



Large-stepsize integrators for charged-particle dynamics over multiple time scales

Ernst Hairer¹ · Christian Lubich² · Yanyan Shi^{2,3,4}

Received: 25 January 2021 / Revised: 9 April 2022 / Accepted: 10 May 2022

© The Author(s) 2022

Abstract

The Boris algorithm, a closely related variational integrator and a newly proposed filtered variational integrator are studied when they are used to numerically integrate the equations of motion of a charged particle in a mildly non-uniform strong magnetic field, taking step sizes that are much larger than the period of the Larmor rotations. For the Boris algorithm and the standard (unfiltered) variational integrator, satisfactory behaviour is only obtained when the component of the initial velocity orthogonal to the magnetic field is filtered out. The particle motion shows varying behaviour over multiple time scales: fast gyrorotation, guiding centre motion, slow perpendicular drift, near-conservation of the magnetic moment over very long times and conservation of energy for all times. Using modulated Fourier expansions of the exact and numerical solutions, it is analysed to which extent this behaviour is reproduced by the three numerical integrators used with large step sizes that do not resolve the fast gyrorotations.

Keywords Charged particle · Strong magnetic field · Boris algorithm · Variational integrator · Filtered variational integrator · Modulated Fourier expansion · Long-term behaviour

✉ Christian Lubich
Lubich@na.uni-tuebingen.de

Ernst Hairer
Ernst.Hairer@unige.ch

Yanyan Shi
Shi@na.uni-tuebingen.de ; shiyanyan1995@lsec.cc.ac.cn

¹ Dept. de Mathématiques, Univ. de Genève, CH-1211, Genève 24, Switzerland

² Mathematisches Institut, Univ. Tübingen, D-72076 Tübingen, Germany

³ LSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

⁴ University of Chinese Academy of Sciences, Beijing 100049, China

1 Introduction

The time integration of the equations of motion of charged particles is a basic algorithmic task for particle methods in plasma physics [2]. In this paper we consider the case of a non-uniform strong magnetic field in the asymptotic scaling known as maximal ordering [4, 21], with a small parameter $\varepsilon \ll 1$ whose inverse corresponds to the strength of the magnetic field. The particle motion then shows different behaviour over multiple time scales:

- fast Larmor rotation over the time scale ε ,
- guiding centre motion over the time scale ε^0 ,
- slow drift perpendicular to the magnetic field over the time scale ε^{-1} ,
- near-conservation of the magnetic moment over time scales ε^{-N} with arbitrary $N > 1$,
- and energy conservation for all times.

In this paper we are interested in using numerical integrators with step sizes h that are much larger than the quasi-period $2\pi\varepsilon$ of the Larmor rotation. We thus have the two small parameters h and ε , which we will assume to be related by

$$0 < \varepsilon \leq h^2 \ll 1. \quad (1.1)$$

We study the behaviour of the numerical integrators over the time scales ε^0 , ε^{-1} , and ε^{-N} for $N > 1$.

The papers [23, 26] are similarly motivated by the objective to numerically integrate charged-particle dynamics accurately while stepping over the fast time scale of Larmor rotation. We are, however, not aware of any rigorous analysis of the error behaviour of numerical integrators in a large-stepsize regime in the existing literature. With an emphasis on different aspects, recent papers on numerical methods for charged-particle dynamics in a strong magnetic field include [5–7, 10–12, 15, 16, 24].

In Section 2 we formulate the equations of motion in the scaling considered here and illustrate the solution behaviour over various time scales.

In Section 3 we describe the three numerical integrators studied in this paper: the *Boris algorithm* [3, 9, 14, 22], a closely related *variational integrator* [15, 25], and a newly proposed *filtered variational integrator*, which only requires a minor algorithmic modification of the standard variational integrator and can be interpreted as the standard variational integrator for a Lagrangian with an anisotropically modified kinetic energy term.

In Section 4 we give modulated Fourier expansions of the exact solution and of the numerical solutions of the three numerical methods used with step sizes (1.1). The differential equations for the dominant modulation functions are the key to understanding the method behaviour over the time scales ε^0 and ε^{-1} for all three methods. For the Boris algorithm and the standard (unfiltered) variational integrator, the initial velocity needs to be modified such that its component perpendicular to the magnetic field is

$O(\varepsilon)$ -small. The complete modulated Fourier expansion will be used for studying the long-time near-conservation of the magnetic moment and energy for the filtered variational integrator.

In Section 5 we obtain $O(h^2)$ error bounds uniformly in ε for all three (formally second-order) numerical methods over the time scale ε^0 . This is not an obvious result for large step sizes (1.1) but here it follows directly from a comparison of the modulated Fourier expansions of the exact and numerical solutions.

In Section 6 we show that all three methods reproduce the perpendicular drift with an $O(h^2)$ or $O(h)$ error over the time scale ε^{-1} . This is again obtained via the modulated Fourier expansions, which also yield an $O(\varepsilon)$ approximation to the perpendicular drift by the solution of a slow differential equation over times $O(\varepsilon^{-1})$.

In Section 7 we consider the long-term energy behaviour. For the standard variational integrator with the modified starting velocity we prove near-conservation of the total energy up to time $O(\varepsilon^{-1})$. For the filtered variational integrator we prove near-conservation of magnetic moment and energy over times ε^{-N} with arbitrary $N > 1$ for non-resonant step sizes, using the Lagrangian structure of the modulation system. Moreover, we show results of numerical experiments for the energy behaviour of the three methods over long times.

The conclusion of our investigation is that the new filtered variational integrator with non-resonant large step sizes (1.1) reproduces the characteristic features well over all time scales, and this is fully explained by our theory. The Boris algorithm and the standard (unfiltered) variational integrator also work remarkably well for large stepsizes (1.1) on the time scales ε^0 and ε^{-1} in accordance with our theory, provided that the initial velocity is modified such that the component perpendicular to the magnetic field is reduced to size $O(\varepsilon)$. With this filtering of the starting velocity, the long-time energy behaviour of the Boris method and the standard variational integrator appears to be better in our numerical experiments than we can explain by theory.

2 Multiple time scales in the continuous problem

We study the time integration of the equations of motion of a charged particle in a strong magnetic field, with position $x(t) \in \mathbf{R}^3$ and velocity $v(t) = \dot{x}(t)$ at time t ,

$$\begin{aligned} \ddot{x}(t) &= \dot{x}(t) \times B(x(t)) + E(x(t)) \\ \text{with } B(x) &= \frac{1}{\varepsilon} B_0 + B_1(x) \quad \text{for } 0 < \varepsilon \ll 1, \end{aligned} \quad (2.1)$$

where B_0 is a fixed vector in \mathbf{R}^3 of unit norm, $|B_0| = 1$. The non-constant magnetic field $B_1(x)$ is assumed to have a known vector potential $A_1(x)$, i.e. $B_1(x) = \nabla_x \times A_1(x)$. This gives $B(x) = \nabla_x \times A(x)$ with the vector potential $A(x) = -\frac{1}{2}x \times B_0/\varepsilon + A_1(x)$. We always assume that $B_1 : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ and $E : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ are smooth with derivatives bounded independently of ε on bounded subsets of \mathbf{R}^3 . The above scaling corresponds to what is known as *maximal ordering* in the literature; see [4, 21].

For the initial position and velocity we always assume boundedness independently of ε :

$$|x(0)| \leq C_0, \quad |\dot{x}(0)| \leq C_1. \quad (2.2)$$

For studying the perpendicular drift, we need further assumptions on B_1 and E that are specified in Section 6. When it comes to studying the long-time energy behaviour, we further assume that the force field has a scalar potential, $E(x) = -\nabla\phi(x)$. The total energy is then

$$H(x, v) = \frac{1}{2}|v|^2 + \phi(x), \quad (2.3)$$

which is conserved along every trajectory and is bounded independently of ε under condition (2.2). We further consider the magnetic moment (rescaled with ε),

$$I(x, v) = \frac{1}{2\varepsilon} \frac{|v \times B(x)|^2}{|B(x)|^3}. \quad (2.4)$$

We note that, with v_\perp denoting the velocity component orthogonal to $B(x)$,

$$I(x, v) = \frac{1}{2} \frac{|v \times B_0 + O(\varepsilon)|^2}{1 + O(\varepsilon)} = \frac{1}{2}|v_\perp|^2(1 + O(\varepsilon)) + O(\varepsilon^2), \quad (2.5)$$

for (x, v) in any region that is bounded independently of ε . The magnetic moment is an adiabatic invariant: it is conserved up to $O(\varepsilon)$ over very long times $t \leq \varepsilon^{-N}$ with arbitrary $N > 1$; see e.g. [1, 15, 19, 20].

In Figure 1 we illustrate the solution behaviour on various time scales. We show the fast Larmor rotation of angular frequency ε^{-1} and amplitude $O(\varepsilon)$ on the time scale ε and the guiding centre motion on the time scale ε^0 in the first picture, and in addition the slow drift perpendicular to the magnetic field on the time scale ε^{-1} in the second picture (here: horizontal drift for the magnetic field in vertical direction). Finally, the third picture shows the long-time near-conservation of the magnetic moment and the conservation of energy. Our objective is to understand how the behaviour on the various time scales can be replicated by numerical methods with large time steps that do not resolve the fast Larmor rotations.

In Figure 1 we take the electromagnetic fields and the vector and scalar potentials as

$$B(x) = \frac{1}{\varepsilon} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} x_1(x_3 - x_2) \\ x_2(x_1 - x_3) \\ x_3(x_2 - x_1) \end{pmatrix} \quad \text{with} \quad A(x) = -\frac{1}{2} \begin{pmatrix} x_2 \\ -x_1 \\ 0 \end{pmatrix} + x_1x_2x_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

$$E(x) = -x \quad \text{with} \quad \phi(x) = \frac{1}{2}|x|^2,$$

and the initial values $x(0) = (0.3, 0.2, -1.4)^\top$ and $\dot{x}(0) = (-0.7, 0.08, 0.2)^\top$.

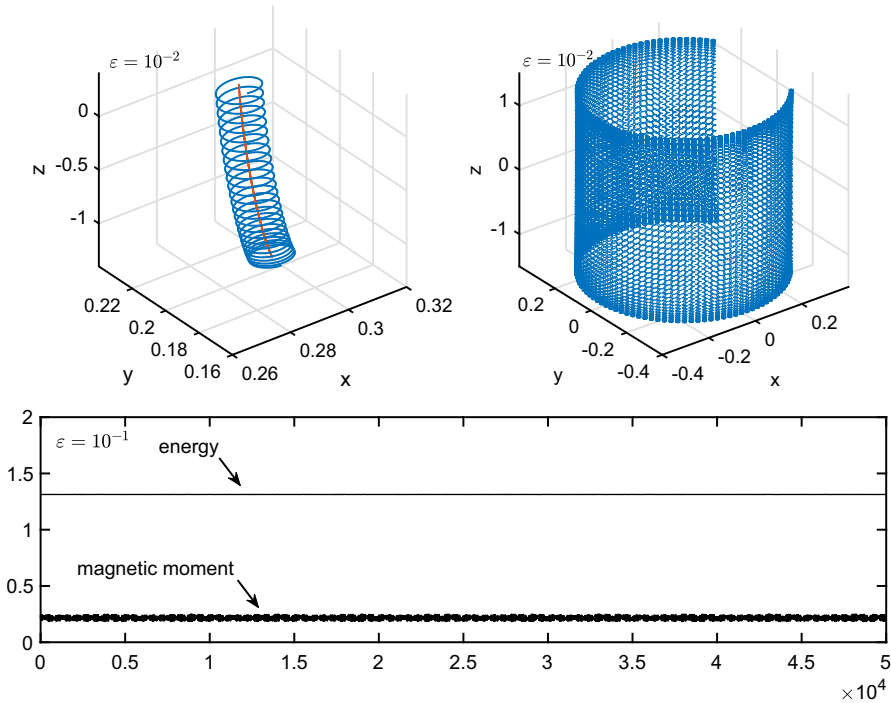


Fig. 1 Trajectories of the particle for $t \leq \pi/2$ (top left) and $t \leq 5/\varepsilon$ (top right), both with $\varepsilon = 10^{-2}$. Energy and magnetic moment for $t \leq \varepsilon^{-4}$ (bottom) for $\varepsilon = 10^{-1}$. The analogous picture for $\varepsilon = 10^{-2}$ instead of $\varepsilon = 10^{-1}$ would show the magnetic moment as a horizontal straight line

Remark 2.1 The results of this paper pertain to the above situation (2.1) – (2.2) of a mildly non-uniform strong magnetic field. For a strongly non-uniform magnetic field $B(x) = \tilde{B}(x)/\varepsilon$ with $\tilde{B}(x)$ independent of ε with non-vanishing gradient, different numerical phenomena arise, which require a different analysis; cf. [15] for a long-time analysis of a variational integrator in a small-step-size regime and [26] for numerical results for an ingeniously modified Boris method used with large step sizes. The modification of the force field proposed in [26] is not needed for the mildly non-uniform situation (2.1) studied in this paper, but it is essential in the case of strongly non-uniform strong magnetic fields.

