Multi-revolution composition methods for highly oscillatory differential equations

Chartier, Philippe; Makazaga, Joseba; Murua, Ander; Vilmart, Gilles

How to cite

CHARTIER, Philippe et al. Multi-revolution composition methods for highly oscillatory differential equations. In: Numerische Mathematik, 2014, vol. 128, n° 1, p. 167–192. doi: 10.1007/s00211-013-0602-0

This publication URL: https://archive-ouverte.unige.ch//unige:41957
Publication DOI: 10.1007/s00211-013-0602-0

Multi-revolution composition methods for highly oscillatory differential equations

Philippe Chartier[1], Joseba Makazaga[2], Ander Murua[2], and Gilles Vilmart[3,4]

November 4, 2013

Abstract

We introduce a new class of multi-revolution composition methods (MRCM) for the approximation of the N^{th}-iterate of a given near-identity map. When applied to the numerical integration of highly oscillatory systems of differential equations, the technique benefits from the properties of standard composition methods: it is intrinsically geometric and well-suited for Hamiltonian or divergence-free equations for instance. We prove error estimates with error constants that are independent of the oscillatory frequency. Numerical experiments, in particular for the nonlinear Schrödinger equation, illustrate the theoretical results, as well as the efficiency and versatility of the methods.

Keywords: near-identity map, highly-oscillatory, averaging, differential equation, composition method, geometric integration, asymptotic preserving.

MSC numbers: 34K33, 37L05, 35Q55.

1 Introduction

In this paper, we are concerned with the approximation of the N-th iterates of a near-identity smooth map by composition methods. More precisely, considering a smooth map $(\varepsilon, y) \mapsto \varphi_\varepsilon(y)$ of the form

$$\varphi_\varepsilon(y) = y + \varepsilon\Theta_\varepsilon(y), \tag{1}$$

we wish to approximate the result of $N = \mathcal{O}(1/\varepsilon)$ compositions of φ_ε with itself

$$\varphi_\varepsilon^N = \underbrace{\varphi_\varepsilon \circ \cdots \circ \varphi_\varepsilon}_{N \text{ times}} \tag{2}$$

with the aid of a method whose efficiency remains essentially independent of ε.

In order to motivate our composition methods, it will be useful to observe that φ_ε can be seen as one step with step-size ε of a first order integrator for the differential equation

$$\frac{dz(t)}{dt} = \Theta_0(z(t)), \tag{3}$$

[1]INRIA Rennes, IRMAR and ENS Cachan Bretagne, Campus de Beaulieu, F-35170 Bruz, France. Philippe.Chartier@inria.fr

[2]Konputazio Zientziak eta A. A. Saila, Informatika Fakultatea, UPV/EHU, E-20018 Donostia – San Sebastián, Spain. Ander.Murua@ehu.es, Joseba.Makazaga@ehu.es

[3]École Normale Supérieure de Cachan, Antenne de Bretagne, INRIA Rennes, IRMAR, CNRS, UEB, av. Robert Schuman, F-35170 Bruz, France. Gilles.Vilmart@bretagne.ens-cachan.fr

[4]Present address: Université de Genève, Section de mathématiques, 2-4 rue du Lièvre, CP 64, 1211 Genève 4, Switzerland

where $\Theta_0(z) = \frac{d}{d\varepsilon}\varphi_\varepsilon(z)\big|_{\varepsilon=0}$, and thus, $\varphi_\varepsilon^N(y)$ may be interpreted as an approximation at $t = N\varepsilon$ of the solution $z(t)$ of (3) with initial condition

$$z(0) = y. \tag{4}$$

A standard error analysis shows that $\varphi_\varepsilon^N(y) - z(N\varepsilon) = \mathcal{O}(\varepsilon H)$ as $H = N\varepsilon \to 0$, which makes clear that, for sufficiently small $H = \varepsilon N$, $\varphi_\varepsilon^N(y)$ could be approximated by one step $\Psi_H(y) \approx z(H)$ of any pth order integrator applied to the initial value problem (3)–(4) within an error of size $\mathcal{O}(H^{p+1} + \varepsilon H)$. In particular, φ_H can be seen as a first order integrator for the ODE (3), and a second order integrator can be obtained as

$$\Psi_H(y) = \varphi_{H/2} \circ \varphi^*_{H/2}(y), \tag{5}$$

where $\varphi^*_\varepsilon := \varphi^{-1}_{-\varepsilon}$ is the *adjoint map* of φ_ε. More generally, one could consider pth order composition integrators of the form [21]

$$\Psi_H(y) := \varphi_{a_1 H} \circ \varphi^*_{b_1 H} \circ \cdots \circ \varphi_{a_s H} \circ \varphi^*_{b_s H}(y) \approx z(H) \tag{6}$$

with suitable coefficients a_i, b_i (see for instance [16, 22] for particular sets of coefficients choosen for different s and p), that would provide an approximation

$$\Psi_H(y) = z(H) + \mathcal{O}(H^{p+1}) = \varphi_\varepsilon^N(y) + \mathcal{O}(H^{p+1} + \varepsilon H)$$

for $H = N\varepsilon \to 0$. However, the accuracy of the approximation is limited by the given value of the problem parameter ε being sufficiently small. Motivated by that, we generalize the approximation (6) by replacing the real numbers a_i, b_i $(j = 1, \ldots, s)$ by appropriate coefficients $\alpha_j(N), \beta_j(N)$ depending on N, chosen in such a way that φ_ε^N is approximated for sufficiently small $H = N\varepsilon$ within an error of size $\mathcal{O}(H^{p+1})$, where the error constant is independent of N, H, ε. We will say that such a method

$$\Psi_{N,H}(y) := \varphi_{\alpha_1(N)H} \circ \varphi^*_{\beta_1(N)H} \circ \cdots \circ \varphi_{\alpha_s(N)H} \circ \varphi^*_{\beta_s(N)H}(y) \tag{7}$$

is an s-stage pth order *multi-revolution composition method* (MRCM) if

$$\Psi_{N,H}(y) = \varphi_\varepsilon^N(y) + \mathcal{O}(H^{p+1}), \quad \text{for} \quad H = N\varepsilon \to 0. \tag{8}$$

For instance, we will see that the second order standard composition method (5) can be modified to give a second order MRCM (7) with $s = 1$, $\alpha_1(N) = (1 + N^{-1})/2$, and $\beta_1(N) = (1 - N^{-1})/2$,

$$\Psi_{N,H}(y) = \varphi_{\alpha_1(N)H} \circ \varphi^*_{\beta_1(N)H}(y) = \varphi_\varepsilon^N(y) + \mathcal{O}(H^3), \quad \text{for} \quad H = N\varepsilon \to 0,$$

where the constant in $\mathcal{O}$ is independent of N and ε. It is interesting to observe that this second order MRCM reduces in the limit case $N \to \infty$ to the standard composition method (5) (a second order integrator for the ODE (3)), which is consistent with the fact that $\varphi_{H/N}^N$ converges to the H-flow of (3) as $N \to \infty$. More generally, any pth order MRCM (7), gives rise to a pth order standard composition method (6) with

$$a_i = \lim_{N\to\infty} \alpha_i(N), \quad b_i = \lim_{N\to\infty} \beta_i(N).$$

In practice, if one wants to approximately compute the map φ_ε^M for a given small value of ε and large positive integers M within a given error tolerance by means of a s-stage pth order MRCM (7), then one should choose a sufficiently small step-size H to achieve the required accuracy, and accordingly choose N as the integer part of H/ε, in order to approximate $\varphi_\varepsilon^M(y)$, for $M = mN$, $m = 1, 2, 3 \ldots$, as

$$\varphi_\varepsilon^{mN}(y) \approx \Psi_{N,H}(y)^m.$$

These approximations will be computed more efficiently than actually evaluating $\varphi_\varepsilon^{mN}(y)$ if such a positive integer N is larger than $2s$. The case where N is not larger than $2s$ is thus treated as follows: we define

$$\Psi_{N,H}(y) := \varphi_\varepsilon^N, \quad \text{for all } 1 \le N < 2s. \tag{9}$$

Since the error of multi-revolution methods essentially depends on H but not on ε, for a prescribed accuracy (which determines H), in the case $N \ge 2s$, the computational cost may be reduced by a factor of $N/(2s) \le H/(2s\varepsilon)$ which increases as ε decreases, while for $N < 2s$ the method is exact and corresponds to N standard compositions of the map φ_ε. This latter case means that the problem where $\varepsilon > H/(2s)$ is not sufficiently highly oscillatory and there is no need to resort to the multi-revolution approach.

We assume that the computational cost of computing $\varphi_\varepsilon^* = \varphi_{-\varepsilon}^{-1}$ is similar to that of computing φ_ε. The main application we have in mind is the time integration of highly-oscillatory problems with a single harmonic frequency $\omega = 2\pi/\varepsilon$. In the numerical examples, we consider in particular problems of the form

$$\frac{d}{dt}y(t) = \frac{1}{\varepsilon}Ay(t) + f(y(t)), \qquad 0 \le t \le T, \qquad y(0) = y_0 \in \mathbb{R}^d, \tag{10}$$

where A is a $d \times d$ constant matrix with $e^A = I$, so that e^{tA} is 1-periodic in time, and where $f : \mathbb{R}^d \to \mathbb{R}^d$ is a given nonlinear smooth function. In this situation, we shall consider φ_ε as the flow with time ε (the period of the unperturbed equation corresponding to $f(y) \equiv 0$) of equation (10). In this case, observe that the adjoint $\varphi_\varepsilon^* = \varphi_{-\varepsilon}^{-1}$ is the flow with time 1 of the above system where A is replaced by $-A$ (equivalently, the flow with time $-\varepsilon$ of (10) with ε replaced by $-\varepsilon$). It is worth stressing that MRCMs can be applied to more general highly-oscillatory problems with a single harmonic frequency. This is the case of any problem that, possibly after a change of variables, can be written into the form

$$\frac{d}{dt}z(t) = g(z(t), t/\varepsilon), \qquad 0 \le t \le T, \qquad z(0) = z_0 \in \mathbb{R}^d, \tag{11}$$

where $g(z, \tau)$ is smooth in z and continuous and 1-periodic in τ. For instance, (10) can be recast into the format (11) with $g(z, \tau) = e^{-\tau A} f(e^{\tau A} z)$ by considering the change of variables $y = e^{tA/\varepsilon} z$. In this more general context, φ_ε will be such that for arbitrary z_0, the solution $z(t)$ of (11) satisfies that $z(\varepsilon) = \varphi_\varepsilon(z_0)$.

It is well known [9, 10] that such a map φ_ε is a smooth near-identity map, and furthermore, that (3) is in this case the first order averaged equation, more precisely,

$$\Theta_0(z) = \left.\frac{d}{d\varepsilon}\varphi_\varepsilon(z)\right|_{\varepsilon=0} = \int_0^1 e^{-At} f(e^{At} z)\, dt.$$

The solution $y(t)$ of the initial value problem (10) sampled at the times $t = \varepsilon M$ will then be given by

$$y(\varepsilon M) = \varphi_\varepsilon^M(y_0),$$

and thus, for an appropriately chosen positive integer N (determined by accuracy requirements and the actual value of ε), we may use a pth order MRCM (7) to compute the approximations

$$y_m = \Psi_{N,H}(y)^m \approx \varphi_\varepsilon^{mN}(y_0) = y(t_m), \quad \text{where} \quad t_m = mH, \quad H = \varepsilon N.$$

The local error estimate (8) then leads by standard arguments to a global error estimate of the form

$$y_m - y(t_m) = \mathcal{O}(H^p), \quad \text{for} \quad t_m = mH \leq T,$$

where the constant in the $\mathcal{O}$-term depends on T but is independent of ε and H.

