ESAIM: M2AN 49 (2015) 695–711 DOI: 10.1051/m2an/2014056

HIGH ORDER NUMERICAL METHODS FOR HIGHLY OSCILLATORY PROBLEMS

DAVID COHEN¹ AND JULIA SCHWEITZER²

Abstract. This paper is concerned with the numerical solution of nonlinear Hamiltonian highly oscillatory systems of second-order differential equations of a special form. We present numerical methods of high asymptotic as well as time stepping order based on the modulated Fourier expansion of the exact solution. In particular we obtain time stepping orders higher than 2 with only a finite energy assumption on the initial values of the problem. In addition, the stepsize of these new numerical integrators is not restricted by the high frequency of the problem. Furthermore, numerical experiments on the modified Fermi–Pasta–Ulam problem as well as on a one dimensional model of a diatomic gas with short-range interaction forces support our investigations.

Mathematics Subject Classification. 34E05, 34E13, 65L20, 65P10.

Received March 4, 2014. Revised November 9, 2014 Published online 8 April 2015.

1. Introduction

We consider the numerical discretisation of second-order Hamiltonian differential equations with highly oscillatory solutions of the special form

$$\ddot{x} + \Omega^2 x = g(x) := -\nabla U(x),\tag{1.1}$$

with the square block matrix

$$\Omega = \begin{pmatrix} 0 & 0 \\ 0 & \omega I \end{pmatrix}, \text{ with } \omega \gg 1,$$

and a smooth nonlinear potential U. This is a Hamiltonian problem with

$$H(x, \dot{x}) = \frac{1}{2} \|\dot{x}\|^2 + \frac{1}{2} \|\Omega x\|^2 + U(x). \tag{1.2}$$

 $Keywords\ and\ phrases.\ Highly\ oscillatory\ differential\ equations,\ multiple\ time\ scales,\ Fermi-Pasta-Ulam\ problem,\ modulated\ Fourier\ expansions,\ high\ order\ numerical\ schemes,\ adiabatic\ invariants.$

Matematik och matematisk statistik, Umeå universitet, 90187 Umeå, Sweden. david.cohen@math.umu.se

² Institut für Angewandte und Numerische Mathematik, Karlsruher Institut für Technologie, 76128 Karlsruhe, Germany. julia.schweitzer@kit.edu

For the above problem, we will only consider initial values satisfying the bounded energy assumption

$$\frac{1}{2} \|\dot{x}(0)\|^2 + \frac{1}{2} \|\Omega x(0)\|^2 \le E \tag{1.3}$$

with a constant E independent of the large parameter ω .

We decompose the solution $x = (x_1, x_2)$ and the nonlinearity $g(x) = (g_1(x), g_2(x))$ according to the blocks of the matrix Ω . One thus has a solution with a slowly varying component x_1 and a rapidly varying one x_2 .

Due to the presence of fast and slow time scales the numerical discretisation of problems of the form (1.1) is a difficult task. One of the most efficient integrators for this kind of highly-oscillatory problems are the trigonometric integrators, see for example the reviews [8,18] and ([16], Chap. XIII) as well as references therein. Other type of numerical schemes were also proposed in the literature. Let us mention (without being exhaustive) the following works: the methods based on averaging techniques from [3], see also [4]; the numerical schemes based on homogenisation from [19]; the multi-scale methods from [22]; the IMEX method from [21], see also the recent publication [20]; and more recently the energy-preserving schemes from [23]. The order of convergence in the time stepsize of these numerical integrators is at most two. The main motivation of the present work is to present and analyse numerical methods for problems (1.1) of higher order of convergence independent of the frequency imposing only a finite energy assumption on the initial values of the problem (1.1).

For the numerical discretisation of (1.1), we will use a modulated Fourier expansion (MFE), see [6,15]. This analytical tool basically decomposes the exact solution of our problem into a slowly varying part and into oscillatory parts. The first terms of this expansion were already used in [5] to develop efficient geometric numerical integrators for (1.1). In particular, symmetric and reversible methods based on the MFE were presented in [5] together with their energy-conservation properties on longtime intervals. However, the order of the methods proposed in [5] was not examined and extensions of these methods to higher order numerical schemes remain to be investigated. This is the main objective of the present paper. At this point, it should be noted that one drawback of the proposed approach is the fact that it may prove difficult to generalise it to more complicated situations such as the semi-discretisation of nonlinear wave equations for example. This mainly comes from the fact that a result like Lemma 2.2 (see below) is not available in this situation. A further difficulty is the absence of a clear separation of the different scales in this situation.

Much in the same spirit as the present work is the series of papers [9–11] on asymptotic-numerical solvers for differential equations with highly oscillatory coefficients. In these references, the special structure (a MFE in fact) of the forcing term present in the problem permits to derive efficient numerical methods. Furthermore, the results from the recent reference [1] are also closely related to the type of problems (1.1) and to the numerical schemes that we propose. The major difference is that the work [1] deals with a model problem for the nonlinear Klein-Gordon equation in the nonrelativistic limit regime, in particular the driven force in the differential equation is a pure power nonlinearity satisfying $f(e^{isy}) = e^{is} f(y)$ for all $s \in \mathbb{R}$. Let us finally mention the recent work [12] which uses the first terms of an asymptotic expansion similar to a MFE to derive numerical schemes for the Klein-Gordon equation in the nonrelativistic limit regime with smooth initial condition. The main difference with our approach is that in [12] a partial differential equation (PDE) with a scalar parameter that induces high oscillations is considered. A complete asymptotic expansion of the solution to this PDE is then derived. Further, the coefficients of this expansion are given as solutions to PDEs which are independent of this parameter and thus the numerical time integration of the Klein-Gordon equation can be carried out without step size restrictions due to this parameter. In the present work, we consider systems of highly oscillatory ordinary differential equations (ODEs) and obtain nonstiff ODEs for the coefficient functions of the MFE (see Eqs. (2.2) and (2.3) below). This then allows us to use high order explicit time stepping methods for the solution to the original problem.