3 Three numerical integrators

We now describe the three numerical integrators for (2.1) that are studied in this paper when applied with large step sizes $h \gg \varepsilon$.

3.1 Boris algorithm

The Boris method, introduced in [3], is the standard integrator for particle-in-cell codes for plasma simulation; see e.g. [2, 8]. Given the position and velocity approximation

$(x^n, v^{n-1/2})$, the algorithm computes $(x^{n+1}, v^{n+1/2})$ as follows, with $B^n = B(x^n)$ and $E^n = E(x^n)$:

$$\begin{aligned} v_+^{n-\frac{1}{2}} &= v^{n-\frac{1}{2}} + \frac{h}{2} E^n \\ v_-^{n+\frac{1}{2}} - v_+^{n-\frac{1}{2}} &= \frac{h}{2} (v_-^{n+\frac{1}{2}} + v_+^{n-\frac{1}{2}}) \times B^n \\ v_-^{n+\frac{1}{2}} &= v_-^{n+\frac{1}{2}} + \frac{h}{2} E^n \\ x^{n+1} &= x^n + h v^{n+\frac{1}{2}}, \end{aligned} \quad (3.1)$$

where the starting value is chosen as $v^{1/2} = v^0 + \frac{h}{2} v^0 \times B^0 + \frac{h}{2} E^0$.

The method has the equivalent two-step formulation

$$\frac{x^{n+1} - 2x^n + x^{n-1}}{h^2} = \frac{x^{n+1} - x^{n-1}}{2h} \times B^n + E^n, \quad (3.2)$$

with the velocity approximation

$$v^n = \frac{x^{n+1} - x^{n-1}}{2h}. \quad (3.3)$$

It is known from [9] that the Boris algorithm is not symplectic unless B is a constant magnetic field. The energy behaviour over long times, which is not fully satisfactory, has been studied in [14] for step sizes with $h|B| \ll 1$, which in our case (2.1) would read $h \ll \varepsilon$ in contrast to (1.1).

In the large-stepsize regime (1.1) the starting velocity needs to be modified. Instead of setting v^0 equal to the initial data $\dot{x}(0)$ we choose v^0 such that its component v_\perp^0 orthogonal to the magnetic field is $O(\varepsilon)$ -small. We propose to take $v^0 = v_\parallel^0 + v_\perp^0$ with

$$v_\parallel^0 = P_0 \dot{x}(0), \quad v_\perp^0 = \varepsilon (v_\parallel^0 \times B_1(x^0) + E(x^0)) \times B_0, \quad (3.4)$$

where $P_0 = B_0 B_0^\top$ is the orthogonal projection in the direction of B_0 . (This choice of v_\perp^0 will be explained in Section 4 right after Theorem 4.2.) Without such a modification of the starting velocity, the Boris algorithm shows highly oscillatory behaviour with a large amplitude proportional to $(h^2/\varepsilon)|v_\perp^0|$; cf. [23].

Since the Boris method with large step size (1.1) and the proposed filtering of the initial velocity will give an approximation to the guiding centre rather than to the oscillatory trajectory, it is reasonable to take the guiding centre approximation $x(0) + \varepsilon \dot{x}(0) \times B_0$ instead of $x(0)$ as the starting position x^0 .

We note that while the one-step map $(x^n, v^{n-1/2}) \mapsto (x^{n+1}, v^{n+1/2})$ is volume-preserving [22], the starting-value map $(x(0), \dot{x}(0)) \mapsto (x^0, v^0)$ and also the map $(x^0, v^0) \mapsto (x^1, v^{1/2})$ are far from volume-preserving for step sizes (1.1).

3.2 Standard variational integrator

The variational integrator to be studied here is constructed in the same way as is done in the interpretation of the Störmer–Verlet method as a variational integrator; see e.g. [17, Chap. VI, Example 6.2] and [25]. The integral of the Lagrangian $L(x, v) = \frac{1}{2}|v|^2 + A(x)^\top v - \phi(x)$ over a time step is approximated in two steps: the path $x(t)$ of positions is approximated by the linear interpolant of the endpoint positions, and the integral is approximated by the trapezoidal rule. This approximation to the action integral is then extremized. With the derivative matrix $A'(x) = (\partial_j A_i(x))_{i,j=1}^3$ and its transpose $A'(x)^\top$, this variational integrator becomes the following:

$$\begin{aligned} \frac{x^{n+1} - 2x^n + x^{n-1}}{h^2} = \\ A'(x^n)^\top \frac{x^{n+1} - x^{n-1}}{2h} - \frac{A(x^{n+1}) - A(x^{n-1})}{2h} + E^n, \end{aligned} \quad (3.5)$$

or equivalently, written as a perturbation to the Boris algorithm and using that $v \times B(x) = A'(x)^\top v - A'(x)v$,

$$\begin{aligned} \frac{x^{n+1} - 2x^n + x^{n-1}}{h^2} = \frac{x^{n+1} - x^{n-1}}{2h} \times B^n + E^n \\ + A'(x^n) \frac{x^{n+1} - x^{n-1}}{2h} - \frac{A(x^{n+1}) - A(x^{n-1})}{2h}. \end{aligned} \quad (3.6)$$

We note that the correction to the Boris method as given in the second line vanishes for linear $A(x)$. In the situation of the magnetic field of (2.1), we can therefore replace A by A_1 in (3.6). The variational integrator coincides with the Boris algorithm in the case of a constant magnetic field ($B_1 \equiv 0$).

This method is again complemented with the velocity approximation (3.3). It can be given a one-step formulation similar to the Boris algorithm, with the correction term of (3.6) added in the second line of (3.1). It is, however, an implicit method, because the vector potential A is evaluated at the new position x^{n+1} .

For the case of a strong magnetic field and for step sizes with $h|B| \leq \text{Const.}$, the variational integrator has been shown to have excellent near-preservation of energy and magnetic moment over very long times [15].

For large step sizes (1.1), the variational integrator requires the same modification of the starting velocity as the Boris method in order to suppress high oscillations of large amplitude in the numerical solution.

3.3 Filtered variational integrator

As a new method to be studied here, we propose the following modification of the variational integrator: with the filter functions

$$\psi(\zeta) = \text{tanch}(\zeta/2) = \frac{\tanh(\zeta/2)}{\zeta/2}, \quad \varphi(\zeta) = \frac{1}{\text{sinh}(\zeta)} = \frac{\zeta}{\sinh(\zeta)},$$

which are even functions and take the value 1 at $\zeta = 0$, and with the skew-symmetric matrix \widehat{B}_0 defined by $-\widehat{B}_0 v = v \times B_0$ for all $v \in \mathbf{R}^3$, we define the filter matrices

$$\begin{aligned} \Psi &= \psi\left(-\frac{h}{\varepsilon} \widehat{B}_0\right) = I + \left(1 - \text{tanc}\left(\frac{h}{2\varepsilon}\right)\right) \widehat{B}_0^2, \\ \Phi &= \varphi\left(-\frac{h}{\varepsilon} \widehat{B}_0\right) = I + \left(1 - \text{sinc}\left(\frac{h}{\varepsilon}\right)^{-1}\right) \widehat{B}_0^2, \end{aligned}$$

where the rightmost expressions are obtained from a Rodriguez formula; see [16, Appendix]. Here, $\text{tanc}(\xi) = \tan(\xi)/\xi$ and $\text{sinc}(\xi) = \sin(\xi)/\xi$. The filter matrices Ψ and Φ are symmetric and act as the identity on vectors in the direction of B_0 .

We put the filter matrix Ψ in front of the right-hand side of (3.5):

$$\begin{aligned} &\frac{x^{n+1} - 2x^n + x^{n-1}}{h^2} = \\ &\Psi\left(A'(x^n)^\top \frac{(x^{n+1} - x^{n-1})}{2h} - \frac{A(x^{n+1}) - A(x^{n-1})}{2h} + E^n\right). \end{aligned} \quad (3.7)$$

This is combined with the velocity approximation

$$v^n = \Phi \frac{x^{n+1} - x^{n-1}}{2h} + \varepsilon \left(1 - \text{sinc}\left(\frac{h}{\varepsilon}\right)^{-1}\right) E^n \times B_0. \quad (3.8)$$

This filtered variational integrator coincides with the filtered Boris algorithm of [16] for the special case of a constant magnetic field $B(x) = B_0/\varepsilon$. If additionally also E is constant, then this method yields the exact position and velocity, as was shown for the filtered Boris algorithm.

For stepsizes h with $\tan(h/(2\varepsilon)) \geq c > 0$, the filter matrix Ψ is positive definite. The above integrator can then be interpreted as a variational integrator corresponding to a discrete Lagrangian where the kinetic energy term has the modified mass matrix Ψ^{-1} . Its eigenvalues corresponding to the eigenvectors orthogonal to B_0 are $1/\text{tanc}(h/(2\varepsilon))$ and are thus proportional to h/ε , which is greater than h^{-1} under condition (1.1). The discrete Lagrangian reads

$$\begin{aligned} L_h(x^n, x^{n+1}) &= \frac{h}{2} (v^{n+1/2})^\top \Psi^{-1} v^{n+1/2} \\ &\quad + h \frac{A(x^n)^\top + A(x^{n+1})^\top}{2} v^{n+1/2} - h \frac{\phi(x^n) + \phi(x^{n+1})}{2}, \end{aligned}$$

where $v^{n+1/2} = (x^{n+1} - x^n)/h$. The standard (unfiltered) variational integrator has the same discrete Lagrangian except for the identity matrix in place of the matrix Ψ^{-1} .

The filtered variational integrator for (2.1) can be written and implemented as the following implicit one-step method:

$$\begin{aligned} v_+^{n-\frac{1}{2}} &= v_+^{n-\frac{1}{2}} + \frac{h}{2} \Psi E^n \\ v_-^{n+\frac{1}{2}} - v_+^{n-\frac{1}{2}} &= h \Psi \left(\frac{1}{2} (v_-^{n+\frac{1}{2}} + v_+^{n-\frac{1}{2}}) \times B^n \right. \\ &\quad \left. + A'_1(x^n) \frac{1}{2} (v_-^{n+\frac{1}{2}} + v_+^{n-\frac{1}{2}}) - \frac{A_1(x^{n+1}) - A_1(x^{n-1})}{2h} \right) \\ v_-^{n+\frac{1}{2}} &= v_-^{n+\frac{1}{2}} + \frac{h}{2} \Psi E^n \\ x^{n+1} &= x^n + h v_-^{n+\frac{1}{2}}. \end{aligned}$$

This can be solved by a fixed-point iteration for x^{n+1} , where a good starting iterate is obtained from a Boris step. The first velocity is chosen as follows: we set $v^{1/2} = \bar{v} + \frac{1}{2} \delta v$ with $h\bar{v} = \frac{1}{2}(x^1 - x^{-1})$ and $h\delta v = x^1 - 2x^0 + x^{-1}$, where in view of (3.8) for $n = 0$,

$$\bar{v} = \Phi^{-1} \left(v^0 - \varepsilon \left(1 - \text{sinc} \left(\frac{h}{\varepsilon} \right)^{-1} \right) E^n \times B_0 \right)$$

and δv is implicitly determined (and computed via fixed-point iteration) from (3.7) with $n = 0$, i.e. from the equation

$$\delta v = h \Psi \left(\bar{v} \times B(x^0) + A'_1(x^0) \bar{v} - \frac{A_1(x^1) - A_1(x^{-1})}{2h} + E(x^0) \right),$$

where $x^{\pm 1} = x^0 \pm h\bar{v} + \frac{1}{2}h\delta v$.

In contrast to the Boris algorithm and the unfiltered variational integrator, we here take the original initial data $v^0 = \dot{x}(0)$ and $x^0 = x(0)$.

4 Modulated Fourier expansions

We give modulated Fourier expansions of the exact solution of (2.1) and the numerical solutions of the three integrators for large step sizes $h^2 \geq c\varepsilon$ (in the following we set the irrelevant positive constant c equal to 1 for simplicity). Analogous expansions for step sizes $h \leq C\varepsilon$ were previously given in [13, 15, 16]; see also [17, Ch. XIII]. In particular, we explicitly state the differential equations for the dominant modulation functions up to $O(\varepsilon^2)$ for the exact solution, and up to $O(h^2)$ for the numerical solutions.

4.1 Modulated Fourier expansion of the exact motion

We write the solution of (2.1) as

$$x(t) \approx \sum_k z^k(t) e^{ikt/\varepsilon} \quad (4.1)$$

with coefficient functions $z^k(t)$ for which all time derivatives are bounded independently of ε .

We diagonalize the linear map $v \mapsto v \times B_0$, which has eigenvalues $\lambda_1 = i$, $\lambda_0 = 0$ and $\lambda_{-1} = -i$ (recall the normalization $|B_0| = 1$). The normalized eigenvectors are denoted $v_1, v_0 = B_0, v_{-1} = \bar{v}_1$. We let $P_j = v_j v_j^*$ be the orthogonal projections onto the eigenspaces. We write the coefficient functions of (4.1) in the basis (v_j) ,

$$z^k = z_1^k + z_0^k + z_{-1}^k, \quad z_j^k(t) = P_j z^k(t).$$

The following theorem is a variant of Theorems 4.1 in [15, 16], proved by the same arguments but in a technically simplified way, since here we have the constant frequency $1/\varepsilon$ and constant projections P_j , as opposed to the state-dependent frequency and projections in [15, 16].

