

## The Second Data Release of the Sloan Digital Sky Survey

Kevork Abazajian<sup>1</sup>, Jennifer K. Adelman-McCarthy<sup>2</sup>, Marcel A. Agüeros<sup>3</sup>, Sahar S. Allam<sup>4</sup>, Kurt S. J. Anderson<sup>5,4</sup>, Scott F. Anderson<sup>3</sup>, James Annis<sup>2</sup>, Neta A. Bahcall<sup>6</sup>, Ivan K. Baldry<sup>7</sup>, Steven Bastian<sup>2</sup>, Andreas Berlind<sup>8,9,10</sup>, Mariangela Bernardi<sup>11</sup>, Michael R. Blanton<sup>10</sup>, John J. Bochanski Jr.<sup>3</sup>, William N. Boroski<sup>2</sup>, John W. Briggs<sup>8</sup>, J. Brinkmann<sup>5</sup>, Robert J. Brunner<sup>12</sup>, Tamás Budavári<sup>7</sup>, Larry N. Carey<sup>3</sup>, Samuel Carliles<sup>7</sup>, Francisco J. Castander<sup>13</sup>, A. J. Connolly<sup>14</sup>, István Csabai<sup>15,7</sup>, Mamoru Doi<sup>16</sup>, Feng Dong<sup>6</sup>, Daniel J. Eisenstein<sup>17</sup>, Michael L. Evans<sup>3</sup>, Xiaohui Fan<sup>17</sup>, Douglas P. Finkbeiner<sup>6</sup>, Scott D. Friedman<sup>18</sup>, Joshua A. Frieman<sup>2,8,9</sup>, Masataka Fukugita<sup>19</sup>, Roy R. Gal<sup>7</sup>, Bruce Gillespie<sup>5</sup>, Karl Glazebrook<sup>7</sup>, Jim Gray<sup>20</sup>, Eva K. Grebel<sup>21</sup>, James E. Gunn<sup>6</sup>, Vijay K. Gurbani<sup>2,22</sup>, Patrick B. Hall<sup>6</sup>, Masaru Hamabe<sup>23</sup>, Frederick H. Harris<sup>24</sup>, Hugh C. Harris<sup>24</sup>, Michael Harvanek<sup>5</sup>, Timothy M. Heckman<sup>7</sup>, John S. Hendry<sup>2</sup>, Gregory S. Hennessey<sup>25</sup>, Robert B. Hindsley<sup>56</sup>, Craig J. Hogan<sup>3</sup>, David W. Hogg<sup>10</sup>, Donald J. Holmgren<sup>2</sup>, Shin-ichi Ichikawa<sup>26</sup>, Takashi Ichikawa<sup>27</sup>, Željko Ivezić<sup>6</sup>, Sebastian Jester<sup>2</sup>, David E. Johnston<sup>8,9,6</sup>, Anders M. Jorgensen<sup>28</sup>, Stephen M. Kent<sup>2</sup>, S. J. Kleinman<sup>5</sup>, G. R. Knapp<sup>6</sup>, Alexei Yu. Kniazev<sup>21</sup>, Richard G. Kron<sup>8,2</sup>, Jurek Krzesinski<sup>5,29</sup>, Peter Z. Kunszt<sup>7,30</sup>, Nikolai Kuropatkin<sup>2</sup>, Donald Q. Lamb<sup>8,31</sup>, Hubert Lampeitl<sup>2</sup>, Brian C. Lee<sup>32</sup>, R. French Leger<sup>2</sup>, Nolan Li<sup>7</sup>, Huan Lin<sup>2</sup>, Yeong-Shang Loh<sup>6,33</sup>, Daniel C. Long<sup>5</sup>, Jon Loveday<sup>34</sup>, Robert H. Lupton<sup>6</sup>, Tanu Malik<sup>7</sup>, Bruce Margon<sup>18</sup>, Takahiko Matsubara<sup>35</sup>, Peregrine M. McGehee<sup>4,36</sup>, Timothy A. McKay<sup>37</sup>, Avery Meiksin<sup>38</sup>, Jeffrey A. Munn<sup>24</sup>, Reiko Nakajima<sup>39</sup>, Thomas Nash<sup>2</sup>, Eric H. Neilsen, Jr.<sup>2</sup>, Heidi Jo Newberg<sup>40</sup>, Peter R. Newman<sup>5</sup>, Robert C. Nichol<sup>11</sup>, Tom Nicinski<sup>2,41</sup>, Maria Nieto-Santisteban<sup>7</sup>, Atsuko Nitta<sup>5</sup>, Sadanori Okamura<sup>42</sup>, William O'Mullane<sup>7</sup>, Jeremiah P. Ostriker<sup>6</sup>, Russell Owen<sup>3</sup>, Nikhil Padmanabhan<sup>43</sup>, John Peoples<sup>2</sup>, Jeffrey R. Pier<sup>24</sup>, Adrian C. Pope<sup>7</sup>, Thomas R. Quinn<sup>3</sup>, Gordon T. Richards<sup>6</sup>, Michael W. Richmond<sup>44</sup>, Hans-Walter Rix<sup>21</sup>, Constance M. Rockosi<sup>3</sup>, David J. Schlegel<sup>6</sup>, Donald P. Schneider<sup>45</sup>, Ryan Scranton<sup>14</sup>, Maki Sekiguchi<sup>46</sup>, Uros Seljak<sup>6</sup>, Gary Sergey<sup>2</sup>, Branimir Sesar<sup>47</sup>, Erin Sheldon<sup>8,9</sup>, Kazu Shimasaku<sup>42</sup>, Walter A. Siegmund<sup>48</sup>, Nicole M. Silvestri<sup>3</sup>, J. Allyn Smith<sup>49,28</sup>, Vernesa Smolčić<sup>47</sup>, Stephanie A. Snedden<sup>5</sup>, Albert Stebbins<sup>2</sup>, Chris Stoughton<sup>2</sup>, Michael A. Strauss<sup>6</sup>, Mark SubbaRao<sup>8,50</sup>, Alexander S. Szalay<sup>7</sup>, István Szapudi<sup>48</sup>, Paula Szkody<sup>3</sup>, Gyula P. Szokoly<sup>51</sup>, Max Tegmark<sup>39</sup>, Luis Teodoro<sup>1</sup>, Aniruddha R. Thakar<sup>7</sup>, Christy Tremonti<sup>17</sup>, Douglas L. Tucker<sup>2</sup>, Alan Uomoto<sup>7,52</sup>, Daniel E. Vanden Berk<sup>14</sup>, Jan Vandenberg<sup>7</sup>, Michael S. Vogeley<sup>53</sup>, Wolfgang Voges<sup>51</sup>, Nicole P. Vogt<sup>4</sup>, Lucianne M. Walkowicz<sup>3</sup>, Shu-i Wang<sup>8</sup>, David H. Weinberg<sup>54</sup>, Andrew A. West<sup>3</sup>, Simon D.M. White<sup>55</sup>, Brian C. Wilhite<sup>8</sup>, Yongzhong Xu<sup>1</sup>, Brian Yanny<sup>2</sup>, Naoki Yasuda<sup>19</sup>, Ching-Wa Yip<sup>14</sup>, D. R. Yocum<sup>2</sup>, Donald G. York<sup>8,31</sup>, Idit Zehavi<sup>17</sup>, Stefano Zibetti<sup>55</sup>, Daniel B. Zucker<sup>21</sup>