The paper is organised as follows. After recalling some results on the MFE in Section 2, we will describe the construction of the numerical schemes based on the first terms of the modulated Fourier expansion following the lines of ([16], Sect. XIII.3.1) and ([5], Chap. 3) in Section 3. In the fourth section, we will derive numerical schemes of order one, two, three and four for (1.1). In the last section we present numerical experiments on the

modified Fermi–Pasta–Ulam problem as well as on a one dimensional model of a diatomic gas with short-range interaction forces.

2. Preparatory results and building-blocks for high order numerical schemes

We present two fundamental results from ([16], Chap. XIII) and [6] that we need in order to derive high order numerical schemes for (1.1). First, let us recall the following result on the modulated Fourier expansion of the exact solution of (1.1).

Theorem 2.1 (Thm. XIII.5.1 of [16]). Consider a solution x(t) of (1.1) which satisfies the bounded-energy condition (1.3) and stays in a compact set K for $0 \le t \le T$. Then, the solution admits an expansion

$$x(t) = y(t) + \sum_{0 < |k| < N} e^{ik\omega t} z^{k}(t) + R_{N}(t)$$

for arbitrary $N \geq 2$, where the remainder term and its derivative are bounded by

$$R_N(t) = \mathcal{O}(\omega^{-N})$$
 and $\dot{R}_N(t) = \mathcal{O}(\omega^{-N})$ for $0 \le t \le T$.

(Remark that one can get slightly sharper bounds for the remainder terms, see the end of the proof of [16], Thm. XIII.5.1, for details). The real-valued functions $y = (y_1, y_2)$ and the complex-valued functions $z^k = (z_1^k, z_2^k)$ together with all their derivatives (up to arbitrary order M) are bounded by

$$y_1 = \mathcal{O}(1),$$
 $z_1^1 = \mathcal{O}(\omega^{-3}),$ $z^k = \mathcal{O}(\omega^{-k-2}),$
 $y_2 = \mathcal{O}(\omega^{-2}),$ $z_2^1 = \mathcal{O}(\omega^{-1}),$

for $k=2,\ldots,N-1$. Moreover $z^{-k}=\bar{z}^k$ for all k. The constants symbolised by the \mathcal{O} -notation are independent of ω and t with $0\leq t\leq T$ but they depend on N,T, on E from (1.3), and on bounds of the derivatives of the smooth nonlinearity g(x). Finally, the asymptotic expansion

$$x_{\star}(t) := y(t) + \sum_{0 < |k| < N} e^{ik\omega t} z^{k}(t)$$
 (2.1)

is called the (truncated) modulated Fourier expansion of the exact solution.

In order to find the modulation functions y(t) and $z^k(t)$, one inserts the expansion (2.1) into the differential equation (1.1), expands the nonlinearity g(x) around the smooth part $(y_1(t), 0)$ and compares the coefficients in front of $e^{ik\omega t}$. This yields a system of nonstiff differential equations for $y_1(t)$ and $z_2^1(t)$ and algebraic relations for the rest of the modulation functions.

As explained in the proof of (Thm. XIII.5.1 in [16]), the functions $y_1(t)$ and $z_2^1(t)$ are given by differential equations of the form

$$\ddot{y}_1 = \sum_{l>0} \omega^{-l} F_{1l}(y_1, \dot{y}_1, z_2^1), \quad \dot{z}_2^1 = \sum_{l>1} \omega^{-l} F_{2l}(y_1, \dot{y}_1, z_2^1), \tag{2.2}$$

and the remaining modulation functions by algebraic relations

$$z_i^k = \sum_{l>0} \omega^{-l} G_{il}^k(y_1, \dot{y}_1, z_2^1). \tag{2.3}$$

Observe that $y_2 = z_2^0$, and that z_i^{-k} is the complex conjugate of z_i^k , so that also G_{il}^{-k} is the complex conjugate of G_{il}^k . For an efficient computation of the functions F_{il} and G_{il}^k , we rely on the recurrence relations given by the following

Lemma 2.2 (Lem. 2.1 of [6]). The functions F_{il} and G_{il}^k defining the nonstiff differential equations (2.2) and algebraic equations (2.3) satisfy the recurrence relations (for $l \ge 0$):

$$F_{1l} = S_1(0, l), \qquad G_{1l}^k = \frac{1}{k^2} \left(\sum_{m+n+j=l-2} \mathcal{L}_m \mathcal{L}_n G_{1j}^k + 2ik \sum_{m+j=l-1} \mathcal{L}_m G_{1j}^k - S_1(k, l-2) \right),$$

$$F_{2l} = \frac{1}{2i} \left(S_2(1, l-1) - \sum_{m+j=l-1} \mathcal{L}_m F_{2j} \right), \quad G_{2l}^k = \frac{1}{k^2 - 1} \left(\sum_{m+n+j=l-2} \mathcal{L}_m \mathcal{L}_n G_{2j}^k + 2ik \sum_{m+j=l-1} \mathcal{L}_m G_{2j}^k - S_2(k, l-2) \right).$$

The sums are over $m \ge 0, n \ge 0, j \ge 0$, and we have used the abbreviation

$$S_i(k,l) = \sum_{m,n\geq 0} \frac{1}{m!n!} \sum_{\substack{\alpha,\beta\\s(\alpha)+s(\beta)=k}} \sum_{\substack{e,f\\s(e)+s(f)=l}} D_1^m D_2^n g_i(y_1,0) \left(G_{1e}^{\alpha}, G_{2f}^{\beta}\right).$$

