


# Fast numerics for the spin orbit equation with realistic tidal dissipation and constant eccentricity

Michele Bartuccelli<sup>1</sup> · Jonathan Deane<sup>1</sup>  · Guido Gentile<sup>2</sup>

Received: 24 May 2016 / Revised: 10 January 2017 / Accepted: 7 February 2017 /

Published online: 13 March 2017

© The Author(s) 2017. This article is published with open access at Springerlink.com

**Abstract** We present an algorithm for the rapid numerical integration of a time-periodic ODE with a small dissipation term that is  $C^1$  in the velocity. Such an ODE arises as a model of spin–orbit coupling in a star/planet system, and the motivation for devising a fast algorithm for its solution comes from the desire to estimate probability of capture in various solutions, via Monte Carlo simulation: the integration times are very long, since we are interested in phenomena occurring on timescales of the order of  $10^6$ – $10^7$  years. The proposed algorithm is based on the high-order Euler method which was described in Bartuccelli et al. (Celest Mech Dyn Astron 121(3):233–260, 2015), and it requires computer algebra to set up the code for its implementation. The payoff is an overall increase in speed by a factor of about 7.5 compared to standard numerical methods. Means for accelerating the purely numerical computation are also discussed.

**Keywords** Fast numerics for ODEs · Spin-orbit problem · Probability of capture · Monte Carlo simulation

## 1 Introduction

Since the discovery that Mercury, uniquely in the solar system, is trapped in a 3:2 spin-orbit resonance with the Sun—that is, it turns three times on its own axis for every two revolutions it makes around the Sun—there have been attempts to explain how this came

---

✉ Jonathan Deane  
j.deane@surrey.ac.uk

Michele Bartuccelli  
m.bartuccelli@surrey.ac.uk

Guido Gentile  
gentile@mat.uniroma3.it

<sup>1</sup> Department of Mathematics, University of Surrey, Guildford GU2 7XH, UK

<sup>2</sup> Dipartimento di Matematica e Fisica, Università Roma Tre, 00146 Rome, Italy

about. Observations going back to the formation of the Solar system are clearly impossible, but simulation of its history is not. To be more specific, what is needed is a ‘spin-orbit model’, a model of the interaction between a satellite (Mercury) and the body that it orbits (the Sun). The model is in the form of a differential equation (ODE). By solving this ODE starting from many randomly chosen initial conditions, the dynamics of the satellite can then be simulated; the various solutions in which it can become captured are counted, and probabilities of capture can thereby be estimated. This is a classic example of the technique of Monte Carlo simulation, and the use of random initial conditions reflects the fact that we do not know the early history of the Solar system: hence, we choose initial conditions at random so as to cover a wide range of possibilities, and from these we estimate capture probabilities.

The context of such a calculation is wider than that of a single, albeit unusual, case in our Solar system however. Orbiting around other stars, there is evidence for the existence of exoplanets. The resonances in which these are trapped are likely to have consequences for the possibility of life having developed there. This underlines the importance of computing capture probabilities in a wider context.

Clearly, dissipation in the spin-orbit model must play an important role, since this is what leads to eventual capture in one of the available attractors. In this paper, we are concerned with a dissipation model proposed in [Noyelles et al. \(2014\)](#), which we refer to throughout as the NFME model. Here, the authors point out how the model that they advocate describes more realistically the effect of tidal dissipation, in comparison with the much simpler constant time lag (CTL) model. The latter has been used extensively in the literature since the seminal paper by [Goldreich and Peale \(1966\)](#)—see for instance ([Correia and Laskar 2004](#); [Celletti and Chierchia 2008, 2009](#); [Bartuccelli et al. 2015](#)) for recent results on the basins of attraction of the principal attractors. A clear drawback of the CTL model is that, for the parameters of the Sun–Mercury system, the main attractor (with 70% probability of being observed) is quasi-periodic, with an approximate velocity ratio of 1.256—such a solution is usually referred to as *supersynchronous* in the literature. As we have seen, Mercury is in fact in a 3:2 spin-orbit resonance, and the CTL model indicates that the probability of this is only about 8%. On the other hand, with the NFME model, much more plausibly, all attractors have mean velocity ratio close to rational values, and about 42% of the initial data are captured by an attractor with mean velocity ratio close to  $3/2$ —see [Table 4](#). The shortcomings of the CTL model have also been stressed in [Ferraz-Mello \(2013, 2015\)](#), where the creep tide theory has been developed. According to such a theory, the mean excess angular velocity of the supersynchronous solution turns out to be very small (proportional to the relaxation factor) in the case of very large viscosity.

The greater realism of the NFME model comes at a cost, however:

1. The NFME model is not smooth; in fact, it is only  $C^1$  in the angular velocity,  $\dot{\theta}$ , as we shall see; it also involves functions with rapid variations.
2. There is considerable detail to take into account in implementing the NFME model mathematically—see [Sect. 2](#).
3. The functions in the model take significantly more computation time to evaluate, and so numerical calculations are comparatively slow.

The first point represents a substantial obstacle to the application of perturbation theory to the problem, as well as for implementing a fast algorithm to solve the ODE. Both of these objectives are important. For instance, perturbation theory is a useful tool for obtaining any kind of analytical results for the dynamics. For the present work, though, we should now discuss the motivation for developing *fast* algorithms.

It is important in Monte Carlo estimation to give bounds on the probabilities obtained. In practice, this boils down to specifying, say, the 95% confidence interval (CI) for the probabilities, and these confidence intervals shrink as  $I$ , the number of initial conditions investigated, grows—specifically, the width of a given CI decreases as  $I^{-1/2}$  (Walpole et al. 1998). This important result is one motivation for developing fast algorithms for the spin-orbit ODE: the more initial conditions we consider, the better defined our probability estimates become. Additionally, realistic values of dissipation in this kind of problem are very small indeed, and so capture typically takes place over a period measured in millions of years. Both of these facts motivate us to develop fast solution algorithms.

The main thrust of this paper is to show how previous work (Bartuccelli et al. 2015) on fast algorithms can be applied to solving the spin-orbit ODE with dissipation given by the NFME model. To this end, we first concisely set out the model for spin-orbit coupling in, for instance, the Sun–Mercury system, as described by Makarov (2012) and further discussed by Noyelles, Frouard, Makarov and Efroimsky in Noyelles et al. (2014). We then propose a fast numerical algorithm, based on the high-order Euler method (HEM) introduced in Bartuccelli et al. (2015), for solving the spin-orbit ODE incorporating the NFME model. Essentially, we need to deal with the time integration of a trajectory in a different way when it is close to a resonance: this requires a careful partition of the phase space into subsets that we call strips, with a different numerical integrator being used in each strip.

After rigorous testing, we conclude that the proposed algorithm is about 7.5 times faster than simply using a general-purpose numerical ODE solver—a useful gain in speed.

In most of the previous related work, for instance, Correia and Laskar (2004), Celletti and Chierchia (2008), Makarov (2012) and Makarov et al. (2012), only about 1000 initial conditions were considered, almost certainly because of computer time limitations. By contrast, we are able to consider more than 50,000 initial conditions here, with the total computation time being of order 2 weeks.

The dynamics of the solutions of the spin-orbit ODE with the NFME model are also interesting, and we discuss these in a separate paper (Bartuccelli et al. 2017).

Throughout this paper, the units for mass, length and time will be kg, km and Earth years, respectively. For ease of cross-referencing, an equation numbered ( $N \cdot x$ ) in this paper is equation ( $x$ ) in Noyelles et al. (2014), modified if necessary, but only for consistency with our notation. When we refer to a ‘standard numerical method’ or ‘numerical method’ for ODE solving, we mean one of the family of well-known general-purpose ODE solvers (e.g. Runge–Kutta, Adams); when we mean HEM, which is also numerical, we refer to it explicitly by name.

## 2 The NFME model

### 2.1 The spin-orbit ODE

The ODE is

$$\ddot{\theta} = \frac{\tau_z^{\text{TRI}}(\theta, t) + \tau_z^{\text{TIDE}}(\dot{\theta})}{\xi M_{\text{planet}} R^2}, \quad (\text{N.1})$$

where  $\theta(t)$  is the angle between the semimajor axis of the planet, assumed to be an ellipsoid, and the major axis of its orbit,  $\xi$  is a measure of the inhomogeneity of the planet (with  $\xi = 2/5$  for a homogeneous sphere),  $M_{\text{planet}}$  is the mass of the planet and  $R$  is its radius. Numerical values for these, and indeed all relevant parameters can be found in Table 1.

