

TECHNISCHE
UNIVERSITÄT
DRESDEN

Analyse lückenhafter, verrauschter Zeitreihen mithilfe der Bayes-Statistik

Bachelor-Arbeit
zur Erlangung des Hochschulgrades
Bachelor of Science
im Bachelor-Studiengang Physik

vorgelegt von

Maximilian Arthus Schanner
geboren am 18.05.1993 in Wolfhagen

Institut für theoretische Physik
Fachrichtung Physik
Fakultät Mathematik und Naturwissenschaften
Technische Universität Dresden
2014

Eingereicht am 30. Juni 2014

1. Gutachter: Prof. Dr. B. Kämpfer
2. Gutachter: Prof. Dr. M. Soffel

Zusammenfassung

Das MCMC-Verfahren der Bayes-Statistik ist eine moderne Methode, nichtlineare Optimierungen vorzunehmen. In der Exoplanetensuche existieren verschiedene Varianten dieser Methode, welche auch für Asteroidenlichtkurven vielversprechende Möglichkeiten eröffnen. In der vorliegenden Arbeit wurde ein MCMC-Verfahren für Asteroidenlichtkurven untersucht und mit der herkömmlichen Methode nach Harris et al. verglichen. Obwohl das neue Verfahren in der Lage ist, die Rotationsperiode zu bestimmen, ist es der herkömmlichen Methode in Genauigkeit, Geschwindigkeit und Anwenderfreundlichkeit unterlegen.

Abstract

Markov-Chain Monte-Carlo is a modern way for nonlinear regression derived from Bayesian statistics. The search for exoplanets provides several applications of mcmc, which are also promising in the field of asteroid lightcurves. During this work a modified version of mcmc was tested and compared to the well known method from Harris et al. While the new algorithm is able to detect the rotation period, it's inferior in accuracy, speed and handiness, in comparison to Harris' method.

Inhaltsverzeichnis

1	Physikalische Situation	3
1.1	Lichtkurvendaten	3
2	Theoretische Grundlagen	7
2.1	Mathematisches Modell	7
2.2	Analysemethoden	8
2.2.1	Methode der kleinsten Quadrate	9
2.2.2	MCMC-Methode	10
2.2.3	Modellvergleich	14
3	Entwicklung eines Analyseverfahrens für Asteroidenlichtkurven	17
3.1	Simulierte Daten	17
3.2	MCMC-Verfahren	18
3.2.1	Optimierung des Algorithmus	20
4	Analyse realer Lichtkurven und Vergleich mit dem Verfahren nach Harris	29
5	Zusammenfassung	35
A	Literaturverzeichnis	39
	Abbildungsverzeichnis	41

Einleitung

Die Analyse von Zeitreihen unterschiedlichster Art spielt in vielen Bereichen der Forschung eine Rolle. In der Astronomie stellen die Lichtkurven einen großen Teil der Zeitreihen dar. Aus diesen kann eine Vielzahl von Informationen gewonnen werden, obwohl sie lediglich eine Helligkeitsveränderung über der Zeit abbilden. Das Ziel dieser Arbeit ist die Analyse von Asteroidenlichtkurven zur Bestimmung der Rotationsperiode des erzeugenden Asteroiden. Dazu soll ein neues Verfahren untersucht werden, welches mit der aktuellen Standardmethode nach Harris et al.^[4] zu vergleichen ist.

Während das Standardverfahren die Methode der kleinsten Quadrate zur Analyse verwendet, wird das neue Verfahren im Kontext der Bayes-Statistik entworfen. Vor allem im Bereich der Exoplanetensuche, allerdings ebenfalls in vielen anderen, auch nicht technischen Wissenschaften, wurden in den letzten Jahren verschiedene Bayes-statistische Analyseverfahren entwickelt und erfolgreich angewendet. Am vielversprechendsten für die Zwecke dieser Arbeit ist die MCMC-Methode (engl.: Markov-Chain Monte-Carlo). Ein zentraler Unterschied zum Verfahren nach Harris et al. ist, dass keine Linearisierungen vorgenommen werden müssen.

Zunächst wird in Kapitel 1 die physikalische Situation erläutert. In Kapitel 2 wird kurz auf die Standardmethode eingegangen, um anschließend einen detaillierten Einblick in die MCMC-Methode zu geben. Die speziellen Probleme, welche bei der Analyse von Asteroidenlichtkurven auftraten, werden in Kapitel 3 aufgezeigt und Modifikationen am MCMC-Verfahren vorgestellt, um diese zu lösen. Abschließend werden in Kapitel 4 beide Verfahren anhand der Analyse von 2 Asteroidenlichtkurven verglichen.

1 Physikalische Situation

Hier soll kurz die vorliegende physikalische Situation beschrieben und dargestellt werden, wie daraus die analysierten Daten gewonnen werden.

Neben Planeten und deren Monde findet man im Sonnensystem eine Vielzahl an Kleinplaneten, zu denen auch die Asteroiden zählen. Beobachtet man einen Asteroiden durch ein Teleskop, erkennt man typischerweise einen Lichtpunkt. Verfolgt man diesen über längere Zeit, so bemerkt man eine Variation in der Helligkeit.^[1] Diese Variation kann auf drei Ursachen zurückgeführt werden:

- Die Entfernung zwischen Asteroid, Erde und Sonne verändert sich und damit die Intensität des reflektierten Lichts.
- Der Phasenwinkel zwischen Asteroid, Erde und Sonne verändert sich und damit der beleuchtete Bereich des Asteroiden.
- Der Asteroid rotiert und verändert damit die zugewandte Reflektionsfläche.

Das Ziel dieser Arbeit ist die Bestimmung der Rotationsperiode. Hierbei sind die übrigen Helligkeitsschwankungen störend und müssen korrigiert werden. Da die Rotationsperiode von Asteroiden üblicherweise im Bereich von $2\text{h} < P \leq 20\text{h}$ liegt, sich die Position des Asteroiden gegenüber Erde und Sonne jedoch deutlich langsamer ändert, entsteht eine Überlagerung der höherfrequenten Rotationskurve mit der niederfrequenten, für einzelne Beobachtungsnächte nahezu konstanten Lichtkurve der Bahnbewegung. Dies führt zu einem Versatz zwischen den Beobachtungsnächten. Eine beispielhafte Lichtkurve ist in Abbildung 1.1 zu sehen. Es existiert ein mathematisches Modell, welches die Entfernungs- und Phaseneinflüsse beschreibt und damit zur Korrektur der Lichtkurve verwendet werden kann (eine genauere Beschreibung erfolgt in Kapitel 2).

1.1 Lichtkurvendaten

Die in dieser Arbeit verwendeten Lichtkurven basieren auf Daten des Minor Planet Center (<http://www.minorplanetcenter.net>). Diese werden im ALCDE-Format (engl. Asteroid Lightcurve Data Exchange) bereitgestellt und liefern zusätzlich zu den eigentlichen Zeitreihen unter anderem auch die benötigten Informationen über die o.g. Entfernungen und den

Phasenwinkel des Asteroiden in der jeweiligen Beobachtungsnacht. Während der Phasenwinkel direkt in Grad entnommen wird, werden die Entfernungen zu Erde und Sonne bereits wie im mathematischen Modell benötigt als

$$\text{UCORMAG} = -5 \cdot \log(r\Delta) \quad (1.1)$$

angegeben. (r bezeichnet die Entfernung Asteroid→Erde und Δ die Entfernung Asteroid→Sonne.) Wie Entfernung und Phasenwinkel die gemessenen Helligkeiten beeinflussen wird mathematisch in Abschnitt 2.1 beschrieben. Hier ist außerdem anzumerken, dass die gemessenen Helligkeiten vom Minor Planet Center als differenzielle Helligkeiten angegeben werden. Diese sind Vergleichshelligkeiten zu von Nacht zu Nacht variierenden Referenzsternen. Dadurch entstehen weitere vertikale Versätze zwischen den einzelnen Beobachtungsnächten.

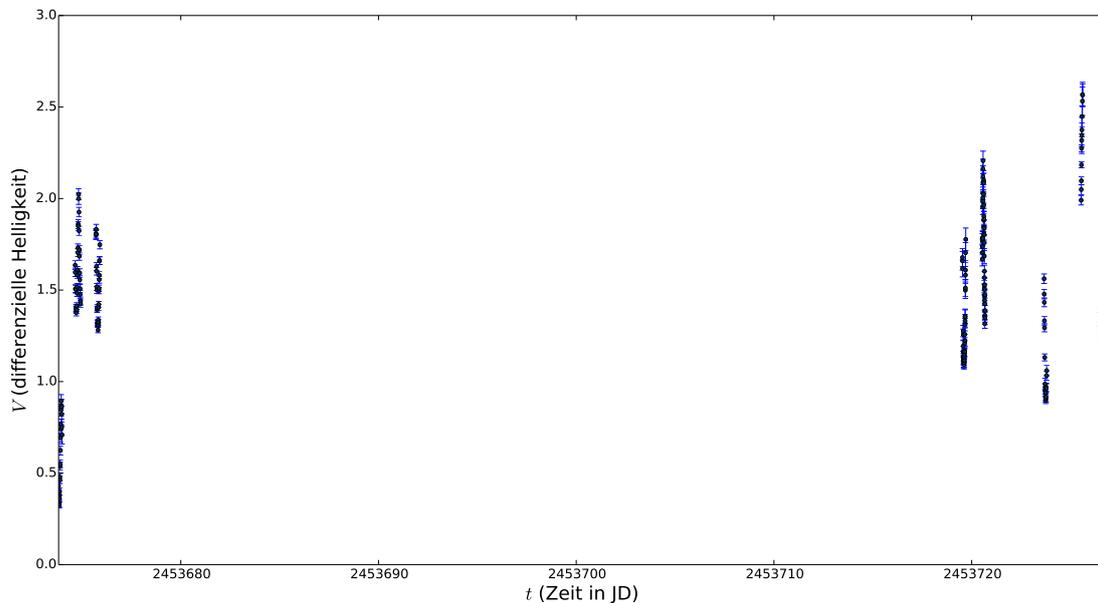


Abbildung 1.1: Beispielhafte Lichtkurve des Asteroiden (2839) Annette. Die Daten wurden aus der Lichtkurvendatenbank des Minor Planet Center entnommen. Man erkennt die beschriebenen vertikalen Versätze zwischen den einzelnen Beobachtungsnächten (Häufungen).

2 Theoretische Grundlagen

Die zu Grunde liegende Rotationsbewegung von Asteroiden erzeugt eine periodische Lichtkurve, die durch eine endliche Fourierreihe beschrieben werden kann. Die dabei definierten Parameter zur Beschreibung der Kurve werden in diesem Kapitel eingeführt. Anschließend werden die beiden in dieser Arbeit verwendeten Analysemethoden vorgestellt.

2.1 Mathematisches Modell

Zur Beschreibung der Rotationslichtkurve des Asteroiden wird angenommen, dieser befinde sich an einem festen Punkt im Raum und rotiere dort. Offensichtlich liegt diese Situation real nicht vor. Dennoch existiert in der Astronomie ein Parameter, die absolute Helligkeit H , welche die Helligkeit des Asteroiden in einer Entfernung von 1AU beim Phasenwinkel 0° angibt. Die Absolute Helligkeit kann nach R. Dymock^[2] und R. K. Buchheim^[3] wie folgt berechnet werden:

$$H = H(\alpha) + 2.5 \log[(1 - G)\Phi_1(\alpha) + G\Phi_2(\alpha)]$$

mit $\Phi_i(\alpha) = \exp\left[-A_i\left(\tan\left(\frac{\alpha}{2}\right)\right)^{B_i}\right]$

und $A_1 = 3.33, A_2 = 1.87, B_1 = 0.63, B_2 = 1.22$

G bezeichnet den Steigungsparameter (engl. slope parameter). Aus H kann durch Umstellen leicht die phasenabhängige Helligkeit $H(\alpha)$ gewonnen werden.

$$H(\alpha) = H - 2.5 \log[(1 - G)\Phi_1(\alpha) + G\Phi_2(\alpha)] \quad (2.1)$$

Die gemessenen Helligkeiten V setzen sich wie folgt zusammen:

$$V = H(\alpha) + 5 \log(r\Delta) + R(t) + \text{Fehler} \quad (2.2)$$

Der zweite Term auf der rechten Seite berücksichtigt die Entfernungen Asteroid→Erde r und Asteroid→Sonne Δ und wird üblicherweise von den gemessenen Helligkeiten subtrahiert, um die Distanz des Asteroiden auf eine astronomische Einheit zu korrigieren. Der dritte Term beschreibt die gesuchte Rotationslichtkurve $R(t)$. In realen Daten können zusätzliche weitere Versätze enthalten sein, bspw. durch variierende Standorte, unterschiedliche Filter etc. Um

diese zu berücksichtigen wird ein weiterer Term eingeführt, in welchem auch die absolute Helligkeit H verarbeitet wird. Jetzt kann das komplette mathematische Modell für ideale, d.h. fehlerfreie, Daten aufgestellt werden:

$$V_{\text{ideal}} - 5 \log(r\Delta) = \sum_{j=1}^n [\bar{H}_j \cdot h_j(t)] - 2.5 \log[(1 - G)\Phi_1(\alpha) + G\Phi_2(\alpha)] + \sum_{l=1}^m [A_l \sin(\omega \cdot l \cdot t) + B_l \cos(\omega \cdot l \cdot t)] \quad (2.3)$$

Hier und im Folgenden bezeichnen

- n die Zahl der Beobachtungsnächte,
- \bar{H}_j den verbleibenden Versatz in der j -ten Beobachtungsnacht inklusive der absoluten Helligkeit H ,
- $h_j(t) = \Theta(t_j - t)\Theta(t - t_{j-1})$,
- t_j das Ende (in Bezug auf den Beobachtungsbeginn) der j -ten Beobachtungsnacht; $t_0 = 0$,
- m die Ordnung der zur Beschreibung verwendeten Fourierreihe,
- A_l, B_l die Fourierkoeffizienten der Lichtkurve,
- ω die Frequenz der Fourierreihe (entspricht der Rotationsfrequenz des Asteroiden),
- $P = \frac{2\pi}{\omega}$ die Periode der Fourierreihe (bzw. der Asteroidenrotation).