Here, $\alpha = (\alpha_1, \dots, \alpha_m)$, $\beta = (\beta_1, \dots, \beta_n)$, $e = (e_1, \dots, e_m)$ as well as $f = (f_1, \dots, f_n)$ are multi-indices with $\alpha_i \neq 0$, β_i arbitrary, $e_i \geq 0$, $f_i \geq 0$, and $(G_{1e}^{\alpha}, G_{2f}^{\beta}) = (G_{1,e_1}^{\alpha_1}, \dots, G_{1,e_m}^{\alpha_m}, G_{2,f_1}^{\beta_1}, \dots, G_{2,f_n}^{\beta_n})$. We use the abbreviation $s(\alpha) = \sum_{i=1}^m \alpha_i$ and similarly for the other multi-indices. We denote by $D_1^m D_2^n g_i(y_1, 0)(G_{1e}^{\alpha}, G_{2f}^{\beta})$ the mth derivative with respect to the first component and nth derivative with respect to the second component of $g_i(y_1, 0)$ evaluated at $(G_{1e}^{\alpha}, G_{2f}^{\beta})$. Furthermore, the operator \mathcal{L}_l applied to a smooth function $G(y_1, \dot{y}_1, z_2)$ is defined by

$$\mathcal{L}_{l}G = \hat{D}_{2}G \cdot F_{1l} + \hat{D}_{3}G \cdot F_{2l} + \begin{cases} \hat{D}_{1}G \cdot \dot{y}_{1} & if \ l = 0\\ 0 & if \ l \geq 1, \end{cases}$$

where $l \ge 0$ and \hat{D}_j denotes the partial derivative with respect to the jth argument of $G(y_1, \dot{y}_1, z_2)$. Finally, we note that

$$\begin{split} G_{10}^k &= 0, & G_{11}^k &= 0 \quad for \; k \neq 0, \\ G_{20}^k &= 0, & G_{21}^k &= 0 \quad for \; k \neq \pm 1, \\ G_{20}^{\pm 1} &= z_2^{\pm 1}, & G_{2l}^{\pm 1} &= 0 \quad for \; l \geq 1. \end{split}$$

Bounds for the functions F_{il} and G_{il}^k can be obtained by ([6], Thm. 2.4). Furthermore, observe that, by tracing the dependence on N in the remainder terms for the asymptotic expansion given by the MFE, one can obtain exponentially small defect for the MFE, see ([6], Thm. 3.2). These results are however not used in the construction of the high order numerical methods proposed in the next section.

3. General construction of the numerical methods and some results

We propose to solve the truncated nonstiff system of differential equations for the modulation functions y_1 and z_2^1 up to asymptotic order N using a standard numerical time stepping scheme, such as an explicit Runge–Kutta method of higher time stepping order p with a (possibly) large time stepsize. To obtain an approximation to the solution of the original highly oscillatory problem (1.1), we insert the numerical approximations of y_1 and z_2^1 into the algebraic equations for the remaining modulation functions and compute the truncated MFE (2.1) of asymptotic order N.

To summarise, we obtain a numerical integrator of asymptotic order N and time stepping order p for (1.1) following the flow chart

- (1) Truncate the ansatz (2.2) and (2.3) after the $\mathcal{O}(\omega^{-N})$ -terms.
- (2) Use the recurrence relations given by the above lemma up to terms of asymptotic order $\mathcal{O}(\omega^{-N})$.

- (3) Solve numerically the nonstiff truncated differential equations (2.2) with an explicit scheme of high time stepping order p.
- (4) Compute the algebraic relations (2.3) and the approximation of the solution of (1.1) using the MFE (2.1).

We would like to point out, that the initial values $y_1(0), \dot{y}_1(0)$ and $z_2^1(0)$ for the differential equations (2.2) are obtained from the conditions

$$x_{\star}(0) = x(0)$$
 and $\dot{x}_{\star}(0) = \dot{x}(0)$.

This yields a nonlinear system of equations, which can be solved by a fixed point iteration as shown in ([16], Chap. XIII.3).

Note, that step (4) of the flow chart is to be understood as a postprocessing, which only needs to be performed at times, where the solution of the original problem is required. The size of the new system is smaller than that of the original one. Therefore, in terms of storage, the method is comparable to other methods applied directly to the original system. However, the right hand side of the new system is more complicated depending on the nonlinearity and the asymptotic order. If the nonlinear interaction is nonlocal, the computational cost may grow considerably. It approximately scales with the largest number of directly interacting variables via the nonlinearity to the power of the asymptotic order times the number of all variables, since the derivatives of the nonlinearity appearing in the equations for the truncated system yield multilinear forms of the respective order.

We have the following error estimates on compact time intervals for the proposed numerical approximations of solutions to problem (1.1).

Proposition 3.1. Under the assumptions of Theorem 2.1, if we solve the truncated system for the modulation functions of asymptotic order N with a numerical time stepping scheme of order p, then the error of the approximation satisfies (here $t_n = nh$)

$$\left\| x(t_n) - y_n - \sum_{0 < |k| < N} e^{ik\omega t_n} z_n^k \right\| \le C_a \omega^{-N} + C_t h^p.$$

In addition, the approximation of the time derivative of the solution satisfies

$$\left\| \dot{x}(t_n) - \dot{y}_n - \sum_{0 < |k| < N} \left(ik\omega z_n^k + \dot{z}_n^k \right) e^{ik\omega t_n} \right\| \le \widetilde{C}_a \omega^{-N} + \widetilde{C}_t h^p,$$

where the approximation to $\dot{z}_{2,n}^1$ is obtained by the differential equation for z_2^1 and the approximations to the remaining modulation functions are computed by differentiation of the algebraic relations. The constants C_a , C_t , \tilde{C}_a and \tilde{C}_t do not depend on h or n. Moreover C_a and \tilde{C}_a do not depend on ω .

Proof. For the proof of the error bound of the solution itself, we insert the exact truncated MFE and use the triangle inequality to obtain

$$\| x(t_n) - y_n - \sum_{0 < |k| < N} z_n^k e^{ik\omega t_n} \| \le \| x(t_n) - x_\star(t_n) \| + \| y(t_n) + \sum_{0 < |k| < N} z^k(t_n) e^{ik\omega t_n} - y_n - \sum_{0 < |k| < N} z_n^k e^{ik\omega t_n} \| .$$

The first term yields the asymptotic part of the error by Theorem 2.1. For y_1 and z_2^1 the numerical time stepping scheme of order p yields approximations $y_{1,n}$ and $z_{2,n}^1$ satisfying $y_1(t_n) - y_{1,n} = \mathcal{O}(h^p)$ as well as $z_2^1(t_n) - z_{2,n}^1 = \mathcal{O}(h^p)$. For a sufficiently smooth function g the error of the approximation to the remaining modulation functions is then at most $\mathcal{O}(h^p)$. This yields the stated error bound.