Typically, the maps φ_μ and φ_μ^* in (7) with $\mu = \alpha_j(N)H$, $\mu = \beta_j(N)H$ $(j = 1, \ldots, s)$ can not be computed exactly. In the context of highly oscillatory systems, and in particular, for systems of the form (10), the actual (approximate) computation of φ_μ can be carried out essentially as a black-box operation: In practice, one may use any available implementation of some numerical integrator to approximate the flow with time 1 of the ODE

$$\frac{d}{dt}u(t) = Au(t) + \mu f(u(t)). \tag{12}$$

In particular, φ_μ may be approximated by applying n steps of step-size $h = 1/n$ of an appropriate splitting method to (12), where n is chosen so as to resolve one oscillation. Let $\Phi_{\mu,h}(y)$ denote the approximation of φ_μ obtained in this way with a qth order splitting method, then the following estimate

$$\Phi_{\mu,h}(y) - \varphi_\mu(y) = \mathcal{O}(\mu^r h^q) \tag{13}$$

will be guaranteed to hold with $r = 1$.

In the most general framework, we shall assume that, if φ_μ can not be computed exactly, then it is approximated by some computable map $\Phi_{\mu,h}$ (depending on a small parameter h that controls the accuracy of the approximation) satisfying the error estimate (13) for some $r \geq 0$ and $q \geq 1$. Observe that one can expect $r \geq 1$ in the right-hand side of (13) if (as in the case of splitting methods for (12),) $\Phi_{\mu,h}$ is constructed so that $\Phi_{0,h}(y) = \varphi_0(y) = y$.

In what follows, the method (7) where the involved maps φ_μ and φ_μ^* are assumed to be computed exactly, will be referred to as *semi-discrete multi-revolution composition methods*. We next define the following fully-discrete version, in the spirit of Heterogenerous multiscale methods (HMM) (see [1, 11, 12]) which combine the application of macro-steps of length H (to advance along the solution of (10)) with the application to (12) of some integrator with micro-steps of size $h = 1/n$ (where n is chosen large enough to resolve each oscillation).

Definition 1.1 (Fully-discrete multi-revolution composition methods)**.** *Let an integer $s \geq 1$ and a family of coefficients $\alpha_j(N), \beta_j(N)$ defined for $j = 1, \ldots, s$ and $N \geq 2s$, and an approximation $\Phi_{\varepsilon,h}$ of φ_ε. For the approximation of φ_ε^N, with $\varepsilon > 0$, $N \geq 1$, we define an s-stage fully-discrete MRCM as the composition*

$$\begin{aligned} \Psi_{N,H,h}(y) &= \Phi_{\alpha_1(N)H,h} \circ \Phi^*_{\beta_1(N)H,h} \circ \cdots \circ \Phi_{\alpha_s(N)H,h} \circ \Phi^*_{\beta_s(N)H,h}(y) \quad \text{if } N \geq 2s, \\ \Psi_{N,H,h}(y) &= \Phi_{\varepsilon,h}^N \quad \text{if } N < 2s, \end{aligned} \tag{14}$$

where $H = N\varepsilon$ and $\Phi^*_{\varepsilon,h} := \Phi^{-1}_{-\varepsilon,h}$ is the adjoint map *of $\Phi_{\varepsilon,h}$.*

When solving a highly-oscillatory problem of the form (10) with standard numerical methods, stability and accuracy requirements induce a step-size restriction of the form $h \leq C\varepsilon$ which renders the computation of a reasonably accurate solution more and more costly and sometimes even untractable for small values of ε. In contrast, we propose to use the approximation $\Psi_{N,H}^m(y) \simeq \varphi_\varepsilon^M(y)$ where $\varepsilon M \leq T$ with T fixed, $m, M, N \in \mathbb{N}$ with $N\varepsilon = H$, $mN = M$. Using for $\Psi_{N,H}$ a MRCM (7) of order p, we prove that the error is $\Psi_{N,H}^m(y) - \varphi_\varepsilon^M(y) = \mathcal{O}(H^p)$ with a constant independent of m, N, ε. Given a prescribed accuracy Tol which induces a stepsize H ($H \geq \varepsilon$), setting N as the integer part of $\varepsilon^{-1}H$, this permits to approximate the solution with the accuracy Tol at a cost proportional to $m = TH^{-1} \sim Tol^{-1/p}$ which does not grow for small values of ε. Using the fully-discrete MRCM (14), the cost is scaled by the additional factor $n = h^{-1}$, again independently of ε.

The general idea of multi-revolution methods has been first considered in astronomy, where ε-perturbation of periodic systems are recurrent, and named as such since these methods approximate many *revolutions* (N periods of time) by only a few (in our approach, $2s$ compositions then accounts for $2s$ revolutions with different values of the perturbation parameter ε). A class of multi-revolution Runge-Kutta type methods has then been studied in the context of oscillatory problems of the form (10) in [3, 4, 23, 25] and [2] was the first systematic analysis of the order conditions of convergence. Closely related methods were considered in [19] and also in [5].

Actually, MRCM are *asymptotic preserving*, a notion introduced in the context of kinetic equations (see [18], and the recent works [20, 14]) and ensuring that a method is uniformly accurate for a large range of values of the parameter ε with a computational cost essentially independent of ε. This is a feature shared by the proposed classes of multi-revolution methods.

The methods introduced in this paper differ from existing other multi-revolution methods in that they are intrinsically geometric, since they solely use compositions of maps of the form φ_μ and φ_μ^{-1}, whose geometric properties are determined by equation (10). In particular, they are *symplectic* if (10) is *Hamiltonian*, *volume-preserving* if (10) is *divergence-free*, and share the same invariants which are independent of ε as the flow of (10). This is also true in the fully-discrete version (14) provided that the micro-integrator $\Phi_{\mu,h}$ used to approximate φ_μ satisfy the required geometric properties.

Deriving general order conditions for (7) requires to compare the Taylor expansions of both sides of $\Psi_{N,H}(y) \simeq \varphi_\varepsilon^N(y)$. Although conceptually easy, the task is rendered very intricate by the enormous number of terms and redundant order conditions naturally arising. Explicit conditions for standard composition methods have been obtained in a systematic way in [24] by using the formalism of B_∞-series and trees. In the situation we consider here, the map φ_ε is the flow with time ε of an ODE that depends on the parameter ε, and consequently does not obey a group law. The question of approximating $\varphi_{\varepsilon/N}^N$ by a composition of the form (7) then makes perfect sense, and this article aims at analyzing the properties and order conditions of such methods.

The paper is organized as follows. In Section 2, we derive the order conditions of the multi-revolution composition methods and perform a global error analysis of the methods. Section 3 presents several methods of orders 2 and 4, and describes how they have been obtained. Section 4 is devoted to numerical experiments aimed at giving a numerical confirmation of the orders of convergence derived in Section 2 and to show the efficiency and versatility of the newly introduced methods.

2 Convergence analysis of MRCMs

In this section, we derive general order conditions for method (7) to approach φ_ε^N. There is a complete analogy with order conditions of standard composition methods, with the exception that the right-hand side of each condition is now depending on N. Prior to addressing the general case, observe that the simplest method $\varphi_H \simeq \varphi_\varepsilon^N$ with $H = N\varepsilon$, corresponds to $s = 1$, $\alpha_1 = 1$ and $\beta_1 = 0$. As shown in the introduction, φ_H is the simplest approximation to φ_ε^N in the following sense,

$$\|\varphi_H(y) - \varphi_\varepsilon^N(y)\| \leq CH^2 \text{ for all } 0 \leq H = N\varepsilon \leq H_0.$$

Constructing high-order compositions soon becomes rather intricate, not to say undoable, unless one uses an appropriate methodology. This is precisely the object of the paper [24] which gives order conditions for standard composition methods explicitly. We will hereafter follow the presentation of [16]. The starting point of this section is the Taylor series expansion[1]

$$\varphi_\varepsilon(y) = y + \varepsilon d_1(y) + \varepsilon^2 d_2(y) + \varepsilon^3 d_3(y) + \dots \tag{15}$$

of the smooth map (1).

2.1 Preliminaries: trees and B_∞-series

In this subsection, we briefly recall the framework of B_∞-series for the study of composition methods of the form

$$\varphi_{\alpha_s\varepsilon} \circ \varphi^*_{\beta_s\varepsilon} \circ \cdots \circ \varphi_{\alpha_1\varepsilon} \circ \varphi^*_{\beta_1\varepsilon}(y) \tag{16}$$

originally developed for the numerical integration of equation (3) and yet completely relevant to the present situation. We thus define T_∞ as the set of rooted trees where each vertex bears a positive integer and we denote $\textcircled{1}, \textcircled{2}, \textcircled{3}, \dots$ the trees with one vertex. Given $\tau_1, \dots, \tau_m \in T_\infty$, we write as

$$\tau = [\tau_1, \dots, \tau_m]_j \tag{17}$$

the tree obtained by attaching the m roots of $\tau_1, \dots, \tau_m$ to a new root with label j. Incidentally, we define $i(\tau) = j$ the label beard by its root, $|\tau| = 1 + |\tau_1| + \dots + |\tau_m|$ its number of vertices, $\|\tau\| = i(\tau) + \|\tau_1\| + \dots + \|\tau_m\|$ the sum of its labels[2] and $\sigma(\tau) = \mu_1!\mu_2! \cdots \sigma(\tau_1) \cdots \sigma(\tau_m)$ its symmetry coefficient, where $\mu_1, \mu_2, \dots$ count equal trees among $\tau_1, \dots, \tau_m$. Now, the B_∞-series associated to a map $a : T_\infty \cup \{\emptyset\} \to \mathbb{R}$ is the formal series

$$B_\infty(a, \varepsilon, y) = a(\emptyset)y + \sum_{\tau \in T_\infty} \frac{\varepsilon^{\|\tau\|}}{\sigma(\tau)} a(\tau) F(\tau)(y)$$

where the so-called "elementary differentials" are maps from $\mathcal{U}$ to $\mathbb{R}^d$ defined inductively by the relations

$$\begin{aligned} F(\textcircled{j})(y) &= d_j(y), \\ F([\tau_1, \dots, \tau_m]_j)(y) &= d_j^{(m)}(y)(F(\tau_1)(y), \dots, F(\tau_m)(y)). \end{aligned}$$

[1] Notice that $d_1(y) = \Theta_0(y)$.

[2] By convention, $|\emptyset| = \|\emptyset\| = 0$.

We immediately see that the Taylor expansion (15) of φ_ε can be considered as a B_∞-series

$$\varphi_\varepsilon(y) = y + \varepsilon d_1(y) + \varepsilon^2 d_2(y) + \varepsilon^3 d_3(y) + \ldots = B_\infty(e_1, \varepsilon, y)$$

with coefficients satisfying $e_1(\tau) = 0$ for all $\tau \in T_\infty$ with $|\tau| > 1$ and $e_1(\emptyset) = 1$, $e_1(\textcircled{j}) = 1$ for all $j \in \mathbb{N}^*$. As immediate is the obtention of the coefficients of the B-series expansion of the exact solution $z(\varepsilon)$ of (3)

$$B_\infty(e_\infty, \varepsilon, y) = z(\varepsilon)$$

with coefficients $e_\infty(\tau)$ recursively defined[3] by

$$e_\infty(\emptyset) = 1, \quad e_\infty(\tau) = \frac{1}{|\tau|} e'_\infty(\tau) \text{ if } i(\tau) = 1, \text{ and } e_\infty(\tau) = 0 \text{ otherwise,} \tag{18}$$

where the prime stands for the following B-series operation: Given $a : \mathcal{T}_\infty \to \mathbb{R}$, the map a' is defined recursively by

$$a'(\textcircled{j}) = 1 \text{ and for all } \tau = [\tau_1, \ldots, \tau_m]_j \in \mathcal{T}_\infty,\ a'(\tau) = a(\tau_1) \cdots a(\tau_m).$$

We now quote the following fundamental result from [24]:

Lemma 2.1. *The following compositions are B_∞-series*

$$\begin{aligned} \varphi^*_{\beta_k \varepsilon} \circ \cdots \circ \varphi_{\alpha_1 \varepsilon} \circ \varphi^*_{\beta_1 \varepsilon}(y) &= B_\infty(b_k, \varepsilon, y) \\ \varphi_{\alpha_k \varepsilon} \circ \varphi^*_{\beta_k \varepsilon} \circ \cdots \circ \varphi_{\alpha_1 \varepsilon} \circ \varphi^*_{\beta_1 \varepsilon}(y) &= B_\infty(a_k, \varepsilon, y) \end{aligned}$$

with coefficients given recursively for all $T \in \mathcal{T}_\infty$ by $a_k(\emptyset) = b_k(\emptyset) = 1, a_0(\tau) = 0$ and

$$b_k(\tau) = a_{k-1}(\tau) - (-\beta_k)^{i(\tau)} b'_k(\tau) \text{ and } a_k(\tau) = b_k(\tau) + \alpha_k^{i(\tau)} b'_k(\tau).$$

In order to eliminate redundant order conditions, we finally fix as in [24] a total order relation $<$ on T_∞ compatible with $|\cdot|$, i.e. such that $u < v$ whenever $|u| < |v|$.

