

Absolute Magnitudes of Turnoff Stars in Globular Clusters

Palomar 13 and Whiting 1

Kathleen Grabowski¹, Matthew Newby¹, Heidi Jo Newberg¹

¹ Rensselaer Polytechnic Institute

May 10, 2013

We characterize the F-turnoff (FTO) stars in two globular clusters in the halo of the Milky Way. Data for the two clusters, Palomar 13 and Whiting 1, recently became available in Data Release 8 (DR8) of the Sloan Digital Sky Survey (SDSS). Based on a histogram of the magnitude distribution, we find that the mean magnitude of FTO stars in Palomar 13 is $M_g = 4.46 \pm 0.26$ and the mean magnitude of FTO stars in Whiting 1 is $M_g = 4.11 \pm 0.30$. These measurements are in good agreement with Newby et al. (2011). Palomar 13 is marginally inconsistent within $2\sigma - 3\sigma$ because fewer of the fainter stars are observed than expected. Since Palomar 13 is within the range of cluster properties previously studied, it was not expected that this cluster would be an outlier. The discrepancy is due to lower than expected completeness in measuring faint stars in the cluster, because the seeing in the Palomar 13 images is worse than in other SDSS images of globular clusters. We continue to find that globular clusters in the halo of the Milky Way all have similar absolute magnitude distributions. We also confirm that Whiting 1 is within the Sagittarius dwarf tidal stream, while Palomar 13 is not.

1 Introduction

Stars are classified by temperature using the letters OBAFGKM listed in order from hottest to coolest. The brightness of a star is measured in magnitudes, where a decrease in magnitude of one unit (for example from 0 to -1) is an increase in brightness by a factor of 2.5. The apparent magnitude measures the brightness that a star appears to us on Earth. Absolute magnitude is a measure of how bright the star would be if it was viewed from 32.6 lightyears (10 parsecs) away. Another stellar characteristic, color, is quantified by the difference between the magnitudes in two filters. The Sloan Digital Sky Survey (SDSS)[1] measures the amount of light from each observed star through five different filters, labeled *ugriz*, representing the ultra-violet, green, red, infrared, and one micron parts of the spectrum.

In this paper we will focus specifically on stars of spectral type F. Stars with this temperature have colors of $0.1 < (g-r)_0 < 0.3$, where the zero subscript denotes that the color has been dereddened using the Schlegel, Finkbeiner, and Davis (1998) extinction map [2], to remove the effects of reddening due to dust

between us and the star.

A color-magnitude diagram, also called a Hertzsprung-Russell (HR) diagram, shows some measure of temperature on the x-axis and some measure of brightness on the y-axis. Brighter stars are towards the top of the diagram, and fainter stars are towards the bottom. See Figure 3 in Newby et al. (2011) for representative HR diagrams of globular clusters in SDSS data [3]. A globular cluster is a gravitationally bound group of hundreds of thousands of stars originally formed from one cloud of gas and dust. However, using high accuracy Hubble Space Telescope data, Piotto et al. (2012) recently showed that many globular clusters have more than one stellar population, visible from a broad subgiant branch (SBG) or distinct separation between multiple SGBs [4].

All of the stars begin on the main sequence of the HR diagram and the energy that makes them shine comes from fusion of hydrogen into helium. Although the more massive stars have more hydrogen fuel to burn, they are so much hotter and brighter than less massive stars that they run out of fuel more quickly. As these hot, bright stars move off the main sequence and eventually die, there is a most massive (most blue and bright) group of stars that is still left in the main sequence of the globular cluster. The stars in this group are known as “turnoff” stars. For most globular clusters, the color (temperature) of the stars is spectral type F (hereafter FTO stars).

Stars that have a known intrinsic brightness (absolute magnitude) are often used as distance indicators in astronomy because the intrinsic brightness can be compared with the observed brightness (apparent magnitude) to get the distance using the inverse square law. FTO stars are not traditionally chosen for distance calculations because they span a range of 1.5 magnitudes or more in brightness. When one absolute magnitude is assumed for the FTO stars in a particular globular cluster, the stars would appear to be spread in distance by a factor of 2. However, if there are enough FTO stars in a sky survey, the statistics of the absolute magnitude distribution can be used to map the volume density of stars without knowing the distance to each individual star. This technique is known as statistical photometric parallax (Newberg et al.(2013)) [5]. For example, Newberg et al. (2002) used F-stars to identify new halo structures [5] from spatial overdensities of SDSS stars in the Galactic halo. Even though A stars are each individually better indicators in the halo than FTO stars, the abundance of FTO stars made it possible to see many more halo substructures than Yanny et al. (2000), who used similar data and techniques but used color-selected A stars [6]. Later, Newberg and Yanny used FTO stars to estimate distances to overdensities [7] and halo streams including the Sagittarius stream [9], the Orphan stream [10], and GD-1 [11].

Some authors have measured distances and spatial stellar densities using a maximum likelihood technique [12] that assumes a Gaussian distribution in magnitude of FTO star absolute magnitudes centered at $M_g = 4.2$ in the green passband with $\sigma = 0.6$. This absolute magnitude distribution was suggested from a study

of some of the known halo substructures, but it was not known how universal the absolute magnitude distribution was. In general, the color and brightness of FTO stars depends on age and chemical composition (metallicity) of the stellar population observed. However, Newby et al. (2011) found that the absolute magnitude distribution of color-selected F stars is intrinsically similar for the eleven globular clusters in the Milky Way stellar halo that were studied. They found that the best fit to the data was a “double-sided” Gaussian distribution with a central absolute magnitude of $M_g = 4.18 \pm 0.008$ [3], a left side standard deviation of $\sigma_\ell = 0.36$, and a right side standard deviation of $\sigma_r = 0.76$. In addition, it was discovered that the observed standard deviation appears to be dependent on apparent magnitude, due to increasing errors in the measured colors of fainter and more distance stars in a given survey. Since the FTO stars were selected based on color, the sample is significantly contaminated by stars of different types as the color errors increase [3].

Newby et al. (2011) used eleven globular clusters in the Northern Galactic Cap from SDSS DR7 in their analysis. They spanned ages from 9.5 to 13.5 Gyr and metallicities from $[\text{Fe}/\text{H}] = -1.17$ to $[\text{Fe}/\text{H}] = -2.30$. The similarity in the absolute magnitudes of the FTO stars was surprising, but is apparently due to the Milky Way Age-Metallicity Relationship (AMR) [13, 14, 15] recently discussed by DeAngeli (2005), Marín-Franch (2009), and Dotter (2011). Younger clusters would normally have a brighter turnoff than older clusters, as fewer high-mass stars would have burned out and left the main sequence. Metal-rich clusters usually have fainter turnoffs. Since metal-rich clusters are generally younger than metal-poor clusters, the effects cancel so that the turnoff distribution is largely the same for globular clusters in the Milky Way. Newby et al. (2011) showed that the similarity in turnoff distributions for old populations in the Galactic halo simplifies measurements for distances but complicates estimates for age. Both age and metallicity affect the turnoff magnitude and color of a cluster.

In this paper, we apply the techniques from Newby et al. (2011) to see if the conclusions hold for two additional globular clusters in the Southern Galactic Cap that became available in a later SDSS data release [16]. One cluster has an age and metallicity within the range for clusters chosen by Newby et al. (2011), while the second cluster is younger and more metal-rich than the previous clusters studied. Both clusters studied are more distant than any of the clusters studied by Newby et al. (2011), so our study tests the predictions for higher color errors than previously studied. We present color-magnitude diagrams and isochrones for each cluster. We also show histograms of the magnitude distribution of color-selected FTO stars, which are compared with the predictions of Newby et al. (2011). The mean value of a histogram of the magnitudes of stars for each cluster also marks the bluest point on the cluster’s isochrone. We conclude that the mean magnitude of FTO stars in these globular clusters in the Milky Way are consistent with the results of Newby et al. (2011), though these clusters are more distant and one is at a higher metallicity than previously

studied.

2 Globular Cluster Data Selection

In the Catalog of Parameters for Milky Way Globular Clusters [17] there are three clusters in the Milky Way halo that are in the Southern Galactic Cap portion of the SDSS footprint. One of the three clusters, NGC 7089, was included in the studies by Newby et al. (2011). The other two are Palomar 13 and Whiting 1.

Palomar 13 was studied using the Keck telescope and High Resolution Echelle Spectrometer (HIRES) by Côté et al.(2002) [18] to estimate a distance of 24.3 kpc, equivalent to the distance given by [17]. Bradford et al. (2011) [19] used the Keck/DEep Imaging Multi-Object Spectrograph (DEIMOS) to determine a spectroscopic metallicity of $[Fe/H]=-1.6$ and an age of 12 Gigayear (Gyr) for Palomar 13. Whiting 1 was originally reported by Whiting et al. (2001)[20] and was subsequently studied by Carraro et al. (2007) [21] using BVI photometry and color-magnitude diagrams to obtain an estimated age of $6.5_{-0.5}^{+1.0}$ Gyr, a metallicity of $Z=0.004$ ($[Fe/H]=-0.65$) and a distance of 29.6 kpc, which is closer than the estimate from [17].

To build our data set for each globular cluster, we used the coordinates from [17] to center the globular cluster in the SkyServer Navigate Tool. We magnified the cluster within the frame so that the cluster was visible and visually distinct from the background. The Right Ascension and Declination (α , δ) values were read from the corners of the image, 0.5 degrees from the center of the clusters, and the square area of stars was selected from the SDSS database. The selection limits for Palomar 13 were $347.0^\circ < \alpha < 364.45^\circ$ and $12.4^\circ < \delta < 13.0^\circ$. The selection limits for Whiting 1 were $30.05^\circ < \alpha < 31.05^\circ$ and $-3.55^\circ < \delta < -3.0^\circ$. We queried the SDSS database over the area chosen for each cluster, selecting objects classified as “star” with colors in the range $-0.3 < (g - r)_0 < 0.6$ to include a wide range of stars, including FTO stars, and $(u - g)_0 > 0.4$ to eliminate quasars [5].

The selected stars are shown in Figure 1. The selected stars that are more than 0.037 degrees from each cluster were included in “background” samples. We removed the stars that were located in an annulus with a width of 0.02 degrees at the boundary of the cluster and background stars to remove contamination from background stars. These stars are not shown in the figure.

3 Isochrone Fits

In Figures 2 and 3 we show color-magnitude diagrams of the stars in each globular cluster, and of the stars in a region around each globular cluster, respectively. The background stars are at different distances, have different ages, and should not group together on a color-magnitude diagram in the same way that globular

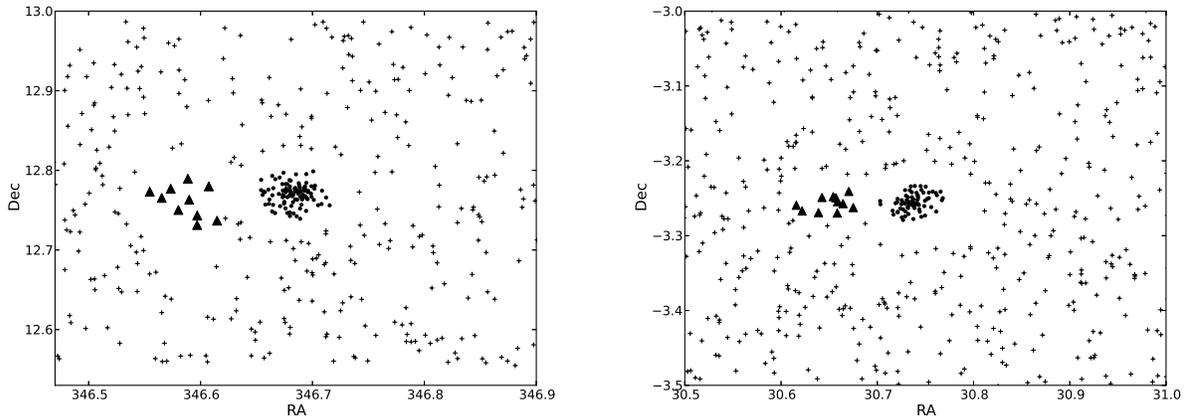


Figure 1: Right Ascension and Declination positions of stars in Palomar 13 on the left and Whiting 1 on the right. In each panel, cluster stars are marked with filled circles, and background stars are marked with plus signs. Stars within an annulus of width 0.02 degrees been discarded at the boundary between the cluster and background stars to prevent contamination in the cluster and background samples. A sample of the background the same angular size as the cluster is marked with solid triangles.

cluster stars do. There is a lack of stars with a magnitude fainter than 22 in the figures, because the fainter stars in globular clusters are obscured by brighter ones [5]. Outside of the globular clusters, the SDSS is efficient at detecting sources to 23rd magnitude in g . The FTO stars attributed to the cluster have a range of about 2 magnitudes. In both panels of Figure 2 and the right panel of Figure 3, most of the stars are bluer than $(g-r)_0=0.4$. In the left panel of Figure 3, there is a higher density of stars redder than $(g-r)_0 = 0.4$

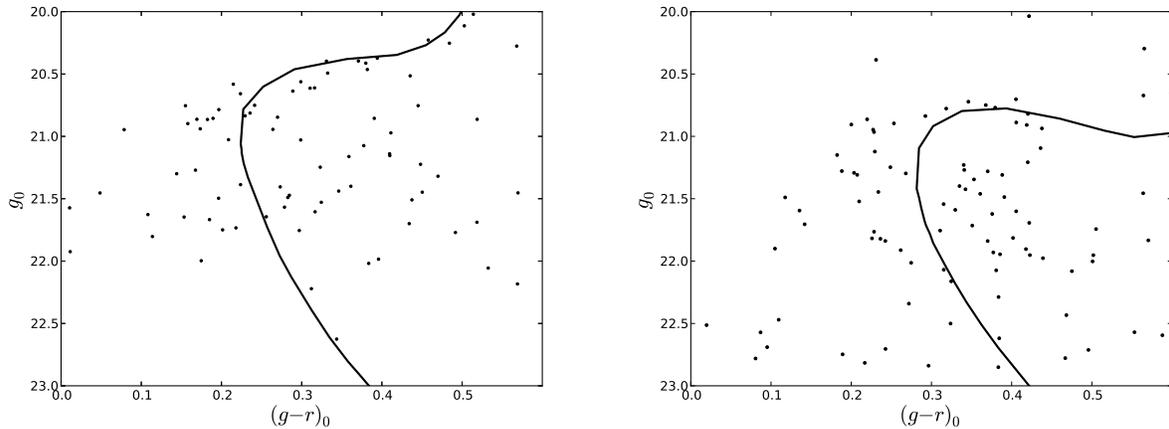


Figure 2: Color-magnitude plots for Palomar 13 (left) and Whiting 1 (right) are shown with isochrones overlaid. We used an age of 12 Gyrs and a metallicity of $Z=0.00039$ for Palomar 13 [19]. We used an estimated age of 6.5 Gyrs and a metallicity $Z=0.004$ for Whiting 1 [21]. The third parameter, distance, was 24.3 kpc [18] for Palomar 13 and 30.1 kpc [17] for Whiting 1. The data has a large scatter in color, but the isochrones fit the data. Because faint stars are overpowered by brighter stars in crowded parts of the sky like globular clusters, there are few stars fainter than the 22nd magnitude.

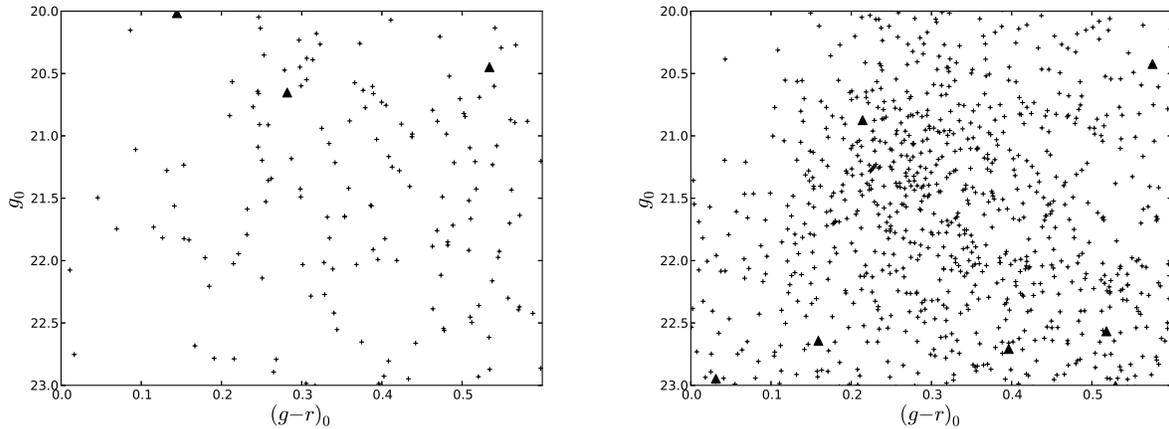


Figure 3: A color-magnitude diagram for background stars near Palomar 13 on the left and near Whiting 1 on the right. The solid triangles show the positions of stars from a sample of the background with the same area as the globular cluster. Not all of the triangles depicted in Figure 4 are visible here because the magnitudes are not in the range shown. The scatter of stars is indicative of a good choice in background. The background of Whiting 1 shows the main sequence of the Sagittarius tidal stream.