For the error of the derivative of the solution, we obtain

$$\begin{aligned} & \left\| \dot{x}(t_n) - \dot{y}_n - \sum_{0 < |k| < N} (\mathrm{i}k\omega z_n^k + \dot{z}_n^k) \mathrm{e}^{\mathrm{i}k\omega t_n} \right\| \\ & \leq \left\| \dot{x}(t_n) - \dot{x}_\star(t_n) \right\| + \left\| \dot{y}(t_n) + \sum_{0 < |k| < N} \left(\mathrm{i}k\omega z^k(t_n) + \dot{z}^k(t_n) \right) \mathrm{e}^{\mathrm{i}k\omega t_n} - \dot{y}_n - \sum_{0 < |k| < N} \left(\mathrm{i}k\omega z_n^k + \dot{z}_n^k \right) \mathrm{e}^{\mathrm{i}k\omega t_n} \right\|. \end{aligned}$$

To treat the term $ik\omega(z^k(t_n)-z_n^k)$, we observe that the modulation functions obtained by the algebraic relations are of the form ω^{-k} times functions of at most order one evaluated at the numerical solutions of the MFE. This therefore cancels the factor of ω and only the error of the time stepping scheme remains. The remaining terms are treated similarly as in the first part of the proof.

Note that the constants C_a and \widetilde{C}_a do not depend on the time stepsize h or on the number of time steps n, since they directly stem from the error of the truncated asymptotic expansion. Furthermore, they are independent of ω by the results of Theorem 2.1. The constants C_t and \widetilde{C}_t are independent of the time stepsize h and the number of time steps n since they are related to the error constant of the numerical method used for the integration of the nonstiff truncated system.

Let us observe that the constants C_t and \widetilde{C}_t may dependent on ω , but only on inverse powers thereof. These constants are the usual error constants of the explicit methods we choose to apply to the nonstiff truncated system. They therefore depend on the coefficients of the numerical method, on the length of the computation interval, the Lipschitz constant as well as some derivatives (depending on the time stepping order p of the used method) of the nonlinearity. Since the nonlinearity g of the original problem is assumed to be sufficiently smooth and independent of ω , the nonlinearity of the new systems, which consists of sums of derivatives of g multiplied with inverse powers of ω , only depends weakly on ω (i.e. on inverse powers of ω). Thus only this weak dependence enters the constants C_t and \widetilde{C}_t . Moreover, the numerical experiments in the following give evidence that there is indeed no essential dependence of the constants C_t and \widetilde{C}_t on ω .

Let us further mention that the proposed numerical methods do not suffer from numerical resonances (when the product of the stepsize and the large frequency ω is close to a nonzero integer multiple of π). Furthermore, the proposed schemes (of asymptotic order greater than one) correctly model the slow energy exchange in the problem, since the evolution of z_2^1 , which describes this slow exchange of energy, is present in our numerical approximation given by the truncated MFE (2.1). This will be illustrated in the numerical experiments presented in Section 5.3.

Finally, as demonstrated in [5], by carefully selecting the numerical integrators for the time integration of the nonstiff system (2.2), one can obtain longtime energy-conservation properties of the numerical solutions using the proposed approach. Such a result for generic integrators is not in the scope of the present work. But, one can still observe excellent longtime energy-conservation, even in this case, in the numerical experiments presented below.

4. High order numerical methods

This section presents the nonstiff systems together with their algebraic relations and the corresponding truncated MFE up to asymptotic order four approximating highly oscillatory problems of the form (1.1). The derivation directly follows from Lemma 2.2 but is rather tedious, thus we only present the results. Throughout the remainder of the section we use the abbreviation $g_i := g_i(y_1, 0)$ for i = 1, 2 for the sake of presentation.

4.1. System of asymptotic order one

In order to obtain an asymptotic approximation of order one for problems of the form (1.1), one needs to solve the nonstiff system (recall that g_1 denotes $g_1(y_1, 0)$)

$$\ddot{y}_1 = g_1$$

and compute the truncated MFE

$$x_{\star}(t) := (y_1(t), 0).$$

In this case, there are no algebraic relations needed yet. Note, that Theorem 2.1 actually only holds for $N \ge 2$. However, we included this system for the sake of completeness. It will be shown later, that the scaling in ω holds, but the solution does not show the proper energy conservation properties.

4.2. System of asymptotic order two

In order to obtain an asymptotic approximation of order two for problems of the form (1.1), one needs to solve the nonstiff system (recall that D_2g_i denotes the derivative with respect to the second component of g_i evaluated at $(y_1, 0)$, here i = 1, 2)

$$\ddot{y}_1 = g_1 + D_2^2 g_1(z_2^1, z_2^{-1}) + \omega^{-2} D_2 g_1 g_2,$$

$$\dot{z}_2^1 = \frac{\omega^{-1}}{2i} D_2 g_2 z_2^1,$$

insert the solution into the algebraic relations

$$y_2 = z_2^0 = \omega^{-2} g_2,$$

$$z_1^1 = 0,$$

and with this, compute the truncated MFE

$$x_{\star}(t) := y(t) + e^{i\omega t}z^{1}(t) + e^{-i\omega t}\bar{z}^{1}(t).$$

4.3. System of asymptotic order three

In order to obtain an asymptotic approximation of order three for problems of the form (1.1), one needs to solve the nonstiff system (recall that D_lg_i denotes the derivative with respect to the lth component of g_i evaluated at $(y_1, 0)$, here i = 1, 2 and l = 1, 2)

$$\ddot{y}_1 = g_1 + D_2^2 g_1 \left(z_2^1, z_2^{-1} \right) + \omega^{-2} D_2 g_1 g_2, \tag{4.3a}$$