Definition 2.2. *(Hall Set). The Hall set corresponding to the order relation $<$ is the subset $H \subset T_\infty$ defined by*

(i) $\forall j \in \mathbb{N}, \textcircled{j} \in H$

(ii) $\tau \in H$ *if and only if there exist* $u, v \in H, u > v$, *such that* $\tau = u \circ v$.

Theorem 2.3. *(Murua and Sanz-Serna [24]) Consider $B(a, \varepsilon, y)$ and $B(b, \varepsilon, y)$ two B_∞-series obtained as compositions of the form* (16) *and let $p \geq 1$. The following two statements are equivalent:*

(i) $\forall \tau \in T_\infty,\ \|\tau\| \leq p,\ a(\tau) = b(\tau),$

(ii) $\forall \tau \in \mathcal{H},\ \|\tau\| \leq p,\ a(\tau) = b(\tau).$

In the usual setting of composition methods, the previous theorem immediately gives the reduced number of order conditions for order p by comparing the B_∞-series $B_\infty(a, \varepsilon, y)$ obtained from (16) and $B_\infty(e_\infty, \varepsilon, y)$. In our context, we have to compare $B_\infty(a, \varepsilon, y)$ with the B_∞-series of φ_ε^N. This is the purpose of the next section.

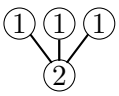
Order 1:	①	$\sum_{k=1}^{s}(\alpha_k+\beta_k)=1$
Order 2:	②	$\sum_{k=1}^{s}(\alpha_k^2-\beta_k^2)=N^{-1}$
Order 3:	③	$\sum_{k=1}^{s}(\alpha_k^3+\beta_k^3)=N^{-2}$
		$\sum_{k=1}^{s}(\alpha_k^2-\beta_k^2)\sum_{\ell=1}^{k}{}'(\alpha_\ell+\beta_\ell)=\frac{N^{-1}-N^{-2}}{2}$
Order 4:	④	$\sum_{k=1}^{s}(\alpha_k^4-\beta_k^4)=N^{-3}$
		$\sum_{k=1}^{s}(\alpha_k^3+\beta_k^3)\sum_{\ell=1}^{k}{}'(\alpha_\ell+\beta_\ell)=\frac{N^{-2}-N^{-3}}{2}$
		$\sum_{k=1}^{s}(\alpha_k^2-\beta_k^2)\Big(\sum_{\ell=1}^{k}{}'(\alpha_\ell+\beta_\ell)\Big)^2=\frac{N^{-1}(1-N^{-1})(2-N^{-1})}{6}$
Order 5:	⑤	$\sum_{k=1}^{s}(\alpha_k^5+\beta_k^5)=N^{-4}$
		$\sum_{k=1}^{s}(\alpha_k^4-\beta_k^4)\sum_{\ell=1}^{k}{}'(\alpha_\ell+\beta_\ell)=\frac{N^{-3}-N^{-4}}{2}$
		$\sum_{k=1}^{s}(\alpha_k^3+\beta_k^3)\sum_{\ell=1}^{k}{}'(\alpha_\ell^2-\beta_\ell^2)=\frac{N^{-3}-N^{-4}}{2}$
		$\sum_{k=1}^{s}(\alpha_k^2-\beta_k^2)\Big(\sum_{\ell=1}^{k}{}'(\alpha_\ell+\beta_\ell)\Big)\Big(\sum_{\ell=1}^{k}{}'(\alpha_\ell^2-\beta_\ell^2)\Big)=\frac{N^{-2}(1-N^{-1})(2-N^{-1})}{6}$
		$\sum_{k=1}^{s}(\alpha_k^2-\beta_k^2)\Big(\sum_{\ell=1}^{k}{}'(\alpha_\ell+\beta_\ell)\Big)^3=\frac{N^{-1}(1-N^{-1})^2}{4}$
		$\sum_{k=1}^{s}(\alpha_k^3+\beta_k^3)\Big(\sum_{\ell=1}^{k}{}'(\alpha_\ell+\beta_\ell)\Big)^2=\frac{N^{-2}(1-N^{-1})(2-N^{-1})}{6}$

Table 1: Fifth-order conditions for MRCMs. The prime attached to a summation symbol indicates that the sum of α_ℓ^j is only from 1 to $k-1$ while the sum of β_ℓ^j remains for 1 to k

.

2.2 Semi-discrete error analysis

Observe that by taking $\alpha_i = N^{-1}, \beta_i = 0$, $i = 1, \ldots, N$ in (16) Lemma 2.1 immediately yields that the composition $\varphi^N_{\varepsilon/N}(y)$ is again a B_∞-series

$$B_\infty(e_N, \varepsilon, y) = \varphi^N_{\varepsilon/N}(y). \tag{19}$$

[3]Notice that $e_\infty(\tau) = 0$ if at least one of its labels is different from 1.

Its coefficients $e_N(\tau)$ can be computed by using the following lemma.

Lemma 2.4. *For all $N \in \mathbb{N}^*$, the coefficients $e_N(\tau)$ of the B_∞-series in (19) satisfy*

$$\forall j \in \mathbb{N}^*, \quad e_N(\textcircled{j}) = N^{1-j},$$
$$\forall \tau = [\tau_1, \dots, \tau_n]_j \in \mathcal{T}_\infty, \quad N^{\|\tau\|} e_N(\tau) = \sum_{k=1}^{N-1} k^{\|\tau_1\|+\dots+\|\tau_n\|} e'_k(\tau).$$

Proof. With $\alpha_i = 1$ and $\beta_i = 0$, $i = 1, \dots, N$, Lemma 2.1 gives $b_k(\tau) = a_{k-1}(\tau)$ and thus

$$a_N(\tau) = \sum_{k=1}^{N} a'_{k-1}(\tau) = \sum_{k=1}^{N-1} a'_k(\tau).$$

Using $B_\infty(e_N, \varepsilon, y) = B_\infty(a_N, \varepsilon/N, y)$ yields $a_N(\tau) = N^{\|\tau\|} e_N(\tau)$ and allows to conclude. □

We obtain for instance $e_N(\textcircled{1}) = 1$, $e_N(\textcircled{2}) = N^{-1}$ and

$$e_N\Big(\begin{smallmatrix}\textcircled{1}\\ |\\ \textcircled{1}\end{smallmatrix}\Big) = \frac{1-N^{-1}}{2}, \quad e_N\Big(\begin{smallmatrix}\textcircled{1}\\ |\\ \textcircled{2}\end{smallmatrix}\Big) = \frac{N^{-1}(1-N^{-1})}{2}, \quad e_N\Big(\begin{smallmatrix}\textcircled{1} & \textcircled{1}\\ & \diagdown\diagup\\ & \textcircled{2}\end{smallmatrix}\Big) = \frac{N^{-1}(1-N^{-1})(2-N^{-1})}{6}.$$

Now, recalling that the map φ_ε can be interpreted as a consistent integrator for equation (3), $\varphi^N_{\varepsilon/N}(y_0)$ converges to its solution $z(\varepsilon)$ for $N \to \infty$ and it is thus expected that the coefficients $e_N(\tau)$ converge to $e_\infty(\tau)$ as $N \to \infty$. This is shown in next proposition.

Consider now the B_∞-series $B_\infty(a, \varepsilon, y)$ associated to a semi-discrete MRCM of the form (7) with $H = \varepsilon$. Writing the order conditions now boils down to comparing the coefficients of $B_\infty(a, \varepsilon, y)$ and $B_\infty(e_N, \varepsilon, y)$ and estimating the remainder term. Next lemma provides estimates of the derivatives of φ^N_ε w.r.t. ε. In order to alleviate the presentation, let us denote for $\rho > 0$, $B_\rho(y_0) = \{y \in \mathbb{R}^d; \|y - y_0\| \le \rho\}$, and for a given function $y \mapsto k(y)$ defined on $B_\rho(y_0)$,

$$\|k\|_\rho := \sup_{y \in B_\rho(y_0)} \|k(y)\| \quad \text{and} \quad \|\partial^n_y k\|_\rho := \sup_{\substack{y \in B_\rho(y_0),\\ \|v_i\| = 1,\, i = 1, \dots, n}} \|\partial^n_y k_\varepsilon(y)\,(v_1, \dots, v_n)\,\|.$$

Note that if $(y, \varepsilon) \mapsto \Theta_\varepsilon(y)$ in (1) is of class C^{p+1} with respect to (y, ε) on the compact set $B_\rho(y_0) \times [-\varepsilon_0, \varepsilon_0]$, then there exist positive constants K and L such that, for all $|\varepsilon| \le \varepsilon_0$

$$\|\partial_y \varphi_\varepsilon\|_\rho \le 1 + \varepsilon L, \qquad \forall k = 2, \dots, p+1, \quad \|\partial^k_y \varphi_\varepsilon\|_\rho \le \varepsilon L,$$
$$\forall 0 \le k + l \le p+1, \quad \|\partial^k_y \partial^l_\varepsilon \varphi_\varepsilon\|_\rho \le K.$$

Lemma 2.5. *Assume that $(y, \varepsilon) \mapsto \Theta_\varepsilon(y)$ is defined and of class C^{p+1} with respect to (y, ε) on $B_{2R}(y_0) \times [-\varepsilon_0, \varepsilon_0]$ for a given $R > 0$ and a given $\varepsilon_0 > 0$. Then, there exists a constant H_0 such that for all ε and $N \ge 1$ with $H = N\varepsilon \le H_0$,*

$$\left\|\partial^{p+1}_\varepsilon \varphi^N_\varepsilon\right\|_R \le C N^{p+1}, \qquad \left\|\partial^{p+1}_H \varphi^N_{H/N}\right\|_R \le C, \tag{20}$$

where C is independent of N and ε.

