Statistical Determination
of the MACHO Mass Spectrum

Cheongho Han
Andrew Gould*

Dept. of Astronomy, The Ohio State University
Columbus, OH 43210
e-mail cheongho@payne.mps.ohio-state.edu
e-mail gould@payne.mps.ohio-state.edu

Abstract

The mass function of 55 Massive Compact Objects (MACHOs) detected toward the Galactic bulge is statistically estimated from Einstein ring crossing times $t_e$. For a Gaussian mass function, the best-fitting parameters are $\langle \log(m/M_\odot) \rangle = -0.79$ and $\sigma_{\log(m/M_\odot)} = 0.44$. If the mass function follows a power-law distribution, the best-fitting mass cut-off and power are $m_{cut} = 0.1 \ M_\odot$ and $p = -2.3$, respectively. Both best determined mass functions are compared with that obtained from Hubble Space Telescope (HST) observations. The power law is favored over the Gaussian and HST mass functions at the $\sim 2.8 \sigma$ and $\sim 4.2 \sigma$ levels, respectively. However, the fact that all the models have very poor fits to the longest four events with $t_e \geq 80$ days remains a puzzle.

submitted to *The Astrophysical Journal* April 12, 1995

Preprint: OSU-TA-5/95

* Alfred P. Sloan Foundation Fellow
1. Introduction

Major efforts to detect MACHOs by observing microlensing events of source stars located in the Galactic bulge have been carried out by the OGLE (Udalski et al. 1994) and MACHO (Alcock et al. 1995) groups. Current and prospective observations can constrain the disk/bulge normalization (Stanek 1994) and the mass distribution of the bulge (Kiraga & Paczyński 1994; Evans 1994; Han & Gould 1995b).

However, it is very difficult to obtain information about the physical parameters of the individual lenses. This is because the only measurable quantity from current observations, the Einstein ring crossing time \( t_e \), depends on a combination of physical parameters of the individual lenses. The Einstein ring crossing time is related to the physical parameters of the lenses by

\[
    t_e = \frac{r_e}{v}; \quad r_e = \left(\frac{4Gm}{c^2} \frac{D_{ol} D_{ls}}{D_{os}}\right)^{1/2}, \tag{1.1}
\]

where \( m \) and \( v \) are the mass and the transverse speed of the lens, \( r_e \) is Einstein ring radius, and \( D_{ol}, D_{os}, \) and \( D_{ls} \) are the distances between the observer, lens, and source.

Several ideas have been proposed to break the degeneracy of \( t_e \). Distance can be measured when the lensing star is bright enough to be detected (Kamionkowski 1994). The ambitious and promising idea of measuring MACHO parallaxes from a satellite would clarify the dynamical motions of MACHOs (Gould 1994b, 1995; Han & Gould 1994a). The MACHO proper motion, \( \mu = v/D_{ol} \), can be measured photometrically (Gould 1994a; Nemiroff & Wickramasinghe 1994) and spectroscopically (Maoz & Gould 1994). When the information from proper motion and parallax measurements are combined, the degeneracy of \( t_e \) can be completely broken, yielding \( m, v, \) and \( D_{ol} \). However, measuring the distance or proper motion of the lens star is possible only under very restricted conditions. Most lenses are likely to be too faint to be detected. A proper motion is measurable only when a lens
crosses over or very close to the face of the source star. Although satellite-based parallaxes are promising, actual measurements are many years off.

However, one can still obtain much information about the mass spectrum of MACHOs from the data on time scales, $t_e$, that are provided by current observations: MACHO (Bennett et al. 1994) and OGLE (Udalski et al. 1994) have detected > 50 events. De Rújula, Jetzer, & Massó (1991) have developed a method of “mass moment” to analyze MACHO masses and Jetzer (1994) has applied this to measure the mean mass ($0.28 \, M_\odot$) of bulge MACHOs using 4 OGLE events. Here we use Maximum Likelihood to estimate the mass function of microlenses which are detected toward the Galactic bulge by combining the latest $t_e$ data from MACHO and OGLE with plausible models for the velocity and spatial distributions of the lenses and sources. We test both Gaussian and power-law mass functions, and find $\langle \log(m/M_\odot) \rangle = -0.79$ and $\sigma_{\log(m/M_\odot)} = 0.45$ for a Gaussian mass function, and $m_{\text{cut}} = 0.1 \, M_\odot$ and $p = -2.3$ for a power law. The determined best-fitting mass functions are compared with the observed mass functions as determined with the Hubble Space Telescope (HST).