Theorem 4.1 *Let $x(t)$ be a solution of (2.1) with an initial velocity bounded independently of ε ($|\dot{x}(0)| \leq C_1$), which stays in a compact set K for $0 \leq t \leq T$ (with K and T independent of ε). For an arbitrary truncation index $N \geq 1$ we then have an expansion*

$$x(t) = \sum_{|k| \leq N} z^k(t) e^{ikt/\varepsilon} + R_N(t)$$

with the following properties:

- (a) *The modulation functions z^k together with their derivatives (up to order N) are bounded as $z_j^0 = O(1)$ for $j \in \{-1, 0, 1\}$, $z_1^1 = O(\varepsilon)$, $z_{-1}^{-1} = O(\varepsilon)$, and for the remaining (k, j) with $|k| \leq N$,*

$$z_j^k = O(\varepsilon^{|k|+1}).$$

They are unique up to $O(\varepsilon^N)$ and are chosen to satisfy $z_{-j}^{-k} = \overline{z_j^k}$. Moreover, $\dot{z}_{\pm 1}^0$ together with its derivatives is bounded as $\dot{z}_{\pm 1}^0 = O(\varepsilon)$.

- (b) *The remainder term and its derivative are bounded by*

$$R_N(t) = O(t^2 \varepsilon^N), \quad \dot{R}_N(t) = O(t \varepsilon^N) \quad \text{for } 0 \leq t \leq T.$$

(c) The functions $z_0^0, z_{\pm 1}^0, z_1^1, z_{-1}^{-1}$ satisfy the differential equations

$$\begin{aligned}\ddot{z}_0^0 &= P_0(\dot{z}_0^0 \times B_1(z_0^0) + E(z_0^0)) + 2P_0 \operatorname{Re}\left(\frac{i}{\varepsilon} z_1^1 \times B_1'(z_0^0) z_{-1}^{-1}\right) + O(\varepsilon^2), \\ \ddot{z}_{\pm 1}^0 &= \pm i\varepsilon P_{\pm 1}(\dot{z}_{\pm 1}^0 \times B_1(z_0^0) + E(z_0^0)) + O(\varepsilon^2), \\ \ddot{z}_{\pm 1}^{\pm 1} &= P_{\pm 1}(\dot{z}_{\pm 1}^{\pm 1} \times B_1(z_0^0)) + O(\varepsilon^2).\end{aligned}$$

All other modulation functions z_j^k are given by algebraic expressions depending on $z_0^0, \dot{z}_0^0, z_1^1, z_{-1}^{-1}$.

(d) Initial values for the differential equations of item (c) are given by

$$\begin{aligned}z_0^0(0) &= x(0) + \varepsilon \dot{x}(0) \times B_0 + O(\varepsilon^2), \\ \dot{z}_0^0(0) &= P_0 \dot{x}(0) - \varepsilon P_0((\dot{x}(0) \times B_0) \times B_1(x(0))) + O(\varepsilon^2), \\ z_{\pm 1}^{\pm 1}(0) &= \mp i\varepsilon P_{\pm 1} \dot{x}(0) + O(\varepsilon^2).\end{aligned}$$

The constants symbolized by the O -notation are independent of ε and t with $0 \leq t \leq T$, but depend on N , on the velocity bound M , on bounds of derivatives of B_1 and E on the compact set K , and on T .

4.2 Resonant modulated Fourier expansion of the Boris algorithm and the standard variational integrator for $h^2 \geq \varepsilon$

When the Boris method is applied to the linear differential equation $\ddot{x} = \dot{x} \times B_0/\varepsilon$ with $|B_0| = 1$ (that is, B_1 and E are not present in (2.1)), then diagonalization of B_0 shows that x^n is a linear combination (with coefficients independent of n) of terms 1 , nh and $e^{\pm i nh \omega}$, where

$$h\omega = 2 \arctan\left(\frac{h}{2\varepsilon}\right).$$

If h/ε is large, then $h\omega$ is close to π . In particular, if $h^2 \geq \varepsilon$, then $h\omega = \pi - \gamma h$ with $\gamma > 0$ bounded independently of h and ε with $h^2 \geq \varepsilon$, and so $e^{\pm i nh \omega} = (-1)^n e^{\mp i nh \gamma}$, where we note that $e^{\mp i t \gamma}$ is a smooth function of t all of whose derivatives are bounded independently of ε and h . In the general case of (2.1), we have the following result.

Theorem 4.2 *Let x^n be the numerical solution obtained by applying either the Boris algorithm or the variational integrator to (2.1) with a stepsize h satisfying*

$$h^2 \geq \varepsilon. \quad (4.2)$$

We assume that the starting velocity v^0 is bounded independently of ε and h and that its component orthogonal to B_0 , i.e. $v_{\perp}^0 = (I - P_0)v^0$, is small:

$$|v_{\perp}^0| \leq c_1 \varepsilon. \quad (4.3)$$

We further assume that the numerical solution x^n stays in a compact set K for $0 \leq nh \leq T$ (with K and T independent of ε and h). For an arbitrary truncation index $N \geq 2$, we then have a decomposition

$$x^n = y(t) + (-1)^n z(t) + R_N(t), \quad t = nh, \quad (4.4)$$

with the following properties:

- (a) The functions $y(t)$ and $z(t)$ together with their derivatives (up to order N) are bounded as $y = O(1)$, $z = O(h^2)$. They are unique up to $O(h^N)$. Moreover, we have $\dot{y} \times B_0 = O(\varepsilon)$ and $z \cdot B_0 = O(h^4)$.
- (b) The remainder term is bounded by

$$R_N(t) = O(t^2 h^N) \quad \text{for } 0 \leq t \leq T.$$

- (c) The functions $y_j = P_j y$ ($j = 0, \pm 1$) and $z_{\pm 1} = P_{\pm 1} z$ satisfy the differential equations

$$\begin{aligned} \ddot{y}_0 &= P_0(\dot{y} \times B_1(y) + E(y)) + O(h^2), \\ \dot{y}_{\pm 1} &= \pm i\varepsilon P_{\pm 1}(\dot{y} \times B_1(y) + E(y)) + O(\varepsilon h^2), \\ \dot{z}_{\pm 1} &= \mp 4i \frac{\varepsilon}{h^2} z_{\pm 1} + O(\varepsilon h^2). \end{aligned}$$

The function $z_0 = P_0 z$ is given by an algebraic expression depending on y , \dot{y}_0 and $z_{\pm 1}$.

- (d) Initial values for the differential equations of item (c) are given by

$$\begin{aligned} y(0) &= x^0 + O(h^2), \\ \dot{y}_0(0) &= P_0(x^0)v^0 + O(h^2), \\ z_{\pm 1}(0) &= \mp \frac{ih^2}{4\varepsilon} P_{\pm 1}\left(v^0 \mp i\varepsilon(P_0 v^0 \times B_1(x^0) + E(x^0))\right) + O(h^4). \end{aligned}$$

The constants symbolized by the O -notation are independent of ε , h and n with $0 \leq nh \leq T$, but depend on the velocity bound, on bounds of derivatives of B_1 and E on the compact set K , and on T .

We note that the differential equations for y agree with those for z^0 of the exact solution up to $O(h^2)$. The differential equations for $z_{\pm 1}$ and for $z_{\pm 1}^{\pm 1}$ of the exact solution differ, but we still have

$$\frac{d}{dt} |z_{\pm 1}|^2 = 2 \operatorname{Re} z_{\pm 1}^* \dot{z}_{\pm 1} = O(\varepsilon |z_{\pm 1}|^2) + O(\varepsilon h^N) = O(\varepsilon h^4),$$

which is to be compared with

$$\frac{d}{dt} |z_{\pm 1}^{\pm 1}|^2 = 2 \operatorname{Re} (z_{\pm 1}^{\pm 1})^* \dot{z}_{\pm 1}^{\pm 1} = O(\varepsilon^3).$$

To obtain an $O(h^2)$ approximation to the guiding centre $z^0(t)$ over bounded time intervals, we run the Boris algorithm with the modified initial velocity $v^0 = P_0 \dot{x}(0)$ instead of $\dot{x}(0)$, or even better, determine $P_{\pm 1} v^0$ such that $z_{\pm 1}(0) = O(h^4)$, which holds true with the proposed choice (3.4).

Proof The bounds of parts (a) and (b) are proved as in previous proofs of modulated Fourier expansions; see e.g. [15] and [17, Ch. XIII]. Here we just show (c) and (d), assuming that the bounds of (a) and (b) are already available.

To derive the differential equations of (c), we insert (4.4) into the two-step formulation of the numerical method, expand $y(t \pm h)$ and $z(t \pm h)$ into Taylor series at t , expand the nonlinear functions B_1 and E at $y(t)$ and separate the terms without and with the factor $(-1)^n$. This gives us the equations

$$\begin{aligned} \ddot{y} + O(h^2) &= \dot{y} \times \frac{B_0}{\varepsilon} + \dot{y} \times B_1(y) + E(y) + O(h^2) \\ -\frac{4}{h^2}z - \ddot{z} + O(h^2) &= -\dot{z} \times \frac{B_0}{\varepsilon} + \dot{z} \times B_1(y) + \dot{y} \times B'_1(y)z + E'(y)z + O(h^2). \end{aligned}$$

In the equation for z we note that also \ddot{z} and the last three terms on the right-hand side are $O(h^2)$ as z and its derivatives are $O(h^2)$, and the indicated $O(h^2)$ terms are then actually $O(h^4)$.

Taking the projection P_0 on both sides of the differential equation for y yields the stated second-order differential equation for y_0 on noting that $P_0(\dot{y} \times B_0) = 0$. Moreover, since $P_{\pm 1}(\dot{y} \times B_0) = \pm i\dot{y}_{\pm 1}$, we obtain

$$\mp \frac{i}{\varepsilon} \dot{y}_{\pm 1} = -\ddot{y}_{\pm 1} + P_{\pm 1}(\dot{y} \times B_1(y) + E(y)) + O(h^2).$$

Differentiating this equation and multiplying with $i\varepsilon$ yields $\ddot{y}_{\pm 1} = O(\varepsilon)$, which is $O(h^2)$ under condition (4.2). So we obtain the stated first-order differential equation for $y_{\pm 1}$.

Taking the projection P_0 in the above equation for z yields $-\frac{4}{h^2}z_0 = O(h^2)$, and hence $z_0 = O(h^4)$. Taking the projections $P_{\pm 1}$ yields

$$-\frac{4}{h^2}z_{\pm 1} = \mp \frac{i}{\varepsilon} \dot{z}_{\pm 1} + O(h^2),$$

which can be rearranged into the stated differential equation for $z_{\pm 1}$.

In view of (4.4) for $n = 0$ and $z(0) = O(h^2)$, we have $y(0) = x^0 + O(h^2)$. Since we obtain by inserting (4.4) for $n = -1, 1$

$$v^0 = \frac{x^1 - x^{-1}}{2h} = \dot{y}(0) - \dot{z}(0) + O(h^2),$$

we obtain the stated expression for $\dot{y}_0(0)$ on taking the projection P_0 . Taking the projections $P_{\pm 1}$ and using the differential equations for $y_{\pm 1}$ and $z_{\pm 1}$, we arrive at the stated expression for $z_{\pm 1}(0)$. \square

4.3 Non-resonant modulated Fourier expansion of the filtered variational integrator for $h^2 \geq \varepsilon$

As the filtered integrator is exact for the linear equation $\ddot{x} = \dot{x} \times B_0/\varepsilon$, it has the same high frequency $1/\varepsilon$. When this integrator is applied to (2.1), it has a modulated Fourier equation that is very similar to that of the exact solution given in Theorem 4.1.