**Table 1** Numerical values for the parameters used in the NFME model for the Sun and Mercury, taken from [Noyelles et al. \(2014\)](#)

Name	Symbol	Numerical value
Parameter values specific to Mercury		
Semimajor axis	$a$	$5.791 \times 10^7$ km
Mean motion	$n$	$26.0879$ rad year $^{-1}$
Planetary radius	$R$	$2.44 \times 10^3$ km
Dimensionless m.o.i. (inhomogeneity)	$\xi = C/(M_{\text{planet}} R^2)$	0.346
Triaxiality	$(B - A)/C$	$9.350 \times 10^{-5}$
Planetary mass	$M_{\text{planet}}$	$3.301 \times 10^{23}$ kg
Unrelaxed rigidity	$\mu$	$7.967 \times 10^{28}$ kg km $^{-1}$ year $^{-2}$
Present-day orbital eccentricity	$e$	0.2056
Andrade time	$\tau_A$	500 year
Maxwell time	$\tau_M$	500 year
Andrade parameter	$\alpha$	0.2
Acceleration constants		
Triaxiality acceleration constant	$\zeta$	$0.09545$ year $^{-2}$
Tidal acceleration constant	$\eta$	$0.03096$ year $^{-2}$
Other parameter values		
Mass of Sun	$M_{\text{star}}$	$1.989 \times 10^{30}$ kg
Gravitational constant	$G$	$6.646 \times 10^{-5}$ kg $^{-1}$ km $^3$ year $^{-2}$
Triaxiality index range	$Q^{\text{TRI}}$	$\{-4, \dots, 6\}$
Tidal index range	$Q^{\text{TIDE}}$	$\{-1, \dots, 7\}$

### 2.2 The triaxiality torque

The triaxiality-caused torque along the  $z$ -axis,  $\mathcal{T}_z^{\text{TRI}}(\theta, t)$ , is a torque exerted on the planet by the gravitational field of the star, arising from the fact that the planet is not a perfect sphere. [Goldreich and Peale \(1966\)](#) and the authors of [Noyelles et al. \(2014\)](#) approximate  $\mathcal{T}_z^{\text{TRI}}$  by its quadrupole part, which is given by

$$\mathcal{T}_z^{\text{TRI}} \approx -\frac{3a^3}{2r^3} (B - A)n^2 \sin 2(\theta - f) = -\frac{3a^3}{2r^3} (B - A)n^2 (\sin 2\theta \cos 2f - \cos 2\theta \sin 2f), \tag{N.5}$$

where the principal moments of inertia of the planet, along the  $x$ ,  $y$  and  $z$  axes, respectively, are  $A < B < C = \xi M_{\text{planet}} R^2$ ;  $n$  is its mean motion;  $r = r(t)$  is the distance between the centres of mass of the star and the planet at time  $t$ ;  $a$  is the semimajor axis of the orbit of the planet; and  $f = f(t)$  is the true anomaly as seen from the star. We further define  $\mathcal{M}(t)$ , the mean anomaly, which is such that  $n = \dot{\mathcal{M}}$ , with a good approximation to  $n$  ([Noyelles et al. 2014](#)) being given by<sup>1</sup>  $n \approx \sqrt{G(M_{\text{planet}} + M_{\text{star}})/a^3}$ , where  $G$  is the gravitational constant

<sup>1</sup> Numerically, this is true: [Noyelles et al. \(2014\)](#) give  $n = 26.0879$  year $^{-1}$  for Mercury and the approximation yields  $n \approx 26.0897$  year $^{-1}$ .

and  $M_{\text{star}}$  is the mass of the star. Hence, aside from a constant of integration which we set to zero,  $\mathcal{M} = nt$ .

The standard procedure from this point is to make the following pair of Fourier expansions:

$$\left(\frac{r(t)}{a}\right)^n \frac{\cos}{\sin}(mf(t)) = \sum_{k \in \mathbb{Z}} X_k^{n,m}(e) \frac{\cos}{\sin}(knt), \tag{N.7, N.8},$$

where  $X_k^{n,m}(e)$  are the Hansen coefficients, which depend on the orbital eccentricity  $e$ . The authors of [Noyelles et al. \(2014\)](#) compute them via [\(Duriez 2007\)](#)

$$X_k^{n,m}(e) = (1 + z^2)^{-n-1} \sum_{g=0}^{\infty} (-z)^g \sum_{h=0}^g C_{g-h,h}^{n,m} J_{k-m+g-2h}(ke), \tag{N.9}$$

where  $z = (1 - \sqrt{1 - e^2})/e$ ,  $J_k(x)$  is the  $k$ -th order Bessel function of the first kind, and

$$C_{r,s}^{n,m} = \binom{n+1+m}{r} \binom{n+1-m}{s}$$

with the binomial coefficients (extended to all integer arguments) being given by

$$\binom{l}{m} = \begin{cases} \frac{l(l-1)\dots(l-m+1)}{m!} & l \in \mathbb{Z}, m \geq 1 \\ 1 & l \in \mathbb{Z}, m = 0 \\ 0 & l \in \mathbb{Z}, m < 0. \end{cases}$$

For convenience, we follow [Noyelles et al. \(2014\)](#) by defining the functions  $G_{lpq}(e) = X_{l-p+q}^{-l-1, l-2p}(e)$ . Then finally,

$$\sigma_z^{\text{TRI}}(\theta, t) = -\frac{3}{2}(B - A)n^2 \sum_{q \in \mathbb{Z}} G_{20q}(e) \sin(2\theta - (q + 2)nt). \tag{N.10}$$

In [Noyelles et al. \(2014\)](#), the sum is over  $q \in Q^{\text{TRI}} = \{-4, \dots, 6\}$ .

### 2.3 The tidal torque

The authors of [Noyelles et al. \(2014\)](#) forcefully point out the problems associated with the constant time lag (CTL) model, which leads to a dissipation term of the form  $\gamma(\dot{\theta} - \omega)$ , with  $\gamma$  and  $\omega$  constants. They argue that this model, among other things, is incompatible with the physically plausible rheology of terrestrial planets, and can also give rise to quasi-periodic solutions for the spin-orbit problem ([Correia and Laskar 2004](#); [Celletti and Chierchia 2009](#); [Bartuccelli et al. 2015](#)).

Tidal dissipation is modelled in [Noyelles et al. \(2014\)](#) by expanding both the tide-raising potential of the star, and the tidal potential of the planet, as Fourier series. The Fourier modes are written  $\omega_{lmpq}$ , where  $l, m, p, q$  are integers. Appendix B of [Noyelles et al. \(2014\)](#) justifies the simplification  $l = m = 2, p = 0$ , which arises from the smallness of the obliquity of the planet (axial tilt), in combination with an averaging argument. This then gives the following approximation for the polar component of the tidal torque:

$$\sigma_z^{\text{TIDE}}(\dot{\theta}) = \frac{3GM_{\text{star}}^2}{2a} \left(\frac{R}{a}\right)^5 \sum_{q, j \in Q^{\text{TIDE}}} G_{20q}(e)G_{20j}(e)H_2(\omega_{220q}) \text{sgn}(\omega_{220q}) \cos[(q - j)\mathcal{M}]. \tag{N.11a}$$

Here,  $Q^{\text{TIDE}} = \{-1, \dots, 7\}$ ;  $H_2(\omega_{220q}) = k_2(\omega_{220q}) \sin |\epsilon_2(\omega_{220q})|$ , where  $k_2$  is a positive definite, even function of the mode  $\omega_{lmpq}$  and  $\epsilon_2$  is an odd function of the mode;  $\text{sgn}()$  is the usual signum function. Hence, overall,  $H_2(x) \text{sgn}(x)$  is an *odd* function of  $x$ . The reason why  $H_2$  is written in this form is that it will turn out to depend on a *fractional power* of  $|\omega_{lmpq}|$ : the oddness of  $H_2(x) \text{sgn}(x)$  is therefore retained without raising a negative number to this fractional power.

The additive error terms for this expression are  $O(e^8\epsilon)$  and  $O(\epsilon R^7/a^7)$ , in which  $\epsilon$ , which is  $O(1)$ , is a phase lag. For Mercury,  $e^8 \approx 10^{-6}$  and  $(R/a)^7 \approx 10^{-31}$ . In [Noyelles et al. \(2014\)](#), the additional approximation is made that, since the terms with  $q \neq j$  oscillate, they naturally affect the detailed dynamics of the planet, but nonetheless, the overall capture probabilities are insensitive to them ([Makarov et al. 2012](#)). Hence, the tidal torque is well approximated by the secular part only, i.e. the  $q = j$  terms; this gives

$$\tau_z^{\text{TIDE}}(\dot{\theta}) = \langle \tau_z^{\text{TIDE}}(\dot{\theta}) \rangle_{l=2} = \frac{3GM_{\text{star}}^2}{2a} \left(\frac{R}{a}\right)^5 \sum_{q \in Q^{\text{TIDE}}} G_{20q}^2(e) H_2(\omega_{220q}) \text{sgn}(\omega_{220q}). \tag{N.11b}$$

Neglecting apsidal and nodal precessions, we have

$$\omega_{lmpq} = (l - 2p + q)n - m\dot{\theta} \tag{N.12}$$

with  $l \geq 2$ ,  $m, p = 0, \dots, l$  and  $q \in \mathbb{Z}$ . For the reasons mentioned earlier, we consider only the mode  $\omega_{220q} = (q + 2)n - 2\dot{\theta}$ .

Finally, we give the NFME model for the function  $H_l(\omega_{lmpq})$ . We use the abbreviation  $\chi = |\omega_{lmpq}|$ , in terms of which

$$H_l(\omega_{lmpq}) = -\frac{3A_l}{2(l-1)} \frac{I'(\chi)\chi}{(\mathcal{R}'(\chi) + A_l\chi)^2 + I'(\chi)^2}, \tag{N.15 mod.}$$

where

$$A_l = \frac{4\pi(2l^2 + 4l + 3)\mu R^4}{3lGM_{\text{planet}}^2},$$

with  $\mu$  being the unrelaxed rigidity;

$$\mathcal{R}'(\chi) = \chi + \chi^{1-\alpha} \tau_A^{-\alpha} \cos(\alpha\pi/2) \Gamma(\alpha + 1) \tag{N.16 mod.}$$

and

$$I'(\chi) = -\tau_M^{-1} - \chi^{1-\alpha} \tau_A^{-\alpha} \sin(\alpha\pi/2) \Gamma(\alpha + 1). \tag{N.17 mod.}$$

In these expressions,  $\Gamma$  is the usual gamma function,  $\tau_A$  and  $\tau_M$  are the Andrade and Maxwell times of the mantle, respectively, and  $\alpha$  is the Andrade parameter.