Wie man sieht enthält das Modell $2m + n + 2$ freie Parameter, von denen $2m + n$ linear und zwei (G und die Frequenz) nichtlinear sind. α, r, Δ, n und die t_j sind keine freien Parameter, da sie durch die Beobachtung vorgegeben sind. Prinzipiell ist auch m selbst ein freier Parameter, sodass zu jedem gewählten m ein eigenes Modell und damit ein eigener Parameterraum gehört.

2.2 Analysemethoden

Die $2m + n + 2$ freien Parameter sollen so gewählt werden, dass sie eine gegebene Lichtkurve am besten beschreiben. Dazu benötigt jede Lösungsmethode ein Maß, welches angibt wie gut ein Modell mit einem Satz von Parametern die Daten wiedergibt. Außerdem sollte jede Methode eine Unsicherheit liefern, die angibt, wie signifikant die bestimmten besten Parameter sind. Wie die zwei in dieser Arbeit verglichenen Methoden versuchen, diese Anforderungen zu erfüllen, wird im Folgenden beschrieben.

2.2.1 Methode der kleinsten Quadrate

Die ursprünglich von Harris et al.^[4] vorgestellte Methode ist ein Standardverfahren zur Analyse von Lichtkurven.^[5] Als Approximationsmaß verwendet sie die kleinsten Quadrate. Betrachtet man eine Reihe von N Messdaten D_k zu den Zeiten t_k mit den Messunsicherheiten E_k , so liefert

$$\chi^2 = \sum_{k=1}^N \left(\frac{D_k - H(t_k)}{E_k} \right)^2 \quad (2.4)$$

ein Maß dafür, wie gut das Modell $H(t)$ mit M freien Parametern die Daten D_k wiedergibt. Ein kleiner Wert für χ^2 entspricht dabei einer guten Approximation. Prinzipiell könnte man für $H(t)$ das mathematische Modell (2.3) einsetzen und den Satz von Parametern suchen, welcher χ^2 minimiert. Für das so entstehende nichtlineare Problem existieren verschiedene Lösungsmöglichkeiten.^[6] Harris et al. gehen jedoch anders vor. Zunächst wird der log-Term des Modells (2.3) verarbeitet. Die deutlich niederfrequenteren Bahnbewegung kann bei der Bestimmung der Rotationskurve durch eine Stufenfunktion genähert werden. Damit entsteht pro Beobachtungsnacht ein weiterer vertikaler Versatz, der durch modifizierte \bar{H} -Terme behandelt werden kann. Hintergrund dieser Vereinfachung ist die Annahme eines über eine Nacht konstanten Phasenwinkels α . Damit entsteht ein neues Modell

$$V_{\text{ideal}} - 5 \log(r\Delta) = \sum_{j=1}^n [\bar{H}_j^* \cdot h_j(t)] + \sum_{l=1}^m [A_l \sin(\omega \cdot l \cdot t) + B_l \cos(\omega \cdot l \cdot t)] \quad (2.5)$$

Dieses Modell hat $M = 2m + n + 1$ freie Parameter. Setzt man nun verschiedene feste Ordnungen \tilde{m} und Frequenzen $\tilde{\omega}$ ein, so entsteht zu jedem Paar $\tilde{m}, \tilde{\omega}$ ein lineares Problem, welches sich sehr viel leichter lösen lässt, als das nichtlineare Problem von oben. (Dazu kann bspw. eine Singulärwertzerlegung verwendet werden.^[6]) Man bezeichnet das Minimum von χ^2 als unkorrigierte Varianz s^2 und $\mathfrak{s}^2 = s^2 / (N - M)$ als Bias-korrigierte Varianz. Als beste Wahl wird das Modell $M_{\tilde{m}, \tilde{\omega}}$ angesehen, welches den kleinsten Wert für s^2 liefert. Während die unkorrigierte Varianz mit zunehmender Ordnung m immer abnehmen wird, wird \mathfrak{s}^2 irgendwann anfangen zu wachsen. Dies zeigt an, dass die zusätzlichen Terme der Fourierreihe nicht mehr signifikant sind.^[4] Um eine Fehlerabschätzung für das gewählte $\tilde{\omega}$ anzugeben, wird die Differenz zwischen $\tilde{\omega}$ und der Frequenz berechnet, welche das zu $\tilde{\omega}$ gehörige s^2 um \mathfrak{s}^2 erhöht. Harris et al. beschreiben außerdem ein Verfahren, womit die als Wichtungsfaktoren wirkenden E_k reskaliert werden können, um die Genauigkeit ihres Algorithmus zu erhöhen.^[4] Anschließend können die Fehler für die Fourierkoeffizienten A_l, B_l gemäß

$$\Delta A = \Delta B = \left(\frac{s^2}{\sum_{k=1}^N \frac{1}{E_k^2}} \right)^{1/2} \quad (2.6)$$

berechnet werden. Die Fehler für die mittleren Helligkeiten \bar{H}_j^* der j -ten Beobachtungsnacht ergeben sich gemäß

$$\Delta \bar{H}_j^* = \left[\frac{1}{N_j - 1} \frac{N}{N - M} \frac{\chi_j^2}{\sum_{k=1}^{N_j} \frac{1}{E_k^2}} \right]^{1/2} \quad (2.7)$$

mit $\chi_j^2 = \sum_{k=1}^{N_j} \left(\frac{D_k - H(t_k)}{E_k} \right)^2$,

wobei N_j die Zahl der Beobachtungen in der j -ten Nacht bezeichnet. (Für $N_j = 1$ ist der obige Ausdruck nicht definiert.) Sind die \bar{H}_j^* bestimmt, kann aus diesen nach einem von Bowell et al. beschriebenen Verfahren der Steigungsparameter G bestimmt werden.^[4]

2.2.2 MCMC-Methode

Die heutzutage im Bereich der Exoplanetensuche häufig verwendete MCMC-Methode (engl.: Markov-Chain Monte-Carlo), welche als Alternative zur Methode von Harris et al. untersucht wird, ist eine Methode der Bayes-Statistik. Daher werden hier zunächst im allgemeinen Kontext die Grundlagen dieser vorgestellt. In Kapitel 3 wird das hier gewonnene Wissen dann auf die Problematik der Asteroidenlichtkurve angewandt.

Im Gegensatz zur “traditionellen” Statistik (engl. auch frequentist statistic) erlaubt die Bayes-Statistik Aussagen über die Plausibilität von Ereignissen, anstatt Vorhersagen über deren Häufigkeit zu treffen.^[7] Der zentrale Ausgangspunkt der Bayes-Statistik ist das Bayessche Gesetz: Bezeichne H_i eine Hypothese (bspw. die Gültigkeit eines Satzes von Parametern, eines Modells etc.), D einen Satz von Daten und I die zur Verfügung stehenden Hintergrundinformationen (bspw. Ergebnisse vorheriger Messungen), dann gilt^[7]

$$p(H_i|D, I) = p(H_i|I) \cdot \frac{p(D|H_i, I)}{p(D|I)}. \quad (2.8)$$

Die linke Seite bezeichnet man als *A-posteriori-Wahrscheinlichkeit* (engl. posterior probability). Rechts steht

- $p(H_i|I)$ für die *A-priori-Wahrscheinlichkeit* (engl. prior probability), also die Plausibilität von H_i ohne das Vorliegen der Daten D ,
- $p(D|H_i, I)$ für die *Likelihood* von H_i , die angibt wie plausibel das Erhalten der Daten D ist, wenn man die Gültigkeit von H_i voraussetzt, und
- $p(D|I) = \sum_i p(H_i|I)p(D|H_i, I)$ für die *globale Wahrscheinlichkeit* der Klasse H (engl. global likelihood), meist eine Normalisierungskonstante.

Mit Hilfe der A-posteriori-Wahrscheinlichkeit lassen sich verschiedene Hypothesen vergleichen oder im Fall von Parameterbestimmung eine Verteilung $X \rightarrow p(X|D, I)$ angeben (X bezeichne einen oder mehrere Parameter), aus der dann über Maximum oder Erwartungswert der wahrscheinlichste (“plausibelste”) Satz von Parametern bestimmt werden kann. Die Fehler können ebenfalls aus der Verteilung entnommen werden, bspw. über deren Halbwertsbreite. Das Ziel der hier verwendeten MCMC-Methode ist demnach die Angabe der A-posteriori-Verteilung für die Parameter des Modells aus Gleichung (2.3) (bzw. Gleichung (2.5), s.u.). Um die einzelnen Parameter - vor allem aber die Frequenz ω - zu bestimmen, bedient man sich der Methode der Marginalisierung. Dadurch kann die A-posteriori-Verteilung eines Parameters mittels Integration aus der Verteilung für alle Parameter gewonnen werden:^[7]

$$p(x_i|D, I) = \int \dots \int p(X|D, I) dx_1 \cdot \dots \cdot dx_{i-1} \cdot dx_{i+1} \cdot \dots \cdot dx_N \quad (2.9)$$

Nun treten, wie man sieht, hochdimensionale Integrale auf, die häufig sehr aufwändig zu lösen sind. Zusätzlich geht im Falle der Parameterbestimmung die Summe in $p(D|I)$ ebenfalls zu einem Integral über, welches aufwändig zu lösen ist. Daher bedarf es alternativer Methoden, um die A-posteriori-Verteilung so zu bestimmen, dass $p(x_i|D, I)$ leicht aus $p(X|D, I)$ gewonnen und die Bestimmung von $p(D|I)$ umgangen werden kann.

Monte-Carlo Integration und Metropolis-Hastings Algorithmus

Eine Möglichkeit, die hochdimensionalen Integrale numerisch zu lösen liegt in der Monte-Carlo Integration (siehe dazu auch *Numerical Recipes*^[6]). Betrachte ein bestimmtes Integral $\int f dV$ über dem Volumen V . Dann wähle gleichverteilt zufällig N Punkte $\{x_1, \dots, x_N\}$ aus V . Nach dem zentralen Theorem der Monte-Carlo Integration gilt dann:^[8]

$$\int f dV \approx V \langle f \rangle \pm V \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{N}}, \quad (2.10)$$

mit

$$\langle f \rangle = \frac{1}{N} \sum_{i=1}^N f(x_i); \quad \langle f^2 \rangle = \frac{1}{N} \sum_{i=1}^N f^2(x_i). \quad (2.11)$$

Der \pm -Term gibt eine Standardabweichung für den Fall normalverteilter Fehler bzw. eine sehr grobe Fehlerabschätzung für andere Fälle an. Aufgrund des Gesetztes der großen Zahlen kann die Monte-Carlo Integration das Integral prinzipiell beliebig genau approximieren, indem man N erhöht.^[9] Die $p(X|D, I)$ sind jedoch häufig stark lokalisiert, d.h. es existieren große Bereiche im Parameterraum, auf denen $p(X|D, I)$ nahezu null ist. Dies führt dazu, dass aufgrund der Gleichverteilung der $\{x_i\}$ viel Zeit verschwendet wird, um Regionen zu untersuchen, die nicht von Interesse sind.^[9] Abhilfe schaffen modifizierte Versionen des Monte-Carlo

Integrationsverfahrens wie das so genannte “importance sampling” (im Fall hochdimensionaler Probleme schwer anzuwenden^[8,9]) oder die Integration mithilfe einer Markov-Kette und dem Metropolis-Hastings Algorithmus (kurz MH-Algorithmus), welche hier vorgestellt wird (siehe auch *Bayesian Logical Data Analysis for the Physical Sciences*^[9]).