Unlike closer globular clusters, it is difficult to pinpoint the turnoff of these clusters because they are close to the magnitude limits of the SDSS photometry and the errors in color are large. Also, much of the main sequence below the turnoff is not visible. Instead of fitting an isochrone to the data, we used the accepted values from other sources for the age and metallicity for each cluster. Following Newby et al. (2011), we used the linear color correction from [3] to calibrate the Padova isochrones [22, 23] to An et al. (2009) [24] results.

For Palomar 13, we used a metallicity of $[\text{Fe}/\text{H}] = -1.7$ and $Z = 0.00039$ and an age of 12 Gyrs [19]. For Whiting 1, we used a metallicity $[\text{Fe}/\text{H}] = -0.65$ and $Z = 0.004$, and an estimated age of the cluster of 6.5 Gyr from photometry measurements [21]. In the HR diagrams for both clusters, there is a concentration of stars at about the right color and magnitude for the cluster turnoff. Due to the large distances of these clusters, the observed color spread of the turnoff is large and the fraction of cluster stars observed is reduced. The color-magnitude diagram for stars in Palomar 13 has few stars with a magnitude fainter than 22 due to the crowding and cluster detection efficiency at faint magnitudes [5]. A decrease in the number of stars fainter than the 22nd magnitude is also evident in the HR diagram for Whiting 1.

The left panel of Figure 2 shows a color-magnitude diagram of Palomar 13 cluster stars where the stars group together and follow the shape of the isochrone. The left panel of Figure 3 shows a color-magnitude diagram of the background stars from Figure 1, showing an even and apparently random distribution of stars. They do not clump around the isochrone. A small area of the background in Figure 1 was chosen to have the same area in the sky as the cluster and was plotted as solid triangles, to show the expected density

of stars that are not part of a cluster. Only three of ten stars are in the portion of the HR diagram shown and they do not follow the same pattern as the color-magnitude diagram for the cluster in Figure 2.

The right panel of Figure 2 shows a color-magnitude diagram of Whiting 1 cluster stars. The isochrone shows that the concentration of stars is near the expected turnoff magnitude. In the right panel of Figure 3, the sample of background chosen to have the same angular size as the clusters are denoted by solid triangles. The background is not part of the cluster, so it is surprising that many of the background stars seem to follow the isochrone.

The magnitude versus color distribution of the background stars selected around Whiting 1 looks similar to the distribution of Whiting 1 itself. It is unusual to find that the stars outside a globular cluster have the same distribution in an HR diagram as the cluster itself. In this case, we believe this is due to the globular cluster's association with the Sagittarius dwarf tidal stream. Law and Majewski (2010)[25] found that Whiting 1 was likely to be part of the trailing tail of the Sagittarius dwarf stream based on angular position, heliocentric distance, and radial velocity. These stars, and possibly the Whiting 1 globular cluster, were ripped away from the Sagittarius dwarf spheroidal galaxy by the tidal forces in the Milky Way.

Though Palomar 13 was originally a candidate for association with the Sagittarius stream, the proper motions from Siegel (2001)[26] led them to conclude that it is not likely that Palomar 13 was part of the Sagittarius stream. We agree with these results, as stars from the Sagittarius dwarf tidal stream do not appear in the background of Palomar 13.

4 Measured versus Expected F-Turnoff Absolute Magnitude Distribution

The globular clusters studied in Newby et al. (2011) had similar μ and σ_ℓ values (where μ is the peak turnoff star absolute magnitude and σ_ℓ is the width of the left side of the Gaussian) regardless of distance, age, or metallicity. This implies that the halo cluster populations are intrinsically similar. There were differences in the measured values of σ_r due only to observational biases in SDSS photometry at faint magnitudes. Equation 12 [3] provides a fourth order polynomial fit to σ_r as a function of distance due to color errors

$$1.741633 + 0.457079d - 0.0250001d^2 + 5.720776 \times 10^{-4}d^3 - 4.7 \times 10^{-6}d^4. \quad (1)$$

This formula is only valid for $10 \leq d \leq 45$ kpc. Errors in colors at large distances reduced the number of true FTO stars detected on the faint side of the histogram, but increased the number of intrinsically redder and fainter stars. The intrinsic value of 0.76 ± 0.04 presented for σ_r is an error-weighted average of the four

closest clusters, which are consistent with each other and have little contamination from red main-sequence stars.

The μ (mean absolute magnitude) determined by Newby et al. (2011) was 4.18 ± 0.008 and σ_ℓ (bright-side σ) was 0.36 ± 0.006 [3] based on the parameter fits to the 11 clusters studied. Figure 16 of Newby et al. (2011) was used to determine the expected faint side width, σ_r , for the two clusters studied here. The faint-side width is affected by the size of the cluster and the distance to the cluster. As the distance to the cluster increases, photometric color errors increase and red main sequence stars leak into the color selection range. At a distance of about 25 kpc away, the contamination from red stars stops increasing as these faint stars begin to fall below the SDSS apparent magnitude detection threshold [3]. The central regions of clusters are often excluded from the SDSS database, as the photometric pipeline was not well-suited for resolving crowded stars. Our two clusters are at such a large distance that only the bright stars are detected, as the fainter stars are washed out or beyond the magnitude limit of the survey.

We apply the same techniques as Newby et al. (2011) to determine whether or not Palomar 13 and Whiting 1 agree with the previously studied clusters. To determine the turnoff using magnitude alone, we made a histogram for each of the globular clusters with bin widths of 0.2 magnitudes. We chose a range of $0 < M_g < 8$, but did not fit to stars with a calculated absolute magnitude brighter than 2 to reduce the chance of binning stars that are not from the cluster.

For each cluster, we used all of the photometric data in Figure 1 to make a histogram of the cluster stars and a separate histogram for the background. We scaled the background histogram to the area of the cluster and subtracted it from the cluster histogram. We then fit a “double-sided” Gaussian distribution to each cluster where the standard deviation is different on each side of the mean. The form of the fit function [3] is

$$G(M_g : \mu, \sigma_\ell, \sigma_r, A) = A \cdot \exp \left[\frac{-(M_g - \mu)^2}{2\sigma_i^2} \right] \quad (2)$$

where $\sigma_i = \sigma_\ell$ if $M_g \leq \mu$ and $\sigma_i = \sigma_r$ if $M_g \geq \mu$. When normalized,

$$A = (2\pi \left[\frac{\sigma_\ell + \sigma_r}{2} \right]^{-1/2}). \quad (3)$$

The parameters $\mu, \sigma_\ell, \sigma_r$, and A are the magnitude at the peak of the population, the standard deviation of the left side, the standard deviation of the right side, and the amplitude, respectively. The amplitude, A , is dependent on the number of stars included in the cluster studies and will vary for each cluster. Since the true Poisson counting errors are given by $\sqrt{N+1} + 1$, the minimum error on a bin height is 2 counts so we do not expect bins with less than 4 counts to significantly affect the fit distribution. A two-stage χ^2

minimization process avoided local minima and accurately determined the values of the parameters. We used a Markov Chain Monte Carlo technique to inspect equilibrium points and then input those points into a gradient descent algorithm to locate the global minimum.

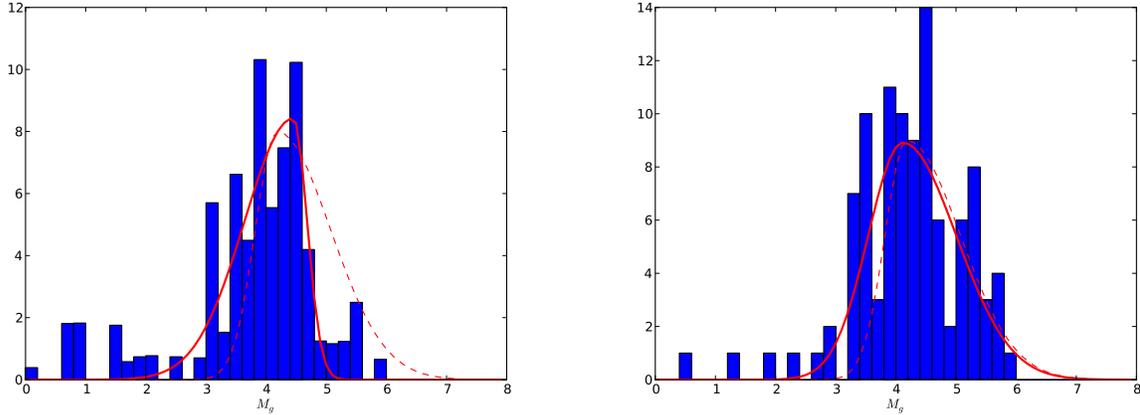


Figure 4: Histogram of absolute magnitudes of Palomar 13 on the left and Whiting 1 on the right. In both plots, the dashed line shows the double-sided Gaussian expected for each cluster with the values obtained from Newby et al. (2011). The solid line in each plot fits a double-sided Gaussian distribution to the histogram.

We present the absolute magnitude and best-fit Gaussian for Palomar 13 in the left panel of Figure 4. The dashed line is the double-sided Gaussian distribution fit anticipated by the values determined by Newby et al. (2011)[3], and the solid line is the double-sided Gaussian fit to Palomar 13 data. The μ value, or the turn-off absolute magnitude, is 4.46 ± 0.26 for the parameter fit. This corresponds to the leftmost and bluest point of the isochrone in Figure 3 at a color $(g - r)_0 = 0.22$. This shows that using the histogram of magnitudes for FTO stars in the cluster yields the same turnoff color and magnitude as the Padova isochrones, given the published age and metallicity. The value predicted from Newby et al. (2011) is $M_g = 4.18 \pm 0.008$. The mean absolute magnitude for this cluster deviates from the value from Newby et al. (2011) by slightly more than 1σ , where σ is the error in M_g .

The measured σ_ℓ and σ_r values for Palomar 13 are 0.81 ± 0.25 and 0.226 ± 0.23 , respectively. The σ_ℓ value for Palomar 13 is much higher than the expected value of 0.36 ± 0.006 [3], but still within 2σ . Though the stars with a magnitude between 0 and 2 were not included in the data analysis to prevent contamination from foreground stars, there is a larger magnitude spread of stars brighter than the mean magnitude than fainter. The expected value of σ_r was 0.89 [3] including the bias against observing faint stars in globular clusters in the SDSS. The value we obtained is almost 3 standard deviations lower than the prediction and there are not very many stars with absolute magnitudes fainter than 5. [5]. The amplitude fit, A, for the double-Gaussian distribution for Palomar 13 is 8.44 ± 2.22 . The amplitude is different for every cluster; the

more cluster stars observed, the larger the amplitude of the Gaussian fit.

The absolute magnitude histogram and the best-fit Gaussian for Whiting 1 are shown in the right panel of Figure 4. The dashed line is the double-sided Gaussian distribution fit anticipated by the values determined by Newby et al. (2011) [3]. The measured and expected Gaussians correspond well on the right side of double-sided Gaussian distributions, but have a larger standard deviation on the left side. The measured mean turnoff absolute magnitude, μ , is 4.11 ± 0.3 . The μ value anticipated by Newby et al. (2011) of 4.18 ± 0.008 is well within the errors. The σ_ℓ parameter for the fit of the data is 0.57 ± 0.25 . The expected value of 0.36 ± 0.006 is consistent with the fit values within the errors. The fit value for σ_r is 0.86 ± 0.27 , compared to the expected value of 0.76 at this distance. Again, this fit value is consistent within errors with the expected value. The amplitude fit, A , for the double-Gaussian distribution for Whiting 1 is 8.91 ± 1.86 .

The expected and observed magnitude distributions for Whiting 1 agree within 1σ error bars. The mean magnitude and the bright side sigma for Palomar 13 agree with the expected values within 2σ . The faint side error is nearly 3σ lower than the expected and there are very few stars with a magnitude fainter than 5. There are fewer faint stars in Palomar 13 than expected. Whiting 1 is the most distant and most metal-rich cluster, so we were surprised that Palomar 13 was the cluster that differs from Newby et al. (2011) results. We suspect the disagreement is the result of poorer seeing in the Palomar 13 image, which affects the width of the point-spread-function. All of the data for Palomar 13 is contained in one run, and the width of the point-spread function in the r filter is over $2''$, which is outside of the usual image quality limit in SDSS. For comparison, the width of the point-spread function in the r filter for stars in Whiting 1 ranges from $1.33''$ to $1.58''$. The poor seeing will increase the chance that fainter stars will be obscured by brighter stars in a globular cluster, and likely leads to the apparent lack of fainter turnoff stars.

5 Age-Metallicity Relationship

Newby et al. (2011) showed that their cluster sample was consistent with the assumed Age-Metallicity Relationship (AMR) for the Galaxy, and that this AMR is the underlying reason for the similar intrinsic turnoff properties of old halo clusters. Since supernovae will enrich star-forming clouds with metals as the Universe ages, the formation time of a cluster should correlate with its metallicity, such that metallicity decreases as age increases. Recent observational evidence [13, 14, 15] suggests that this relationship is more complicated for globular clusters in the Milky Way, with several clusters having very old ages (~ 13 Gyr) but a wide range of metallicities. These constant-age clusters are thought to be coeval with the Milky Way's initial formation event, which is thought to have consisted of rapid star formation and metal enrichment in a short span of time [15]. Figure 17 of Newby et al. (2011) illustrated that their studied clusters were consistent

with the decreasing-metallicity-with-age-trend (the AMR), and we reproduce that plot, with Whiting 1 and Palomar 13 added, as Figure 5.

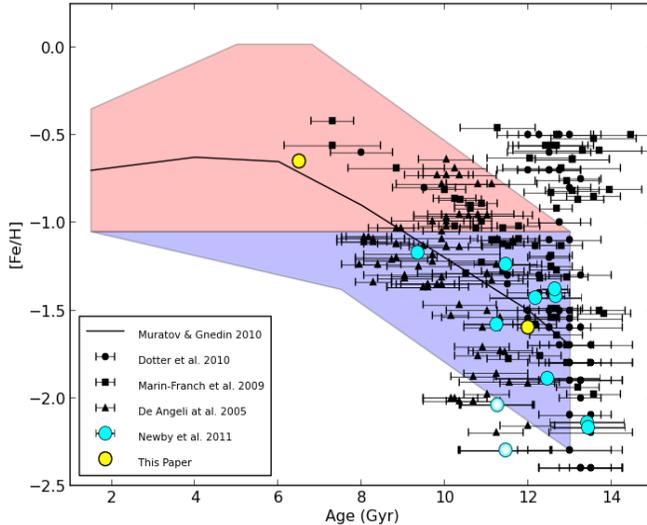


Figure 5: Here Figure 17 from Newby et al. (2011) is reproduced, with permission, with the addition of our new clusters. The plot of age versus metallicity for Milky Way globular clusters includes different sources, where clusters studied by Newby et al. (2011) are denoted by blue circles and the two clusters from this study indicated by yellow circles. The mean value of the theoretical AMR from Muratov and Gnedin (2010) is indicated by the solid line. The shaded areas show the spread of metallicities for a given age as shown in Muratov & Gnedin (2010), where blue and red correspond to old and young globular cluster populations, respectively.

Clusters that lie along the halo cluster AMR, a theoretical model for which was recently presented in Muratov & Gnedin (2010), have remarkably similar turnoff star properties. As a cluster ages, the turnoff color becomes fainter and redder as brighter, bluer stars evolve off of the main sequence. Metal-poor stars, however, are bluer and brighter than analogous metal-rich stars. Newby et al. (2011) remarked that along the average halo cluster AMR, these competing properties almost entirely canceled each other out, resulting in intrinsically similar turnoff colors and absolute magnitudes for old halo globular clusters. We see from Figure 5 that our clusters are surprisingly consistent with the halo cluster AMR.