### 2.4 Parameter values for the Sun/Mercury system

In [Table 1](#), we give numerical values appropriate to the Sun and Mercury for the parameters introduced so far. We standardise the units to kg for mass, km for length and (Earth) years for time.

### 2.5 The triaxiality and tidal angular accelerations

In order to study the spin-orbit problem further, it is convenient to introduce two quantities with the dimensions of angular acceleration

$$\ddot{\theta}^{\text{TRI}}(\theta, t) = \mathcal{I}_z^{\text{TRI}}(\theta, t)/\xi M_{\text{planet}} R^2 \quad \text{and} \quad \ddot{\theta}^{\text{TIDE}}(\dot{\theta}) = \mathcal{I}_z^{\text{TIDE}}(\dot{\theta})/\xi M_{\text{planet}} R^2.$$

Taken individually, these quantities are not accelerations *per se*—more correctly, they might be described as the triaxiality/tidal contributions to the angular acceleration, since, when computed on the solution  $\theta(t)$ , their sum is  $\ddot{\theta}$ , as we shall see from Eq. (3)—but for brevity, we refer to them as the triaxiality and tidal accelerations. We then define

$$\zeta = \frac{3(B - A)n^2}{2\xi M_{\text{planet}} R^2} \approx 0.09545 \text{ year}^{-2} \quad \text{so that}$$

$$\ddot{\theta}^{\text{TRI}}(\theta, t) = -\zeta \sum_{q \in \mathcal{Q}^{\text{TRI}}} G_{20q}(e) \sin(2\theta - (q + 2)nt), \tag{1}$$

and

$$\eta = \frac{3GM_{\text{star}}^2 R^5}{2a^6 \xi M_{\text{planet}} R^2} \cdot \frac{3A_l}{2(l - 1)} = \frac{3\pi(2l^2 + 4l + 3)}{l(l - 1)} \frac{\mu M_{\text{star}}^2 R^7}{\xi M_{\text{planet}}^3 a^6} \approx 0.03096 \text{ year}^{-2}$$

for  $l = 2$ , so that

$$\ddot{\theta}^{\text{TIDE}}(\dot{\theta}) = -\eta \sum_{q \in \mathcal{Q}^{\text{TIDE}}} G_{20q}^2(e) P_2(|\omega_{220q}|) \text{sgn}(\omega_{220q}), \tag{2}$$

where

$$P_2(\chi) = \frac{I'(\chi)\chi}{(\mathcal{R}'(\chi) + A_2\chi)^2 + I'(\chi)^2}$$

and  $\mathcal{R}'(\chi)$  and  $I'(\chi)$  are defined in Eq. (N.17 mod.).

In terms of these angular accelerations, the spin-orbit ODE becomes

$$\begin{aligned} \ddot{\theta} + \zeta \sum_{q \in \mathcal{Q}^{\text{TRI}}} G_{20q}(e) \sin(2\theta - (q + 2)nt) \\ + \eta \sum_{q \in \mathcal{Q}^{\text{TIDE}}} G_{20q}^2(e) P_2(|(q + 2)n - 2\dot{\theta}|) \text{sgn}((q + 2)n - 2\dot{\theta}) = 0 \end{aligned} \tag{3}$$

and this is our starting point for the rest of the paper.

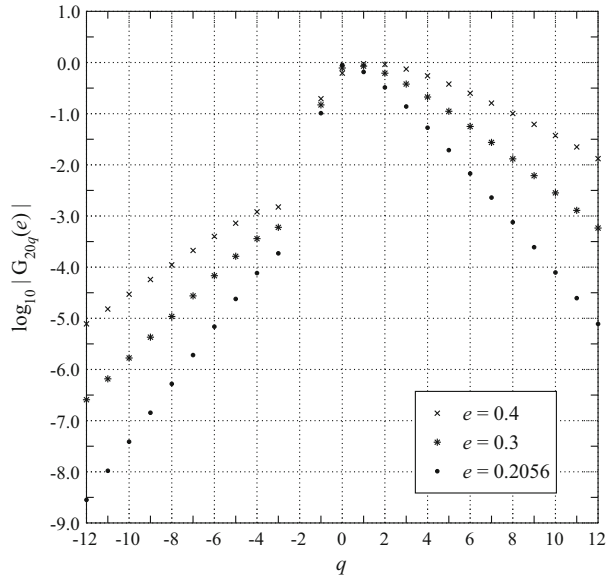
### 3 The NFME model with practical values

We now look at the behaviour and magnitude of the various functions that go to make up the NFME model, using the values for the Sun and Mercury given in Table 1.

#### 3.1 Hansen coefficients

In Fig. 1, we plot the Hansen coefficients relevant to the problem,  $G_{20q} = X_{q+2}^{-3,2}$ , for  $e = 0.2056, 0.3$  and  $0.4$  and for  $q = -12, \dots, 12$ . These were computed from Eq. (N.9) by truncating the first sum at  $p = 120$  and using 20 significant figures for computation—these values allow the computation of the coefficients to a much higher precision than necessary.

**Fig. 1** Plot of  $\log_{10} |G_{20q}(e)|$  for  $e = 0.2056, 0.3, 0.4$  (the latter two values being for illustration) and  $q = -12, \dots, 12$ . Note that, for all values of  $e$  considered, (i) only  $G_{2,0,-1}$  is negative, and (ii)  $G_{2,0,-2} = 0$  and is therefore not plotted



### 3.2 The triaxiality acceleration

We first make a simple estimate of the order of magnitude of  $\ddot{\theta}^{\text{TRI}}(\theta, t)$ . Treating  $\theta$  and  $t$  as independent variables, it is clear from Eq. (N.10) that  $\ddot{\theta}^{\text{TRI}} \in [-D, D]$  for all  $\theta, t$ , where

$$D(e) = \zeta \sum_{q \in \mathcal{Q}^{\text{TRI}}} |G_{20q}(e)|. \tag{4}$$

We find  $D(0.2056) = 0.2096$ ,  $D(0.3) = 0.3016$  and  $D(0.4) = 0.4396 \text{ year}^{-2}$ .

In Fig. 2, we give a plot of  $\ddot{\theta}^{\text{TRI}}(\theta_0 + \dot{\theta}_0 t, t)$  for  $e = 0.2056$ ,  $\theta_0 = 2.0$ ,  $\dot{\theta}_0 = 29.0$ , the latter two values being (almost arbitrarily) chosen to approximate a solution starting from an angular velocity slightly greater than  $n$ . Note that  $\ddot{\theta}^{\text{TRI}} \in [-0.19, 0.19]$ , which is consistent with the estimate, from Eq. (4), of  $[-0.21, 0.21]$  when  $e = 0.2056$ . As we shall see, for example from Fig. 3, the triaxiality term dominates the tidal term by several orders of magnitude for most of the time.

### 3.3 The tidal acceleration

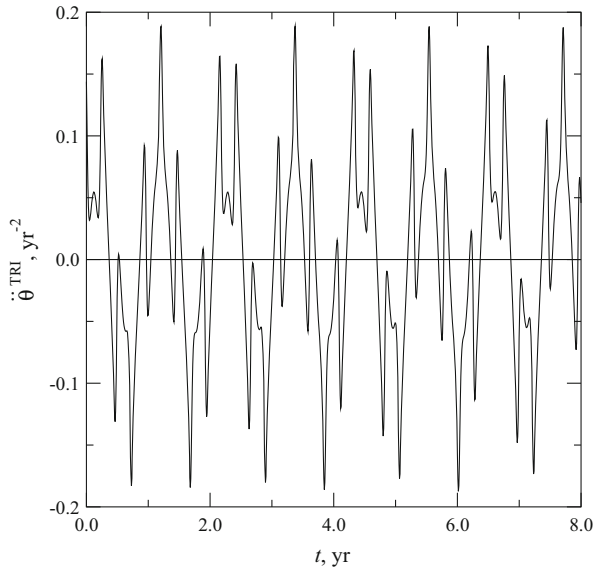
We now discuss the tidal acceleration  $\ddot{\theta}^{\text{TIDE}}(\dot{\theta})$ . Throughout this section, we take  $e = 0.2056$ . Figures 3, 4 and 5 show the tidal acceleration, plotted vertically, versus the relative rate of rotation,  $\dot{\theta}/n$ .

This function is illustrated in Fig. 3, which shows  $\ddot{\theta}^{\text{TIDE}}(\dot{\theta})$ , plotted over the entire range of interest,  $\dot{\theta}/n \in [-1, 5]$ . The dominant features here are the ‘kinks’, which occur at  $\dot{\theta}/n = 1 + q/2$  for  $q = -1, \dots, 7$ , this being the range of the sum,  $\mathcal{Q}^{\text{TIDE}}$ , defining the tidal acceleration.

