Mikrosoczewkowanie grawitacyjne

P. Rozwadowski[1]

Uniwersytet Warszawski

[1]Uniwersytet Warszawski, Obserwatorium Astronomiczne, Warszawa, Al. Ujazdowskie 4

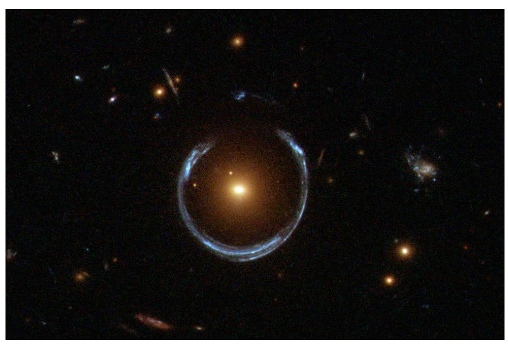
Streszczenie

W poniższej pracy przedstawiono zjawisko mikrosoczewkowania grawitacyjnego oraz metodę wyznaczania parametrów krzywej paczyńskiego i dopasowanie jej do obserwowanych zmian blasku. Posłużono się i porównano algorytmy Levenberga-Marquardta, BFGS oraz Powell'a dla wybranych soczewek z projektu OGLE z 2015 roku.

Słowa kluczowe: soczewkowanie grawitacyjne, optymalizacja, OGLE

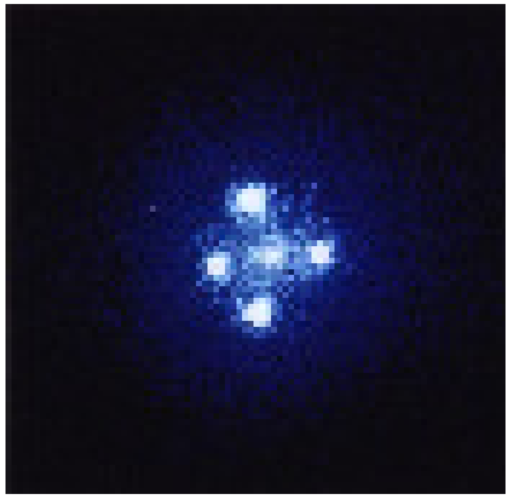
1. Wstęp

Zgodnie z ogólną teorią względności Einstein'a (OTW) masywne ciała powodują zakrzywienie czasoprzestrzeni, a więc także zakrzywienie promieni świetlnych (poruszających się po liniach geodezyjnych - tj. najkrótszych drogach pomiędzy dwoma punktami w przestrzeni metrycznej). Soczewką może być zarówno gwiazda, czarna dziura jak i cała gromada galaktyk. Niewielkie zniekształcenie i przesunięcie obrazu w wyniku tego zjawiska nosi nazwę słabego soczewkowania. Z silnym soczewkowaniem mamy do czynienia, gdy obraz jest znacznie zniekszałcony - przybiera obraz łuku (Pierścień Einsteina) bądź powstaje obraz wielokrotny (Krzyż Einsteina).W poniższej pracy przeanalizowano zagadnienie mikrosoczewkowania czyli zakrzywienia przestrzeni przez obiekt rozmiaru gwiazdy.



Rys. 1.— LRG 3-757 - pierścień Einsteina

OGLE (The Optical Gravitational Lensing Experiment - Eksperyment Soczewkowania Grawitacyjnego) to projekt naukowy realizowany przez naukowców z Obserwatorium Astronomicznego Uniwersytetu Warszawskiego, którego celem jest detekcja zjawisk mikrosoczewkowania grawitacyjnego. Badania prowadzone są za pomocą polskiego teleskopu w Las Campanas w Chile. Do sukcesów projektu należą m.in. odkrycia pierwszych mikrosoczewek w kierunku centrum Drogi Mlecznej, odkrycia pozasłonecznych układów planetarnych czy też oszacowania liczby obiektów MACHO. W części czwartej projektu rejestruje się do 2000 zjawisk mikrosoczewkowania rocznie.