Theorem 4.3 *Let x^n be a solution of the filtered variational integrator applied to (2.1) with a stepsize h satisfying*

$$h^2 \geq \varepsilon \quad (4.5)$$

and, for some $N \geq 1$, the non-resonance conditions

$$\begin{aligned} \left| \sin\left(\frac{kh}{2\varepsilon}\right) \right| &\geq c > 0, \quad \left| \cos\left(\frac{kh}{2\varepsilon}\right) \right| \geq c > 0 \quad (k = 1, \dots, N), \\ \left| \tan\left(\frac{kh}{2\varepsilon}\right) - \tan\left(\frac{h}{2\varepsilon}\right) \right| &\geq c > 0 \quad (k = 2, \dots, N), \end{aligned} \quad (4.6)$$

where c is a positive constant. We assume that the initial velocity $v^0 = \dot{x}(0)$ is bounded independently of ε and h , as in (2.2). We further assume that the numerical solution x^n stays in a compact set K for $0 \leq nh \leq T$ (with K and T independent of ε and h). We then have an expansion, at $t = nh$,

$$x^n = \sum_{|k| \leq N} z^k(t) e^{ikt/\varepsilon} + R_N(t) \quad (4.7)$$

with the following properties:

- (a) *The bounds of parts (a) of Theorem 4.1 for the modulation functions are valid also in this case, except $z_0^k = O(h\varepsilon^{|k|})$ for $|k| \geq 1$.*
- (b) *The remainder at $t = nh$ is bounded, for arbitrary $M > 1$, by*

$$P_0 R_N(t) = O(t^2 h^M) + O(t^2 \varepsilon^N), \quad P_{\pm 1} R_N(t) = O(t^2 \varepsilon h^{M-1}) + O(t^2 \varepsilon^N).$$

- (c) *The functions $z_0^0, z_{\pm 1}^0, z_1^1, z_{-1}^{-1}$ satisfy the differential equations*

$$\begin{aligned} \ddot{z}_0^0 &= P_0(\dot{z}_0^0 \times B_1(z_0^0) + E(z_0^0)) + O(h^2), \\ \ddot{z}_{\pm 1}^0 &= \pm i\varepsilon P_{\pm 1}(\dot{z}_0^0 \times B_1(z_0^0) + E(z_0^0)) + O(\varepsilon h), \\ \ddot{z}_{\pm 1}^{\pm 1} &= \frac{\varepsilon}{h} \sin\left(\frac{h}{\varepsilon}\right) P_{\pm 1}(z_{\pm 1}^{\pm 1} \times B_1(z_0^0)) + O(\varepsilon^2). \end{aligned}$$

All other modulation functions z_j^k are given by algebraic expressions depending on $z_0^0, \dot{z}_0^0, z_1^1, z_{-1}^{-1}$.

(d) Initial values for the differential equations of item (c) are given by

$$\begin{aligned} z^0(0) &= x^0 + O(h^2), \\ \dot{z}_0^0(0) &= P_0 v^0 + O(h^2), \\ z_{\pm 1}^{\pm 1}(0) &= \mp i\varepsilon P_{\pm 1} v^0 + O(\varepsilon h). \end{aligned}$$

The constants symbolized by the O -notation are independent of ε and t with $0 \leq t \leq T$, but depend on M and N , on the velocity bound (2.2), on bounds of derivatives of B_1 and E on the compact set K , and on T .

Proof Parts (a) and (b) are again proved as in previous proofs of modulated Fourier expansions; see e.g. [15] and [17, Ch. XIII]. Here we only show (c) and (d), assuming that the bounds of (a) and (b) are already available.

To derive the differential equations of (c), we insert (4.7) into the two-step formulation (3.7) (or equivalently (3.6) with an extra factor Ψ on the right-hand side), expand $z^k(t \pm h)$ into a Taylor series at t , use Lemma 5.1 of [16] to expand the first and second-order difference quotients for $z^k(t)e^{ikt/\varepsilon}$ for $0 < |k| \leq N$, and expand B_1 and E at $z^0(t)$. We then separate the terms multiplying $e^{ikt/\varepsilon}$ for $|k| \leq N$. Moreover, we consider the components $z_j^k = P_j z^k$ for $j = 0, \pm 1$.

For $k = 0$, $j = 0$ we obtain

$$\ddot{z}_0^0 + O(h^2) = P_0 \left((\dot{z}_0^0 + O(h^2)) \times B_1(z^0) + E(z^0) + O(\varepsilon^2/h) \right),$$

where the $O(h^2)$ terms result from the Taylor expansions of the second and first order difference quotients of z^0 , and the (smaller) $O(\varepsilon^2/h)$ term results from the Taylor expansion of B_1 and E at z^0 and the bound $z^k = O(\varepsilon^{|k|})$. This yields the first equation of (c).

For $k = 0$, $j = 1$ we obtain

$$\begin{aligned} \ddot{z}_1^0 + O(\varepsilon h^2) &= \frac{2\varepsilon}{h} \tan\left(\frac{h}{2\varepsilon}\right) \left(\frac{i}{\varepsilon} (\dot{z}_1^0 + O(h^2)) \right. \\ &\quad \left. + P_1 \left((\dot{z}_0^0 + O(h^2)) \times B_1(z^0) + E(z^0) + O(\varepsilon^2/h) \right) \right). \end{aligned}$$

We solve this equation for \dot{z}_1^0 , which appears in the dominant term with a factor h^{-1} , and recall that $|\tan(h/(2\varepsilon))| \geq c > 0$ by the non-resonance condition (4.6). Using that \ddot{z}_1^0 and its higher derivatives are $O(\varepsilon)$ by part (a), this yields

$$\dot{z}_1^0 = i\varepsilon P_1 \left(\dot{z}_0^0 \times B_1(z^0) + E(z^0) \right) + O(\varepsilon h),$$

which is the differential equation for z_1^0 stated in (c). The case $j = -1$ is obtained by taking complex conjugates.

For $k = 1$, $j = 1$ we find for $y_1^1(t) = z_1^1(t)e^{it/\varepsilon}$, using Lemma 5.1 of [16] and the $O(\varepsilon)$ bound for z_1^1 and its derivatives of part (a),

$$\begin{aligned} & \frac{y_1^1(t+h) - 2y_1^1(t) + y_1^1(t-h)}{h^2} \\ &= e^{it/\varepsilon} \left(-\frac{4}{h^2} \sin^2\left(\frac{h}{2\varepsilon}\right) z_1^1(t) + \frac{2i}{h} \sin\left(\frac{h}{\varepsilon}\right) \dot{z}_1^1(t) + O(\varepsilon) \right) \end{aligned}$$

and

$$\frac{y_1^1(t+h) - y_1^1(t-h)}{2h} = e^{it/\varepsilon} \left(\frac{i}{h} \sin\left(\frac{h}{\varepsilon}\right) z_1^1(t) + \cos\left(\frac{h}{\varepsilon}\right) \dot{z}_1^1(t) + O(\varepsilon h) \right).$$

The latter formula yields

$$\begin{aligned} & P_1 \tanh\left(\frac{h}{2\varepsilon} \widehat{B}_0\right) \left(\frac{y_1^1(t+h) - y_1^1(t-h)}{2h} \times \frac{B_0}{\varepsilon} \right) \\ &= \frac{2\varepsilon}{h} \tan\left(\frac{h}{2\varepsilon}\right) \frac{i}{\varepsilon} e^{it/\varepsilon} \left(\frac{i}{h} \sin\left(\frac{h}{\varepsilon}\right) z_1^1(t) + \cos\left(\frac{h}{\varepsilon}\right) \dot{z}_1^1(t) + O(\varepsilon h) \right) \\ &= e^{it/\varepsilon} \left(-\frac{4}{h^2} \sin^2\left(\frac{h}{2\varepsilon}\right) z_1^1(t) + \frac{2i}{h} \tan\left(\frac{h}{2\varepsilon}\right) \cos\left(\frac{h}{\varepsilon}\right) \dot{z}_1^1(t) + O(\varepsilon) \right) \end{aligned}$$

and similarly

$$\begin{aligned} & P_1 \tanh\left(\frac{h}{2\varepsilon} \widehat{B}_0\right) \left(\frac{y_1^1(t+h) - y_1^1(t-h)}{2h} \times B_1(z^0(t)) \right) \\ &= e^{it/\varepsilon} \frac{2\varepsilon}{h} \tan\left(\frac{h}{2\varepsilon}\right) P_1 \left(\frac{i}{h} \sin\left(\frac{h}{\varepsilon}\right) z_1^1(t) \times B_1(z^0(t)) + O(\varepsilon) \right). \end{aligned}$$

We insert the modulated Fourier expansion (4.7) into the two-step formulation of the filtered variational integrator, which we write as (3.6) with the extra filter factor Ψ on the right-hand side, and we collect the terms with the factor $e^{it/\varepsilon}$. The dominant terms after projecting with P_1 are given by the above formulas. The remaining term on the right-hand side (as in the second line of (3.6) but multiplied with Ψ) is of size $(2\varepsilon/h) \tan(h/2\varepsilon) \cdot O(\varepsilon + h^2) = O(h\varepsilon)$ for $h^2 \geq \varepsilon$ under condition (4.6). We thus obtain

$$\begin{aligned} & -\frac{4}{h^2} \sin^2\left(\frac{h}{2\varepsilon}\right) z_1^1(t) + \frac{2i}{h} \sin\left(\frac{h}{\varepsilon}\right) \dot{z}_1^1(t) + O(\varepsilon) \\ &= -\frac{4}{h^2} \sin^2\left(\frac{h}{2\varepsilon}\right) z_1^1(t) + \frac{2i}{h} \tan\left(\frac{h}{2\varepsilon}\right) \cos\left(\frac{h}{\varepsilon}\right) \dot{z}_1^1(t) + O(\varepsilon) \\ &+ \frac{2\varepsilon}{h} \tan\left(\frac{h}{2\varepsilon}\right) P_1 \left(\frac{i}{h} \sin\left(\frac{h}{\varepsilon}\right) z_1^1(t) \times B_1(z^0(t)) + O(\varepsilon) \right). \end{aligned}$$

Here the dominant terms are the first terms on the left-hand and the right-hand sides, which are the same and thus cancel. The dominant terms then become the terms

containing the factor $(2i/h)z_1^1(t)$. Since a calculation shows that we have, with $\xi = h/(2\varepsilon)$ for short,

$$\sin(2\xi) - \tan(\xi) \cos(2\xi) = (\tan(2\xi) - \tan(\xi)) \cos(2\xi) = \tan(\xi),$$

the above equation yields the differential equation for z_1^1 as stated in part (c) of the theorem. The result for z_{-1}^{-1} is obtained by taking complex conjugates.

The formulae for the initial values are obtained by the same arguments as in the proof of Theorem 4.2, using here that $(x^1 - x^{-1})/(2h)$ is related to v^0 by (3.8) for $n = 0$. \square

5 Time scale ε^0 : error bounds for position and parallel velocity

Comparing the modulated Fourier expansions of the numerical solution with that of the exact solution, we obtain the following error bounds from Theorems 4.1–4.3.

Theorem 5.1 *Consider applying the Boris method, the variational integrator and the filtered variational integrator to (2.1) over a time interval $0 \leq t \leq T$ (with T independent of ε) using a stepsize h with*

$$h^2 \geq \varepsilon.$$

Suppose that the conditions of Theorem 4.2 are satisfied in the case of the Boris method and the variational integrator (in particular, small perpendicular starting velocity: $v_\perp^0 = O(\varepsilon)$), and that the conditions of Theorem 4.3 are satisfied in the case of the filtered variational integrator (in particular, the non-resonance conditions (4.6) and bounded initial velocity (2.2)). For each of the three methods, the errors in position x and parallel velocity $v_\parallel = P_0 v$ at time $t_n = nh \leq T$ are then bounded by

$$|x^n - x(t_n)| \leq Ch^2, \quad |v_\parallel^n - v_\parallel(t_n)| \leq Ch^2 \quad (t_n \leq T),$$

where C is independent of ε , h and n with $h^2 \geq \varepsilon$ and $nh \leq T$ (but depends on T).

Proof The result is obtained by representing the exact and numerical solutions by their modulated Fourier expansions and using the bounds and differential equations of the modulation functions as given in Theorems 4.1–4.3. Note that the differential equations of the dominating modulation functions for the three methods and for the exact solution coincide up to defects of size $O(h^2)$, which lead to an $O(h^2)$ error in the positions. Inserting the modulated Fourier expansion of the numerical solution into the formula for the approximate velocity v^n for each method and comparing with the time-differentiated modulated Fourier expansion of the exact solution then yields the $O(h^2)$ error bound for the parallel velocity. \square

Remark 5.2 For $h^2 \sim \varepsilon$, the above error bounds are thus $O(\varepsilon)$. For all three methods, the error bounds remain in general $O(\varepsilon)$ also for smaller stepsizes $h \sim \varepsilon$. This can be