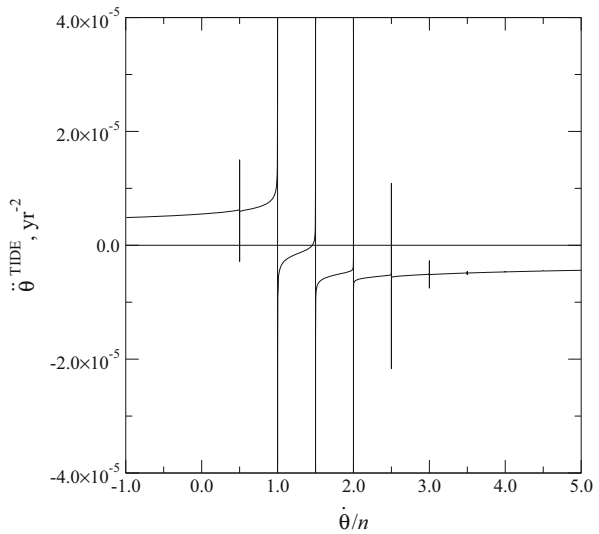
The five ‘kinks’ at which  $\ddot{\theta}^{\text{TIDE}}$  changes sign are those corresponding to  $q = -1, \dots, 3$ , and, for the purpose of comparison, these are plotted over the same narrow range of  $\dot{\theta}/n = 1 + q/2 \pm 10^{-4}$ , and on the same vertical scale,  $\pm 7 \times 10^{-4}$ , in Fig. 4.



**Fig. 2** An illustrative plot of the triaxiality angular acceleration,  $\ddot{\theta}^{\text{TRI}}(\theta, t)$  as a function of time, with  $e = 0.2056$ , and  $\theta(t) = 2 + 29t$



**Fig. 3** The tidal angular acceleration,  $\ddot{\theta}^{\text{TIDE}}(\dot{\theta})$ , with  $e = 0.2056$ , as given by Eq. (2). ‘Kinks’ occur at values  $\dot{\theta}/n = 1/2, 1, 3/2, 2, \dots, 9/2$ , although those at 4, 9/2 are too small to see on this scale. Note that the angular acceleration does not change sign at a kink for  $\dot{\theta}/n = 3, 7/2, 4, 9/2$

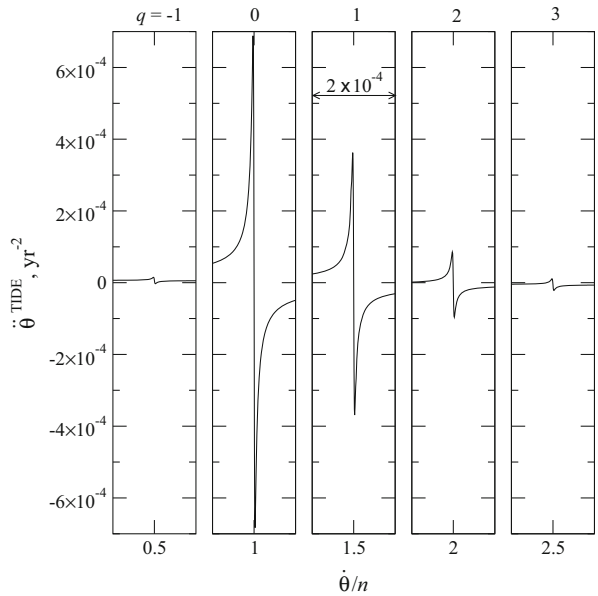


The fact that  $\ddot{\theta}^{\text{TIDE}}$  is continuous, even at a kink, is implied in Fig. 5, in which  $\ddot{\theta}^{\text{TIDE}}$  has been plotted against  $\dot{\theta}/n$  for  $\dot{\theta}/n \in [3/2 - 4 \times 10^{-5}, 3/2 + 4 \times 10^{-5}]$ . However, the derivative  $n d\ddot{\theta}^{\text{TIDE}}/d\dot{\theta}$  displays a cusp at the values of  $\dot{\theta}/n$  corresponding to a kink, as is apparent from the definition of  $H(\omega_{lmpq})$  in Eq. (N.15 mod.), in particular from its dependence on  $\chi = |\omega_{220q}| = |(q + 2)n - 2\dot{\theta}|$ . This is illustrated in Fig. 6.

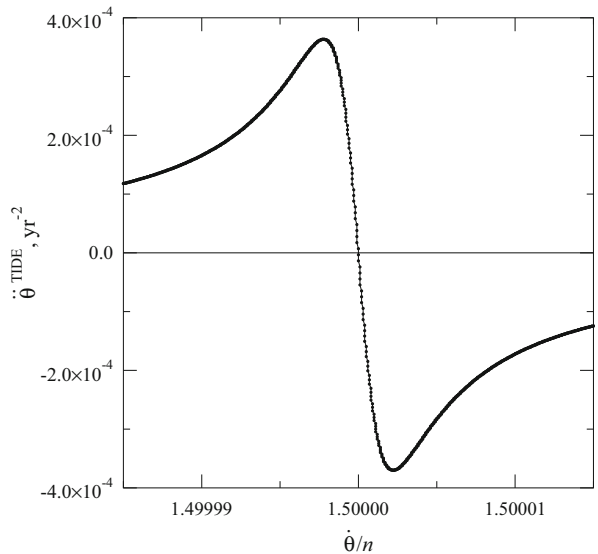
### 4 A fast algorithm

We find qualitatively different solutions to the differential equation (3), depending on the initial conditions: as well as periodic solutions, some of the solutions we find appear, numerically

**Fig. 4** Magnified version of the ‘kinks’ corresponding to  $q = -1, 0, 1, 2, 3$ , as shown in Fig. 3. The width of each plot is  $2 \times 10^{-4}$  and the vertical range is  $\pm 7 \times 10^{-4} \text{ year}^{-2}$



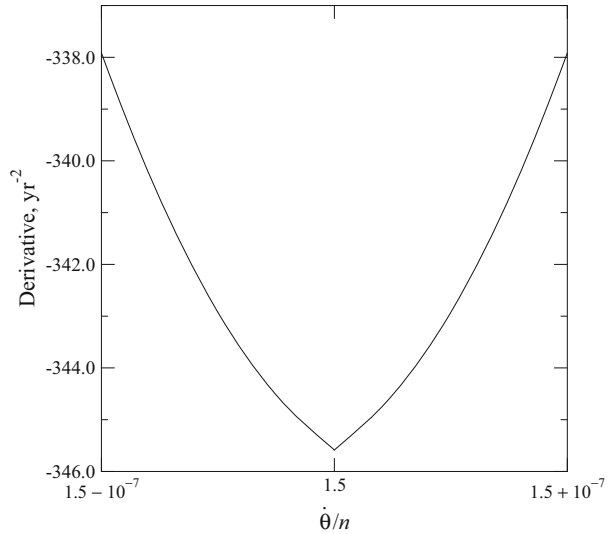
**Fig. 5** Further magnified version of the ‘kink’ around  $\dot{\theta}/n = 3/2$ , suggesting that the derivative is finite for all values of  $\dot{\theta}/n$



at least, not to be simply periodic—we shall have more to say about these solutions in Bartuccelli et al. (2017). Here, we just point out that we distinguish the different captured solutions from each other solely by their mean  $\dot{\theta}$  values,  $\langle \dot{\theta} \rangle$ , defined by  $\langle \dot{\theta} \rangle = \lim_{T \rightarrow \infty} T^{-1} \int_0^T \dot{\theta}(t) dt$ ; and that capture can take place in solutions for which a term in the sum (N.10), the triaxiality torque, is approximately zero, *i.e.* for  $\theta \approx (q + 2)nt/2$ ,  $q \in Q^{\text{TRI}}$ . Hence, we only expect solutions for which  $\langle \dot{\theta}/n \rangle \approx -1, -1/2, 1/2, \dots, 7/2$  and 4.

In order to compute probabilities of capture by the different solutions, with small confidence intervals, via Monte Carlo simulation, many solutions to Eq. (3), starting from random

**Fig. 6** The derivative of the tidal angular acceleration with respect to the relative rate of rotation, that is,  $n \, d\dot{\theta}^{\text{TIDE}}/d\dot{\theta}$ , showing the cusp at  $\dot{\theta}/n = 3/2$ . Since  $\ddot{\theta}^{\text{TIDE}}$  is  $C^1$ , its first derivative exists for all  $\dot{\theta}$



initial conditions in the set  $\mathcal{Q} = [0, \pi] \times [\dot{\theta}_{\min}, \dot{\theta}_{\max}]$ , must be computed. Many are needed because, for a given confidence level, the width of the confidence interval is proportional to  $I^{-1/2}$ , where  $I$  is the number of solutions computed: more simulations narrow the confidence interval, but rather slowly.

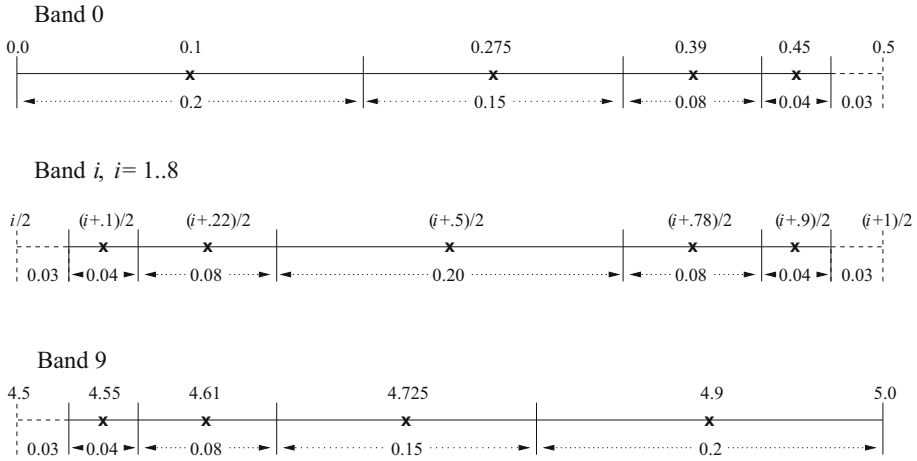
We take  $\mathcal{Q} = [0, \pi] \times [0, 5n]$  in what follows. The triaxiality acceleration depends on  $2\theta$ , and so we need only consider  $\theta_0 \in [0, \pi]$ ; and the right-most kink is at  $\dot{\theta} = 9n/2$ , so we choose the maximum value of  $\dot{\theta}_0$  to be somewhat greater than this.

