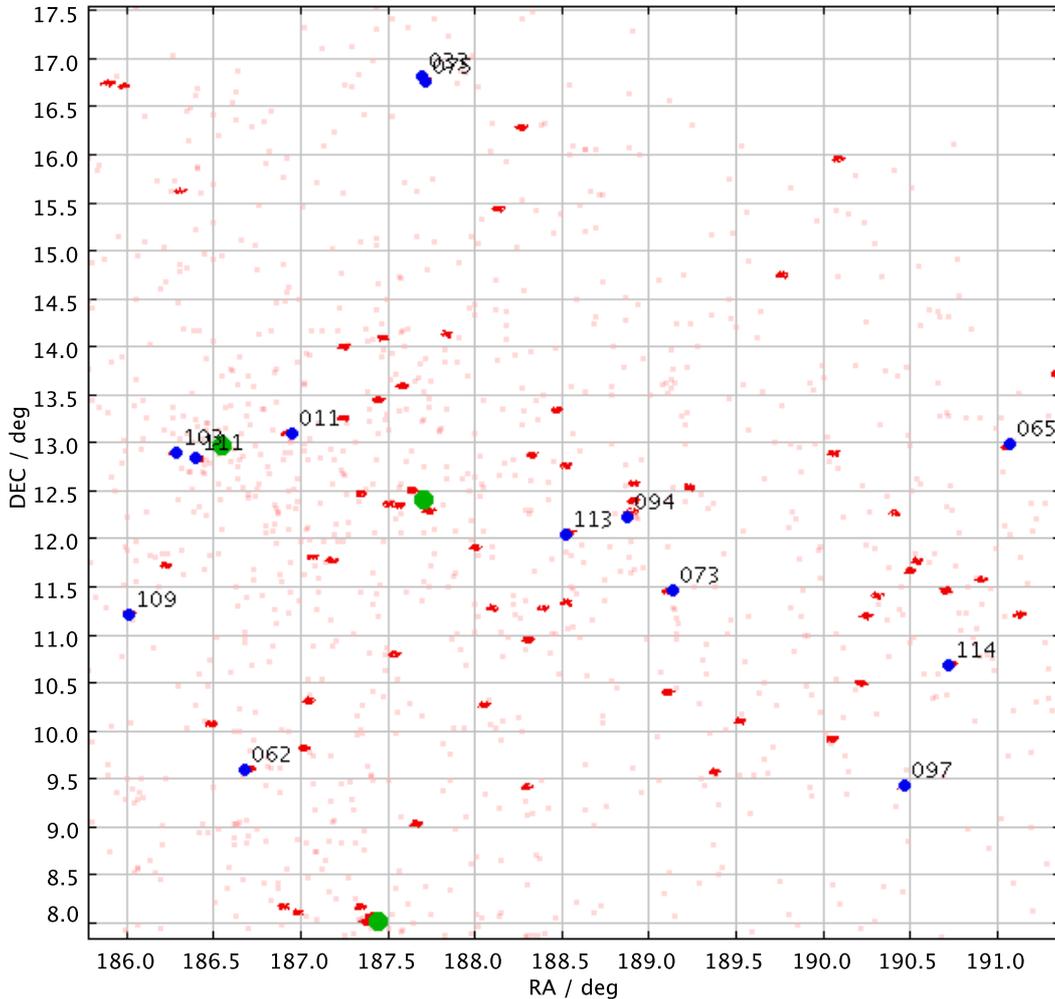


Figure 10: A zoomed in version of figure 9 using the same colors. Numbers here are the numbers we have given the dwarfs in appendix E. At the top there are two dots for which the numbers are not readable because the numbers got printed over each other. The left one is source 33, the right one is 75. At the left near M86 the numbers are not perfectly readable either; the left one is 103, the right one 111.

It seems that there might be four subgroups we can identify from figure 10: a central one around M86 (11, 103, 111), another central one east of M87 (73, 94, 113), the two sources in the north (33, 75) and the rest of the sources that do not seem especially clustered.

When we study the parameters for these galaxies hardly anything can be said to really discriminate the groups from each other, other than color. It seems that the two central groups, either by statistical fluke or by physical reasons, are on average a bit bluer than the other groups and also have a higher RMS in color values, i.e. have a broader color distribution.

We did a statistical test to see if we could translate this hunch into something quantitative, but a two-tailed Kolmogorov-Smirnov test (aiming for a significance level of 0.9) (Wall and Jenkins, 2003) resulted in rejecting the hypothesis that the distributions might be two separate ones. We can thus from now on consider the previous grouping physically irrelevant as there were no other big differences in parameters.

Now then let us see what we can say about the distribution of dwarfs in the survey as a whole. We started by examining parameters as function of (projected) distance to M87, one of the dominant galaxies in the cluster. The parameters are actually pretty uncorrelated with this as well. The same goes for distance to M86 and M49, the two other dominant clusters.

Our sample is thus a relatively smooth distribution (positionally) as far as photometric and morphological parameters are concerned. What we might say though is that the distribution seems to be somewhat centrally concentrated, i.e. around M87 and especially M86. Perhaps we should not put too much weight in this as the ACSVCS sampling inherently causes central abundancies (the majority of images is taken in the central region), but more on this in section 5.

4.3.2 Parameter-parameter relations

Thus having briefly covered distribution, we continue our parameter analysis with a comparison to relations from Graham and Guzmán (2003) and other parameter-parameter relations. The first result of these comparisons is in figure 11. In these we compare four parameters: absolute B-band magnitude M_B , mean B-band surface magnitude within the effective radius $\langle \mu \rangle_{e,B}$, central B-band surface magnitude $\mu_{0,B}$ and color $g - z$. Figure 11(d) also displays a relationship found in Graham and Guzmán (2003), which they saw as a unifying relationship for early type galaxies and dwarf ellipticals.

In figure 11(a) there is quite a lot of spread in the other data compared to our data. This is due largely to differences in surface brightness profiles. Galaxies with steep profiles have a higher amount of light within the center and thus have a lower effective radius, resulting in a lower (brighter) surface magnitude. This hypothesis is supported by extra data from SDSS. By using the ratio of `petroR50` to `petroR90` from SDSS as inverse concentration index C (Shimasaku et al., 2001; Strateva et al., 2001) you can indeed show a C gradient in the direction of the scattering (left to right in figure 11(a)).

As we mentioned in section 3.1, we should not simply assume the `MU_MAX` parameter to be the central surface brightness, especially now that we are looking at low surface brightness, patchy sources for which the patches could actually have higher surface brightness than the center and thus be used for the `MU_MAX` parameter. A first look at figures 11(b) and 11(d) hints at there being a problem with this parameter compared to the Binggeli and Jerjen (1998) data. We thus took a look at our source images and determined actual *central* surface brightnesses for each dwarf candidate by hand. The result of this is in figure 12.

4.3.3 Luminosity function

Finally we will compare the luminosity function of the dwarfs we found to our expected Schechter luminosity function. In figure 13 the magnitude distribution for our sample is compared to the expected one from section 1.2. Although there is some offset from expectation as seen in figure 13(a), when taken in larger magnitude bins as in figure 13(b) a remarkable similarity with the expected distribution can be displayed.