Proof. For $\tilde{y}_0 \in B_R(y_0)$ and denoting $M := \sup_{|\varepsilon|\leq\varepsilon_0} \|\Theta_\varepsilon\|_{2R}$, we have

$$\|\varphi_\varepsilon^N(\tilde{y}_0) - y_0\| \leq \sum_{k=1}^{N} \|\varphi_\varepsilon^k(\tilde{y}_0) - \varphi_\varepsilon^{k-1}(\tilde{y}_0)\| + \|\tilde{y}_0 - y_0\| \leq R + N\varepsilon M$$

as long as the iterates $\varphi_\varepsilon^i(\tilde{y}_0)$ and $\varphi_\varepsilon^i(y_0)$ remain in $B_{2R}(y_0)$ for $0 \leq i \leq N$. Hence, if $N\varepsilon \leq H_0 := \min(R/M, \varepsilon_0)$ then $\|\varphi_\varepsilon^k\|_R \leq 2R$ for all $k = 0, \ldots, N$. Under this assumption, we now wish to prove by induction on n, that

$$\forall n = 1, \ldots, p+1, \quad \left\|\partial_\varepsilon^n \varphi_\varepsilon^N\right\|_R \leq C_n N^n \tag{21}$$

for some constants C_n independent of N, ε. Now, given a smooth function $g : B_{2R}(y_0) \to \mathbb{R}^d$ of class C^{p+1}, Faà di Bruno's formula reads

$$\partial_\varepsilon^k (g \circ \varphi_\varepsilon^N) = \sum_{\mathbf{m}\in\mathbb{N}^k,\, \sigma(\mathbf{m})=k} B_{\mathbf{m}}\, g^{(|\mathbf{m}|)} \circ \varphi_\varepsilon^N \Big((\partial_\varepsilon^1 \varphi_\varepsilon^N)^{m_1}, \ldots, (\partial_\varepsilon^k \varphi_\varepsilon^N)^{m_k} \Big)$$

where the sum is over all multi-indices $\mathbf{m} = (m_1, \ldots, m_k)$ of $\mathbb{N}^k$ such that $k = \sigma(\mathbf{m}) := \sum_{j=1}^k j m_j$ and where $|\mathbf{m}|$ denotes $m_1 + \ldots + m_k$ and

$$B_{\mathbf{m}} = \frac{k!}{m_1! 1!^{m_1} \cdots m_k! k!^{m_k}}.$$

We now use the differentiation formula

$$\partial_\varepsilon^n (\varphi_\varepsilon \circ \varphi_\varepsilon^N) = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\, \partial_\varepsilon^k (\partial_\mu^{(n-k)} \varphi_\mu \circ \varphi_\varepsilon^N)\Big|_{\mu=\varepsilon}$$

and take $g = \partial_\mu^{(n-k)} \varphi_\mu \Big|_{\mu=\varepsilon}$ in Faà di Bruno's formula. This yields

$$\partial_\varepsilon^n (\varphi_\varepsilon^{N+1}) = \sum_{\substack{0\,\leq\, k\,\leq\, n,\\ \mathbf{m}\,\in\,\mathbb{N}^k,\,\sigma(\mathbf{m})\,=\,k}} \frac{n!}{k!(n-k)!}\, B_{\mathbf{m}}\, \Big(\partial_y^{|\mathbf{m}|} \partial_\varepsilon^{n-k} \varphi_\varepsilon \Big) \circ \varphi_\varepsilon^N \Big((\partial_\varepsilon^1 \varphi_\varepsilon^N)^{m_1}, \ldots, (\partial_\varepsilon^n \varphi_\varepsilon^N)^{m_n} \Big)$$

Hence, using the induction assumption, we get the estimates

$$\begin{aligned}
\left\|\partial_\varepsilon^n \varphi_\varepsilon^{N+1}\right\|_R &\leq \|\partial_\varepsilon^n \varphi_\varepsilon\|_{2R} + \sum_{\substack{1\,\leq\, k\,\leq\, n-1,\\ \mathbf{m}\,\in\,\mathbb{N}^k,\,\sigma(\mathbf{m})\,=\,k}} \frac{n!}{k!(n-k)!}\, B_{\mathbf{m}}\, \|\partial_y^{|\mathbf{m}|} \partial_\varepsilon^{n-k} \varphi_\varepsilon\|_{2R} \prod_{j=1}^{k} \|\partial_\varepsilon^j \varphi_\varepsilon^N\|_R^{m_j} \\
&\quad + \sum_{\mathbf{m}\,\in\,\mathbb{N}^n,\,\sigma(\mathbf{m})\,=\,n,\,m_n\,=\,0} B_{\mathbf{m}}\, \|\partial_y^{|\mathbf{m}|} \varphi_\varepsilon\|_{2R} \prod_{j=1}^{n} \|\partial_\varepsilon^j \varphi_\varepsilon^N\|_R^{m_j} + \|\partial_y \varphi_\varepsilon\|_{2R} \|\partial_\varepsilon^n \varphi_\varepsilon^N\|_R \\
&\leq K + K\tilde{C}_n \sum_{k=1}^{n-1} N^k + \varepsilon n \hat{C}_n L N^n + (1+\varepsilon L) \|\partial_\varepsilon^n \varphi_\varepsilon^N\|_R \\
&\leq n K \tilde{C}_n (N+1)^{n-1} + n \hat{C}_n L H_0 N^{n-1} + (1+\varepsilon L) \|\partial_\varepsilon^n \varphi_\varepsilon^N\|_R \\
&\leq \bar{C}_n (N+1)^{n-1} + (1+\varepsilon L) \|\partial_\varepsilon^n \varphi_\varepsilon^N\|_R
\end{aligned}$$

where the constants $\tilde{C}_n$ and $\hat{C}_n$ are defined as

$$\tilde{C}_n = \max_{k=1,\dots,n-1} \sum_{\substack{\mathbf{m}\in\mathbb{N}^k \\ \sigma(\mathbf{m})=k}} B_{\mathbf{m}} \prod_{j=1}^{k} C_j^{m_j} \text{ and } \hat{C}_n = \sum_{\substack{\mathbf{m}\in\mathbb{N}^{n-1} \\ \sigma(\mathbf{m})=n}} B_{\mathbf{m}} \prod_{j=1}^{n-1} C_j^{m_j}$$

and $\bar{C}_n = n\max(K, K\tilde{C}_n, \hat{C}_n L H_0)$. Finally, using a standard discrete Gronwall lemma and $H = N\varepsilon \le H_0$ yields

$$\|\partial_\varepsilon^n(\varphi_\varepsilon^N)\|_R \le \bar{C}_n \sum_{k=1}^{N} (1+\varepsilon L)^{N-k} k^{n-1} \le \bar{C}_n N^n e^{L\varepsilon N} \le \bar{C}_n N^n e^{LH_0}$$

which allows to conclude the proof of the first estimate in (20) by choosing $C_n = \bar{C}_n e^{LH_0}$ in (21). The second estimate is straightforwardly obtained through a change of variables. □

We may now state the main result for the local error of the semi-discrete MRCM (7).

Theorem 2.6. *Consider a semi-discrete MRCM* (7)-(9) *and assume further that its coefficients* $\alpha_i(N)$, $\beta_i(N)$, $i = 1,\dots,s$ *are bounded with respect to* N *for all* $N \ge 2s$ *and satisfy*

$$a(\tau) = e_N(\tau), \quad \text{for all } \tau \in \mathcal{H} \text{ with } \|\tau\| \le p, \tag{22}$$

for a given order $p \ge 1$. *Then, there exist* $C, H_0 > 0$ *such that for all* $H \le H_0$,

$$\|\Psi_{N,H} - \varphi_\varepsilon^N\|_R \le C H^{p+1}$$

where $H = N\varepsilon$ *and the constants* C, H_0 *are independent of* N, ε.

Proof. Consider the two B_∞-series $B_\infty(a, H, y)$ and $B_\infty(e_N, H, y)$ associated respectively to the semi-discrete MRCM (7) and to $\varphi_{H/N}^N(y)$ in (19). It follows from Theorem 2.3 that these B_∞-series formally coincide up to order H^p. A Taylor expansion of $\Psi_H(y) - \varphi_{H/N}^N(y)$ with integral remainder thus leads to

$$\Psi_{N,H}(y) - \varphi_{H/N}^N(y) = \int_0^H \frac{1}{p!}(H-s)^p \frac{\partial^{p+1}\Psi_s}{\partial s^{p+1}}(y)ds - \int_0^H \frac{1}{p!}(H-s)^p \frac{\partial^{p+1}\varphi_{s/N}^N}{\partial s^{p+1}}(y)\,ds.$$

The derivative $\frac{\partial^{p+1}\varphi_{s/N}^N}{\partial s^{p+1}}$ is bounded by Lemma 2.5. Given that coefficients α_j, β_j are uniformly bounded with respect to N, $\frac{\partial^{p+1}\Psi_s}{\partial s^{p+1}}$ is bounded as well. We conclude using $\varphi_{H/N}^N(y) = \varphi_\varepsilon^N(y)$. □

We report in Table 1 order conditions up to order 5 as derived above. Note that Lemma 2.4 implies (by induction) that the value of $N^{\|\tau\|}e_N(\tau)$ is independent of the labels of the nodes of a given tree $\tau \in \mathcal{T}_\infty$. This explains why similar right-hand sides $e_N(\tau)$ are obtained for trees where only labels differ.

Remark 2.7. *For* $N \to \infty$, *it can be shown that the coefficients of the* B_∞*-series* (19) *satisfy* $e_N(\tau) \to e_\infty(\tau)$. *This means that the order conditions* (22) *reduce for* $N \to \infty$ *to the classical order conditions of standard composition methods* (16) *for the approximation of the flow of* (3), *the averaged problem as* $\varepsilon \to 0$.

Remark 2.8. *It can be shown (see [26] and the recent work [9]) that the highly-oscillatory solution of* (10) *is, at times which are integer multiples of the oscillatory period* ε*, asymptotically close to the solution of non-stiff averaged ODEs. Precisely, for all* $p \geq 1$*, there exists a smooth ODE of the form*

$$\frac{dz}{dt} = \Theta_0(z) + \varepsilon\, G_1(z) + \cdots + \varepsilon^{p-1} G_{p-1}(z), \quad z(0) = y_0 \tag{23}$$

whose solution $z(t)$ *satisfies* $z(M\varepsilon) - y(M\varepsilon) = \mathcal{O}(\varepsilon^p)$ *if* $M\varepsilon \leq T$*, for all integer values of* M*, where the constant in* $\mathcal{O}$ *is independent of* ε*. A similar statement can be made in the general case of an arbitrary smooth near-identity map, where* φ_ε *can be interpreted as a one-step integrator for the ODE (3), and (23) is the modified equation of (3) associated to* φ_ε *considered in backward error analysis of one step integrators [16, Chap. IX]. In addition, applying a pth order MRCM* (7) *to* (10) *(that is, (8) holds for integer values of* N*), it can be proved that for* $0 < \varepsilon \leq H$ *with* $H \to 0$*,*

$$\Psi_{H/\varepsilon,H}(y_0) - z(H) = \mathcal{O}(H^{p+1}).$$

Note that in the above convergence estimate, H/ε *in not necessarily an integer, which means that* $\Psi_{N,H}$ *in (7) also makes sense for non-integer values of* N*.*

2.3 Fully-discrete error analysis

In this subsection, we derive convergence estimates for fully-discrete MRCMs (14). We highlight once again that this is essential in view of applications because the exact computation of the map φ_ε is not available in general and has to be approximated by a map $\Phi_{h,\varepsilon}$.

Theorem 2.9. *Assume that the hypotheses of Theorem 2.6 are fulfilled. Consider a fully-discrete MRCM* (14) *where the basic map* $\Phi_{h,\varepsilon}$ *is assumed to satisfy the accuracy estimate* (13) *for given* q *and* r*. Then there exist* $C, H_0, h_0 > 0$ *such that*

$$\|\Psi_{N,H,h} - \varphi_\varepsilon^N\|_R \leq C(H^{p+1} + H^r h^q)$$

for all $H = N\varepsilon \leq H_0$*,* $h \leq h_0$ *and the constants* C, H_0, h_0 *re independent of* N, ε, h*.*

As a consequence of Theorem 2.9, by standard arguments in the convergence analysis of one-step integrators, one gets a global error estimate for the numerical approximations $y_m = \Psi_{N,H,h}(y_{m-1})$ of problem (10) of the form

$$y_m - y(mH) = \mathcal{O}(H^p + H^{r-1}h^q) \quad \text{for} \quad m\,H \leq T.$$

For the proof of Theorem 2.9, we recall the following classical discrete Gronwall estimate.