$$\dot{z}_{2}^{1} = \frac{\omega^{-1}}{2i} D_{2} g_{2} z_{2}^{1} + \frac{\omega^{-2}}{4} D_{1} D_{2} g_{2} \left(\dot{y}_{1}, z_{2}^{1}\right), \tag{4.3b}$$

insert the solution into the algebraic relations

$$y_2 = z_2^0 = \omega^{-2} g_2,$$

$$z_1^1 = -\omega^{-2} D_2 g_1 z_2^1,$$

and with this, compute the truncated MFE

$$x_{\star}(t) := y(t) + e^{i\omega t}z^{1}(t) + e^{-i\omega t}\bar{z}^{1}(t).$$

4.4. System of asymptotic order four

Finally, in order to obtain an asymptotic approximation of order four for problems of the form (1.1), one needs to solve the nonstiff system (recall that D_lg_i denotes the derivative with respect to the lth component

of g_i evaluated at $(y_1, 0)$, here i = 1, 2 and l = 1, 2)

$$\ddot{y}_{1} = g_{1} + D_{2}^{2}g_{1}\left(z_{2}^{1}, z_{2}^{-1}\right) + \omega^{-2}D_{2}g_{1}g_{2} + \frac{1}{4}D_{2}^{4}g_{1}\left(z_{2}^{1}, z_{2}^{1}, z_{2}^{-1}, z_{2}^{-1}\right)
+ \omega^{-2}\left\{D_{2}g_{1}D_{2}^{2}g_{2}\left(z_{2}^{1}, z_{2}^{-1}\right) + D_{2}^{3}g_{1}\left(z_{2}^{1}, z_{2}^{-1}, g_{2}\right) - 2\operatorname{Re}D_{1}D_{2}g_{1}\left(D_{2}g_{1}z_{2}^{-1}, z_{2}^{1}\right)\right\}
+ \omega^{-4}\left\{D_{2}g_{1}\left(D_{2}g_{2}g_{2} - D_{1}g_{2}g_{1} - D_{1}^{2}g_{2}\left(\dot{y}_{1}, \dot{y}_{1}\right)\right) + \frac{1}{2}D_{2}^{2}g_{1}\left(g_{2}\left(y_{2}, 0\right), g_{2}\right)\right\},$$

$$\dot{z}_{2}^{1} = \frac{\omega^{-1}}{2\mathrm{i}}D_{2}g_{2}z_{2}^{1} + \frac{\omega^{-2}}{4}D_{1}D_{2}g_{2}\left(\dot{y}_{1}, z_{2}^{1}\right) + \frac{\omega^{-1}}{4\mathrm{i}}D_{2}^{3}g_{2}\left(z_{2}^{1}, z_{2}^{1}, z_{2}^{-1}\right)
+ \frac{\omega^{-3}}{2\mathrm{i}}\left\{-D_{1}g_{2}D_{2}g_{1}z_{2}^{1} + D_{2}^{2}g_{2}\left(z_{2}^{1}, g_{2}\right) - \frac{1}{4}D_{1}D_{2}g_{2}\left(g_{1}, z_{2}^{1}\right)
- \frac{1}{4}D_{1}^{2}D_{2}g_{2}\left(\dot{y}_{1}, \dot{y}_{1}, z_{2}^{1}\right) + \frac{1}{4}D_{2}g_{2}\left(D_{2}g_{2}z_{2}^{1}\right)\right\},$$
(4.5b)

insert the solution into the algebraic relations

$$\begin{split} y_2 &= z_2^0 = \omega^{-2} g_2 + \omega^{-2} D_2^2 g_2 \left(z_2^1, z_2^1 \right), \\ z_1^1 &= -\omega^{-2} D_2 g_1 z_2^1 - 2 \mathrm{i} \omega^{-3} D_1 D_2 g_1 \left(\dot{y}_1, z_2^1 \right), \\ z_1^2 &= -\frac{\omega^{-2}}{8} D_2^2 g_1 \left(z_2^1, z_2^1 \right), \\ z_2^2 &= -\frac{\omega^{-2}}{6} D_2^2 g_2 \left(z_2^1, z_2^1 \right), \end{split}$$

and with this, compute the truncated MFE

$$x_{\star}(t) := y(t) + e^{i\omega t}z^{1}(t) + e^{-i\omega t}\bar{z}^{1}(t) + e^{2i\omega t}z^{2}(t) + e^{-2i\omega t}\bar{z}^{2}(t).$$

5. Numerical experiments

We conclude this paper with numerical experiments on the modified Fermi–Pasta–Ulam problem and on a one dimensional model of a diatomic gas in order to illustrate the accuracy as well as the energy conservation properties of the proposed integrators.

5.1. The modified Fermi-Pasta-Ulam problem

Let us consider the modified Fermi–Pasta–Ulam problem, as described in ([16], Sect. I.5.1), *i.e.* with Hamiltonian

$$H(p_1, p_2, q_1, q_2) = \frac{1}{2} \sum_{i=1}^{3} (p_{1,i}^2 + p_{2,i}^2) + \frac{\omega^2}{2} \sum_{i=1}^{3} q_{2,i}^2 + U(q_1, q_2),$$

where

$$U(q_1, q_2) = \frac{1}{4} \left\{ (q_{1,1} - q_{2,1})^4 + \sum_{i=1}^2 (q_{1,i+1} - q_{2,i+1} - q_{1,i} - q_{2,i})^4 + (q_{1,3} + q_{2,3})^4 \right\}.$$

The total energy H is exactly conserved and this problem also possesses an almost-invariant, namely the oscillatory energy

$$I(p_2, q_2) = \frac{1}{2} \sum_{i=1}^{3} p_{2,i}^2 + \frac{\omega^2}{2} \sum_{i=1}^{3} q_{2,i}^2.$$

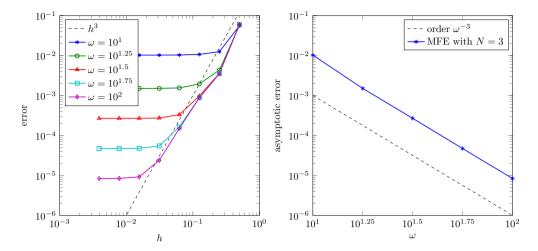


FIGURE 1. The errors of the numerical approximations using the terms of the MFE up to asymptotic order N=3 solved with a Runge–Kutta scheme of time stepping order p=3 for the modified Fermi–Pasta–Ulam problem is plotted versus the time stepsize (left) and the parameter ω (right). The reference lines have slope 3. In the left picture, the different lines correspond to different values of ω between 10 and 100 growing from top to bottom.