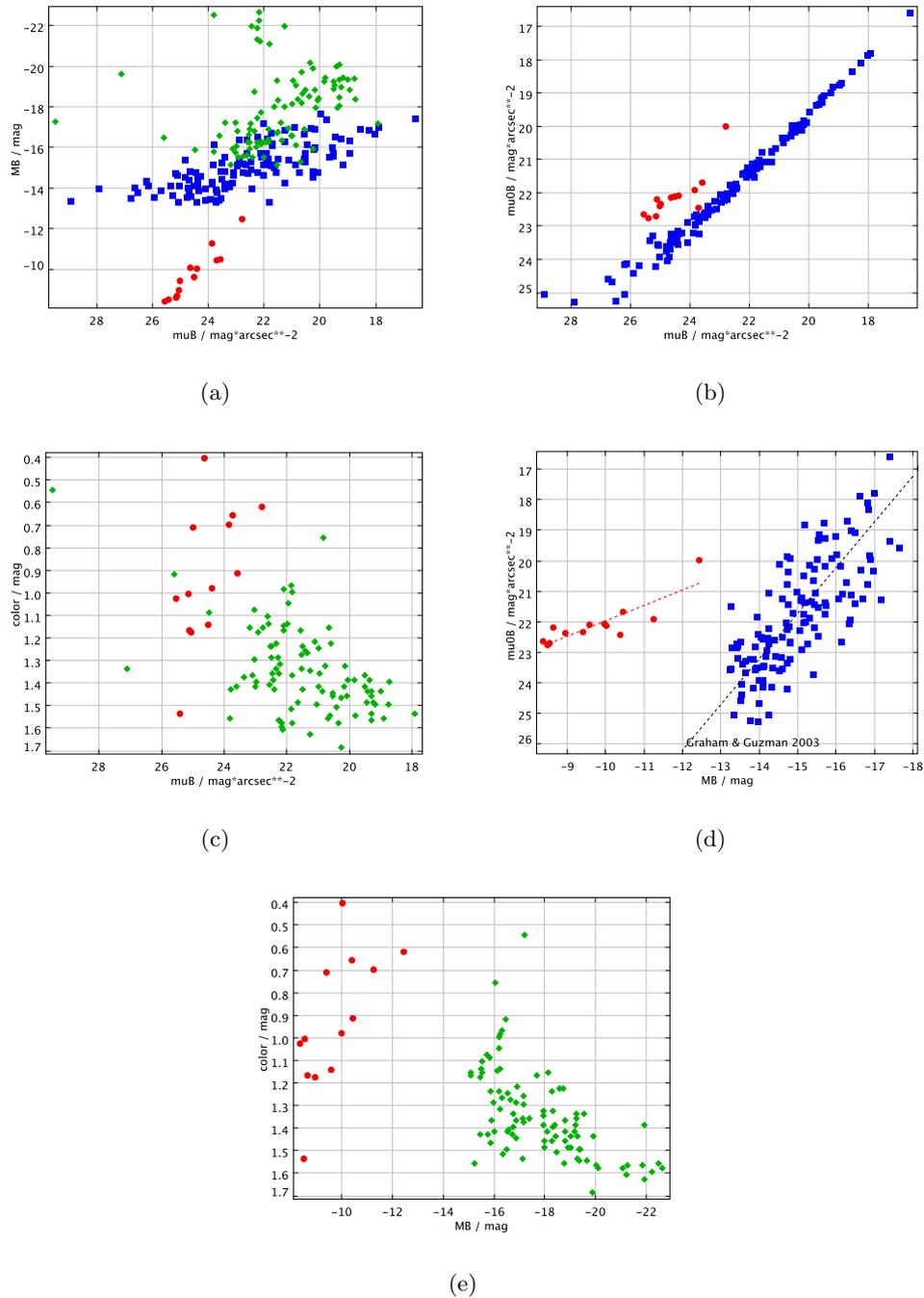


Figure 11: Parameter-parameter relations, adding our dwarf data and the data from the ACSVCS project to the data in Graham and Guzmán (2003) (which are in turn taken from Binggeli and Jerjen (1998)). Our data are in red, green data are from Ferrarese et al. (2006) and blue data are from Binggeli and Jerjen (1998).

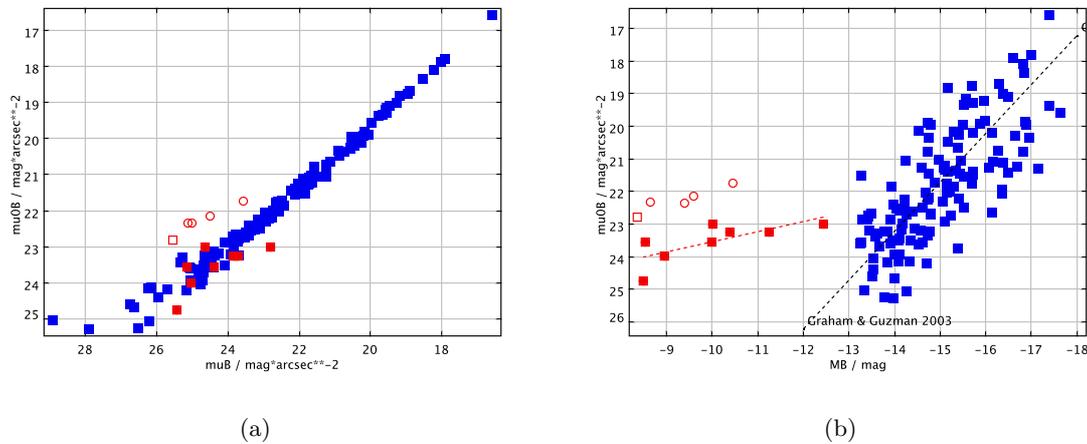


Figure 12: μ_0 relations, determined by hand. Our data are in red, blue data are from Binggeli and Jerjen (1998). The open circles are galaxies we classified as dSph,N in section 4.2.

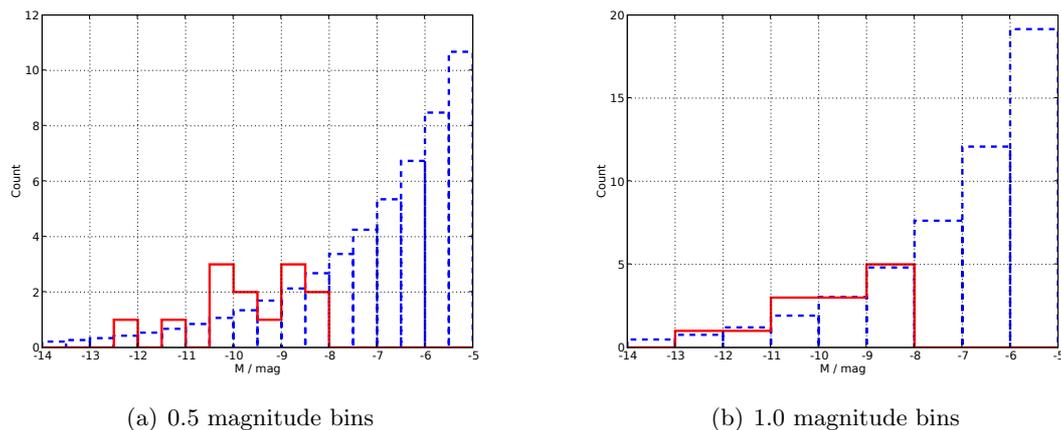


Figure 13: Luminosity function comparison. Our dwarf data are in red, the expected distribution is in blue dashed lines. In figure 13(b) a magnitude bin of 1.0 was used whereas a bin of 0.5 was used for figure 13(a) (as in figure 1).

5 Discussion

After analysis comes synthesis; what can we deduce from the data at hand? We ask ourselves four guiding questions in this discussion:

1. What are the color-magnitude relations? How do our galaxies compare to brighter ones? Do they differ or do they not and why?
2. What is the color dispersion? Color dispersion correlates with age dispersion; what does this mean for faint dwarf galaxy ages?
3. Are there star-forming dwarfs in our sample? These can be traced by being very blue.

4. Where are the dwarfs in our set? How are they distributed spacially? Do their parameters vary with position?

Let us start with the latter question, as that is already largely answered in section 4.3.1. The parameters do not vary much with position for as far as our sample is concerned, although this does not say much about dwarf parameter variability in general as our sample is far too small for such conclusions. There is the case of the somewhat larger color dispersion in the central galaxies, but as a statistical test shows these are no significant differences.