We now describe a fast algorithm for solving Eq. (3), which is based on the high-order Euler method (HEM), described in detail in Bartucelli et al. (2015). This method has to be adapted for the NFME model because of the discontinuities in the second derivative of  $\dot{\theta}^{\text{TIDE}}(\dot{\theta})$ —that is, at the centres of the kinks—that occur at  $\dot{\theta}/n = 1/2, 1, 3/2, \dots, 9/2$ . The adaptation we use requires that the subset  $\mathcal{Q}$  of the  $(\theta, \dot{\theta})$  phase plane be split into strips, all with  $\theta \in [0, \pi)$ , but with a variety of ranges of  $\dot{\theta}$ . This splitting is needed in order to meet the error criterion. A strip can be of type H (HEM), when it is far enough removed from the kinks that the HEM can be used, and type N (Numerical), surrounding a kink, where a suitable numerical method has to be used. In a type N strip, there is a possibility that the ODE is stiff, and the numerical method chosen should take this into account.

The details of this splitting are given in Fig. 7, in which  $\dot{\theta}/n$  is plotted horizontally, with type H regions being shown as continuous lines, type N as dashed lines. The strips are grouped into ten bands, one band covering the region between two kinks. For instance, Band 0 consists of four type H strips, centred on  $\dot{\theta}/n = 0.1, 0.275, 0.39$  and  $0.45$  and with widths  $0.2, 0.15, 0.08$  and  $0.04$ , respectively; and one type N strip, where  $\dot{\theta}/n \in [0.47, 0.5]$ ,  $\dot{\theta} = 0.5n$  being the position of the first kink. The reasoning behind the choice of these values is given in Sects. 4.1 and 4.2. The expansion point  $\theta_e$  for the series solution is always the centre of the strip, and is shown by ‘x’ in Fig. 7.

### 4.1 Region H: HEM applies

The HEM, which is a fixed timestep implementation of the Frobenius method, can in principle be used to solve Eq. (3) in regions of the phase plane,  $(\theta, \dot{\theta})$ , where both the functions  $\ddot{\theta}^{\text{TRI}}$



**Fig. 7** The different strips used for H (HEM, *solid line*) and N (numerical, *dashed line*) computation. Numerical values of  $\hat{\theta}/n$  are shown. A ‘x’ shows the expansion point for a series solution. The width of a region in terms of  $\hat{\theta}/n$  over which the series solution is valid, for all  $\theta$ , to within the error bounds (see text) is shown below the line. The strips are grouped into bands. In Bands 0 and 9, there are four H strips and one N strip; in the remaining bands, there are five H, and two N strips

and  $\ddot{\theta}^{\text{TIDE}}$  are analytic. In other words, the HEM is a numerically implemented analytical continuation of  $\theta(t)$  on a set of overlapping discs of fixed radius in the complex  $t$ -plane (hence, fixed timestep), where derivatives of  $\theta$  of any order are computed from the differential equation (3). Now, from its definition,  $\ddot{\theta}^{\text{TRI}}(\theta, t)$  is an entire function for all  $\theta, t$ , so this is no bar to using the HEM. However,  $\ddot{\theta}^{\text{TIDE}}(\hat{\theta})$  is  $C^1$  in  $\hat{\theta}$ , with its second derivative being undefined at  $\hat{\theta}/n = 1/2, 1, 3/2, \dots, 9/2$ . Hence, we approximate  $\ddot{\theta}^{\text{TIDE}}$  by its Taylor series of degree  $D_{\text{TID}}$  about a point  $\hat{\theta} = \hat{\theta}_e$ , bearing in mind that this will only be good for values of  $\hat{\theta}$  far from the kinks. In the work reported here, our error criterion is satisfied with  $D_{\text{TID}} = 25$  for all strips.

With these provisos, the HEM works well. We briefly review the method here; full details can be found in Bartuccelli et al. (2015).

Let the state vector  $\mathbf{x}(t) = (\theta(t), \dot{\theta}(t))$  and define  $t_i = ih, i \in \mathbb{N}_0$ , where  $h > 0$  is a timestep. The ODE allows us to compute, recursively, derivatives of  $\theta(t)$  of all orders, far from the kinks. Let the  $j$ -th time derivative of  $\theta(t)$  be written as  $\theta^{(j)}(t)$ . We then write the degree- $D_s$  series solution to Eq. (3) as

$$\mathbf{x}(t_i) = \mathbf{p}_i(\mathbf{x}(t_{i-1})) = \mathbf{x}(t_{i-1}) + \sum_{j=1}^{D_s} \frac{h^j}{j!} \mathbf{f}_j(\mathbf{x}(t_{i-1}), t_{i-1}) + O(h^{D_s+1}), \quad (5)$$

where  $D_s > 1$  ( $D_s = 1$  gives the Euler method), and  $\mathbf{f}_j(\mathbf{x}, t) = (\theta^{(j)}(t), \theta^{(j+1)}(t))$ . Note that  $\mathbf{f}(\mathbf{x}, t)$  depends on  $t$  as well as  $\mathbf{x}$  because the ODE is non-autonomous. For the same reason,  $\mathbf{p}$  depends on  $i$ . To satisfy the error criterion, the following values of  $D_s$  are used in the various bands:

Band	0	1	2	3	4	5	6	7	8	9
$D_s$	16	15	15	14	14	14	15	16	17	17

Equation (5) allows us to advance the solution by a time  $h$ . Our immediate objective, however, is to compute the Poincaré map for the ODE in an efficient way. Let us denote the period

in  $t$  of  $\dot{\theta}^{\text{TRI}}$  as  $T_0 = 2\pi/n$  (which is one ‘Mercury year’), and put  $\mathbf{x}_k = \mathbf{x}(kT_0)$ . Then, the Poincaré map,  $\mathbf{P}$ , is defined by  $\mathbf{x}_{k+1} = \mathbf{P}(\mathbf{x}_k)$ . Iteration of  $\mathbf{P}$  many times starting from a given initial condition  $\mathbf{x}_0 = (\theta(0), \dot{\theta}(0))$  enables us to generate a sequence of ‘snapshots’ of the state variables as the system evolves, from which we can deduce the attractor in which the system is eventually captured. For details of capture detection, see below. Using a large number  $I$  of initial conditions, one can then estimate the probability of capture in any of the possible steady-state solutions.

In order to satisfy the error criterion, timestep  $h$  needs to be sufficiently small. As with all the parameters mentioned so far, there is a trade-off between speed and accuracy; a choice of  $h = T_0/M$  with  $M = 24$  is a good compromise in practice. Finally, then, the map  $\mathbf{P}$  can be built up from the functions  $\mathbf{p}_i$  by  $\mathbf{P}(\mathbf{x}) = \mathbf{p}_M \circ \mathbf{p}_{M-1} \circ \dots \circ \mathbf{p}_1(\mathbf{x})$ , and this is the HEM.

The expressions for  $\mathbf{p}_i$  are derived by computer algebra and are initially very large, but most of the terms are negligible and hence can be pruned away. By ‘term’ we mean here a polynomial in  $\dot{\theta}$ —these are then multiplied by powers of  $\sin 2\theta, \cos 2\theta$  in order to make up  $\mathbf{p}_i$ . In practice, every term whose magnitude is less than  $10^{-16}$  at the largest and smallest values of  $\dot{\theta}$  within a strip is deleted. The value of  $10^{-16}$  is chosen because the usual double precision arithmetic is carried out to around 16 significant figures. The resulting expression is then converted to Horner form (Press et al. 1992) for efficient evaluation. Typically, after pruning has been carried out, the expressions for  $\mathbf{p}_i$  contain 20–35 terms and are typically of the form:  $\theta$ -component =  $\theta + r_1 + sr_2 + cr_3 + scr_4 + s^2r_5$ ,  $\dot{\theta}$ -component =  $r_6 + sr_7 + cr_8 + scr_9 + s^2r_{10} + s^2cr_{11} + s^3r_{12}$ , where the  $r_k$  are polynomials in  $\dot{\theta}(t_{i-1})$  of degree up to about 12, and  $s = \sin 2\theta, c = \cos 2\theta$ . An explicit example follows; this is the  $\theta$ -component of  $\mathbf{p}_1(x, y)$ , where  $(x, y) = (\theta(0), \dot{\theta}(0))$ . It is valid for the middle strip of Band 3 (see Fig. 7), that is, for  $\dot{\theta} \in [1.55, 1.95]$ , and integrates the  $\theta$  component of the solution from  $t = 0$  to  $t = T_0/24 \approx 0.0100353$ . The expression, in Horner form, is

$$\begin{aligned} \theta(T_0/24) = & x + 0.458 + (0.0100 + (-0.304e - 12 + (0.490e - 13 + (-0.357e - 14 \\ & + (0.754e - 15 + (-0.605e - 16 \\ & + (0.137e - 16 + (-0.115e - 17 + (0.267e - 18 + (-0.228e - 19 \\ & + 0.542e - 20y)y)y)y)y)y)y)y)y)y \\ & + (-0.950e - 5 + (0.345e - 8 + (0.316e - 9 + (-0.929e - 13 \\ & + (-0.421e - 14 + 0.124e - 17y)y)y)y) \\ & + (-0.261e - 11 + (-0.477e - 12 + (0.487e - 15 + 0.295e - 16y)y) \\ & + 0.150e - 15s)s) \\ & + (-0.344e - 6 + (-0.632e - 7 + (0.208e - 10 \\ & + (0.126e - 11 + (-0.333e - 15 - 0.121e - 16y)y)y)y) \\ & + (0.299e - 10 + (-0.480e - 13 - 0.436e - 14y)y)s)c, \end{aligned}$$

evaluated to 3 significant figures (in practice, it would need to be evaluated to 18 s.f.). Note that there are 48 strips and 24 integrators in each strip, one for each timestep, and two sets of integrators, one for  $\theta$  and one for  $\dot{\theta}$ . Hence, the total number of integrators is  $48^2 = 2304$ .