2. Models

We adopt a bar-structured Galactic bulge with the double exponential disk. The Galactic bulge is modeled by a “revised COBE” model which is based on the COBE model (Dwek et al. 1994) except for the central part of the bulge. In the inner $\sim 600 \, \text{pc}$ of the bulge, we adopt the high central density Kent (1992) model since the COBE model does not match very well with observations in this region. The disk is assumed to have an exponential distribution with vertical and radial scale heights of $H_z = 325 \, \text{pc}$ and $H_R = 3.5 \, \text{kpc}$, respectively (Bahcall 1986). The models are discussed in detail by Han & Gould (1995a; 1995b).

In our model, disk MACHOs are assumed to have a Gaussian velocity distribution represented by $f(v_i) \propto \exp \left[ -\frac{(v_i - \bar{v}_i)^2}{2\sigma_i^2} \right]$, $i = x, y, z$. Here the coordinates ($x, y, z$) have their center at the Galactic center and $x$ and $z$ axes point to the
Sun and the north Galactic pole, respectively. We adopt the disk component velocity distribution as $v_{z,disk} = 0 \text{ km s}^{-1}$, $\sigma_{z,disk} = 20 \text{ km s}^{-1}$, $v_{y,disk} = 220 \text{ km s}^{-1}$, and $\sigma_{y,disk} = 30 \text{ km s}^{-1}$. The velocity of the barred bulge model is deduced from the tensor virial theorem (Binney & Tremaine 1987; Han & Gould 1995b) with resulting velocity dispersion components $(\sigma_x', \sigma_y', \sigma_z') = (115.7, 90.0, 78.6) \text{ km s}^{-1}$.

Here, the coordinates $(x', y', z')$ are aligned along the axes of the triaxial Galactic bulge; the longest axis to the $x'$ direction and shortest to the $z'$ direction. The projected velocity dispersions are then computed by $\sigma_x'^2 = \sigma_x'^2 \cos^2 \theta + \sigma_y'^2 \sin^2 \theta$, $\sigma_y'^2 = \sigma_y'^2 \sin^2 \theta + \sigma_y'^2 \cos^2 \theta$, and $\sigma_z = \sigma_z'$, where $\theta \sim 20^\circ$ is the angle at which one views the major axis of the bulge. The resulting values are $(\sigma_x, \sigma_y, \sigma_z) = (113.0, 93.0, 78.6) \text{ km s}^{-1}$. We assumed a nonrotating bulge, but see Blum (1995).

### 3. Mass Spectrum

We have so far described the models of velocities, $\mathbf{v}$, and lens locations, $D \equiv D_{ol}D_{ls}/D_{os}$, which are the two parameters required for the complete determination of the time scale $t_e$ for a given mass $m$. The next step is to model mass spectrum of lenses. We test three mass functions. In the first model, we assume that $\log m$ is Gaussian distributed:

$$f(\log m) = \frac{1}{\sqrt{2\pi}\sigma_{\log m}^2} \exp \left[-\frac{(\log m - \langle \log m \rangle)^2}{2\sigma_{\log m}^2}\right], \quad (3.1)$$

where $m$ is expressed in units of $M_\odot$. The two free parameters in this mass function are the mean and standard deviation of the logarithmic mass: $\langle \log m \rangle$ and $\sigma_{\log m}$. We also test the standard power-law mass function,

$$f(m) = Km^p \Theta(m - m_{cut}), \quad (3.2)$$

where $\Theta$ is the Heaviside step function, and $K \equiv (p + 1)m_{cut}^{-p-1}$. Two different expressions for the mass function $f(m)$ and $f(\log m)$ are related by $f(m)dm =$
f(\log m) d\log m$. In addition, the mass model based on observations with the HST (hereafter referred to as the Hubble model) is tested and compared with the power-law and Gaussian mass functions. The details of the Hubble model are described in §4.

First, we compute the distribution of time scales $f(t'_e)$ for a fixed value of mass (e.g. $1 \, M_\odot$). To find $f(t'_e)$ one should weight by the transverse speed $v$ and the cross-section (i.e. Einstein ring radius $r_e$). This is because events with faster transverse speeds and larger cross-sections are more likely to occur. Then the distribution function of $t'_e$ is computed by

$$f(t'_e) = \xi \int_0^{d_{max}} dD_{os} n(D_{os}) \int_0^{D_{cl}} dD_{cl} \rho(D_{cl}) D^{1/2} \tag{3.3}$$

$$\times \int dv_y \int dv_z v f(v_y, v_z) \delta \left( t'_e - \sqrt{4GM_\odot D/cv} \right),$$

where $D \equiv D_{cl}D_{ls}/D_{os}$, and $\xi$ is a normalization factor to be discussed below. We assume that the bulge is cut off at 4 kpc from the Galactic center, and therefore $d_{max} = 12$ kpc. The increase of volume element, and thus increase in total number of stars in the volume is assumed to be compensated by the decrease of observable stars due to decreasing detectability with distance. This is equivalent to the parameter $\beta = -1$ model of Kiraga & Paczyński (1994). We take disk self-lensing events into consideration by setting the lower limit of source stars to be 0 in equation (3.3).