- 
- <sup>1</sup>Theoretical Division, MS B285, Los Alamos National Laboratory, Los Alamos, NM 87545
- <sup>2</sup>Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510
- <sup>3</sup>Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195
- <sup>4</sup>New Mexico State University, Department of Astronomy, P.O. Box 30001, Dept 4500, Las Cruces, NM 88003
- <sup>5</sup>Apache Point Observatory, P.O. Box 59, Sunspot, NM 88349
- <sup>6</sup>Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544
- <sup>7</sup>Center for Astrophysical Sciences, Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218
- <sup>8</sup>Department of Astronomy and Astrophysics, The University of Chicago, 5640 S. Ellis Ave., Chicago, IL 60637
- <sup>9</sup>Center for Cosmological Physics, The University of Chicago, 5640 South Ellis Ave., Chicago, IL 60637
- <sup>10</sup>Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, New York, NY 10003
- <sup>11</sup>Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213
- <sup>12</sup>Department of Astronomy University of Illinois 1002 W. Green Street, Urbana, IL 61801
- <sup>13</sup>Institut d'Estudis Espacials de Catalunya/CSIC, Gran Capita 2-4, 08034 Barcelona, Spain
- <sup>14</sup>Department of Physics and Astronomy, University of Pittsburgh, 3941 O'Hara St., Pittsburgh, PA 15260
- <sup>15</sup>Department of Physics, Eötvös University, Budapest, Pf. 32, Hungary, H-1518
- <sup>16</sup>Institute of Astronomy, School of Science, University of Tokyo, 2-21-1 Osawa, Mitaka, Tokyo 181-0015, Japan
- <sup>17</sup>Steward Observatory, 933 N. Cherry Ave, Tucson, AZ 85721
- <sup>18</sup>Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218
- <sup>19</sup>Institute for Cosmic Ray Research, University of Tokyo, 5-1-5 Kashiwa-no-Ha, Kashiwa City, Chiba 277-8582, Japan
- <sup>20</sup>Microsoft Research, 455 Market Street, Suite 1690, San Francisco, CA 94105
- <sup>21</sup>Max-Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany
- <sup>22</sup>Lucent Technologies, 2000 Lucent Lane, Naperville, IL 60566
- <sup>23</sup>Dept. Mathematical & Physical Sciences, Japan Women's University 2-8-1, Mejirodai, Bunkyo-ku, Tokyo 112-8681, Japan
- <sup>24</sup>U.S. Naval Observatory, Flagstaff Station, P.O. Box 1149, Flagstaff, AZ 86002-1149
- <sup>25</sup>US Naval Observatory, 3540 Mass Ave NW, Washington, DC 20392
- <sup>26</sup>National Astronomical Observatory, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
- <sup>27</sup>Astronomical Institute, Tohoku University, Aramaki, Aoba, Sendai 980-8578, Japan
- <sup>28</sup>ISR-4, MS D448, Los Alamos National Laboratory, Los Alamos, NM 87545
- <sup>29</sup>Mt. Suhora Observatory, Cracow Pedagogical University, ul. Podchorazych 2, 30-084 Cracow, Poland
- <sup>30</sup>CERN, IT Division, 1211 Geneva 23, Switzerland

## ABSTRACT

The Sloan Digital Sky Survey has validated and made publicly available its Second Data Release. This data release consists of 3324 square degrees of five-band (*ugriz*) imaging data with photometry for over 88 million unique objects, 367,360 spectra of galaxies, quasars, stars and calibrating blank sky patches selected over 2627 degrees of

- 
- <sup>31</sup>Enrico Fermi Institute, The University of Chicago, 5640 S. Ellis Ave., Chicago, IL 60637
- <sup>32</sup>Lawrence Berkeley National Laboratory, One Cyclotron Rd., Berkeley CA 94720-8160
- <sup>33</sup>Center for Astrophysics and Space Astronomy, University of Colorado, Boulder, CO 80309
- <sup>34</sup>Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QJ, United Kingdom
- <sup>35</sup>Department of Physics and Astrophysics Nagoya University Chikusa, Nagoya 464-8602, Japan
- <sup>36</sup>SNS-4, MS H820, Los Alamos National Laboratory, Los Alamos, NM 87545
- <sup>37</sup>Department of Physics, University of Michigan, 500 East University Ave., Ann Arbor, MI 48109
- <sup>38</sup>Institute for Astronomy Royal Observatory Blackford Hill Edinburgh EH9 3HJ Scotland
- <sup>39</sup>Department of Physics, University of Pennsylvania, Philadelphia, PA 19104
- <sup>40</sup>Department of Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, NY 12180
- <sup>41</sup>CMC Electronics Aurora 43W752 Route 30 Sugar Grove, IL 60554
- <sup>42</sup>Department of Astronomy and Research Center for the Early Universe, School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan
- <sup>43</sup>Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544
- <sup>44</sup>Physics Department, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, NY 14623-5603
- <sup>45</sup>Department of Astronomy and Astrophysics, the Pennsylvania State University, University Park, PA 16802
- <sup>46</sup>Japan Participation Group, c/o Institute for Cosmic Ray Research, University of Tokyo, 5-1-5 Kashiwa-no-Ha, Kashiwa City, Chiba 277-8582, Japan
- <sup>47</sup>University of Zagreb, Department of Physics, Bijenička cesta 32, 10000 Zagreb, Croatia
- <sup>48</sup>Institute for Astronomy, 2680 Woodlawn Road, Honolulu, HI 96822
- <sup>49</sup>University of Wyoming, Dept. of Physics & Astronomy, Laramie, WY 82071
- <sup>50</sup>Adler Planetarium and Astronomy Museum 1300 Lake Shore Drive Chicago, IL 60605
- <sup>51</sup>Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse 1, D-85741 Garching, Germany
- <sup>52</sup>Carnegie Observatories 813 Santa Barbara St. Pasadena, CA 91101
- <sup>53</sup>Department of Physics, Drexel University, Philadelphia, PA 19104
- <sup>54</sup>Department of Astronomy, Ohio State University, Columbus, OH 43210
- <sup>55</sup>Max Planck Institute for Astrophysics, Karl Schwarzschildstrasse 1, D-85748 Garching, Germany
- <sup>56</sup>Code 7215, Remote Sensing Division Naval Research Laboratory 4555 Overlook Ave SW Washington, DC 20392

this area, and tables of measured parameters from these data. The imaging data reach a depth of  $r \approx 22.2$  (95% completeness limit for point sources) and are photometrically and astrometrically calibrated to 2% rms and 100 milli-arcsec rms per coordinate, respectively. The imaging data have all been processed through a new version of the SDSS imaging pipeline, in which the most important improvement since the last data release is fixing an error in the model fits to each object. The result is that model magnitudes are now a good proxy for point spread function (PSF) magnitudes for point sources, and Petrosian magnitudes for extended sources. The spectroscopy extends from 3800Å to 9200Å at a resolution of 2000. The spectroscopic software now repairs a systematic error in the radial velocities of certain types of stars, and has substantially improved spectrophotometry. All data included in the SDSS Early Data Release and First Data Release are reprocessed with the improved pipelines, and included in the Second Data Release. Further characteristics of the data are described, as are the data products themselves and the tools for accessing them.

*Subject headings:* Atlases—Catalogs—Surveys

## 1. The Sloan Digital Sky Survey

The Sloan Digital Sky Survey (SDSS; York et al. 2000) is an imaging and spectroscopic survey of the high Galactic latitude sky visible from the Northern hemisphere. The principal survey goals are to measure the large-scale distribution of galaxies and quasars and to produce an imaging and spectroscopic legacy for the astronomical community. The SDSS data have been used in well over 200 refereed papers to date on subjects ranging from the colors of asteroids (Ivezić et al. 2002) to magnetic white dwarfs (Schmidt et al. 2003) to structures in the Galactic halo (Newberg et al. 2003) to the star-formation history of galaxies (Kauffmann et al. 2003) to Type II quasars (Zakamska et al. 2003) to the large-scale distribution of galaxies (Pope et al. 2004; Tegmark et al. 2004). The survey uses a dedicated 2.5m telescope with a three-degree field of view at Apache Point Observatory, New Mexico. A 120 mega-pixel camera (Gunn et al. 1998) images in five broad bands ( $u, g, r, i$  and  $z$ ; Fukugita et al. 1996; Stoughton et al. 2002) on clear moonless nights of good seeing. These data are photometrically calibrated using an auxiliary 20-inch telescope with a  $40' \times 40'$  imager, which determines the photometricity of each night (Hogg et al. 2001), and measures the extinction and photometric zeropoint using a network of standard stars (Smith et al. 2002).

The imaging data are processed through a series of pipelines that locate and measure the properties of all detected objects (Lupton et al. 2001) and carry out photometric and astrometric calibration (Pier et al. 2003). From the resulting catalogs of objects, complete catalogs of galaxies (Eisenstein et al. 2001; Strauss et al. 2002) and quasar candidates (Richards et al. 2002) are selected for spectroscopic followup, and are assigned to spectroscopic tiles of diameter 3 degrees (Blanton et al. 2003). Spectroscopy is performed on nights with moonlight, mild cloud cover, and/or poor

seeing using a pair of double spectrographs with coverage from 3800–9200Å, and resolution  $\lambda/\Delta\lambda$  of roughly 2000. A plug plate for each tile holds 640 optical fibers of 3'' entrance aperture which feed the spectrographs, together with eleven coherent fiber bundles to image guide stars. Because of the diameter of the cladding holding the optical fibers, spectroscopy cannot be carried out for objects separated by less than 55'' on a given plate.