There is one thing that is interesting positionally and that is the fact that about half of our dwarfs are found in the dense central regions of the Virgo cluster. This central concentration of dwarfs might be explained away by either ACSVCS oversampling in these regions or the fact that the center simply holds a larger concentration of galaxies overall, but it would also corroborate the harassment model of galaxy formation, in which high concentrations are the cause of dwarf galaxy creation. This hypothesis would be strengthened by the fact that there is a large part of the Virgo cluster infalling into the rest of the cluster, gravitationally led by M86, and that this is happening quite close to the central region where we found most of our dwarfs. In fact, three of our sources lie extremely close to M86. M86's tangential direction is probably southward (Rangarajan et al., 1995). Proximity and perhaps shock waves might cause the effect we are seeing. Also, if it has been moving southward for a long time already it might have created the two northern dwarfs. All this is highly speculative though. Central clustering might also have been caused by M87, hosting the dominant mass concentration of the Virgo cluster, but showing any correlation between M87's position and that of our dwarfs is even less feasible than correlating positions to M86; they are simply too much spread out over the region. To resolve this issue we tried to find a simple statistical test to see if the dwarfs could simply have been drawn from the underlying distribution or if they are probably from different distributions. We were unable though to find one suitable for within the remaining timespan of this project. Some possible approaches would be 1) determining two-point correlation functions and comparing these to previously found dwarf populations (perhaps in other clusters as well), 2) determining luminosity functions per bin of area on the sky by taking only the sources in that bin; then we can compare this per bin to the number of found dwarfs and 3) making detailed (number) surface density maps of the region based on the VCC and simply take the values of the density at all dwarf positions. We can then see if indeed they predominantly reside in dense regions. These would thus prove useful starting points for further study, although we would probably need more than 13 dwarfs to really make good use of the second and third tests.

So we then take a look at the parameter-parameter relations, as distributional considerations seem a lost cause. Using the color-magnitude relations, i.e. figures 11(c) and 11(e), we shall first try to answer the first two questions of our guiding questions above. First we might notice that although there is no overlap between the different datasets in figure 11(c), the color values seem to disperse in perfect accordance with the trend set by the other data, both in upper and lower limits. This is a good indication that we are indeed looking at dwarf galaxies, as that is what the other data are as well (starting from $M_B = -18$ of course), and that they are Virgo cluster members (although the large scatter in color reduces the robustness of these statements). As we then look at what is actually there we establish that two trends are apparent: 1) fainter galaxies are on average bluer than brighter ones and 2) the scatter in color increases towards fainter galaxies. The first effect can be explained by positing that fainter galaxies tend to be more metal-poor (Bothun et al., 1984; Poggianti et al., 2001) and

are thus bluer (Ferguson and Binggeli, 1994; Jerjen et al., 2004). The second one is a bit trickier. As Smecker-Hane and McWilliam (1999) and Poggianti et al. (2001) suggest, the color scatter at the faint end might be caused by different star formation histories. Smecker-Hane and McWilliam (1999) point out that in Local Group dSphs there are no two galaxies that have exactly the same history. Rather, they display a range of combinations of differently aged stellar populations indicating that indeed their formation histories largely depend not on some possible intrinsic factor, but on environmental factors. The bluest dwarfs could thus consist of a relatively young population, the reddest ones of older populations, whereas the intermediate ones could be a combination of populations of different ages, indicating multiple epochs of star formation for the galaxy. On a sidenote: an interesting effect observed by Poggianti et al. (2001) is that there seems to be a bimodality in the metallicities of faint galaxies. This might hint at two different mechanisms of dwarf formation (squelching and harassment possibly). Unfortunately we cannot directly show this as we have no metallicities, but our scatter in color could also be caused partly by this difference in metallicity (as metallicity correlates to color) and thus our results do not contradict their results. To conclude this part of the discussion: our large color scatter (RMS of 0.3 mags and a total span of ~ 1 mag) seems consistent with a combination of these effects (metallicity (= redness) correlating with luminosity and dwarf galaxies consisting of differently aged populations (where age also determines color)). To actually determine ages we would need more data (preferably indicators of metallicity) to disentangle the age and metallicity effects.

Next up: central surface brightness relations. A look at figures 11(b) and 11(d) indicates that this parameter may not be defined correctly in one of the datasets. Indeed this becomes apparent when we redefine our μ_0 not as SE MU_MAX but rather by determining the central pixel value by hand from the images. This in fact makes perfect sense; for our patchy galaxies the maximum surface brightness is more often than not different from the central one. The result of this redefinition is found in figure 12. Here we see in figure 12(a) that the $\langle \mu \rangle_e$ versus μ_0 relation used for the Binggeli and Jerjen (1998) data (derived in appendix B) is indeed correct. The largest offsets from Binggeli and Jerjen (1998) are found for the galaxies we identified as being nucleated in section 4.2, which also makes sense as indeed their central surface brightnesses would be heightened due to the bright core. Applying this redefinition to figure 11(d) as well we obtain figure 12(b). Here we see that while the data now seem to produce a better overlap with the Binggeli and Jerjen (1998) data, we still see a bending towards higher central surface brightnesses for our dwarfs with respect to the Binggeli and Jerjen (1998) dwarfs. The relationship from Graham and Guzmán (2003, figure 9), $M_B = \frac{2}{3}\mu_0 - 29.5$ uniting ellipticals and dwarf ellipticals, thus does not seem to hold for our galaxies. This could be caused by the fact that the relationship from Graham and Guzmán (2003) is actually based on dE galaxies only, whereas our data contain no dEs at all, but rather dSphs and dIrrs. We might then be looking at a dichotomy between dEs and Es on the one hand and dSph and dIrr on the other. Whereas Graham and Guzmán (2003) are trying to close the gap between the former, we could possibly conclude on creating a gap between the former and the latter. Of course all of this could simply be an effect caused by the detection limits of our survey; we lack extremely low surface brightness sources. The same goes for the relationship in figure 11(a), which apart from the scatter shows that our distribution is offset towards the bright end (the right) of the plot relative to the general apparent correlation of the other data. So, considering that these parameters seem to be biased by detection limits (and are uncorrected for concentration index, see second paragraph of section 4.3.2) we will abstain from drawing any definite conclusions here, but the above discussion on μ_0 might be

an interesting starting point for further research.

Then back to our guiding questions: on to number three. Star formation is traced by the blue color of massive hot OB stars, which are short-lived and thus only show up in places where star formation has just taken place (or is taking place). Herefore you would expect patches of blue to trace star forming regions. In our galaxies the closest we get to a blue color are our g-band images, so for objects for which bright patches are visible in the g-band that are not there (or are significantly less pronounced) in the z-band we might say that star formation could be taking place. Using this criterion only source 11 could possibly be said to have star formation. For this source we see a few patches of blue in the south-western regions. Source 73 might also be said to have a little star formation as it is bluer all-over and has a small blue peak in the south-east, but these facts could also be interpreted as the galaxy being younger or more metal-poor and the peak could be some background galaxy, piercing through the optically thin dwarf. Herefore we only assume source 11 to have star formation. Noteworthy is the fact that this galaxy is quite close to M86 as well as M87 (it is about in between the both of them) and is therefore situated in a very active region of the cluster, as both these galaxies host large 'subclusters' that are heavily interacting (Binggeli, 1999). It might be that recent interactions with one of these subcluster members have indeed caused some of the possible star formation in source 11.

Finally we can give a short comment on the luminosity distribution of figure 13. When properly binned ('proper' because our dataset is way too small to justify using 0.5 mag bins) we see a remarkable similarity to the expected distribution. This however does not imply that the used faint-end slope value of $\alpha = 1.5$ is indeed the correct Virgo Cluster value, as we have not taken incompleteness of our dataset into account. As mentioned in section 3.2 the extremely red and blue dwarfs might have been filtered out. Because these galaxies are faint (otherwise they would have been detected in both bands and thus included in the dataset), if present, they would increase the number of faint end galaxies, which would lead to a lower α . We can then conclude that our findings corroborate currently widely supported values for Virgo of $\alpha > -1.5$.

In the end we must conclude that we cannot say much with absolute certainty. Unfortunately it seems our sample is severely magnitude limited, thus prohibiting us from properly extrapolating known relationships from e.g. Graham and Guzmán (2003) into low magnitude realms. Also, the focus on small faint galaxies using a specific set of SExtractor parameters might have been the cause of our limited results, for as far as the apparent lack of low surface brightness sources is concerned. Of course our aim was not to find extremely low surface brightness dwarfs (i.e. dwarf galaxies with surface magnitude far below the sky value), so it was to be expected that finding other weak dwarfs would be hard, as weak dwarfs inherently seem to possess low surface brightness (if we are to believe other studies like the ones we compared to).

Nonetheless we were able to explain the results found in figures 11 and 13 by using age, metallicity and luminosity function based arguments, so it is likely that indeed we did find some new Virgo cluster dwarf galaxies. This alone can be seen as an achievement given the number of background galaxies present in the unfiltered lists. This project was thus basically a test-case in astronomical data mining, which will likely turn out to be one of the greatest astronomical challenges of the coming years. As the amount of available data will keep on growing faster and faster we will need to sharpen our tools and minds to be able to find those needles in the haystack.