The plot of age versus $[\text{Fe}/\text{H}]$ metallicity in Figure 5 includes the mean metallicity at a range of ages [27] plotted as a solid line, and we find that Palomar 13 is close to the AMR with an age of 12 Gyr and $[\text{Fe}/\text{H}]=1.7$ in the metal-poor region [19] and Whiting 1, with $[\text{Fe}/\text{H}]=-0.65$ and an age of 6.5 Gyr [21], is also close to the solid line, but in the metal-rich region.

We plot the Padova isochrones of clusters in this study in Figure 6 against theoretical isochrones for similar age and metallicity values along the average cluster AMR. The filled circles correspond to the isochrone for

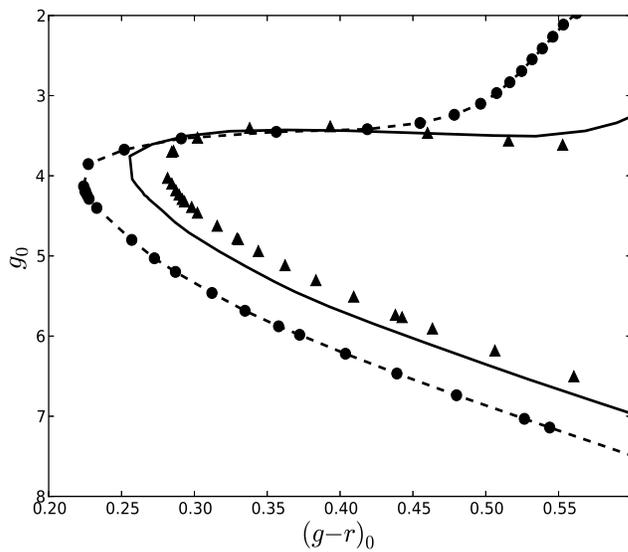


Figure 6: We plot the Padova isochrones from Figure 2 for Palomar 13 and Whiting 1 in addition to two theoretical isochrones chosen from Figure 18 of Newby et al. (2011) to compare the turnoff locations of combinations of age and metallicity. The black triangles correspond to the isochrone for Whiting 1 with an age of 6.5 Gyrs and a metallicity $[\text{Fe}/\text{H}]$ of -0.65. The dashed line overlapped by filled circles represents a cluster 12 Gyrs old with a metallicity $[\text{Fe}/\text{H}]$ of -1.7, which are the estimates for Palomar 13. The solid line represents a cluster 8 Gyrs old with a metallicity of -0.9. From left to right, the age decreases and metallicity increases.

Palomar 13, which lies almost exactly on the average AMR line and so coincides exactly with the dashed black isochrone, which is a hypothetical cluster on the AMR with $[\text{Fe}/\text{H}] = -1.7$ and an age of 12 Gyr. The triangles correspond to the isochrone created for Whiting 1. The central solid line is an isochrone for a cluster 8 Gyr and $[\text{Fe}/\text{H}] = -0.9$, which is the oldest and most metal-poor cluster analyzed by Newby et al. (2011). Though the ages range from 12 Gyrs to 6.5 Gyrs, the color and magnitude of the FTO stars are similar, hardly differing in turnoff magnitude, and the color differs by less than 0.1 in $(g-r)_0$. As the turnoff colors approach the edge of the color selection box for FTO stars, Whiting 1 may represent the very youngest of clusters for which the Newby et al. (2011) method can be applied.

6 Conclusions

In this paper, we analyze two globular clusters in the Milky Way halo using the techniques of Newby et al. (2011). We create a histogram of the absolute magnitudes for each cluster. We fit a double-sided Gaussian to each histogram and compared it to the results for parameters reported by Newby et al. (2011).

For Palomar 13, the absolute magnitude histogram fit parameters μ , σ_ℓ , and σ_r are 4.46 ± 0.26 , 0.81 ± 0.025 , and 0.2226 ± 0.23 , respectively. Similarly, the magnitude histogram fit parameters for Whiting 1 are $\mu =$

4.11 ± 0.3 , $\sigma_\ell = 0.57 \pm 0.25$, and $\sigma_r = 0.86 \pm 0.27$. The Whiting 1 parameters agree with the expectations from Newby et al. (2011). The Palomar 13 results are marginally different, at the two and three sigma level.

We believe the smaller number of faint stars observed in Palomar 13 is due to larger seeing in the SDSS image for that cluster, and not due to an intrinsic difference in the absolute magnitude distribution of turnoff stars in that cluster.

The similarity of the Whiting 1 results allows us to extend the region of applicability of the Newby et al. (2011) results to halo globular clusters as young as 7 Gyr, metallicities as high as $[\text{Fe}/\text{H}] = -0.7$, and to distances as high as 30 kpc.

We support the conclusion of [25] that Palomar 13 is not a member of the Sagittarius dwarf tidal stream, while Whiting 1 is within the Sagittarius dwarf tidal stream.

7 Acknowledgements

We thank the anonymous reviewer for insightful feedback to improve the clarity of the paper.

This research is based upon work supported by the National Science Foundation under Grant No. AST 10-09670.

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, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, 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, Max Planck Institute for Extraterrestrial Physics, 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.

References

- [1] York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, *AJ*, 120, 1579

- [2] Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525
- [3] Newby, M., Newberg, H. J., Simones, J., Cole, N., & Monaco, M. 2011, *ApJ*, 743, 187
- [4] Piotto, G., Milone, A. P., Anderson, J., et al. 2012, arXiv:1208.1873
- [5] Newberg, H. J., Yanny, B., Rockosi, C., et al. 2002, *ApJ*, 569, 245
- [6] Yanny, B., Newberg, H. J., Kent, S., et al. 2000, *ApJ*, 540, 825
- [7] Newberg, H. J., & Yanny, B. 2006, *Journal of Physics Conference Series*, 47, 195
- [8] Yanny, B., Newberg, H. J., Grebel, E. K., et al. 2003, *ApJ*, 588, 824
- [9] Yanny, B., Newberg, H. J., Johnson, J. A., et al. 2009, *ApJ*, 700, 1282
- [10] Newberg, H. J., Willett, B. A., Yanny, B., & Xu, Y. 2010, *ApJ*, 711, 32
- [11] Willett, B. A., Newberg, H. J., Zhang, H., Yanny, B., & Beers, T. C. 2009, *ApJ*, 697, 207
- [12] Cole, N., Newberg, H. J., Magdon-Ismail, M., et al. 2008, *ApJ*, 683, 750
- [13] De Angeli, F., Piotto, G., Cassisi, S., et al. 2005, *AJ*, 130, 116
- [14] Marín-Franch, A., Aparicio, A., Piotto, G., et al. 2009, *ApJ*, 694, 1498
- [15] Dotter, A., Sarajedini, A., & Anderson, J. 2011, *ApJ*, 738, 74
- [16] Aihara, H., Allende Prieto, C., An, D., et al. 2011, *ApJS*, 193, 29
- [17] Harris, W. E. 1996, *AJ*, 112, 1487
- [18] Côté, P., Djorgovski, S. G., Meylan, G., Castro, S., & McCarthy, J. K. 2002, *ApJ*, 574, 783
- [19] Bradford, J. D., Geha, M., Muñoz, R. R., et al. 2011, *ApJ*, 743, 167
- [20] Whiting, A. B., Hau, G. K. T., & Irwin, M. 2002, *ApJS*, 141, 123
- [21] Carraro, G., Zinn, R., & Moni Bidin, C. 2007, *A&A*, 466, 181
- [22] Marigo, P., Girardi, L., Bressan, A., et al. 2008, *A&A*, 482, 883
- [23] Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. 2000, *A&AS*, 141, 371
- [24] An, D., Pinsonneault, M. H., Masseron, T., et al. 2009, *ApJ*, 700, 523
- [25] Law, D. R., & Majewski, S. R. 2010, *ApJ*, 718, 1128

- [26] Siegel, M. H., Majewski, S. R., Cudworth, K. M., & Takamiya, M. 2001, *AJ*, 121, 935
- [27] Muratov, A. L., & Gnedin, O. Y. 2010, *ApJ*, 718, 1266

Particular Solution to a Time-Fractional Heat Equation

Simon P. Kelow ^{*} (SPK39@nau.edu)

Kevin M. Hayden [†] (Kevin.Hayden@nau.edu)

July 21, 2013

Abstract

When the derivative of a function is non-integer order, e.g. the $1/2$ derivative, one ventures into the subject of fractional calculus. The time-fractional heat equation is a generalization of the standard heat equation as it uses an arbitrary derivative order close to 1 for the time derivative. We present a particular solution to an initial-boundary-value time-fractional heat equation problem and compare the properties of the solutions when the time derivative order is varied. Orders of the time-fractional heat equation which return solutions that display physically impossible characteristics are also considered. One observed property is slightly less exponential decay of heat for the solutions of non-integer order time derivatives greater than one. Another property is solutions with non-integer order time derivatives less than one exhibit slightly greater exponential decay. Both of these comparisons are made with respect to the standard heat equation solution, where the order of the time derivative is one.

Keywords: Fractional Calculus; Time-Fractional Heat Equation; Time-Fractional Diffusion Equation.

^{*}Department of Astronomy and Physics, Northern Arizona University, Flagstaff, AZ

[†]Faculty Advisor, Department of Mathematics and Statistics, Northern Arizona University, Flagstaff, AZ

1 Introduction

In the last few decades, fractional differential equations have been found to be successful models of real life phenomenon (see e.g. [1, 2, 4] and their contained references). Motivated by the quote, "... we may say that Nature works with fractional time derivatives." [7], we investigate the initial-boundary-value problem (IBVP) for a time-fractional heat equation. Ultimately we seek to graph the solutions for when the standard heat equation, $D_t u = D_x^2 u$, is generalized to the time-fractional heat equation, $D_t^\alpha u = D_x^2 u$, for several values of α .

We proceed by first providing the overview for the IBVP, and solving it with the standard heat equation. Next, we solve the same IBVP with the generalized time-fractional heat equation, and solutions for multiple derivative orders are graphed for comparison with the standard solution. Finally, the solutions displaying physically unrealistic properties are plotted.

2 IBVP Heat Equation Solution

The standard heat equation is $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$. For this equation the temperature is represented by u , which is a function of time, t , and space, x . To make comparison easier, we use the derivative operator D to write the heat equation with the following notation:

$$D_t u = D_x^2 u. \tag{1}$$

Consider an ideal one-dimensional rod of length L with the boundary conditions

$$u(t, 0) = u(t, L) = 0 \tag{2}$$

for $t \geq 0$ and an initial condition with a parabolic profile of heat distribution given by

$$u(0, x) = -\frac{4a}{L^2}x^2 + \frac{4a}{L}x \tag{3}$$

shown in Figure 1. Note the initial condition satisfies the boundary conditions and has a maximum temperature of $u(0, L/2) = a$.

The solution is then assumed to be of the form $u(t, x) = T(t)X(x)$. Plugging this into

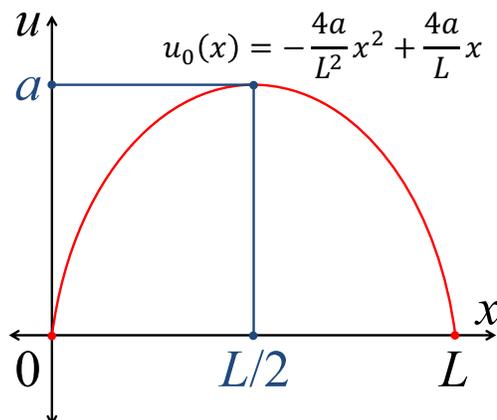


Figure 1: *Initial heat distribution along the rod.*

equation (1) yields $T'(t)X(x) = T(t)X''(x)$. Since this equation is separable, there exists a constant of separation, $C \in \mathbb{R}$, such that

$$\frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)} = C. \quad (4)$$

To obtain a non-trivial, real solution, assume $C < 0$. This will also satisfy the property of temperature decay to the defined boundary conditions. Thus we let $C = -\lambda^2$ for $\lambda \in \mathbb{R}$ and we must solve the ordinary differential equations (ODEs):

$$\begin{cases} X''(x) = -\lambda^2 X(x) \\ T'(t) = -\lambda^2 T(t). \end{cases} \quad (5)$$

These ODEs have the solutions $X(x) = C_1 \cos(\lambda x) + C_2 \sin(\lambda x)$ and $T(t) = C_3 e^{-\lambda^2 t}$. Reassigning the constants leads to the general solution to the heat equation (1)

$$u(t, x) = C_1 \cos(\lambda x) e^{-\lambda^2 t} + C_2 \sin(\lambda x) e^{-\lambda^2 t}. \quad (6)$$

Moving onto the particular solution, we start by satisfying the boundary conditions (2). By inspection, the condition $u(t, 0) = 0$ requires $C_1 = 0$. The only non-trivial way to satisfy $u(t, L) = 0$ is by solving $\sin(\lambda L) = 0$. Thus

$$\lambda_n = \frac{n\pi}{L}, \quad (7)$$

for $n \in \mathbb{N}$, yields infinitely many solutions to the heat equation. The sum of such solutions is given by the Fourier series solution

$$u(t, x) = \sum_{n=0}^{\infty} C_n \sin(\lambda_n x) e^{-\lambda_n^2 t}. \quad (8)$$

The initial condition (3) is used to determine the Fourier coefficients (C_n) given by

$$C_n = \frac{2}{L} \int_0^L \left(-\frac{4a}{L^2} x^2 + \frac{4a}{L} x \right) \sin(\lambda_n x) dx. \quad (9)$$

Integrating and using the properties of sine and cosine will yield

$$C_n = \begin{cases} 0, & n \text{ even} \\ \frac{32a}{n^3 \pi^3}, & n \text{ odd.} \end{cases} \quad (10)$$

Ignoring the zero-valued even-indexed terms in equation (8), we have the particular solution to the IBVP heat equation (1-3):

$$u(t, x) = \sum_{n \in \mathbb{N}, \text{ odd}} \frac{32a}{n^3 \pi^3} \sin(\lambda_n x) e^{-\lambda_n^2 t}, \quad (11)$$

where we now have

$$\lambda_n = \frac{n\pi}{L}. \quad (12)$$

The main properties of this solution do not change for differing values of maximum initial heat a and rod length L . Using $a = L = 1$, the particular solution can be seen in Figure 2.

3 IBVP Time-Fractional Heat Equation Solution

The time-fractional heat equation generalizes the time derivative of the standard heat equation. Consider the equation

$$D_t^\alpha u = D_x^2 u \quad (13)$$

where $\alpha \in [1 - \delta, 1 + \delta] \subset \mathbb{R}$ and $\delta > 0$ is relatively small. From Luchko [5], uniqueness and existence are proven with similar general conditions to what we are considering. However,

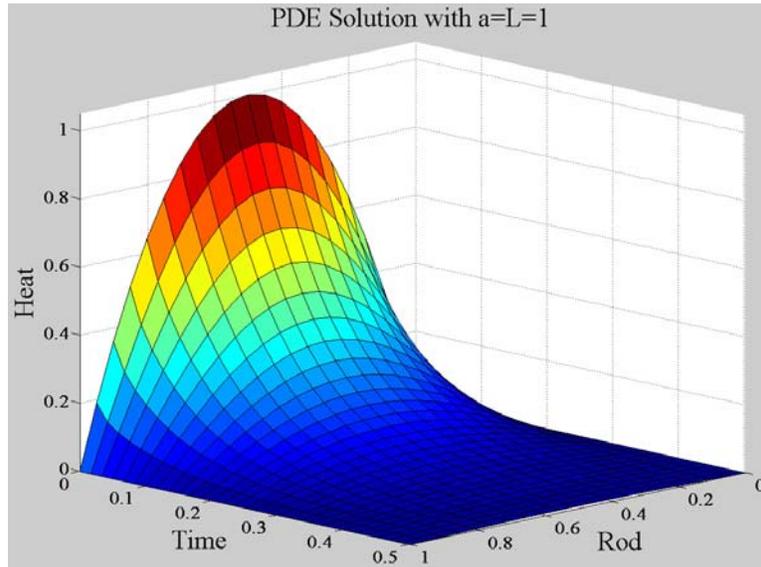


Figure 2: A graph of the particular solution to the IBVP heat equation with $a = L = 1$.

we emphasize that we seek to visualize what the differences in the solutions for equation (13) when α is varied.