## 2. The Second Data Release

A high-level overview of the SDSS may be found in York et al. (2000), while many of the details of the software and data products may be found in Stoughton et al. (2002; hereafter the EDR paper). The latter paper also describes our Early Data Release (EDR), which consisted mostly of data taken during our commissioning period. The First Data Release was made public in April 2003; Abazajian et al. (2003; hereafter the DR1 paper) describe these data and give further details and updates on the quality of the data and its processing. The current paper describes the Second Data Release of the SDSS (DR2), which was made available to the public on 15 March 2004. The properties of DR2 are summarized in Table 1. The DR2 footprint is defined by all non-repeating survey-quality imaging runs within the a priori defined elliptical survey area in the Northern Galactic Cap, and three stripes in the Southern Galactic Cap (York et al. 2000) obtained prior to 1 July 2002, and the spectroscopy associated with that area obtained before that date. In fact, 34 square degrees of DR2 imaging data in the Northern Galactic Cap lie outside this ellipse. While the DR2 scans do not repeat a given area of sky, they do overlap to some extent, and the data in the overlaps are included in DR2 as well. The DR2 includes reprocessing of all data included in DR1, and those data in EDR that pass our data-quality criteria for the official survey.

The sky coverage of the imaging and spectroscopic data that make up DR2 is given in Figure 1. The effective areas of the two are 3324 and 2627 square degrees, respectively. The natural unit of imaging data is a *run*; the DR2 contains data from 105 runs in the **best** database, and 105 runs in the **target** database (**best** and **target** are defined in § 3). Similarly, the natural unit of spectroscopic data is a *plate* of 640 spectra each (of which 32 are devoted to background sky determination); the DR2 contains data from 574 plates.

The DR2 data are available via links off the web site <http://www.sdss.org/dr2>, which also includes extensive technical information about the SDSS data, and should be checked for errata and caveats to the data. The types of data that are available are described in detail in the DR1 and EDR papers, and on the above web site. There are two principal ways to access the data. The first, the Data Archive Server (DAS), allows one to download the FITS files containing the imaging and spectroscopic catalogs, the images themselves, and the spectra. This is the appropriate place to go to download large quantities of data in bulk. The second option, the Catalog Archive Server (CAS), allows one to perform database queries by object attributes and to obtain finding charts of given regions of sky and plots of the spectra. The CAS also provides pointers to the survey images and spectra in FITS format.

The EDR paper (Stoughton et al. 2002) describes the SDSS data in detail. We do not repeat that description here, but put emphasis on changes since DR1 and new-found problems in imaging (§ 3), spectroscopy (§ 4), and target selection (§ 5). An Appendix describes the conversion between magnitudes, fluxes, and counts in the imaging data.

### 3. The SDSS Imaging Data

The SDSS imaging pipelines have evolved as the survey has progressed, leading to continual improvements in the measured quantities. However, this means that much of the spectroscopic targetting is based on reductions carried out with old versions of the imaging pipelines. For this reason, we release two versions of the imaging data for each area of the sky. The **target** version is that which was used for defining spectroscopic targets of a given region, while the **best** version uses the latest version of the imaging pipelines. In some cases, improved data (e.g., with better seeing) were taken in a given region of sky after targetting was done, in which case the **target** and **best** data are independent runs. A total of 183 square degrees of sky are different runs between **target** and **best**, the majority along the Equatorial Stripe in the Fall sky.

The quality of the imaging data is described in the EDR and DR1 papers; the distribution of seeing, sky brightness, and depth for the DR2 data are no different than for DR1, nor are the accuracies of the astrometric and photometric calibrations. The **best** reductions do, however, have some substantial improvements over what was included in DR1. We now describe these improvements, and give further caveats of problems that have since come to light.

#### 3.1. Model Magnitudes

The computation of model magnitudes in the DR1 and EDR processing had a serious bug. The image of each object detected in any of the five bands in the imaging data is fit to de Vaucouleurs ( $I(\theta) \propto \exp[-(\theta/\theta_0)^{1/4}]$ ) and exponential ( $I(\theta) \propto \exp[-\theta/\theta_0]$ ) radial profiles of arbitrary axis ratio and inclination, convolved with the local Point Spread Function (PSF). However, these fits used an incorrect model of the PSF, which caused systematic errors in the fit parameters, especially for objects of small scale size  $\theta_0$  (i.e., close to a PSF). For each of these model fits, the code determined an aperture correction to force the exponential and de Vaucouleurs magnitudes to equal the PSF magnitude for stars; this correction was then applied to *all* objects. Because of this software error, this aperture correction was large, 0.2 magnitudes. Thus the model magnitudes of galaxies were systematically too large by typically 0.2 magnitudes. In the mean, exponential scale lengths **rexp** were overestimated by  $\sim 0.1'' - 0.2''$  for objects with **rexp**  $< 6''$ , while de Vaucouleurs scale lengths **rdev** were overestimated by a factor 1.25 for large objects, and as much as a factor of 2 for objects with **rdev**  $< 0.5''$ .

This error has been fixed in the latest version of the pipeline, and has been extensively tested

Table 1. Characteristics of the SDSS Second Data Release (DR2)

<b>Imaging</b>	
Footprint area	3324 deg <sup>2</sup>
Imaging catalog	88 million unique objects
Magnitude limits: <sup>a</sup>	
<i>u</i>	22.0
<i>g</i>	22.2
<i>r</i>	22.2
<i>i</i>	21.3
<i>z</i>	20.5
Median PSF width	1.4'' in <i>r</i>
RMS photometric calibration errors:	
<i>r</i>	2%
<i>u – g</i>	3%
<i>g – r</i>	2%
<i>r – i</i>	2%
<i>i – z</i>	3%
Astrometry	< 0.1'' rms absolute per coordinate
<b>Spectroscopy</b>	
Footprint area	2627 deg <sup>2</sup>
Wavelength Coverage	3800–9200Å
Resolution	1800–2100
Signal-to-noise ratio	> 4 per pixel at <i>g</i> = 20.2
Wavelength calibration	< 5 km sec <sup>-1</sup>
Redshift accuracy	30 km sec <sup>-1</sup> rms for Main galaxies
Number of spectra	367,360

<sup>a</sup>95% completeness for point sources in typical seeing; 50% completeness numbers are typically 0.4 mag fainter (DR1 paper).

with simulations. The code now also takes the best fit exponential and de Vaucouleurs fits in each band and asks for the linear combination of the two that best fits the image. The coefficient (clipped between zero and one) of the de Vaucouleurs term is stored in the quantity `fracDeV`<sup>57</sup>. This allows us to define a *composite flux*:

$$F_{composite} = \text{fracDeV} \times F_{deV} + (1 - \text{fracDeV}) \times F_{exp}, \quad (1)$$

where  $F_{deV}$  and  $F_{exp}$  are the de Vaucouleurs and exponential fluxes (*not* magnitudes) of the object in question. The magnitude derived from  $F_{composite}$  is referred to below as the `cmode1` magnitude (as distinct from the `model` magnitude, based on the better-fitting of the exponential and de Vaucouleurs models in the  $r$  band; see the EDR paper).

With these changes in place, there is now excellent agreement between `cmode1` and Petrosian magnitudes of galaxies, and `cmode1` and PSF magnitudes of stars (Figure 2). `cmode1` and Petrosian magnitudes are not expected to be identical, of course; as Strauss et al. (2002) describe, the Petrosian aperture excludes the outer parts of galaxy profiles, especially for elliptical galaxies. As a consequence, there is an offset of 0.05–0.1 mag between `cmode1` and Petrosian magnitudes of bright galaxies, depending on the photometric bandpass and the type of galaxy. The rms scatter between model and Petrosian magnitudes at the bright end is now between 0.05 and 0.08 magnitudes, depending on bandpass; the scatter between `cmode1` and Petrosian magnitudes for all galaxies is smaller, 0.03 to 0.05 magnitudes. For comparison, the code that was used in the EDR and DR1 had scatters of 0.1 mag and greater, with much more significant offsets.