6 Conclusions

We summarize the scientific results from this project.

- We found 13 previously unstudied (i.e. not catalogued in NED; some of them were in SDSS though) objects that we suspect to be small, faint dwarf galaxy members of the Virgo Cluster;
 - Magnitude limits taken into account, we can say that the galaxies are in reasonable agreement with magnitude - surface brightness relations from the literature (perhaps hinting at a higher slope at lower magnitudes, but to properly conclude this we would really need more data from extremely low surface brightness galaxies). Note though that one of the selection criteria were that they conform to this relation, so this is not a surprising result.
 - Other relations such as color versus magnitude and surface brightness also agree with data from the literature, thus suggesting that indeed our galaxies are Virgo cluster dwarf galaxies.
 - Qualitatively classifying the 13 objects we were able to place them in a generally used scheme for dwarf galaxy subclasses.
- Due to the small size of our sample we cannot draw any definite conclusions regarding positional dependences of parameters.
- There seems to be a small overdensity of dwarfs in the region of the M87 and M86 subclusters, though whether this is caused by the distribution of the ACSVCS data or by physical reasons we cannot say. Further statistical study (for which a larger dwarf sample is desirable) may resolve this issue.
- We observe a large color scatter (RMS of 0.3 mags and a total span of ~ 1 mag), which seems consistent with a combination of two effects from the literature: 1) metallicity (= redness) correlating with luminosity and 2) dwarf galaxies consisting of differently aged populations (where age also determines color), implying that indeed dwarf galaxies are primarily formed by environmental influences.
- We have identified possible star formation in one of our dwarf galaxies (number 11), which is located in the middle of the M87 / M86 interaction region, suggesting interactional causes of the star formation.
- The luminosity distribution of our dwarfs suggests a faint-end luminosity function slope of $\alpha < -1.5$, which is in accordance with current literature.

We must keep in mind that all photometric data in this project is produced by SExtractor, which is a great tool for detection of sources, but still lacks in photometric precision. We have herefore drawn our conclusions with caution and basically only argue on a statistical basis. A next step, and a useful starting point for further research, would be to get more precise photometric data on our dwarfs by using a tool like GalFit. In our impatience we did indeed experiment with GalFit, but we ran into time-consuming issues of fine tuning for some bright galaxies and computing time for our large initial dataset and thus eventually abstained from using it for this project.¹⁸

¹⁸In Häußler et al. (2007) we found a particularly interesting study of optimally using the combination of SExtractor and GalFit on large datasets using a package called Galapagos. Another option is of course using

7 Acknowledgements

Some/all of the data presented in this paper were obtained from the Multimission Archive at the Space Telescope Science Institute (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-HST data is provided by the NASA Office of Space Science via grant NAG5-7584 and by other grants and contracts.

Graphs and histograms in this document were created with either TOPCAT or Matplotlib (a Python module). The images of sources are screenshots from ds9.

One of the things I have learned during this project is that science is best done with the help of others. I owe my thanks to the following people:

- My main supervisor, Reynier Peletier, for overall counseling and support during all of the Klein Onderzoek.
- Reynier Peletier and Edwin Valentijn for coming up with the idea of the Klein Onderzoek.
- Alister Graham for sending me the data used in Graham and Guzmán (2003).
- Gijs Verdoes Kleijn, John McFarland and Ewout Helmich in particular and the rest of the Astro-WISE team in general for helping me out with Astro-WISE related trouble.
- Gijs Verdoes Kleijn and Edwin Valentijn for inviting me to the Lorentz Center workshop on Astro-WISE, where I was given the opportunity to present my research.
- Eite Tiesinga for giving me the gigabytes I needed on the NAS-drive (Network Attached Storage; a box full of harddrives) he was testing.
- Hugo Buddelmeijer for helping out with Python related stuff (PyFITS, PyRAF etc.).
- Gert Sikkema for helping out with Source Extractor.
- Jarno Ruwen for lending me his copy of Binney and Merrifield (1998) from which I learned a great part of the theoretical background on (dwarf) galaxies.
- Femke van der Honing for her overall love and support during this never-ending project.
- My fellow students and others at the Kapteyn for the inspirational hours at the coffee table.

the Astro-WISE implementation of GalFit.

References

- J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, C. Allende Prieto, K. S. J. Anderson, S. F. Anderson, J. Annis, N. A. Bahcall, C. A. L. Bailer-Jones, I. K. Baldry, J. C. Barentine, B. A. Bassett, A. C. Becker, T. C. Beers, E. F. Bell, A. A. Berlind, M. Bernardi, M. R. Blanton, J. J. Bochanski, W. N. Boroski, J. Brinchmann, J. Brinkmann, R. J. Brunner, T. Budavári, S. Carliles, M. A. Carr, F. J. Castander, D. Cinabro, R. J. Cool, K. R. Covey, I. Csabai, C. E. Cunha, J. R. A. Davenport, B. Dilday, M. Doi, D. J. Eisenstein, M. L. Evans, X. Fan, D. P. Finkbeiner, S. D. Friedman, J. A. Frieman, M. Fukugita, B. T. Gänsicke, E. Gates, B. Gillespie, K. Glazebrook, J. Gray, E. K. Grebel, J. E. Gunn, V. K. Gurbani, P. B. Hall, P. Harding, M. Harvanek, S. L. Hawley, J. Hayes, T. M. Heckman, J. S. Hendry, R. B. Hindsley, C. M. Hirata, C. J. Hogan, D. W. Hogg, J. B. Hyde, S.-i. Ichikawa, v. Ivezić, S. Jester, J. A. Johnson, A. M. Jorgensen, M. Jurić, S. M. Kent, R. Kessler, S. J. Kleinman, G. R. Knapp, R. G. Kron, J. Krzesinski, N. Kuropatkin, D. Q. Lamb, H. Lampeitl, S. Lebedeva, Y. S. Lee, R. F. Leger, S. Lépine, M. Lima, H. Lin, D. C. Long, C. P. Loomis, J. Loveday, R. H. Lupton, O. Malanushenko, V. Malanushenko, R. Mandelbaum, B. Margon, J. P. Marriner, D. Martínez-Delgado, T. Matsubara, P. M. McGehee, T. A. McKay, A. Meiksin, H. L. Morrison, J. A. Munn, R. Nakajima, J. Neilsen, Eric H., H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nieto-Santisteban, A. Nitta, S. Okamura, R. Owen, H. Oyaizu, N. Padmanabhan, K. Pan, C. Park, J. Peoples, John, J. R. Pier, A. C. Pope, N. Purger, M. J. Raddick, P. Re Fiorentin, G. T. Richards, M. W. Richmond, A. G. Riess, H.-W. Rix, C. M. Rockosi, M. Sako, D. J. Schlegel, D. P. Schneider, M. R. Schreiber, A. D. Schwoppe, U. Seljak, B. Sesar, E. Sheldon, K. Shimasaku, T. Sivarani, J. A. Smith, S. A. Snedden, M. Steinmetz, M. A. Strauss, M. SubbaRao, Y. Suto, A. S. Szalay, I. Szapudi, P. Szkody, M. Tegmark, A. R. Thakar, C. A. Tremonti, D. L. Tucker, A. Uomoto, D. E. Vanden Berk, J. Vandenberg, S. Vidrih, M. S. Vogeley, W. Voges, N. P. Vogt, Y. Wadadekar, D. H. Weinberg, A. A. West, S. D. M. White, B. C. Wilhite, B. Yanny, D. R. Yocum, D. G. York, I. Zehavi, and D. B. Zucker. The sixth data release of the sloan digital sky survey. *The Astrophysical Journal Supplement Series*, 175(2):297–313, April 2008.
- N. Benítez, H. Ford, R. Bouwens, F. Menanteau, J. Blakeslee, C. Gronwall, G. Illingworth, G. Meurer, T. J. Broadhurst, M. Clampin, M. Franx, G. F. Hartig, D. Magee, M. Sirianni, D. R. Ardila, F. Bartko, R. A. Brown, C. J. Burrows, E. S. Cheng, N. J. G. Cross, P. D. Feldman, D. A. Golimowski, L. Infante, R. A. Kimble, J. E. Krist, M. P. Lesser, Z. Levay, A. R. Martel, G. K. Miley, M. Postman, P. Rosati, W. B. Sparks, H. D. Tran, Z. I. Tsvetanov, R. L. White, and W. Zheng. Faint galaxies in deep advanced camera for surveys observations. *The Astrophysical Journal Supplement Series*, 150(1):1–18, January 2004.
- E. Bertin and S. Arnouts. SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement*, 117:393–404, June 1996.
- B. Binggeli. The virgo cluster - home of m87. In H.-J. Röser and K. Meisenheimer, editors, *The radio galaxy Messier 87 : proceedings of a workshop held at Ringberg Castle, Tegernsee, Germany, 15-19 September 1997*, pages 9–... New York, Springer, 1999.
- B. Binggeli and H. Jerjen. Is the shape of the luminosity profile of dwarf elliptical galaxies an useful distance indicator? *Astronomy and Astrophysics*, 333:17–26, May 1998.