Lemma 2.10. *Let* $(\phi_j, \psi_j), j = 1, \ldots, k$*, be* k *couples of maps satisfying for* $\rho, \nu > 0$

$$\|\phi_j(y) - \psi_j(y)\| \leq \rho, \qquad \|\phi_j(y_1) - \phi_j(y_2)\| \leq (1+\nu)\|y_1 - y_2\|,$$

for all $j = 1, \ldots, k$ *and all* y, y_1, y_2*. Then,*

$$\|\phi_k \circ \cdots \circ \phi_1(y) - \psi_k \circ \cdots \circ \psi_1(y)\| \leq e^{\nu k} k\rho.$$

Proof. Let $a_j = \phi_k \circ \cdots \circ \phi_{k-j+1}$, $b_j = \psi_j \circ \cdots \circ \psi_1$. We have

$$\begin{aligned} a_k(y) - b_k(y) &= \sum_{j=0}^{k-1} a_{k-j-1} \circ \phi_{j+1} \circ b_j(y) - a_{k-j-1} \circ \psi_j \circ b_j(y) \\ \|a_k(y) - b_k(y)\| &\leq \sum_{j=0}^{k-1} (1+\nu)^{k-j-1} \|\phi_{j+1} \circ b_j(y) - \psi_{j+1} \circ b_j(y)\| \leq k e^{\nu k} \rho \end{aligned}$$

where we used the estimate $\sum_{j=0}^{k-1}(1+\nu)^{k-j-1} \leq \sum_{j=0}^{k-1} e^{\frac{j}{k}\nu k} \leq k \int_0^1 e^{\nu k t} dt \leq k e^{\nu k}$. □

Proof of Theorem 2.9. We use the estimate

$$\|\Psi_{N,H,h} - \varphi_\varepsilon^N\|_R \leq \|\Psi_{N,H,h} - \Psi_{N,H}\|_R + \|\Psi_{N,H} - \varphi_\varepsilon^N\|_R$$

From Theorem 2.6, we have $\|\Psi_{N,H} - \varphi_\varepsilon^N\|_R \leq C H^{p+1}$. The next estimate

$$\|\Psi_{N,H,h} - \Psi_{N,H}\|_R \leq C H^r h^q$$

is a consequence of Lemma 2.10 with $k = 2s$, $\rho = C H^r h^q$ (using (13) with ε replaced by $\alpha_j(N)H$ and $\beta_j(N)H$), $\nu = \mathcal{O}(\varepsilon)$ being a Lipsitz constant for the near-identity map φ_ε, and $\phi_{2j-1} = \varphi_{\alpha_j(N)H}$, $\phi_{2j} = \varphi^*_{\beta_j(N)H}$, $\psi_{2j-1} = \Phi_{h,\alpha_j(N)H}$, $\psi_{2j} = \Phi^*_{h,\beta_j(N)H}$. □

2.4 Application to second order highly oscillatory systems

The multi-revolution approach can be applied to second order highly oscillatory systems of the form

$$q_1''(t) = g_1(q_1(t), q_2(t), q_1'(t)), \quad q_2''(t) = \frac{1}{\varepsilon^2}\Omega^2 q_2(t) + g_2(q_1(t), q_2(t), q_1'(t)), \quad 0 \leq t \leq T, \quad (24)$$

where we distinguish q_1 and q_2, respectively the smooth and highly oscillatory components. We denote $y_0 = (q_1(0), q_2(0), q_1'(0), q_2'(0))$ the initial condition, Ω is a constant positive-definite symmetric matrix with eigenvalues in $2\pi\mathbb{N}^*$ and g_1, g_2 are given smooth vector fields which do not depend on the highly oscillatory velocity $q_2'(t)$. Such a form often appears for highly oscillatory mechanical systems and an example is presented in Section 4.1. The main observation is that system (24) can be cast in the form (10) by setting $y = (q_1, q_2, q_1', q_2')^T$ and

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \eta I \\ 0 & 0 & 0 & 0 \\ 0 & -\eta^{-1}\Omega & 0 & 0 \end{pmatrix}, \qquad f(y) = \begin{pmatrix} q_1' \\ 0 \\ g_1(q_1, q_2, q_1') \\ g_2(q_1, q_2, q_1') \end{pmatrix} \quad (25)$$

where $\eta = \varepsilon$ is a fixed parameter. Then, the MRCM yields an uniformly accurate integrator for (24) as stated in the following theorem for semi-discrete MRCMs. We highlight that a convergence estimate for fully-discrete MRCMs could be derived analogously (see Theorem 2.9).

Theorem 2.11. *Consider $y = (q_1, q_2, q_1', q_2')$ the exact solution of (24) where we assume that the vector fields g_1, g_2 have the regularity C^{p+1}. Consider φ_ε the flow map with time ε of (10)-(25) where $\eta = \varepsilon$ is fixed, and $\Psi_{N,H}$ a MRCM (7) satisfying the hypotheses of Theorem 2.6. Then, there exist $C, H_0 > 0$ such that for all $H \leq H_0$, $N \geq 2s$, and $mH \leq T$,*

$$\|\Psi_{N,H}^m(y_0) - y(mH)\| \leq CH^p,$$

where $H = N\varepsilon$ and the constants C, H_0 are independent of N, ε, m, η.

Proof. Using the variation of constant formula $y(t) = y_0 + \varepsilon \int_0^t e^{(t-s)A} f(e^{sA} y(s))ds$ and the identity

$$e^{sA} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \cos(s\Omega) & 0 & \eta\Omega^{-1}\sin(s\Omega) \\ 0 & 0 & 0 & 0 \\ 0 & -\eta^{-1}\Omega\sin(s\Omega) & 0 & \cos(s\Omega) \end{pmatrix},$$

we obtain that the near-identity map φ_ε (which depends on η), is smooth with derivatives with respect to ε bounded independently of $0 < \eta \leq \eta_0$, thanks to the fact that g_1, g_2 in (24) are independent of the highly oscillatory velocities q_2' for which a factor η^{-1} appears in e^{sA}. We may thus apply Theorem 2.6 to deduce the local error estimate $\|\Psi_{N,H}(y_0) - y(H)\| \leq CH^{p+1}$. The global error estimate follows from standard arguments. □

3 Effective construction of MRCMs

The simplest method of order 1 is obtained simply for $s = 1, \alpha_1 = 1, \beta_1 = 0$ in (16),

$$\varphi_H(y) = \varphi_\varepsilon^N(y) + \mathcal{O}(H^2).$$

For order 2, there exist a unique solution with $s = 1$, given by $\alpha_1 = (1 + N^{-1})/2$ and $\beta_1 = (1 - N^{-1})/2$,

$$\varphi_{\alpha_1 H} \circ \varphi_{-\beta_1 H}^{-1}(y) = \varphi_{(H+\varepsilon)/2} \circ \varphi_{-(H-\varepsilon)/2}^{-1}(y) = \varphi_\varepsilon^N(y) + \mathcal{O}(H^3).$$

For order 3, there do not exist real solutions with $s = 2$. We directly consider order 4, for which there are 7 order conditions to be satisfied. It turns out that there exists a family of solutions with $s = 3$, i.e. with only 6 free parameters α_j, β_j. We consider the following solution for $N = \infty$ given by with

$$\alpha_1 = \beta_1 = \alpha_3 = \beta_3 = \frac{1}{4 - 2 \cdot 2^{1/3}}, \quad \alpha_2 = \beta_2 = \frac{1}{2} - 2\alpha_1.$$

The idea is then to set $\delta = 1/N$ and to search for continuous function $\alpha_j(\delta^{-1}), \beta_j(\delta^{-1})$ defined for $\delta \in [0, (2s)^{-1}]$ and which coincide with the above coefficients for $\delta = 0$. This calculation is made by a continuation method.

However, it is known that for standard composition methods ($N = \infty$), the composition methods with minimal number of compositions are not the most efficient in general. We thus increment the parameter s and construct a family of MRCMs of order $p = 4$ with $s = 4$ where we choose to minimize the sum of the squares of the coefficients. This yields the following optimization problem with constraints: find $\delta \mapsto (\alpha_i(\delta^{-1}), \beta_i(\delta^{-1})), i = 1, \ldots, s$ minimizing

$\sum_{k=1}^{s}(\alpha_k(\delta^{-1})^2 + \beta_k(\delta^{-1})^2)$ and fulfilling the order conditions up to order p. This is done using a standard optimization package. For a practical implementation, we consider a set of $K = 33$ Chebyshev points $\delta_k, k = 1, \ldots, K$ sampling the interval $[0, (2s)^{-1}]$ and for which we compute the corresponding coefficients $\alpha_k(\delta_i^{-1}), \beta_k(\delta_i^{-1}), k = 1, \ldots, k$. This calculation is made once for all and stored. We then use Chebyshev interpolation to recover the coefficients α_i, β_i for any value of $\delta = 1/N \in [0, (2s)^{-1}]$. The number K of sample points has been chosen to guarantee that the Chebyshev interpolation error is smaller than the machine precision.

Remark 3.1. *Notice that multi-revolution composition methods with complex coefficients can also be considered (see [6, 17] in the context of standard composition methods). For instance, the fourth order conditions to achieve order* 3 *for a multi-revolution composition method have a complex solution for* $s = 2$*, given for all* $N \geq 2$ *by:*

$$\alpha_1(N) = \overline{\alpha_2(N)} = \frac{1}{4} + \frac{1}{2N} + i\frac{\sqrt{3 - 12/N^2}}{12}, \qquad \beta_1(N) = \overline{\beta_2(N)} = \frac{1}{4} - \frac{1}{2N} + i\frac{\sqrt{3 - 12/N^2}}{12}.$$

4 Numerical experiments

The aim of this part is to obtain a numerical confirmation of the orders of convergence given above and to demonstrate the efficiency of MRCMs. The first problem, which is a modification of the Fermi-Pasta-Ulam problem [13], is directly of the form (10) and serves classically in the literature as a test problem to measure the error behavior of the various methods for integrating single-frequency highly oscillatory systems. The second test problem is borrowed from the PDE literature and requires to be discretized with a spectral method: we aim with this example at illustrating the qualitative properties of MRCMs.

4.1 A Fermi-Pasta-Ulam like problem

In this subsection, we consider a problem taken from [16], which is a single-frequency modification of the Fermi-Pasta-Ulam problem often used to test methods for highly-oscillatory problem. Its Hamiltonian function is given by

$$E_\eta(p, q) = \frac{1}{2}\sum_{k=1}^{3}(p_{1,k}^2 + p_{2,k}^2) + \frac{1}{2\eta^2}\sum_{k=1}^{3} q_{2,k}^2 + V(q), \tag{26}$$

where $\eta > 0$ is a fixed small parameter and considering the quartic interaction potential

$$V(q) = \frac{1}{4}((q_{1,1} - q_{2,1})^4 + (q_{1,2} - q_{2,2} - q_{1,1} - q_{2,1})^4 + (q_{1,3} - q_{2,3} - q_{1,2} - q_{2,2})^4 + (q_{1,3} + q_{2,3})^4).$$

Using the time transformation $\hat{t} = t/(2\pi)$, this problem can be cast in the form (24) with $q_j = (q_{j,1}, q_{j,2}, q_{j,3})^T$, $j = 1, 2$, $\Omega = 2\pi I$ is the 3×3 identity matrix and $g_j(q_1, q_2, q_1') = -2\pi\nabla_{q_j}V(q)$ is the gradient of V with respect to q_j, $j = 1, 2$ (multiplied by -2π). Following the methodology in Section 2.4, we have that φ_μ is the flow with time 1 associated to the Hamiltonian function $2\pi\mu H_{\mu,\eta}(p, q)$ with η fixed and

$$H_{\mu,\eta}(p, q) = \frac{1}{\mu}\sum_{k=1}^{3}\Big(\frac{\eta}{2}p_{2,k}^2 + \frac{1}{2\eta}q_{2,k}^2\Big) + \Big(\frac{1}{2}\sum_{k=1}^{3} p_{1,k}^2 + V(q)\Big),$$

so that a convenient choice of $\Phi_{\mu,h}$ may be the composition of n steps of step-size $h = 1/n$ of a Strang splitting method applied to the splitting, as written above, into fast and slow contributions of the Hamiltonian (so that each period of the fast term are covered with n steps of the splitting method). In all the numerical experiments we present for that example, we have considered the second order Strang splitting method iterated n times with constant stepsize $h = 1/n$ for the definition of the basic map $\Phi_{\varepsilon,h}$ in the fully-discrete MRCMs (14).