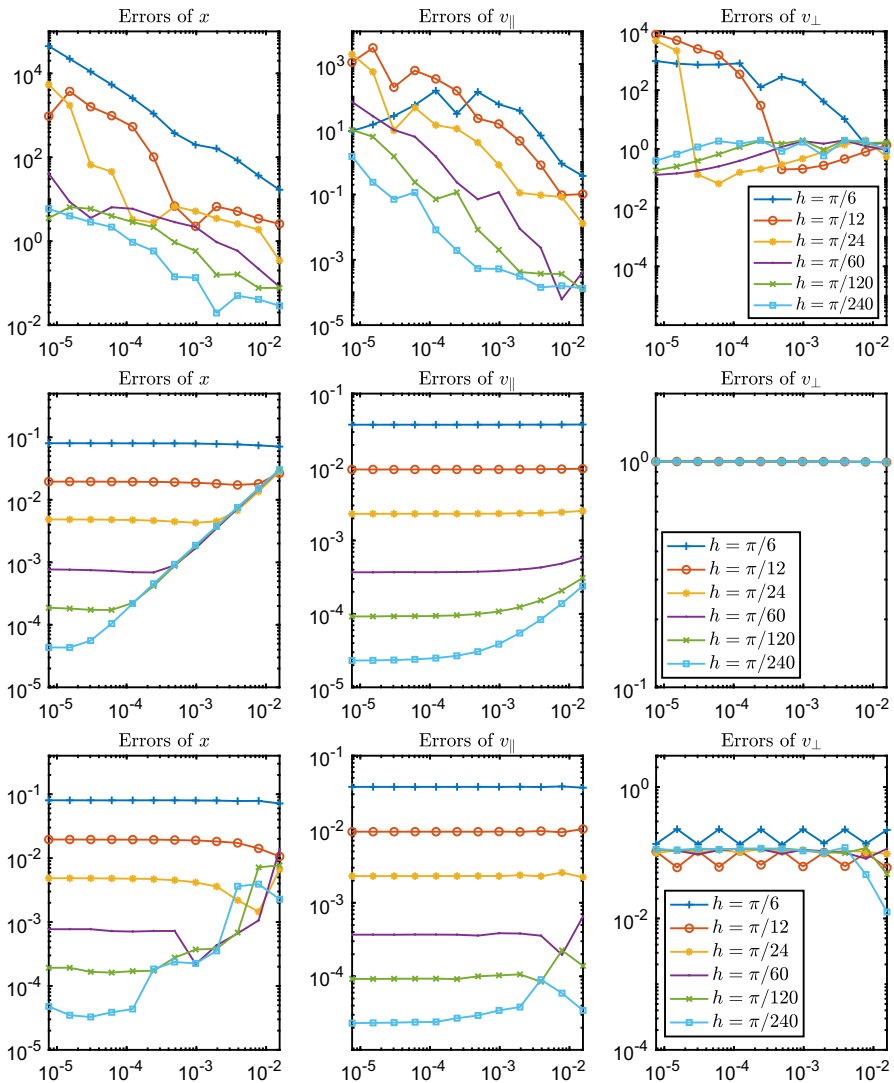


Fig. 2 Global error vs. ε ($\varepsilon = 1/2^j$, $j = 6, \dots, 17$) with different h for the Boris algorithm with starting values $x(0)$, $v(0)$ (top row), with modified starting values (3.4) (centre row), and for the filtered variational integrator with starting values $x(0)$, $v(0)$ (bottom row)

shown by comparing the modulated Fourier expansions for such stepsizes, as given in [15] for the standard variational integrator. The filtered Boris method of [16], used with $h \sim \varepsilon$, has an $O(\varepsilon^2)$ error in the position and the parallel velocity, and an $O(\varepsilon)$ error in the perpendicular velocity.

Numerical experiment. For the example of Section 2, Figure 2 shows the relative errors in x , v_{\parallel} and v_{\perp} at time $t = \pi/2$ versus ε for various step sizes h for three numerical approaches:

- (i) in the top row for the Boris algorithm with the original initial data as starting values,
- (ii) in the centre row for the Boris algorithm with modified starting values (3.4),
- (iii) in the bottom row for the filtered variational integrator with the original initial data as starting values.

For any step size h , the errors in x and v_{\parallel} increase roughly proportionally to h^2/ε when $\varepsilon \rightarrow 0$ in case (i), whereas in cases (ii) and (iii) the errors tend to a constant error level proportional to h^2 .

6 Time scale ε^{-1} : perpendicular drift

6.1 Perpendicular drift of the exact motion

We let $P_{\parallel} = P_0 = B_0 B_0^{\top}$ be the orthogonal projection onto the span of B_0 , and $P_{\perp} = P_1 + P_{-1} = I - P_{\parallel}$ the orthogonal projection onto the plane orthogonal to B_0 . We decompose $x \in \mathbf{R}^3$ as

$$x = x_{\parallel} + x_{\perp} \quad \text{with} \quad x_{\parallel} = P_{\parallel}x, \quad x_{\perp} = P_{\perp}x.$$

We assume that (with slight abuse of notation for B_1)

$$B_1(x) = B_1(x_{\perp}) + \varepsilon B_2(x), \quad E(x) = E_{\perp}(x_{\perp}) + E_{\parallel}(x) + \varepsilon E_2(x), \quad (6.1)$$

with $E_{\perp} \cdot B_0 = 0$ and $E_{\parallel} \times B_0 = 0$, and where the functions B_1 , B_2 and E_{\perp} , E_{\parallel} , E_2 on the right-hand side and all their derivatives are bounded independently of ε . We thus only allow a weak dependence of the magnetic field and the perpendicular electric field on x_{\parallel} . We then have the following result.

Theorem 6.1 *Let $x(t)$ be a solution of (2.1) with (6.1), with an initial velocity bounded independently of ε ($|\dot{x}(0)| \leq M$), which stays in a compact set K for $0 \leq t \leq c\varepsilon^{-1}$ (with K and c independent of ε). Then, the solution $y_{\perp}(t)$ of the initial-value problem for the slow differential equation*

$$\dot{y}_{\perp}(t) = \varepsilon E_{\perp}(y_{\perp}(t)) \times B_0, \quad y_{\perp}(0) = x_{\perp}(0), \quad (6.2)$$

remains $O(\varepsilon)$ -close to the perpendicular component of $x(t)$ over times $O(\varepsilon^{-1})$:

$$|x_{\perp}(t) - y_{\perp}(t)| \leq C\varepsilon, \quad 0 \leq t \leq c/\varepsilon. \quad (6.3)$$

The constant C is independent of ε and t with $0 \leq t \leq c/\varepsilon$, but depends on the initial velocity bound M , on bounds of derivatives of B_1 and E on the compact set K , and on c .

Remark 6.2 It is well known in the physical literature (going back to [20, Eq. (13)]) that the perpendicular velocity is largely determined by the $E \times B$ term, as is justified by

averaging techniques; see also, e.g., [10, Eq. (6)] in the numerical literature. An $O(\varepsilon)$ bound over times $O(\varepsilon^{-1})$ as in (6.3) was recently proved in [12] in the more restricted setting of a constant magnetic field ($B_{\parallel} \equiv 0$) and an electric field with $E_{\parallel} \equiv 0$.

Proof The proof uses the modulated Fourier expansion of Theorem 4.1, in particular the differential equations for $z_{\pm 1}^0$ and $z_{\pm 1}^{\pm 1}$ in part (c), and the familiar argument of Lady Windermere's fan [18]. We structure the proof into four parts (a)–(d).

(a) Over the (short) time interval $0 \leq t \leq 1$, Theorem 4.1 yields that

$$x_{\perp}(t) = z_{\perp}^0(t) + z_1^1(t)e^{it/\varepsilon} + z_{-1}^{-1}(t)e^{-it/\varepsilon} + O(\varepsilon^2),$$

where $z_{\perp}^0(t) = z_1^0(t) + z_{-1}^0(t)$ and $z_{\pm 1}^{\pm 1}(t)$ satisfy the differential equations

$$\begin{aligned}\dot{z}_{\perp}^0 &= \varepsilon \left((\dot{z}_{\parallel}^0 + \dot{z}_{\perp}^0) \times B_1(z_{\perp}^0) + E_{\perp}(z_{\perp}^0) \right) \times B_0 + O(\varepsilon^2), \\ \dot{z}_1^1 &= P_1(z_1^1 \times B_1(z_{\perp}^0)) + O(\varepsilon^2),\end{aligned}$$

and $z_{-1}^{-1} = \overline{z_1^1}$. We note that $\dot{z}_{\parallel}^0 = \dot{z}_0^0 = \dot{x}_{\parallel} + O(\varepsilon)$, because we have $\frac{d}{dt}(z_0^1 e^{it/\varepsilon}) = (iz_0^1/\varepsilon + \dot{z}_0^1)e^{it/\varepsilon} = O(\varepsilon)$. Moreover, the implicit differential equation for z_{\perp}^0 can be solved for \dot{z}_{\perp}^0 to yield

$$\dot{z}_{\perp}^0 = \varepsilon \left(\dot{z}_{\parallel}^0 \times B_1(z_{\perp}^0) + E_{\perp}(z_{\perp}^0) \right) \times B_0 + O(\varepsilon^2).$$

(b) On every time interval $n \leq t \leq n+1$ (with $n \leq c/\varepsilon$) we can do the same and, denoting by $y_{\perp}^{[n]}$ the function z_{\perp}^0 on this interval and by $z_1^{[n]}$ the function z_1^1 , we have

$$x_{\perp}(t) = y_{\perp}^{[n]}(t) + 2 \operatorname{Re} \left(z_1^{[n]} e^{it/\varepsilon} \right) + O(\varepsilon^2), \quad n \leq t \leq n+1,$$

where $y_{\perp}^{[n]}$ and $z_1^{[n]}$ solve the initial value problems

$$\begin{aligned}\dot{y}_{\perp}^{[n]} &= \varepsilon \left(\dot{x}_{\parallel} \times B_1(y_{\perp}^{[n]}) + E_{\perp}(y_{\perp}^{[n]}) \right) \times B_0, \\ y_{\perp}^{[n]}(n) &= x_{\perp}(n) - 2 \operatorname{Re} \left(z_1^{[n]}(n) e^{in/\varepsilon} \right),\end{aligned}$$

and

$$\begin{aligned}\dot{z}_1^{[n]} &= P_1(z_1^{[n]} \times B_1(y_{\perp}^{[n]})), \\ z_1^{[n]}(n) &= -i\varepsilon P_1 \dot{x}(n).\end{aligned}$$

We consider these initial value problems on the time interval $n \leq t \leq c/\varepsilon$. By the error bound of the modulated Fourier expansion on the interval $[n, n+1]$ (in particular at $t = n+1$) as stated by Theorem 4.1 (b), by the essential uniqueness of the coefficient functions of the modulated Fourier expansion and their bounds as stated by Theorem 4.1 (a), and by the approximation of the modulation functions z_{\perp}^0

and z_1^1 on the interval $[n, n + 1]$ by the functions $y_\perp^{[n]}$ and $z_1^{[n]}$ defined above, which has an $O(\varepsilon^2)$ error because of Theorem 4.1 (c) and (d), we obtain

$$\begin{aligned} y_\perp^{[n]}(n+1) &= y_\perp^{[n+1]}(n+1) + O(\varepsilon^2), \\ z_1^{[n]}(n+1) &= z_1^{[n+1]}(n+1) + O(\varepsilon^2). \end{aligned}$$

In view of the factor ε in front of the right-hand side of the differential equations for $y_\perp^{[n+1]}$ and $y_\perp^{[n]}$, this estimate implies that

$$y_\perp^{[n+1]}(t) - y_\perp^{[n]}(t) = O(\varepsilon^2), \quad n+1 \leq t \leq c/\varepsilon.$$

Moreover, taking the inner product of the differential equation for $z_1^{[n]}$ with $z_1^{[n]}$ shows that

$$\frac{d}{dt} |z_1^{[n]}|^2 = 2 \operatorname{Re} \overline{z_1^{[n]}}^\top \dot{z}_1^{[n]} = 0,$$

and hence

$$|z_1^{[n]}(t)| = |z_1^{[n]}(n)|, \quad n \leq t \leq c/\varepsilon.$$

(c) Next we study the difference between $y_\perp^{[0]}(t)$ and $y_\perp(t)$ of (6.2). We have

$$\begin{aligned} y_\perp^{[0]}(t) - y_\perp(t) &= \left(y_\perp^{[0]}(0) - y_\perp(0) \right) + \varepsilon \int_0^t \left(E_\perp(y_\perp^{[0]}(s)) - E_\perp(y_\perp(s)) \right) \times B_0 \, ds \\ &\quad + \varepsilon \int_0^t \left(\dot{x}_\parallel(s) \times B_1(y_\perp^{[0]}(s)) \right) \times B_0 \, ds. \end{aligned}$$

The difference of the initial values is $O(\varepsilon)$, and the last integral term is bounded using partial integration:

$$\begin{aligned} &\varepsilon \int_0^t \left(\dot{x}_\parallel(s) \times B_1(y_\perp^{[0]}(s)) \right) \times B_0 \, ds \\ &= \varepsilon \left(x_\parallel(t) \times B_1(y_\perp^{[0]}(t)) - x_\parallel(0) \times B_1(y_\perp^{[0]}(0)) \right) \times B_0 \\ &\quad - \varepsilon \int_0^t \left(x_\parallel(s) \times \frac{\partial B_1}{\partial x_\perp}(y_\perp^{[0]}(s)) \dot{y}_\perp^{[0]}(s) \right) \times B_0 \, ds. \end{aligned}$$

This is $O(\varepsilon)$ for $0 \leq t \leq c/\varepsilon$, because x_\parallel is bounded by assumption and $\dot{y}_\perp^{[0]}(s) = O(\varepsilon)$. With a Lipschitz bound of E and the Gronwall lemma, this yields that the difference between $y_\perp^{[0]}(t)$ and $y_\perp(t)$ of (6.2) is bounded by

$$y_\perp^{[0]}(t) - y_\perp(t) = O(\varepsilon), \quad 0 \leq t \leq c/\varepsilon.$$

(d) With the above estimates we obtain, for $n \leq t \leq n+1 \leq c/\varepsilon$,

$$\begin{aligned} x_{\perp}(t) - y_{\perp}(t) &= \left(x_{\perp}(t) - y_{\perp}^{[n]}(t) - 2 \operatorname{Re} \left(z_1^{[n]}(t) e^{it/\varepsilon} \right) \right) + 2 \operatorname{Re} \left(z_1^{[n]}(t) e^{it/\varepsilon} \right) \\ &\quad + \sum_{j=0}^{n-1} \left(y_{\perp}^{[j+1]}(t) - y_{\perp}^{[j]}(t) \right) + \left(y_{\perp}^{[0]}(t) - y_{\perp}(t) \right) \\ &= O(\varepsilon) + O(\varepsilon) + O(n\varepsilon^2) + O(\varepsilon) = O(\varepsilon), \end{aligned}$$

which is the stated result. \square

6.2 Perpendicular drift of numerical approximations

For the Boris algorithm with large step size (1.1) and a small perpendicular component of the starting velocity we obtain the following result from Theorem 4.2.