Indeed, this quantity is nearly preserved over times that are exponentially long in the high frequency ω along the exact solution of the modified Fermi–Pasta–Ulam problem [6]. The initial values for the second order differential equation (1.1) (using the notations $q = (q_1, q_2) = (x_1, x_2) = x$ and similarly for the momenta p) obtained with the above Hamiltonian function are given by

$$q_{1,1}(0) = 1$$
, $p_{1,1}(0) = 1$, $q_{2,1}(0) = 1/\omega$, $p_{2,1}(0) = 1$

and zero for the remaining ones. These initial values satisfy the bounded energy assumption (1.3).

5.2. Numerical study of the asymptotic and time stepping order

In order to demonstrate the behaviour of the error of our numerical method, we exemplary consider the system (4.3) of asymptotic order N=3 for the modified Fermi–Pasta–Ulam problem and solve it with an explicit Runge–Kutta method of time stepping order p=3. For different values of ω we compute the solution up to time $T_{\rm end}=1$ for varying time stepsizes h and measure the relative errors at $T_{\rm end}$ using a high resolution simulation of the original system (1.1) solved with the Störmer–Verlet scheme as reference solution. In the left part of Figure 1, the error is plotted versus the time stepsize h. It shows third order convergence in the time stepsize up to the point, where the asymptotic error dominates. Note that the error curves are almost identical for the different values of ω before the saturation occurs. Therefore, the pre-asymptotic convergence in the time stepsize, namely the constant C_t , is indeed independent of ω . In the right part of Figure 1 the error is plotted against the values of ω always computed with time stepsizes small enough, that the error is dominated by the asymptotic part. We again observe the predicted scaling in ω^{-1} .

5.3. Numerical study of the energy conservation

Next, we use the numerical approximations to the solution $x_{\star}(t)$ of the truncated systems of asymptotic order N from one to four to compute the total energy H - 0.8 (for a better display), the oscillatory energy I together with its components on the time interval [0,100] for a fixed (large) time stepsize $h = 10^{-2}$. The nonstiff differential equations are all solved with an explicit Runge–Kutta scheme of time stepping order p = 3. The results are shown in Figure 2 along with the energies computed by the Störmer–Verlet method applied to

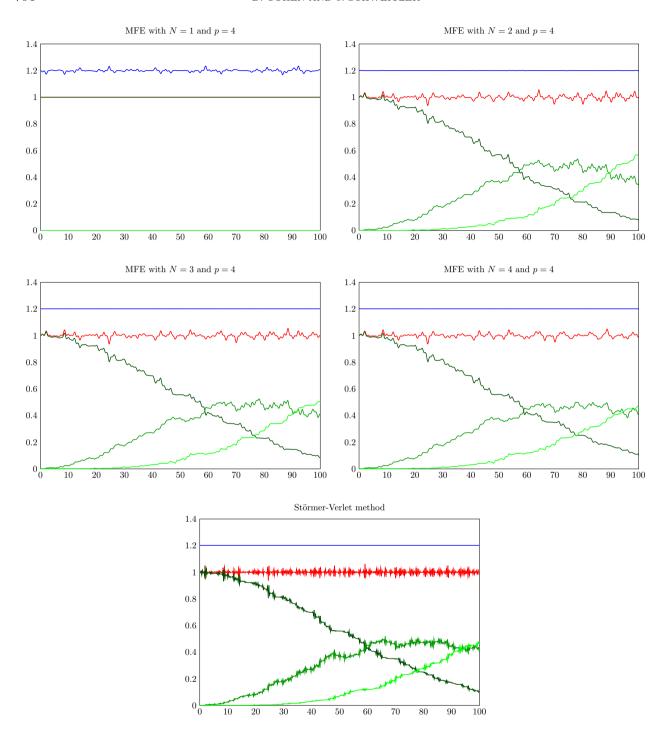


FIGURE 2. Exchange of energy between stiff springs for the modified Fermi–Pasta–Ulam problem solved using the MFE with different asymptotic orders and the Störmer–Verlet method. Blue: total energy H-0.8. Red: oscillatory energy I. Shades of green: components of oscillatory energy I_i . (In color online)

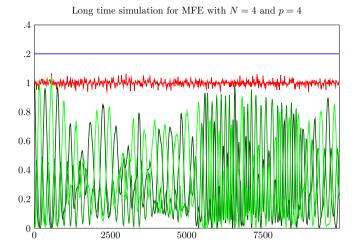


FIGURE 3. Exchange of energy between stiff springs for the modified Fermi–Pasta–Ulam problem on a long time interval solved with the method of asymptotic order N=4 and time stepping order p=4. Blue: total energy H-0.8. Red: oscillatory energy I. Shades of green: components of oscillatory energy I_i . (In color online)

the original system (1.1) with a small stepsize $h = 10^{-4}$. One can observe that the modelling of the exchange of energy gets better by increasing the number of modulation functions present in the method, namely the asymptotic order N.

A numerical experiment with the same parameters as above but on a much longer time interval $[0, 10\,000]$ is presented in Figure 3. Excellent energy preservation is observed for the numerical solution of the truncated system (4.5) of asymptotic order N=4 solved with an explicit Runge–Kutta method of time stepping order p=4.