At the second stage, we compute the actual time scale distribution $f(t_e)$ of MACHOs whose masses are distributed by a function $g(m)$. The distribution $f(t_e)$ is obtained by convolving two functions $f(t'_e)$ and $g(m)$:

$$f(t_e) = \eta(t_e) \int_{m_{cut}}^{m_{up}} dm \, m^{1/2} g(m) \int dt'_e f(t'_e) \delta(t_e - t'_e \sqrt{m}), \tag{3.4}$$

where $\eta(t_e)$ is the detection efficiency, and the factor $m^{1/2}$ is included to weight
events by their cross section which is $r_e \propto m^{1/2}$. In the computation of $f(t_e)$, we assume an upper mass limit $m_{up} = 10 \, M_\odot$. The distribution of observed time scale, $f_{obs}(t_e)$, is obtained from data measured by MACHO (Bennett et al. 1994) and OGLE (Udalski et al. 1994) groups. Out of 57 events detected, 45 by MACHO and 12 by OGLE, our data set includes all of the time scales except the event caused by a binary lens. Since the binary lensing event was detected by both groups, the total number of time scales is $N_{tot} = 55$.

Using the assumed distributions of $v$, $D$, and $m$, we find the best-fitting parameters of the mass functions; $(\log m)$ and $\sigma_{\log m}$ for the Gaussian and $m_{cut}$ and $p$ for the power-law mass spectrum. The best-fitting mass function is obtained by comparing model and observed time scale distributions. For this, we use the maximum likelihood test in which the statistic, $\ln L$, is computed by

$$\ln L = \sum_{i=1}^{N_{tot}} \ln f(t_{e,obs,i}).$$

(3.5)

Prior to the computation of $\ln L$, the test distribution $f(t_e)$ constructed from equations (3.3) and (3.4) is corrected by the detection efficiency $\eta(t_e)$ using the functions provided by Bennett et al. (1994) for MACHO and Zhao, Spergel, & Rich (1994) for OGLE. Since the efficiencies for each group are different, the likelihood should be computed separately and summed:

$$\ln L = \sum_{j} \sum_{i=1}^{N_j} \ln f_j(t_{e,obs,i}),$$

(3.6)

where the subscripts $j = 1, 2$ represent values of MACHO and OGLE, respectively. In this test, one determines the best-fitting distribution that maximizes $\ln L$. The likelihood statistic is related to the uncertainty by $\Delta(\ln L) = \sigma^2/2$. Since our concern is finding best-fitting mass spectrum not the total amount of optical depth or frequency, we leave the overall normalization $\xi$ as a variable so that the expected number of events matches with that of the observed events.
4. Result and Discussion

The results of the ln L computation in parameter space are shown as contour maps in Figure 1 for power-law (upper panel) and Gaussian (lower panel) mass functions, respectively. In the figure, the best-fitting positions in parameter space are marked with ‘x’ and the contours are drawn at $1\sigma$, $2\sigma$, and $3\sigma$ levels. For the Gaussian mass function, we find the best-fitting parameters of $\langle \log m \rangle = -0.79$ and $\sigma_{\log m} = 0.44$. The best-fitting parameters for the power-law distribution are $m_{\text{cut}} = 0.1 \, M_\odot$ and $p = -2.3$. Therefore, most of the lenses are more massive than hydrogen-burning mass limit with the power-law mass function. By contrast, a significant fraction of lenses are expected to be brown dwarfs ($m < 1 \, M_\odot$) under the Gaussian mass function.

Gould, Bahcall, & Flynn (1995) used HST to measure the luminosity function of local Galactic disk stars. They then applied mass-luminosity relation from Henry & McCarthy (1990) to obtain a local mass function. Even though there are difficulties in comparing the mass functions determined from MACHOs and HST observation because MACHOs are a mixture of both disk and bulge populations, it is interesting to apply the Hubble mass function to the MACHO data. The Hubble mass function is shown and compared with the best-fitting power-law and Gaussian mass functions in Figure 2. In the Hubble model, we have included the white dwarf (WD) population which is modeled based on the observation of McMahan (1989). The mass function of WD is normalized using the mass density $\rho_{WD} \sim 0.005 \, M_\odot \, \text{pc}^{-3}$ (Bahcall 1984). Luminous massive stars would have a low scale height and thus only nearby objects could contribute to microlensing events seen toward Baade’s window, $b \sim -4^\circ$. Then it is expected one would detect these objects due to both their closeness and high luminosity. However, there are no events in which the lens is bright enough to be easily detected. Therefore, we cut off the Hubble mass function at the upper limit of $\sim 2.2 \, M_\odot$.