The `cmode1` and PSF magnitudes of stars are forced to be identical in the mean by aperture corrections; this was true in older versions of the pipeline. The rms scatter between model and PSF magnitudes for stars is much reduced, from 0.03 mag to 0.02 mag, the exact values depending on bandpass. In the EDR and DR1, star-galaxy separation was based on the difference between model and PSF magnitudes (cf. the discussions in the EDR paper, Scranton et al. 2002, and Strauss et al. 2002). We now do star-galaxy separation to set the `type` parameter in the pipeline outputs using the difference between `cmode1` and PSF magnitudes, with the threshold at the same value (0.145 magnitudes).

Given the excellent agreement between `cmode1` magnitudes and PSF magnitudes for point sources, and between `cmode1` magnitudes and Petrosian magnitudes (albeit with intrinsic offsets due to aperture corrections) for galaxies, the `cmode1` magnitude is now an adequate proxy to use as a universal magnitude for all types of objects. As it is approximately a matched aperture to a galaxy, it has the great advantage over Petrosian magnitudes, in particular, of having close to optimal noise properties.

For measuring *colors* of extended objects, however, we continue to recommend using the model (not the `cmode1`) magnitudes; the colors of galaxies were almost completely unaffected by the DR1

---

<sup>57</sup>Due to an accident of the history this parameter is misleadingly termed `fracPSF` in the flat files of the DAS.



software error (cf., the discussion in § 5). The model magnitude is calculated using the best-fit parameters in the  $r$  band, and applies it to all other bands; the light is therefore measured consistently through the same aperture in all bands.

### 3.2. Other Substantive Changes to the Imaging Pipelines

- The behavior of the deblender of overlapping images has been further improved since the DR1; these changes are most important for bright galaxies of large angular extent ( $\theta > 1'$ ). In the EDR, and to a lesser extent in the DR1, bright galaxies were occasionally “shredded” by the deblender, i.e., interpreted as two or more objects and taken apart. With improvements in the code that finds the center of large galaxies in the presence of superposed stars, and the deblending of stars superposed on galaxies, this shredding now rarely happens. Indeed, inspections of several hundred NGC galaxies shows that the deblend is correct in 95% of the cases; most of the exceptions are irregular galaxies of various sorts.
- The PSF is measured from atlas images roughly  $7''$  across for stars; any error in the sky level determined from these images couples to spatial variability of the PSF by the Karhunen-Loève expansion used to model the PSF. This manifested itself in systematic offsets between the PSF and model magnitudes of stars of several hundredths of a magnitude, even with the fixes in the model magnitude code described above. This zero-point term in the PSF is now explicitly suppressed.
- The pixel size is  $0.396''$ , giving well-sampled images for the typical seeing of  $1''$  or more. On rare occasions when the seeing became much better than  $0.9''$  (FWHM), the undersampling causes the code that found stars suitable for determining the PSF to miss many objects, yielding an incorrect PSF and therefore poor stellar photometry (the seeing was never good enough in the runs included in DR1, so this error was not triggered). Changes to the thresholds for the selection of PSF stars have solved this problem.
- Astrometry for each object is referred to the reference frame of the  $r$ -band images. However, for objects of extreme color which are undetected in the  $r$  band (for example, cool brown dwarfs and  $z > 5.7$  quasars), DR1 had an error in the astrometric transformation from the detection band to the  $r$ -band, resulting in positional errors of several arcseconds. This problem is fixed in DR2, and the positions of objects not detected in the  $r$ -band are now correct.
- The EDR and DR1 match each SDSS object to the nearest object in USNO-A2.0 (Monet et al. 1998), using a 30 arcsec matching radius. USNO-A2.0 provides positions at a single epoch (no proper motions are provided), based on POSS-I plates. Proper motions are then calculated based on the SDSS and POSS-I positions, with a typical time baseline of 50 years. For motions greater than  $\sim 40$  mas year<sup>-1</sup>, corresponding to separations between the SDSS and USNO-A2.0 positions of greater than 2 arcsec, contamination by false

matches becomes significant and rises with increasing motion/separation (see the DR1 web site, <http://www.sdss.org/dr1>, for a fuller discussion). The DR2 reductions use USNO-B1.0 (Monet et al. 2003), which provides positions and proper motions based on various Schmidt photographic surveys (primarily POSS-I and POSS-II in the area of sky covered by SDSS). Each SDSS object is matched to the nearest USNO-B1.0 object within 1 arcsec, after first converting the USNO-B1.0 positions to the epoch of the SDSS observations. This eliminates nearly all of the false matches, yielding much cleaner samples of high proper motion stars. The USNO-B1.0 proper motion is then given for each matching SDSS object. More sophisticated techniques, using SDSS astrometry to recalibrate the USNO-B1.0 astrometry and then recalculate the proper motions based on both SDSS and USNO-B1.0 positions, are discussed by Munn et al. (2004) and Gould & Kollmeier (2004).

- When an image is saturated in the SDSS imaging data, the CCD wells overflow and a bleed trail results. However, the total number of electrons associated with the object, bleed trail and all, still at least approximately reflects the brightness of the object. For objects for which the flag `HAS_SATUR_DN` is set in a given band, the imaging pipeline includes the counts associated with the bleed trail of saturated objects in flux measurements. In particular, the fiber, Petrosian, PSF, and model magnitudes include this light, and it is added to the central value of the radial profile (i.e., `profMean[0]`). As the pipeline works on a single frame at a time, bleed trails that cross frame boundaries will not be properly accounted for. In addition, the fluxes of close pairs of saturated stars whose saturated regions overlap will not be correct.

### 3.3. Newly Discovered Problems in the Imaging Data

To the best of our knowledge, none of the problems described below is so severe as to make any substantive change to the conclusions of science papers using SDSS data.

- The  $u$  filter has a natural red leak around 7100 Å (cf. Smith et al. 2002), which is supposed to be blocked by an interference coating. However, under the vacuum in the camera, the wavelength cutoff of the interference coating has shifted redward (see the discussion in the EDR paper), allowing some of this red leak through. The extent of this contamination is different for each camera column. It is not completely clear if the effect is deterministic; there is some evidence that it is variable from one run to another with very similar conditions in a given camera column. Roughly speaking, however, this is a 0.02 magnitude effect in the  $u$  magnitudes for mid-K stars (and galaxies of similar color), increasing to 0.06 magnitude for M0 stars ( $r - i \approx 0.5$ ), 0.2 magnitude at  $r - i \approx 1.2$ , and 0.3 magnitude at  $r - i = 1.5$ . There is a large dispersion in the red leak for the redder stars, caused by three effects:
  - The differences in the detailed red leak response from column to column, beating with the complex red spectra of these objects.
  - The almost certain time variability of the red leak.

- The red-leak images on the  $u$  chips are out of focus and are not centered at the same place as the  $u$  image because of lateral color in the optics and differential refraction—this means that the fraction of the red-leak flux recovered by the PSF fitting depends on the amount of centroid displacement.

To make matters even more complicated, this is a *detector* effect. This means that it is not the real  $u - i$  and  $u - z$  which drive the excess, but the instrumental colors (i.e., including the effects of atmospheric extinction), so the leak is worse at high airmass, when the true ultraviolet flux is heavily absorbed but the infrared flux is relatively unaffected. Given these complications, we cannot recommend a specific correction to the  $u$ -band magnitudes of red stars, and warn the user of these data about over-interpreting results on colors involving the  $u$  band for stars later than K.

- There is a slight and only recently recognized downward bias in the determination of the sky level in the photometry, at the level of roughly 0.1 DN per pixel. This is apparent if one compares large-aperture and PSF photometry of faint stars; the bias is of order 29 mag arcsec<sup>-2</sup> in  $r$ . This, together with scattered light problems in the  $u$  band, can cause of order 10% errors in the  $u$ -band Petrosian fluxes of large galaxies.
- The SDSS photometry is intended to be on the AB system (Oke & Gunn 1983), by which a magnitude 0 object should have the same counts as a flat-spectrum source of  $F_\nu = 3631$  Jy. However, this is known not to be exactly true; the SDSS photometric zeropoints are slightly off the AB standard. We are continuing our effort to pin down these offsets. Our present estimate, based on comparison to the STIS standards of Bohlin, Dickinson, & Calzetti (2001) and confirmed by SDSS photometry and spectroscopy of fainter hot white dwarfs, is that the  $u$ -band zeropoint is in error by 0.04 mag,  $u_{AB} = u_{SDSS} - 0.04$  mag, and that  $g$ ,  $r$ , and  $i$  are close to AB. These statements are certainly not precise to better than 0.01 mag; in addition, they depend critically on the system response of the SDSS 2.5-meter, which was measured by Doi et al. (2004, in preparation). The  $z$ -band zeropoint is not as certain at this time, but there is mild evidence that it may be shifted by about 0.02 mag:  $z_{AB} = z_{SDSS} + 0.02$  mag. The large shift in the  $u$  band was expected because the adopted magnitude of the SDSS standard BD+17° 4708 in Fukugita et al. (1996) was computed at zero airmass, thereby making the assumed  $u$  response bluer than that at the mean airmass. We intend to give a fuller report on the SDSS zeropoints, with uncertainties, in the near future. For further discussion of the conversion between magnitudes and physical units, see Appendix A.
- About 0.3% of the DR2 imaging footprint area (about 300 out of 100,000 fields, or 10 deg<sup>2</sup>) for DR2 are marked as ‘holes’, indicated in the CAS by setting quality = 5 (HOLE) in the ‘Field’ table. These are areas of sky where no objects are cataloged, and researchers interested in structure statistics of galaxy or star distributions may wish to mask out these holes from their coverage map. Roughly half of these fields include a very bright star (generally  $r < 5$ ) or a very large galaxy or globular cluster, causing the object deblending in the photometric