Anstatt unabhängige und gleichverteilte Stichproben aus V zu ziehen, durchläuft der MH-Algorithmus den Parameterraum mit einer Art Zufallsbewegung (engl. random walk). Dabei wird eine Kette direkt aus der zu approximierenden Verteilung $p(X|D, I)$ generiert, die o.g. Markov-Kette.^[9] Um den MH-Algorithmus anwenden zu können bedarf es zunächst einer “Angebotsverteilung” $q(Y|X_t)$ (engl. proposal distribution), welche einen Kandidaten für den nächsten Schritt X_{t+1} der Kette liefert. Als nächstes wird eine Rate $\alpha(X_t, Y)$ berechnet, die angibt, mit welcher Wahrscheinlichkeit Y als nächster Schritt X_{t+1} akzeptiert wird:

$$\alpha(X_t, Y) = \min\{1, r\} \quad \text{mit} \quad r = \frac{p(Y|D, I) q(X_t|Y)}{p(X_t|D, I) q(Y|X_t)} \quad (2.12)$$

r wird als Metropolis-Rate bezeichnet. Hier wird ersichtlich, wie der MH-Algorithmus das Problem der Bestimmung von $p(D|I)$ löst. Zähler und Nenner im ersten Faktor sind proportional zur globalen Wahrscheinlichkeit, damit entfällt $p(D|I)$ aus dem Verfahren. Jetzt kann der MH-Algorithmus aufgestellt werden:

1. Definiere einen Anfangszustand X_0 und setze $t = 0$.
2. Wiederhole:
 - i) Ziehe einen Kandidaten Y aus $q(Y|X_t)$.
 - ii) Berechne $\alpha(X_t, Y)$.
 - iii) Ziehe eine Zufallsvariable u gleichverteilt aus $(0, 1)$.
 - iv) Wenn $u \leq \alpha(X_t, Y)$ setze $X_{t+1} = Y$ andernfalls setze $X_{t+1} = X_t$.
 - v) Erhöhe t um 1.

Die mit der Markov-Kette generierte Verteilung konvergiert unter drei Bedingungen gegen die gesuchte A-posteriori-Verteilung $p(X|D, I)$:^[9]

1. Die Kette muss *irreduzibel* sein, d.h. sie muss in der Lage sein, alle Punkte des Parameter-raums mit positiver Wahrscheinlichkeit zu erreichen.
2. Sie muss *aperiodisch* sein.
3. Sie muss *umkehrbar* sein.

Diese Anforderungen werden durch eine geeignete Wahl von $q(Y|X_t)$ erfüllt. Ergebnis des MH-Algorithmus ist eine Kette $\pi(X)$ von Zuständen X_i im Parameterraum, deren Histogramm

$p(X|D, I)$ approximiert. Auch die marginalisierten Verteilungen $p(x_i|D, I)$ können leicht gewonnen werden, man muss lediglich aus $\pi(X)$ ein Histogramm für x_i erstellen.^[9] Wichtig ist zu erwähnen, dass die Kette zu Beginn eine Art “Einschwingphase” (engl. burn in period) benötigt. Allgemein werden die ersten 1.000-5.000 Zustände der Kette als nicht signifikant betrachtet und verworfen.^[9]

Für den MH-Algorithmus existieren zahlreiche Modifikationen, welche ihn für bestimmte Situationen optimieren. Auch die in dieser Arbeit betrachteten Asteroidenlichtkurven erfordern Veränderungen der ursprünglichen Version (siehe Kapitel 3). Die für diese Arbeit relevanten Erweiterungen werden nun vorgestellt.

Parallele Temperierung

Wie bei vielen nichtlinearen Optimierungsverfahren kann es auch beim MH-Algorithmus vorkommen, dass er in lokalen Extrema (in diesem Falle Maxima) “hängen bleibt.” Ursache dafür ist die zur Verteilung proportionale Metropolis-Rate, die für Punkte X_t innerhalb des Maximums und Kandidaten Y außerhalb sehr klein wird. Um diesem Problem entgegenzuwirken kann ein Temperierungsverfahren angewendet werden (siehe wieder *Bayesian Logical Data Analysis for the Physical Sciences*^[9]). Das hier angewandte Verfahren der parallelen Temperierung (engl. parallel tempering) durchläuft den MH-Algorithmus gleichzeitig zu n_T verschiedenen “Temperaturen.” Betrachte erneut die Metropolis-Rate r :

$$r = \frac{p(Y|D, I) q(X_t|Y)}{p(X_t|D, I) q(Y|X_t)} = \frac{p(Y|I) p(D|Y, I) q(X_t|Y)}{p(X_t|I) p(D|X_t, I) q(Y|X_t)} \quad (2.13)$$

Während der erste und letzte Faktor im rechten Ausdruck unberührt bleiben, wird im mittleren Term ein Temperierungsexponent $\beta = 1/T$ eingeführt:

$$r(\beta) = \frac{p(Y|I) p(D|Y, I)^\beta q(X_t|Y)}{p(X_t|I) p(D|X_t, I)^\beta q(Y|X_t)} \quad (2.14)$$

Es entstehen damit für “heißere” Durchläufe deutlich flachere Verteilungen, sodass ein Herausspringen aus den Maxima mit zunehmender Temperatur wahrscheinlicher wird und damit weiter ausgebreitete Ketten $\Pi(X, \beta)$ entstehen. Man erkennt, dass $\Pi(X, \beta = 1)$ genau der Ursprungskette entspricht ($\Pi(X, 1) = \pi(X)$) und damit wieder gegen die Verteilung $p(X|D, I)$ “konvergiert.” Um nun Informationen zwischen der Ursprungsverteilung $\pi(X)$ und den anderen Verteilungen $\Pi(X, \beta)$ auszutauschen, wird an bestimmten Punkten in der Zufallsbewegung ein Wechsel der Zustände durchgeführt. Dazu wird zufällig eine der ersten $n_T - 1$ Verteilungen ausgewählt und eine “Monte-Carlo Akzeptanzrate” r_M berechnet, nach der dann wie oben entschieden wird, ob ein Wechsel zwischen der gewählten Verteilung i und der Nachbarverteilung

$i + 1$ durchgeführt wird. $X_{t,i}$ bezeichne den Zustand t der Verteilung i .

$$r_M = \min \left\{ 1, \left(\frac{p(D|X_{t,i}, I)}{p(D|X_{t,i+1}, I)} \right)^{\beta_{i+1} - \beta_i} \right\} \quad (2.15)$$

Damit wird der Algorithmus von oben wie folgt erweitert und die Schritte i) bis v) für alle n_T Verteilungen durchgeführt (r oben ist durch $r(\beta)$ aus Gleichung (2.14) zu ersetzen):

- vi) Ziehe eine weitere Zufallsvariable v gleichverteilt aus $(0, 1)$.
- vii) Wenn $v \leq n_s$ wähle gleichverteilt eine der ersten $n_T - 1$ Verteilungen i ; sonst weiter bei i). n_s ist ein Parameter über den reguliert werden kann, wie oft ein Wechsel erfolgen soll.
- viii) Berechne r_M für die gewählte Verteilung i .
- ix) Ziehe eine dritte Zufallsvariable w gleichverteilt aus $(0, 1)$.
- x) Wenn $w \leq r_M$ tausche die Zustände $X_{t,i}$ und $X_{t,i+1}$.

Neben dem Einfluss auf ein mögliches “Steckenbleiben” hat die parallele Temperierung einen weiteren Vorteil. Aus den verschiedenen Verteilungen $\Pi(X, \beta)$ lassen sich die für einen Modellvergleich benötigten Bayes-Faktoren berechnen (siehe 2.2.3)!

Adaptive Algorithmen

Die Angebotsverteilung $q(Y|X_t)$ hat einen starken Einfluss auf die Effizienz des MH-Algorithmus, daher ist es nötig diese geschickt zu wählen. Alternativ kann eine flexible Angebotsverteilung gewählt werden, die im Laufe des Verfahrens angepasst wird, um die Effizienz zu steigern. Dabei kann die Markov-Eigenschaft der Kette verloren gehen, d.h. die oben genannten Kriterien genügen nicht mehr, um die Konvergenz der Kette gegen die A-posteriori-Verteilung zu sichern. Es existieren zahlreiche dieser so genannten adaptiven Verfahren,^[10,11] von denen im Laufe der Arbeit mehrere getestet wurden (siehe Kapitel 3).

2.2.3 Modellvergleich

Wie bereits unter 2.1 erwähnt, liefern unterschiedliche Ordnungen m_i der Fourierreihe konkurrierende Modelle zur Beschreibung der Daten. Im herkömmlichen Verfahren liefern die jeweiligen χ^2 -Werte eine Angabe dafür, welches Modell die Daten am besten beschreibt. Die Bayes-Statistik liefert ebenfalls ein Verfahren zur Bestimmung des besten Modells. Dabei werden jeweils zwei Modelle miteinander verglichen (siehe *Bayesian Logical Data Analysis for the Physical Sciences*^[12]). Betrachte zwei konkurrierende Modelle M_i . Dann ist die A-posteriori-Wahrscheinlichkeit für Modell i durch

$$p(M_i|D, I) = p(M_i|I) \frac{p(D|M_i, I)}{p(D|I)} \quad (2.16)$$

gegeben und der Quotient

$$O_{ij} = \frac{p(M_i|D, I)}{p(M_j|D, I)} \quad (2.17)$$

gibt an, wieviel wahrscheinlicher (“plausibler”) Modell i gegenüber Modell j ist. Üblicherweise definiert man noch die Bayes-Faktoren B_{ij} , die der zentrale Teil der O_{ij} sind, da häufig zu Beginn kein Modell dem anderen vorzuziehen ist und damit der erste Faktor in Gleichung (2.18) gleich eins wird.

$$O_{ij} = \frac{p(M_i|I) p(D|M_i, I)}{p(M_j|I) p(D|M_j, I)} = \frac{p(M_i|I)}{p(M_j|I)} B_{ij} \quad (2.18)$$

Die zur Berechnung der B_{ij} benötigten Terme $p(D|M_{i/j}, I)$ lassen sich aus der parallelen Temperierung gewinnen und begründen damit einen weiteren Vorteil dieser Erweiterung! Es gilt:^[9]

$$\ln[p(D|M_i, I)] = \int \langle \ln[p(D|X, M_i, I)] \rangle_\beta d\beta \quad (2.19)$$

mit $\langle \ln[p(D|X, M_i, I)] \rangle_\beta = \frac{1}{n} \sum_t \ln[p(D|X_{t,\beta}, M_i, I)]$

Die Terme in der Summe auf der rechten Seite lassen sich schnell aus den Zuständen der Kette $\Pi(X, \beta)$ berechnen, denn “unbewusst” wurden oben, beim Erzeugen der Kette, bereits die Informationen $I' = M_i, I$ benutzt, also alle hier vorliegenden Informationen I eingeschlossen, sowie die Gültigkeit des Modells M_i , vorausgesetzt! Um nun die $p(D|M_{i,j}, I)$ zu berechnen, muss die Funktion $\langle \ln[p(D|X, M_i, I)] \rangle_\beta$ aus den β_i bestimmt werden, wozu bereits eine einfache Interpolation durch Polynome ausreicht.

3 Entwicklung eines Analyseverfahrens für Asteroidenlichtkurven

Nachdem bisher die theoretischen Grundlagen dieser Arbeit erläutert wurden, wird in diesem Kapitel auf den Kern der Arbeit eingegangen. Es wurden verschiedene Erweiterungen des ursprünglichen Metropolis-Hastings Algorithmus untersucht. Nachdem die aufgetretenen Probleme aufgezeigt wurden, werden die erfolgreichsten Modifikationen hier genauer vorgestellt und mit simulierten Daten getestet.

3.1 Simulierte Daten

An einigen Stellen in diesem Kapitel werden simulierte Daten genutzt. Um diese zu gewinnen, wurde ein für alle simulierten Daten identischer, fiktiver Beobachtungsplan erstellt. Anhand der zeitlichen Änderung der Phase bekannter Asteroiden wurde die Phasenänderung innerhalb des fiktiven Beobachtungszeitraums generiert. Es wurde ein nach Dymock typischer G -Wert von 0.15 verwendet, um den Einfluss auf die Lichtkurve zu simulieren.^[2] Da für gemessene Helligkeiten zunächst der Entfernungseinfluss über die zugehörigen UCORMAG Terme korrigiert wird, wurden die Testdaten direkt für eine Entfernung von 1AU simuliert, d.h. $UCORMAG = -5 \log(r\Delta) = 0$. Um die Rotationskurve zu simulieren, wurden Fourierreihen fester Ordnung verwendet. Messunsicherheiten wurden durch normalverteiltes Rauschen simuliert. Der Zeitplan mit dem zugehörigen julianischen Datum und der Phase der jeweiligen "Beobachtungsnacht" findet sich in Tabelle 3.1.

Datum	Julianisches Datum	Phase
18.04.2014, 04:48h	2456765.70000	12.71°
18.04.2014, 23:48h	2456766.49167	12.92°
20.04.2014, 03:48h	2456767.65833	13.23°
22.04.2014, 22:48h	2456770.45000	13.97°

Tabelle 3.1: Fiktiver Zeitplan zur Simulation von Lichtkurvendaten und Phase der jeweiligen Beobachtungsnacht.