Theorem 6.3 *Under the assumptions of Theorem 4.2 (in particular (4.2)–(4.3)), and provided that the numerical solution x^n of the Boris method stays in a compact set K for $0 \leq t \leq c\varepsilon^{-1}$ (with K and c independent of ε and h), the solution $y_{\perp}(t)$ of the initial-value problem for the slow differential equation (6.2) remains $O(h^2)$ -close to the perpendicular component of x^n over times $O(\varepsilon^{-1})$:*

$$|x_{\perp}^n - y_{\perp}(t_n)| \leq Ch^2, \quad 0 \leq t_n = nh \leq c/\varepsilon. \quad (6.4)$$

The constant C is independent of ε and h and n with $0 \leq nh \leq c/\varepsilon$, but depends on the initial velocity bound, on bounds of derivatives of B_1 and E on the compact set K , and on c .

Proof The proof uses Theorem 4.2 and Lady Windermere's fan in the same way as in the proof of Theorem 6.1, without any additional difficulty. We therefore omit the details. \square

Analogous results hold true also for the standard and filtered variational integrators, for the latter with non-resonant stepsizes (4.6), using the corresponding modulated Fourier expansions as given in Theorems 4.2 and 4.3. We note that for the filtered variational integrator we do not need the smallness assumption (4.3) for the perpendicular component of the velocity required for the Boris and standard variational integrators, but the mere boundedness of the initial velocity suffices for the filtered variational integrator. However, in view of the $O(\varepsilon h)$ remainder term (instead of $O(\varepsilon h^2)$) in the differential equation for $z_{\pm 1}^0$ in part (c) of Theorem 4.3, the error bound of x_{\perp}^n for the filtered variational integrator is only $O(h)$ instead of $O(h^2)$.

Numerical experiment. For the example of Section 2 and for the methods (i)–(iii) of the numerical experiments of Section 5, Figure 3 shows the projection of the computed particle trajectory onto the plane perpendicular to $B_0 = e_3$ up to time $T = 5/\varepsilon$, for the fixed step size $h = 10^{-2}$ and three values of ε . The exact solution has a gyroradius of $O(\varepsilon)$, which is too small to be visible in the figure. It is observed

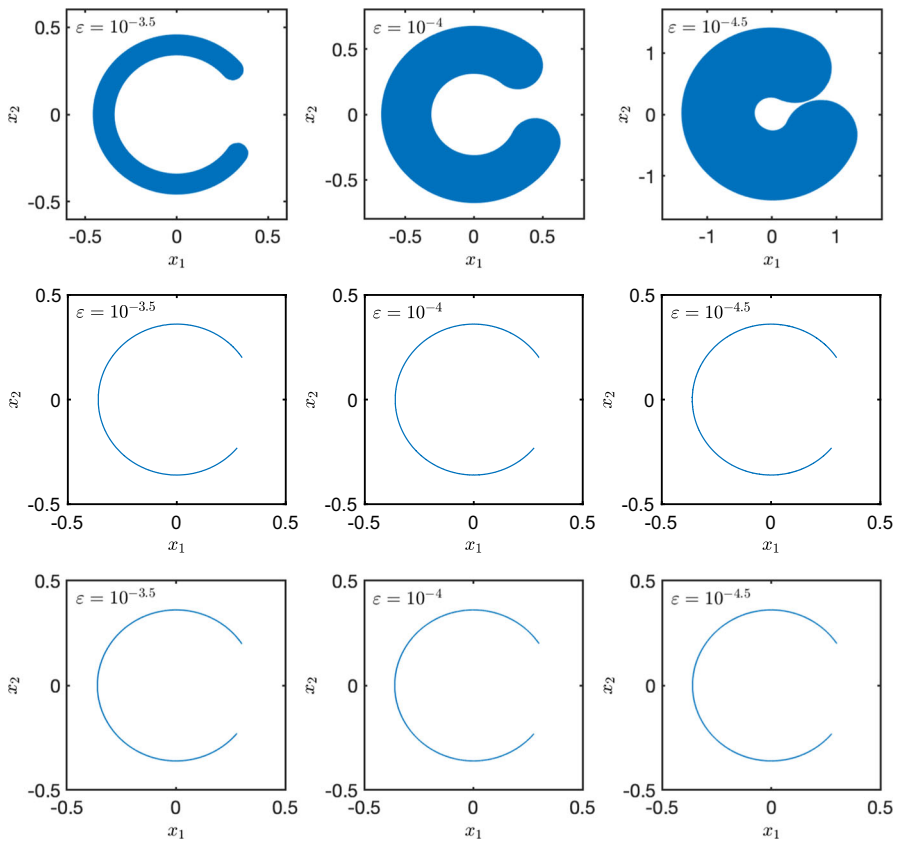


Fig. 3 Particle trajectory for times $t \leq 5/\varepsilon$ projected onto the perpendicular plane as computed by the Boris algorithm with starting values $x(0)$, $v(0)$ (top row), with modified initial values (3.4) (centre row), and by the filtered variational integrator with starting values $x(0)$, $v(0)$ (bottom row). The step size used is $h = 10^{-2}$ in all cases

that the Boris algorithm with the original initial velocity as starting velocity shows a substantially enlarged gyroradius for $h \gg \varepsilon$, while after modifying the starting velocity to (3.4), the Boris algorithm shows correct results. The same behaviour is observed also for the standard variational integrator (not shown here, since the pictures are indistinguishable). In contrast, the filtered variational integrator shows correct results both for the original initial values (as shown) and for the modified starting velocity (not shown here).

7 Long-term near-conservation of magnetic moment and energy

7.1 Time scale ε^{-1} : Standard variational integrator

For the standard (unfiltered) variational integrator with step sizes (1.1) and the modified starting velocity (3.4) we can show energy conservation up to $O(h^2)$ over time ε^{-1} ,

provided that $h^6 \leq \varepsilon$. We do not have, and do not expect, such a result for the Boris algorithm in a non-uniform magnetic field (2.1).

Theorem 7.1 *Under the assumptions of Theorem 4.2, and provided that the numerical solution x^n of the variational integrator with step size (1.1) and starting velocity (3.4) stays in a compact set K for $0 \leq t \leq c\varepsilon^{-1}$ (with K and c independent of ε and h), the total energy (2.3) remains $O(h^2)$ -close to the initial energy over times $c \min(\varepsilon^{-1}, h^{-6})$:*

$$|H(x^n, v^n) - H(x^0, v^0)| \leq Ch^2, \quad 0 \leq nh \leq c \min(\varepsilon^{-1}, h^{-6}). \quad (7.1)$$

Moreover, with the modified initial velocity, the magnetic moment (2.4) remains $O(\varepsilon^2)$ small over times $c\varepsilon^{-1}$:

$$|I(x^n, v^n)| \leq C\varepsilon^2, \quad 0 \leq nh \leq c\varepsilon^{-1}. \quad (7.2)$$

The constants C are independent of ε and h and n with $0 \leq nh \leq c/\varepsilon$, but depend on bounds of derivatives of B_1 and E on the compact set K , and on c .

Proof The proof uses Theorem 4.2 and arguments from the proof of Proposition 6.2 in [13]. We first consider the energy behaviour over a short time interval of length 1, over which we can apply Theorem 4.2. With $D = d/dt$ and the shift operator e^{hD} , with $\delta(\zeta) = \zeta - \zeta^{-1}$ and $\rho(\zeta) = \zeta - 2 + \zeta^{-1}$, and with the expansions $\delta(e^h)/(2h) = (1 + \alpha_2 h^2 + \alpha_4 h^4 + \dots)$ and $\rho(e^h)/h^2 = (1 + \beta_2 h^2 + \beta_4 h^4 + \dots)$, we insert the decomposition (4.4) into the two-step formulation (3.2) of the numerical method and obtain the equation for the function $y(t)$ in (4.4) as

$$\begin{aligned} \ddot{y} + \beta_2 h^2 y^{(4)} + \beta_4 h^4 y^{(6)} + \dots &= (\dot{y} + \alpha_2 h^2 y^{(3)} + \alpha_4 h^4 y^{(5)} + \dots) \times \frac{B_0}{\varepsilon} \\ &+ A_1'(y)^\top \frac{\delta(e^{hD})}{2h} y - \frac{\delta(e^{hD})}{2h} A_1(y) - \nabla \phi(y) + O(|z|^2) + O(h^N), \end{aligned} \quad (7.3)$$

where the left-hand side contains only *even*-order derivatives of y , and the right-hand side contains only *odd*-order derivatives of y . We multiply both sides of (7.3) with \dot{y}^\top . The multiplied left-hand side is the time derivative of an expression in which the appearing second and higher derivatives of y can be substituted as functions of (y, \dot{y}) via the differential equation for y in part (c) of Theorem 4.2; cf. [14]. On the right-hand side we have

$$\begin{aligned} &-\dot{y}^\top \varepsilon^{-1} \widehat{B}_0 (\alpha_2 h^2 y^{(3)} + \alpha_4 h^4 y^{(5)} + \dots) \\ &+ \dot{y}^\top \frac{1}{2h} \left(A_1'(y)^\top \delta(e^{hD}) y - \delta(e^{hD}) A_1(y) \right) - \frac{d}{dt} \phi(y) + O(|z|^2) + O(h^N). \end{aligned} \quad (7.4)$$

The first term is $O(h^2)$ because $\dot{y}_\perp = \dot{y}_1 + \dot{y}_{-1}$ and its derivatives are $O(\varepsilon)$ by Theorem 4.2. Since \widehat{B}_0 is a skew-symmetric matrix, the first term is again the time derivative of an expression in which the appearing second and higher derivatives of y

can be substituted as functions of (y, \dot{y}) ; cf. [14]. The same holds true for the second term, as is shown in the proof of Proposition 6.2 of [13].

We have thus found a function $H_h(x, v)$ with the properties that uniformly for all x in a bounded domain and all bounded v with $v_\perp = O(\varepsilon)$ we have

$$H_h(x, v) = H(x, v) + O(h^2), \quad (7.5)$$

$$\frac{d}{dt} H_h(y(t), \dot{y}(t)) = O(|z(t)|^2) + O(h^N). \quad (7.6)$$

We now consider the equation for z . With the starting velocity (3.4) we have $|z(0)| \leq c_0 h^4$ for some constant c_0 ; see part (d) of Theorem 4.2. The differential equation for $z_\perp = z_1 + z_{-1}$ can be written as

$$\dot{z}_\perp = \frac{4\varepsilon}{h^2} z_\perp \times B_0 + O(\varepsilon |z_\perp|) + O(\varepsilon h^N).$$

Multiplying this equation with $2(z_\perp)^\top$ and noting that $2(z_\perp)^\top \dot{z}_\perp = (d/dt)|z_\perp|^2$, we obtain

$$\frac{d}{dt} |z_\perp|^2 \leq C\varepsilon |z_\perp|^2 + O(\varepsilon h^N),$$

which shows that $|z_\perp(t)| \leq e^{\tilde{c}\varepsilon t} |z_\perp(0)| + O(t\varepsilon h^N)$. Moreover, from the proof of Theorem 4.2 we have $|z_0(t)| \leq Ch^2 |z_\perp(t)|$. Patching many short time intervals of length 1 together as in part (b) of the proof of Theorem 6.1, we find that on each of these intervals up to time $c\varepsilon^{-1}$ (but not on longer time intervals $\varepsilon^{-\alpha}$ with $\alpha > 1$ because of the $e^{\tilde{c}\varepsilon t}$ exponential growth of our bound of z_\perp), we can apply Theorem 4.2 and the oscillatory component z on the interval remains of size $O(h^4)$. By (7.6) we thus have

$$H_h(y(t), \dot{y}(t)) = H_h(y(0), \dot{y}(0)) + O(th^8).$$

(Different to Section 6, we now do not put a superscript on y and z to designate the interval of length 1 in which t lies). Together with

$$H(x^n, v^n) = H_h(x^n, v^n) + O(h^2) = H_h(y(t_n), \dot{y}(t_n)) + O(h^2),$$

this yields the stated result for the energy.