- B. Binggeli, A. Sandage, and G. A. Tammann. Studies of the virgo cluster. ii - a catalog of 2096 galaxies in the virgo cluster area. *Astronomical Journal*, 90:1681–1759, September 1985.
- J. Binney and M. Merrifield. *Galactic Astronomy*. Princeton Series in Astrophysics. Princeton University Press, 1998.
- M. R. Blanton, J. Dalcanton, D. Eisenstein, J. Loveday, M. A. Strauss, M. SubbaRao, D. H. Weinberg, J. Anderson, John E., J. Annis, N. A. Bahcall, M. Bernardi, J. Brinkmann, R. J. Brunner, S. Burles, L. Carey, F. J. Castander, A. J. Connolly, I. Csabai, M. Doi, D. Finkbeiner, S. Friedman, J. A. Frieman, M. Fukugita, J. E. Gunn, G. S. Hennessy, R. B. Hindsley, D. W. Hogg, T. Ichikawa, v. Ivezić, S. Kent, G. R. Knapp, D. Q. Lamb, R. F. Leger, D. C. Long, R. H. Lupton, T. A. McKay, A. Meiksin, A. Merelli, J. A. Munn, V. Narayanan, M. Newcomb, R. C. Nichol, S. Okamura, R. Owen, J. R. Pier, A. Pope, M. Postman, T. Quinn, C. M. Rockosi, D. J. Schlegel, D. P. Schneider, K. Shimasaku, W. A. Siegmund, S. Smee, Y. Snir, C. Stoughton, C. Stubbs, A. S. Szalay, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, B. Yanny, N. Yasuda, and D. G. York. The luminosity function of galaxies in sdss commissioning data. *The Astronomical Journal*, 121(5):2358–2380, May 2001.
- G. D. Bothun, W. Romanishin, S. E. Strom, and K. M. Strom. A possible relationship between metal abundance and luminosity for disk galaxies. *Astronomical Journal*, 89: 1300–1309, September 1984.
- N. Brosch. The dwarf galaxy population of the virgo cluster. *eprint arXiv:astro-ph/0206205*, June 2002. Invited review for the Ringberg Castle workshop on the Virgo Cluster.
- M. Capaccioli. Photometry of early-type galaxies and the $r \exp 1/4$ law. In H. G. Corwin and L. Bottinelli, editors, *The world of galaxies; Proceedings of the Conference, Paris, France, Apr. 12-14, 1988*, pages 208–227. New York, Springer-Verlag, 1989.
- P. Côté, J. P. Blakeslee, L. Ferrarese, A. Jordán, S. Mei, D. Merritt, M. Milosavljević, E. W. Peng, J. L. Tonry, and M. J. West. The acs virgo cluster survey. i. introduction to the survey. *The Astrophysical Journal Supplement Series*, 153(1):223–242, July 2004.
- N. Cross, S. P. Driver, W. Couch, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole, M. Colless, C. Collins, G. Dalton, K. Deeley, R. De Propris, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, S. Moody, P. Norberg, J. A. Peacock, B. A. Peterson, I. Price, M. Seaborne, W. Sutherland, H. Tadros, and K. Taylor. The 2df galaxy redshift survey: the number and luminosity density of galaxies. *Monthly Notices of the Royal Astronomical Society*, 324(4): 825–841, July 2001.
- I. Csabai, T. Budavári, A. J. Connolly, A. S. Szalay, Z. Györy, N. Benítez, J. Annis, J. Brinkmann, D. Eisenstein, M. Fukugita, J. Gunn, S. Kent, R. Lupton, R. C. Nichol, and C. Stoughton. The application of photometric redshifts to the sdss early data release. *The Astronomical Journal*, 125(2):580–592, Februari 2003.
- E. A. Evstigneeva, M. D. Gregg, M. J. Drinkwater, and M. Hilker. Internal properties of ultracompact dwarf galaxies in the virgo cluster. *The Astronomical Journal*, 133(4):1722–1740, April 2007.

- E. A. Evstigneeva, M. J. Drinkwater, C. Y. Peng, M. Hilker, R. De Propris, J. B. Jones, S. Phillipps, M. D. Gregg, and A. M. Karick. Structural properties of ultra-compact dwarf galaxies in the fornax and virgo clusters. eprint arXiv:0804.4353, April 2008.
- H. C. Ferguson and B. Binggeli. Dwarf elliptical galaxies. *Astronomy and Astrophysics Review*, 6(1-2):67–122, November 1994.
- L. Ferrarese, P. Côté, A. Jordán, E. W. Peng, J. P. Blakeslee, S. Piatek, S. Mei, D. Merritt, M. Milosavljević, J. L. Tonry, and M. J. West. The acs virgo cluster survey. vi. isophotal analysis and the structure of early-type galaxies. *The Astrophysical Journal Supplement Series*, 164(2):334–434, June 2006.
- A. W. Graham. A k-band central disc surface brightness correlation with scalelength for early-type disc galaxies, and the inclination correction. *Monthly Notices of the Royal Astronomical Society*, 326(2):543–552, September 2001.
- A. W. Graham and S. P. Driver. A concise reference to (projected) sérsic $r1/n$ quantities, including concentration, profile slopes, petrosian indices, and kron magnitudes. *Publications of the Astronomical Society of Australia*, 22(2):118–127, 2005.
- A. W. Graham and R. Guzmán. Hst photometry of dwarf elliptical galaxies in coma, and an explanation for the alleged structural dichotomy between dwarf and bright elliptical galaxies. *The Astronomical Journal*, 125(6):2936–2950, June 2003.
- B. Häussler, D. H. McIntosh, M. Barden, E. F. Bell, H.-W. Rix, A. Borch, S. V. W. Beckwith, J. A. R. Caldwell, C. Heymans, K. Jahnke, S. Jogee, S. E. Koposov, K. Meisenheimer, S. F. Sánchez, R. S. Somerville, L. Wisotzki, and C. Wolf. Gems: Galaxy fitting catalogs and testing parametric galaxy fitting codes: Galfit and gim2d. *The Astrophysical Journal Supplement Series*, 172(2):615–633, October 2007.
- B. W. Holwerda. Source extractor for dummies v5, Dec 2005.
- H. Jerjen, B. Binggeli, and K. C. Freeman. Surface br photometry of newly discovered dwarf elliptical galaxies in the nearby sculptor and centaurus a groups. *The Astronomical Journal*, 119(2):593–608, February 2000.
- H. Jerjen, B. Binggeli, and F. D. Barazza. Distances, metallicities, and ages of dwarf elliptical galaxies in the virgo cluster from surface brightness fluctuations. *The Astronomical Journal*, 127(2):771–788, February 2004.
- A. Jordán, J. P. Blakeslee, E. W. Peng, S. Mei, P. Côté, L. Ferrarese, J. L. Tonry, D. Merritt, M. Milosavljević, and M. J. West. The acs virgo cluster survey. ii. data reduction procedures. *The Astrophysical Journal Supplement Series*, 154(2):509–517, October 2004.
- R. G. Kron. Photometry of a complete sample of faint galaxies. *The Astrophysical Journal Supplement Series*, 43:305–325, June 1980.
- T. Lisker, E. K. Grebel, and B. Binggeli. Virgo cluster early-type dwarf galaxies with the sloan digital sky survey. iv. the color-magnitude relation. *Astronomical Journal*, 135(1):380–399, January 2008.