Rys. 2.— Obraz kwazara Q2237+030 - Krzyż Einsteina

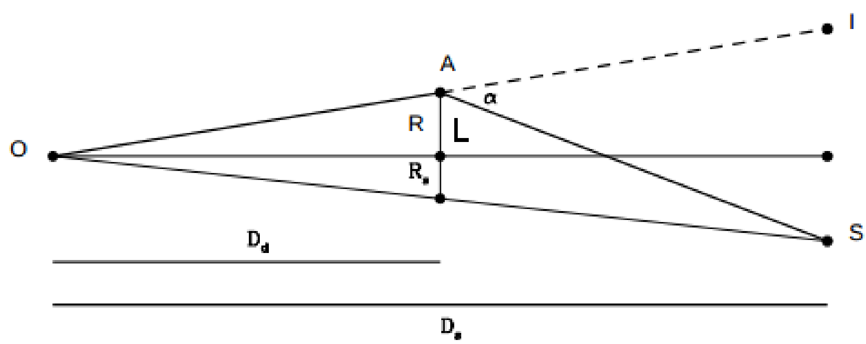
W rozdziale drugim omówiono mikrosoczewkowanie przez punktowe (dla obserwatora) masy. W rozdziale trzecim opisano trzy wybrane algorytmy dopasowywania modelu do danych obserwacyjnych. Wyniki dopasowania przedstawiono w rozdziale czwartym.

2. Mikrosoczewkowanie przez punktowe soczewki

Odległe soczewki są dla obserwatora punktowymi masami. Aby doszło do zjawiska na jednej linii musi znaleźć się obserwator (O), soczewka (L) oraz obiekt soczewkowany (S). Schemat geometryczny mikrosoczewkowania pokazano na rysunku 3. Promienie świetlne są ugięte o kąt α. Obserwator widzi obraz pozorny (I) zamiast rzeczywistego (S). Zgodnie z OTW kąt α wyraża się wzorem

$$\alpha = \frac{4GM}{Rc^2}, \tag{1}$$

gdzie G to stała grawitacji, M to masa soczewki, c to prędkość światła. Z perspektywy obserwatora pojawiają się dwa pozorne obrazy I_1 i I_2 (rysunek 3). W trakcie trwania zjawiska obiekty S, L i O przemieszczają się względem siebie przez co obraz pozorny zmienia swoją jasność w zależności od odległości kątowej od soczewki (rysunek 5). Całość trwa od kilku godzin do kilkuset dni w zależności od ułożenia się względem siebie obiektów. Typowe odległości między S oraz L są rzędu milisekund łuku co przy obecnych metodach jest niemożliwe do bezpośredniego zaobserwowania. Jednak możliwe jest obserwowanie wzmocnienia jasności a następnie pociemnienie obiektu soczewkowanego.



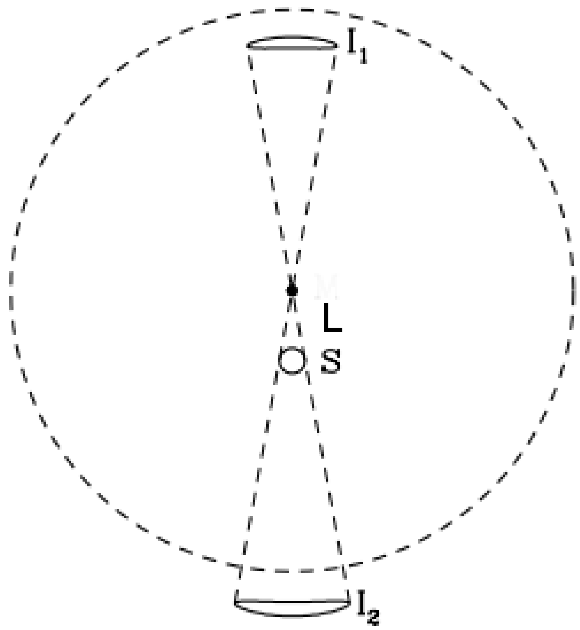
Rys. 3.— Schemat soczewkowania grawitacyjnego. O - obserwator, L - soczewka, S - obiekt soczewkowany, I - obraz pozorny.