To conclude this subsection, we compute the maximum error of the total energy H given in (1.2) on the interval [0, 100] as a function of $h\omega$ for a fixed stepsize h = 0.02. Figure 4 (left column) displays the results for our method of asymptotic as well as time stepping order N = p = 3 and the trigonometric integrator ([17] and method (D) in ([16], Chap. XIII))

$$x_{n+1} - 2\cos(h\Omega)x_n + x_{n-1} = h^2\Psi g(\Phi x_n)$$
 (5.1)

with filter functions $\Psi(\xi) = \text{sinc}^2(\xi/2)$ and $\Phi(\xi) = \text{sinc}(\xi)(1 + 1/3\sin^2(\xi/2))$, where $\text{sinc}(\xi) = \sin(\xi)/\xi$. No resonances are observed for the numerical integrator based on the MFE except in the regime $h\omega$ close to zero. But this regime is not of interest in the present article. Similar results are obtained for the maximum error of the oscillatory energy I along numerical solutions given by the proposed methods in the left column of Figure 4. Finally, we would like to comment on the fact that particular choices of the filter functions present in (5.1) could also lead to trigonometric integrators that do not suffer from resonances in the oscillatory and/or the total energy of the problem as seen in the last two rows of Figure 4. These plots were obtained using method (E) in ([16], Chap. XIII), that is with filter functions $\Psi(\xi) = \text{sinc}^2(\xi)$ and $\Phi(\xi) = 1$ and method (G) in [14], that is with filter functions $\Psi(\xi) = \text{sinc}(\xi)$. Similar results are also obtained with the modified trigonometric integrators presented in [20]. To make this comparison with the trigonometric methods even clearer, we also present these errors in a logarithmic scale in Figure 5.

5.4. Runtime comparison

Finally, we show some runtime comparisons in Figure 6. To that end, we solve the modified Fermi-Pasta-Ulam problem with $\omega = 500$ on the time interval [0, 10] with the trigonometric method (E) as well as the

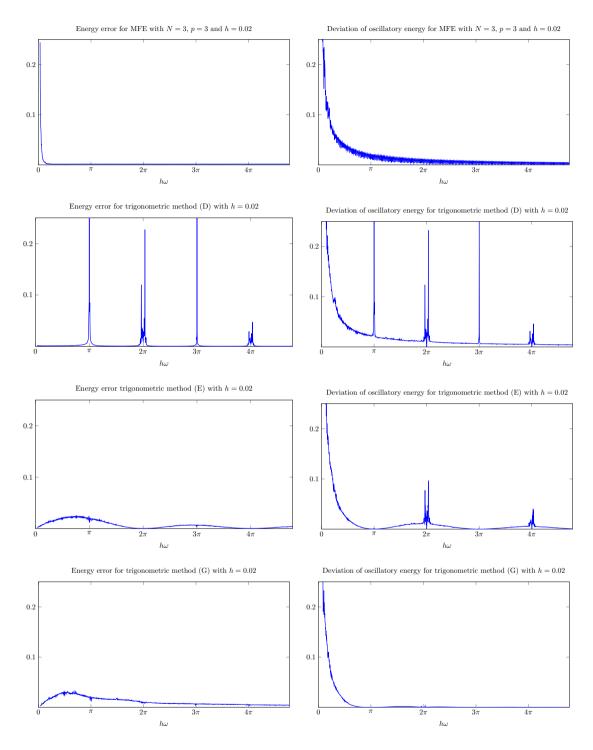


FIGURE 4. Maximum error of the total energy H (left) and maximum deviation of the oscillatory energy I (right) on the interval [0, 100] for MFE with N = p = 3 and trigonometric methods (D), (E) and (G).

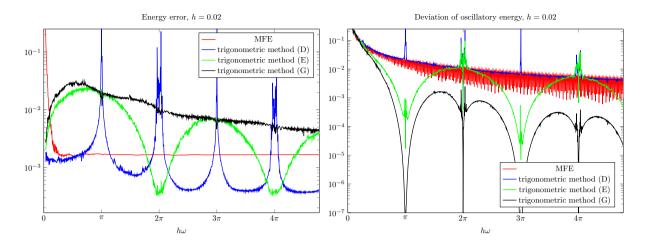


FIGURE 5. Maximum error of the total energy H (left) and maximum deviation of the oscillatory energy I (right) on the interval [0, 100] for MFE with N = p = 3 and trigonometric methods (D), (E) and (G) in logarithmic scale.

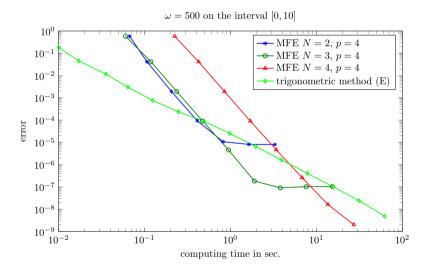


FIGURE 6. Runtime comparison between the MFE-based methods with N=2,3,4 and p=4 and the trigonometric method (E).

MFE-based methods of asymptotic order N=2,3,4 using an explicit Runge–Kutta method of time stepping order p=4. Each method uses a constant time stepsize and is applied several times for different time stepsizes along its corresponding curves ($h=2^{-k}$, MFE N=2: $k=1,\ldots,7$, MFE N=3: $k=1,\ldots,9$, MFE N=4: $k=1,\ldots,8$, Trigonometric: $k=2,\ldots,15$). For the reference solution we solved the original highly oscillatory system with the trigonometric method (E) using a very fine time stepsize $h_{\rm ref}=2^{-16}$.

It can be seen that if the time stepsize is chosen such that the time stepping error matches the asymptotic error there is some potential in the MFE methods to beat the trigonometric method, which is the state of the

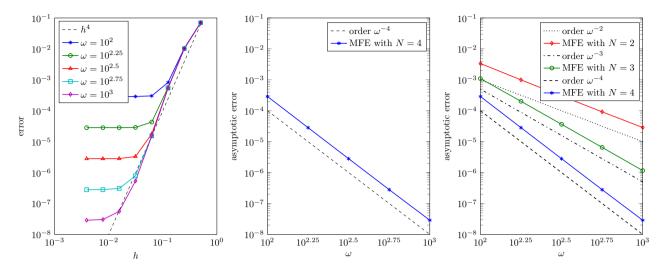


FIGURE 7. The errors of the numerical approximations using the terms of the MFE up to asymptotic order N=4 solved with a Runge–Kutta scheme of time stepping order p=4 for the diatomic model using a Lennard–Jones potential is plotted *versus* the time stepsize (left) and the parameter ω (middle). The reference lines have slope 4. In the left picture, the different lines correspond to different values of ω between 100 and 1000 growing from top to bottom. Asymptotic orders for the MFE systems with N=2,3,4 (right).

art numerical scheme for this kind of problem. Note that a further reduction of the time stepsize is inefficient, since the error is dominated by the asymptotic error and thus cannot be improved any more.

