

The Distribution, Classification, and Color Evolution of Galaxies Since $z \sim 0.5$ in the Sloan Digital Sky Survey

A Major Qualifying Project

Submitted to the Faculty

Of the

Worcester Polytechnic Institute

In Partial Fulfillment of the Requirements for the

Degree of Bachelor of Science

By

Tingting Liu

April 26, 2012

Acknowledgement

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, University of Cambridge, University of Florida, the French Participation Group, the German Participation Group, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

Abstract

The Sloan Digital Sky Survey (SDSS) is a large imaging and spectroscopic survey. We took the imaging and spectral data of a sample of galaxies at redshift from 0 to 1 from the SDSS database and aimed at studying the evolution of their rest-frame luminosity and color with redshift. We found that the galaxy distribution is bimodal in all colors in all redshift bins, from which we can classify galaxies into red and blue types. Further, this bimodal distribution has become redder towards present, which suggests the aging of galaxies. Additionally, we found that galaxy luminosity increases with redshift and the number ratio of red to blue galaxies increases with luminosity.

Table of Contents

Acknowledgement	1
Abstract.....	2
1. Problem Statement.....	6
2. Background	7
2.1. The Sloan Digital Sky Survey	7
2.1.1. Overview	7
2.1.2. The Third Sloan Digital Sky Survey and the Eighth SDSS Data Release.....	7
2.2. Main Instruments Used in SDSS.....	8
2.2.1. Telescope	8
2.2.2. Camera	8
2.3. SDSS Imaging Data	9
2.4. SDSS Spectroscopic Data.....	11
2.5. SDSS Image Processing.....	11
2.5.1. The Image Processing Pipelines	11
2.5.2. Star/Galaxy Classification.....	12
2.6. Two of the Coordinate Systems Used by SDSS	12
2.6.1. The Equatorial Coordinate System	12
2.6.2. The Great Circle Coordinate System	13
2.7. The Hertzsprung–Russell Diagram	13
2.8. Morphological Types of Galaxies	14
2.8.1. Overview	14
2.8.2. Spiral Galaxies	15
2.8.3. Elliptical Galaxies.....	15
2.8.4. The Hubble Sequence	16
2.9. Galaxy Color Diagrams	18
2.10. Conclusion.....	19
3. Methodology.....	20
3.1. Overview	20
3.2. The Structured Query Language and SDSS CASJobs	20
3.2.1. Overview	20
3.2.2. Query Basics.....	20

3.3.	Galaxy Search Criteria	23
3.4.	Processing Data.....	24
3.4.1.	SDSS Algorithms	24
3.4.2.	Calculating Rest-Frame Luminosity.....	25
3.4.3.	Magnitude and Color	26
4.	Results.....	27
4.1.	Redshift Distribution of Galaxies Since $z \sim 1$	27
4.2.	Color Distribution of Galaxies Since $z \sim 0.5$	28
4.3.	The Evolution of Galaxies Since $z \sim 0.2$	33
5.	Conclusions	35
	References	36

Table of Figures

Figure 2.1: SDSS-III DR8 Coverage (Credit: SDSS). The white regions are covered by the SDSS Legacy survey, and the red and green regions are covered by SEGUE-1 and SEGUE-2, respectively.	7
Figure 2.2: SDSS Telescope at the Apoche Point Observatory, New Mexico (Credit: SDSS).....	8
Figure 2.3: An illustration of the Arrangement of the CCDs and Five Filters on the SDSS-III Camera (Credit: SDSS).	8
Figure 2.4: The SDSS-III Camera Filter Throughput Curves (Credit: SDSS).....	9
Figure 2.5: Schematic of the Arrangement of the CCDs and Filters on the SDSS Camera (Credit: SDSS). Each row has a different filter.....	10
Figure 2.6: Imaging Run 756: A Small Section of One Run (Credit: SDSS). A “run” is produced by one scan of the SDSS camera.	10
Figure 2.7: Equatorial Coordinate System. One of the intersections points of equator and ecliptic, labeled γ , is the zero point of Right Ascension.....	13
Figure 2.8: One Form of the Hertzsprung-Russell Diagram	14
Figure 2.9: A Spiral Galaxy – the Whirlpool Galaxy (Credit: NASA/ESA).....	15
Figure 2.10: A Barred Spiral Galaxy – NGC 1300 (Credit: NASA/ESA).....	15
Figure 2.11: An Elliptical Galaxy – ESO 325-G004 (Credit: NASA).....	16
Figure 2.12: The Hubble Sequence. “E” is for ellipticals, “SO” is for lenticulars, and “S” and “SB” are for non-barred and barred spirals, respectively.	16
Figure 2.13: A Lenticular Galaxy – the Spindle Galaxy (Credit: NASA).....	17
Figure 2.14: An Irregular Galaxy – the Large Megellanic Cloud (Credit: NASA)	17
Figure 2.15: Galaxy Positions in the SDSS Color Space. Ellipticals, starforming, and non-starforming spirals have different positions in the four-dimensional SDSS color space.	18
Figure 2.16: Histogram of Galaxy Populations. Ellipticals, starforming, and non-starforming spirals have different population distributions for different SDSS colors	19
Figure 4.1: Number Fraction-Redshift Relation of Galaxies.....	27
Figure 4.2: Number Fraction-Look Back Time Relation of Galaxies	28
Figure 4.3: Color Distribution of Galaxies for U-G with 0.05 Redshift Bins	29
Figure 4.4: Color Distribution of Galaxies for G-R with 0.05 Redshift Bins	29
Figure 4.5: Color Distribution of Galaxies for R-I with 0.05 Redshift Bins.....	30
Figure 4.6: Color Distribution of Galaxies for I-Z with 0.05 Redshift Bins	30
Figure 4.7: Bimodal Fit of Galaxies of Redshift from 0 to 0.2	31
Figure 4.8: Bimodal Fit of Galaxies of Redshift from 0.2 to 0.5	31
Figure 4.9: Distribution of Galaxies in U-G for All Redshift Bins	32
Figure 4.10: Evolution of Mean U-G with Redshift	32
Figure 4.11: Evolution of Red/Blue Galaxy Number Ratio with Redshift	33
Figure 4.12: Galaxies Redshift-Color Evolution Since $z \sim 0.2$	33
Figure 4.13: Galaxies Redshift-Color Relation per Luminosity Interval.....	34

1. Problem Statement

The main objective of this project is to study the evolution of galaxies both as a function of the galaxy properties, and as a function of cosmic time. The luminosities of galaxies are plotted against their rest-frame colors, as well as redshifts.

It is observed that the star-forming processes taking place in a galaxy contribute to its color as a whole. It is therefore hypothesized that younger and older galaxies would have different positions on the color plot.

Next, since there are more younger galaxies in the younger universe, or the farther observable universe, than in the older, or nearer universe, it is expected that the luminosity and color of a galaxy will also be a function of its redshift, which is used here as a measurement of position of the galaxy in the time frame of the universe.

2. Background

2.1. The Sloan Digital Sky Survey

2.1.1. Overview

The Sloan Digital Sky Survey (SDSS) is an imaging and spectroscopic survey using an optical telescope at Apache Point Observatory (APO), New Mexico to map the universe on a large scale. A 120 mega-pixel CCD camera images 1.5 square degrees¹ of sky at a time, operating with five filters: ultraviolet, green, red, and two infrareds. Since its launch in April 2000, SDSS has obtained images covering more than one third of the sky (14,555 square degrees) and has obtained spectral data for 9274 square degrees of the sky containing 521,990 stars, 860,836 galaxies, and 116,003 quasars.

2.1.2. The Third Sloan Digital Sky Survey and the Eighth SDSS Data Release

The Third Sloan Digital Sky Survey (SDSS-III) is the third phase of the SDSS which began in August 2008 and will continue through July 2014. SDSS-III includes four separate surveys: the Baryon Oscillation Spectroscopic Survey (BOSS), the Sloan Extension for Galactic Understanding and Exploration 2 (SEGUE-2), the Apache Point Observatory Galactic Evolution Experiment (APOGEE), and the Multi-Object APO Radial Velocity Exoplanet Large-area Survey (Marvels).

SDSS-III's first data release, Data Release 8 (DR8) was made available to the scientific community and the general public in January 2011, which contains all the imaging data so far taken by the SDSS, as well as new spectroscopic data taken for the SEGUE-2 project.

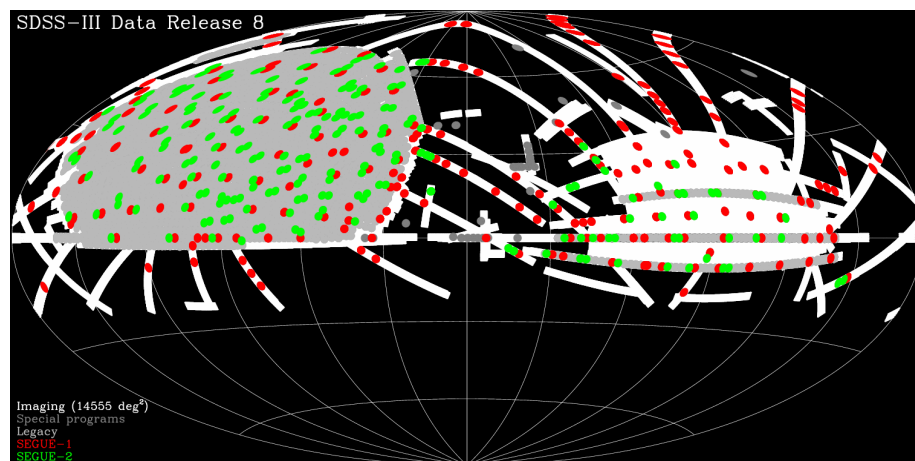


Figure 2.1: SDSS-III DR8 Coverage (Credit: SDSS).
The white regions are covered by the SDSS Legacy survey,
and the red and green regions are covered by SEGUE-1 and SEGUE-2, respectively.