- B. Moore, G. Lake, T. Quinn, and J. Stadel. On the survival and destruction of spiral galaxies in clusters. *Monthly Notices of the Royal Astronomical Society*, 304(3):465–474, April 1999.
- M. Odenkirchen, E. K. Grebel, D. Harbeck, W. Dehnen, H.-W. Rix, H. J. Newberg, B. Yanny, J. Holtzman, J. Brinkmann, B. Chen, I. Csabai, J. J. E. Hayes, G. Hennessy, R. B. Hindsley, v. Ivezić, E. K. Kinney, S. J. Kleinman, D. Long, R. H. Lupton, E. H. Nielsen, A. Nitta, S. A. Snedden, and D. G. York. New insights on the draco dwarf spheroidal galaxy from the sloan digital sky survey: A larger radius and no tidal tails. *The Astronomical Journal*, 122(5):2538–2553, November 2001.
- C. Pavlovsky. *Advanced Camera for Surveys Instrument Handbook for Cycle 16*. Baltimore: STScI, 7.1 edition, December 2006.
- S. Phillipps, Q. A. Parker, J. M. Schwartzberg, and J. B. Jones. Dwarf spheroidal galaxies in the virgo cluster. *Astrophysical Journal Letters*, 493:59–62, February 1998.
- B. M. Poggianti, T. J. Bridges, B. Mobasher, D. Carter, M. Doi, M. Iye, N. Kashikawa, Y. Komiyama, S. Okamura, M. Sekiguchi, K. Shimasaku, M. Yagi, and N. Yasuda. A photometric and spectroscopic study of dwarf and giant galaxies in the coma cluster. iii. spectral ages and metallicities. *The Astrophysical Journal*, 562(2):689–712, December 2001.
- F. V. N. Rangarajan, D. A. White, H. Ebeling, and A. C. Fabian. The morphology and metal abundance of m86 from rosat pspc and hri observations: dust destruction in supersonic ram-pressure stripping. *Monthly Notices of the Royal Astronomical Society*, 277(3):1047–1062, December 1995.
- J. I. Read and G. Gilmore. Mass loss from dwarf spheroidal galaxies: the origins of shallow dark matter cores and exponential surface brightness profiles. *Monthly Notices of the Royal Astronomical Society*, 356(1):107–124, January 2005.
- S. Sabatini, J. Davies, R. Scaramella, R. Smith, M. Baes, S. M. Linder, S. Roberts, and V. Testa. The dwarf lsb galaxy population of the virgo cluster - i. the faint-end slope of the luminosity function. *Monthly Notice of the Royal Astronomical Society*, 341(3):981–992, May 2003a.
- S. Sabatini, S. Roberts, and J. Davies. Dwarf lsb galaxies and their environment: The virgo cluster, the ursa major cluster, isolated galaxies and voids. *Astrophysics and Space Science*, 285(1):97–106, 2003b.
- A. Sandage, B. Binggeli, and G. A. Tammann. Studies of the virgo cluster - part five - luminosity functions of virgo cluster galaxies. *Astronomical Journal*, 90:1759–1771, September 1985.
- P. Schechter. An analytic expression for the luminosity function for galaxies. *The Astrophysical Journal*, 203:297–306, January 1976.
- K. Shimasaku, M. Fukugita, M. Doi, M. Hamabe, T. Ichikawa, S. Okamura, M. Sekiguchi, N. Yasuda, J. Brinkmann, I. Csabai, S.-I. Ichikawa, Z. Ivezić, P. Z. Kunszt, D. P. Schneider, G. P. Szokoly, M. Watanabe, and D. G. York. Statistical properties of bright galaxies in the sloan digital sky survey photometric system. *The Astronomical Journal*, 122(3):1238–1250, September 2001.

- T. Smecker-Hane and A. McWilliam. Ages and chemical abundances in dwarf spheroidal galaxies. In I. Hubeny, S. Heap, and R. Cornett, editors, *Spectrophotometric Dating of Stars and Galaxies, ASP Conference Proceedings*, volume 192, pages 150–158. Astronomical Society of the Pacific, 1999.
- M. Stiavelli, B. W. Miller, H. C. Ferguson, J. Mack, B. C. Whitmore, and J. M. Lotz. The nuclear cusp slopes of dwarf elliptical galaxies. *The Astronomical Journal*, 121(3):1385–1394, March 2001.
- I. Strateva, v. Ivezić, G. R. Knapp, V. K. Narayanan, M. A. Strauss, J. E. Gunn, R. H. Lupton, D. Schlegel, N. A. Bahcall, J. Brinkmann, R. J. Brunner, T. Budavári, I. Csabai, F. J. Castander, M. Doi, M. Fukugita, Z. Györy, M. Hamabe, G. Hennessy, T. Ichikawa, P. Z. Kunszt, D. Q. Lamb, T. A. McKay, S. Okamura, J. Racusin, M. Sekiguchi, D. P. Schneider, K. Shimasaku, and D. York. Color separation of galaxy types in the sloan digital sky survey imaging data. *The Astronomical Journal*, 122(4):1861–1874, October 2001.
- J. L. Tonry, A. Dressler, J. P. Blakeslee, E. A. Ajhar, A. B. Fletcher, G. A. Luppino, M. R. Metzger, and C. B. Moore. The sbf survey of galaxy distances. iv. sbf magnitudes, colors, and distances. *The Astrophysical Journal*, 546(2):681–693, January 2001.
- R. B. Tully, R. S. Somerville, N. Trentham, and M. A. W. Verheijen. Squelched galaxies and dark halos. *The Astrophysical Journal*, 569(2):573–581, April 2002.
- S. van den Bergh. A catalogue of dwarf galaxies. *Publications of the David Dunlap Observatory*, 2(5):147–150, 1959.
- L. van Zee, E. J. Barton, and E. D. Skillman. Stellar populations of dwarf elliptical galaxies: Ubvri photometry of dwarf elliptical galaxies in the virgo cluster. *The Astronomical Journal*, 128(6):2797–2814, December 2004.
- C. J. Walcher, J. W. Fried, A. Burkert, and R. S. Klessen. About the morphology of dwarf spheroidal galaxies and their dark matter content. *Astronomy and Astrophysics*, 406:847–854, August 2003.
- J. Wall and C. Jenkins. *Practical Statistics for Astronomers*. Cambridge University Press, 2003.

A Research description (Dutch)

Titel: Dwergstelsels in de Virgo-ACS Survey

Begeleiders: E.A. Valentijn en R.F. Peletier

Het onderzoek: De Virgo-ACS is een survey van 100 vroeg-type stelsels in de Virgo cluster in 2 banden: Gunn g en I. Bepaalde aspecten van deze data zijn geanalyseerd door het Virgo-ACS team in de VS (zie <http://www1.cadc-ccda.hia-ihp.nrc-cnrc.gc.ca/community/ACSVCS/index.html>). In hun studie is de populatie van dwergen nogal onderbelicht gebleven. Om hier wat aan te doen is het volgende onderzoek bedacht: vind met AstroWise alle dwergstelsels in de plaatjes van de ACS Survey. Dit zijn niet alleen de hierboven genoemde 100 stelsels, maar ook veel andere, vooral zwakkere stelsels. Bepaal vervolgens de fysische parameters (helderheid, kleur, assenverhouding, schaallengte, morfologisch type enz.), en ga vervolgens deze fysische parameters bestuderen als functie van afstand tot het centrum van de Virgo cluster. Het doel is te bepalen hoe dwergstelsels evolueren, en in hoeverre hun omgeving daarvoor verantwoordelijk is.

Tijdsduur: 3 maanden (20 EC's)

Praktische uitvoering: Het onderzoek bestaat uit de volgende stappen:

1. Inlezen in de literatuur
2. De gegevens van de Virgo-ACS survey ophalen uit het ACS archief en inlezen in AstroWise
3. Sextractor draaien om alle melkwegstelsels te vinden
4. Scheiden van cluster- en achtergrondstelsels (+ enkele voorgrondstelsels) op basis van hun parameters
5. Galfit draaien photometrische parameters te bepalen en die te vertalen in fysische parameters
6. Parameters bestuderen als functie van afstand tot het centrum van de cluster
7. Bepalen wat dit voor consequenties heeft voor het vormen van dwergstelsels.