Consider the same boundary and initial conditions (2-3) as well as a solution of the form $u(t, x) = T(t)X(x)$. Then the procedure of finding the solution does not change until we solve the differential equations (DEs)

$$\begin{cases} X''(x) = -\lambda^2 X(x) \\ D_t^\alpha T(t) = -\lambda^2 T(t), \end{cases} \quad (14)$$

which are analogous to equations (5). From Miller and Ross [3] or Luchko [5], the solution to the fractional differential equation (FDE) $D_t^\alpha T(t) = -\lambda^2 T(t)$ is the Mittag-Leffler function:

$$T(t) = \sum_{k=0}^{\infty} \frac{(-\lambda^2 t^\alpha)^k}{\Gamma(\alpha k + 1)}, \quad (15)$$

where Γ is the standard gamma function.

The general solution can be written at this point, but note from Section 2 all the unknowns of the IBVP are determined by the X portion of the equation. Therefore only the exponential

part in solution (11) changes (to the Mittag-Leffler function), and we have the particular solution

$$u(t, x) = \sum_{n \in \mathbb{N}, \text{ odd}} \left[\frac{32a}{n^3 \pi^3} \sin(\lambda_n x) \sum_{k=0}^{\infty} \frac{(-\lambda_n^2 t^\alpha)^k}{\Gamma(\alpha k + 1)} \right], \quad (16)$$

where

$$\lambda_n = \frac{n\pi}{L}. \quad (17)$$

4 Solution Comparison

Now that we have the particular solution to the time-fractional heat equation (16), a plot comparing its solutions for various α values is desired. Unfortunately the Mittag-Leffler function part of the solution, equation (15), converges slowly. The convergence of this portion of our solution is slow enough that it computationally diverges for almost any $t > 0$. In lieu of using high precision numerical calculations, a “less rigorous” solution was found.

The pattern of the standard derivative of the exponential function $D_t^n [e^{rt}] = r^n e^{rt}$ where $n \in \mathbb{N}$ can be generalized to the fractional derivative as

$$D_t^\alpha [e^{rt}] = r^\alpha e^{rt}, \quad (18)$$

where $\alpha \in \mathbb{R}$, see [6] and contained references. Using this derivative and recalling the FDE is $D_t^\alpha T(t) = -\lambda_n^2 T(t)$, it can be seen that

$$T(t) = e^{\sqrt[\alpha]{-\lambda_n^2} t} \quad (19)$$

is a solution. By substituting solution (19) into equation (16), we have a new solution which can be numerically computed. Our alternate solution is

$$u(t, x) = \sum_{n \in \mathbb{N}, \text{ odd}} \frac{32a}{n^3 \pi^3} \sin(\lambda_n x) e^{\sqrt[\alpha]{-\lambda_n^2} t}, \quad (20)$$

where λ_n is as in equation (17).

Note the potential problem of a negative number inside an even indexed root in the expression $e^{\sqrt[\alpha]{-\lambda_n^2} t}$. This leads to a potentially imaginary and non-unique solution. Since we are

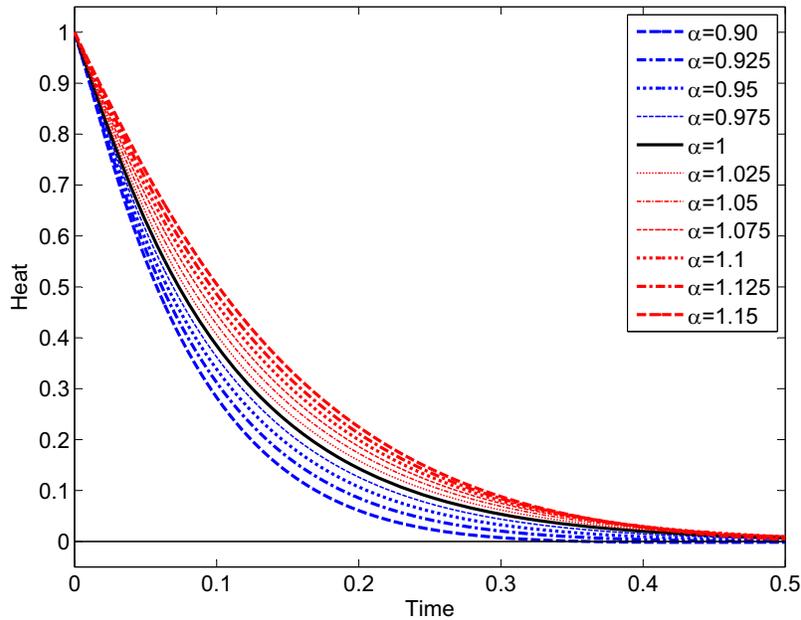


Figure 3: *Heat decay from the IBVP time-fractional heat equation at $x = L/2$ with $a = L = 1$.*

mainly interested in comparing the graphs of the solutions for multiple α values, we ignored the non-uniqueness property by only evaluating the real part of this expression. We were then able to create Figure 3, which is the heat decay for the rod at $x = L/2$ for several α values. All of the plotted solutions in Figure 3 are “physical realistic,” in that they satisfy the boundary conditions, are free of negative temperature values, and don’t display an increase in heat of the rod in the absence of any external heating factors.

In addition to the physical solutions plotted in Figure 3, there are several solutions outside of this range which provide physically unrealistic solutions. Figure 4 shows a plot of these various solutions.

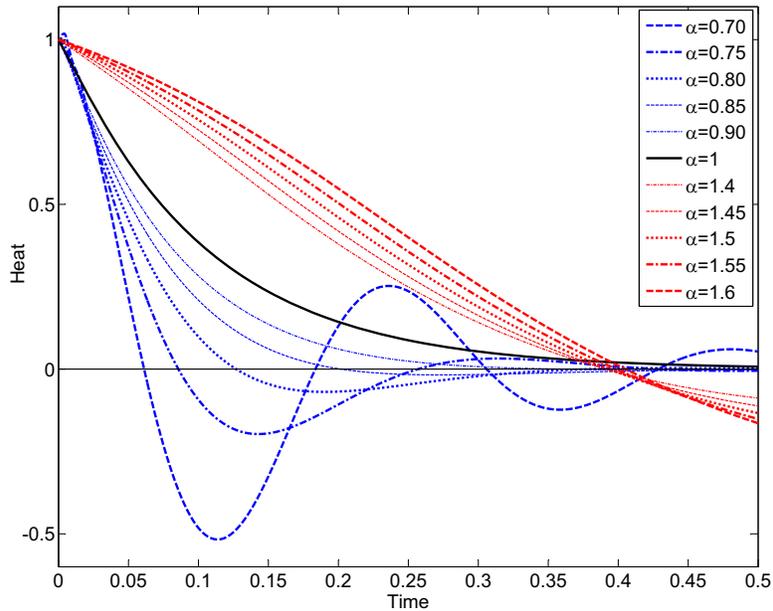


Figure 4: *Solutions for heat decay from the IBVP time-fractional heat equation at $x = L/2$ with $a = L = 1$ which display physically unrealistic properties.*

5 Conclusions and Future Work

From Figure 3, there are several important properties to be seen:

- A continuous change in α appears to yield a continuous change in the heat decay profile of u .
- For the values $\alpha > 1$, there is slightly less exponential decay in heat compared to the heat decay for $\alpha = 1$.
- For the values $\alpha < 1$, there is slightly greater exponential decay in heat compared to the heat decay for $\alpha = 1$.
- All solutions sufficiently close to α satisfy the boundary conditions, and display physically realistic properties.

Therefore we conclude that the time-fractional heat equation is a physically legitimate generalization of the standard heat equation that might be used for values $\alpha \approx 1$. As expected

however, when the value of α strays too far from $\alpha = 1$ the time-fractional heat equation will exhibit properties which are no longer realistic for describing the heat decay of an object.

Ideas for future work:

- Work with equation (16) to get the Mittag-Leffler portion of the solution to numerically converge. Then one can compare the unique solution (16) and confirm that the solution (20) is a viable alternative to work with.
- Do equations (16) or (20) solve the wave equation when $\alpha = 2$? If so, what are the IVBP conditions for the wave equation to have one of these as a solution? That is to say, can the heat equation “morph” into the wave equation through the time-fractional generalization?
- Solve the FDE $D_t^\alpha u = kD_x^\beta u$. The solution of this can be used to generalize the heat equation, the Laplace’s equation, and the wave equation for appropriate values of α , β and k (near there classical values) and initial/boundary conditions.
- Does a fractional heat equation solution compare to empirical data with better accuracy/precision than the standard heat equation solution under certain conditions or with certain materials?

6 Acknowledgements

This project received funding from the 2011-2012 Arizona/NASA Space Grant, an undergraduate research internship program. Also, a thank you to Kevin Hayden (my mentor/advisor) for all your time and help.

References

- [1] R. Hilfer, *Applications of Fractional Calculus in Physics*, World Scientific, Singapore, 2000
- [2] A.A. Kilbas, H.M. Srivastava, and J.J. Trujillo, *Theory and Applications of Fractional Differential Equations*, Elsevier, Amsterdam, 2006
- [3] K.S. Miller, and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, A Wiley-Interscience Publication, 1993

- [4] I. Podlubny, *Fractional Differential Equations*, Mathematics in Science and Engineering, vol. 198, 1999
- [5] Y. Luchko, *Some uniqueness and existence results for the initial-boundary-value problems for the generalized time-fractional diffusion equation*, Computers and Mathematics with Applications, vol. 59, 2010, 1766-1772
- [6] M. Kleinz, and T. Osler, *A Child's Garden of Fractional Derivatives*, Mathematical Association of America, vol. 31, Number 2, pp. 82-89, 2000
- [7] S. Westerlund, *Dead matter has memory!*, Physica Scripta, vol. 43, 1991, 174-179

Effect of Temperature on Pulse Wave Velocity and Arterial Compliance

Tucker Frawley^a and T. Brian Bunton^a

Dept. of Chemistry and Physics^a, Coastal Carolina University, Conway, SC 29528

The compliance of a material is the ability of the material to expand or contract. A low arterial compliance will lead to high blood pressure and, ultimately, cardiovascular disease. The velocity of the pulse wave traveling through an artery is a function of the compliance of the artery that it is traveling through. Therefore, the compliance of an artery can be indirectly observed by measuring several pulse wave velocities in the artery. This experiment investigates the effect, if any, of temperature on arterial compliance. Two electrocardiographs (ECGs) were taken from healthy human subjects at the right wrist to right elbow, before and after an application of a temperature gradient. For one study, the subject's arm was put into water of a significantly higher temperature than the body. For the other study, the subject's arm was put into water of a significantly lower temperature than the body. From the ECGs taken, the relative pulse wave velocities before and after the applications of the local temperature gradients were determined. It was determined that the local heating resulted in an increase in the pulse wave velocity in the artery, and the application of the local cooling resulted in a decrease in the pulse wave velocity in the artery.

Keywords: compliance, pulse wave velocity, cardiovascular, thermoregulation

1. Background and Introduction

According to estimates taken in 2006, over 81.1 million people in the United States suffer from at least one form of cardiovascular disease. Of these people, 76.4 million of them have high blood pressure¹. One of the underlying sources of high blood pressure, and of many cardiovascular diseases as well, is low arterial compliance². The compliance of a blood vessel is defined as the ability of the vessel to expand, or contract, to best accommodate a particular volume and a particular hydrostatic pressure of blood³. In

more general terms, arterial compliance is the elasticity of an artery. One way to observe arterial compliance is by observing the velocity of the pulse wave moving through an artery. The pulse wave is the increase of pressure radiating through the arteries that occurs with each contraction of the left ventricle of the heart⁴. Because the pulse wave velocity is a function of arterial compliance, a relative arterial compliance can be determined by recording several pulse wave velocities through that artery. This allows one to identify individuals who have a low arterial compliance and are susceptible to

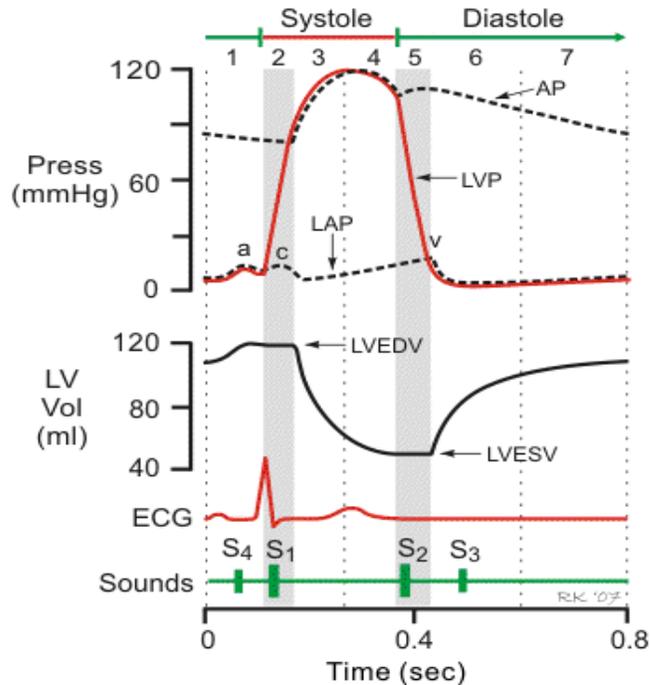


Figure 1. Changes in pressure (P) and volume (V) of left ventricle (LV, solid lines) and aorta (LA, dotted lines) during cardiac cycle⁶. The left ventricle's volume varies between its end-diastolic volume (EDV) to its end-systolic volume (ESV) during one cycle.

hypertension and other cardiovascular diseases.

The cardiovascular system is made up of two parts, the pulmonary circulation portion and the systemic circulation portion. This experiment focuses on the systemic portion of the cardiovascular cycle. The systemic portion circulates oxygen-rich blood throughout the body. It does this by transferring the oxygenated blood in the left atrium to the left ventricle, where the blood is then pumped through the aorta into systemic arteries. The blood is pumped into these arteries by the contraction of the left ventricle⁵. This phase of the cardiac cycle, known as systole, increases the pressure inside the ventricle, which ejects the blood through the aorta to the lower pressure arteries. The blood flow is driven through the arteries by this transient movement of pressure called the pulse wave. When all of the blood is ejected from the left ventricle, the

aortic valve closes, and the systolic phase of the cardiac cycle ends. The left ventricle is then refilled with oxygenated blood from the left atrium, lowering its pressure in the process. This phase of the cardiac cycle, characterized by low blood pressure, is known as the diastole⁶. The change in volume and pressure of the left ventricle during the cardiac cycle can be seen in Figure 1.

During the pulse wave's travel through the systemic arteries, it will encounter a pulse wave reflective site at some point. An example of one of these sites is a major branching point such as the renal artery and femoral artery junction. After encountering this reflective site, the pulse wave is reflected in the opposite direction so that it returns to the heart. Under normal circumstances, the pulse wave returns to the heart after the aorta has closed⁷. This is during the diastolic phase of the cardiac cycle, so diastolic pressure is increased. Because of constructive

interference, the magnitude of the diastole is thus increased⁸. However, the pulse wave may return to the heart before the aorta has closed, during the systolic phase of the cardiac cycle⁷. This increases systolic pressure and consequently increases the magnitude of the systole.

The timing of the return of this pulse wave to the heart is determined by the pulse wave's velocity, which is a function of arterial compliance⁸. This pulse wave velocity (PWV) can be determined by dividing the distance traveled in an artery (Δx) by the time it takes for the wave to travel that distance, or the pulse transit time (PTT)⁹.

$$PWV = \frac{\Delta x}{PTT} \quad (1)$$

The pulse wave velocity can also be calculated by using Eq. (2), the Bramwell-Hill equation, which relates the pulse wave velocity to pressure and volume of the artery. It does this using the change in pressure of the artery (ΔP), the volume (V) and change of volume of the artery (ΔV), and the density of the blood (ρ)⁹.

$$PWV = \sqrt{\frac{V \cdot \Delta P}{\rho \cdot \Delta V}} \quad (2)$$

The speed at which the wave travels in the arteries is heavily based on the elasticity of the arteries⁸. Arterial compliance (C) is the change in arterial blood volume (ΔV) due to a given change in arterial blood pressure (ΔP)¹⁰.

$$C = \frac{\Delta V}{\Delta P} \quad (3)$$

Eq. (3) can be inserted into Eq. (2), showing that the pulse wave velocity and the square root of the compliance of an artery are inversely related¹⁰.

$$PWV = \sqrt{\frac{V}{\rho \cdot C}} \quad (4)$$

An artery with a high compliance will stretch out to allow the pulse wave to pass. This stretching of the artery causes the pulse wave to travel slower than it would without the stretching. In contrast, an artery with low compliance and high arterial stiffness will not stretch out when the pulse wave passes. Therefore, because the square root of arterial compliance and pulse wave velocity are inversely related, the pulse wave will travel faster through arteries with a low compliance⁸. At optimum performance, the arteries' compliance will allow the pulse wave to return to the heart at such a time that diastolic pressure will increase. However, a relatively low compliance will make the pulse wave travel faster and return to the heart more quickly. The pulse wave will return during the systolic phase of the cardiac cycle, increasing the systolic pressure⁷. This will result in higher cardiac work for this low arterial compliance. This indicates that a high arterial compliance is much more beneficial than a low arterial compliance¹¹.