¹ A square degree is a non-SI unit measure of solid angle. 1 square degree = $(\pi/180)^2$.

2.2. Main Instruments Used in SDSS

2.2.1. Telescope

The dedicated SDSS telescope is a 2.5m f/5 reflective optical telescope located at the Apache Point Observatory in New Mexico. It has an unvignetted² 3° field of view.

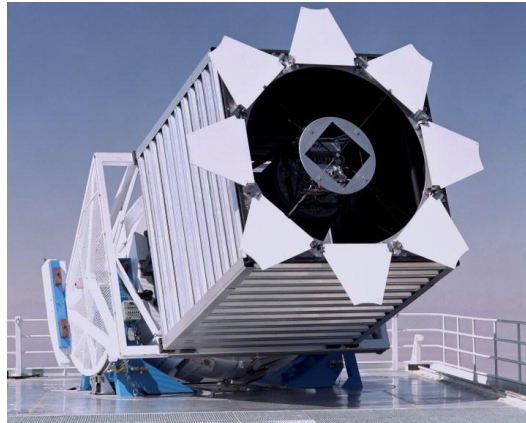


Figure 2.2: SDSS Telescope at the Apache Point Observatory, New Mexico (Credit: SDSS)

2.2.2. Camera

The imaging camera collects imaging data using a 5x6 array of 2048x2048 pixel CCDs on the focal plane of the telescope, aligned with a 5x6 array of filters; the five filters, r, i, u, z, and g, cover the respective rows of the array.

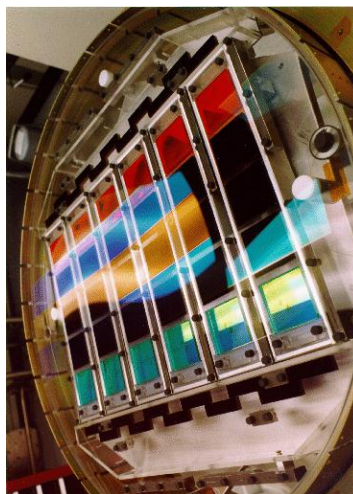


Figure 2.3: An illustration of the Arrangement of the CCDs and Five Filters on the SDSS-III Camera (Credit: SDSS).

² Vignetting is the reduction of image brightness or saturation at the periphery compared to the center of the image.

U (ultraviolet)	G (green)	R (red)	I (near-infrared)	Z (near-infrared)
3551 Å	4686 Å	6166 Å	7480 Å	8932 Å

Table 2.1: Central Wavelengths of Filters

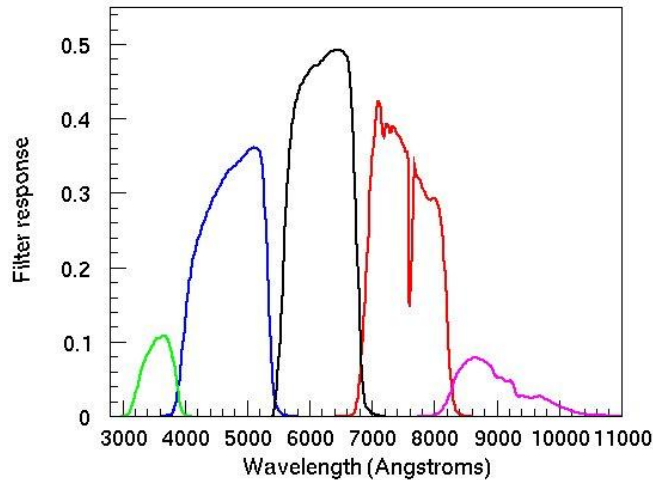


Figure 2.4: The SDSS-III Camera Filter Throughput Curves (Credit: SDSS)

The instrument operates in a drift scan mode: as the telescope moves along great circles, the camera reads the CCDs along the columns at the same rate, producing images of a given object in five bands. Since it takes an object 54.1 seconds to move from one end of a CCD to another, the effective exposure time is 54.1 seconds in each filter. Due the gaps between rows of filters and the time it takes to move from one row to the next, images are taken at 71.7 second intervals. Because there are also gaps between the columns of CCDs, a pass overlaps from its neighbor in such way that the two passes combine to form a solid area.

2.3. SDSS Imaging Data

Three instruments are used for the acquisition of imaging data: the imaging camera mounted on the SDSS telescope, the photometric telescope (a 0.5-meter telescope used for monitoring the atmospheric transmission³), and a 10 micron all-sky scanner (to detect clouds).

³ The phenomenon that the atmosphere absorbs some of electromagnetic radiation, while letting through some others.

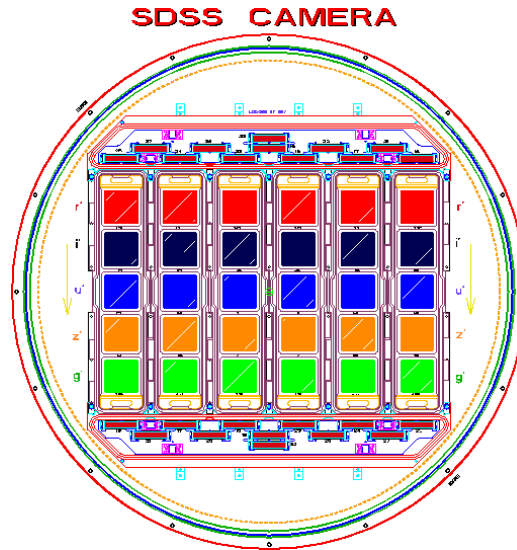


Figure 2.5: Schematic of the Arrangement of the CCDs and Filters on the SDSS Camera (Credit: SDSS). Each row has a different filter.

As mentioned in Section 2.2.2, the camera operates in a drift scan mode. Each scan is called a “run”, which is identified by an integer number from 94 to 8162. Each of the six columns of CCDs is known as a camera column (“camcol”), each being 2048 pixels wide. Each camcol is split into a series of 1489x2048 pixel “fields”, overlapping with the adjacent field by 128 pixels; each field thus has 1361 unique rows.

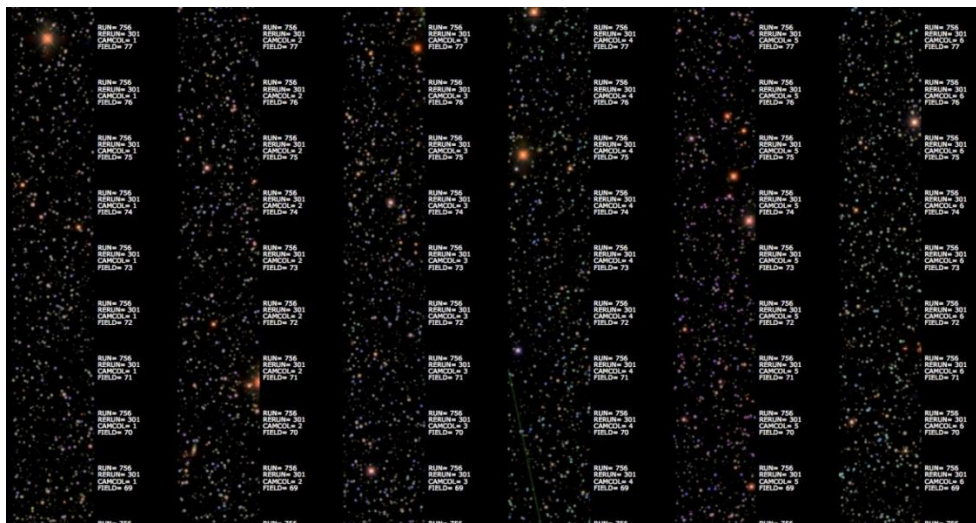


Figure 2.6: Imaging Run 756: A Small Section of One Run (Credit: SDSS). A “run” is produced by one scan of the SDSS camera.

The five rows of CCDs, corresponding to the five rows of filters, take images in the time sequence r, i, u, z, and g; however, in the database catalog, the data are presented in the order of increasing wavelength, i.e. u, g, r, i, and z, and the colored images is a composite of g, r, and i bands.

2.4. SDSS Spectroscopic Data

Spectra data of an object is identified by its plate number, the night on which it is observed, and its fiber number. Each spectra file, also stored in FITS format, is named by the convention “spPlate-pppp-mmmmm.fits”, where pppp is the four-digit plate number and mmmmm is the five-digit MJD number, identifying the night of the observation.

The “plate” is a circular aluminum plate on the focal plane; the SDSS Legacy survey, the largest spectroscopic survey of the SDSS, used around 1800 plates, while SEGUE-1 has 407 plates and SEGUE-2 204. Each has holes drilled on it, into which 640 optical fibers are plugged. During each spectroscopic exposure, typically 900s, two spectrographs (one covering wavelengths from 3800Å to 6100Å, the other from 5900Å to 9100Å) collect spectra from 420 points on the focal plane on two CCD sensors.

The same object may be observed on multiple nights and the fibers can be plugged and unplugged by the observatory staff, and therefore the same fiber number may not correspond to the same object.

Additionally, spectral data are wavelength-calibrated, having been converted, using the IAU standard, from air wavelength (AIR) to vacuum wavelength (VAC), which will be stored in the file.

2.5. SDSS Image Processing

2.5.1. The Image Processing Pipelines

Imaging data from the APO are processed with a photometric pipeline often referred to as “Photo”, which is a combination of pipelines consisting of the Serial Stamp Collecting pipeline (SSC), the Postage Stamp Pipeline (PSP), and the Frames pipeline.