3.2 MCMC-Verfahren

Modell im MCMC-Verfahren

Wie schon mehrfach erwähnt ist das Ziel dieser Arbeit die Bestimmung der Rotationsperiode $P = 2\pi/\omega$. Aus dem Zusammenhang zwischen Frequenz und Periode wird unter Einbezug von Gleichung (2.3) bzw. (2.5) ersichtlich, dass es sinnvoller ist, die A-posteriori-Verteilung für die Frequenz zu ermitteln und aus den gewonnenen Informationen die Periode zu bestimmen, denn die Frequenz geht direkt in das Modell ein und wird damit linear untersucht, während die Periode über einen $1/P$ -Zusammenhang einfließt und daher nichtlinear untersucht wird. Zusätzlich wird auch hier die Annahme getroffen, der Phasenwinkel sei über die jeweilige Beobachtungsnacht konstant. Dies hat zwei Ursachen. Zum einen liefert das Minor Planet Center differenzielle Helligkeiten, die als Vergleichswerte zu Referenzsternen angegeben werden. Damit entstehen von Nacht zu Nacht mitunter sehr große Versätze und der log-Zusammenhang in (2.3) ist nicht direkt rekonstruierbar. Harris et al. verwenden korrigierte Daten, sodass wie o.g. G aus den \bar{H}_j^* berechnet werden kann. Zum anderen ist die Änderung des Phasenwinkels innerhalb einer Nacht nicht bekannt, sodass der Phasenwinkeleinfluss mathematisch durch eine Stufenfunktion beschrieben wird, deren Einfluss auf die Rotationsperiode vernachlässigbar ist. Statt also G als zusätzlichen Parameter zu untersuchen, werden die auf 1AU korrigierten Daten auf ein "Einheitsniveau" gelegt, indem für jede Nacht die mittlere Helligkeit berechnet und von den Messwerten abgezogen wird. Dies hat den Vorteil, dass der Bereich in dem die \bar{H}_j^* untersucht werden müssen deutlich reduziert und der Parameterraum um eine Dimension verkleinert wird. Demnach ist das im folgenden MCMC-Verfahren verwendete Modell analog zu dem aus Gleichung (2.5).

Um zu überprüfen, ob ein bestimmter Algorithmus ausreichend effizient arbeitet und die gesamte A-posteriori-Verteilung liefert, wurde ein vereinfachtes Testsignal benutzt, welches keinen vertikalen Versatz aufweist und als "Rotationskurve" einen einfachen Sinus besitzt. Anschließend wurde versucht, dieses Signal mit einer Fourierreihe zweiter Ordnung zu rekonstruieren. Die A-posteriori-Verteilung der Frequenz eines Signals erster Ordnung, welches mit einer Reihe zweiter Ordnung überprüft wird, enthält zwei deutliche Spitzen. Eine bei der "echten" Frequenz, dann entsprechen die Koeffizienten A_1 und B_1 denen des Signals und die Koeffizienten A_2 und B_2 verschwinden, sowie eine weitere Spitze bei der halben Frequenz, bei der sich die Koeffizienten entsprechend umgekehrt verhalten. Die \bar{H}_j^* sollten für beide Spitzen verschwinden, da kein vertikaler Versatz vorliegt, d.h. die marginalisierten Verteilungen sollten nur eine Spitze um null aufweisen! Diese Verteilung eignet sich gut, um das Verhalten von modifizierten MH-Algorithmen zu testen, da sie wie beschrieben charakteristische Eigenschaften besitzt. Wird keine oder nur eine der Spitzen gefunden, erforscht der getestete Algorithmus den Parameterraum nicht gut!

Vorwissen

Wie aus Gleichung (2.8) und (2.12), sowie aus den darauf aufbauenden Gleichungen ersichtlich wird, ist es notwendig zu Beginn des Verfahrens die A-priori-Verteilung $p(X|I)$ der Parameter zu kennen und zu spezifizieren, was die Hintergrundinformationen I sind. Wissenschaftler, bei denen dieses Vorwissen identisch ist, sollten immer zu den selben Erkenntnissen gelangen! Dies macht einen wichtigen Aspekt beim Vergleich von Analysen im Rahmen der Bayes-Statistik aus. Die A-priori-Verteilung gibt wie o.g. die Wahrscheinlichkeitsverteilung der Parameter an, bevor die Daten D bekannt sind. Würde man bspw. einen Asteroiden beobachten, dann eine Analyse durchführen und den selben Asteroiden später erneut beobachten und im Rahmen des selben Modells analysieren, so würde man die A-posteriori-Verteilung der ersten Analyse als A-priori-Verteilung der zweiten Analyse ansetzen. Da zu Beginn jedoch noch keine Ergebnisse vorliegen, muss die A-priori-Verteilung als so genannte nicht-informative A-priori-Verteilung (engl. non-informative prior) gewählt werden. Dazu eignet sich eine Gleichverteilung^[13] und es ergibt sich

$$p(X|I) = \prod_i p(x_i|i) = \prod_i \frac{1}{x_{i,\max} - x_{i,\min}} \quad (3.1)$$

mit $x_i \in \{\omega, A_1, B_1, \dots, A_m, B_m, \bar{H}_1^*, \dots, \bar{H}_n^*\}$.

Man erkennt, dass $p(X|I)$ für alle Zustände der Markov-Kette (und auch für die Angebote Y) identisch ist. Damit wird der erste Faktor in (2.12) zu eins. Als Information I wird hier, neben sonstigem Hintergrundwissen wie der Kenntnis von Phase und UCORMAG für die jeweilige Beobachtungsnacht etc., die Gültigkeit des jeweiligen Modells nach (2.5) verwendet.

Likelihood

Ebenfalls benötigt wird die Likelihood-Funktion für die Parameter des entsprechenden Modells (2.5). Diese kann nach Gregory für deterministische Modelle wie das hier vorliegende wie folgt berechnet werden:^[14]

$$p(D|X, I) = \prod_{i=1}^N p(E_i|X, I) \quad (3.2)$$

$p(E|X, I)$ bezeichnet die Wahrscheinlichkeitsverteilung für die Abweichungen der Messwerte D von den Vorhersagen $F(t|X)$ des Modells (Fehler). In dieser Arbeit werden normalverteilte

Fehler angenommen, sodass sich für die Likelihood-Funktion

$$\begin{aligned} p(D|X, I) &= \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} \cdot \exp \left[-\left(\frac{D_i - F(t_i|X)}{\sigma_i} \right)^2 \right] \\ &= \left(\prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} \right) \cdot \exp \left[-\sum_{i=1}^N \left(\frac{D_i - F(t_i|X)}{\sigma_i} \right)^2 \right] \end{aligned} \quad (3.3)$$

mit $F(t|X) = V_{\text{ideal}} - 5 \log(r\Delta)$ aus Gleichung (2.5)

ergibt. Als σ_i werden die vom Minor Planet Center gelieferten Messunsicherheiten angenommen.

Angebotsverteilung

Zuletzt wird noch eine Angebotsverteilung $q(Y|X_t)$ benötigt. Zu Beginn der Arbeit wurde diese als $M = 2m + n + 1$ -dimensionale Normalverteilung mit fester Varianz und ohne Korrelation gesetzt.

$$q(Y|X_t) = N(X_t, \Sigma) = \frac{1}{(2\pi)^{M/2} |\Sigma|^{1/2}} \cdot \exp \left[-\frac{1}{2} (Y - X_t)^\top \Sigma^{-1} (Y - X_t) \right] \quad (3.4)$$

mit $\Sigma = \text{diag}(\sigma)$ und dem Varianzvektor $\sigma = (\sigma_1, \dots, \sigma_M)$

Dies hat den Vorteil, dass $q(Y|X_t)$ symmetrisch und damit der letzte Faktor in (2.12) gleich eins ist ($q(Y|X_t) = q(X_t|Y)$).

3.2.1 Optimierung des Algorithmus

Zunächst wurde der ursprüngliche MH-Algorithmus mit den oben angegebenen Verteilungen getestet. Es wurden Ketten mit 1.000.000 Zuständen erzeugt, die ersten 5.000 wurden jeweils verworfen (s.o.). Die für den Algorithmus wichtigen Parameter finden sich in Tabelle 3.2. Die Akzeptanzrate lag in der Größenordnung von 0.001%! Das entspricht einer Kette mit nur etwa 10 verschiedenen Zuständen, die damit weder signifikant, noch besonders effizient berechnet ist. Das erste große Problem liegt also in der Erhöhung der Akzeptanzrate und damit der Effizienz des Algorithmus. Wie unter 2.2.2 beschrieben, existieren verschiedene adaptive Verfahren, welche die Effizienz des MH-Algorithmus erhöhen.

Parameter	ω in $\frac{1}{h}$	A_l, B_l in mag	\bar{H}_j^* in mag
Startwert	0.2	0	0
Standardabweichung	0.5	0.2	0.2
Maximum	3	1	1
Minimum	0.1	-1	-1

Tabelle 3.2: Parameter des ursprünglichen MH-Algorithmus. Besonders zu bemerken ist die maximale Frequenz von $3/h$, welche auf die von Pravec et al.^[17] gefundene obere Schranke von 12 Rotationen/Tag zurückzuführen ist.

Adaptive Algorithmen

Die Akzeptanzrate des MH-Algorithmus hängt sehr stark von der Angebotsverteilung ab. Dabei kann sie hauptsächlich durch deren Skalenparameter (bspw. die Varianz im Falle einer Normalverteilung) beeinflusst werden.^[15] Kleinere Skalen (und damit kleinere Schrittweiten) führen zu höheren Akzeptanzraten, jedoch auch zu einer langsameren Untersuchung der Zielverteilung $p(X|D, I)$. Wählt man die Rate zu hoch, so werden zu viele Angebote Y verworfen, sodass der Algorithmus ebenfalls ineffizient wird. Nach Gregory^[9] (mit Bezug auf Roberts et al.^[15]) liegt die optimale Akzeptanzrate für hochdimensionale Modelle bei etwa 25% und für ein- oder zweiparametrische Modelle bei etwa 50%. Die Akzeptanzrate durch simples ausprobieren verschiedener Schrittweiten auf den optimalen Wert zu heben gelang nicht. Daher wurde als erstes der von Roberts vorgestellte Adaptive Metropolis-Algorithmus^[11] getestet. Bei diesem wird eine alternative Angebotsverteilung

$$q(Y|X_t) = (1 - \xi)N(X_t, (2.38)^2\Sigma_n/M) + \xi N(X_t, (0.1)^2\mathbf{1}_M/M) \quad (3.5)$$

verwendet. Grundlage dieses Algorithmus ist die Annahme, dass die Angebotsverteilung $q(Y|X_t) = N(X_t, (2.38)^2\Sigma/M)$ für hochdimensionale Fälle optimal ist, wenn Σ die Kovarianzmatrix der Verteilung $p(X|D, I)$ bezeichnet. Σ_n bezeichnet die Kovarianzmatrix der Verteilung, die zum Zeitpunkt t aus $\pi(X)$ gewonnen werden kann. Diese eignet sich, um Σ zu approximieren und sollte sich dieser schrittweise annähern. ξ ist eine kleine positive Konstante (z.B. $\xi = 0.05$ nach Roberts und Rosenthal^[11]). Die Kovarianzmatrix Σ_n wird während des Durchlaufs mehrfach berechnet, wozu zuerst die aktuelle Verteilung $p_t(X|D, I)$ aufgestellt werden muss. Dies wird aufwändiger, je weiter der Algorithmus fortschreitet. Das gesamte Verfahren wurde als zu rechenaufwändig erachtet und daher verworfen. Stattdessen wurden zwei adaptive Verfahren innerhalb des Gibbs-Sampling untersucht.

Adaptiver Metropolis-Algorithmus innerhalb des Gibbs-Sampling

Das Gibbs-Sampling ist eine Variation des MH-Algorithmus. Dabei wird statt des gesamten Zustands X immer nur eine Komponente x_i getauscht und i zufällig oder zyklisch innerhalb

von M variiert. Die Zufallsbewegung wird also immer nur in eine Richtung ausgeführt, anstatt "diagonal." Als Angebotsverteilung $q(y_i|X_t)$ wird beim Gibbs-Sampling ursprünglich die von den anderen Parametern $\{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_M\} \equiv X^-$ abhängige Verteilung $p(x_i|X^-, D, I)$ (engl. full conditional distribution) verwendet, welche leichter aufzustellen ist als $p(X|D, I)$. Außerdem führt diese spezielle Wahl dazu, dass die Metropolis-Rate zu eins wird.^[6] Dennoch ist das Erzeugen von Angeboten aus $p(x_i|X^-, D, I)$ für hohe Dimensionen und komplizierte Modelle aufwändig (eine Möglichkeit wäre die Inversionsmethode). Daher wird eine einfache Angebotsverteilung $q(y_i|X_t)$ aufgestellt und wie oben die Metropolis-Rate berechnet, anhand derer das Angebot angenommen oder verworfen wird. Dieses Verfahren wird Metropolis-Algorithmus innerhalb des Gibbs-Sampling genannt und es existieren verschiedene adaptive Methoden, welche dessen Effizienz erhöhen.