Much research has already been done studying the effect of factors on long-term behavior of arterial compliance. A well-known factor affecting arterial compliance is aging. Studies have shown that as an individual gets older, arteries lose their compliance and become much stiffer. With this stiffening, the arteries lose their ability to distend in a response to increased pressure. This leads to an increased blood pressure as individuals age, resulting in an eventual increase in cardiovascular risk¹². The stiffening of the arteries can be attributed to the increase in the number of collagen fibers surrounding the arteries in elder individuals. Collagen fibers lose their elasticity as age increases and stiffen up around the arteries. This hinders the arteries' ability to distend when needed¹³.

Another factor affecting arterial compliance that has been extensively studied is physical exercise. In a study done

determining the effect of physical activity on arterial compliance, Tanaka and colleagues found that aerobic exercise training can prevent the stiffening of arteries¹³. The study showed that this regular physical exercise could not increase arterial compliance, but it would slow the normal loss of compliance developed from aging¹².

2. Theory

This experiment seeks to observe the effect of temperature on pulse wave velocities, indirectly determining its effect, if any, on arterial compliance. Extensive previous research has been done investigating the effect of temperature on the human body and its thermoregulation that would suggest that an increase in the temperature of the body will cause an increase in arterial cross-section in that area of the body¹⁴. The human body has an extensive positive/negative feedback system that allows it to regulate its body temperature¹⁵. In all warm-blooded animals, including humans, vasodilation occurs in response to an increase in the temperature of the ambient environment⁶. Vasodilation is increase in the internal diameter of a blood vessel that results from relaxation of smooth muscle within the wall of the vessel, thus causing an increase in blood flow¹⁶. This vasodilation occurs in an effort to divert the flow of heated blood to the surface of the skin. The vasodilation is a function of homeostasis and will lead to a decrease in the temperature of the surface of the body⁶.

Using the Bramwell-Hill equation, Eq. (4), the pulse wave velocity can be calculated in terms of the volume of an artery (V), the compliance of the artery (C), and the density of the blood flowing through the artery (ρ). The volume of the artery can be replaced by the cross-sectional area of the artery (A) to obtain a pulse wave velocity factor (PWV factor)¹⁷.

$$PWV \text{ factor} = \sqrt{\frac{A}{\rho \cdot C}} \quad (5)$$

An increase in temperature should increase the diameter of an artery which would, consequently, increase the cross-sectional area of the artery. Likewise, a decrease in temperature should decrease the diameter of an artery. This would decrease the cross-sectional area of the artery.

The pulse wave velocity can also be calculated in terms of the distance traveled by the pulse wave (Δx) and the pulse transit time (PTT), as defined in Eq. (1). A new equation can be obtained by rearranging Eq. (1) to solve for the distance traveled by the pulse wave (Δx).

$$\Delta x = PWV \cdot PTT \quad (6)$$

We can use this equation together with Eq. (5) to eliminate the distance traveled by the pulse wave. We assume no change in the arterial length as a consequence of the applied temperature gradient; further study to support such an assumption is encouraged. We are now able to compare the cross sectional area (A_i), arterial compliance (C_i), and pulse transit time (PTT_i) before an event to the cross sectional area (A_f), arterial compliance (C_f), and pulse transit time (PTT_f) after the event.

$$\sqrt{\frac{A_i}{\rho \cdot C_i}} \cdot PTT_i = \sqrt{\frac{A_f}{\rho \cdot C_f}} \cdot PTT_f \quad (7)$$

This equation can then be solved for the pulse transit time after an event (PTT_f).

$$PTT_f = \sqrt{\frac{C_f / C_i}{A_f / A_i}} \cdot PTT_i \quad (8)$$

From Eq. (8), it can be determined that if the pulse transit time after an event (PTT_f) is greater than the pulse transit time before an event (PTT_i), then the value under the square root bracket must be greater than one. In order

for this to happen, the cross-sectional area of the artery must increase proportionally less than the arterial compliance increases. Conversely, if the pulse transit time after an event (PTT_f) is less than the pulse transit time before an event (PTT_i), then the value under the square root bracket must be between zero and one. In order for this to happen, the cross-sectional area of the artery must decrease proportionally less than the arterial compliance decreases.

For this experiment, the mitigating event was the application of local heating and cooling to subjects' arms. If the pulse transit time after the application of local heating (PTT_f) is greater than the pulse transit time before (PTT_i), then the arterial compliance must have increased. However, if PTT_f is less than PTT_i , then the application's effect on arterial compliance cannot be determined. This would require either the cross sectional area to increase due to the event, the arterial compliance to decrease due to the event, or both. Because previous research has shown that the cross sectional area of an artery will increase due to an application of heat with a higher temperature relative to the body¹⁵, the arterial compliance cannot be determined in this case. For the application of local cooling, a positive ΔPTT (defined as $PTT_f - PTT_i$) would mean an effect on arterial compliance cannot be determined, while a negative value would denote a decrease in compliance.

3. Experiment

Data collection took place over the span of one year at Coastal Carolina University. Data was taken from thirty total subjects. These subjects included males and females ages 18 to 50 and were in excellent health. They volunteered to have their data collected for no benefits. The materials used for the collection of data included a laptop with LoggerPro software, a Vernier LabPro sensor, a Vernier ECG sensor, and Kendall Q-Trace

resting ECG electrodes. For the local heating, a hot whirlpool (set to 110°F) was used. For the local cooling, a cold whirlpool (set to 53°F) was used.

Experiments were begun by placing electrodes on the right wrist, just below the right elbow, and just below the left elbow of the sitting subject. For each experiment, the positive lead was placed on the electrode on the right wrist, the negative electrode was placed on the electrode just below the right elbow, and the reference lead was placed on the electrode just below the left elbow. This set-up recorded measurements in the radial artery. After all three leads were attached, the subject was asked to relax, and data was collected in LoggerPro for a time interval of ten seconds. The leads and electrodes were then removed from the subject, and the subject was given the application of local heating or cooling. The subjects dipped their right arms up to just above the elbow in the whirlpool. The subjects were asked to keep their arm in the whirlpool for two minutes. The three electrodes were then replaced. With the leads attached, the subject was again asked to relax and data was collected in LoggerPro for ten seconds.

After the completion of data collection, the ECG graphs taken from each subject were then analyzed in LoggerPro. From the graphs, the pulse transit times for each experiment (before and after the local heating or cooling) were determined. This was done by taking the time elapsed from the peak of the R wave to the peak of the T wave on the ECG. The peaks of these waves can be seen in Figure 2. A sample of data collected from the experiment can be seen in Figure 3.

For each subject, five to ten pulse transit times before and after the local heating or cooling were measured. Average PTT's before and after were then determined for each subject. Finally, a ΔPTT was determined for each subject.

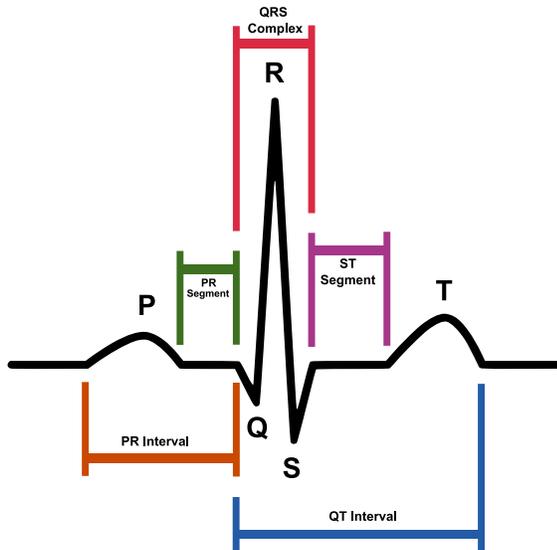


Figure 2. A schematic representation of a normal ECG¹⁸.

4. Results

Thirty total subjects were tested. Twenty-three subjects had their data taken with local heating. However, only fifteen of the twenty-three sets of data collected were deemed functional and were used for analysis. The remaining eight samples contained faulty data and were subsequently not used. Some of the data contained very large amplitudes compared to the functional data used. Other faulty data did not resemble an ECG at all and contained no visible waves or peaks. Possible reasons for the faulty data could be attributed to the possibility of an equipment malfunction, recent physical activity of a subject, or possibly human error in the set-up of the equipment. Fifteen subjects had their data taken after local cooling. All fifteen samples taken were deemed functional and were used for analysis.

Of the data that was used for the local heating, the results were consistent with each other. From the ECGs, pulse transit times before the application ranged from 180 ms to 240 ms and averaged 209 ms. Pulse transit times after the application ranged from 210 ms to 300 ms and averaged 231 ms. Fourteen

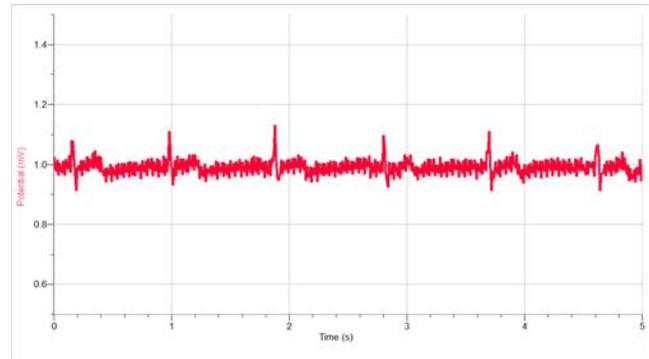


Figure 3. Sample of ECG data collected in LoggerPro.

of the fifteen samples of data resulted in a positive value of ΔPTT , and one sample had no change in the pulse transit times before and after the local heating. The values of ΔPTT ranged from 0 ms to 42 ms and averaged a difference of 22 ms (Table 1).

Of the data that was used for local cooling, the results were again consistent with each other. From the ECGs, pulse transit times before the application ranged from 210 ms to 240 ms and averaged 222 ms. Pulse transit times after the application ranged from 190 ms to 250 ms and averaged 210 ms. Thirteen of the fifteen samples of data resulted in a negative value of ΔPTT , and two samples of data resulted in a positive value of ΔPTT . The values of ΔPTT ranged from -30 ms to 44 ms and averaged a difference of negative -11 ms. The data for each subject can be seen in Table 2.

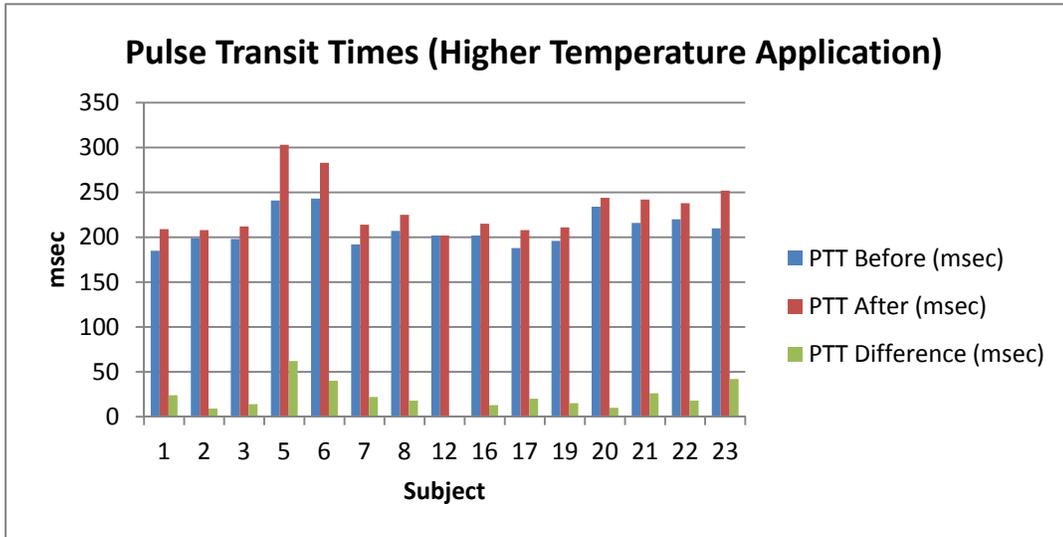


Figure 4. Graph of pulse transit time averages and differences for local heating.

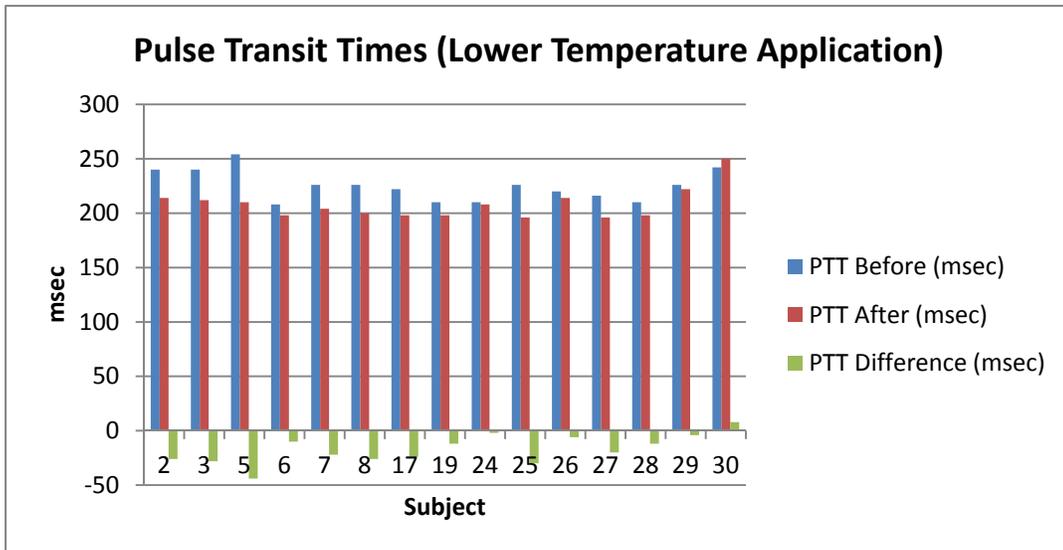


Figure 5. Graph of pulse transit time averages and differences for local cooling.

Subject	PTT Before (msec)	PTT After (msec)	PTT Difference (msec)
1	185	209	24
2	199	208	9
3	198	212	14
5	241	303	62
6	243	283	40
7	192	214	22
8	207	225	18
12	202	202	0
16	202	215	13
17	188	208	20
19	196	211	15
20	234	244	10
21	216	242	26
22	220	238	18
23	210	252	42

Table 1: Data collected for each subject given local heating.

Subjec	PTT Before	PTT After	PTT Difference
2	240	214	-26
3	240	212	-28
5	210	254	-44
6	208	198	-10
7	226	204	-22
8	226	200	-26
17	222	198	-24
19	210	198	-12
24	210	208	-2
25	226	196	-30
26	220	214	-6
27	216	196	-20
28	210	198	-12
29	226	222	-4
30	242	250	8

Table 2: Data collected for each subject given local cooling.

5. Conclusion

The effect of temperature on arterial compliance could be determined from the data collected in the experiment. From the samples used, it was determined that as temperature increases, arterial compliance increases also. The fifteen samples of data after local heating were consistent with each other in showing that as temperature increases, the pulse wave velocity decreases. This, consequently, means the arterial compliance has increased. Fourteen of the fifteen samples of data after local cooling were consistent with each other in showing that as temperature decreases, the pulse wave velocity increases, and arterial compliance has decreased. The one data point with positive Δ PTT can neither confirm nor deny the hypothesis. The results of both local heating and cooling together are consistent with each other in showing that as temperature increases, the compliance of an artery increases also.