SSC collects the atlas images (the pixels as part of a detected object) of stars for astrometric calibration, which is done by the astrometric pipeline (Astrom) by mapping CCD pixel coordinates to celestial coordinates (Section 2.6.1). Next, PSP determines the point spread function (PSF) of each object; the PSF, which is the response of the atmosphere, telescope, and instrument to a star, varies with position, time, and filter. Finally, the Frames pipeline finds objects by running a PSF-matched filter over the images, “deblends” (separate overlapping images) them, and measures their properties. The Frames outputs are stored as Flexible Image Transport System (FITS, a standard data format used in astronomy) files, available through the SDSS-III database server, the Science Archive Server.

2.5.2. Star/Galaxy Classification

The photometric pipeline classifies objects as point-like (“star”) or extended (“galaxy”) based on the difference between the PSF magnitude (*psfMag*) and the composite model magnitude (*cmodelMag*), measured through each of the five filters. The PSF magnitude is the optimal measure of the flux of a star, determined by fitting a PSF model describing a star to the object. A model magnitude is the equivalent for a galaxy, and the *cmodel* magnitude is the best-fitting linear combination of two profiles of an object, the de Vaucouleurs model and the exponential model, describing elliptical and spiral galaxies, respectively.

The pure de Vaucouleurs profile and the pure exponential profile are two special cases of the Sersic profile which describes how the surface brightness *I* of a galaxy varies as a function of apparent distance *r* from its center and has the form

$$\ln I(r) = \ln I_0 - k * (r)^{\frac{1}{n}}$$

where *I*₀ is the intensity at *r*=0 (center of the galaxy), and *n* is the Sersic index, the values for most galaxies are in the range $\frac{1}{2} < n < 10$, depending on their ellipticity: *n*=4 yields the de Vaucouleurs model, which gives a good description of an elliptical galaxy, while *n*=1 yields the exponential model that best describes spiral and dwarf elliptical galaxies.

An object is classified as galaxy if

$$psfMag - cmodelMag > 0.145$$

and is classified as star if otherwise.

2.6. Two of the Coordinate Systems Used by SDSS

2.6.1. The Equatorial Coordinate System

The equatorial coordinate system is a kind of celestial coordinate system commonly used in astronomy to map celestial objects. It projects the Earth’s equator and geographic poles onto the celestial sphere, which is an imaginary sphere of arbitrarily large radius, concentric with the Earth and rotating around the same axis; the projections of the terrestrial equator and poles are their celestial equivalents.

One of the coordinates is the right ascension (RA, symbol α), the celestial equivalent of the longitudinal angle. Like the terrestrial longitude, RA measures an angle to the east of a zero point, which is known as the first point of Aries, the position of the center of the Sun at vernal equinox. It is convention to measure RA in hours, minutes, seconds; one hour of RA is equal to 1/24 of 360 degrees, which is 15 degrees.

The other coordinate is the declination (dec, symbol δ), the celestial latitude angle. Points on the celestial equator has a declination of 0, points north of the equator has positive declinations, and points south of the equator has negative declinations.

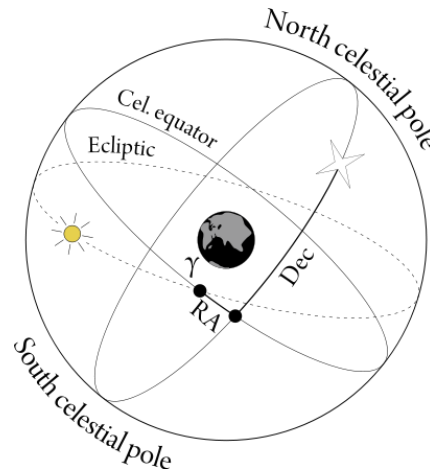


Figure 2.7: Equatorial Coordinate System.
One of the intersections points of equator and ecliptic, labeled γ , is the zero point of Right Ascension.

Additionally, though the equatorial coordinates of a celestial object stay roughly constant, over long periods of time, they are affected by the precession and nutation of the Earth, as well as the proper motion of the object itself. Therefore, when comparing long-term observations (several decades or even centuries), it is necessary to specify a moment in time as the reference point, known as epoch. The current epoch used is J2000.0, for 12:00 Terrestrial Time on January 1, 2000 in the Gregorian calendar.

2.6.2. The Great Circle Coordinate System

SDSS uses another coordinate system, the Great Circle Coordinate System, for the initial collection of data. The two parameters, μ and ν , are the longitude and latitude in the spherical coordinate system whose equator is along the center of the “stripe” in question. The images formed by two scans (Section 2.2.2) are called “strips”, and two “strips” comprise a “stripe”.

2.7. The Hertzsprung–Russell Diagram

Around 1910, two astronomers, Ejnar Hertzsprung and Henry Russell, made significant contribution to the understanding of stellar evolution through the plot of stars now known as the Hertzsprung-Russell diagram (H-R diagram). The H-R diagram is a logarithmic plot of luminosity of stars versus effective surface temperature, and it classifies different types of stars according to their masses: white dwarfs, main sequence stars, and giants.

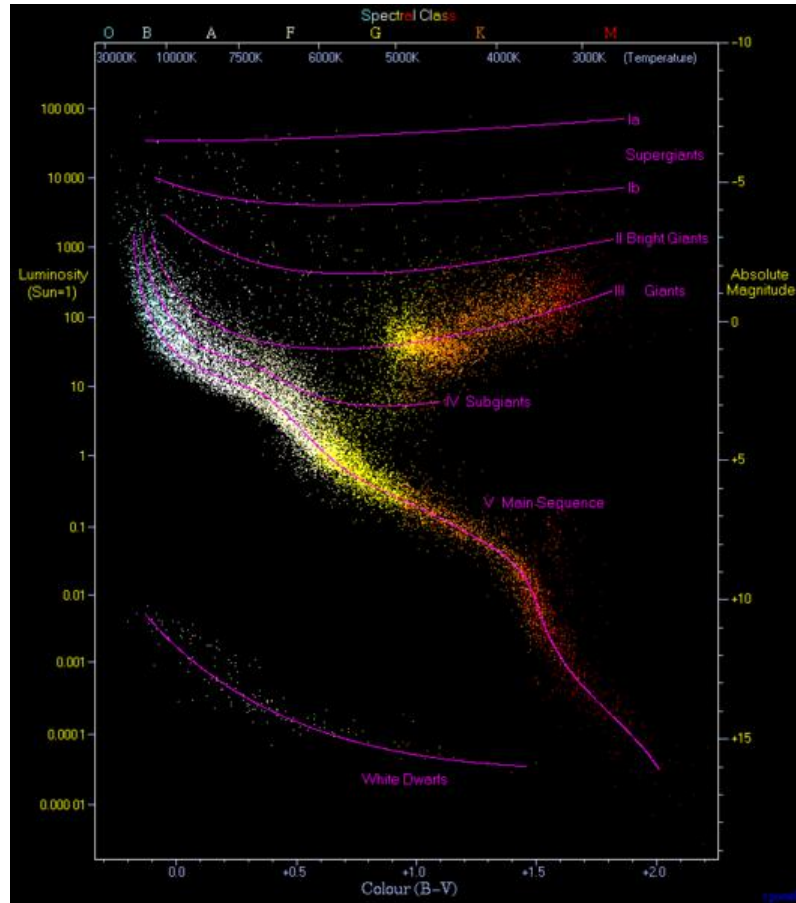


Figure 2.8: One Form of the Hertzsprung-Russell Diagram

The H-R diagram provides an inspiration for this thesis: the classification and evolution of galaxies can also be studied in a similar fashion, by plotting luminosity against color.

Moreover, we are interested in galaxy's luminosity-color relation at different cosmic times and its evolution over time.

2.8. Morphological Types of Galaxies

2.8.1. Overview

Classified according to their morphologies, there are three main types of galaxies in the universe: spiral galaxies, elliptical galaxies, and irregular galaxies. This thesis is mainly interested in the first two types.

2.8.2. Spiral Galaxies

The components of a spiral galaxy are a stellar bulge, a flat stellar and gaseous disk, and a halo of stars, gas, and dark matter. The spiral arms are where gas densities are high in the disk, and the high gas density gives rise to star formation.



Figure 2.9: A Spiral Galaxy – the Whirlpool Galaxy (Credit: NASA/ESA)



Figure 2.10: A Barred Spiral Galaxy – NGC 1300 (Credit: NASA/ESA)

2.8.3. Elliptical Galaxies

In many ways, elliptical galaxies resemble the bulge of spiral galaxies. They are some of the most massive galaxies in the universe, and it is explained that elliptical galaxies are formed as a result of the collision and merger of spiral galaxies. During the galaxy collision, gas clouds

collapse, and vigorous star formations take place. After the merger, however, the density of interstellar medium is low, so low that star forming processes rarely take place. As a result, elliptical galaxies consist of older stars.

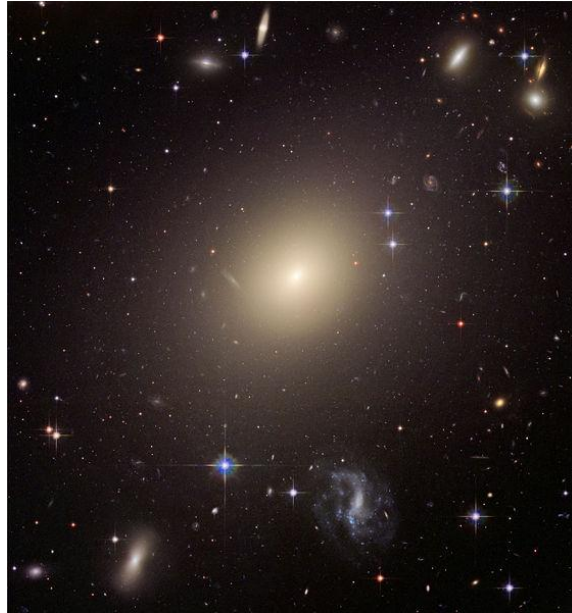


Figure 2.11: An Elliptical Galaxy – ESO 325-G004 (Credit: NASA)

2.8.4. The Hubble Sequence

The Hubble Sequence, named after its inventor Edwin Hubble, is a commonly used morphological classification scheme for galaxies.