pipeline to time out. While no catalog information is available for these fields, the corrected image is available. Data of sufficiently poor quality can also be marked as a hole: very poor seeing (significantly worse than  $2''$  FWHM), glitches in the telescope tracking, and non-photometric data. There are also a few small gaps, also marked as holes, which fall between two adjacent SDSS scans.

- The  $u$  chip in the third column of the camera is read out on two amplifiers. On occasion, electronic problems on this chip caused one of the two amplifiers to fail, meaning that half the chip has no detected objects on it. This was a problem for only two of the 105 imaging runs included in DR2: run 2190, which includes a total of 360 frames in two separate contiguous pieces on strip 12N (centered roughly at  $\delta = +5^\circ$  in the North Galactic Cap; NGC), and run 2189, which includes 76 frames on stripe 36N near the northern boundary of the contiguous area in the NGC. The relevant frames are flagged as bad in the `quality` flag; in addition, individual objects in this region have the  $u$  band flagged as `NOTCHECKED_CENTER` (or, for objects which straddle the boundary between the two amplifiers, `LOCAL_EDGE`). Richards et al. (2002) describe how the quasar selection algorithm handles such data; the net effect is that no quasars are selected by the  $ugri$  branch of the algorithm for these data.

## 4. The SDSS Spectroscopic Data

### 4.1. Improvements to Spectrophotometric Calibration

There have been three substantial improvements to the algorithms which photometrically calibrate the spectra (Tremonti et al. 2004, in preparation): (1) improved matching of observed standard stars to models; (2) tying the spectrophotometry directly to the observed fiber magnitudes from the photometric pipeline; and (3) no longer using the “smear” exposures.

*Analysis of spectroscopic standard stars:* As described in the EDR paper, each spectroscopic plate contains 16 spectrophotometric standard stars, chosen by their colors to be F subdwarf stars. In the EDR and DR1 `Spectro2d` calibration pipelines, fluxing was achieved by *assuming* that the mean spectrum of the stars on each half-plate was equivalent to a synthetic composite F8 subdwarf spectrum from Pickles (1998). In the reductions included in DR2, the spectrum of each standard star is spectrally typed by comparing with a grid of theoretical spectra generated from Kurucz model atmospheres (cf. Kurucz 1992) using the spectral synthesis code `SPECTRUM` (Gray & Corbally 1994; Gray, Graham, & Hoyt 2001). The flux calibration vector is derived from the average ratio of each star (after correcting for Schlegel et al. [1998] reddening) and its best-fit model.

Unlike the EDR and DR1, the final calibrated DR2 spectra are *not* corrected for foreground Galactic reddening (a relatively small effect; the median  $E(B - V)$  over the survey is 0.034). This may be changed in future data releases.

*Improved Comparison to Fiber Magnitudes:* We now compute the absolute calibration by tying

the  $r$ -band fluxes of the standard star spectra to the fiber magnitudes output by the latest version of the photometric pipeline. The latest version now corrects fiber magnitudes to a constant seeing of  $2''$ , and includes the contribution of flux from overlapping objects in the fiber aperture; these changes greatly improve the overall data consistency.

*Smears:* As the EDR paper describes, “smear” observations are low signal-to-noise ratio (S/N) spectroscopic exposures made through an effective  $5.5'' \times 9''$  aperture, aligned with the parallactic angle. Smears were designed to account for object light excluded from the  $3''$  fiber due to seeing, atmospheric refraction and object extent. However, extensive experiments comparing photometry and spectrophotometry calibrated with and without smear observations have shown that the smear correction provides improvements only for point sources (stars and quasars) with very high S/N. For extended sources (galaxies) the spectrum obtained in the  $3''$  fiber aperture is calibrated to have the total flux and spectral shape of the light in the smear aperture. This is undesirable, for example, if the fiber samples the bulge of a galaxy, but the smear aperture includes much of its disk. For this reason, we do *not* apply the smear correction to the data in DR2.

To the extent that all point sources are centered in the fibers in the same way as are the standards, our flux calibration scheme corrects the spectra for losses due to atmospheric refraction without the use of smears. Extended sources are likely to be slightly over-corrected for atmospheric refraction. However, most galaxies are quite centrally concentrated and more closely resemble point sources than uniform extended sources. In the mean, this overcorrection makes the  $g - r$  color of the galaxy spectra too red by  $\sim 1\%$ .

The left panel of Figure 3 compares  $r$  fiber magnitudes with those synthesized from the spectra of all DR2 objects with spectral S/N per pixel greater than 5. For point sources alone, this rms difference is 0.040 magnitudes, a 45% improvement over DR1. For extended sources, the effect of the smears was to give a systematic offset between spectroscopic and fiber magnitudes of up to a magnitude; with the DR2 reductions, this trend is gone. The slight offset of the mean from zero is a seeing effect.

The right panel of Figure 3 compares the  $g - r$  and  $r - i$  colors derived from the spectra and photometry; the scatter is  $\sim 40\%$  lower in DR2 than in DR1. The few percent offset of the colors from zero is an indication that there are small residual errors in our spectrophotometry, perhaps due to errors in the theoretical models used to calibrate the standard stars, or to offsets between our photometric system and a true AB system (see the discussion at the end of § 3).

To evaluate our spectrophotometry over smaller scales, of order  $100 \text{ \AA}$ , we compared the calibrated spectra of a sample of 166 hot DA white dwarfs drawn from Kleinman et al. (2004) to theoretical models. DA white dwarfs are useful for this comparison because they have simple hydrogen atmospheres that can be accurately modeled (e.g., Finley, Koester, & Basri 1997). Figure 4 shows the results of dividing each spectrum by its best fit model. The median of the curves shows a net residual of order 2% at the bluest wavelengths. This is a major improvement over DR1 where the residuals were of order 15% at  $4300 \text{ \AA}$  due to the mismatch between the observed standard stars

and the assumed model.

## 4.2. Problematic Spectroscopic Plates

Table 2: Plates with known problems

Plate ID	MJD	Problem
302	51668	Electronic Noise
338	51694	Electronic Noise
339	51692	Electronic Noise
341	51690	Electronic Noise
342	51691	Electronic Noise
343	51692	Electronic Noise
344	51693	Electronic Noise
345	51690	Electronic Noise
346	51693	Electronic Noise
349	51699	Electronic Noise
350	51691	Electronic Noise
426	51882	Red light leak
504	52316	Many rejected pixels
721	52228	Many rejected pixels
761	52266	Many rejected pixels
769	52282	Many rejected pixels
770	52282	Many rejected pixels
775	52295	Many rejected pixels
778	52337	Many rejected pixels

A small number of plates, listed in Table 2, suffered from a variety of minor problems. The CCD frames for several plates suffered from a transient electronic problem in the left amplifier of the red camera in Spectrograph 2 that randomly injected noise into either a single pixel or a cluster of pixels. These events mimic cosmic rays and are largely eliminated by the standard data processing pipeline, but they might leave artifacts in the reduced spectra. Another set of plates labeled “Many Rejected Pixels”, suffered from having the spectrograph collimator improperly focused. This problem induced a mismatch between the flatfields and the spectra themselves, causing the optimal extraction process to reject an excessive number of pixels. Comparing overlapping objects from adjacent plates confirms that the redshifts from these problematic plates are unbiased, but the spectra themselves should not be used for precision work or spectrophotometry. Finally, during the exposure of one plate, light from an LED somewhere on the telescope found its way to the spectrographs, resulting in an artificial excess of light centered roughly at 6500Å; the

spectrophotometry of this plate is quite poor.