6. Acknowledgements

The authors thank Daniel Skripkariuk for inspiration for the project¹⁹ and the staff of the Coastal Carolina University Athletic Department for use of resources.

7. References

[1] American Heart Association. "About High Blood Pressure (HBP)." Accessed March 3, 2013. http://www.heart.org/HEARTORG/ConditionC/HighBloodPressure/AboutHighBloodPressure_uAb/About-High-Blood-Pressure_UCM_002050_Article.jsp.

[2] Najjar, Samer S. et al. "Pulse Wave Velocity Is an Independent Predictor of the Longitudinal Increase in Systolic Blood Pressure and of Incident Hypertension in the Baltimore Longitudinal Study of Aging."

Journal of the American College of Cardiology 51 (2008): 1377-1383. Accessed March 3, 2013. doi: 10.1016/j.jacc.2007.10.065.

[3] Millodot, Michel. *Dictionary of Optometry and Visual Science*, 7th ed. Edinburgh: Butterworth-Heinemann. Elsevier, 2009.

[4] *The American Heritage Medical Dictionary*, 2nd ed. revised.

[5] Anderson, Robert M. *The Gross Physiology of the Cardiovascular System*. Tucson: Racquet Press, 1993. Accessed May 5, 2011. <http://cardiovascular.cx>.

[6] Klabunde, Richard E. *Cardiovascular Physiology Concepts*, 2nd ed. Baltimore: Lippincott Williams & Wilkins, 2011. Accessed May 5, 2011. <http://cvphysiology.com>.

[7] The UCSF Academic Geriatric Resource Center. *Pulse Wave Velocity*. Accessed May 5, 2011. <http://agrc.ucsf.edu>.

[8] Bramwell, J. Crighton, and A.V. Hill. "The Velocity of the Pulse Wave in Man." *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, 93 (1922): 298-306.

[9] Asmar, Roland, et al. "Assessment of Arterial Distensibility by Automatic Pulse Wave Velocity Measurement." *Hypertension* 26 (1995): 485-490. Accessed May 5, 2011. doi:10.1161/01.HYP.26.3.485.

[10] Tozzi, Piergiorgio; Corno, Antonio; and Hayoz, Daniel. "Definition of Arterial Compliance." *American Journal of Physiology – Heart and Circulatory Physiology* 278 (2000): H1407. Accessed May 5, 2011.

[11] Rhoades, Rodney A. and Bell, David R. *Medical Physiology: Principles for Clinical Medicine*, 3rd ed. Baltimore: Lippincott Williams & Wilkins, 2009.

[12] Joyner, Michael J. "Effect of Exercise on Arterial Compliance." *Circulation* 102 (2000), 1214-1215.

[13] Tanaka, Hirofumi et al. "Aging, Habitual Exercise, and Dynamic Arterial Compliance." *Circulation* 102 (2000): 1270-1275.

[14] Charkoudian, Nisha. "Skin Blood Flow in Adult Human Thermoregulation: How It Works, When It Does Not, and Why." *Mayo Clinic Proceedings* 78 (2003): 603-612.

[15] Sessler, Daniel I. "Perianesthetic Thermoregulation and Heat Balance in Humans." *The FASEB Journal* 7 (1993), 638-644.

[16] *The Gale Encyclopedia of Medicine*, 4th ed.

[17] Hast, Jukka. "Self-Mixing Interferometry and Its Applications in Noninvasive Pulse Detection." MSc, University of Oulu, 2003.

[18] SinusRhythmLabels.svg. Accessed December 11, 2011.

<http://commons.wikimedia.org/wiki/File:SinusRhythmLa.svg>.

[19] Skripkariuk, Daniel. "Development of a Pulse Wave Velocity Lab." Report, University of Western Ontario, 2009.

High Peak Power VCSELs in Short Range LIDAR Applications

Neil E. Newman,¹ Duncan C. Spaulding,¹ Graham Allen,¹ Mohamed A. Diagne,^{1,2}

¹Department of Physics, Astronomy, and Geophysics, Connecticut College, New London, Connecticut 06320

²Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02420

A technique for short-range LIDAR that monitors high-peak-power pulses rather than uniform square pulses is demonstrated using a commercially available Optek OPV310 VCSEL with a 10 μ m aperture operating at a wavelength of 850nm for applications in LIDAR-based active defense systems. Using 10ns pulses at a 0.1% duty cycle and 60mW of peak power, a target one meter away exhibiting Lambertian reflectance was detected with a 3:1 minimum signal-to-noise ratio in a narrow-field LIDAR setup using a 28dB amplifier. At 0.75m, using the same target and no signal amplification, a 2:1 minimum signal to noise ratio was achieved in a wide-field setup. These results establish the viability of commercially available low power VCSEL devices for LIDAR.

Many current applications of LIDAR involve simple range-finding for imaging¹; however, short-range LIDAR systems for countermeasure deployment are not well developed. A system used for detection of fast-moving projectiles, such as those from improvised explosive devices (IEDs) and explosive formed penetrators (EFPs) must be able to detect and track targets at short ranges (about one meter) and successfully deploy countermeasures. By optimizing the electrical circuitry, the output power of a VCSEL can be enhanced^{2,3}, for use in wide-field LIDAR, which relies on the divergence of light from a point source to detect over a wide area. The wide-field method is preferred over concentrating the light with optics for the narrow-field because the latter method is incapable of detecting any projectiles whose paths do not cross through the beam; thus, a narrow-field LIDAR system is fundamentally flawed for projectile detection. Furthermore, commercially available VCSELs are not typically suitable for short-range detection, as many are low power output devices that are designed for use in fiber optics or other applications with low losses, whereas LIDAR is a high loss application¹. Since the number of photons that a laser device is able to emit in a short amount of time is limited^{4,5}, and because only a low percentage of photons that are emitted reflect back to the detector, enhancing the capability of a LIDAR system involves increasing the maximum output of the laser and/or the detector sensitivity and responsiveness.

VCSELs heat up through extended operation if stressed beyond their design capabilities, eventually resulting in lowered conversion

efficiency⁶. However, using pulses with pulse widths closer to the thermal time (around ten nanoseconds) can significantly reduce this effect by reducing the accumulation of heat at the device junction level. Furthermore, a lower duty cycle can account for dissipation of any accrued thermal energy⁴. In doing so, it is possible to put much more energy into the device than the design would otherwise allow under CW conditions. By reducing pulse width and duty cycle, the peak output power can be increased (see Fig. 1). Alongside these higher output powers, a larger, focused beam spot was implemented in the LIDAR test to determine the basic parameters required to later optimize a VCSEL for wide-field detection.

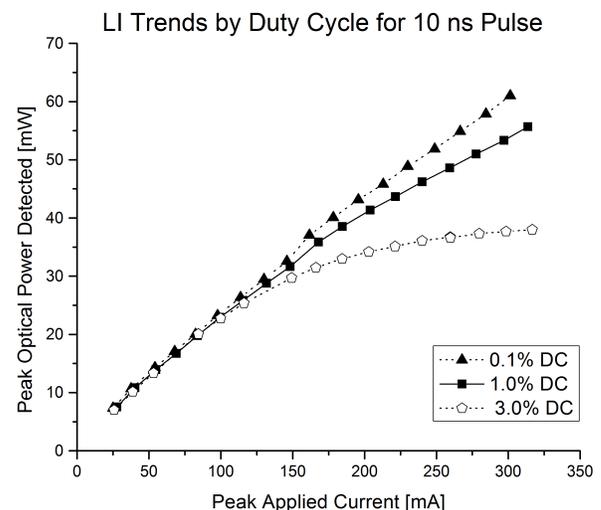


Figure 1. LI chart of OPV310 peak output power vs peak input current.

Parasitics within the electrical circuit network amounted to increased electrical impedance,

which distorted the input signals⁵. Additional impedance was introduced by a number of factors including the packaging of the VCSELs, the coaxial connections between devices, and the fact that the impedance of the system varied for different pulse repetition rates⁴. A rectifier diode was employed to limit the negative reflected current pulses from the impedance mismatch before these reflected pulses could sink into the device and cause failure due to high reverse bias. The use of the rectifier diode thereby increased the performance of the VCSEL and alleviated the need for strict impedance matching by preventing pulses with inverted voltage from entering the devices.

An Optek OPV310 VCSEL was used for this procedure with a Thorlabs DET210 photodiode for detection and optical power measurement. The OPV310 VCSEL emits at 850 nm with a 24-degree divergence angle and has a maximum output of 1.5 mW under CW conditions⁷. The DET210 is a 1 GHz, 0.8-mm² silicon-PIN photodiode with a spectral response from 200 to 1100 nm⁸. The small aperture of the photodiode required the use of a 6 cm diameter double-convex lens to focus the reflected light into the detector. Before each test, the detector and lens positions were optimized for maximum signal detection. For a pulse width of 10 ns, three different duty cycles were investigated to find the peak performance without thermal rollover. At a duty cycle of 3%, this thermal rollover was noticeable as the output power leveled off with increasing current, while a 0.1% duty cycle showed a nearly linear increase (see Fig. 1). The maximum peak output power attained was 61mW with a duty cycle of 0.1%, while the maximum peak output powers of the 1% and 3% duty cycles were 55mW and 38mW, respectively. It should be noted that higher powers may be possible on this device at the 0.1% duty cycle, but the Agilent 8104AA pulse generator used was limited to a 10.5-volt maximum output.

With the setup configured for wide-field detection and the target distance set at 0.75±0.03 m, the system was able to detect a Lambertian block with a 0.05 m by 0.05 m surface facing the device when the laser was operated at the maximum power output. The property of Lambertian reflectance on the surface of this block

ensures isotropic scattering from incident light on the block. The results obtained using the Lambertian target were similar to those from a block of aluminum and from a block of copper when each was oriented to reflect light away from the detector. (see Fig. 2) The signal-to-noise ratios in the configuration with each the Lambertian block and the off-angle copper were approximately 2:1. (see Fig. 2) The DET210 photodiode has a broad spectral response⁸, and during testing was shown to respond better to the laser light when room lighting was deactivated. Therefore, the use of a filter to isolate the narrow spectrum of the OPV310 could increase the detected optical power and, in turn, would increase the resolution of the system further. Additionally, currently available Geiger-mode avalanche photodiodes would be more likely to respond to a lower number of photons than what was returned in the 0.75 m test, and thusly, may realize the goal of 1 m detection with the OPV310 at the utilized output power⁹. Therefore, the number of photons incident on the block at this reduced range was sufficient for this LIDAR demonstration. The maximum output power was not great enough to do wide-field LIDAR at a range greater than 0.75 m. However, it was possible to focus the beam spot to approximately 2 cm in diameter on a target area at a one-meter range with a plano-convex lens located at the laser aperture. With the Lambertian block placed at a target range of 1.00±0.03 m, a calculated distance of 0.996±0.047 m was obtained. (see Table 1)

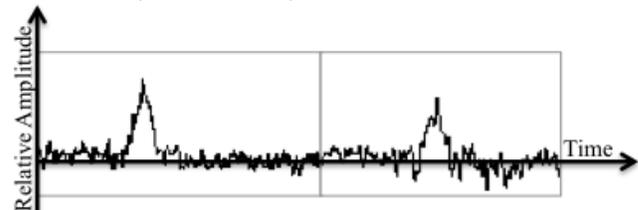


Figure 2. Detector response for wide-field LIDAR at 0.75 m. Left image shows the response from the Lambertian block, the right shows the response for a flat copper block angled 45 degrees from the detector.

Actual Range (cm±3cm)	Calculated Range (cm)
90	90.9 ± 3.6
100	99.6 ± 4.7
110	110.1 ± 5.9
120	119.4 ± 3.4
130	130.5 ± 3.1
140	140.8 ± 4.8

Table 1. LIDAR measurements for select ranges.

A ZFL-1000H 28dB amplifier was used in order to find the minimum output power of the VCSEL necessary to produce a reliable signal for the narrow-field arrangement. The amplifier was applied to the output from the photodetector in conjunction with noise modeling and filtering. The minimum triggering signal was found when the VCSEL output power was reduced to 30mW, with a signal to noise ratio of about 2:1. When using the maximum achieved output of the VCSEL in the narrow-field setup, the signal-to-noise ratio was no less than 3:1. (see Fig. 3) Given the peak output power limitations of commercial VCSELs, even under pulse conditions, detecting projectiles in a wide-field of view at a one-meter distance would require increasing either the output power of the device or the detector sensitivity further.

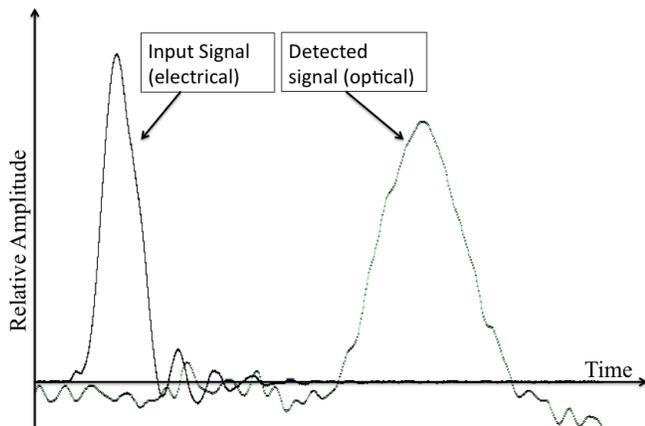


Figure 3. Sample LIDAR capture. The thinner pulse on the left side is the input current pulse to the VCSEL. The wider pulse on the right is the light returned to the detector. This image is from the 1m narrow-field LIDAR test, with the peak-to-peak delay corresponding to the calculated distance. Note that the amplitudes of the two waves are not scaled identically.

Further investigation is necessary to determine whether carrier saturation in the cavity of the VCSEL has been achieved for the 0.1% duty cycle, and also whether the duty cycle could be further refined in the 0.1% to 1% range while maintaining similar output levels. For the application of projectile detection from explosive or shaped charges, a typical speed of 2 km/s¹⁰ would be possible for detection ranges under one meter, as the target would not scale a large distance between pulses. However, for the maximum recorded shaped charge projectile speed

on the order of 12 km/s¹⁰, the distance scaled between pulses could be too great for effective detection and velocity measurement of a target at the optimum pulse width and duty cycle of this system. It may be possible to repeat this exercise with multiple devices and a form of time-division multiplexing to enhance the effective pulse repetition frequency (PRF) of the system while maintaining the same output power. With a higher PRF, the time between pulses would be decreased, allowing for accurate measurements to be taken of higher speed projectiles.

While commercial VCSELs are not typically suitable for wide-field LIDAR, it was possible to use one as a demonstration platform for pulse amplitude optimization and LIDAR testing. This VCSEL was able to output a maximum peak power of 61mW at a 10-nanosecond pulse width and 0.1% duty cycle with the use of a rectifier diode. Additionally, this power was sufficient to demonstrate effectiveness for narrow-field LIDAR scenarios, and should be applicable to wide-field LIDAR for capable devices. Efforts are underway to design and produce wavelength beam combining for custom-built large arrays of single mode VCSELs to provide more power under pulse conditions and to cover a wider field of view in short-range LIDAR applications. In addition, more sensitive detectors than the DET210 photodiode could further enhance the signal to noise ratio.

Acknowledgements

This work was supported by the National Science Foundation (Awards 0925545), and by the Connecticut College Dean of Faculty Funds.

References

- ¹ Geske, J., C. Wang, M. MacDougal, R. Stahl, D. Follman, H. Garrett, T. Meyrath, D. Snyder, E. Golden, J. Wagener, and J. Foley. "High power VCSELs for miniature optical sensors," *Proc. of SPIE* 7615 (2010): 76150E-1–11, doi: 10.1117/12.847184.
- ² Geske, J., M. MacDougal, G. Cole and D. Snyder. "High-power VCSELs for smart munitions," *Proc. of SPIE* 6287 (2006): 628703-1–12, doi: 10.1117/12.679296.