Do opisu zmian jasności blasku można posłużyć się krzywą zaproponowaną przez Bohdana Paczyńskiego. Jest to funkcja jasności obserwowanej w zależności od czasu

$$m = m_{baseline} - 2.5 \log_{10}(f_s(A(u) - 1) + 1). \tag{2}$$

Zawiera ona pięć niezależnych parametrów, które należy dopasować do obserwacji. Pierwszym z nich jest jasność obiektu, kiedy nie zachodzi zjawisko mikroczewkowania $m_{baseline}$. f_s to stosunek jasności soczewki do całkowitej jasności. A jest funkcją u i mówi o powiększeniu jasności:

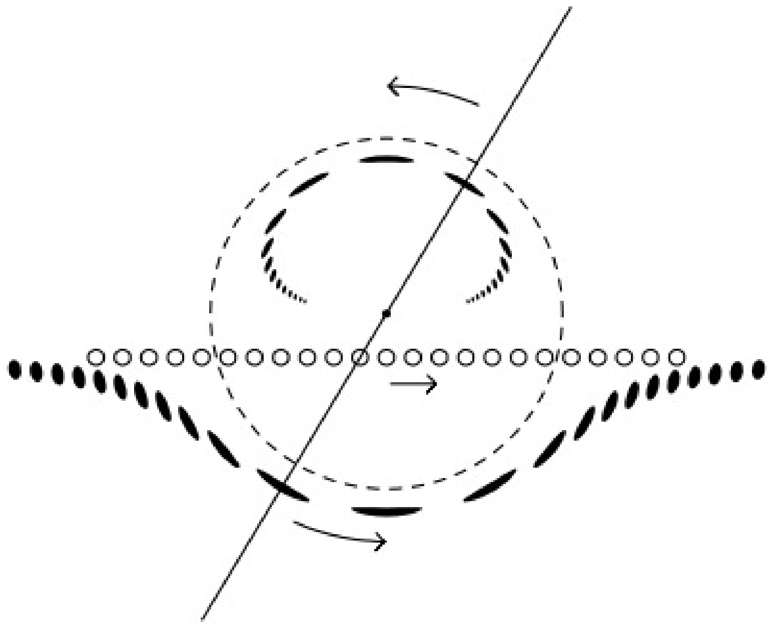
$$A(u) = \frac{u^2 + 2}{u(u^2 + 4)^{\frac{1}{2}}} \tag{3}$$



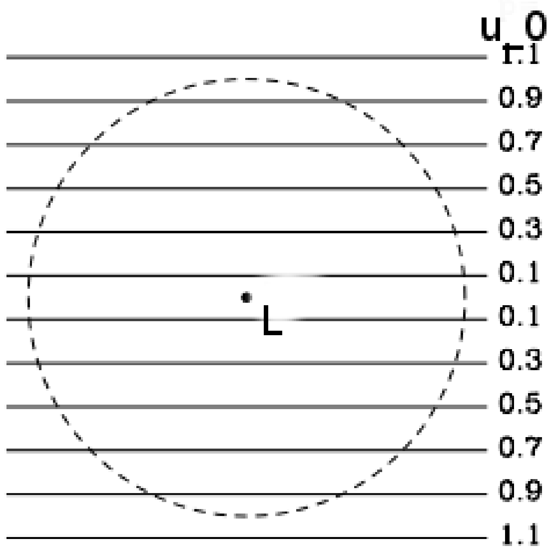
Rys. 4.— Obraz soczewkowania grawitacyjnego z perspektywy obserwatora. L - soczewka, S - obiekt soczewkowany (niewidoczny), I_1 oraz I_2 - obrazy pozorne. Przerywaną kreską zaznaczono promień Einsteina. Obraz I_1 znajduje się w jego wnętrzu, a I_2 na zewnątrz.