We now describe the error criterion used. The final version of the code for computing the Poincaré map via the HEM is compared to a high-accuracy, standard numerical computation of the same thing. The full expression for  $\dot{\theta}^{\text{TIDE}}$  is used in the accurate numerical computation, not its series approximation. The numerical algorithm used is the standard Runge–Kutta method as implemented in the computer algebra software Maple, computing to 25 significant figures and with absolute and relative error tolerances of  $10^{-15}$ . The numerical and HEM

computations of  $\mathbf{P}(\mathbf{x})$  are then compared, for each of the type H strips, using 250 uniformly distributed random values of  $\mathbf{x}$  in each strip. The comparison gives the maximum value of the modulus of the difference between each component, computed both ways, over the 250 random points. The maximum difference observed over all strips is assumed to be representative of the overall maximum difference. Its values are about  $3 \times 10^{-14}$ ,  $1.4 \times 10^{-13}$  in the  $\theta$ - and  $\hat{\theta}$ -components, respectively.

### 4.2 Region N: numerical method must be used

From Fig. 7, it can be seen that the HEM can be used for about 89% of  $\hat{\theta} \in [0, 5n]$ , but in the strips surrounding the kinks,  $\theta \in [0, \pi)$ ,  $\hat{\theta}/n \in [(q + 2)/2 - 0.03, (q + 2)/2 + 0.03]$ ,  $q \in \mathcal{Q}^{\text{TIDE}}$ , the type N strips, a purely numerical method has to be used.

It is possible that the ODE may be stiff in type N strips, so we choose two numerical methods and compare the results. The methods used are: (1) an explicit, order 7 Runge–Kutta (RK) method due to Dormand and Price, as described in Hairer et al. (1993), and (2) the Adams method/backward differentiation formulae (BDF) (Radhakrishnan and Hindmarsh 1993), with the ability to switch automatically between them. The Adams method is a variable-order (up to 12 in the implementation used here), explicit predictor–corrector method, which, along with RK, is suitable for non-stiff problems, whereas BDF is suitable for stiff problems.

In practice, with the parameters in Table 1, BDF is not used—the problem turns out to be insufficiently stiff—so the comparison between (1) and (2) above comes down to comparing the RK and Adams methods. The implementation of Adams used (Radhakrishnan and Hindmarsh 1993) is approximately 1.6 times slower than RK for this problem, but the probabilities obtained from integration starting from the same 3200 random points in  $\mathcal{Q}$  are in good agreement—see Table 2—so, for type N strips, we choose RK, for which we fix both the absolute and relative error tolerances to be  $2 \times 10^{-14}$ . This value is chosen to be comparable with the error entailed by polynomial interpolation.

The probabilities obtained with RK and Adams are not identical, neither should we expect them to be. The long-term fate of a given trajectory depends very sensitively on the details of its computation, implemented for very long integration times. What is important in the end

**Table 2** Comparison of probability of capture by the eight attractors with  $\langle \hat{\theta}/n \rangle = 1/2, 1, \dots, 4$

Runge–Kutta			Adams		
$\langle \hat{\theta}/n \rangle$	Probability (%)	95% CI	$\langle \hat{\theta}/n \rangle$	Probability (%)	95% CI
1/2	0.8125	0.3110	1/2	0.8750	0.3227
1	27.44	1.546	1	27.38	1.545
3/2	43.44	1.717	3/2	43.81	1.719
2	22.03	1.436	2	21.91	1.433
5/2	5.063	0.7596	5/2	5.094	0.7618
3	1.094	0.3604	3	0.750	0.2989
7/2	0.031	(0.0612)	7/2	0.094	(0.1060)
4	0.094	(0.1060)	4	0.094	(0.1060)

These were computed using the Runge–Kutta and the Adams method in N-type strips, with HEM being used elsewhere. The same 3200 uniformly distributed random initial conditions were used in both cases. Confidence intervals in brackets are unreliable since there are too few data points in the case of these attractors

**Table 3** Timings for the computation of  $10^5$  iterations of the Poincaré map,  $\mathbf{P}$ , in various circumstances

Method	Strip	Mean time for $10^5$ iterations of $\mathbf{P}$ , CPU-sec
Runge–Kutta, slow $\ddot{\theta}^{\text{TIDE}}$	N	61.1
Runge–Kutta, fast $\ddot{\theta}^{\text{TIDE}}$	N	20.5
HEM	H	0.313

These are: type H or type N strips, and, in type N strips, with or without the fast computation of  $\ddot{\theta}^{\text{TIDE}}$  described in Sect. 4.2. The mean was taken over 895/105 random initial conditions for type H/N strips respectively. If required, timings for the Adams method can be obtained by multiplying Runge–Kutta timings by 1.6

is the probabilities themselves, and Table 2 shows these to be robust against the algorithm used to solve the ODE.

Computation time can be saved by efficient calculation of  $\ddot{\theta}^{\text{TRI}}$  and  $\ddot{\theta}^{\text{TIDE}}$ . The expression for  $\ddot{\theta}^{\text{TRI}}$  can be converted into a polynomial of degree 1 in  $\cos 2\theta$  and  $\sin 2\theta$ , and degree 8 in  $\cos nt/2$  and  $\sin nt/2$ . In Horner form, this polynomial can be evaluated efficiently, using 2 sin/cos evaluations, 16 addition and 36 multiplication operations.

Evaluation of  $\ddot{\theta}^{\text{TIDE}}(\dot{\theta})$  in the obvious way is computationally expensive, since it requires the calculation of one fractional power per term in Eq. (N.11b)—nine in all. A more efficient way to evaluate it is:

- Case 1** if  $\dot{\theta}$  is far from a kink, then use a pre-computed Chebyshev polynomial fit (Press et al. 1992) to  $\ddot{\theta}^{\text{TIDE}}$ —the function is very smooth here;
- Case 2** if  $\dot{\theta}$  is close to a kink, compute the contribution to  $\ddot{\theta}^{\text{TIDE}}$  from that kink exactly, according to the appropriate single term in the sum (N.11b), with the effect of the remaining kinks being replaced by a Chebyshev polynomial fit.

Hence, at most one fractional power is computed per evaluation of  $\ddot{\theta}^{\text{TIDE}}$ . In practice, we use a polynomial only (Case 1 above), of degree 25, unless  $2\dot{\theta}/n$  is within  $\pm 0.08$  of an integer, when Case 2 applies. In Case 2, we use a polynomial of degree 7 to fit the remaining terms. The resulting absolute error is no more than  $4 \times 10^{-14}$ . For comparison, the ratio of the CPU-time taken to evaluate  $\ddot{\theta}^{\text{TIDE}}$  directly, and via polynomial fitting, is about 5.1.

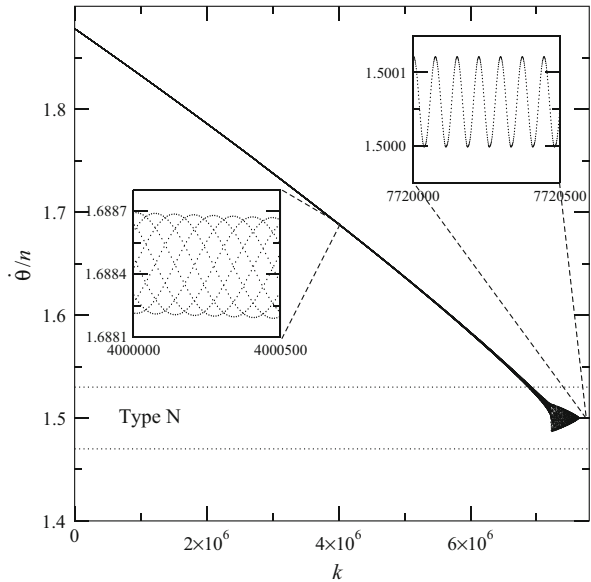
A comparison of the timings in different circumstances is given in Table 3, for which 1000 random initial conditions were used. The CPU-time taken to iterate  $\mathbf{P}$   $10^5$  times starting from each of these was measured, with any data in which the trajectory moved from a type H strip to one of type N, or vice versa, being rejected. As mentioned earlier in this section, the Adams method, for this problem at least, is about 1.6 times slower than RK.