The long-term smallness of the magnetic moment follows from (2.5) and the relation $v_\perp^n = \dot{y}_\perp(t_n) - (-1)^n \dot{z}_\perp(t_n) + O(\varepsilon h^2)$. This yields $v_\perp^n = O(\varepsilon)$ by the differential equations in part (c) of Theorem 4.2 for $y_{\pm 1}$ and $z_{\pm 1}$, which contain a factor ε on the right-hand side. These functions are again patched together over many short intervals as is done in the proof of Theorem 6.1. \square

7.2 Time scale ε^{-N} for $N > 1$: Filtered variational integrator

We have the following result on the long-term near-conservation of magnetic moment and energy by the filtered variational integrator with non-resonant large step sizes (4.5) with (4.6).

Theorem 7.2 *Let $M > N$ be arbitrary positive integers. Under the assumptions of Theorem 4.3 (in particular (4.5)–(4.6) and an initial velocity bounded independently of ε), and provided that the numerical positions x^n of the filtered variational integrator stay in a compact set K for $0 \leq t \leq c \varepsilon^{-N}$ (with K and c independent of ε and h), the magnetic moment and the total energy along the numerical solution (x^n, v^n) remain almost conserved over such long times:*

$$\begin{aligned} |I(x^n, v^n) - I(x^0, v^0)| &\leq Ch \\ |H(x^n, v^n) - H(x^0, v^0)| &\leq Ch \end{aligned} \quad \text{for } 0 \leq t \leq c \min(h^{-M}, \varepsilon^{-N}).$$

The constant C is independent of ε and h and n with $0 \leq nh \leq c \min(h^{-M}, \varepsilon^{-N})$, but depends on the initial velocity bound, on bounds of derivatives of B_1 and E on the compact set K , on c , and on the choice of M and N .

Proof The proof uses arguments that are very similar to the proofs of Theorems 2.2 and 2.3 of [15] on the long-term near-conservation properties of the standard variational integrator for step sizes $h \leq c\varepsilon$. We therefore only indicate the main steps in the proof, which are marked as items (i)–(iv) below.

To simplify the expressions for the remainder terms, we assume in the following the mild stepsize restriction $h^m \leq \varepsilon$ for some fixed $m > 2$ and we choose $M \geq mN$. This is only done for ease of presentation and allows us to cover the time scale ε^{-N} . Without this assumption we arrive at the stated time scale $\min(h^{-M}, \varepsilon^{-N})$.

(i) (*Lagrangian structure of the modulation equations*; cf. [15, (5.23)]) Over a time interval of length 1 we consider the modulation functions $z^k(t)$ of Theorem 4.3 multiplied with the corresponding highly oscillatory exponentials:

$$y^k(t) = z^k(t)e^{ikt/\varepsilon} \text{ for } |k| \leq N \text{ and } y^k(t) = 0 \text{ for } |k| > N.$$

We write $\mathbf{y} = (y^k)_{k \in \mathbb{Z}}$ and define the extended potentials

$$\begin{aligned} \mathcal{U}(\mathbf{y}) &= \sum_{\substack{0 \leq m \leq N \\ s(\alpha)=0}} \frac{1}{m!} \phi^{(m)}(y^0) \mathbf{y}^\alpha \\ \mathcal{A}(\mathbf{y}) &= (\mathcal{A}_k(\mathbf{y}))_{k \in \mathbb{Z}} = \left(\sum_{\substack{0 \leq m \leq N \\ s(\alpha)=k}} \frac{1}{m!} A^{(m)}(y^0) \mathbf{y}^\alpha \right)_{k \in \mathbb{Z}}, \end{aligned}$$

where the sums are taken over all multi-indices $\alpha = (\alpha^1, \dots, \alpha^m)$ with $\alpha^j \in \mathbb{Z} \setminus \{0\}$ with prescribed sum $s(\alpha) = \alpha^1 + \dots + \alpha^m$, and where we use the notation

$\phi^{(m)}(y^0)\mathbf{y}^\alpha = \phi^{(m)}(y^0)(y^{\alpha^1}, \dots, y^{\alpha^m})$ for the m -linear m th derivative of ϕ evaluated at $(y^{\alpha^1}, \dots, y^{\alpha^m})$, and analogously for $A^{(m)}(y^0)\mathbf{y}^\alpha$. The terms for $m = 0$ are to be interpreted as $\phi(y^0)$ and $A(y^0)$.

The system of modulation equations of the filtered variational integrator can then be written, up to $O(\varepsilon^N)$, as the discrete Euler-Lagrange equations corresponding to the discrete Lagrangian

$$\begin{aligned} \mathcal{L}_h(\mathbf{y}^n, \mathbf{y}^{n+1}) &= \frac{h}{2}(\mathbf{v}^{n+1/2})^*(\Psi^{-1} \otimes I)\mathbf{v}^{n+1/2} \\ &\quad + \frac{h}{2}(\mathcal{A}(\mathbf{y}^n) + \mathcal{A}(\mathbf{y}^{n+1}))^*\mathbf{v}^{n+1/2} - \frac{h}{2}(\mathcal{U}(\mathbf{y}^n) + \mathcal{U}(\mathbf{y}^{n+1})) \end{aligned}$$

with $\mathbf{v}^{n+1/2} = (\mathbf{y}^{n+1} - \mathbf{y}^n)/h$, which differs from that of the standard variational integrator only by the modified kinetic energy term with Ψ^{-1} . We thus have

$$\Psi^{-1}\delta_h^2 y^k = \sum_{j \in \mathbb{Z}} \left(\frac{\partial \mathcal{A}_j}{\partial y^k}(\mathbf{y}) \right)^* \delta_{2h} y^j - \delta_{2h} \mathcal{A}_k(\mathbf{y}) - \left(\frac{\partial \mathcal{U}}{\partial y^k}(\mathbf{y}) \right)^* + O(\varepsilon^N), \quad (7.7)$$

where $\delta_{2h} f(t) = (f(t+h) - f(t-h))/(2h)$ and $\delta_h^2 f(t) = (f(t+h) - 2f(t) + f(t-h))/h^2$ denote the first-order and second-order symmetric difference quotients, respectively.

(ii) (*Almost-invariant close to the magnetic moment*; cf. [15, Theorem 5.2]) With the group action $S(\lambda)\mathbf{y} = (e^{ik\lambda} y^k)_{k \in \mathbb{Z}}$ (for $\lambda \in \mathbf{R}$), we have

$$\mathcal{U}(S(\lambda)\mathbf{y}) = \mathcal{U}(\mathbf{y}), \quad \mathcal{A}(S(\lambda)\mathbf{y}) = S(\lambda)\mathcal{A}(\mathbf{y}) \quad \text{for all } \lambda.$$

Differentiation with respect to λ (at $\lambda = 0$) yields

$$\begin{aligned} \sum_{k \in \mathbb{Z}} ik \frac{\partial \mathcal{U}}{\partial y^k}(\mathbf{y}) y^k &= 0 \\ \sum_{j \in \mathbb{Z}} ij \frac{\partial \mathcal{A}_k}{\partial y^j}(\mathbf{y}) y^j &= ik \mathcal{A}_k(\mathbf{y}) \quad \text{for } k \in \mathbb{Z}. \end{aligned}$$

Multiplying (7.7) with $-ik(y^k)^*$, summing over k and using these relations yields that the function

$$\begin{aligned} \mathcal{I}_h[\mathbf{y}](t) &= -\frac{i}{\varepsilon h} \sum_k k y^k(t)^* \Psi^{-1} y^k(t+h) \\ &\quad + \frac{i}{2\varepsilon} \sum_k k \left(\mathcal{A}_k(\mathbf{y}(t))^* y^k(t+h) - y^k(t)^* \mathcal{A}_k(\mathbf{y}(t+h)) \right) \end{aligned}$$

satisfies

$$\mathcal{I}_h[\mathbf{y}](t) - \mathcal{I}_h[\mathbf{y}](t-h) = O(h\varepsilon^{N+1}) \quad (7.8)$$

and is thus an almost-invariant of the modulation system. Using the bounds of the modulation functions, we find that

$$\begin{aligned}\mathcal{I}_h[\mathbf{y}](t) = & -\frac{i}{\varepsilon h} \left(y^1(t)^* \Psi^{-1} y^1(t+h) - y^{-1}(t)^* \Psi^{-1} y^{-1}(t+h) \right) \\ & + \frac{i}{2\varepsilon^2} \left(y^1(t)^* (y^1(t+h) \times B_0) - y^{-1}(t)^* (y^{-1}(t+h) \times B_0) \right) + O(\varepsilon).\end{aligned}$$

Here, a calculation shows that the first term equals $(1 + \cos(h/\varepsilon))|z_1^1|^2/\varepsilon^2 + O(h)$, and the second term equals $-\cos(h/\varepsilon)|z_1^1|^2/\varepsilon^2 + O(h)$. So we obtain

$$\mathcal{I}_h[\mathbf{y}](t) = \frac{1}{\varepsilon^2} |z_1^1(t)|^2 + O(h).$$

On the other hand, since

$$\begin{aligned}v^n &= \Phi \frac{x^{n+1} - x^{n-1}}{2h} + O(\varepsilon) \\ &= \left(I + (1 - \operatorname{sinc}(h/\varepsilon)^{-1}) \hat{B}_0^2 \right) \frac{x^{n+1} - x^{n-1}}{2h} + O(\varepsilon) \\ &= \dot{z}^0(t) + \frac{i}{\varepsilon} \left(z_1^1(t) e^{it/\varepsilon} - z_{-1}^1(t) e^{-it/\varepsilon} \right) + O(h) \quad \text{at } t = nh\end{aligned}$$

and $\dot{z}^0(t) \times B_0 = O(\varepsilon)$, we find that

$$I(x^n, v^n) = \frac{1}{2} |v^n \times B_0|^2 + O(\varepsilon) = \frac{1}{\varepsilon^2} |z_1^1(t)|^2 + O(h).$$

So we obtain that the magnetic moment along the numerical solution is $O(h)$ -close to the almost-invariant:

$$I(x^n, v^n) = \mathcal{I}_h[\mathbf{y}](nh) + O(h). \quad (7.9)$$

(iii) (*Almost-invariant close to the total energy*; cf. [15, Theorem 5.3]) Multiplying (7.7) with $(\dot{y}^k)^*$ and summing over k gives

$$\begin{aligned}\sum_k (\dot{y}^k)^* \Psi^{-1} \delta_h^2 y^k - \sum_k \left(\frac{d}{dt} \mathcal{A}_k(\mathbf{y})^* \delta_{2h} y^k - (\dot{y}^k)^* \delta_{2h} \mathcal{A}_k(\mathbf{y}) \right) + \frac{d}{dt} \mathcal{U}(\mathbf{y}) \\ = O(\varepsilon^N).\end{aligned} \quad (7.10)$$

The arguments in the proof of Theorem 5.3 in [15] show that each of the three terms on the left-hand side is a total differential up to $O(\varepsilon^N)$. So there exists a function

$$\mathcal{H}_h[\mathbf{y}](t) = \mathcal{K}_h[\mathbf{y}](t) + \mathcal{M}_h[\mathbf{y}](t) + \mathcal{U}[\mathbf{y}](t),$$

where the time derivatives of the three terms on the right-hand side equal the three corresponding terms on the left-hand side of (7.10), and we have

$$\frac{d}{dt} \mathcal{H}_h[\mathbf{y}](t) = O(\varepsilon^N).$$

We now determine the dominant part of $\mathcal{H}_h[\mathbf{y}]$. We find

$$\begin{aligned} \mathcal{K}_h[\mathbf{y}] &= \frac{1}{2}(\dot{z}^0)^* \Psi^{-1} \dot{z}^0 + \frac{2(z^1)^* \Psi^{-1} z^1}{h^2} \left(\frac{h}{\varepsilon} \sin(h/\varepsilon) - 2 \sin^2(h/2\varepsilon) \right) + O(h) \\ &= \frac{1}{2}|\dot{z}^0|^2 + \frac{2|z_1^1|^2}{\tan^2(h/2\varepsilon)h^2} \left(\frac{h}{\varepsilon} \sin(h/\varepsilon) - 2 \sin^2(h/2\varepsilon) \right) + O(h) \\ &= \frac{1}{2}|\dot{z}^0|^2 + \frac{|z_1^1|^2}{\varepsilon^2} (1 + \cos(h/\varepsilon)) + O(h) \\ \mathcal{M}_h[\mathbf{y}] &= -\cos(h/\varepsilon) \frac{|z_1^1|^2}{\varepsilon^2} + O(h) \\ \mathcal{U}[\mathbf{y}] &= \phi(z^0) + O(h). \end{aligned}$$

Thus we have

$$\mathcal{H}_h[\mathbf{y}](t) = \frac{1}{2}|\dot{z}^0(t)|^2 + \frac{|z_1^1(t)|^2}{\varepsilon^2} + \phi(z^0(t)) + O(h). \quad (7.11)$$

On the other hand, from the formula for v^n in (ii) we have, at $t = nh$,

$$\frac{1}{2}|v^n|^2 = \frac{1}{2}|\dot{z}^0(t)|^2 + \frac{|z_1^1(t)|^2}{\varepsilon^2} + O(h).$$