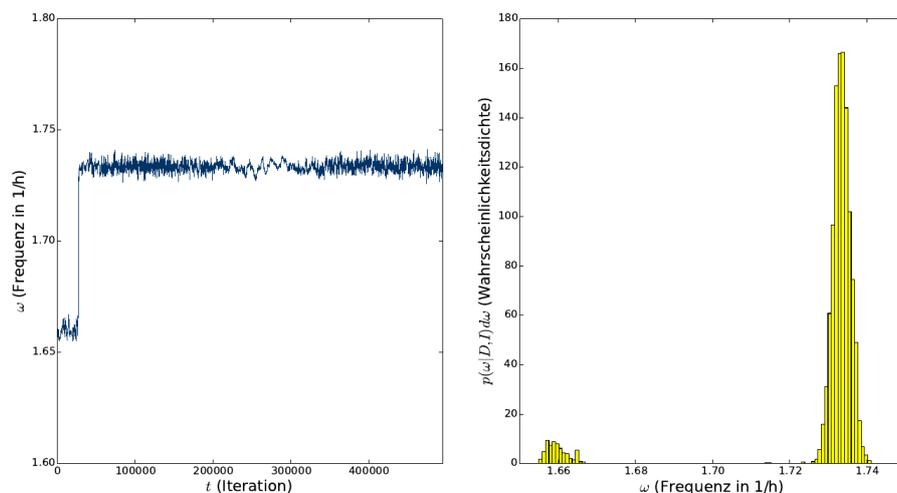
Das erste untersuchte Verfahren wurde ebenfalls von Roberts vorgestellt^[11] und arbeitet wie folgt. Der Angebotszustand Y wird aus X_t berechnet, indem zur i -ten Komponente von X_t ein Schritt Δx_i addiert wird, welcher normalverteilt mit Erwartungswert 0 und Varianz σ_i^2 erzeugt wird. Zu Beginn werden die σ_i fest gewählt (bspw. wie die des ursprünglichen Verfahrens aus Tabelle 3.2). Alle 50 Schritte nach der Einschwingphase wird die Akzeptanzrate in der jeweiligen Komponente überprüft und die Varianz entsprechend modifiziert. Die Varianz sollte erhöht werden, wenn die Rate zu hoch und gesenkt werden, wenn die Rate zu niedrig ist. Da das Verfahren jetzt eindimensional ist, liegt die optimale Akzeptanzrate bei 44%.^[11] Die neue Varianz $\sigma_i'^2$ berechnet sich gemäß

$$\sigma_i' = \exp[\ln(\sigma_i) \pm \delta(t)] \tag{3.6}$$

mit $\delta(t) = \min\{0.01, \sqrt{50/t}\}$.

Nach Roberts sichern die Eigenschaft $\delta(t) \rightarrow 0$ ($t \rightarrow \infty$) und eine (in der Praxis unrelevante) Schranke für die σ_i die Konvergenz der Kette.

Es wurden wieder Ketten mit 1.000.000 Zuständen erzeugt, von denen die ersten 5.000 Zustände verworfen wurden. Als Startwerte wurden die in Tabelle 3.2 gegebenen Startwerte des ursprünglichen MH-Algorithmus gewählt. Die Akzeptanzrate der einzelnen Parameter lag nach 200.000 Schritten im Bereich von 40-50%, am Ende lag die durchschnittliche Akzeptanzrate bei etwa 44%! Eine Übersicht über den Verlauf der Rate während des Durchlaufs gibt Abbildung 3.2. Die erzeugte marginalisierte Verteilung der Frequenz ist in Abbildung 3.1 zu sehen, man erkennt, dass keine der gesuchten Spitzen gefunden wurde. Damit konnte die erste untersuchte Modifikation das Effizienzproblem des Ursprungsalgorithmus lösen, der Parameterraum wird jedoch noch nicht ausreichend untersucht.

**Abbildung 3.1:**

Links: Die Bewegung der Kette im Parameterraum bezüglich der Frequenz ω . Zu sehen ist nur der Ausschnitt des Parameterraums, in dem die zwei lokalen Extrema erforscht wurden. **Rechts:** Marginalisierte A-posteriori-Verteilung der Frequenz. Die gesuchten Spitzen liegen bei etwa 0.6/h und 0.3/h. Beide Grafiken gehören zum adaptiven Metropolis-Algorithmus innerhalb des Gibbs-Sampling nach Roberts.

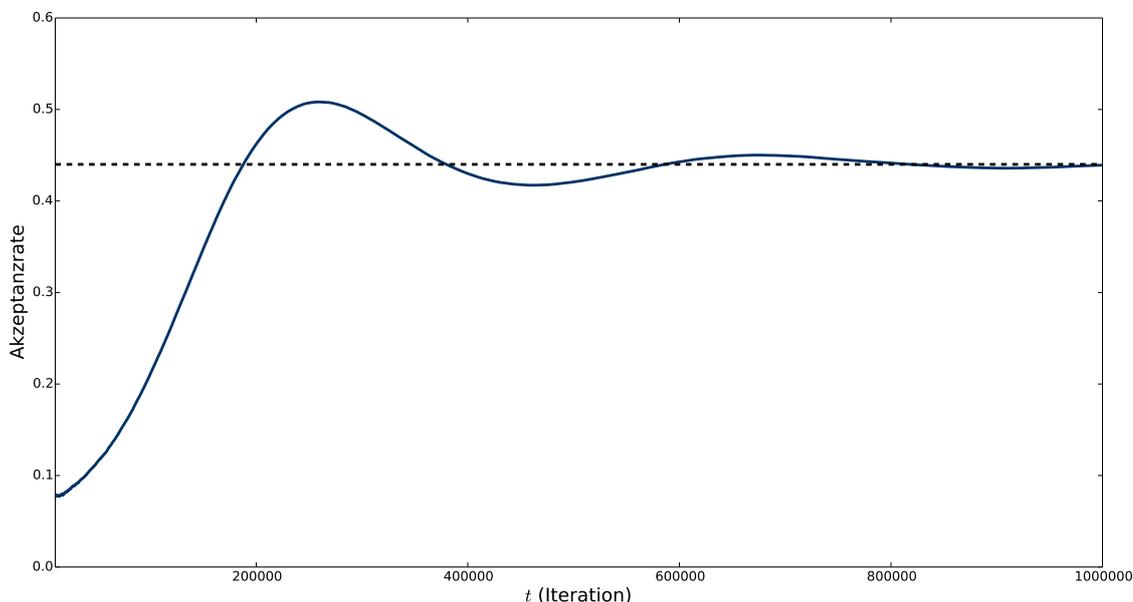


Abbildung 3.2: Verhalten der durchschnittlichen Akzeptanzrate während eines Durchlaufs der adaptiven Methode im Gibbs-Sampling nach Roberts (durchgezogene Linie). Wie man sieht, nähert sich die Rate über den Durchlauf der optimalen Rate von 0.44 (gestrichelte Linie) an.

Das zweite Verfahren (nach Ford^[10]) verwendet eine Normalverteilung als Angebotsverteilung.

$$q(y_i|X_t, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left[-\frac{(x_i - y_i)^2}{2\sigma_i^2} \right] \quad (3.7)$$

Wieder werden die σ_i zunächst fest gewählt und nach der Einschwingphase modifiziert, um eine gewünschte Akzeptanzrate von $\psi_o = 0.44$ zu erreichen. Dazu wird wie folgt vorgegangen: Nach jedem Schritt wird die Akzeptanzrate ψ_i der jeweiligen Komponente berechnet. Wenn $(\psi_i - \psi_o)^2 > S_i\psi_o(1 - \psi_o)/N_i$, wird σ_i angepasst. Dabei ist S_i ein Parameter, der angibt, wie oft die σ_i angepasst werden sollen (zu Beginn ist $S_i = 2$) und N_i die Zahl der Schritte, seit σ_i das letzte Mal angepasst wurde. Wenn die σ_i angepasst werden sollen, werden sie mit $C = \max\{0.01, (\psi_i/\psi_o)^\phi\}$ multipliziert. Dabei ist $\phi = 1$ für $0.5\psi_o < \psi_i$, $\phi = 1.5$ für $0.2\psi_o < \psi_i \leq 0.5\psi_o$ und $\phi = 2$ für $\psi_i \leq 0.2\psi_o$. Wenn ein σ_i verändert wurde, wird die Kette in den Ausgangszustand X_0 zurückgesetzt und S_i um eins erhöht. Sobald ein ψ_i innerhalb von 10% der gewünschten Akzeptanzrate liegt, wird es fixiert. Nachdem alle σ_i angepasst wurden, läuft der Metropolis-Algorithmus innerhalb des Gibbs-Sampling normal ab, sodass die Konvergenzkriterien nicht verletzt sind.

Erneut wurden Ketten mit 1.000.000 Zuständen erzeugt, die ersten 5.000 verworfen und die Startparameter und Startvarianzen aus Tabelle 3.2 gewählt. Da nach der Kalibrationsphase kein Einfluss mehr auf die Akzeptanzrate genommen wird, weicht sie gegen Ende häufig stärker von der gewünschten Rate ψ_o ab. So lag die Akzeptanzrate der Koeffizienten und der \bar{H}_j^* am Ende bei 50-70%, die der Frequenz jedoch nur bei 0.1-1%. Die Effizienz konnte also gegenüber dem Ursprungsverfahren verbessert werden, jedoch nicht so gut wie mit dem ersten Verfahren. Die mit dem zweiten Verfahren erzeugte marginalisierte Verteilung der Frequenz findet sich in Abbildung 3.3. Auch hier ist zu erkennen, dass der Parameterraum nicht ausreichend untersucht wurde. Damit teilen sich beide Modifikationen dieses Problem. Da das Verfahren nach Roberts das Effizienzproblem jedoch deutlich besser lösen konnte, wurde für dieses die unter 2.2.2 bereits vorgestellte parallele Temperierung eingeführt.

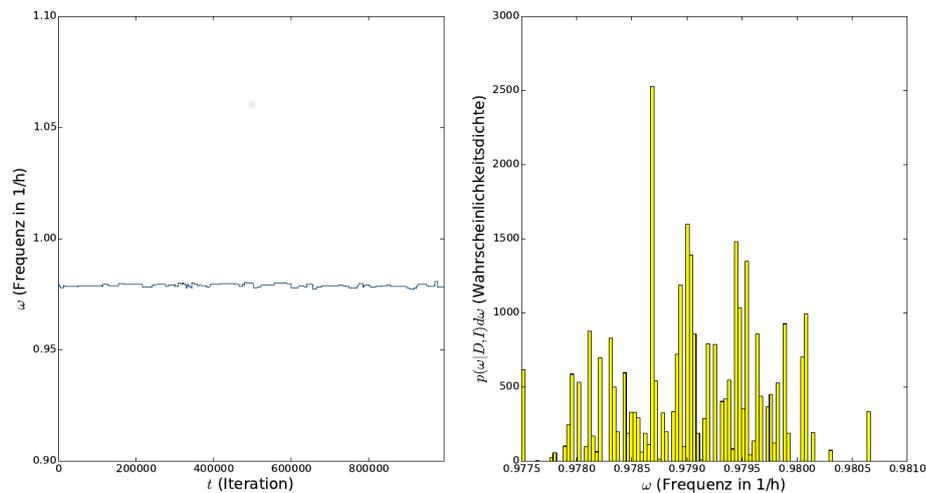


Abbildung 3.3:

Links: Die Bewegung der Kette im Parameterraum bezüglich der Frequenz ω . Zu sehen ist nur der Ausschnitt des Parameterraums, in dem das lokale Extremum erforscht wurde.

Rechts: Marginalisierte A-posteriori-Verteilung der Frequenz. Beide Grafiken gehören zum adaptiven Metropolis-Algorithmus innerhalb des Gibbs-Sampling nach Ford. Auffällig ist, dass die Verteilung deutlich "haariger" wirkt, als in Abbildung 3.1. Dies ist auf die sehr viel geringere Akzeptanzrate in der Frequenz zurückzuführen!