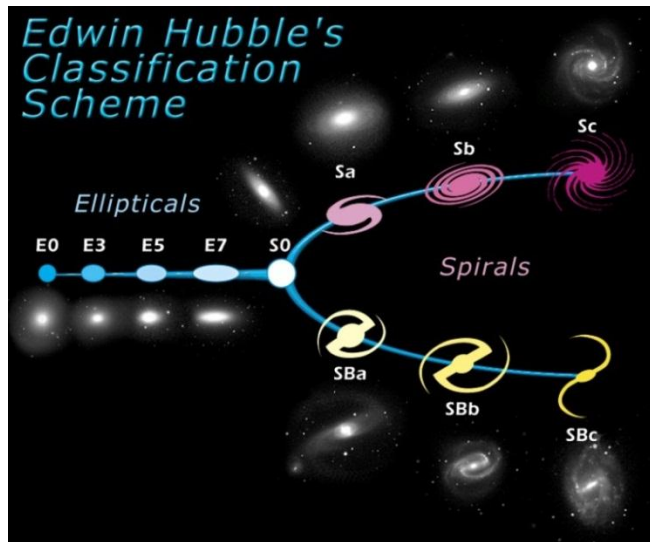


Figure 2.12: The Hubble Sequence. “E” is for ellipticals, “S0” is for lenticulars, and “S” and “SB” are for non-barred and barred spirals, respectively.

According to this scheme, galaxies are classified into three groups according to their morphology: ellipticals (denoted by letter E), lenticulars (S0), and spirals (Sa/SBa, Sb/SBb, Sc/SBc). Irregular galaxies, namely, those that lack regular structure, do not fit into the Hubble Sequence.



Figure 2.13: A Lenticular Galaxy – the Spindle Galaxy (Credit: NASA)



Figure 2.14: An Irregular Galaxy – the Large Magellanic Cloud (Credit: NASA)

2.9. Galaxy Color Diagrams

In our early effort to understand the correlation between the star forming processes in a galaxy and the color indices of the galaxy as a whole, it was discovered that, due to the different levels of star forming processes in different types of galaxies, they have different positions in the four dimensional SDSS color-color space (u-g, g-r, r-i, and i-z).

An small sample (~30) of star-forming galaxies are randomly selected from the SDSS database (identified through their characteristic hydrogen emission lines), and samples of spiral (with no vigorous star-forming taking place) and elliptical galaxies as classified by the NASA/IPAC Extragalactic Database are also selected, after being cross-referenced with the SDSS data. A color-color plot of them is shown as follows.

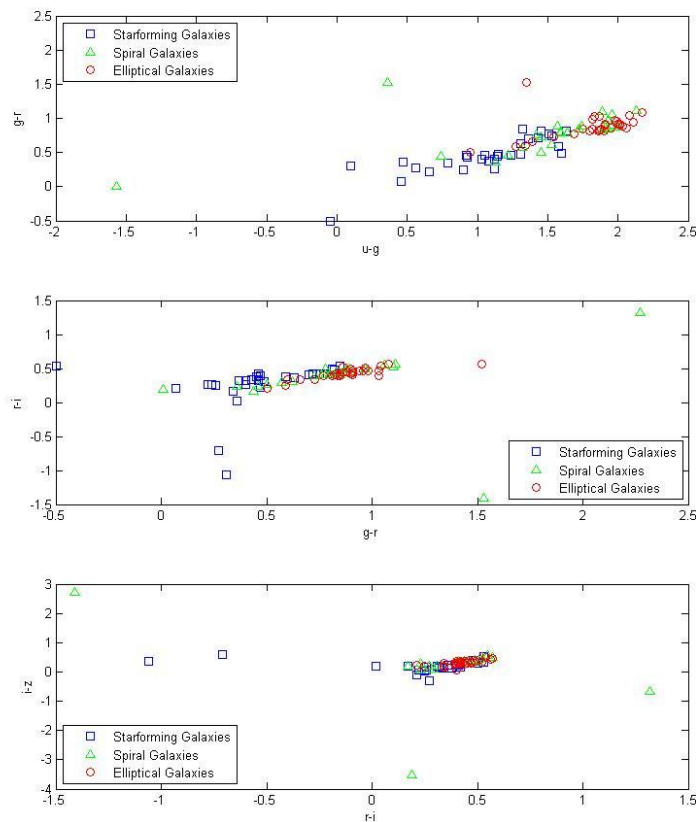


Figure 2.15: Galaxy Positions in the SDSS Color Space.
Ellipticals, starforming, and non-starforming spirals
have different positions in the four-dimensional SDSS color space.

It can be seen from the plot that the three types of galaxies tend to have different positions in the color plot. A histogram of the sample of galaxies also shows the difference of galaxies populations.

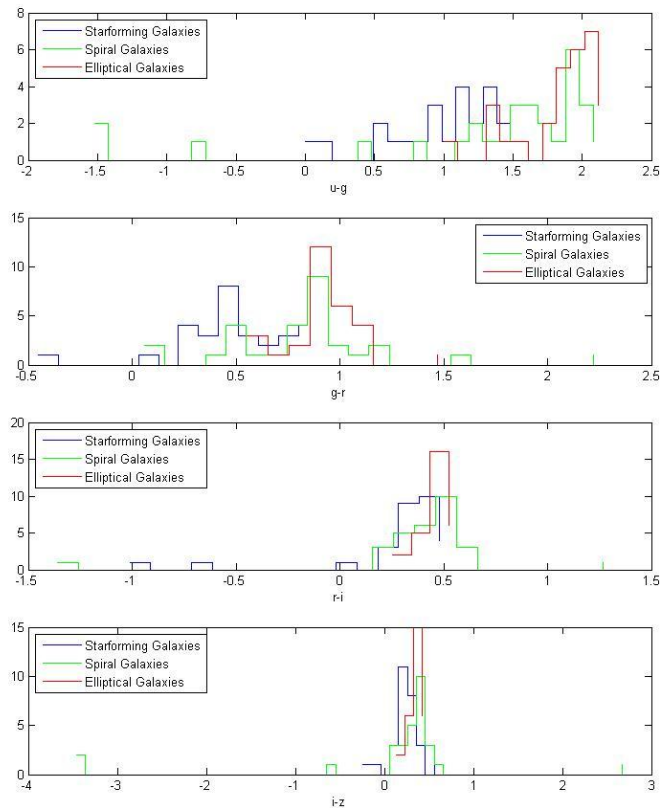


Figure 2.16: Histogram of Galaxy Populations.
Ellipticals, starforming, and non-starforming spirals
have different population distributions for different SDSS colors

The color diagram implies that instead of classify galaxies according to their morphologies, like the Hubble Sequence, it is possible to classify them according to their colors, since colors indicate their starforming activities. Further, star formation rate changes as the galaxy evolve, which enables us to study the color evolution of galaxies.

2.10. Conclusion

It has been observed that different levels of star-forming processes exist in galaxies of different ages, which results in the differences in galaxy spectra and colors. It has also been recognized that the galaxy populations are different in different cosmic times. By connecting the characteristics of galaxies and their positions in the time frame of the universe, we hope to classify galaxies according to their colors and understand their color evolution over cosmic time.

3. Methodology

3.1. Overview

The SDSS database is an ideal source of data for the large-scale study of the universe, thanks to its unprecedentedly large amount of data. So far, the SDSS surveys have covered 14,555 square degrees of the sky⁴, have taken imaging data of 469,053,874 objects in total (among which 208,478,448 are galaxies), and have reached as far as redshift $z=6$.

This section summarizes the method of acquiring data by submitting a query to the SDSS database Catalog Archive Server through the user interface CASJobs, the use of the Structured Query Language, data selection criteria, and the process of data.

3.2. The Structured Query Language and SDSS CASJobs

3.2.1. Overview

Data can be retrieved from the SDSS Catalog Archive Server, which can be accessed through CASJobs, by submitting a query written in a programming language known as the Structured Query Language (SQL). The database returns a table of records selected according to the query, which can be saved in format of csv, xml, or FITS⁵.

The SDSS data are organized in different “tables”, each table containing a different type of data, such as spectra (“SpecObj”) and images (“PhotoObj”). The SDSS database contains 63 tables.

In each table, a “column” presents a certain type of data, such as the object ID number (“objID”) and magnitudes. The number of columns in different tables varies.

3.2.2. Query Basics

The most basic SQL query consists of three “blocks”: “select”, “from”, and “where”. “Select” selects certain columns from the table(s) specified by the “where” block. The search criteria are specified by the “where” block, where different criteria are separated by the logical operators “and”, “or”, and “not”. One can also specify a certain range of the data by using the conditional operators: =, >, <, >=, <=, <>, and “between ... and ...”. Further, if only the first n records are needed, one can use the command “select top n ” in the “select” block.

⁴ The total number of square degrees of a sphere is $129600/\pi$, or approximately 41,253 square degrees.

⁵ FITS, Flexible Image Transport System, the most commonly used digital file format in astronomy.

Since the data for the same objects are stored in different table, multiple tables can be joined together by requiring that the ID numbers of the objects are the same. Additionally, when selecting from multiple tables, the columns of data must be specified in the form “table.column”.