Natomiast u wyraża się poprzez u_0 - bezwymiarowy parametr zderzenia mówiący o najmniejszej kątowej odległości między L i S, t_0 - czas maksymalnego pojaśnienia, t_E - skala czasowa zjawiska. Na rysunku 6 przedstawiono przykładowe odległości kątowe względem soczewki, a na rysunku 7 zmiany jasności tym wywołane.

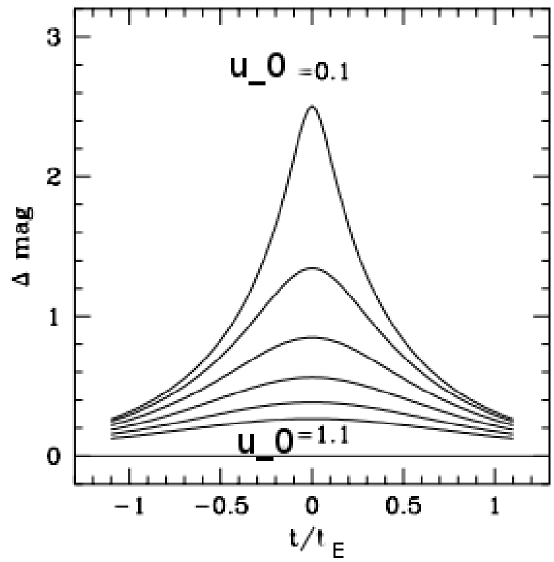
$$u = \sqrt{u_0^2 + \Big(\frac{t - t_0}{t_E}\Big)^2} \tag{4}$$



Rys. 5.— Obraz soczewkowania grawitacyjnego z perspektywy obserwatora w różnych momentach trwania zjawiska.



Rys. 6.— Przykładowe odległości u_0 zględem soczewki L.



Rys. 7.— Przykładowe powiększenia jasności w zależności od parametru u_0.

3. Dopasowanie parametrów krzywej do obserwacji

Głównym zadaniem tej pracy jest dopasowanie punktów pomiarowych (czasu obserwacji i jasności obserwowanej) do wyżej wymienionej krzywej. Poszukujemy pięciu parametrów $m_{baseline}$, f_s, u_0, t_0, t_E, której najlepiej opisują zjawisko. W tym celu posłużymy się trzema algorytmami optymalizacyjnymi: Levenberga-Marquardta, BFGS oraz Powell'a. Przy czym skupimy się głównie na pierwszym ze względu na jego popularność, szybkość działania i to, że daje on bardzo zadowalające rezultaty. Program do dopasowywania krzywej zaimplementowano w język programowania Python.

3.1. Metoda Levenberga-Marquardta

Jest to iteracyjny algorytm łączący metodę największego spadku oraz metodę Gaussa-Newtona. Celem algorytmu jest znalezienie takiego wektora parametrów β, który minimalizuje funkcjonał

$$S(\beta) = \Sigma_{i=1}^{m}\Big[y_i - f(x_i, \beta)\Big]^2, \tag{5}$$

mając zadane punkty pomiarowe (x_i, y_i). Na początek należy wybrać wektor startowy β. W każdym kroku iteracyjnym wektor β zastępujemy przez kolejne przybliżenie $\beta + \delta$. W celu wyznaczenia δ funkcję $f(x_i, \beta + \delta)$ przybliżamy przez

$$f(x_i, \beta + \delta) \approx f(x_i, \beta) + J_i\delta, \tag{6}$$

gdzie J_i jest gradientem względem β

$$J_i = \frac{\partial f(x_i, \beta)}{\partial \beta} \tag{7}$$

W minimum $S(\beta)$ gradient S względem δ będzie równy zeru. Tak więc w pierwszym przybliżeniu f mamy

$$S(\beta + \delta) \approx \Sigma_{i=1}^{m}\Big(y_i - f(x_i, \beta) - J_i\delta\Big)^2, \tag{8}$$