### 4.3 Capture test

We now describe the test used to detect capture of a solution. Figure 8 shows  $\dot{\theta}_k = \dot{\theta}(kT_0)$  plotted against  $k$  for  $0 \leq k \leq 7.8 \times 10^6$ . In this case, the initial condition was  $\mathbf{x}_0 = (0, 49)$  and capture took place after about  $7.7 \times 10^6$  iterations of the Poincaré map, which corresponds to  $7.7 \times 10^6 T_0 = 1.85 \times 10^6$  year.

There are several ways that capture could be detected. We choose to divide the  $\dot{\theta}_k$  dataset into blocks of length  $L$  and compute the least squares gradient of each block. As can be seen from Fig. 8, this gradient will be negative pre-capture, and close to zero post-capture.

**Fig. 8** An illustration of capture dynamics:  $\dot{\theta}_k$ , with initial condition  $\mathbf{x}_0 = (0, 49) = (0, 1.878n)$ , is plotted against  $k$ , for  $0 \leq k \leq 7.8 \times 10^6$ . In this case, capture takes place in a solution with  $(\dot{\theta}/n) \approx 3/2$ . Two small regions are shown on a magnified scale. On the *left*, we see the pre-capture dynamics and on the *right*, post-capture. The HEM was used for all computations outside the strip  $1.47 \leq \dot{\theta}/n \leq 1.53$ , marked ‘Type N’



In detail the test is as follows. Let  $\bar{y}_j$  be the mean value of  $\dot{\theta}_k/n$  over the  $j$ -th block of length  $L$ , and let  $m_j$  be the least squares gradient of  $\dot{\theta}_k$  plotted against  $k$  in block  $j$ . Capture is deemed to have taken place when

$$|2\bar{y}_j - [2\bar{y}_j]| < \varepsilon_i \quad \text{and} \quad |m_j| < \varepsilon_m \quad \text{for } K \text{ successive blocks.}$$

Here,  $[x]$  is the nearest integer to  $x$ , and the factor of two occurs in the first expression because capture can take place at integer or half-integer values of  $\dot{\theta}/n$ . In practice, we choose  $L = 10,000$ ,  $K = 8$ ,  $\varepsilon_i = 10^{-3}$  and  $\varepsilon_m = 3 \times 10^{-7}$ . The speed of a probability computation depends on the choice of these parameters.

Figure 9 shows a typical plot of  $m_j$  versus block number  $j$ , from  $j = 0$  until capture, which takes place at  $j = 770$ .

### 5 A probability of capture computation

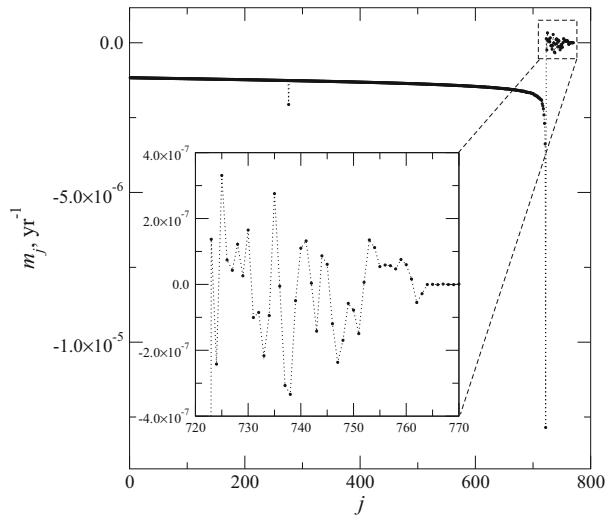
The main purpose of this paper has been to show how the computation of capture probabilities can be done relatively fast. We now wish to show how this works in practice, by producing probability of capture data for a particular case, a case chosen to allow the algorithm to exhibit its effectiveness. To this end, we now give the results of a calculation of probability of capture for the parameters in Table 1, which are of intrinsic interest as well as being a useful illustration.

As in Bartuccelli et al. (2015), we use the CPU-sec as a unit of time, which is defined in terms of the following sum:

$$S(m) = \sum_{i=1}^m \frac{(i+1)(i+3)}{i(i+2)(i+4)(i+6)},$$



**Fig. 9** An illustration of the capture test in practice. The least squares gradient of  $\hat{\theta}_k, m_j$ , versus the block number,  $j$ , for the trajectory in Fig. 8. Blocks are of length 10,000 iterations of the Poincaré map and if one of the criteria for capture is that  $|m_j| < 3 \times 10^{-8}$  for eight successive blocks, then this happens in blocks 763–770 inclusive. Hence, capture in this case is detected after  $7.7 \times 10^6$  iterations of the Poincaré map



whose evaluation requires  $5m$  multiplications and  $6m - 1$  additions. We define 1 CPU-sec to be the CPU-time taken to evaluate  $S(N_c)$ , where  $N_c = 5.96 \times 10^7$ , which is the CPU-time taken to do this computation on the computer used to do some of the calculations in this paper. The time taken can vary according to circumstances, e.g. the loading, the type of processor and so on. Hence, care has to be taken in the codes to scale the CPU-sec appropriately for the particular hardware used: the computation of  $S(N_c)$  is timed after every successful capture, and the CPU-sec scaling factor is updated on each occasion.

One other point to note is that a computation of probability is trivial to parallelise. The initial conditions in  $\mathcal{Q}$  are generated by a pseudo-random number generator, and a different sequence of pseudo-random numbers can be produced just by changing the seed. Hence, by running  $N$  copies of the programme on  $N$  separate processors with  $N$  different seeds,  $N$  times the number of initial conditions can be investigated at the same time.

The CPU-time taken to integrate until capture depends strongly on the initial condition, and in particular, the proportion of the integration that is carried out in type N strips (which is a slow process) as opposed to type H strips (where it is fast). By considering 57,600 random initial conditions, we find the following:

Mean time to capture: 1156 CPU-sec, with standard deviation 1190 CPU-sec

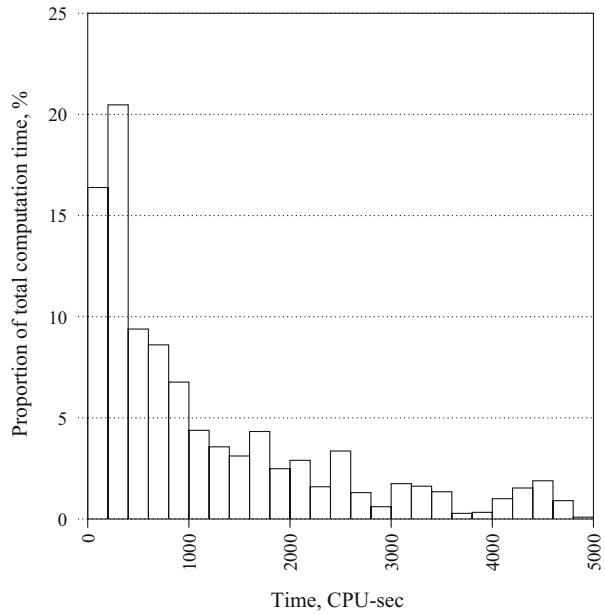
Maximum time to capture: 5161 CPU-sec; minimum time to capture: 1.188 CPU-sec.

Total iterations in type N strips:  $\sim 1.6 \times 10^{11}$ ; total iterations in type H strips:  $\sim 1.2 \times 10^{12}$ .

From the above, we see that overall, about  $1.4 \times 10^{12}$  iterations of the Poincaré map were needed to estimate the probability of capture for 57,600 initial conditions, and that 12% of these were in type N strips, with the remaining 88% being in type H strips. Since capture can only take place in the vicinity of a kink, iteration until capture always requires some iteration in N strips. The values given above have a sensitive dependence on the capture parameters.

We can now estimate the factor by which our approach speeds up a typical probability of capture computation. Let  $T_{\text{mix}}(k)$  be the CPU-time taken to iterate the Poincaré map  $k$  times using HEM in type H strips and the chosen numerical method in type N strips, and let

**Fig. 10** Histogram showing the probability that the computation time for iteration to capture lies in various ranges. The interpretation is that, for instance, a CPU-time of between 0 and 200 CPU-sec is required to establish the capture of about 16.3% of trajectories (*leftmost bar*). Four values of CPU-time >5000 CPU-sec were excluded from the histogram, and so the dataset contains 57,596 points. The modal value of CPU-time is in the range 200–400 CPU-sec



$T_{\text{num}}(k)$  be the time taken when using the numerical method everywhere. Then, using the data above and from Table 3, we estimate that

$$\frac{T_{\text{mix}}(k)}{T_{\text{num}}(k)} \approx \frac{k(0.12 \times 20.5 + 0.88 \times 0.313)10^{-5}}{20.5k \times 10^{-5}} \approx \frac{2}{15}.$$

Our approach therefore speeds up this computation by a factor of about 7.5.

A histogram of capture times is given in Fig. 10. Let  $\tau(\mathbf{x}_0)$  be the number of CPU-sec required to iterate the Poincaré map, starting from  $\mathbf{x}_0$ , until the capture criterion of Sect. 4.3 is met. In order to produce Fig. 10, 57,600 random values of  $\mathbf{x}_0 \in \mathcal{Q}$  were generated and  $\tau(\mathbf{x}_0)$  was computed for each. This figure is a histogram of the proportions of the values of  $\tau(\mathbf{x}_0)$  that lie in the ranges 0–200, 200–400, ..., 4800–5000 CPU-sec.