B Derivations

B.1 Average surface magnitude within the effective radius: $\langle \mu \rangle_e$

The surface brightness is independent of distance and is thus a great measurable quantity, giving direct access to a fundamental property of the object. In its non-logarithmic form it is defined as flux per surface area:

$$s = f/A \quad (21)$$

Being astronomers though, we take our portions in logarithms. The usual definition of an apparent magnitude as derived from a flux is:

$$m = -2.5 \log(f) + m_{\text{zeropoint}} \quad (22)$$

We then define the surface magnitude as simply:

$$\mu \equiv -2.5 \log(f/A) + m_{\text{zeropoint}} \quad (23)$$

Which can be rewritten as:

$$\mu = m_{\text{zeropoint}} - 2.5 \log(f) + 2.5 \log(A) = m + 2.5 \log(A) \quad (24)$$

For completeness' sake we now show that indeed this is a distance independent measure.

$$\mu = m + 2.5 \log(\text{area}) = m + 2.5 \log(A[\text{arcsec}^2]) \quad (25)$$

$$m - M = 5 \log(d/10\text{pc}) \quad (26)$$

$$\mu = M + 2.5 \log\left(\frac{d^2}{(10\text{pc})^2} A[\text{arcsec}^2]\right) = M + 2.5 \log\left(\frac{d^2}{(10\text{pc})^2} \pi(r[\text{arcsec}])^2\right) \quad (27)$$

$$\tan(r[\text{degrees}]) = r[\text{pc}]/d \sim r[\text{degrees}] = \frac{r[\text{arcsec}]}{3600} \quad (28)$$

$$\mu = M + 5 \log\left(\pi 3600 \frac{r[\text{pc}]}{10\text{pc}}\right) = M + 5 \log(\pi 360 r[\text{pc}]) \quad (29)$$

And this only depends on luminosity and physical size of the object, so there you have it.

Now, the surface area of a galaxy is a quite poorly defined quantity, so we will just have to choose one that best suits our needs. For this section, as the title implies, this will be the area within an effective radius, in our case the half light radius (i.e. the radius within which half of the galaxy's light is contained). In this SExtractor based project we will use the `FLUX_RADIUS` parameter for this (which is given in pixels), combined with the plate scale of $0.05''/\text{pixel}$, as we will want to calculate μ in units of $\text{mag}/\text{arcsec}^2$. So we define effective radius r_e :

$$r_e = 0.05 \text{ FLUX_RADIUS} \quad (30)$$

For our apparent magnitude we use SE parameter `MAG_BEST`. Because we are talking about the magnitude within the effective radius in which half the light is contained we will have to subtract half of the flux from this total magnitude.

$$m_{1/2} = -2.5 \log(f/2) + m_{\text{zeropoint}} = m + 2.5 \log(2) \approx \text{MAG_BEST} + 0.75 \quad (31)$$

We then finally arrive at the definition of $\langle\mu\rangle_e$, the average surface magnitude within the effective radius:

$$\langle\mu\rangle_e \equiv m_{1/2} + 2.5 \log(\pi r_e^2) = \text{MAG_BEST} + 2.5 \log(\pi \text{FLUX_RADIUS}^2 0.05^2) + 0.75 \quad (32)$$

At least this is what you would get if you were looking at your galaxy face-on. Most galaxies are not oriented that conveniently though, so we will need to make a slight adjustment. According to Graham (2001, paragraph 3) this correction is:

$$\mu_{\text{inc}} = \mu_{\text{face-on}} + 2.5C \log(b/a) \quad (33)$$

Here b/a is the ratio of the semiminor to semimajor axis of the disc. For an optically thick disc $C = 0$, and for an optically thin disc $C = 1$ (Graham, 2001). We will assume $C = 1$, because the small dwarfs we are looking for are indeed likely to be optically thin. In SE elliptical parameters are given by parameters A and B. The final equation then for an uncalibrated, but inclination-corrected magnitude $\langle\mu\rangle_e$ is:

$$\langle\mu\rangle_e = \text{MAG_BEST} + 2.5 \log(\pi \text{FLUX_RADIUS}^2 0.05^2) + 0.75 + 2.5 \log(A/B) \quad (34)$$

B.2 From μ_0 and Sérsic n to $\langle\mu\rangle_e$

The data from Binggeli and Jerjen (1998) only list central surface brightness and Sérsic profile parameter n . As we need $\langle\mu\rangle_e$ for our comparisons we need to convert these values. For this we need two relations for Sérsic profiles from Graham and Driver (2005):

$$\mu(R) = \mu_e + \frac{2.5b_n}{\ln(10)} \left((R/R_e)^{1/n} - 1 \right) \quad (35)$$

$$\langle\mu\rangle_e = \mu_e - 2.5 \log(f(n)) \text{ where } f(n) = \frac{ne^{b_n}}{b_n^{2n}} \Gamma(2n) \quad (36)$$

Filling in equation 35 for $R = 0$ we get:

$$\mu(0) \equiv \mu_0 = \mu_e - \frac{2.5b_n}{\ln(10)} \Rightarrow \mu_e = \mu_0 + \frac{2.5b_n}{\ln(10)} \quad (37)$$

Using this in equation 36 we obtain:

$$\langle\mu\rangle_e = \mu_0 + \frac{2.5b_n}{\ln(10)} - 2.5 \log \left(\frac{ne^{b_n}}{b_n^{2n}} \Gamma(2n) \right) \quad (38)$$

To ease calculation we use an approximation for b_n from Capaccioli (1989), which is pretty accurate within $0.5 < n < 8$, i.e. in our range as well:

$$b_n \approx 1.9992n - 0.3271 \quad (39)$$

Unfortunately this relationship can not be inverted to obtain a Sérsic n from μ_0 and $\langle\mu\rangle_e$, because the Gamma function is an integral in which information is lost.

B.3 Dependence of $\langle \mu \rangle_e - \mu_0$ to Sérsic n

What we can do though (continuing on the last remark in the previous paragraph) is numerically calculate the difference between $\langle \mu \rangle_e$ and μ_0 . Rewriting equation 38 we have:

$$\langle \mu \rangle_e - \mu_0 = \frac{2.5b_n}{\ln(10)} - 2.5 \log \left(\frac{ne^{b_n}}{b_n^{2n}} \Gamma(2n) \right) \quad (40)$$

The result of calculating this using Capaccioli (1989)'s relation for n over its accurate range ($0.5 < n < 8$) is shown in figure 14.

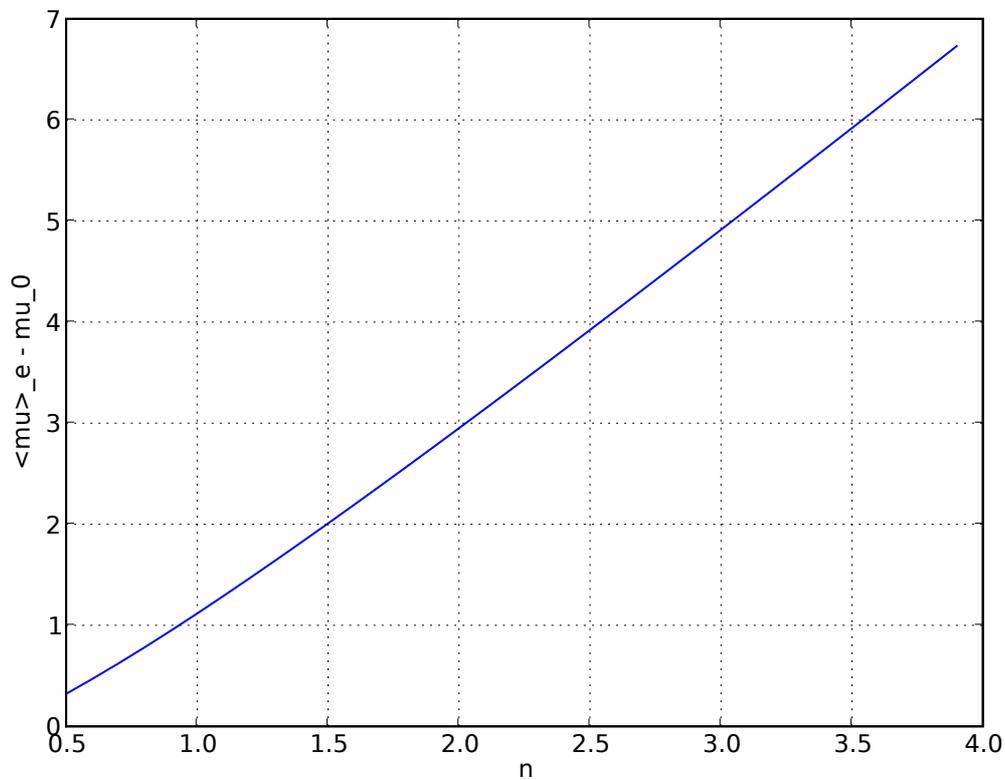


Figure 14: $\langle \mu \rangle_e - \mu_0$ versus Sérsic n , calculated numerically.