We have integrated the problem for different values of η with initial conditions

$$q(0) = (1, 0, 0, \eta, 0, 0)^T, \quad p(0) = (1, 0, 0, 1, 0, 0)^T \tag{27}$$

with the fully-discrete versions of the MRCMs obtained in Section (3) over η^{-2} periods of the stiff springs (that is, for a time interval of length $T = 2\pi\eta^{-1}$). We compute the errors in positions and momenta by comparing the results with a "reference solution" computed very accurately using the Deuflhard method (see e.g. [16]) with a very small constant step size. In Figure 1, the global errors at time $t = 2\pi$ for the components $q_{1,1}, q_{1,2}, p_{1,1}, p_{1,2}$ versus the number of evaluations of the map φ_μ (computed with a high accuracy) are displayed for the

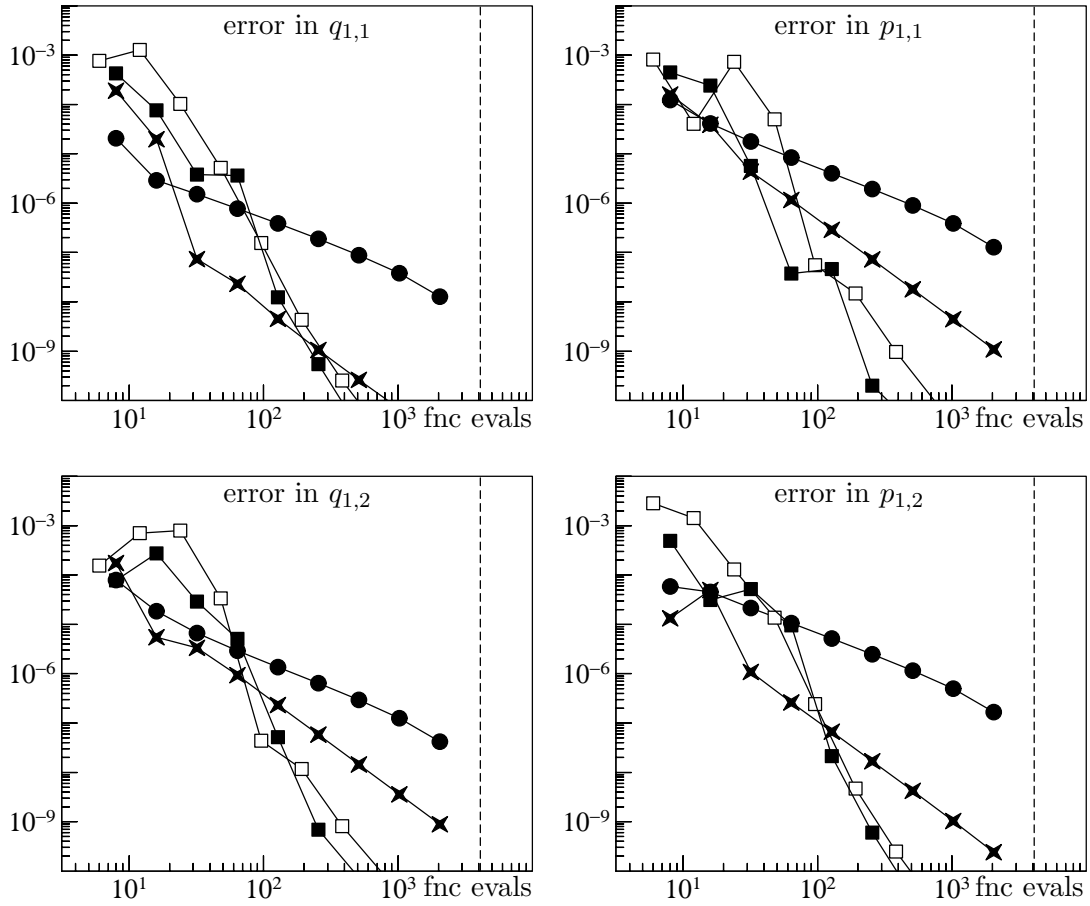


Figure 1: Problem (26) with $\eta = 2^{-12}$ and initial conditions (27). Errors of multi-revolution composition methods at time $t = 2\pi$ as functions of the number of evaluations of φ_μ (for many values of the parameter N). Methods of orders 1 (circles), 2 (stars), 4 ($s = 3$ with white squares), 4 ($s = 4$ with black squares).

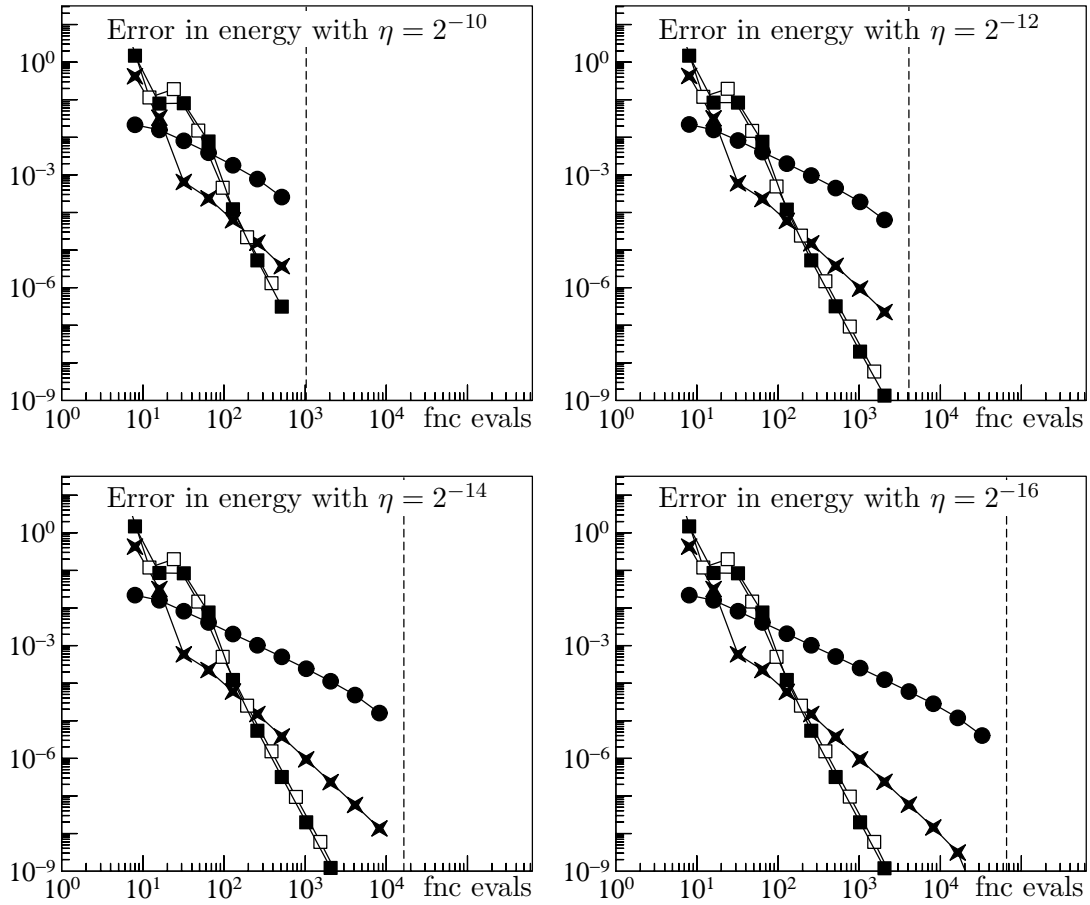


Figure 2: Problem (26) for different values of η and initial conditions (27). Errors in energy at time $t = 2\pi$ versus number of evaluations of φ_μ. Methods of orders 1 (circles), 2 (stars), 4 ($s = 3$ with white squares), 4 ($s = 4$ with black squares).

case $\eta = 2^{-12}$. We consider the semi-discrete MRCMs of orders $1, 2, 4$ for many values of the parameter N. We observe the expected lines of slope 1,2,4. Notice that the MRCM of order 4 with $s = 4$ (black squares) has a better accuracy compared to the one with minimum value $s = 3$ (white squares), as predicted in Section 3. The vertical dotted lines indicate the cost "fnc evals" $= N$ of the naive computation φ_η^N, for which the computational advantage of MRCMs vanishes.

Analogously, in Figure 2, the error in energy is displayed for the cases $\eta = 2^{-j}$ for $j = 10, 12, 14, 16$. This numerical experiment illustrates the uniform accuracy and robustness with respect to the oscillatory parameter of MRCMs.

In Figure 3, the error in energy of the approximation obtained with the fully-discrete MRCM of order 4 ($s = 4$) versus the number of evaluations of $\Phi_{\varepsilon,h}$ is displayed for different values of the micro-step $h = 1/n$ where n is the number of steps of the Strang splitting for evaluating $\Phi_{\varepsilon,h}$. Here, the vertical dotted lines in all figures corresponds to $N = \eta^{-1}$ evaluations per macro-step which corresponds to effectively computing φ_ε^N (or $\Phi_{\mu,h}^N$) instead of applying an s-stage MRCM requiring $2s$ evaluations of the basic map $\Phi_{\mu,h} \approx \varphi_\mu$. This experiment illustrates that simultaneous refinements of the macro and micro stepsizes H, h

is needed for fully-discrete MRCMs to converge, as predicted by Theorem 2.9.

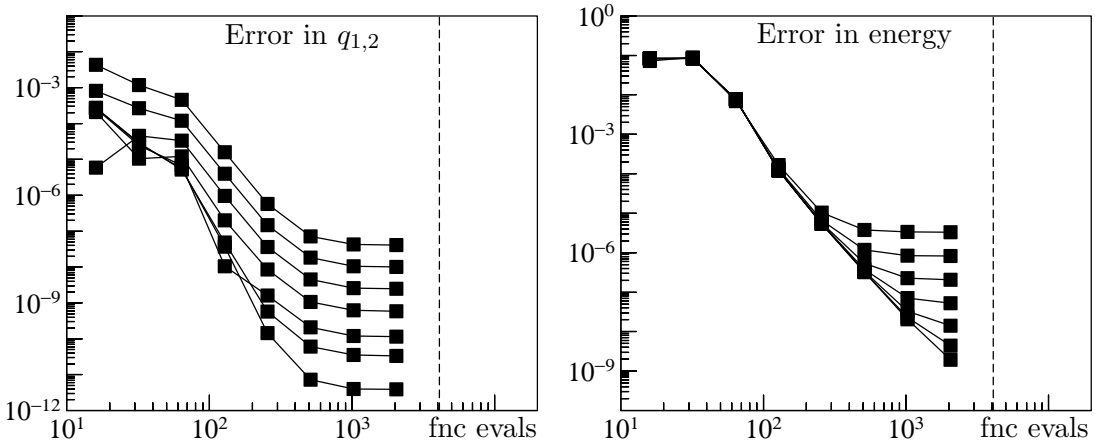


Figure 3: Multi-revolution method of order 4 ($s = 4$) for the Hamiltonian (26) with $\eta = 2^{-12}$. Error in energy in the multi-revolution approximation versus number of evaluations of $\Phi_{\varepsilon,h}$ approximating φ_ε (final time $t = 2\pi$). The lines correspond respectively to $h = 1/n = 2^{-j-1}$, $j = 1, \ldots, 7$ (from top to bottom).