Parallele Temperierung

Wie oben bereits erläutert, werden mit der parallelen Temperierung flachere Verteilungen zeitgleich mit der Ursprungsverteilung untersucht und mit diesen Informationen ausgetauscht, sodass der Parameterraum deutlich breiter untersucht wird, als zuvor. Entscheidend ist die Wahl geeigneter Temperierungsexponenten. Eine übliche Wahl sind logarithmisch verteilte Exponenten zwischen 1 für die ursprüngliche Verteilung und einem Minimalwert, der eine ausreichend flachere Verteilung liefert.^[16] Hier wurden je 5 Ketten der Länge 1.000.000 mit einem minimalen Exponenten von $\beta_{\min} = 0.001$ erzeugt. Die ersten 5.000 Zustände wurden wie üblich verworfen. Die hier gut geeignete Wahl von β_{\min} wurde durch ausprobieren gefunden. Als Startwerte wurden zunächst wieder die aus Tabelle 3.2 verwendet. Verschiedene Durchläufe zeigten jedoch, dass das modifizierte Verfahren gelegentlich in einer Umgebung des Startwertes stecken blieb. Eine zufällige Wahl der Startkoeffizienten und -versätze innerhalb der Grenzen, statt genau in der Mitte konnte erfolgreich verwendet werden, um dieses Verhalten zu unterbinden. Abbildung 3.4 zeigt die mit beiden Modifikationen (Verfahren nach Roberts und parallele Temperierung) erzeugte marginalisierte Verteilung der Frequenz.

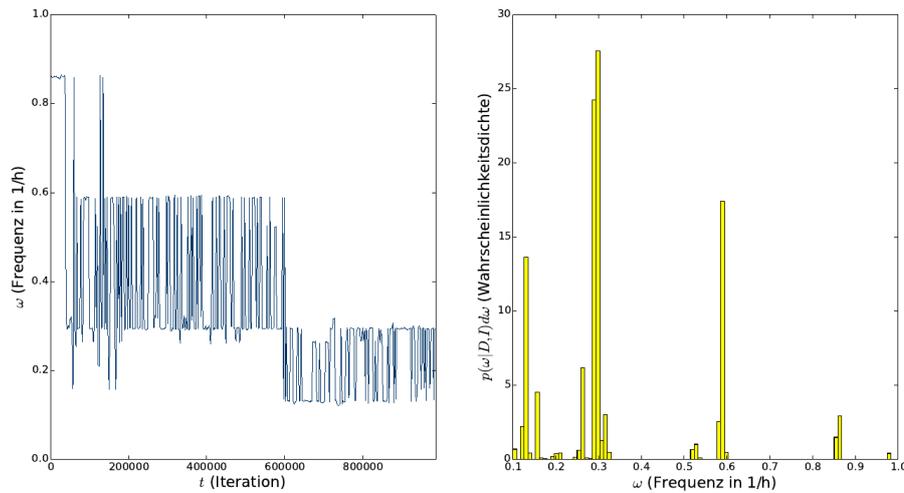


Abbildung 3.4:

Links: Die Bewegung der Kette im Parameterraum bezüglich der Frequenz ω . Zur besseren Übersicht wurde nur jeder 2.500ste Wechsel gezeichnet. Es ist deutlich zu erkennen, dass ein sehr viel größerer Bereich erforscht wurde, als in den Abbildungen 3.1 und 3.3. Es ist davon auszugehen, dass keine höherfrequenten Spitzen existieren und daher die relevanten Bereiche des Parameterraums ausreichend gut untersucht wurden.

Rechts: Marginalisierte A-posteriori-Verteilung der Frequenz. Die beiden erwarteten Spitzen bei der echten Frequenz des Testsignals (etwa 0,6/h) und der halben Frequenz (demnach etwa 0,3/h) sind deutlich zu erkennen. Die dritte deutliche Spitze bei 0,13/h ist auf eine niederfrequente Überlagerung mit dem Versatz durch Beobachtungspausen im Zeitplan zurückzuführen. Beide Grafiken gehören zum Verfahren, welches sowohl die adaptive Methode nach Roberts als auch die parallele Temperierung nutzt.

Der modifizierte Algorithmus die Tests am stark vereinfachten Signal bestanden hatte, wurde er an weiteren simulierten Daten getestet. Dabei wurde eine weitere Modifikation am Verfahren vorgenommen. Da die A-posteriori-Verteilung häufig einige stark lokalisierte Spitzen enthält, die weniger gut erforscht werden, ist es kompliziert daraus die Parameter und deren Unsicherheiten zu gewinnen. Deshalb wurden nach dem finden der A-posteriori-Verteilung der Frequenz weitere Durchläufe ohne parallele Temperierung gestartet, welche nur deutlich eingeschränkte Frequenzbereiche untersuchten. Die Bereiche wurde um die höchsten Spitzen der A-Posteriori-Verteilung gewählt. So konnten die Parameter und deren Unsicherheiten effektiv und sinnvoll bestimmt werden, ohne auf einen Einblick in die volle A-Posteriori-Verteilung der Frequenz zu verzichten.

Während der Tests wurden auch einige Modellvergleiche nach dem unter 2.2.3 beschriebenen Verfahren durchgeführt. Die Ergebnisse eines Tests sind in den folgenden Abbildungen zu sehen.

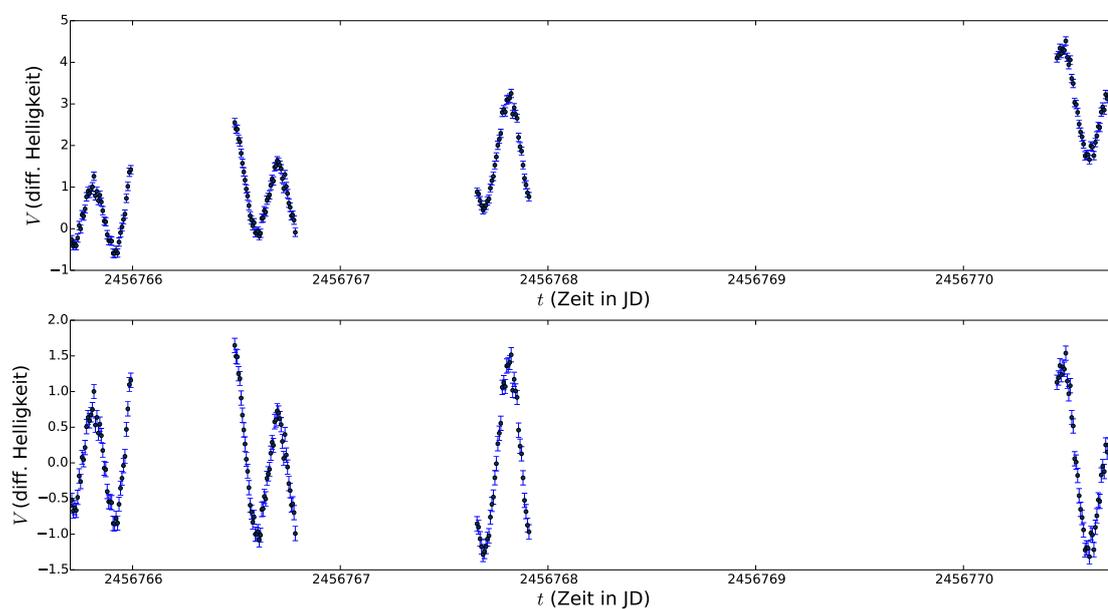


Abbildung 3.5:

Oben: Simulierte Lichtkurve mit Offsets nach dem in Tabelle 3.1 gegebenen Zeitplan. Die verwendete Periode beträgt 10.67h. Es wurde eine Fourierreihe 2. Ordnung erzeugt.

Unten: Die selbe Lichtkurve, wie sie vom Verfahren verwendet wird. Die einzelnen Beobachtungsnächte wurden auf ein Einheitsniveau um null gelegt (siehe Abschnitt 3.2 oben).

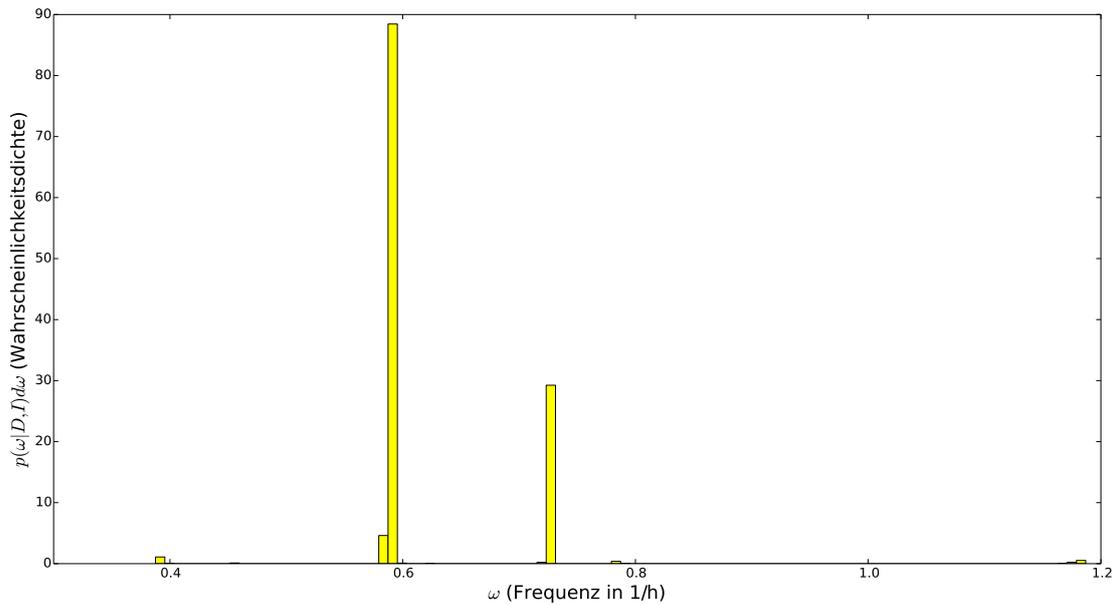


Abbildung 3.6: Nach dem modifizierten Verfahren erzeugte A-Posteriori-Verteilung des Signals aus Abbildung 3.5. Die echte Frequenz liegt bei 0.589 1/h. Dort liegt hier die wahrscheinlichste Spitze. Daher wurde im Parameterbestimmungsverfahren ein Bereich von 0.5 1/h bis 0.7 1/h untersucht. Die daraus konstruierte Lichtkurve findet sich in Abbildung 3.7. Die zweite Spitze bei etwa 0.7 1/h wurde ebenfalls untersucht, wie zu erwarten ist, entstand dabei aber keine sinnvolle Lichtkurve.

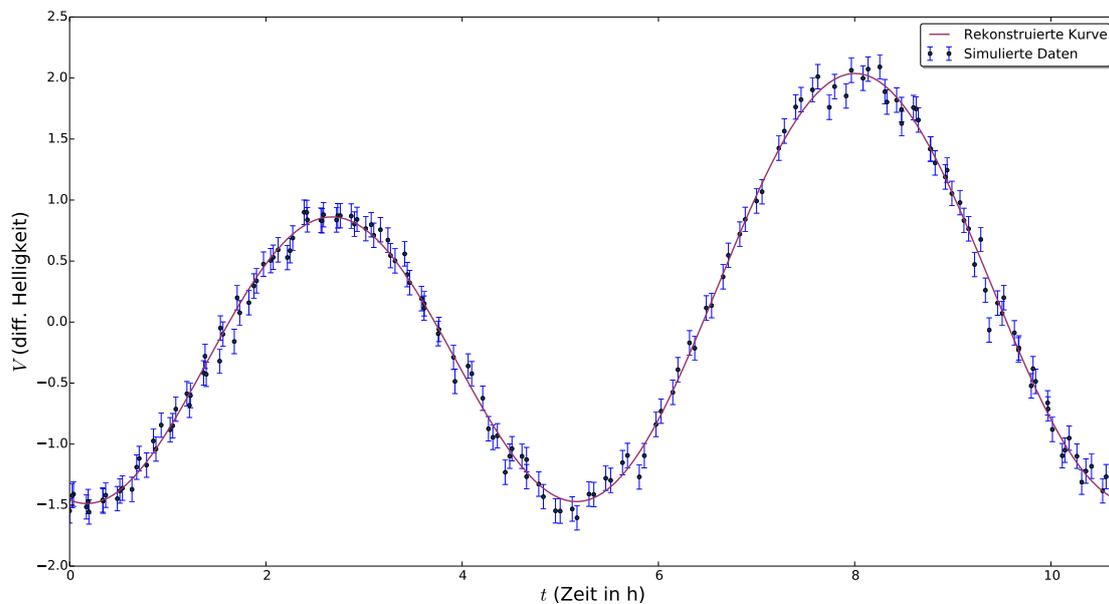


Abbildung 3.7: Die aus den Ergebnissen des Parameterbestimmungsverfahrens mit eingeschränktem Frequenzbereich konstruierte Lichtkurve mit den verwendeten Daten aus Abbildung 3.5. Man erkennt, dass die Kurve sehr gut rekonstruiert wurde!

4 Analyse realer Lichtkurven und Vergleich mit dem Verfahren nach Harris

Zum Abschluss dieser Arbeit wurden die Lichtkurven der Asteroiden (122) Gerda und (252) Clementina analysiert. Für beide Asteroiden existieren Analysen mit dem Verfahren nach Harris et al. (siehe Abbildung 4.1 und 4.6). Die Analysen wurden von der Astronomical Society of Las Cruces bereitgestellt und von F. Pilcher durchgeführt. Anhand von (122) Gerda wird gezeigt, wie eine Analyse mit dem in Kapitel 3 vorgestellten Verfahren abläuft. Anschließend werden die Resultate der Analyse beider Asteroiden mit den Resultaten des Verfahrens nach Harris et al. verglichen.