Biorąc pochodną po δ otrzymujemy

$$(J^T J)\delta = J^T[y - f(\beta)], \tag{9}$$

gdzie J to macierz Jacobiego gdzie i-ty wiersz to J_i, f to $f(x_i, \beta)$, a y to y_i. Równanie to pozwala obliczyć δ. Levenberg zaproponował by zmodyfikować ten wzór

$$(J^T J + \lambda I)\delta = J^T[y - f(\beta)], \tag{10}$$

gdzie I to macierz jednostkowa. Parametr λ koryguje każdą iterację. Jeśli spadek S jest szybki to metoda jest zbliżona do Gaussa-Newtona, podczas gdy iteracje dają niesatysfakcjonującą zbieżność λ jest zwiększana dając metodę zbliżoną do metody najszybszego spadku. W przypadku dużej wartości λ $J^T J$ nie jest używane w ogóle. W tym celu Marquardt zmodyfikował formułę do postaci

$$(J^T J + \lambda \mathrm{diag}(J^T J))\delta = J^T[y - f(\beta)]. \tag{11}$$

3.2. Metoda BFGS

Algorytm Broyden-Fletcher-Goldfarb-Shanno jest metodą quasi newtonowską pozwalającą zminimalizować S. Jest ona uogólnieniem metody siecznych na przypadek wielowymiarowy. S rozwijamy w szereg Taylora do drugiego stopnia

$$S(\beta+\delta) \approx f(\beta) + \nabla f(\beta)^T \delta + \frac{1}{2}\delta^T H \delta, \tag{12}$$

gdzie ∇f oznacza gradient, a H Hesjan. Z rozwinięcia gradientu otrzymujemy pierwszy krok

$$\delta = -H^{-1}\nabla f(\beta). \tag{13}$$

Kolejny hesjan w tej metodzie obliczany jest iteracyjnie według wzoru

$$H = H_{old} + \frac{yy^T}{y^T\delta} - \frac{H_{old}\delta(H_{old}\delta)^T}{\delta^T H_{old}\delta}. \tag{14}$$

Oraz w kolejnych krokach $\beta = \beta_{old} + \delta$.

3.3. Metoda Powella

Algorytm Powella jest bezgradientową metodą bezpośredniego poszukiwania minimum funkcji. Polega ona na tworzeniu nowych kierunków poszukiwania minimum na podstawie poprzednich. Na początek wybierany jest punkt początkowy β oraz zbiór N (N =dim(β)) liniowo niezależnych kierunków poszukiwań s_i, na przykład w bazie kanonicznej e_i. Następnie poruszając się wzdłuż s_1 znajdujemy punkt y_1 i z niego szukamy minimum wśród pozostałych kierunków by potem znaleźć y_2 wzdłuż kolejnego s_1. Otrzymujemy nowy kierunek sprzężony do s_1: $d = y_2$-y_1. Jeśli $||d||$ jest dostatecznie małe lub nowo powstałe kierunki s_1,s_2,...,s_{N-1},d są liniowo zależne to algorytm kończy pracę. W przeciwnym wypadku $s_j = s_{j-1}$ oraz $s_1 = \frac{d}{||d||}$ i ponownie szukamy minimum.