For example: find galaxies for which SDSS has spectra that have g magnitude between 17 and 17.4 and redshift less than 0.05. For each galaxy, retrieve the object ID, the five magnitudes, the redshift, the plate/MJD number, and the fiber number:

```
Select top 10
galaxy.objID,
galaxy.u, galaxy.g, galaxy.r, galaxy.i, galaxy.z,
specObj.z, specObj.fiberID,
plateX.plate, plateX.mjd

from
galaxy, specObj, plateX

where
galaxy.objID=specObj.bestObjID AND
specObj.plateID=plateX.plateID AND
specObj.class='galaxy' AND
galaxy.g BETWEEN 17 AND 17.4 AND
specObj.z < 0.05
```

One can do a simple operation on the returned data by applying an aggregate function. For example, min(x), max(x), avg(x), stdev(x) return, respectively, the minimum, maximum, average, and standard deviation of the returned column x; count(x) or count(*) returns the number of records in column x or the whole table searched.

One can also rename the returned column by using “... as ...” in “select” block. For example, the following command retrieves the minimum, maximum, and average declination of one of the SDSS's equatorial stripes:

```
select
min(dec) as min_dec,
max(dec) as max_dec,
avg(dec) as avg_dec

from
photoObj

where
run = 5112
```

One can group different types of data by a criterion specified in the “group by” block. For example, the following command finds the numbers of records in the specObj table that

have redshift between 0.5 and 1 and groups the results to show how many of the objects are stars, galaxies, quasars, or other:

```
select
  class, count(z) as num_redshift
from
  specObj
where
  z BETWEEN 0.5 AND 1
group by
  class
```

One can also sort results in ascending or descending order by adding an “order by” block. For example, the following command finds the redshift of the nearest galaxy whose spectrum was measured by the SDSS with high confidence:

```
select
  p.objID,
  s.z as redshift,
  p.u, p.r, p.g, p.i, p.z
from
  photoObj p, specObj s
where
  p.objID = s.bestObjID
  and s.class = 'galaxy'
  and s.z > 0
  and p.objID <> 0
  and s.zWarning = 0
order by
  s.z
```

In addition, one can apply one of the SDSS functions on the data, and the program returns an output. The function line is included in the “from” block. For example, the following command returns the RA, Dec, and magnitudes of galaxies near the position ra=140, dec=20.

```
select
  p.ObjID, p.ra, p.dec, p.u, p.g, p.r, p.i, p.z
from
  photoObj p, dbo.fGetNearbyObjEq(140,20,5) n
where
  p.objID = n.objID
  AND p.type = 3
```

3.3. Galaxy Search Criteria

The following types of the data are selected for the purpose of this study:

The “objID”: an object is identified by its objID and we required that the number is non-zero (“objID <> 0”);

Redshift “z”: we require that the Doppler shift of galaxy is between 0 and 1 (“s.z between 0 and 1”) and it is measured with high confidence (“zWarning = 0”);

Absolute magnitudes “modelMag_U”, “modelMag_G”, “modelMag_R”, “modelMag_I”, and “ modelMag_Z”: for measuring colors of extended objects (such as galaxies), the use of model magnitudes (modelMag) instead of composite model magnitudes (cModelMag) are commended by SDSS. As it explains, “the modelMag 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.”

We also select the five model fluxes from the database “modelFlux_u”, “modelFlux_g”, “modelFlux_r”, “modelFlux_i”, “modelFlux_z”.

Finally, we randomly select 10% of the galaxies (“htmID*37 & 0x000000000000FFFF < (650*0.1)”); that returns 84,422 galaxies that match the criteria above.

The command used to select galaxies samples is as follows:

```
SELECT
  g.objID,
  s.z AS spectroscopic_redshift,
  g.modelMag_u AS u, g.modelMag_g AS g, g.modelMag_r AS r, g.modelMag_i AS i,
  g.modelMag_z AS z,
  g.modelFlux_u, g.modelFlux_g, g.modelFlux_r, g.modelFlux_i, g.modelFlux_z,
  p.z AS photometric_redshift,
  p.absMagU AS U, p.absMagG AS G, p.absMagR AS R, p.absMagI AS I, p.absMagZ
  AS Z
  INTO mydb.restframe_magnitude

FROM
  galaxy g,
  specObj s,
  Photoz p

WHERE
  g.objID = s.bestObjID AND
  g.objID = p.objID AND
  s.class = 'galaxy' AND
  g.objID <> 0 AND
  s.z BETWEEN 0 AND 1 AND
  s.zWarning = 0 AND
  g.htmID*37 & 0x000000000000FFFF < (650*10)
```


3.4. Processing Data

3.4.1. SDSS Algorithms

Using the magnitudes and flux densities obtained, we applied the following functions to process the data before analyzing them.

Function Name	Returns...	Parameters
fCosmoAgeOfUniverse	the time interval between a particular redshift and the beginning of the universe	@z: redshift; @OmegaM: matter density; @OmegaL: dark energy density; @OmegaR: radiation density; @Omega0: dark energy state equation @h_0: Hubble Constant
fCosmoLookBackTime	the time interval between the present time and a particular redshift	(Same as Above)
fCosmoDI	the luminosity distance at a given redshift	(Same as Above)

Table 3.1: Functions and Their Outputs

Though there are SDSS functions available to calculate cosmic time and look back time, the following equations can be used when approximate values are needed.

For an object with relativistic Doppler shift z and recessional velocity v :

$$1 + z = \sqrt{\frac{1 + \frac{v}{c}}{1 - \frac{v}{c}}}$$

which gives the radial velocity

$$v = c \frac{(1 + z)^2 - 1}{(1 + z)^2 + 1}$$

Given the Hubble's Law

$$v = H_0 D$$

where H_0 is the present Hubble constant which is approximately 73 km/s/Mpc, we can therefore calculate the proper distance D in Mpc. Expressing the distance in terms of light years using the following conversion

$$1 \text{ Mpc} = 3.26 \text{ Gly}$$

we can find the look back time in Giga years. To get the age of universe at that time, simply subtract the look back time from the age of the universe, which is estimated to be 13 billion years.

3.4.2. Calculating Rest-Frame Luminosity

For an object with a non-zero redshift, the transformation between observed (e.g. apparent magnitudes) and rest-frame (e.g. absolute magnitudes) photometric measurements is known as the “K correction”, and it was described in detail in Humason, Mayall & Sandage 1956 and Oke & Sandage 1968, as well as in Hogg et al 2002. The K correction factor K_{QR} is defined by

$$m_R = M_Q + DM + K_{QR}$$

where m_R is the apparent magnitude observed in bandpass R, M_Q is absolute magnitude emitted in bandpass Q (the spectrum has been shifted from Q to R due to the redshift), and DM is the distance modulus, a distance measure relating apparent and absolute magnitudes.

Though the magnitudes obtained from the SDSS are absolute and therefore have already been converted to the rest frames of the galaxies, we need to do K correction for the fluxes in the five bands. For an object at redshift z , the flux and luminosity densities per unit frequency are related by

$$L(\nu_e) = 4\pi D_L^2 \frac{f(\nu_0)}{1+z}$$

Where D_L is the luminosity distance, and $\nu_e = (1+z)\nu_0$.

	Ultraviolet d_u	Green d_g	Red d_r	Infrared d_i	Infrared d_z
Effective Width (Å)	6000	1400	1400	1500	1200

Table 3.2: Effective Wavelengths and Widths of SDSS Filters (Gunn, SDSS Project Book)

Given the flux densities f per unit wavelength, we are able to approximate the flux of a galaxy F (erg/s/m²) summed over all SDSS bands using the effective filter widths.

$$F = 3.631 * 10^{-25} * 10^{10} * c * \left(\frac{f_u}{d_u} + \frac{f_g}{d_g} + \frac{f_r}{d_r} + \frac{f_i}{d_i} + \frac{f_z}{d_z} \right)$$

where the factor of $3.631 * 10^{-25}$ is due to the unit conversion that 1 nanomaggie = $3.631 * 10^{-6}$ Jansky = $3.631 * 10^{-25}$ erg/s/m²/Hz, and that 1 Å = 10^{-9} m. From here we can calculate the bolometric luminosity (erg/s)

$$L = 4\pi(D_L * 3.1 * 10^{22})^2 * F * (1 + z)$$

3.4.3. Magnitude and Color

The brightness of an object is indicated by its magnitude, which is calculated from its flux. For example, in the commonly used Pogson magnitude system, the apparent magnitude is given by

$$m = -2.5 * \log_{10}\left(\frac{f}{f_0}\right)$$

where f_0 is a reference flux in the band x , and f is the observed flux in the same band. It is worth pointing out that, due to the minus sign, the brighter an object is, the lower its magnitude.

SDSS, however, does not adopt this system. In order to produce better measurements for low signal-to-noise objects, SDSS uses the arcsinh magnitude system by Lupton et al., and it is defined as follows

$$\mu(x) = -a[\sinh^{-1}\left(\frac{x}{2b}\right) + \ln b]$$

where x is the dimensionless normalized flux f/f_0 as in the Pogson magnitude, constant $a = 2.5/\ln(10)$, and b is a “softening constant” for each filter.

To indicate the “color” of an object, we simply subtract the magnitudes in two band passes. For the purpose of the project, we produce U-G, G-R, R-I, and I-Z. Again, according to the definition of magnitude, a larger U-G means that the object is more luminous in the G band than it is in the U band.

4. Results

4.1. Redshift Distribution of Galaxies Since $z \sim 1$

The 84,422 galaxies returned by the SDSS CASJobs are 10% of the galaxies of redshift up to 1 so far available in the SDSS database, and their population distribution is shown in the Fig.4.1. The distribution of galaxies that first increases with redshift then decreases has been both theoretically discussed (Shanks, T. et al, 1984) and observed (Nagashima, M. et al. 2001). Despite the non-smooth distribution due to large-scale fluctuations, we can treat this sample as a uniformly selected sample of galaxies covering a wide range of types.