(122) Gerda

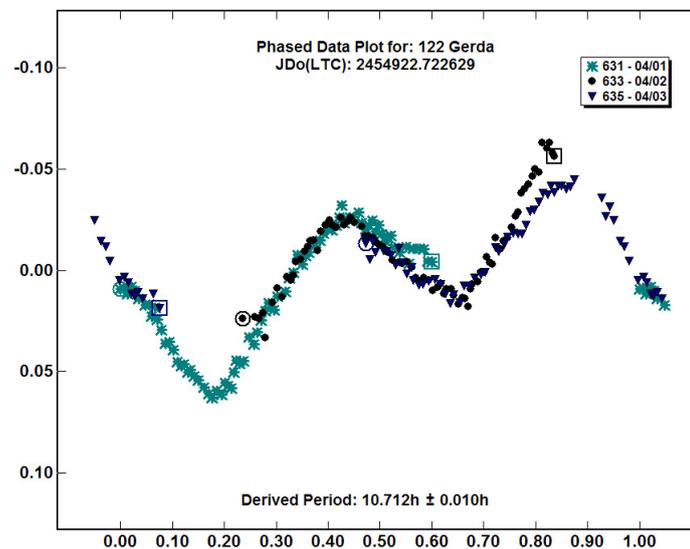


Abbildung 4.1: Resultat der Analyse der Lichtkurve von (122) Gerda mit dem Verfahren nach Harris et al. Die Analyse wurde von F. Pilcher durchgeführt und das Ergebnis von der Astronomical Society of Las Cruces bereitgestellt. Die Grafik ist zu finden unter <http://aslc-nm.org/122Gerda.jpg> (zuletzt überprüft am 26.06.2014)

Der Asteroid (122) Gerda wurde wie folgt mit dem modifizierten MCMC-Verfahren analysiert. Eine detaillierte, schrittweise Beschreibung des Verfahrens findet sich in Kapitel 5. Zunächst

wurden je 5 unterschiedlich temperierte Ketten mit 500.000 Zuständen für Ordnungen im Bereich von $m = 2, \dots, 6$ erzeugt. Anschließend wurden sukzessive Modellvergleiche durchgeführt, um die plausibelste Ordnung zu finden. Diese wurde bei $m = 4$ lokalisiert. Die A-posteriori-Verteilung für Ordnung 4 findet sich in Abbildung 4.2. Die deutlichen Spitzen wurden anschließend mit einer nicht temperierten Kette untersucht und die entsprechenden Lichtkurven konstruiert (Abbildungen 4.3 bis 4.5). Man erkennt, dass die wahrscheinlichste Spitze bei 0.27 1/h eine nicht gut gestützte Lichtkurve liefert, deren hohe Wahrscheinlichkeit nur durch ein nicht sinnvolles ‐Zusammenschieben‐ der Messwerte erklärt wird. Die Spitze bei 0.39 1/h liefert eine gut gestützte Lichtkurve. Ihr entspricht eine Rotationsperiode von $15.77 \pm 0.20\text{h}$. Die Spitze bei 0.58 1/h liefert die Lichtkurve, welche auch vom Verfahren nach Harris et al. gefunden wird, und gibt als Rotationsperiode die selbe Periode von $10.69 \pm 0.9\text{h}$ an. Wie man sieht, ist der angegebene Unsicherheitsbereich, welcher aus der Breite der Spitze in der A-posteriori-Verteilung nach einer Untersuchung im eingeschränkten Frequenzbereich gefunden wurde, deutlich größer als der von F. Pilcher mit der herkömmlichen Methode bestimmte Bereich von $10.712 \pm 0.010\text{h}$.

Das Ergebnis der Lichtkurvenanalyse von (122) Gerda nach dem vorgestellten Verfahren ist die Lichtkurve in Abbildung 4.5, da sie sinnvoll durch Messwerte gestützt wird und etwas wahrscheinlicher ist, als die zur Spitze bei 0.39 1/h gehörige. Dennoch ist das neue Verfahren prinzipiell gescheitert, denn nur ein ‐manuelles Aussortieren‐ der wahrscheinlichsten Kurve lieferte überhaupt ein sinnvolles Resultat. Sowohl die herkömmliche, als auch die mit dem neuen Verfahren nach einem manuellen Eingriff bestimmte Lichtkurve sind gut durch Messdaten gestützt, sodass an dieser Stelle lediglich festzustellen ist, dass die Unsicherheiten nach dem bekannten Verfahren deutlich geringer sind.

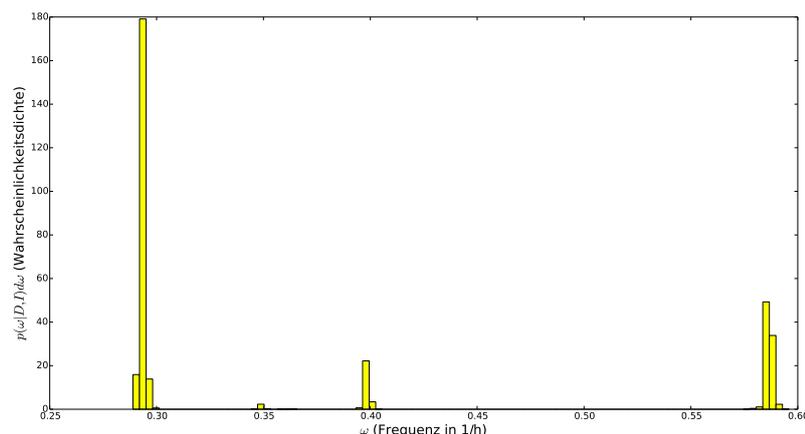


Abbildung 4.2: Erzeugte A-posteriori-Verteilung zur Fourierreihe 4. Ordnung aus der Analyse von (122) Gerda. Man erkennt 3 deutliche Spitzen, deren zugehörige Lichtkurven in den Abbildungen 4.3 bis 4.5 zu finden sind.

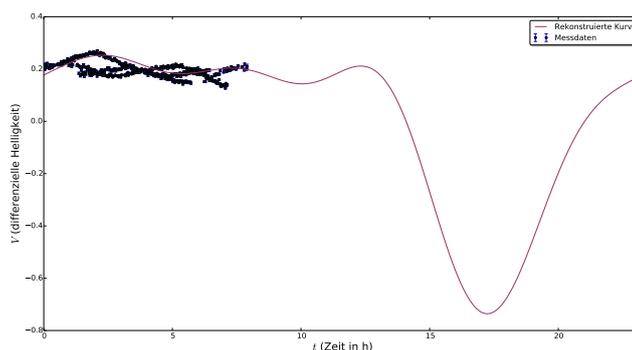


Abbildung 4.3: Lichtkurve zur Spitze um 0.3 1/h . Die Kurve ist nicht gut durch Datenpunkte gestützt und wird daher als nicht signifikant betrachtet, obwohl sie der höchsten Spitze in der Verteilung 4.2 entspricht.

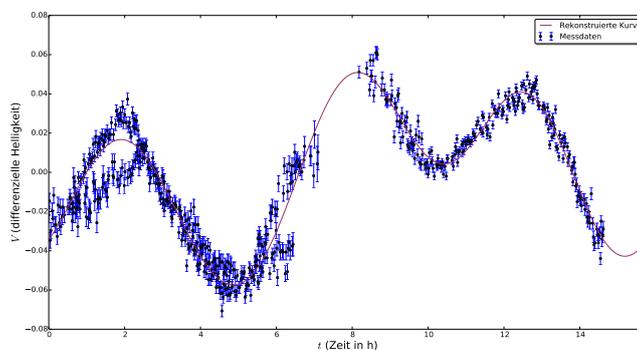


Abbildung 4.4: Lichtkurve zur Spitze um 0.4 1/h . Die ermittelte Rotationsperiode beträgt $15.77 \pm 0.20 \text{ h}$.

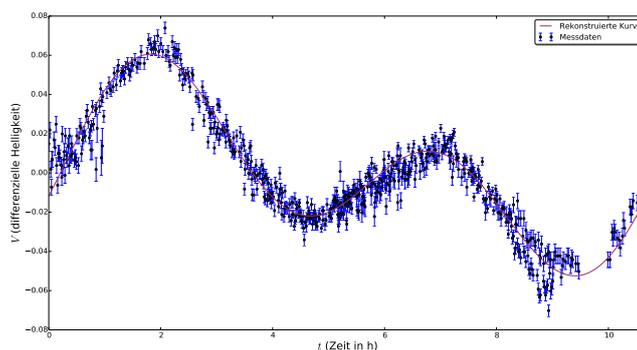


Abbildung 4.5: Lichtkurve zur Spitze um 0.6 1/h . Die Kurve entspricht einer Phasenverschiebung der Kurve in Abbildung 4.1 und gibt damit das Resultat der Analyse nach Harris wieder. Da die zugehörige Spitze in der Verteilung 4.2 wahrscheinlicher ist, als die bei 0.4 1/h , stimmen die Resultate der herkömmlichen Methode und der neuen Methode überein.

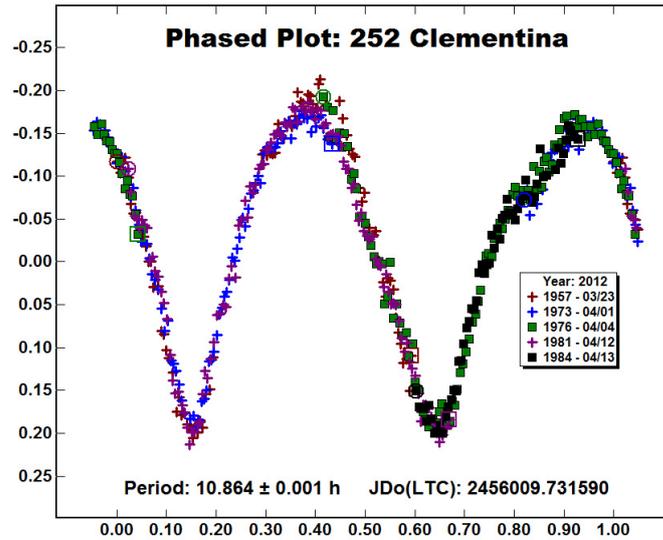


Abbildung 4.6: Resultat der Analyse der Lichtkurve von (252) Clementina mit dem Verfahren nach Harris et al. Die Analyse wurde von F. Pilcher durchgeführt und das Ergebnis von der Astronomical Society of Las Cruces bereitgestellt. Die Grafik ist zu finden unter <http://aslc-nm.org/252Clementina.jpg> (zuletzt überprüft am 26.06.2014)

(252) Clementina

Die Analyse der Lichtkurve von (252) Clementina mit dem neuen Verfahren lieferte $m = 3$ als plausibelste Ordnung der Fourierreihe. Abbildung 4.7 zeigt die zur 3. Ordnung gehörige A-posteriori-Verteilung. Sie weist nur eine Spitze bei etwa 0.6 1/h auf. Die zu dieser Spitze gehörige Lichtkurve ist in Abbildung 4.8 zu sehen. Man erkennt, dass die Lichtkurven beider Verfahren (Abbildungen 4.8 und 4.6) übereinstimmen. Die mit dem MCMC-Verfahren bestimmte Rotationsperiode liegt bei $10.860 \pm 0.20 \text{ h}$. Damit ist das Verfahren nach Harris et al. auch im Fall von (252) Clementina genauer als das neue Verfahren.

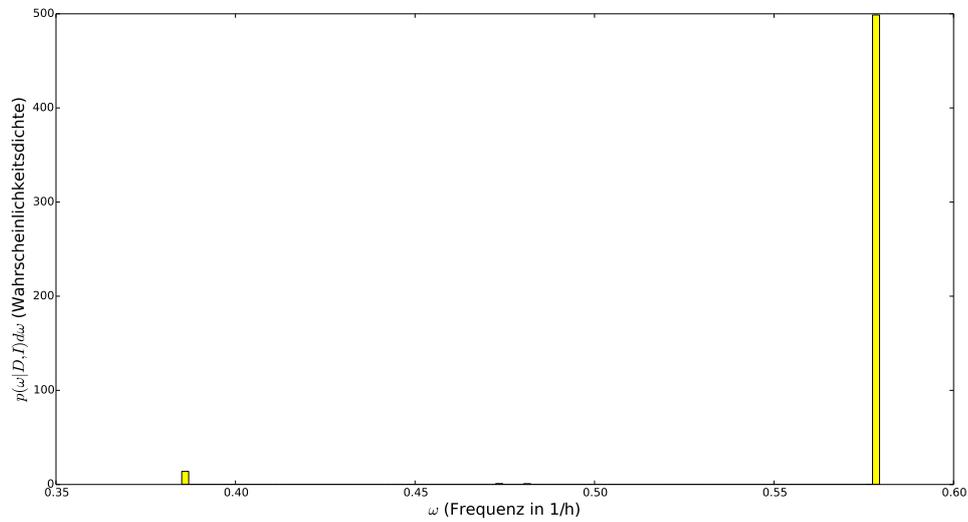


Abbildung 4.7: Erzeugte A-posteriori-Verteilung zur Fourierreihe 3. Ordnung aus der Analyse von (252) Clementina. Wie man sieht liegt nur eine deutliche Spitze vor, deren zugehörige Lichtkurve in Abbildung 4.8 zu sehen ist.