### 4.3. Stellar Radial Velocities

Spectra for approximately 35,000 Galactic stars of all common spectral types, targetted both by the quasar target selection algorithm, and in directed stellar programs, are available with DR2. Radial velocities (RVs) are stored as redshifts and were measured by cross-correlation to a set of stellar templates. Repeat observations of spectroscopic plates show that the stellar radial velocities are reproducible to roughly  $5 \text{ km s}^{-1}$  for stars brighter than about  $r \sim 18$ . However, the DR2 cross-correlation procedure used introduces small additional systematic errors in addition to possible dispersion errors, depending on spectral type; these systematics are of order  $10 \text{ km s}^{-1}$  or less for stars of spectral type A through K with S/N per resolution element  $> 10$ . For white dwarfs and low S/N A stars, lines are too broad for accurate RV determination, while M dwarf radial velocities are also less reliable due to prominent molecular bands. New for DR2 are measures of the centers and depths of the Ca triplet lines (8500, 8544, 8664 Å); these may be useful for more refined radial velocities of M stars.

We note that in the EDR and DR1, zeropoint errors in the templates resulted in quoted RVs for low metallicity F stars that were too large by  $20 \text{ km s}^{-1}$ , while the velocities of A-type stars were too large by  $49 \text{ km s}^{-1}$  (Yanny et al. 2004). These problems were uncovered through an on-going process involving observations of known RV standards and cross checks with other model and template fitting techniques; we hope to continue to improve the tabulated stellar properties in future data releases or as value added data product releases.

### 4.4. Mismatches Between the Spectra and Photometric Data

Each spectroscopic plate includes 32 sky fibers, placed in regions where the imaging data do not include a detected object; by definition, there is no photometric object associated with these. Similarly, there are 437 fibers among the 367,360 spectra in DR2 which were broken at the time of observation; neither a spectrum nor a photometric object is associated with them. Of the remainder, there are 62 objects for which the matching of fibers to objects cannot be reconstructed with any confidence, and therefore whose right ascensions and declinations are uncertain. For these objects, the right ascension and declination are listed as  $-9999$ .

Errors in the deblending algorithm in the `target` reductions caused spectroscopy to be carried out occasionally on non-existent objects (e.g., diffraction spikes of bright stars or satellite trails), or incorrectly shredded fragments of large galaxies (at  $z = 0.01$ , a full 10% of galaxy targets are such fragments). Many of these objects no longer exist in the `best` imaging reductions with its improvements to the deblender. In other cases, the photometric pipeline timed out during the `best` imaging reductions in fields for which `target` imaging proceeded without problem, so that the `best`

photometry is missing for bona-fide objects. This predominantly happens in fields close to a few very bright stars. We expect to recover objects from these “timeout holes” in future data releases.

A total of 663 spectroscopic objects therefore do not have a counterpart in the `best` images, 0.2% of the total. Of these, 80 (including the 62 unmapped fibers) can only be retrieved from the table `specObjA11` in the database. The remaining 583 objects are contained in the default spectroscopic table `specObj`, but will not be found by queries requesting both photometric and spectroscopic information.

#### 4.5. Redshift and Classification Correctness

As described in the DR1 paper, we have compared the results of two independent codes to measure the redshifts of the spectra. These codes are the pipeline used for the official SDSS reductions, whose redshifts are based on cross-correlation and emission-line fitting, and an independent pipeline which uses a  $\chi^2$  method to fit templates directly to spectra (e.g., Glazebrook et al. 1998). Only 1.7% of the objects in DR2 not included in DR1 show gross differences in redshift and/or classification between the two codes. We examined the spectra of all these cases by eye. One third of those discrepancies are for very low S/N spectra, for which no redshift is determinable; in the vast majority of these cases, both pipelines correctly indicated that they had failed. Of the remainder, the redshifts or spectral classifications in the official reductions were clearly incorrect in 0.3% of the spectra, a few hundred objects in total (many of them intrinsically interesting objects such as extreme broad-absorption line quasars, superpositions of objects, and other oddities). A list of corrections will be posted to the DR2 web site when it is completed. A similar exercise was carried out for the DR1 data (and a similar error rate was found); the resulting corrections are incorporated into the database.

### 5. Target Selection

With the change in the model magnitude code (§ 3.1), the mean  $g - r$  and  $r - i$  model colors of galaxies have shifted by about 0.005 magnitudes. Because the target selection for luminous red galaxies (LRGs; Eisenstein et al. 2001) is very sensitive to color, this would have increased the number density of targets by about 10%. Instead, we shifted the LRG color cuts to compensate; in addition, improved star-galaxy separation allows tighter cuts on the model-PSF quantity by which stars are rejected. Here we give the updated equations 2 and 3 respectively, from Eisenstein et al. (2001):

$$c_{\perp} = (r - i) - (g - r)/4.0 - 0.177; \tag{2}$$

$$c_{\parallel} = 0.7(g - r) + 1.2[(r - i) - 0.177]; \tag{3}$$

equations 4 and 8 for Cut I:

$$r_{\text{Petro}} < 13.116 + c_{\parallel}/0.3; \tag{4}$$



$$r_{\text{PSF}} - r_{\text{model}} > 0.24; \quad (5)$$

and Equations 10, 11, and 13 for Cut II:

$$c_{\perp} > 0.449 - (g - r)/6; \quad (6)$$

$$g - r > 1.296 + 0.25(r - i); \quad (7)$$

$$r_{\text{PSF}} - r_{\text{model}} > 0.4. \quad (8)$$

This new version of LRG target selection is applied to the **best** region of sky reduced with the latest version of the imaging pipeline. It is of course not applied retroactively to the **target** version of the sky, which used older versions of the pipeline.

Due to other subtle differences in the photometric pipeline and the calibration, these changes will not exactly reproduce the selection criteria actually used when spectroscopy was carried out. Indeed, defining an LRG sample based on the **best** reductions will result in large spectroscopic incompleteness because so many objects are close to the boundaries. Instead, one should use the **target** photometry and adjust the calibrations of that relative to the **best** calibration. Of course, if one is interested in photometric properties of single objects, then we recommend the **best** photometry.

The selection of the main galaxy sample (Strauss et al. 2002) is based on Petrosian magnitudes, which have not changed substantially with the latest versions of the pipelines. Thus the magnitude limit for this sample,  $r_{\text{Petro}} < 17.77$ , has not changed. The improvements to the model magnitudes have allowed us to tighten the star-galaxy separation in galaxy target selection; the code uses the same cut as LRG Cut I (equation 5). Note that some of the EDR and DR1 data were selected with other photometric limits; see the discussion in Appendix A of Tegmark et al. (2004) for details.

As described in the EDR and DR1 papers and by Schneider et al. (2003), the quasar spectroscopic target selection algorithm has evolved in the history of the SDSS. The final version described by Richards et al. (2002) went into effect just after the last of the DR1 data were taken. Thus all DR2 data not included in DR1 (i.e., spectroscopic plate numbers greater than and including 716) use exactly the algorithm described in Richards et al. (2002). The most important change implemented at that time was the addition of sharp color cuts for high-redshift quasars.

## 6. The Future

As the name implies, DR2 is the second of a series of releases of what will eventually be the entire Sloan Digital Sky Survey. The third data release, DR3, is planned for late 2004. DR3 will include all SDSS survey-quality data taken through June 2003, and will include of order a 50% increment over DR2 in spectroscopy and imaging. We expect it to use essentially the same software used for the processing of DR2.

Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>.

The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, Princeton University, the United States Naval Observatory, and the University of Washington. We would like to dedicate this paper to the memory of J. Beverly Oke, whose work on instrumentation and photometric calibration over the decades was crucial for the conception and development of the SDSS.