Apart from the main galaxy samples mostly up to $z \sim 0.25$, SDSS also selects a sample of Luminous Red Galaxies (LRGs), most of which are of redshift from 0.25 to 0.5.

Due to the number density of the galaxy sample, we choose to focus on galaxies since $z \sim 0.5$, and especially since $z \sim 0.2$, which are part of the main galaxies sample.

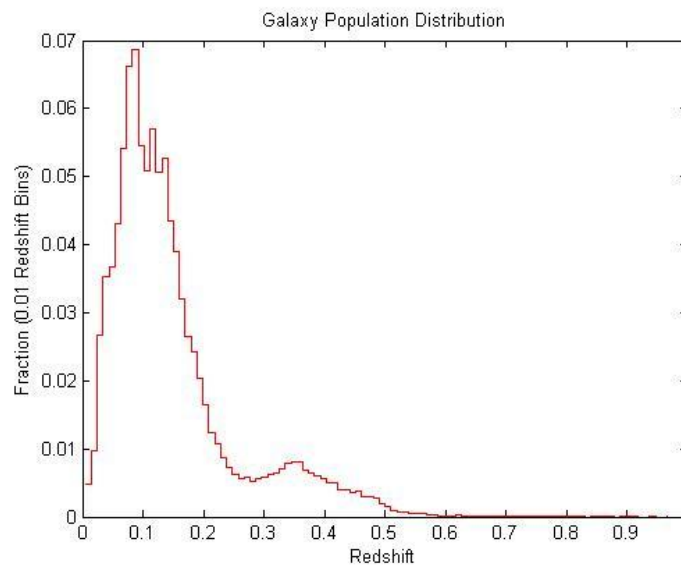


Figure 4.1: Number Fraction-Redshift Relation of Galaxies

We can also express the redshift as cosmic time, in which case the galaxies are dated back to about 8 billion years ago, and the main galaxies sample are from about 3 billion years ago.

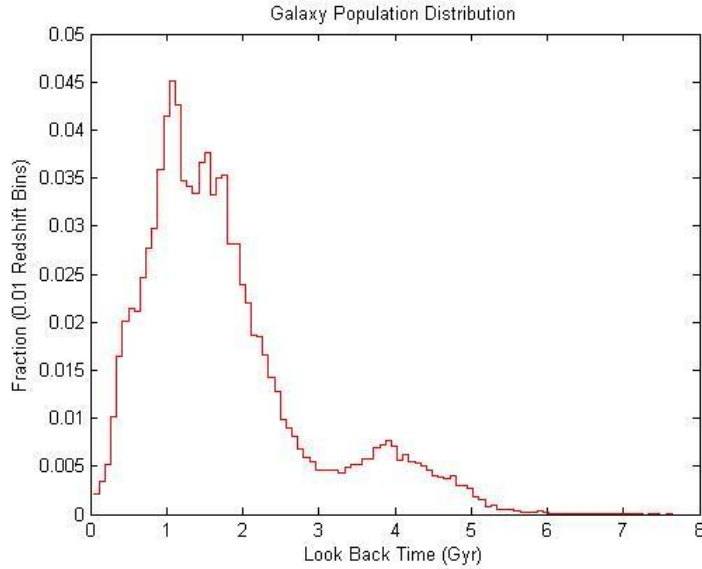


Figure 4.2: Number Fraction-Look Back Time Relation of Galaxies

4.2. Color Distribution of Galaxies Since $z \sim 0.5$

We plotted the number of galaxies, normalized with the total number of galaxies in each redshift bin (with 0.05 redshift bins), against the SDSS colors U-G, G-R, R-I, I-Z. We find that color distribution of galaxies is bimodal for all redshift bins for all colors (Fig. 4.3 through Fig. 4.6). However, we choose to focus on the color U-G (Fig.4.3) since it brackets the 4000\AA Balmer break. It marks the end of starforming activities and is therefore sensitive to metallicity and the aging of stars.

We also find that the peaks of galaxies shift towards red (U-G becomes larger) towards present time (smaller z). This is consistent with our hypothesis that as galaxies are aging over time, they become redder as they age, since there is less star-forming taking place and there are heavier elements in them.