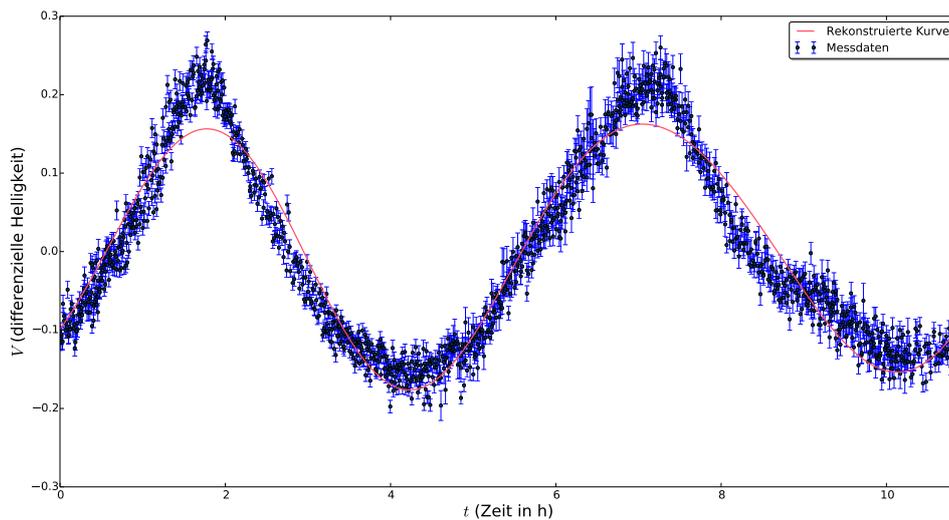


Abbildung 4.8: Lichtkurve zur Analyse von (252) Clementina mit dem MCMC-Verfahren. Die bestimmte Rotationsperiode liegt bei 10.860 ± 0.020 h.

5 Zusammenfassung

Das Ziel der vorliegenden Arbeit war es, ein neues, Bayes-statistisches Verfahren zur Bestimmung der Rotationsperiode von Asteroiden zu testen. Dazu wurde zunächst die mathematische Beschreibung der Lichtkurve durch modifizierte Fourierreihen vorgestellt und dargelegt, wie das herkömmliche Verfahren die entscheidenden Parameter dieser bestimmt. Anschließend wurde anhand von Tests mit simulierten Daten eine Methode entwickelt, welche die A-posteriori-Verteilung der Rotationskurvenparameter für eine feste Ordnung der Fourierreihe bestimmt. Diese arbeitet nach folgendem Algorithmus:

1. Als erstes werden die benötigten Verteilungen aufgestellt. Als *A-priori-Verteilung* wird eine Gleichverteilung verwendet;

$$p(X|I) = \prod_i p(x_i|i) = \prod_i \frac{1}{x_{i,\max} - x_{i,\min}}$$

mit $x_i \in \{\omega, A_1, B_1, \dots, A_m, B_m, \bar{H}_1^*, \dots, \bar{H}_n^*\}$.

Die *Likelihood* wurde nach Gregory typisch für ein deterministisches Modell konstruiert;

$$p(D|X, I) = \left(\prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma_i}} \right) \cdot \exp \left[\sum_{i=1}^N \left(\frac{D_i - F(t_i|X)}{\sigma_i} \right)^2 \right]$$

mit $F(t|X) = V_{\text{ideal}} - 5 \log(r\Delta)$ aus Gleichung (2.5) .

Da ein adaptives Verfahren innerhalb des Gibbs-Sampling genutzt wird, bildet eine eindimensionale Normalverteilung mit variabler Varianz die *Angebotsverteilung*.

2. Als nächstes werden die Startparameter und -varianzen gewählt. Als Startwerte für die Koeffizienten A_l, B_l und Versätze \bar{H}_j^* werden zufällige Werte innerhalb der Grenzen generiert, die Startfrequenz liegt bei 0.5 1/h. Die Grenzen der jeweiligen Parameter und deren Startvarianzen $\xi_{i,0}^2$ finden sich in Tabelle 3.2.
3. Dann kann mit dem eigentlichen Algorithmus begonnen werden. Dazu wird $t = 0$ gesetzt und je nach Länge der Kette (in dieser Arbeit wurden meist 500.000 Zustände generiert) folgendes wiederholt:
 - i) Alle M Schritte wird eine Permutationsmenge $\{\sigma(M)\}$ generiert, aus der sukzessive ein aktueller Parameter i gewählt wird. (M bezeichne die Zahl der Parameter des

Modells.) Dann wird zur i -ten Komponente des aktuellen Zustands ein kleiner Wert Δ_i addiert, welcher aus einer Normalverteilung $N(0, \xi_i)$ generiert wird. Der Zustand $X_t + (0, \dots, \Delta_i, \dots, 0)$ wird als Angebotszustand Y gewählt.

ii) Aus

$$r(\beta) = \frac{p(Y|I) p(D|Y, I)^\beta q(X_t|Y)}{p(X_t|I) p(D|X_t, I)^\beta q(Y|X_t)}$$

wird die Metropolis Rate berechnet und mit einer Zufallszahl v , die gleichverteilt aus $(0, 1)$ gezogen wird verglichen. Ist $v \leq r(\beta)$ wird $X_{t+1} = Y$, andernfalls $X_{t+1} = X_t$ gesetzt.

iii) Alle 50 Schritte wird die jeweilige Akzeptanzrate aller Parameter i , d.h. das Verhältnis der Schritte mit $v \leq r(\beta)$ zur Gesamtschrittzahl in i , mit der gewünschten Rate von 0.44 verglichen. Ist die Rate zu hoch, werden die ξ_i erhöht, andernfalls gesenkt. Dazu setze $\xi_i = \xi'_i$ wobei

$$\xi'_i = \exp[\ln(\xi_i) \pm \delta(t)]$$

mit $\delta(t) = \min\{0.01, \sqrt{50/t}\}$.

iv) Die bisherigen Schritte werden für 5 Verteilungen gleichzeitig durchlaufen. Dabei werden verschiedene Temperierungsexponenten $\beta \in \{1, 0.18, 0.03, 0.005, 0.001\}$ verwendet! Zusätzlich wird in jedem Schritt eine Zufallszahl w gleichverteilt aus $(0, 1)$ gezogen und mit 0.05 verglichen. Ist die Zahl kleiner, wird Information unter den Ketten ausgetauscht. So wird im Durchschnitt alle 20 Schritte innerhalb der Ketten "kommuniziert." Um Information auszutauschen wird zufällig eine ganze Zahl j aus $[0, 4]$ gezogen. Dann wird

$$r_M = \min \left\{ 1, \left(\frac{p(D|X_{t,j}, I)}{p(D|X_{t,j+1}, I)} \right)^{\beta_{j+1} - \beta_j} \right\}$$

berechnet und eine letzte Zufallszahl z aus $(0, 1)$ gezogen. Ist $z \leq r_M$ werden die Zustände $X_{t,j}$ und $X_{t,j+1}$ getauscht.

Das beschriebene Verfahren ist in der Lage die Rotationsperiode aus Asteroidendaten zu bestimmen. Dabei treten jedoch teilweise Probleme auf, da die Daten "zusammengeschoben" werden, was keine sinnvollen Resultate produziert (siehe auch Abbildung 4.3). Neben dieser Schwierigkeit, die durch manuelles Aussortieren der nicht signifikanten Lösungen bewältigt werden kann, liefert das neue Verfahren nicht die Genauigkeit der herkömmlichen Methode. Zuletzt ist die Umsetzung des Algorithmus wie er hier entwickelt wurde wesentlich komplizierter, als der nach Harris et al. und arbeitet wesentlich langsamer.

Eine Optimierung der vielen Parameter des Verfahrens, sowie ein Test auf ausreichende Unterstützung durch Daten könnten die nicht sinnvollen Lösungen eliminieren und die Genauigkeit erhöhen. Dennoch ist nicht abzusehen, dass eine MCMC-Methode mit dem von Harris et al. entwickelten Verfahren konkurrieren kann, da sich keine Vorteile gegenüber aktuell etablierten Umsetzungen wie der im als Standard anerkannten Programm MPO Canopus ergeben.

A Literaturverzeichnis

- [1] B. D. Warner, *A Practical Guide to Lightcurve Photometry and Analysis*, (Springer Science+Business Media, Inc., New York, 2006), Kapitel 1 “Getting Started”
- [2] R. Dymock, *The H and G magnitude system for asteroids*, J.Br.Astron.Assoc 117, Juni 2007
- [3] R. K. Buchheim, *Methods and Lessons Learned Determining The H-G-Parameters Of Asteroid Phase Curves*, Mai 2010
- [4] A. W. Harris, J. W. Young et al., *Photoelectric Observations of Asteroids 3, 24, 60, 261, and 863*, 30. November 1987
- [5] Siehe Kapitel 11 in *A Practical Guide to Lightcurve Photometry and Analysis*^[1], “Period Analysis”
- [6] W. H. Press et al., *Numerical Recipes (Third Edition)*, (Cambridge University Press, Cambridge, 2007), Kapitel 15 “Modeling of Data”
- [7] P. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences*, (Cambridge University Press, Cambridge, 2005), Kapitel 1 “Role of probability theory in science”
- [8] Siehe Kapitel 7 in *Numerical Recipes*^[6], “Random Numbers”
- [9] Siehe Kapitel 12 in *Bayesian Logical Data Analysis for the Physical Sciences*^[7], “Markov chain Monte Carlo”
- [10] E. B. Ford, *Improving the Efficiency of Markov Chain Monte Carlo for Analyzing the Orbits of Extrasolar Planets*, 29. Nezember 2005
- [11] G. O. Roberts, J. S. Rosenthal, *Examples of Adaptive MCMC*, Januar 2008
- [12] Siehe Kapitel 3 in *Bayesian Logical Data Analysis for the Physical Sciences*^[7], “The how-to of Bayesian inference”
- [13] A. Gelman, *Prior distribution*, (in “Encyclopedia of Environmetrics,” John Wiley & Sons, Ltd, Chichester), 2002

- [14] Siehe Kapitel 4 in *Bayesian Logical Data Analysis for the Physical Sciences*^[7], “Assigning probabilities”
- [15] G. O. Roberts et al., *Weak convergence and optimal scaling of random walk metropolis algorithms*, (in “The Annals of Applied Probability”, Vol. 7), 1997
- [16] S. E. Dosso et al., *Parallel tempering for strongly nonlinear geoacoustic inversion*, 19. September 2012
- [17] P. Pravec et al., *Asteroid Rotations*

Abbildungsverzeichnis

1.1 Beispielhafte Lichtkurve des Asteroiden (2893) Annette	4
1.2 Vergrößerung der Lichtkurve aus 1.1	5
3.1 Resultate mit dem Verfahren nach Roberts	23
3.2 Verhalten der durchschnittlichen Akzeptanzrate während eines Durchlaufs der Methode nach Roberts	23
3.3 Resultate mit dem Verfahren nach Ford	25
3.4 Resultate mit dem temperierten Verfahren nach Roberts	26
3.5 Simulierte Lichtkurve	27
3.6 A-posteriori-Verteilung der simulierten Lichtkurve	28
3.7 Rekonstruierte Lichtkurve aus der Simulation	28
4.1 Analyse von (122) Gerda mit dem Verfahren nach Harris et al.	29
4.2 A-posteriori-Verteilung zur Fourierreihe 4. Ordnung der Daten zu (122) Gerda	30
4.3 Lichtkurve zur Spitze um 0.3 1/h in 4.2	31
4.4 Lichtkurve zur Spitze um 0.4 1/h in 4.2	31
4.5 Lichtkurve zur Spitze um 0.6 1/h in 4.2	31
4.6 Analyse von (252) Clementina mit dem Verfahren nach Harris et al.	32
4.7 A-posteriori-Verteilung zur Fourierreihe 3. Ordnung der Daten zu (252) Clementina	33
4.8 Lichtkurve zur Spitze in 4.7	33

Erklärung

Hiermit erkläre ich, dass ich diese Arbeit im Rahmen der Betreuung am Institut für theoretische Physik ohne unzulässige Hilfe Dritter verfasst und alle Quellen als solche gekennzeichnet habe.

Maximilian Arthus Schanner
Dresden, Juni 2014