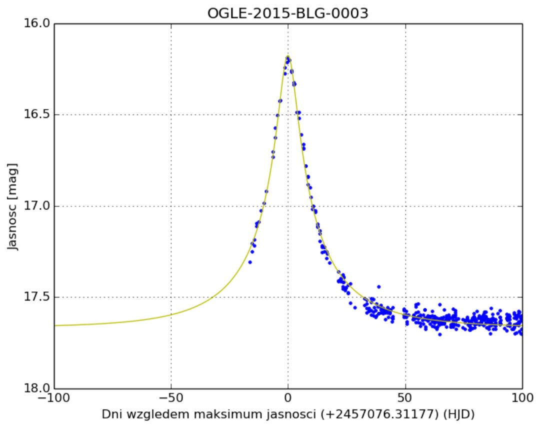
4. Wyniki

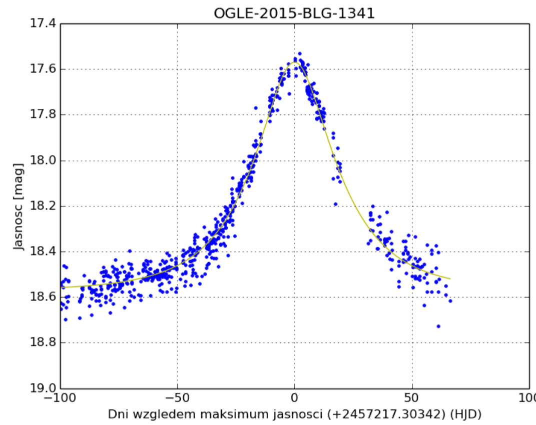
W tabeli 1 podano wyniki dopasowania pięciu przykładowych zjawisk do modelu metodą Levenberga-Marquardta oraz dwóch metodą BFGS oraz Powella. Na rysunkach poniżej przedstawiono obserwacje wraz z naniesionym, dopasowanym modelem metodą Levenberga-Marquardta.

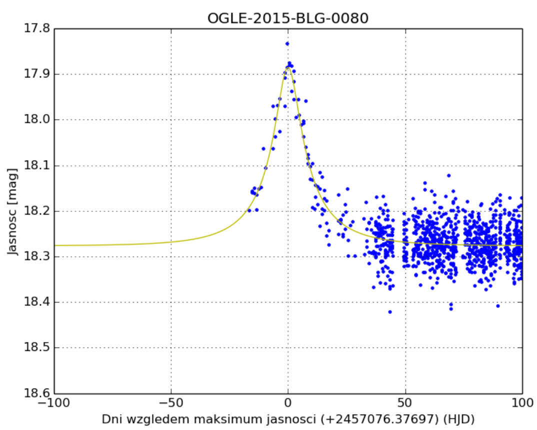
Tabela 1: Dopasowane parametry krzywej dla wybranych zjawisk z projektu OGLE

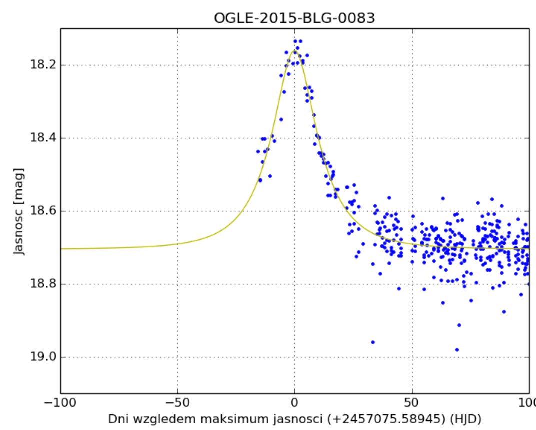
OGLE-2015-BLG-0003	Metoda Levenberga-Marquardta	+/−	Metoda BFGS	Metoda Powella
$m_{baseline}$	17.66745	0.00060	17.67	17.67
f_s	0.304	0.020	0.30	0.30
u_0	0.0938	0.0064	0.094	0.094
t_0	2.4571e+06	0.080	2.4571e+06	2.4571e+06
t_E	40.98	1.87	40.90	40.90
OGLE-2015-BLG-1341	Metoda Levenberga-Marquardta	+/−	Metoda BFGS	Metoda Powella
$m_{baseline}$	18.5735	0.0010	18.57	18.57
f_s	0.99	0.17	1.00	0.96
u_0	0.423	0.051	0.42	0.41
t_0	2.4572e+06	0.18	2.4572e+06	2.4572e+06
t_E	31.88	2.25	31.88	32.42