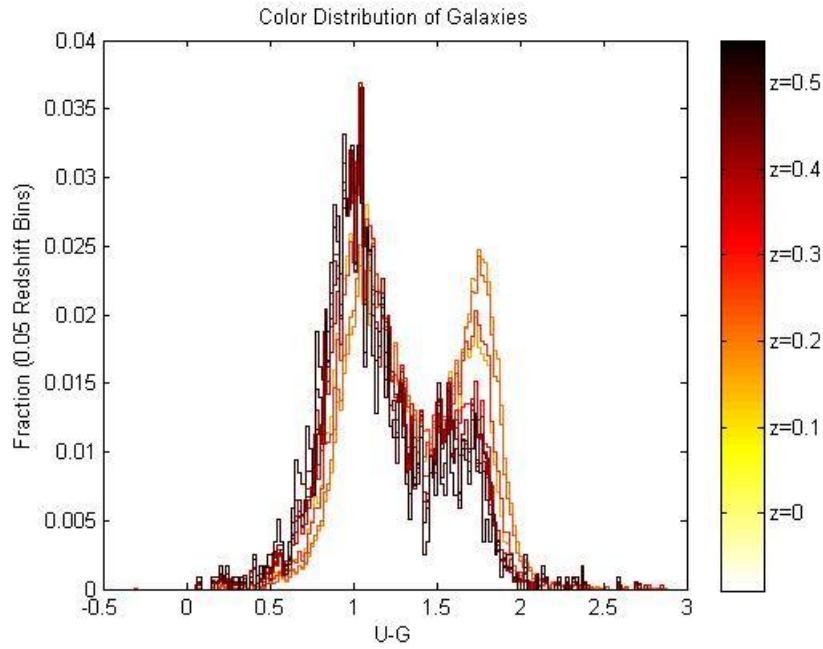


Figure 4.3: Color Distribution of Galaxies for U-G with 0.05 Redshift Bins

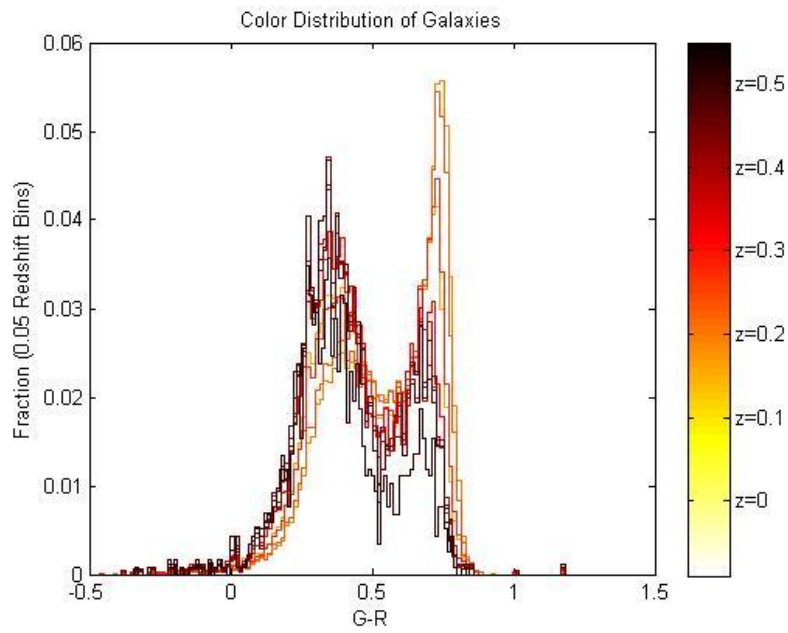


Figure 4.4: Color Distribution of Galaxies for G-R with 0.05 Redshift Bins

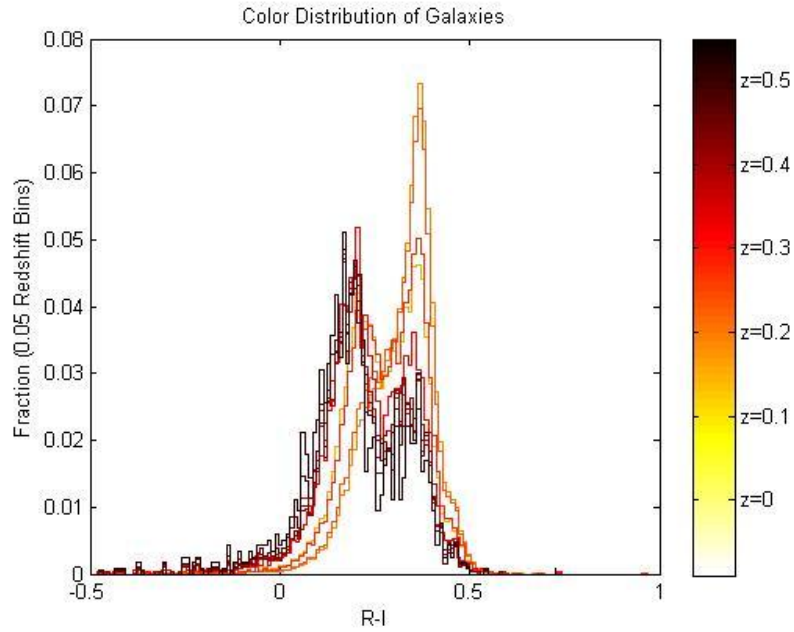


Figure 4.5: Color Distribution of Galaxies for R-I with 0.05 Redshift Bins

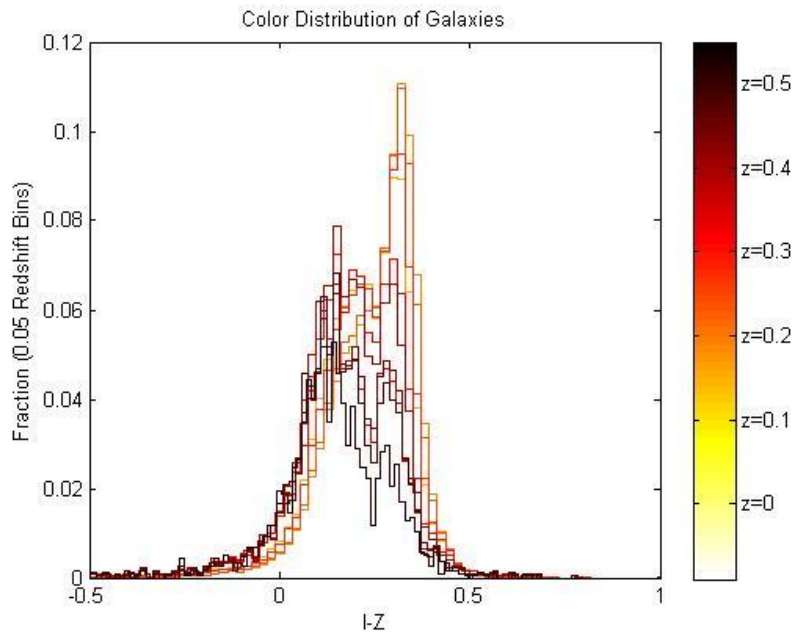


Figure 4.6: Color Distribution of Galaxies for I-Z with 0.05 Redshift Bins

We can fit the distribution with a bimodal distribution. It is a composite of two Gaussian distributions, with one of them being the distribution of blue galaxies (smaller mean U-G value),

the other being red galaxies (larger U-G). Blue galaxies are “late-type” galaxies, responding to Sb, Sc, and Irr types. Red galaxies, on the other hand, are morphologically “early-type” galaxies, responding to E, S0, and Sa in the Hubble Sequence.

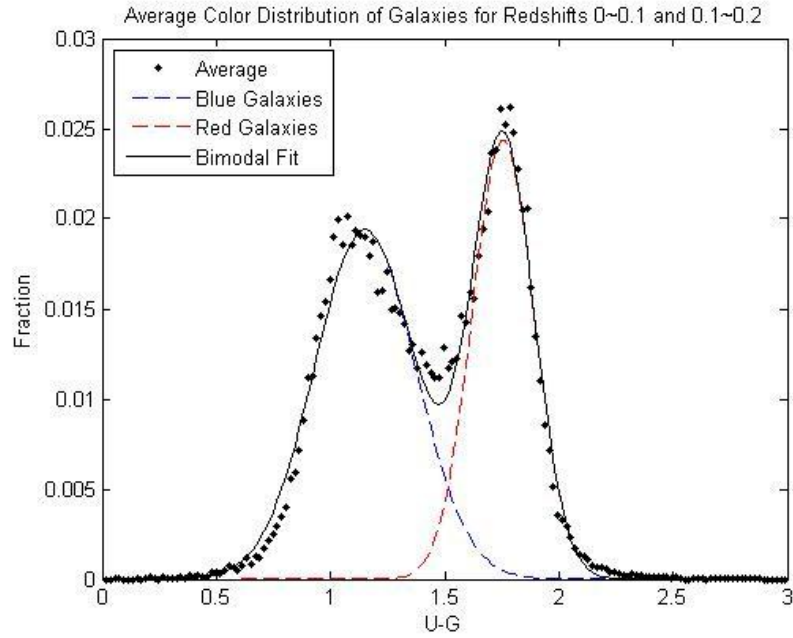


Figure 4.7: Bimodal Fit of Galaxies of Redshift from 0 to 0.2

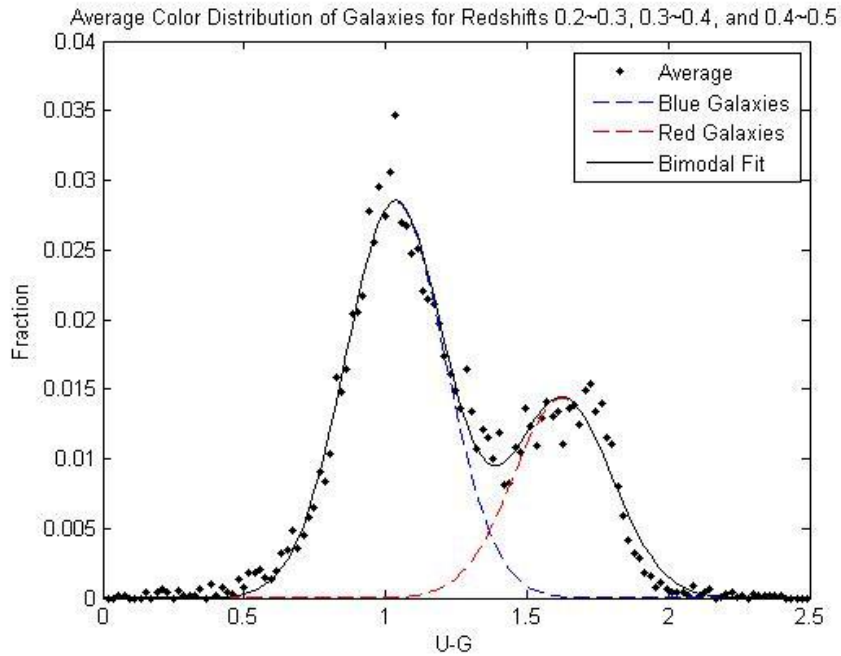


Figure 4.8: Bimodal Fit of Galaxies of Redshift from 0.2 to 0.5

Fig. 4.9 shows the bimodal distribution of U-G color in all 10 redshift bins.

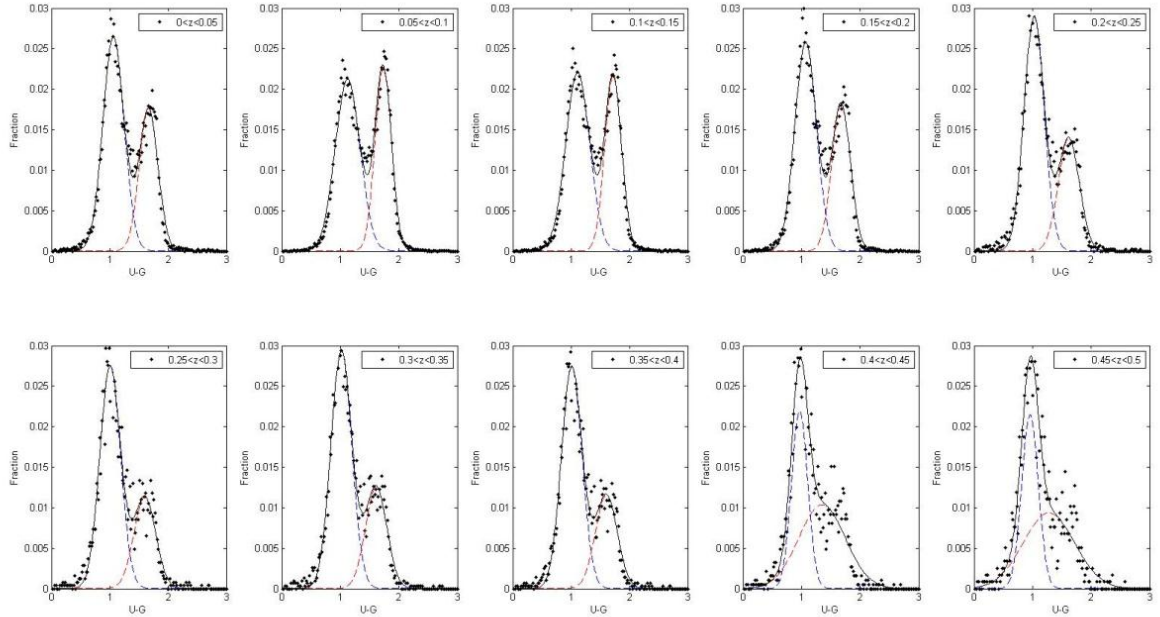


Figure 4.9: Distribution of Galaxies in U-G for All Redshift Bins

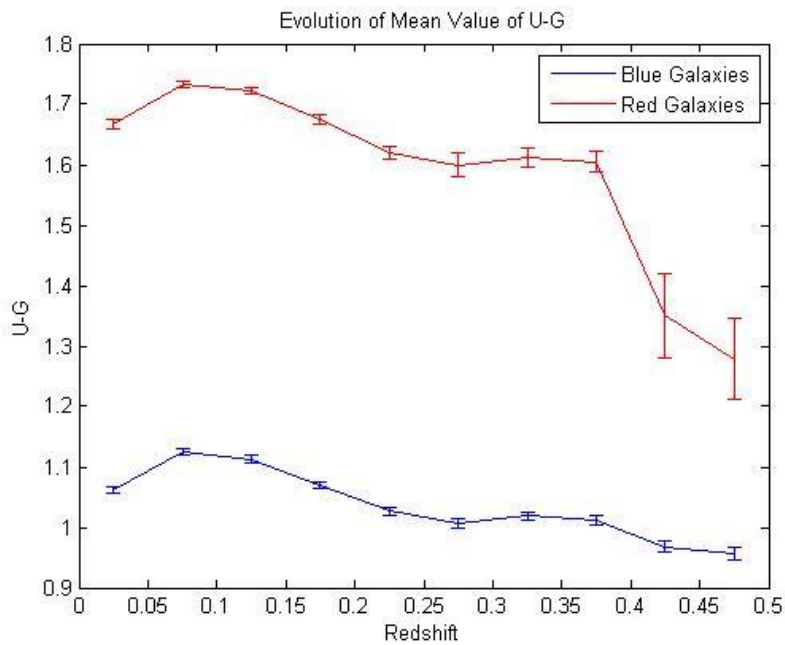


Figure 4.10: Evolution of Mean U-G with Redshift

Fig. 4.10 shows that the mean values of U-G of galaxies tend to become redder towards present, due to the aging of galaxies.