The best-fitting time scale distributions $f(t_e)$ for the three mass functions are shown in Figure 3 and are compared with the observed distribution represented by
histograms in each figure. The power-law mass function fits the observation better than both the Gaussian and Hubble mass functions. Both observed and best-fitting $f(t_e)$ are magnified and shown in the smaller boxes inside each panel so that the differences between observation and the best fits of each mass functions can be compared more easily in the long-event region. The power-law fit is better than the Gaussian by $\sim 2.8 \sigma$ and than Hubble by $\sim 4.2 \sigma$. However, the determined distributions do not match well with the observations for very long events ($t_e \geq 70$ days) regardless of the assumed mass functions. The expected number of events with $t_e > 70$ days are 0.825 and 1.10, for the power-law and Hubble mass functions, respectively. Thus, the Poisson probabilities of observing 4 or more long events are $\sim 1\%$ for the power-law and $\sim 2\%$ for the Hubble mass function. The probability is much lower for the Gaussian. Even though only 4 events have time scale longer than 70 days, these long events are important because they are responsible for the significant fraction of optical depth measured toward the bulge.

Long time-scale events can be produced under various conditions. First, the long events might be caused by a mass function in which a larger fraction of lenses are in the higher mass range. If the mass function is a power-law, this can be achieved by a higher $m_{\text{cut}}$ and lower $p$. In Figure 4, the expected time-scale distributions for various lower values of power (upper panel) and higher mass cut-off (lower panel) are shown and compared to that of best-fitting distribution. However, as one or both of the conditions are satisfied to explain the long events, the fit deviates seriously in the low mass region where the majority of events are located. Therefore, the observed bimodal pattern of $f_{\text{obs}}(t_e)$ is not likely to be due to a lower power or higher mass cut-off than our determination. Another explanation would be a MACHO population with low transverse speed. For example, if the velocity dispersion of bulge population is very low, a MACHO will cross slowly over the Einstein ring, thus producing very long events. However, this explanation does not seem to be plausible because the velocity dispersion of the bulge is observationally well constrained. Another possibility is that longer events might be caused if a lens is composed of a binary system which is so closely-spaced that
one would detect it as a single lens instead of a binary lens. If the fraction of close binary systems is significant, the observed time scale distribution would be significantly different from that expected with an assumed single star mass distribution. However, the fraction of binary stars whose separations are less than 2 AU, which is a typical Einstein ring radius, is < 25% for G type stars in the solar neighborhood (Duquennoy & Mayor 1991). Therefore, in the computation we do not take into consideration the modification of the time scale distribution due to binary systems. Finally, another explanation would be a bimodal mass distribution with the second population composed of heavy dark MACHOs (e.g. black holes, neutron stars, or white dwarfs) with low transverse velocity. If for example there was a kinematically cold population of these objects, they would have a low scale height and hence would be observed mostly near the sun for sources near Baade’s window, i.e. a few degrees from the Galactic plane. Such a population would then have a very low transverse speed.

It is curious that the best-fitting power-law matches well with that of the Salpeter mass function valid in the mass range $M > 1 \, M_\odot$. Determination of the mass function in the solar neighborhood by Miller & Scalo (1979) indicated that the increase in the number of stars becomes shallower (decreasing $p$) at lower mass than the classical Salpeter estimate of $p \sim -2.35$ for stars with $m > 1 \, M_\odot$. Their estimate is $p \sim -1.4$ in the mass range $0.1 \, M_\odot \leq m \leq 1 \, M_\odot$. By contrast, our determined power $p = -2.3$ coincides with the classical value, and would seem to imply that the classical Saipeter mass function extends to lower mass objects. Note, however, that the Hubble function, in which the number of stars decreases as mass decreases after passing the maxima at $m \sim 0.5 \, M_\odot$, is inconsistent with such a power law.

**Acknowledgement:** We would like to thank M. Everett for making very helpful comments and suggestions. This work is supported by a grant AST 9420746 from the NSF.
REFERENCES

10. Evans, N. W. 1994, preprint