Finally, in Figure 4, we plot, for $\eta = 2^{-12}$, the evolutions of the stiff spring energies

$$I_k = \frac{1}{2} p_{2,k}^2 + \frac{1}{2\eta^2} q_{2,k}^2, \quad k = 1, 2, 3,$$

the adiabatic invariant $I = I_1 + I_2 + I_3$ and the energy $E_\eta(p, q) - 0.7\,\eta$ on a time interval of length $T = 2\pi\eta^{-1}$ for $\eta = 2^{-12}$. We observe excellent energy conservation and energy exchanges for the MRCM methods of orders $1, 2, 4$ compared to the reference solution. This reference solution is computed with the standard Deuflhard method with constant step size $h = \eta$, totalizing about $1.1 \cdot 10^8$ steps (recall that a stepsize h comparable to the oscillatory period is needed for standard highly oscillatory integrators). In comparison, notice that the total number of micro steps (Strang splitting) for each of the considered MRCMs is $2snT\eta^{-1}N^{-1} \simeq 1.0 \cdot 10^6$, which identical in all cases due to the chosen parameters N, n and the stage number s of the MRCMs. What is striking in this experiment is that the multi-revolution composition approach yields errors in the energy exchanges lower than 5 percents for the methods of order 2 and 4 with a computational cost reduced by a factor 100 compared to the standard reference integrator.

Remark 4.1. *We have applied our MRCMs to the Hamiltonian $E_\eta(p, q)$ in (26), by considering for each value of η, a system of the form (10) that reduces to the original problem when $\varepsilon = 2\pi\eta$. This way, all the considerations made in the introduction for the application of MRCMs to systems of the form (10) apply directly to that case. Note that the considered family of near-to-identity φ_ε is different (although we do not reflect it in the notation) for each particular value of η.*

This is not however the only way to use MRCMs for the numerical integration of that Hamiltonian problem. For instance, the near-to-identity map φ_ε could simply be defined as the flow with time $2\pi\varepsilon$ of the Hamiltonian function $E_\varepsilon(p, q)$, in which case the convergence theory in Section 2 would also apply. This approach may seem attractive because the map φ_ε becomes

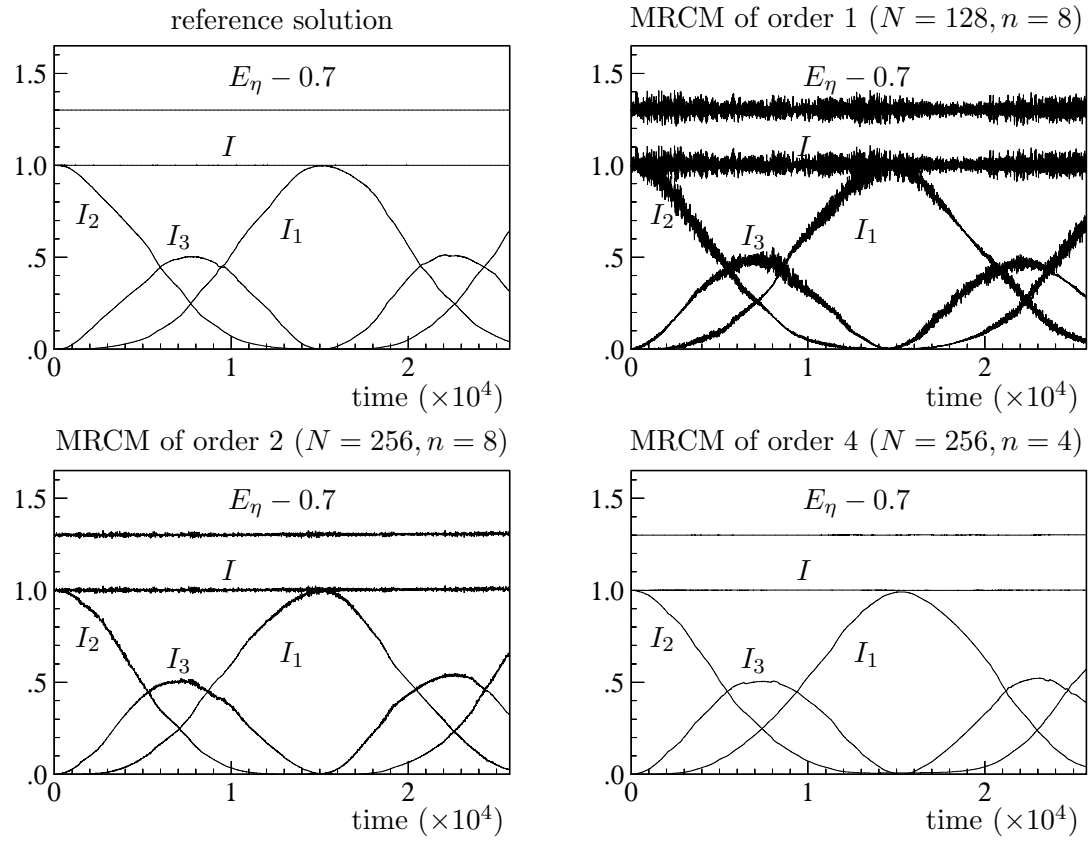


Figure 4: FPU-like problem (26) with $\eta = 2^{-12}$. Energy exchanges on the time interval $(0, 2\pi\eta^{-1})$. Multi-revolution methods of orders $1, 2, 4$. Reference solution computed with constant stepsize $h = \eta$ by the standard Deuflhard method.

symmetric with respect to ε (i.e. $\varphi_\varepsilon \circ \varphi_{-\varepsilon}(y) = y$) which would simplify considerably the order conditions (similarly to the case of standard composition methods based on a symmetric basic integrator). However, our numerical tests applied to (26) *and similar to Fermi-Pasta-Ulam like problems indicate poor performances of the derived methods on time intervals of size $\mathcal{O}(\varepsilon^{-1})$. This seems to be related to the fact that the first order averaged ODE (3) corresponding to that particular choice of φ_ε has unbounded solutions.*

4.2 Application to the cubic nonlinear Schrödinger equation

Highly oscillatory problems of the form (10) are in particular obtained by appropriate discretization in space of several Hamiltonian partial differentiation equations, such as nonlinear versions of wave equation and Schrödinger equation. In this section, we present some numerical experiments of the application of MRCMs to numerically integrate a problem considered in [7] and originally analyzed by B. Grébert and C. Villegas-Blas in [15]. It consists of a nonlinear Schrödinger equation in the with one-dimensional torus with a cubic nonlinearity $|u|^2 u$ multiplied by an inciting term of the form $2\cos(2x)$,

$$\begin{aligned} i\partial_t u &= -\Delta u + 2\varepsilon \cos(2x)|u|^2 u, \quad t \geq 0, \quad u(t,\cdot) \in H^s(\mathbb{T}_{2\pi}) \\ u(0,x) &= \cos x + \sin x. \end{aligned} \tag{28}$$

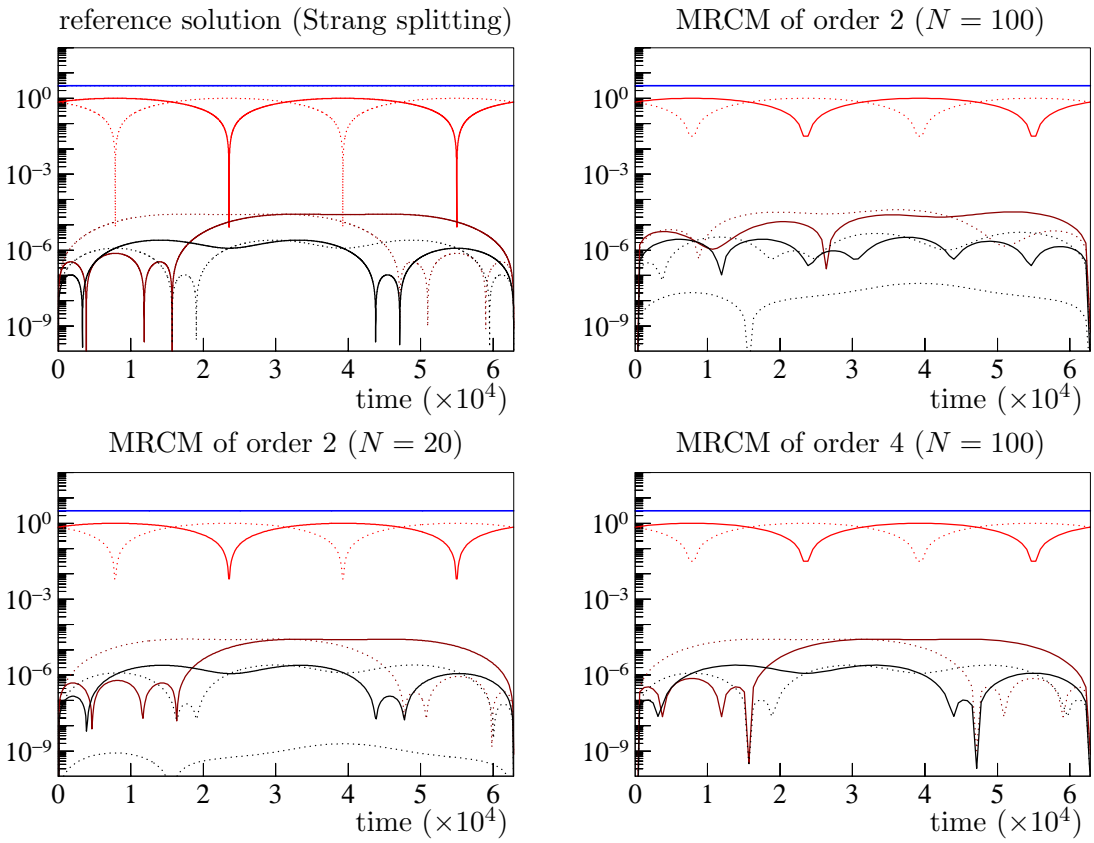


Figure 5: Nonlinear Schrödinger problem (28) with $\varepsilon = 10^{-4}$ on the time interval $(0, 2\pi\varepsilon^{-1})$. Plot of the actions $|\xi_j(t)|$, for $j = 1, 3, 5$ (solid lines) and for $j = -1, -3, -5$ (dotted lines) with colors red ($|j| = 1$), brown ($|j| = 3$), black ($|j| = 5$). The micro stepsize is $h = 2\pi/n$ with $n = 100$.

The problem is known to have a unique global solution in all Sobolev spaces $H^s(\mathbb{T}_{2\pi})$ for $s \geq 0$. A pseudospectral approximation of the form

$$u(t,x) \approx \sum_{k=-\ell}^{\ell} \xi_k(t) e^{ikx}$$

may be obtained by determining the approximate Fourier modes $\xi_k(t)$ as the solution with appropriate initial values of a semidiscrete version of equation (28)

$$\frac{d}{dt}\xi_k = -ik^2\xi_k + \varepsilon\, f_k(\xi_{-\ell}, \ldots, \xi_{-1}, \xi_0, \xi_1, \ldots, \xi_\ell), \quad k = -\ell, \ldots, -1, 0, 1, \ldots, \ell. \tag{29}$$

Clearly, the system of ODEs (29) can be recast into the format (10) by rescaling time (that is, by rewriting the system in terms of the new time variable $\hat{t} = \frac{\varepsilon}{2\pi}t$). The following nonlinear phenomenon is proved in [15].

Theorem 4.2. *[15] Consider the Fourier expansion $u(t,x) = \sum_{k\in\mathbb{Z}} \xi_k(t)e^{ikx}$ of the solution of (28). For all ε small enough, one has for all $|t| \leq \varepsilon^{-9/8}$ the following estimates:*

$$|\xi_1(t)|^2 = \frac{1+\sin(2\varepsilon t)}{2} + \mathcal{O}(\varepsilon^{1/8}), \qquad |\xi_{-1}(t)|^2 = \frac{1-\sin(2\varepsilon t)}{2} + \mathcal{O}(\varepsilon^{1/8}).$$

These estimates imply that the energy remains essentially concentrated in Fourier modes $+1$ and -1 and that these modes exchange their energy periodically (with period π/ε). This effect is named "beating effect" in [15]. Another interesting part of the dynamics of this system concerns the modes 3 and -3 whose energies scale like ε^2 and may be regarded for this reason as a "finer" component of the dynamics.