## REFERENCES

- Abazajian, K. et al. 2003, *AJ*, 126, 2018 (DR1 paper)
- Blanton, M.R., Lin, H., Lupton, R.H., Maley, F.M., Young, N., Zehavi, I., & Loveday, J. 2003, *AJ*, 125, 2276
- Bohlin, R. C., Dickinson, M. E., & Calzetti, D. 2001, *AJ*, 122, 2118
- Eisenstein, D.J. et al. 2001, *AJ*, 122, 2267
- Finley, D.S., Koester, D., & Basri, G. 1997, *ApJ*, 488, 375
- Fukugita, M., Ichikawa, T., Gunn, J.E., Doi, M., Shimasaku, K., & Schneider, D.P. 1996, *AJ*, 111, 1748
- Glazebrook, K., Offer, A.R., & Deeley, K. 1998, *ApJ*, 492, 98
- Gray, R.O. & Corbally, C.J. 1994, *AJ*, 107, 742
- Gray, R.O., Graham, P.W., & Hoyt, S.R. 2001, *AJ*, 121, 2159
- Gould, A., & Kollmeier, J. A. 2004, *ApJS*, in press (astro-ph/0310842)
- Gunn, J.E. et al. 1998, *AJ*, 116, 3040
- Hogg, D.W., Finkbeiner, D.P., Schlegel, D.J., & Gunn, J.E. 2001, *AJ*, 122, 2129
- Ivezić, Ž. et al. 2002, *AJ*, 124, 2943
- Kauffmann, G. et al. 2003, *MNRAS*, 341, 54

- Kleinman, S. et al. 2004, ApJ, in press (astro-ph/0402209)
- Kurucz, R.L. 1992, in *The Stellar Populations of Galaxies*, IAU Symposium 149, edited by Beatriz Barbuy and Alvio Renzini (Dordrecht: Kluwer), 225
- Lupton, R.H., Gunn, J.E., Ivezić, Ž., Knapp, G.R., Kent, S., & Yasuda, N. 2001, in *Astronomical Data Analysis Software and Systems X*, edited by F. R. Harnden Jr., F. A. Primini, and H. E. Payne, ASP Conference Proceedings, 238, 269
- Lupton, R.H., Gunn, J.E., & Szalay, A.S. 1999, AJ, 118, 1406
- Monet, D. G., et al. 1998, USNO-A2.0 (11 CD-ROMs), (Flagstaff, U. S. Naval Observatory)
- Monet, D. G., et al. 2003, AJ, 125, 984
- Munn, J. A., et al. 2004, AJ, in press
- Newberg, H.J. et al. 2003, ApJ, 596, L191
- Pickles, A.J. 1998, PASP, 110, 863
- Pier, J.R., Munn, J.A., Hindsley, R.B., Hennessy, G.S., Kent, S.M., Lupton, R.H., & Ivezić, Ž. 2003, AJ, 125, 1559
- Pope, A.C. et al. 2004, ApJ, in press (astro-ph/0401249)
- Richards, G.T. et al. 2002, AJ, 123, 2945
- Schlegel, D.J, Finkbeiner, D.P., & Davis, M. 1998, ApJ, 500, 525
- Schmidt, G.D. et al. 2003, ApJ, 595, 1101
- Schneider, D.P. et al. 2003, AJ, 126, 2579
- Scranton, R. et al. 2002, ApJ, 579, 48
- Smith, J.A. et al. 2002, AJ, 123, 2121
- Stoughton, C. et al. 2002, AJ, 123, 485 (EDR paper)
- Strateva, I.V. et al. 2001, AJ, 122, 1861
- Strauss, M.A. et al. 2002, AJ, 124, 1810
- Tegmark, M. et al. 2004, ApJ, in press (astro-ph/0310725)
- Yanny, B. et al. 2004, ApJ erratum, in press (astro-ph/0301029)
- York, D.G. et al. 2000, AJ, 120, 1579

Zakamska, N. et al. 2003, AJ, 126, 2125

### A. Conversion between Magnitudes, Counts, and Fluxes

As discussed in the EDR paper, the SDSS catalogs report asinh magnitudes (Lupton, Gunn, & Szalay 1999) instead of the conventional logarithmic magnitudes. The two magnitude definitions differ only for objects detected at low signal-to-noise with the SDSS imaging camera. While asinh magnitudes produce meaningful colors for such objects, their magnitudes cannot be converted into Janskys or absolute magnitudes in the same way as those of brighter objects. We therefore give detailed instructions here for the conversion between SDSS asinh magnitudes, Janskys, and imaging camera counts. For reference, the definition of asinh magnitudes is

$$m = -\frac{2.5}{\ln 10} \left[ \operatorname{asinh} \left( \frac{f/f_0}{2b} \right) + \ln(b) \right], \quad (\text{A1})$$

where  $f_0$  is the photometric zero point of each filter and  $b$  is a softening parameter as given in Table 3 (repeated from the EDR paper). The asinh magnitude differs by more than 1% from the conventional magnitude for objects with flux less than about  $10bf_0$ . For objects brighter than this, conversion to Jansky and absolute magnitude can be done in the same way as for conventional magnitudes with less than 1% error.

To the extent that the SDSS photometry is on the AB system (§ 3.3), the photometric zeropoint  $f_0$  is given by 3631 Jy (Fukugita et al. 1996). With this and the assumption that the object’s spectrum is flat in  $f_\nu$ , asinh magnitudes can be converted into Janskies by inverting equation (A1). A correction for the spectral shape can be obtained from synthetic photometry using the SDSS sensitivity curves which are available for download from the instruments/imager section of the SDSS web site.

To convert between asinh magnitudes and counts  $N$  on the SDSS imaging camera, use Equation A1, replacing  $f/f_0$  as follows:

$$\frac{f}{f_0} \rightarrow \frac{N}{t_{\text{exp}}} 10^{0.4(aa+kk \times \text{airmass})} \quad (\text{A2})$$

where  $aa$ ,  $kk$ , and  $\text{airmass}$  are the photometric zeropoint, extinction term and airmass, respectively, for the object’s field and filter from the `field` table or `tsField` file (these tables also contain the gain value), and  $t_{\text{exp}} = 53.907456$  s, the exposure time for each SDSS pixel.

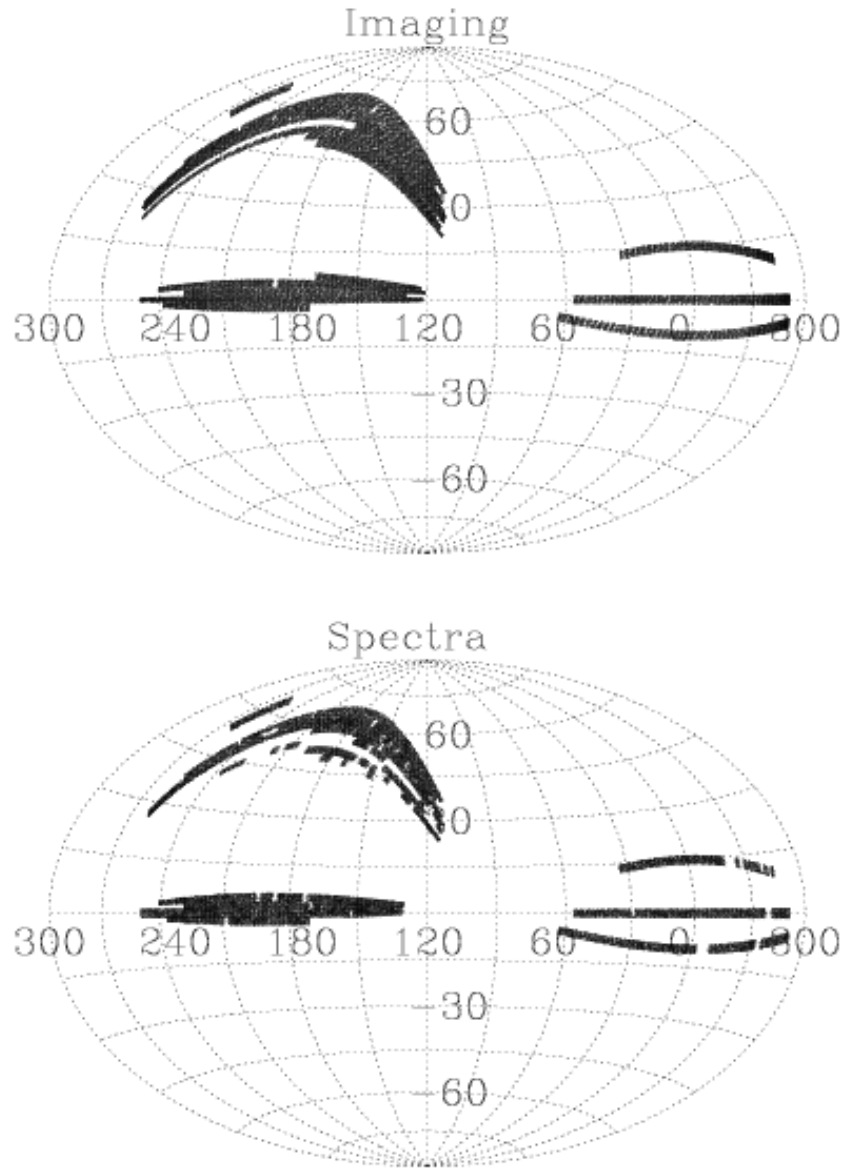


Fig. 1.— The distribution on the sky of the imaging scans and spectroscopic plates included in DR2. This is an Aitoff projection in equatorial coordinates. The total sky area covered by the imaging is 3324 square degrees, and by the spectroscopy is 2627 square degrees.