C Software setup

C.1 Astro-WISE

There are basically two ways to use Astro-WISE at the Kapteyn Institute: you can use the system-wide installed version or you can get your own version (in Astro-WISE language (or actually CVS language) an installation is usually referred to as a *checkout*; so we have the system-wide checkout and the personal checkout). The latter method has the advantage that you can edit Astro-WISE code and add your own. Here is how to setup these methods. In this I am assuming you have already obtained a username and password and know which Astro-WISE project(s) you will want to work on.

C.1.1 Kapteyn version

If you are a beginner to Astro-WISE you will probably just want to use the system-wide checkout. This has the advantage that you do not have to take care of updates yourself; it just works. To start using this system though a few things need to be set into place.

First you will need to tell your shell how to start the Astro-WISE environment (`awe`; a modified Python shell with which you can control the system). To do this in `csh` add the following lines to your `.cshrc`:

```
setenv AWEPIPE /Software/users/astro-wise/awe/current
alias awi "source ~aworacle/bin/coraenv"
```

Activate this new `.cshrc` by issuing the command `source ~/.cshrc` or by opening a new terminal. Then all that is left is to create a configuration file that will automatically log you in when you start the `awe`:

```
mkdir ~/.awe
cat << EOF > .awe/Environment.cfg
[global]
database_user : awpbos
database_password : xxxxxxxx
project : ACS@HST
EOF
```

For `project` you can use any project you want, probably the one you are most active in. You can always change this inside the `awe`. After this you are ready to fire the `awe` up with the command `awi && awe`. The `awi` command will initialize the environment so `awe` can run properly.

C.1.2 Custom checkout

More advanced users will probably want to install their own checkout. You will then be able to modify the system to your every needs (given that you are sufficiently proficient in the Python programming language). At the moment Astro-WISE is setup in such a way that in principle everything is possible. This also means that powerusers will have to take a certain amount of responsibility; you should not change your code in such a way that the results that will be stored in the database are totally inconsistent with what another user would expect.

In Astro-WISE others should be able use the products of your work, so make sure it's clear what they will use.

To get your own checkout you can use CVS:

```
mkdir $awe
cd $awe
cvs -d :pserver:anoncvs@cvs.astro-wise.org:/cvsroot login
password> astrowize
cvs -d :pserver:anoncvs@cvs.astro-wise.org:/cvsroot checkout -r AWBASE awe
```

This will download the latest 'stable' version of Astro-WISE to a directory named `$awe`. Configuration is then pretty much the same as for the system-wide checkout, with the exception that you should change the 'AWEPIPE' variable in `.cshrc` to the directory you just installed `awe` in.

A drawback from this checkout is that you'll have to update it yourself from time to time. To keep track of major updates you can subscribe to the mailinglist on the Astro-WISE website. To do the actual updates give the following commands:

```
cd $awe
cvs -q update -dPA
```

C.2 GalPhot in IRAF

The GalPhot package was originally written for use in IRAF. To use this version you can install it in your personal IRAF directory by the following procedure:

1. Download the package from <http://www.strw.leidenuniv.nl/~franx/galphot>.
2. Unpack to your IRAF directory (usually `~/iraf`):

```
mkdir ~/iraf/galphot
tar -xzf galphot.softlinks.tar.gz ~/iraf/galphot
```

3. Add the following lines to your IRAF `login.cl` file:

```
reset galphot      = [FULL PATH TO IRAF DIRECTORY]/galphot/
task galphot.pkg   = galphot$galphot.cl
```

After this you can load the GalPhot package in IRAF by entering `galphot`.

D Astro-WISE

D.1 Philosophy and personal contributions

In the past year I have worked extensively with parts of the image and analysis pipelines of Astro-WISE (co-adding, SExtractor, GalFit, GalPhot). Having thus gained a lot of user experience I felt obliged to share with the team my comments on and ideas for improvement of the system.

As an aside, I just want to make clear that I strongly support the philosophy behind the system. The analysis of extremely large datasets, for which Astro-WISE was originally built, will likely be a logistic nightmare without having a solid framework in which data can be stored and progress can be monitored by all participants of the research. My comments were therefore not intended to raise any questions as to the principal usefulness of a system like Astro-WISE. Rather they were focussed on what I still found missing during my research. Astro-WISE is built up with multiple-user robustness in mind; it tries to do the things it does as consistently as possible. For a single user though this sometimes throws up unneeded barriers when you need to stray from preset paths.

With all this in mind I have talked to several team members during and after the project and I must say that the talks were very constructive. Based on these talks the following list of features and fixes has now been implemented or put on the development agenda.

- *SourceList* functionality will be extended to include multiple outputs ("vector" output) and some other useful parameters will be added.
- VOTable output for *SourceLists* and *AssociateLists* will become an easily accessible method in the AWE prompt, just as it already is now in the online Database Viewer.
- Some *GalFit* related additions will be made to the AW Howtos and FAQ concerning fixes for bad fits and constraining parameters.
- A new image object (something like GalFitResidualFrame) will be added for GalFit/GalPhot residual images (thus eliminating the need for the trick in section 2.3.1).
- I reported some minor bugs in the *Image Server* and *Distributed Processing Unit* systems, which were subsequently fixed.

D.2 Python scripts

In the course of this project we have produced quite a few lines of Python code for interfacing with Astro-WISE and performing other large tasks. To include these would about double the page count of this report. If you are interested in using or seeing them please contact us at pbos@astro.rug.nl.

E Source images

In this appendix we show the images of 120 sources that were left over after background subtraction filtering in section 4.1. The images were made by first retrieving FITS cutouts from the Astro-WISE image server. These FITS cutouts were then converted to bitmaps using IRAF, `ds9` and ImageMagick. In IRAF we used the display task to display in `ds9` using the following commands for the "equally scaled" images:

```
disp g 1 zscale=no z1=0 z2=1 zrange=no ztrans="log"
disp z 2 zscale=no z1=0 z2=1 zrange=no ztrans="log"
```

For the "custom scaled" images we used different values for `z1`, ranging from -0.1 to 0.05.

From `ds9` we saved the resulting images and then used ImageMagick's `montage` routine to group the images and label them with their filenames (which we use as source numbers to refer to them). For the "custom scaled" images we also added a red border for clarity. The dwarf-images were given a green border. The command used for the first images:

```
montage -label %t -tile 3x8 -frame 5 -geometry +4+4 [input] out.png
```

For the "custom scaled" images:

```
montage -label %t -tile 3x7 -frame 5 -mattecolor red \
-geometry +4+4 [input] out.png
```

And for the dwarf images:

```
montage -label %t -tile 2x4 -frame 5 -mattecolor lightgreen \
-geometry +4+4 [input] out.png
```

The first of the following images display all the sources (grey frames), equally color-scaled (i.e. by the scale in figure 15). After that come the ones we put a custom scale on, i.e. differing for each source (the red frames). Finally we display the final dwarf selection (green frames).



Figure 15: Logarithmic color scale in `ds9` for source images. The values were renormalized to fit the fixed value range; 0 stands for a pixel value of 0 ADU/sec and 200 stands for a pixel value of 1 ADU/sec.