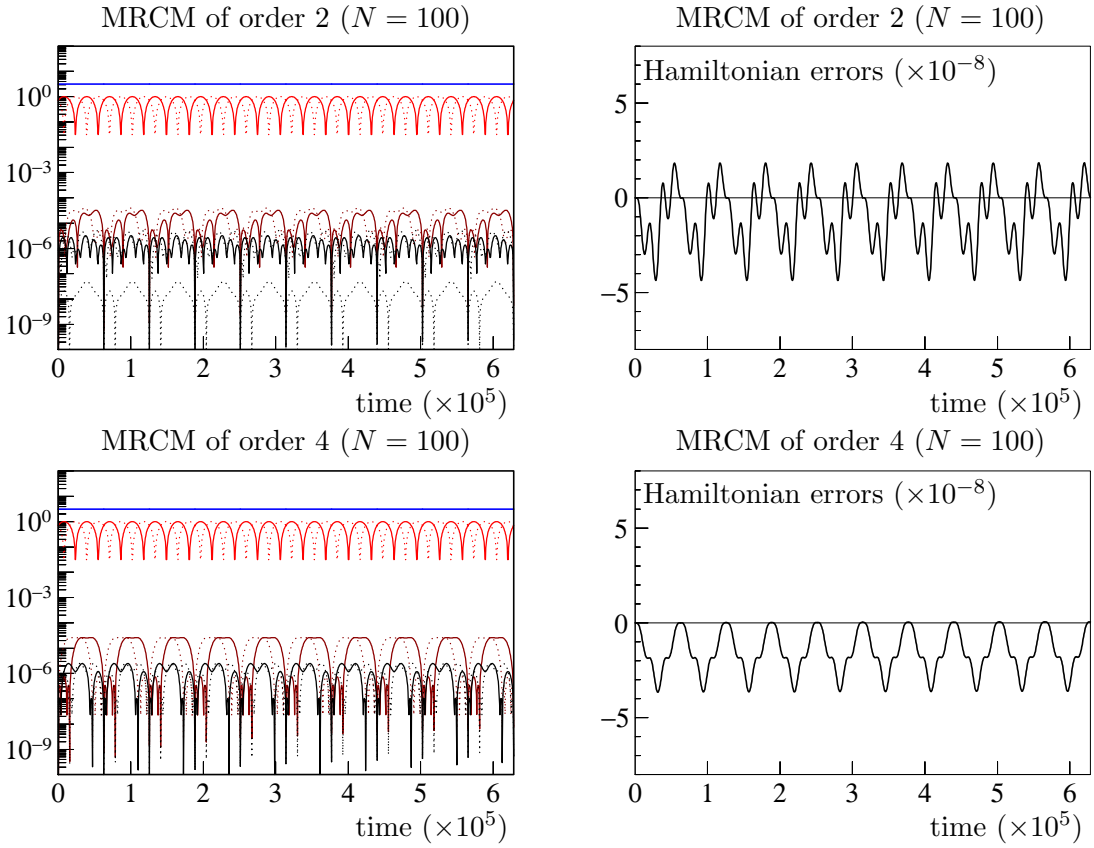


Figure 6: Nonlinear Schrödinger problem (28) with $\varepsilon = 10^{-4}$ on the time interval $(0, 20\pi\varepsilon^{-1})$. Left pictures: actions $|\xi_j(t)|$, for $j = 1, 3, 5$ (solid lines) and for $j = -1, -3, -5$ (dotted lines) with colors red ($|j| = 1$), brown ($|j| = 3$), black ($|j| = 5$). Right pictures: corresponding Hamiltonian errors. The micro stepsize is $h = 2\pi/n$ with $n = 100$.

We have applied our MRCMs to (28) as indicated in Section 1, using the Strang splitting method between linear and nonlinear parts, with step size $h = 2\pi/100$ as the micro-integrator $\Phi_{\mu,h}$. This permits to conserve exactly the mass $\sum_j |\xi_j|^2$ (up to round-off errors). All MRCMs capture the beating effect whereas only methods of order 2 with moderate N and of order 4 with possibly larger N provide satisfactory approximation of modes $+3$ and -3. These facts can be clearly observed in Figure 5 which have been obtained by simulating equation (28) on the time interval $(0, 2\pi\varepsilon^{-1})$ with composition methods of orders 2 ($s = 1$) and 4 ($s = 4$): we have represented in logarithmic scale the modes $|\xi_j|$ for $j = \pm 1$, ± 3, ± 5, ± 7 (notice that the even modes $\xi_{2j}, j \geq 0$ are zero). Here, we consider a spectral Fourier discretization with modes $\xi_j, j = -64, \ldots, 64$. Modes ± 1 are of order $\mathcal{O}(1)$ and the beating effect is well-reproduced by all methods. The energy (thick blue line at the top) is well conserved again by all methods. Modes ± 3 are of order $\mathcal{O}(\varepsilon)$, i.e. $|\xi_{\pm 3}|^2 = \mathcal{O}(\varepsilon^2)$, and are well-captured for

the second-order MRCM with moderate values of N or with the fourth-order MRCM with $N = 100$. Although we do not give here theoretical error estimates for PDEs, the qualitative behavior of the dynamics of equation (28) is clearly well reproduced for a computational cost that is significantly smaller as compared to Strang splitting by its own. Here, for $\varepsilon = 10^{-4}$, the cost is reduced by a factor 10 for the second-order MRCM and by a factor 16 for the fourth-order MRCM. In Figure 6, we further investigate the behavior of the methods on a time interval ten times larger. We observe that the beating effect is still well captured (left pictures, while an excellent energy conservation (without drift) can be observed (right pictures).

5 Conclusion

We have presented a class of multi-revolution methods based on compositions for highly oscillatory systems. We proved uniform convergence estimates independently of the stiffness of the high oscillations which permits to use large time steps.

We point out that our convergence analysis applies only to bounded times intervals of size $\mathcal{O}(1)$ with respect to the ε. Notice however that, although to proposed schemes are not symmetric (expect the symmetric method with $s = 1$ and the conjugate symmetric method with $s = 2$), the coefficients of the implemented methods are $\mathcal{O}(N^{-1})$ close to those of a symmetric method. This means that the excellent energy conservation results [16, Chap. 5] of symmetric methods applied to reversible systems remain valid for the proposed methods on longer time intervals of length $T = \mathcal{O}(\varepsilon^{-1})$.

The versatility and robustness of the multi-revolution approach is illustrated on various problems including the nonlinear Schrödinger equation. For the extension to infinite dimensional problems, we mention the recent works [8] in the context of the linear Schrödinger equation and [27] in the context of stochastic nonlinear systems.

References

[1] A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numer.*, 21:1–87, 2012.

[2] M. Calvo, L. O. Jay, J. I. Montijano, and L. Rández. Approximate compositions of a near identity map by multi-revolution Runge-Kutta methods. *Numer. Math.*, 97(4):635–666, 2004.

[3] M. Calvo, J. I. Montijano, and L. Rández. A family of explicit multirevolution Runge-Kutta methods of order five. In *Analytic and numerical techniques in orbital dynamics (Spanish) (Albarracín, 2002)*, Monogr. Real Acad. Ci. Exact. Fís.-Quím. Nat. Zaragoza, 22, pages 45–54. Acad. Cienc. Exact. Fís. Quím. Nat. Zaragoza, Zaragoza, 2003.

[4] M. Calvo, J. I. Montijano, and L. Rández. On explicit multi-revolution Runge-Kutta schemes. *Adv. Comput. Math.*, 26(1-3):105–120, 2007.

[5] M. P. Calvo, P. Chartier, A. Murua, and J. M. Sanz-Serna. Numerical stroboscopic averaging for ODEs and DAEs. *Appl. Numer. Math.*, 61(10):1077–1095, 2011.

[6] F. Castella, P. Chartier, S. Descombes, and G. Vilmart. Splitting methods with complex times for parabolic equations. *BIT*, 49(3):487–508, 2009.

[7] F. Castella, P. Chartier, F. Méhats, and A. Murua. Stroboscopic averaging for the nonlinear Schrödinger equation. Technical report, Sept. 2012.

[8] P. Chartier, F. Mehats, and M. Thalhammer. Convergence analysis of multi-revolution composition time-splitting pseudo-spectral methods for Schrdinger equations. Part I. the linear case. *In preparation*, 2013.

[9] P. Chartier, A. Murua, and J. M. Sanz-Serna. Higher-order averaging, formal series and numerical integration I: B-series. *Found. Comput. Math.*, 10(6):695–727, 2010.

[10] P. Chartier, A. Murua, and J. M. Sanz-Serna. A formal series approach to averaging: exponentially small error estimates. *Discrete Contin. Dyn. Syst.*, 32(9):3009–3027, 2012.

[11] W. E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003.

[12] W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden. Heterogeneous multiscale methods: a review. *Commun. Comput. Phys.*, 2(3):367–450, 2007.

[13] E. Fermi, J. Pasta, and S. Ulam. Technical Report LA-1940, Los Alamos, 1955. Later published in E. Fermi: Collected Papers Chicago 1965 and Lect. Appl. Math. 15 143 1974.

[14] F. Filbet and S. Jin. A class of asymptotic-preserving schemes for kinetic equations and related problems with stiff sources. *J. Comput. Phys.*, 229(20):7625–7648, 2010.

[15] B. Grébert and C. Villegas-Blas. On the energy exchange between resonant modes in nonlinear Schrödinger equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 28(1):127–134, 2011.

[16] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2010. Structure-preserving algorithms for ordinary differential equations, Reprint of the second (2006) edition.

[17] E. Hansen and A. Ostermann. High order splitting methods for analytic semigroups exist. *BIT*, 49(3):527–542, 2009.

[18] S. Jin. Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations. *SIAM J. Sci. Comput.*, 21(2):441–454, 1999.

[19] U. Kirchgraber. An ODE-solver based on the method of averaging. *Numer. Math.*, 53(6):621–652, 1988.

[20] M. Lemou and L. Mieussens. A new asymptotic preserving scheme based on micro-macro formulation for linear kinetic equations in the diffusion limit. *SIAM J. Sci. Comput.*, 31(1):334–368, 2008.

[21] R. I. McLachlan. On the numerical integration of ordinary differential equations by symmetric composition methods. *SIAM J. Sci. Comput.*, 16:151–168, 1995.

[22] R. I. McLachlan and G. R. W. Quispel. Splitting methods. *Acta Numerica*, 11:341–434, 2002.

[23] B. Melendo and M. Palacios. A new approach to the construction of multirevolution methods and their implementation. *Appl. Numer. Math.*, 23(2):259–274, 1997.

[24] A. Murua and J. M. Sanz-Serna. Order conditions for numerical integrators obtained by composing simpler integrators. *Philos. Trans. Royal Soc. London ser. A*, 357:1079–1100, 1999.

[25] L. R. Petzold, L. O. Jay, and J. Yen. Numerical solution of highly oscillatory ordinary differential equations. In *Acta numerica, 1997*, volume 6 of *Acta Numer.*, pages 437–483. Cambridge Univ. Press, Cambridge, 1997.

[26] J. Sanders, F. Verhulst, and J. Murdock. *Averaging methods in nonlinear dynamical systems*. Springer, New York, second edition, 2007.

[27] G. Vilmart. Weak second order multi-revolution composition methods for highly oscillatory stochastic differential equations with additive or multiplicative noise. *Submitted*, 2013.