The energy along the numerical solution is therefore

$$H(x^n, v^n) = \frac{1}{2}|v^n|^2 + \phi(x^n) = \frac{1}{2}|\dot{z}^0(t)|^2 + \frac{|z_1^1(t)|^2}{\varepsilon^2} + \phi(z^0(t)) + O(h).$$

and hence we have

$$H(x^n, v^n) = \mathcal{H}_h[\mathbf{y}](t) + O(h).$$

(iv) (*From short to long time intervals*; cf. [15, Section 4.5], [17, Section XIII.7]). The stated long-time near-conservation results are now obtained by patching together the short-time near-conservation results of (ii) and (iii) over many intervals of length 1, via an often-used argument that involves the uniqueness up to $O(\varepsilon^{N+1})$ of the modulation functions. \square

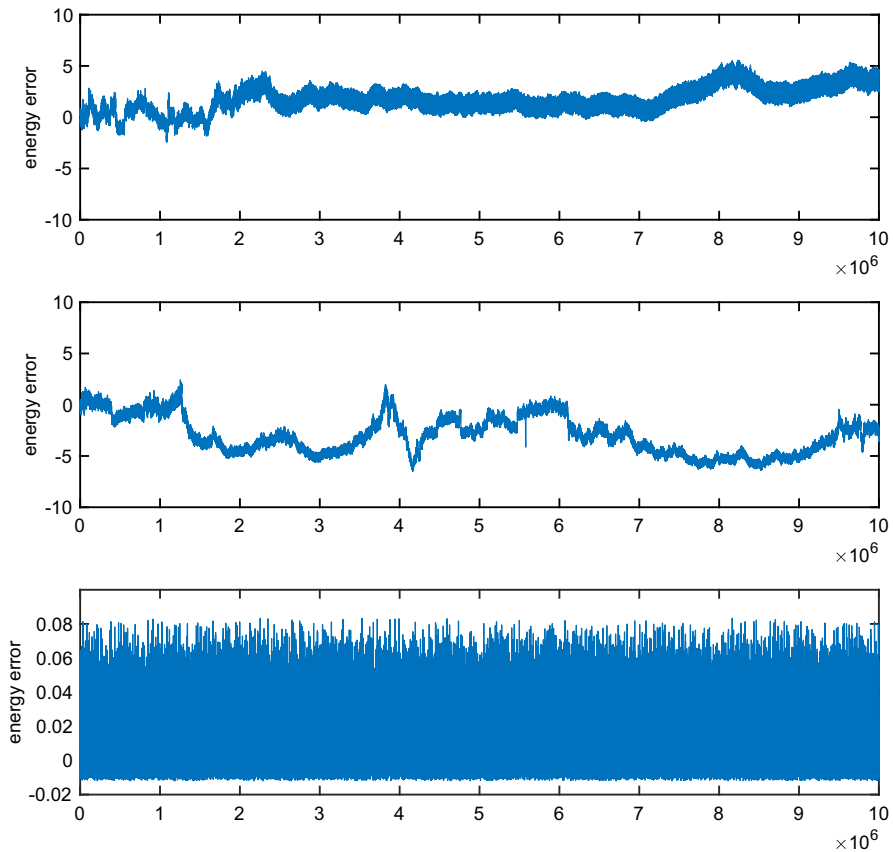


Fig. 4 Energy error $H(x_n, v_n) - H(x_0, v_0)$ along the numerical solutions of the Boris algorithm (top), of the standard variational integrator (centre) and of the filtered variational integrator (bottom), obtained with $\varepsilon = 10^{-4}$ and $h = 10^{-2}$

Numerical experiment. We illustrate the energy behaviour of the numerical methods for the magnetic field

$$B(x) = \frac{1}{\varepsilon} \begin{pmatrix} 1 \\ 0 \\ 0.5 \end{pmatrix} + \begin{pmatrix} x_2 - x_3 \\ x_1 + x_3 \\ x_2 - x_1 \end{pmatrix}$$

and the scalar potential $\phi(x) = x_1^3 - x_2^3 + \frac{1}{5}x_1^4 + x_2^4 + x_3^4$. We take the initial values $x(0) = (0, 1, 0.1)^\top$, $v(0) = (0.09, 0.05, 0.2)^\top$.

We apply the three numerical integrators of Section 3 with $\varepsilon = 10^{-4}$, step size $h = 10^{-2}$, and final time $T = 10^7$. Figure 4 shows the energy error $H(x_n, v_n) - H(x_0, v_0)$ along the numerical solutions of the Boris algorithm, the standard variational integrator and the filtered variational integrator, taking the initial values $x(0), v(0)$ as starting values for all three methods.

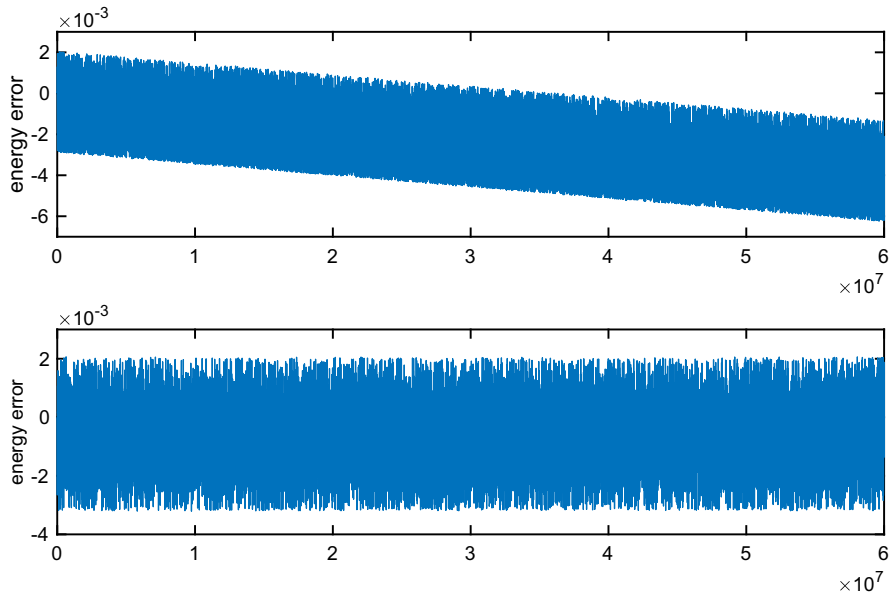


Fig. 5 Energy error $H(x_n, v_n) - H(x_0, v_0)$ along the numerical solutions of the Boris algorithm (top) and of the standard variational integrator (bottom) obtained with modified initial values (3.4) and step size $h = 10^{-2}$, for $\varepsilon = 10^{-4}$

The energy errors of the Boris algorithm and the variational integrator (top and centre picture) appear to behave randomly. Running several trajectories corresponding to random perturbations of the initial data of magnitude 10^{-14} showed energy errors that look like random walks with a deviation of magnitude 10 for $t \leq 10^6$. For larger times, some of the trajectories showed blow-up behaviour. A similar random walk behaviour was observed also for the deviation of the magnetic moment, but the deviation was less than 10^{-1} for $t \leq 10^7$.

In contrast, the energy error of the filtered variational integrator oscillates with a small amplitude without drift (bottom picture of Figure 4). The error $I(x_n, v_n) - I(x_0, v_0)$ of the magnetic moment along the numerical solution of the filtered variational integrator has a very similar behaviour (not shown here).

If we apply the Boris algorithm and the standard variational integrator with *modified initial values* (3.4), then the magnetic moment remains small over very long time, oscillating between 0 and approximately $2 \cdot 10^{-6}$ over the whole time interval. In this case of modified initial velocity, we observe very good near-conservation of energy for the variational integrator while there is a linear drift for the Boris algorithm. However, this drift becomes dominant over the small oscillations in the energy only for times $\sim 10^7$; see Figure 5.

Acknowledgements This work was partially supported by the Swiss National Science Foundation, grant No. 200020_192129, and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID258734477 – SFB 1173. The work by Yanyan Shi was done at the University of Tübingen

during her one-year research stay, which was funded by a scholarship provided by the University of the Chinese Academy of Sciences (UCAS).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Benettin, G., Sempio, P.: Adiabatic invariants and trapping of a point charge in a strong nonuniform magnetic field. *Nonlinearity* **7**(1), 281 (1994)
2. Birdsall, C.K., Langdon, A.B.: *Plasma Physics via Computer Simulation*. Taylor and Francis Group, New York (2005)
3. Boris, J. P.: Relativistic plasma simulation-optimization of a hybrid code. *Proceeding of Fourth Conference on Numerical Simulations of Plasmas* (November 1970), 3–67
4. Brizard, A.J., Hahm, T.S.: Foundations of nonlinear gyrokinetic theory. *Rev. Modern Phys.* **79**(2), 421–468 (2007)
5. Chartier, P., Crouseilles, N., Lemou, M., Méhats, F., Zhao, X.: Uniformly accurate methods for Vlasov equations with non-homogeneous strong magnetic field. *Math. Comp.* **88**(320), 2697–2736 (2019)
6. Chartier, P., Crouseilles, N., Lemou, M., Méhats, F., Zhao, X.: Uniformly accurate methods for three dimensional Vlasov equations under strong magnetic field with varying direction. *SIAM J. Sci. Comput.* **42**(2), B520–B547 (2020)
7. Crouseilles, N., Lemou, M., Méhats, F., Zhao, X.: Uniformly accurate particle-in-cell method for the long time solution of the two-dimensional Vlasov-Poisson equation with uniform strong magnetic field. *J. Comput. Phys.* **346**, 172–190 (2017)
8. Derouillat, J., Beck, A., Pérez, F., Vinci, T., Chiamello, M., Grassi, A., Flé, M., Bouchard, G., Plotnikov, I., Aunai, N., et al.: Smilei: A collaborative, open-source, multi-purpose particle-in-cell code for plasma simulation. *Computer Physics Commun.* **222**, 351–373 (2018)
9. Ellison, C.L., Burby, J.W., Qin, H.: Comment on "Symplectic integration of magnetic systems": A proof that the Boris algorithm is not variational. *J. Comput. Phys.* **301**, 489–493 (2015)
10. Filbet, F., Rodrigues, L.M.: Asymptotically stable particle-in-cell methods for the Vlasov-Poisson system with a strong external magnetic field. *SIAM J. Numer. Anal.* **54**(2), 1120–1146 (2016)
11. Filbet, F., Rodrigues, L.M.: Asymptotically preserving particle-in-cell methods for inhomogeneous strongly magnetized plasmas. *SIAM J. Numer. Anal.* **55**(5), 2416–2443 (2017)
12. Filbet, F., Rodrigues, L.M., Zakerzadeh, H.: Convergence analysis of asymptotic preserving schemes for strongly magnetized plasmas. *Numer. Math.* **149**(3), 549–593 (2021)
13. Hairer, E., Lubich, C.: Symmetric multistep methods for charged particle dynamics. *SMAI J. Comput. Math.* **3**, 205–218 (2017)
14. Hairer, E., Lubich, C.: Energy behaviour of the Boris method for charged-particle dynamics. *BIT* **58**, 969–979 (2018)
15. Hairer, E., Lubich, C.: Long-term analysis of a variational integrator for charged-particle dynamics in a strong magnetic field. *Numer. Math.* **144**(3), 699–728 (2020)
16. Hairer, E., Lubich, C., Wang, B.: A filtered Boris algorithm for charged-particle dynamics in a strong magnetic field. *Numer. Math.* **144**(4), 787–809 (2020)
17. Hairer, E., Lubich, C., Wanner, G.: *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed. Springer Series in Computational Mathematics 31. Springer-Verlag, Berlin, 2006
18. Hairer, E., Nørsett, S. P., Wanner, G.: *Solving Ordinary Differential Equations I. Nonstiff Problems*, 2nd ed. Springer Series in Computational Mathematics 8. Springer, Berlin, 1993

19. Kruskal, M.: The gyration of a charged particle. *Rept. PM-S-33 (NYO-7903)*, Princeton University, Project Matterhorn (1958)
20. Northrop, T. G.: *The adiabatic motion of charged particles*. Interscience Tracts on Physics and Astronomy, Vol. 21. Interscience Publishers John Wiley & Sons New York-London-Sydney, 1963
21. Possanner, S.: Gyrokinetics from variational averaging: existence and error bounds. *J. Math. Phys.* **59**, 8 (2018), 082702, 34
22. Qin, H., Zhang, S., Xiao, J., Liu, J., Sun, Y., Tang, W. M.: Why is Boris algorithm so good? *Physics of Plasmas* **20**, 8 (2013), 084503.1–4
23. Ricketson, L. F., Chacón, L.: An energy-conserving and asymptotic-preserving charged-particle orbit implicit time integrator for arbitrary electromagnetic fields. *J. Comput. Phys.* (2020), 109639
24. Wang, B., Zhao, X.: Error estimates of some splitting schemes for charged-particle dynamics under strong magnetic field. *SIAM J. Numer. Anal.* **59**(4), 2075–2105 (2021)
25. Webb, S.D.: Symplectic integration of magnetic systems. *J. Comput. Phys.* **270**, 570–576 (2014)
26. Xiao, J., Qin, H.: Slow manifolds of classical Pauli particle enable structure-preserving geometric algorithms for guiding center dynamics. *Computer Physics Comm.* **265**, 107981 (2021)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.