-
- ³ Miller, M., M. Grabherr, R. King, R. Jager, R. Michalzik, and K.J. Ebeling. "Improved output performance of high-power VCSELs," *Selected Topics in Quantum Electronics, IEEE Journal of*, 7,2 (2001): 210–216.
- ⁴ Michalzik, R., M. Grabherr and K.J. Ebeling. "High-power VCSELs: modeling and experimental characterization", *Proc. of SPIE* 3286 (1998): 206–219.
- ⁵ Moench, H., J. Baier, S. Gronenborn, J. Kolb, M. Miller, P. Pekarski, M. Schemmann and A. Valster. "Advanced characterization techniques for high power VCSELs", *Proc. of SPIE* 7615 (2010): 76150G-1–11, doi: 10.1117/12/839953.
- ⁶ Grabherr, M., R. Jager, M. Miller, C. Thalmaier, J. Herlein, R. Michalzik, K.J. Ebeling. "Bottom-emitting VCSEL's for high-CW optical output power," *Photonics Technology Letters, IEEE*, 10,8 (1998): 1061–1063.
- ⁷ "Optek OPV310 Data Sheet," Optek, Inc., Accessed May 10, 2012, <http://www.optekinc.com/datasheets/OPV314.pdf>.
- ⁸ "DET210 Data Sheet," Thorlabs, Accessed May 10, 2012, <http://www.thorlabs.com/thorcat/2200/2201-S01.pdf>.
- ⁹ Aull, B.F., A. H. Loomis, D. J. Young, R. M. Heinrichs, B. J. Felton, P. J. Daniels, and D. J. Landers, "Geiger- mode avalanche photodiodes for three-dimensional imaging," *Lincoln Lab. J*, 13 (2002): 335–350.
- ¹⁰ Kennedy, D.R., "History of the Shaped Charge Effect: The First 100 Years," March 1990, *Defense Technical Information Center*, Accessed June 30, 2012, www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA220095.

Spacecraft Approaching Technique

Rey Carvajal¹

Harriet L Wilkes Honors College of Florida Atlantic University

Abstract

Using already established equations of motion for particles under a gravitational force, we analyze the motion of two spacecrafts in the same circular orbit approaching each other if one of the craft were to speed up or slow down. The purpose of this paper is to analyze and develop equations of motion for the transfer of spacecrafts from circular orbits to elliptical orbits with the intention of spacecraft approach. We can come up with a table of velocities that allow for a number of safe approaches for the spacecraft to take. We found that projecting a safe approach requires that we know the new speed of the transfer spacecraft, the desired change in distance between the two spacecrafts, and the magnitude of the radius vector of the perigee for the elliptical transfer.

Introduction

Newton's second law of motion coupled with the principles of kinematics allows for the use of the methods of conservation of momentum and energy to relate force, mass, velocity and time. The application of conservation of momentum and energy in explaining the motion of a particle is extremely useful when dealing with conservative forces such as gravity, since acceleration no longer becomes a necessary relation to analyze the motion of said particle.

Walter Hohmann, in his 1925 book *The Accessibility of Celestial Bodies*, calculated an elliptical orbit transfer between circular coplanar orbits of different radii around a planet. Hohmann used half of an ellipse to accomplish the transfer. Hohmann's research on transfers between orbits around celestial bodies was a result of his fascination with outer space travel. He knew that to travel between and around planets, spacecrafts would require large amounts of fuel, so he discovered the most efficient way to transfer spacecrafts from a circular orbit to an elliptical orbit and then back to another circular orbit.

When in a circular orbit around a celestial body, if a spacecraft slows down or speeds up due to the thrust of its rockets, it enters an elliptical orbit with the point of thrust being the new apogee or perigee respectively. This paper analyzes a similar technique to what Hohmann theorized, except that the spacecraft entering the elliptical orbit will complete one or a number of orbits before re-entering its original circular orbit. The new period of rotation of the space craft in the elliptical orbit will be shorter than the period for the spacecraft still in the circular orbit, so that when the spacecraft

returns to its original circular orbit, it will be closer to the other spacecraft.

Theoretical Application

The initial conditions of spacecrafts A and B are that both are moving in the same circular orbit.

Spacecraft B will either slow down or speed up, thus entering an elliptical orbit. As shown in figure one θ_1 is the angle to A when B reaches the fixed line $L-L^1$. Equations of motion state that the velocity required to have a circular orbit around a planet (Earth) is

$$v_{cir} = \sqrt{\frac{GM}{r_a}} ,$$

and the period (P_A) of travel for the orbit is

$$P_A = \frac{2\pi r_a}{v_{cir}} .$$

*A table of all notations can be found in Appendix 1
The final condition of spacecraft B is that it has ended the first elliptical orbit. Spacecraft A is closer to B. We first must express spacecraft B's period with respect to the spacecraft A's period as shown:

$$P_B = P_A - \frac{\theta_1 - \theta_2}{\omega_A} , \quad \Delta t = \frac{\Delta \theta}{\omega_A} ,$$

$$P_B = P_A - \Delta t .$$

The equations we are interested in that relate to an elliptical orbit around a planet are as follows:

$$P_B = \frac{2\pi ab}{h} , \quad a = \frac{r_a + r_p}{2} , \quad b = \sqrt{r_a r_p}$$

$$h = r_a v_a .$$

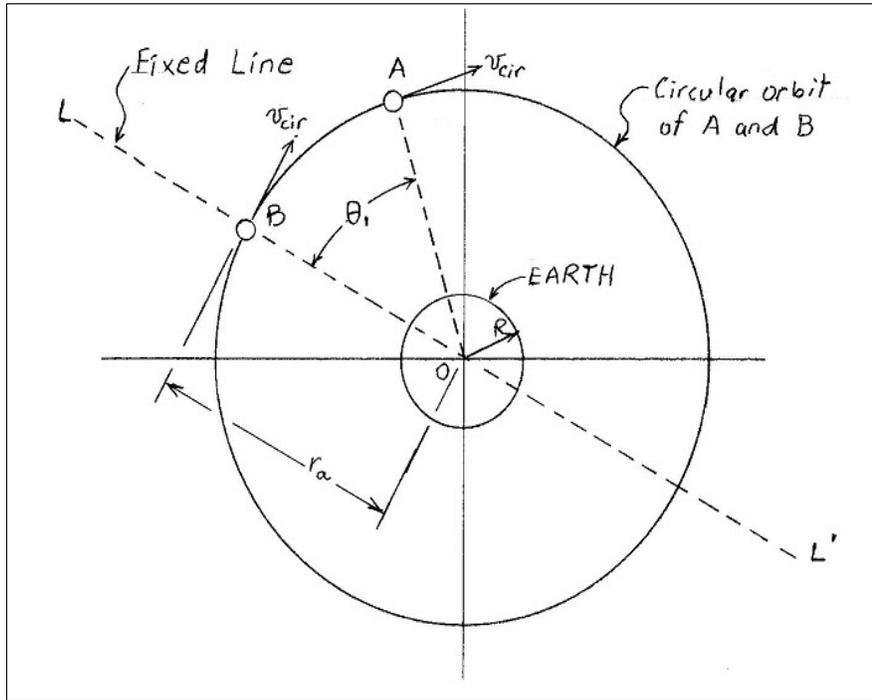


Figure 1. Spacecraft B is beginning its transfer to a full elliptical orbit around planet Earth.

To solve for the new velocity at the apogee, we apply the conservation of energy principle at the apogee and perigee,

$$E_a = E_p \text{ ,}$$

the points with the greatest and smallest distances from the focus (planet). As seen in Figure 2, sub-index *a* refers to the apogee and sub-index *p* refers to the perigee. We then expand the relation to distinguish kinetic from potential energy:

$$T_a + U_a = T_p + U_p$$

$$\frac{1}{2} v_a^2 - \frac{GM}{r_a} = \frac{1}{2} v_p^2 - \frac{GM}{r_p} \text{ .}$$

We then apply the principle of conservation of angular momentum at the apogee and perigee:

$$L_{o,a} = L_{o,p}$$

$$r_a m v_a \sin(\theta_a) = r_p m v_p \sin(\theta_p) \text{ ,}$$

where θ_a is equal to θ_b at 90 degrees.

The above equation represents the magnitude of the cross product between the radial vector and the velocity vector. For both the perigee and apogee the angle between these two vectors is ninety degrees so the magnitude of the total angular momentum at these two points is only dependent on the magnitude of the velocity and radius as shown :

$$r_a v_a = r_p v_p$$

Taking our equations of energy and angular momentum we substitute

$$v_p = \frac{r_a v_a}{r_p} \text{ into}$$

$$v_a^2 - v_p^2 = \frac{2GM}{r_a} - \frac{2GM}{r_p}$$

and solving velocity to obtain:

$$v_a = \sqrt{\frac{2GM}{(r_a + r_p)} * \frac{r_p}{r_a}} \text{ .}$$

We now substitute circular, orbital velocity

$$v_{cir} = \sqrt{\frac{GM}{ra}} \text{ into } P_A = \frac{2\pi r_a}{v_{cir}} \text{ to get}$$

$$P_A = \frac{2\pi r_a^{\frac{3}{2}}}{\sqrt{GM}} \text{ .}$$

Then by substituting P_A of the above equation into

$$P_B = P_A - \Delta t \text{ we obtain:}$$

$$\frac{2\pi ab}{h} = \frac{2\pi r_a^{\frac{3}{2}}}{\sqrt{GM}} - \Delta t$$

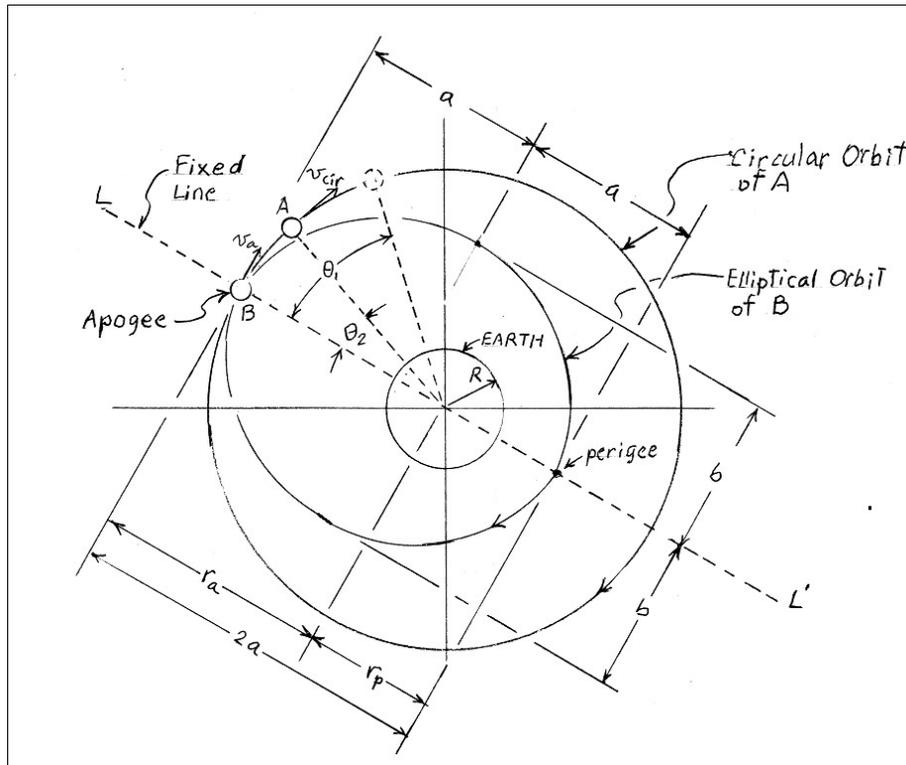


Figure 2. The final conditions of spacecraft B is that it has arrived to the origin of the full elliptical orbit. Spacecraft A is closer to spacecraft B now.

Now that we have only terms of Δt , r_p , and r_a

$$r_p = \left[\frac{\sqrt{2GM}}{\pi} \left(\frac{2\pi r_a^{\frac{3}{2}}}{\sqrt{GM}} - \Delta t \right) \right]^{\frac{2}{3}} - r_a$$

we can solve for r_p . This simplifies the calculation for how close spacecraft B can approach a planet on its elliptical path so as to lessen the time lag between its circular orbit and the circular orbit of spacecraft A.

Analysis

We have made a graph of a family of three approaching spacecraft curves (Figure 3) for planet Earth at altitudes of 100, 200 and 300 km using different Δt of 4, 10, 20, 40, 60, 80, 100, 120, and 140s. We limit the extent of deceleration for spacecraft B to where it's perigee is just above the atmosphere of earth

$$R + atm_{depth} = 6.44 \times 10^6 \text{ m}$$

There is a difference in time between the transition to an elliptical orbit and the transition to a circular orbit for a spacecraft. This is due to the nature of a

rocket's propulsion, for mass is expended to increase or decrease in speed. We will show that this difference in mass has a minimal effect on the Approaching Technique depicted thus far.

Disclaimer-For simplicity in this paper the following assumptions hold:

1. The Change of speed at the apogee (because of the rocket firing) occurs in a very short period of time.
2. The radius at the apogee (r_a) remains constant during the rocket firing.
3. Only the rocket thrust is the unbalanced force during the rocket firing (virtually no gravity).

First, we analyze the decreasing in speed of spacecraft B at the apogee (Figure 4):

By applying Newton's 2nd Law we can relate the reverse thrust of the

$$\text{rocket to the force: } -Thrust = ma$$

We expand the relation into a differential and introduce the rate of burn into the equation

$$-\frac{dm}{dt} v_{rel} = m \frac{dv}{dt}, \frac{dm}{dt} = R_{burn}$$

Altitude (km)	r_a (m)	v_{cir} (m/s)	Δt (s)	r_p (m)	v_a (m/s)	Limitations
100	6.47×10^6	7853	4	6.463×10^6	7851	
			10	6.453×10^6	7848	
			20	6.437×10^6	7843	
			40	6.403×10^6	N/A	$r_p < r_{p,min}$
200	6.57×10^6	7793	4	6.563×10^6	7791	
			10	6.553×10^6	7789	
			20	6.537×10^6	7784	
			40	6.504×10^6	7774	
			60	6.471×10^6	7764	
			80	6.437×10^6	7754	
300	6.67×10^6	7735	4	6.663×10^6	7733	
			10	6.654×10^6	7730	
			20	6.637×10^6	7725	
			40	6.604×10^6	7716	
			60	6.571×10^6	7706	
			80	6.538×10^6	7696	
			100	6.505×10^6	7686	
			120	6.472×10^6	7676	
			140	6.439×10^6	N/A	$r_p < r_{p,min}$

Table 1. Table of Safe Approaches for Spacecraft B

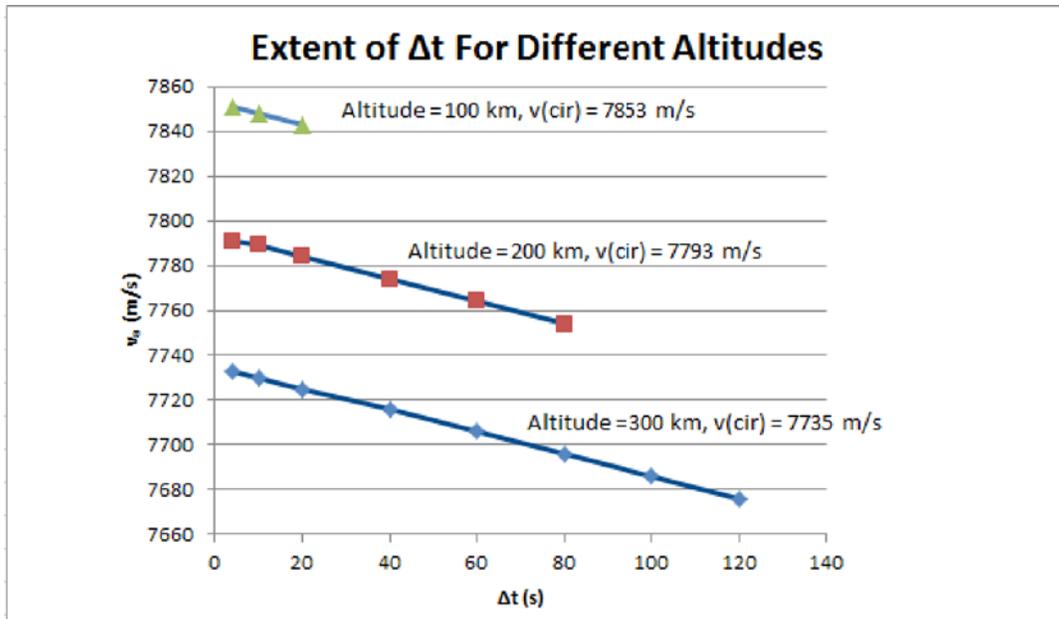


Figure 3. The graph shows that the maximum differences in periods are accomplished at higher altitudes. The maximum difference in periods at 300 km of altitude is six times greater than that at 100 km.