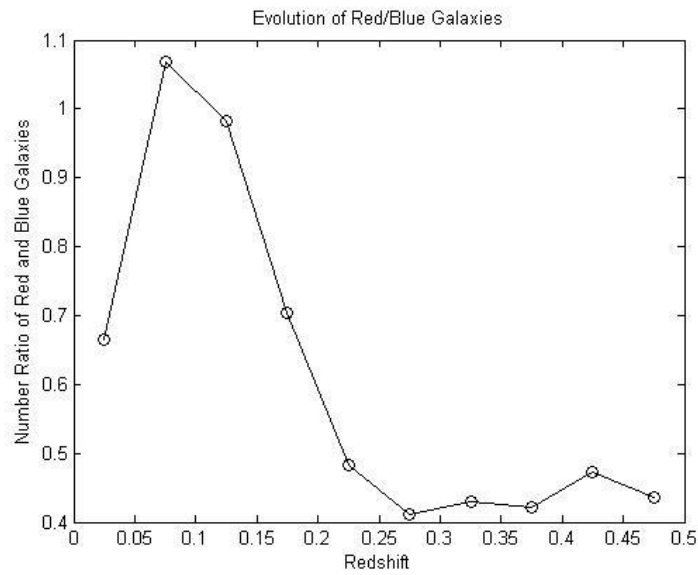


Figure 4.11: Evolution of Red/Blue Galaxy Number Ratio with Redshift

Fig.4.11 shows that the number ratio of red galaxies to blue galaxies has been increasing from $z \sim 0.3$ to $z \sim 0.05$. This is also consistent with the aging of galaxies over time.

4.3. The Evolution of Galaxies Since $z \sim 0.2$

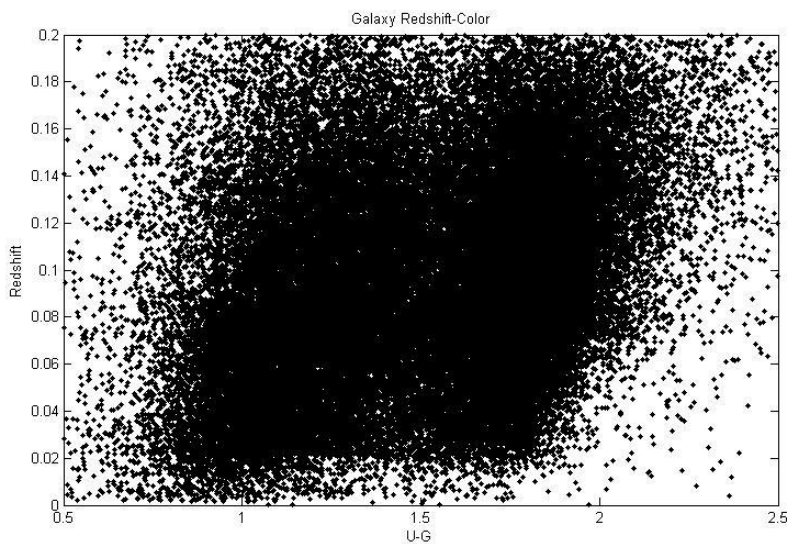


Figure 4.12: Galaxies Redshift-Color Evolution Since $z \sim 0.2$

Figure 4.12 shows the redshift-color relation of galaxies since $z \sim 0.2$. We know from previous discussions that the Gaussian distribution of galaxies with smaller mean value represents blue galaxies, while the one with larger mean value represents red galaxies.

However, in Fig. 4.12 there seems to be a tendency for galaxies to become redder towards the past, while to the contrary, Figures 4.10 and 4.11 suggest that galaxies become redder towards present.

In Fig. 4.13, we replot the redshift-color relation of galaxies with a luminosity interval of $0.05 \cdot 10^{40} W$. We find that in each luminosity interval, the galaxy distribution is bimodal, and the red-to-blue galaxies ratio increase with luminosity. This explains the tilted bimodal distribution in Fig.4.12.

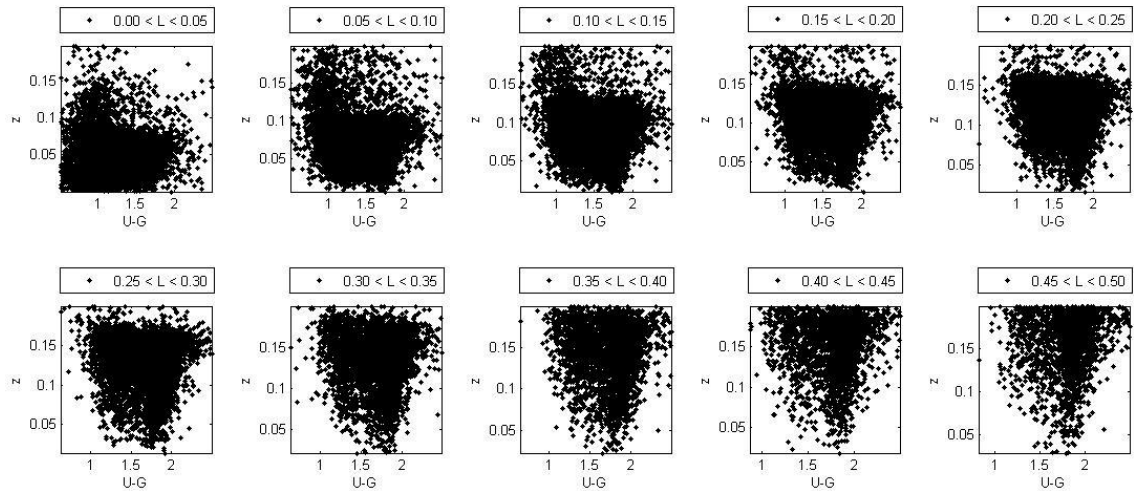


Figure 4.13: Galaxies Redshift-Color Relation per Luminosity Interval

5. Conclusions

We took a sample of galaxies located at $z=1$ or closer from the Sloan Digital Sky Survey database and wanted to study their luminosity-color relation evolving over cosmic time or with redshift.

We hypothesized that galaxies with different levels of star-forming activities would have different positions in the four-dimensional SDSS color space, and therefore we should be able to classify them according to their colors. We first plotted the histograms of galaxy distributions in all four colors for all redshifts (with redshift bin 0.1), and found that the distributions are bimodal. One of the Gaussian distributions has a smaller U-G mean value, which corresponds to the blue or late-type galaxies; the other Gaussian distribution has a larger mean value, which corresponds to the red or early-type galaxies (Fig.4.3-4.6).

Further, the mean values of the Gaussian distribution have been increasing since $z=0.5$ to present, which shows the evolution of galaxy color and agrees with our hypothesis that galaxies are aging over time (Fig.4.10). In addition, the number ratio of red to blue galaxies has also increased since $z=0.2$ (Fig.4.11), which is also consistent with the aging of galaxies.

Additionally, when plotting redshift ranging from 0 to 0.2, we also found that galaxy luminosity increases with redshift, and the red-blue number ratio of galaxies increases with luminosity (Fig.4.13). However, additional work is needed to understand the stated luminosity-redshift and luminosity-number ratio relations.

References

- [1] Aihara H. et al., “The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III” The Astrophysical Journal Supplement, Volume 193, Issue 2, article id. 29 (2011)
- [2] Bell, E. F. et al., “Nearly 5000 Distant Early-Type Galaxies in COMBO-17: A Red Sequence and Its Evolution Since $z \sim 1$ ” The Astrophysical Journal, 608: 752-767, 2004 June 20
- [3] Choudhuri, A.R., 2010, Astrophysics for Physicists, Cambridge University Press, Cambridge, UK, Chap.1.3
- [4] Fukugita, M. et al., “The Sloan Digital Sky Survey Photometric System” The Astronomical Journal, 111: 1748-1756, 1996 April
- [5] Gunn, J. E. et al., “The Sloan Digital Sky Survey Photometric Camera” The Astronomical Journal, 116: 3040-3081, 1998 December
- [6] Gunn, J. E. et al., “The 2.5 m Telescope of the Sloan Digital Sky Survey” The Astronomical Journal, 131: 2332-2359, 2006 April
- [7] Hogg D. W. et al., “The K Correction” arXiv:astro-ph/0210394v1, 17 Oct, 2002
- [8] Komatsu, E. et al., “Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation” The Astrophysical Journal Supplement, 180:330-376, 2009
- [9] Lupton, R. H., et al., “A Modified Magnitude System that Produces Well-Behaved Magnitudes, Colors, and Errors Even for Low Signal-To-Noise Ratio Measurements” The Astronomical Journal, 118: 1406-1410, 1999 September
- [10] Maoz, D., 2007, Astrophysics in A Nutshell, Princeton University Press, Princeton, New Jersey, Chap.5
- [11] Mateus, A. et al., “Semi-Empirical Analysis of Sloan Digital Sky Survey Galaxies – II. The Bimodality of the Galaxy Population Revisited” Monthly Notices of the Royal Astronomical Society, 370: 721-737, 2006
- [12] Stoughton C. et al., “Sloan Digital Sky Survey: Early Data Release” The Astronomical Journal, 123: 485-548, 2002 January
- [13] Strateva, I. et al., “Color Separation of Galaxy Types in the Sloan Digital Sky Survey Imaging Data” The Astronomical Journal, 122: 1861-1874, 2001 October
- [14] York, D. G. et al., “The Sloan Digital Sky Survey: Technical Summary” The Astronomical Journal, 120: 1579-1587, 2000 September