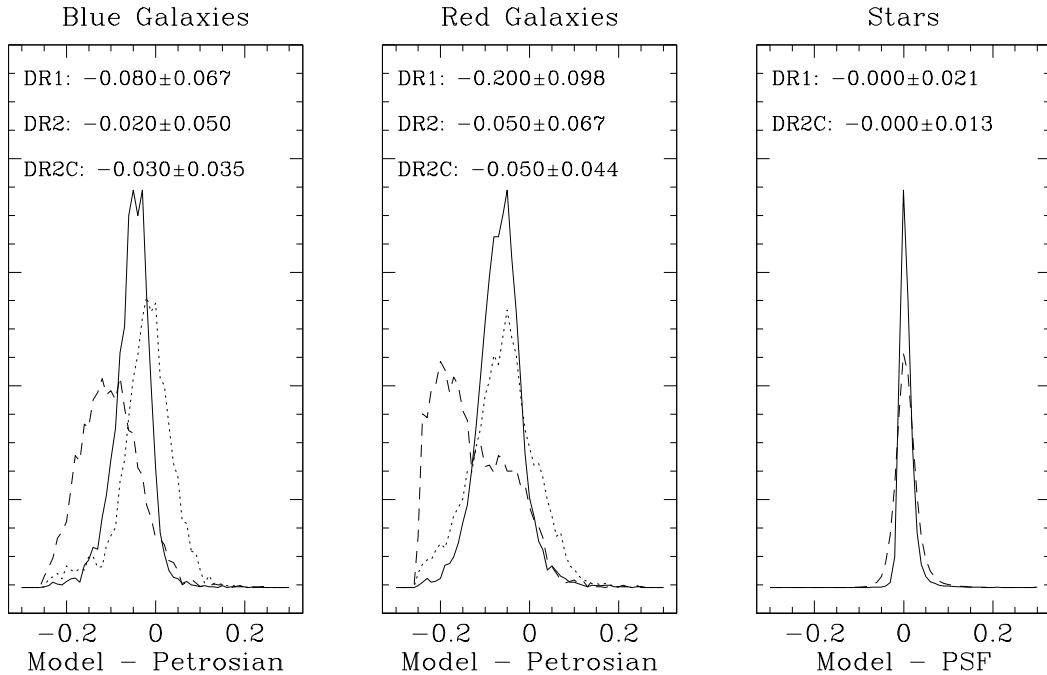


Fig. 2.— Improvements in model magnitudes for stars and galaxies in the DR2 reductions. The first panel shows the distribution of differences between  $r$  band model and Petrosian magnitudes for red ( $u - r > 2.22$ ; Strateva et al. 2001) galaxies brighter than  $r_{Petro} = 19$ ; the three curves are for the old (DR1) reductions (dashed), the current reductions using model magnitudes (dotted; DR2), and the current reductions using `cmodel` magnitudes (solid; DR2C). The mode and standard deviation (based on the interquartile range) of each distribution are given. The bias in model magnitudes in the DR1 reductions is apparent. The second panel shows the same quantities for blue ( $u - r < 2.22$ ) galaxies. The third panel shows the difference between `cmodel` and PSF magnitudes for  $r_{PSF} < 20$  stars, in the DR1 (dotted) and DR2 (solid) reductions; the width of the distribution has decreased by 40% with the new reductions.

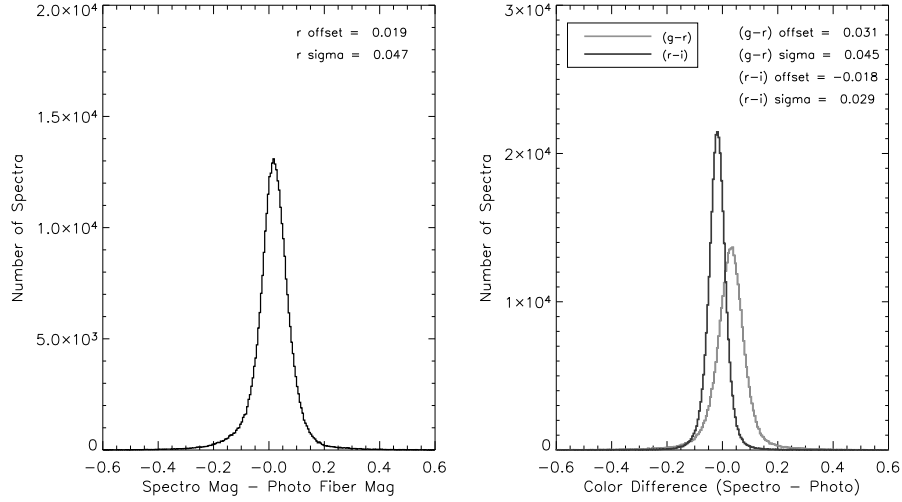


Fig. 3.— Comparison of synthetic  $r$  magnitudes and  $g-r$ ,  $r-i$  colors synthesized from the spectra with photo fiber magnitudes. We have included all objects in DR2 with S/N per pixel  $> 5$ .

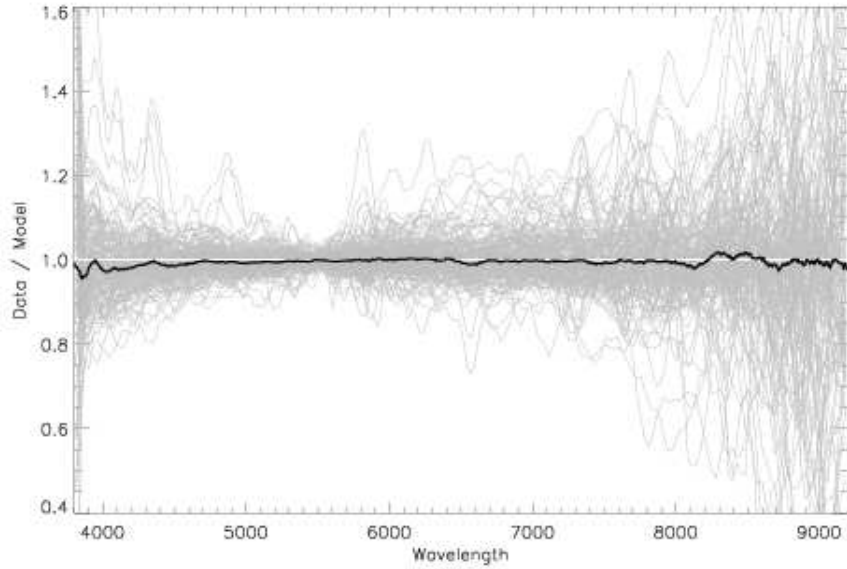


Fig. 4.— Comparison of white dwarf spectra and models. The grey lines represent 166 individual spectra divided by their best fit model. The heavy line is the median. The equivalent median residuals in DR1 were of order 15% at 4300Å; they are now of order a few percent.

Table 3. Asinh Magnitude Softening Parameters

Band	$b$	Zero flux magnitude ( $m(f/f_0 = 0)$ )	$m(f/f_0 = 10b)$
$u$	$1.4 \times 10^{-10}$	24.63	22.12
$g$	$0.9 \times 10^{-10}$	25.11	22.60
$r$	$1.2 \times 10^{-10}$	24.80	22.29
$i$	$1.8 \times 10^{-10}$	24.36	21.85
$z$	$7.4 \times 10^{-10}$	22.83	20.32

Note. — These values of the softening parameter  $b$  are set to be approximate  $1\sigma$  of the sky noise; thus, only low signal-to-noise ratio measurements are affected by the difference between asinh and Pogson magnitudes. The final column gives the asinh magnitude associated with an object for which  $f/f_0 = 10b$ ; the difference between Pogson and asinh magnitudes is  $< 1\%$  for objects brighter than this.