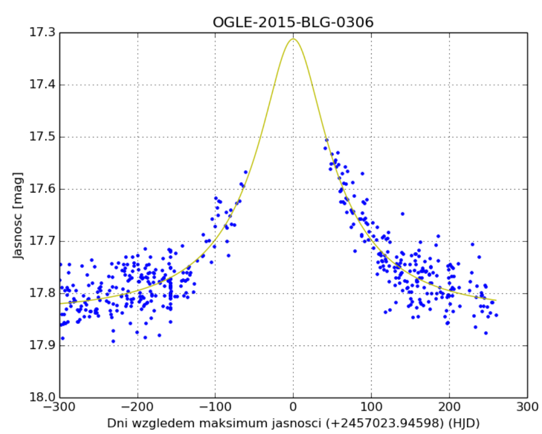
OGLE-2015-BLG-0080	Metoda Levenberga-Marquardta	+/−
$m_{baseline}$	18.27705	0.00044
f_s	0.102	0.036
u_0	0.194	0.060
t_0	2.4571e+06	0.33
t_E	27.18	5.57
OGLE-2015-BLG-0083	Metoda Levenberga-Marquardta	+/−
$m_{baseline}$	18.70577	0.00089
f_s	0.42	0.23
u_0	0.41	0.16
t_0	2.4571e+06	0.37
t_E	20.49	5.04
OGLE-2015-BLG-0306	Metoda Levenberga-Marquardta	+/−
$m_{baseline}$	17.8347	0.0013
f_s	0.177	0.064
u_0	0.227	0.107
t_0	2.4570e+06	1.56
t_E	162.40	27.43


OGLE-2015-BLG-0003
Jasnosc [mag]
Dni wzgledem maksimum jasnosci (+2457076.31177) (HJD)


OGLE-2015-BLG-1341
Jasnosc [mag]
Dni wzgledem maksimum jasnosci (+2457217.30342) (HJD)


OGLE-2015-BLG-0080
Jasnosc [mag]
Dni wzgledem maksimum jasnosci (+2457076.37697) (HJD)


OGLE-2015-BLG-0083
Jasnosc [mag]
Dni wzgledem maksimum jasnosci (+2457075.58945) (HJD)


OGLE-2015-BLG-0306
Jasnosc [mag]
Dni wzgledem maksimum jasnosci (+2457023.94598) (HJD)

5. Podsumowanie

W powyższej pracy przedstawiono zjawisko mikrosoczewkowania grawitacyjnego oraz sposoby dopasowania obserwacji do modelu punktowej soczewki. Spośród trzech algorytmów (Levenberga-Marquardta, BFGS, Powella) najwygodniejszą i najefektywniejszą jest metoda Levenberga-Marquardta. Jest ona szybko zbieżna i najlepiej oszacowuje pożądane parametry.

Rozwijanie metod obserwacji mikrosoczewek w przyszłości pomoże zwiększyć ilość odkrytych planet pozasłonecznych. Poza tym jest ono bardzo obiecującym narzędziem w badaniu ciemnej materii, a także badania struktury galaktyk.

Bibliografia

Paczyński, B. "Gravitational microlensing in the local group" (1996)

Schneider, P., Kochanek, C., Wambsganss, J., "Gravitational Lensing: Strong, Weak and Micro"

Nocedal, Jorge; Wright, Stephen J. (2006). Numerical Optimization, 2nd Edition.

Shanno, David F. (July 1970), "Conditioning of quasi-Newton methods for function minimization", Math. Comput. 24 (111): 647–656

Powell, M. J. D. (1964). "An efficient method for finding the minimum of a function of several variables without calculating derivatives". Computer Journal 7 (2): 155–162