Finally, we give two estimates of the probabilities themselves in Table 4. The 95% CI (Walpole et al. 1998) is defined such that the probability of capture in a given attractor  $A$  lies in the interval  $[\hat{p} - \Delta p, \hat{p} + \Delta p]$ , with 95% confidence. Here,  $\Delta p = 1.96\sqrt{\hat{p}(1 - \hat{p})/I}$ , in which  $I = 5.76 \times 10^4$  is the number of initial conditions for the main results, and  $\hat{p}$  is the proportion of these initial conditions that end up in  $A$ . For these values, in the worst case, which is  $\hat{p} = 1/2$ ,  $\Delta p \approx 0.41\%$ . We also give, for comparison, an estimate of the same probabilities but obtained by using RK only, with  $I = 3600$ . The smaller number of points is due to the much slower computation speed when HEM is not used, but the probabilities obtained are in agreement. The mean computer time to capture for the slow computation was 9055 CPU-sec—*c.f.* 1156 CPU-sec when HEM is used, the latter being quicker by a factor of 7.8, in close agreement with the claimed speed-up factor of 7.5.

**Table 4** Probability of capture by various attractors with  $(\dot{\theta}/n) = 1/2, 1, \dots, 4$ , computed in two ways: (a), columns 2–4, using the Runge–Kutta method in N-type strips, with HEM being used elsewhere; and (b), columns 5–7, using Runge–Kutta everywhere

Mean velocity, $(\dot{\theta}/n)$	(a) 57,600 points, RK and HEM			(b) 3600 points, RK only		
	Number	Probability, $\hat{p}$ (%)	95% CI	Number	Probability, $\hat{p}$ (%)	95% CI
1/2	349	0.606	0.06	16	0.444	0.22
1	16292	28.29	0.37	1056	29.33	1.5
3/2	24344	42.26	0.40	1485	41.25	1.6
2	13252	23.01	0.34	808	22.44	1.4
5/2	2775	4.818	0.18	202	5.611	0.75
3	473	0.8212	0.07	28	0.778	0.29
7/2	89	0.1545	0.03	4	0.111	(0.11)
4	26	0.0451	0.02	1	0.028	(0.05)
9/2	0	0.0000	0.00	0	0	0

Confidence intervals in brackets are not meaningful owing to the small number of points. For (a) and (b), respectively, 57,600 and 3,600 uniformly distributed random initial conditions in  $\mathcal{Q} = [0, \pi] \times [0, 5n]$  were used. Fewer points were used in (b) because of the significantly longer computation time needed. As expected, 9/2 was never observed

## 6 Conclusions

The main purpose of this paper has been to show how the computation of probability of capture in various solutions to the spin-orbit ODE, in the case of constant eccentricity  $e$  and with a particular model of dissipation (Noyelles et al. 2014), can be accelerated.

Estimation of these probabilities requires us to solve the ODE (3), and bearing in mind that this is periodic in time with period  $T_0 = 2\pi/n$ , our objective is to compute the Poincaré map, which advances the state variables  $\mathbf{x} = (\theta, \dot{\theta})$  by a time  $T_0$ , as fast as possible. This enables us to generate a sequence of values  $\mathbf{x}_k$ , from which salient information about the dynamics can be deduced. In particular, we have in mind the application of this work to estimation of probability of capture in various orbits, and the confidence intervals of these probabilities shrink as more orbits are investigated: hence the need for a fast algorithm.

The dissipation term,  $\ddot{\theta}^{\text{TIDE}}(\dot{\theta})$ , which is a  $C^1$  function of the angular velocity  $\dot{\theta}$ , and additionally varies rapidly in the ranges of  $\dot{\theta}$  around the so-called kinks, complicates the computation, and requires the use of a standard numerical ODE solver for some of the time. We point out ways in which these purely numerical computations can be carried out efficiently, by streamlining the calculation of the triaxiality and tidal accelerations. However, in about 88% of the phase plane the high-order Euler method (Bartuccelli et al. 2015) can be used, and this speeds up the computation significantly. Additionally, the triaxiality term, being a smooth function of  $\theta$  and  $t$ , is taken into account in the HEM at almost no additional computational cost.

The present case should be contrasted with the constant time lag model, in which dissipation is just proportional to  $\dot{\theta} - \omega$  with  $\omega$  a constant. In the light of its simplicity, this has been used in many publications, for instance, Bartuccelli et al. (2015) and references therein. In that case, only a single set of integrators  $\mathbf{p}_i$ , defined in Eq. (5), was needed to build up the Poincaré map. In other words, there was only one strip, which was the whole of  $\mathcal{Q}$ , the subset of the phase plane considered, and the HEM could be used everywhere.

Compare that with the current case using the parameters in Table 1, which relate to Mercury and the Sun. Here,  $\mathcal{Q}$  has to be divided into 48 strips and an integrator set up for each. Setting up the codes for the HEM in each strip, which is done by computer algebra, itself takes significant time, but the payoff is an increase in speed by a factor of approximately 65 in the strips where HEM can be used, compared to using a standard numerical method.

A probability of capture computation requires a capture detection algorithm, and one has been described. It is based on the fact that after capture, the angular velocity has no underlying decay: a captured solution has a close to constant mean value of  $\dot{\theta}_k = \dot{\theta}(kT_0)$  for all  $k$  greater than the value at which capture takes place. Typically, the mean is taken over  $10^4$  successive  $k$  values.

A test run of our algorithm, in which 57,600 initial conditions were iterated until capture took place, reveals that the overall speed of computation is faster by a factor of about 7.5 compared to using a standard numerical algorithm alone.

An important question left unanswered in this paper is ‘What is the nature of the solutions in which capture takes place?’ The answer turns out to be more complicated than expected: periodic solutions with mean  $\dot{\theta}/n \approx -1, -1/2, 1/2, \dots, 4$  have been computed (high-accuracy numerics are needed). There is also numerical evidence for the existence of attracting solutions whose period is not a small integer multiple of  $T_0$ . The dynamics of solutions to Eq. (3), both pre- and post-capture, will be subject of a future publication (Bartuccelli et al. 2017).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Bartuccelli, M.V., Deane, J.H.B., Gentile, G.: The high-order Euler method and the spin-orbit model: a fast algorithm for solving differential equations with small, smooth nonlinearity. *Celest. Mech. Dyn. Astron.* **121**(3), 233–260 (2015). doi:[10.1007/s10569-014-9599-7](https://doi.org/10.1007/s10569-014-9599-7)
- Bartuccelli, M.V., Deane, J.H.B., Gentile, G.: Periodic and quasi-periodic attractors for the spin-orbit model with a realistic tidal torque. (2017) **(in preparation)**
- Celletti, A., Chierchia, L.: Measures of basins of attraction in spin-orbit dynamics. *Celest. Mech. Dyn. Astron.* **101**(1–2), 159–170 (2008)
- Celletti, A., Chierchia, L.: Quasi-periodic attractors in celestial mechanics. *Arch. Ration. Mech. Anal.* **191**(2), 311–345 (2009)
- Correia, A.C.M., Laskar, J.: Mercury’s capture into the 3/2 spin-orbit resonance as a result of its chaotic dynamics. *Nature* **429**, 848–850 (2004)
- Duriez, L.: Cours de mécanique Céleste classique. (2007). [http://lal.univ-lille1.fr/mecacel\\_duriez/CoursMCEcr\\_Duriez](http://lal.univ-lille1.fr/mecacel_duriez/CoursMCEcr_Duriez)
- Ferraz-Mello, S.: Tidal synchronization of close-in satellites and exoplanets. A rheophysical approach. *Celest. Mech. Dyn. Astron.* **116**, 109–150 (2013)
- Ferraz-Mello, S.: Tidal synchronization of close-in satellites and exoplanets: II. Spin dynamics and extension to Mercury and exoplanet host stars. *Celest. Mech. Dyn. Astron.* **122**, 359–389 (2015)
- Goldreich, P., Peale, S.: Spin-orbit coupling in the solar system. *Astron. J.* **71**(6), 425–438 (1966)
- Hairer, E., Nørsett, S.P., Wanner, G.: Solving Ordinary Differential Equations I. Nonstiff Problems, 2nd revised edition. Springer Series in Computational Mathematics, Springer, Berlin (1993). ISBN 978-3-540-56670-0. Code available from <http://www.unige.ch/~hairer/software.html>
- Makarov, V.V., Bergeha, C., Efroimsky, M.: Dynamical evolution and spin-orbit resonances of potentially habitable exoplanets: the case of GJ 581d. *Astrophys. J* **761**(2), 83 (2012)

- Makarov, V.V.: Conditions of passage and entrapment of terrestrial planets in spin-orbit resonances. *Astrophys. J* **752**(1), 73 (2012)
- Noyelles, B., Frouard, J., Makarov, V.V., Efroimsky, M.: Spin-orbit evolution of Mercury revisited. *Icarus* **241**, 26–44 (2014)
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: *Numerical Recipes in C*. Cambridge University Press, Cambridge (1992). ISBN 0-521-43108-5
- Radhakrishnan, K., Hindmarsh, A.C.: Description and Use of LSODE, the Livermore Solver for Ordinary Differential Equations, NASA Reference Publication 1327, Lawrence Livermore National Laboratory Report UCRL-ID-113855. (1993). <https://computation.llnl.gov/casc/nsde/pubs/u113855>. Code available from <http://lh3lh3.users.sourceforge.net/solveode.shtml>
- Walpole, R.E., Myers, R.H., Myers, S.L.: *Probability and Statistics for Engineers and Scientists*. Prentice Hall, Upper Saddle River (1998). ISBN 0-13-840208-9