If we substitute the following expression into our equation

$$-R_{burn} v_{rel} = m \frac{dv}{dt} ,$$

and then express the new mass in terms of the initial mass minus the mass burned in transition,

$$\left(-\frac{R_{burn}}{m} \right) v_{rel} = \frac{dv}{dt} , m = m_i - R_{burn} t ,$$

then we arrive to a differential equation that can be analytically integrated:

$$\left(-\frac{R_{burn}}{m_i - R_{burn} t} \right) v_{rel} = \frac{dv}{dt} .$$

We integrate the relation:

$$-R_{burn} v_{rel} \int \frac{dt}{m_i - R_{burn} t} = \int dv$$

$$v_{rel} \ln(m_i - R_{burn} t) = v + c .$$

When $t = 0$

$$v = v_{cir} ; c = v_{rel} \ln(m_i) - v_{cir} ,$$

and we can then obtain the initial conditions

$$v = v_{cir} + v_{rel} \ln \left(\frac{m_i - R_{burn} t}{m_i} \right)$$

For the final conditions, where spacecraft B is slowing down, we first find the time of thrust by making these substitutions $v = v_a, t = t_{th}$ and solving for t

$$v_a = v_{cir} + v_{rel} \ln \left(\frac{m_i - R_{burn} t_{th}}{m_i} \right)$$

$$t_{th} = \frac{m_i}{R_{burn}} \left(1 - e^{\frac{(v_a - v_{cir})}{v_{rel}}} \right) .$$

The mass of the propellant used and the mass of spacecraft B at the end of the thrust are respectively:

$$m_{pro} = R_{burn} t_{th}$$

$$m_f = m_i - m_{pro} .$$

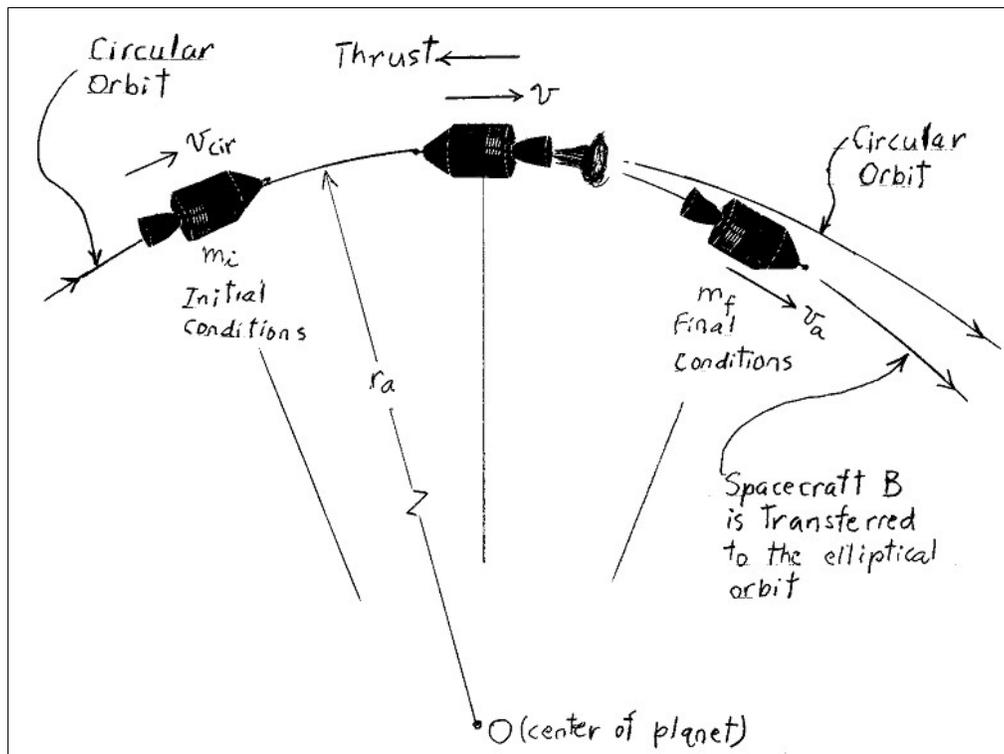


Figure 4. Spacecraft B is reducing its speed by firing the rocket in reverse position. The firing occurs between the initial and final conditions shown.

We use the negative expression for thrust when increasing the speed of spacecraft B at the apogee (Figure 5):

$$Thrust = ma$$

$$-v_{rel} \ln(m_i - R_{burn} t) = v + c$$

where initial conditions are $t=0$, and $v=v_a$. At final conditions $v=v_{cir}$, $t=t_{th}$ and the circular velocity is:

$$v_{cir} = v_a + v_{rel} \ln\left(\frac{m_i}{m_i - R_{burn} t_{th}}\right)$$

We now solve for the time of thrust, as before, and obtain the similar equation

$$t_{th} = \frac{m_i}{R_{burn}} \left(1 - e^{\frac{(v_a - v_{cir})}{v_{rel}}}\right)$$

Calculation

The expressions analyzing time of thrust for an increase and a decrease of speed for spacecraft B are equal to each other. However, because of a difference in mass between the two transfers of the rocket, the time for each will not be the same.

We now look at a specific case where spacecrafts A and B are in a circular orbit 300 km above the Earth and spacecraft B is 400 km away from spacecraft A. Spacecraft B first slows down and enters an elliptical orbit. Spacecraft B then enters back into the original circular orbit at the apogee after completing two elliptical turns. At the end of this, spacecraft B will be 5 km away from spacecraft A. Spacecraft B has a mass (including propellant) of 1200 kg, a rate-of-burn of 4 kg/s, and a relative velocity of 1500 m/s. Data for spacecraft A is not required for these set of calculations. The depth of the atmosphere is 70km. The mass of the Earth (M) will be considered 5.98×10^{24} kg, the radius (R) 6.37×10^6 m, and the gravitational constant (G) $6.673 \times 10^{-11} \text{ Nm}^2/\text{kg}^2$.

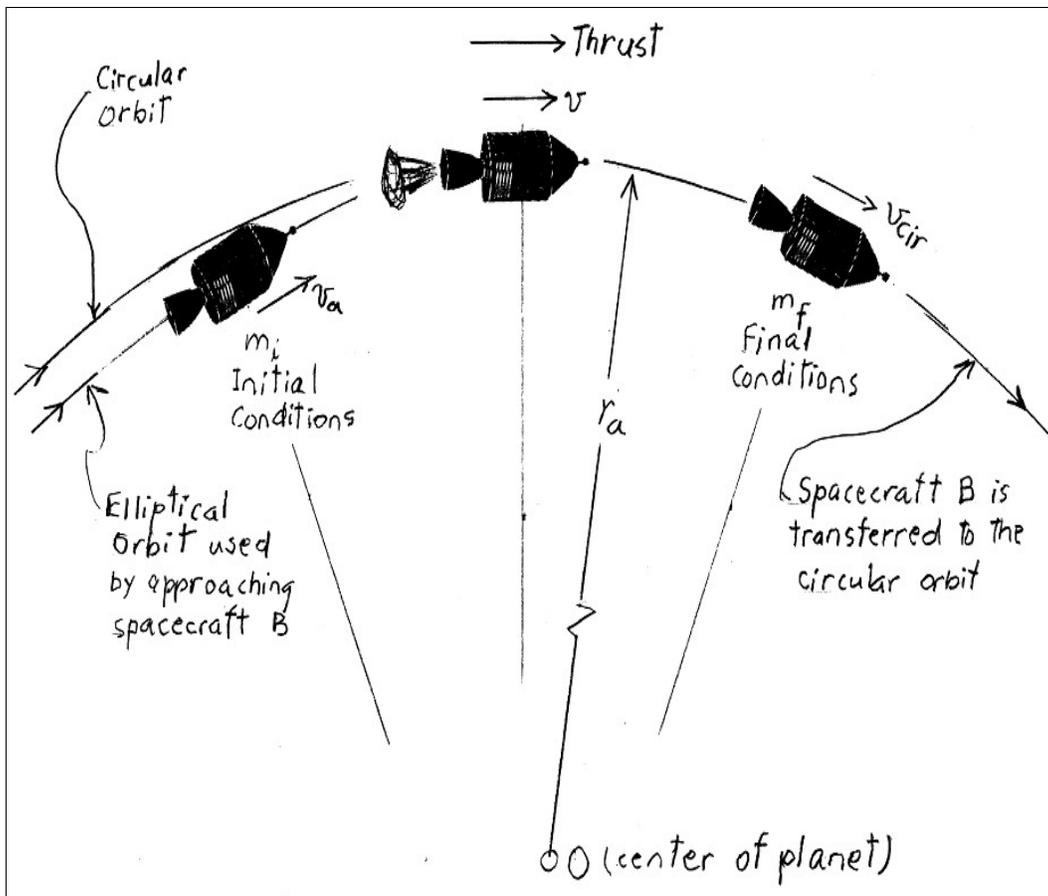


Figure 5. Spacecraft B is increasing its speed to enter back into the circular orbit by firing the rocket in the straight position. Firing occurs between initial and final conditions shown.

The magnitude of the radius vector of the apogee is

$$r_a = R + \text{altitude} = 6.67 \times 10^6 \text{ m} .$$

The circular velocity at 300 km altitude is:

$$v_{cir} = \sqrt{\frac{GM}{r_a}} = 7735 \text{ m/s}$$

The difference between the periods of the circular orbit and the elliptical orbit is:

$$\Delta t = \frac{d - d^1}{n_t v_{cir}}$$

Where d is the distance between A and B (400km), d^1 is the final distance between A and B (5km), and n_t is the number of elliptical turns spacecraft B takes, and Δt is 25.533s.

The perigee radius vector is calculated to be

$$r_p = \left[\frac{\sqrt{2GM}}{\pi} \left(2\pi \frac{(6.67 \times 10^6 \text{ m})^{\frac{3}{2}}}{\sqrt{GM}} - 25.533 \text{ s} \right) \right]^{\frac{2}{3}} - 6.67 \times 10^6 \text{ m}$$

$$r_p = 6.628058 \times 10^6 \text{ m} .$$

We then verify that spacecraft B does not encroach into the atmosphere when it is at the perigee:

$$r_{p, \text{min}} = R + \text{atm}_{\text{depth}} = 6.37 \times 10^6 \text{ m} + 7 \times 10^4 \text{ m}$$

$$r_{p, \text{min}} = 6.44 \times 10^6 \text{ m}$$

$$r_p > r_{(p, \text{min})}$$

$6.628058 \times 10^6 \text{ m} > 6.440 \times 10^6 \text{ m}$, the spacecraft does not reach the border of the atmosphere.

The speed of spacecraft B at the apogee is

$$v_a = \sqrt{\frac{2GM}{(6.670 \times 10^6 \text{ m} + 6.628058 \times 10^6 \text{ m})}} * \frac{6.62858 \times 10^6 \text{ m}}{6.67 \times 10^6 \text{ m}}$$

$$v_a = 7722.3 \text{ m/s} .$$

The time of thrust for spacecraft B to enter the elliptical orbit is

$$t_{th} = \frac{m_i}{R_{burn}} \left(1 - e^{\frac{(v_a - v_{cir})}{v_{rel}}} \right)$$

$$t_{th} = \frac{1200 \text{ kg}}{4 \text{ kg/s}} \left(1 - e^{\frac{(7722.3 \text{ m/s} - 7735 \text{ m/s})}{1500 \text{ m/s}}} \right)$$

$$t_{th} = 2.53 \text{ s} .$$

The time of thrust during transfer from circular to elliptical is greater than for elliptical to circular because of the loss of mass during the first transfer. This is shown in the next set of calculations. The new mass of spacecraft B after entering elliptical orbit is

7

$$m_f = m_i - m_{pro}$$

$$m_{pro} = R_{burn} t_{th} , \quad m_{pro} = 10.12 \text{ kg}$$

$$m_f = 1189.88 \text{ kg} .$$

The time of thrust for spacecraft B to enter back into original circular orbit is

$$t_{th} = \frac{1189.88 \text{ kg}}{4 \text{ kg/s}} \left(1 - e^{\frac{(7722.3 \text{ m/s} - 7735 \text{ m/s})}{1500 \text{ m/s}}} \right) ,$$

which as shown, is smaller than the thrust for the first transfer

$$t_{th} = 2.51 \text{ s} < t_{th} = 2.53 \text{ s} .$$

The final mass of spacecraft B after entering original circular orbit is

$$m_{pro} = R_{burn} t_{th} , \quad m_{pro} = 10.04 \text{ kg}$$

$$m_f = m_i - m_{pro} = 1189.88 \text{ kg} - 10.04 \text{ kg}$$

$$m_f = 1179.84 \text{ kg} .$$

Conclusion

The effects of fuel lost from the two transfers creates a total time difference of 0.02 seconds between the predicted period for the elliptical orbit of B and the circular orbit of A. This time difference depends on the rate of burn, R_{burn} , and is small for our set of data. As to the effectiveness of the approaching technique, this can be summarized by a relation we will call the **Sensitivity of Approach:**

$$\text{Sensitivity of Approach} = \frac{d - d^1}{v_{\text{cir}} - v_a} = \frac{400\text{km} - 5\text{km}}{7735\text{m/s} - 7722.3\text{m/s}} = 15.55 \frac{\text{km}}{\text{ms}^{-1}}$$

The term refers to the difference in distance between spacecraft A and B before and after the technique per the speed reduction of spacecraft B. The distance between spacecrafts A and B is reduced in 15.55km per 1 m/s in speed reduction of spacecraft B. This sensitivity of approach is very high for our sample calculations, and this method of approach is recommended for spacecrafts separated at great distances and only as a previous step to final coupling.

Appendix I

R, M = radius and mass of the Earth in m, kg
 G = gravitational constant, $6,673 \times 10^{-11} \text{Nm}^2/\text{kg}^2$
 v_{cir} = speed of spacecraft in a circular orbit in m/s
 r_a = longer radius vector in m
 r_p = shorter radius vector in m
 P_A = period of spacecraft A in s
 a, b = Semimajor and semiminor axis of the elliptical orbit in m
 P_B = period of spacecraft B in s
 ω_A = angular speed of spacecraft A in rad/s
 Δt = difference between spacecraft periods in s
 h = angular momentum per unit of mass of spacecraft B with respect to O, in m^2/s
 v_a = speed (at the apogee) that should be given to spacecraft B to transfer to an elliptical orbit, in m/s
 m = mass of spacecraft B in kg
 m_i, m_f = initial and final masses of spacecraft B during any transfer, in kg
 Thrust = push created by operating rocket of spacecraft B, in N
 a = acceleration of spacecraft B in m/s^2
 t = time in s
 R_{burn} = rate of propellant burning in kg/s
 v_{rel} = speed of exhaust propellant gases relative to

spacecraft B, in m/s

t_{th} = duration of the thrust created by the exhausting gases, in s

m_{prop} = mass of the burned propellant during any transfer, in kg

d = initial distance between spacecraft A and B in m

d^1 = final distance between spacecraft A and B in m

n_t = number of elliptical turns of spacecraft B

$\text{atm}_{\text{depth}}$ = depth of atmosphere in m

Acknowledgments

A special thanks goes out to my professor and academic counselor Dr. Meredith Blue for her support and encouragement in writing this paper, and for the solid guidance she has provided me with over the past three years. I would also like to thank the Harriet L. Wilkes Honors College of Florida Atlantic University for sponsoring my paper's presentation at the 2012 Florida Collegiate Honors Council Conference.

References

- [1] Beer, Ferdinand P., and E. Russell Johnston. "Kinetics of Particles: Energy and Momentum." *Vector mechanics for engineers: dynamics*. 5th ed. New York: McGraw-Hill, 1988. 729-826. Print.
- [2] Celletti, Alessandra, and Perozzi, Ettore. 2007. "The Accessibility of Celestial Bodies" In *Celestial Mechanics The Waltz of the Planets*. 158-161, edited by R.A. Marriot. Praxis Publishing, Chichester, UK.
- [3] Walter Hohmann. 1960. *The Attainability of Heavenly Bodies*, Washington: NASA Technical Translation F-44
- [4] Resnick, Robert, and David Halliday, and Kenneth S. Krane. 1991. *Physics, 4th Edition, Vol.1*. John Wiley & Sons, Inc, New York.