From Figure 6 it can also be observed, that the method with asymptotic order N=3 is the most efficient choice, since the system to solve is hardly more complicated than the system for the method of asymptotic order N=2 where the latter lacks accuracy.

5.5. One dimensional model of a diatomic gas with short-range interaction forces

Our next example describes a one dimensional model of a diatomic gas with short-range interaction forces, see [2,5]. The Hamiltonian for four diatomic molecules reads

$$H(x,\zeta,p,\pi) = \frac{1}{2} \sum_{l=1}^{2} \left(p_l^2 + \pi_l^2 + \omega^2 \zeta_l^2 \right) + \sum_{l=1}^{3} U \left(x_l + \zeta_l - x_{l-1} - \zeta_{l-1} \right),$$

where the fixed end particles satisfies $x_0 = \zeta_0 = \zeta_3 = 0$ and $x_3 = 12$. We further consider a Lennard–Jones potential $U(s) = s^{-12} - s^{-6}$. As initial values, we choose $x = (4, 4 + 2^{1/6})$, $\zeta = (1/\omega, 1/100\,000)$, p = (0, 1.1) and $\pi = (1.2, 0.1)$.

We first consider the various errors of the proposed numerical methods. To do this, we consider the MFE system of asymptotic order N=4 for the above problem and solve it with an explicit Runge–Kutta method of time stepping order p=4. For different values of ω we compute the solution up to time $T_{\rm end}=1$ for varying time stepsizes h and measure the relative errors at $T_{\rm end}$ using a high resolution simulation of the original system (1.1) solved with the Störmer–Verlet scheme with a stepsize $h_{\rm ref}=10^{-6}$ as reference solution. The expected convergence rates from Proposition 3.1 can be observed in Figure 7 (left and middle pictures). Next, in order to observe the asymptotic orders of the numerical methods presented in the previous section, we solve the MFE systems for N=2,3 and N=4. The corresponding orders of convergence can be observed in the right picture of Figure 7.

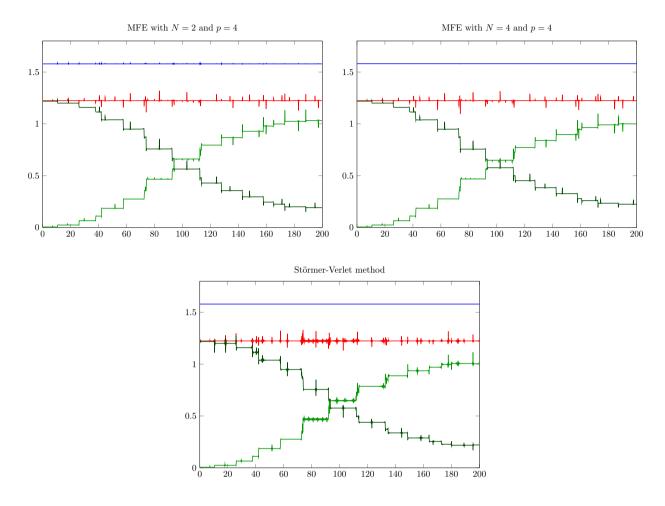


FIGURE 8. Time evolution of the different energies present in the diatomic model with a Lennard–Jones potential solved using the MFE with different asymptotic orders and the Störmer–Verlet method. Blue: total energy H. Red: oscillatory energy I. Shades of green: components of oscillatory energy I_i . (In color online)

Finally, we plot the different energies present in this diatomic gas model along the numerical solutions given by the proposed numerical methods using a time stepsize $h = 10^{-2}$. The reference energies are given by the Störmer–Verlet method with stepsize $h_{\rm ref} = 10^{-6}$. One can observe in Figure 8 that also for this example the exchange of energies, the conservation of the total and the almost-conservation of the oscillatory energy is very well described by the MFE methods.

6. Conclusion and further research

Using a modulated Fourier expansion, we have derived high order numerical algorithms for the time discretisation of second-order differential equations with highly oscillatory solutions. To do this, only a finite energy assumption on the initial values of the problem as well as smoothness of the nonlinearity present in the equation were used. Furthermore, these new numerical methods do not suffer from stepsize restrictions imposed by the high frequency of the problem. Finally, our convergence results and the main properties of these new high

order numerical methods were illustrated on the modified Fermi–Pasta–Ulam problem and on a model of a diatomic gas.

We would like to mention that the techniques presented here could be used to derive high order numerical methods for the discretisation of highly-oscillatory systems with multiple frequencies (as those studied in [7]) or for the numerical discretisation of Hamiltonian PDEs (using the MFE presented in [13] for example). This is however not an easy task (if possible at all). Indeed, the main challenges that need to be overcome are

- (i) The absence of a clear separation of the different scales in such problems as opposed to problems of the form (1.1) as studied in the present publication.
- (ii) The absence of a result such as Lemma 2.2 for such problems.

We will investigate these questions in the future.

Acknowledgements. We greatly appreciate the referees's comments on earlier versions. Furthermore, we thank Ludwig Gauckler, Volker Grimm, Marlis Hochbruck, Tobias Jahnke and Christian Lubich for interesting discussions.

This work was supported by UMIT Research Lab at Umeå University, the Lars Hierta Memorial Foundation and the German Research Foundation (DFG) via GRK 1294.

References

- [1] W. Bao, X. Dong and X. Zhao, Uniformly accurate multiscale time integrators for highly oscillatory second order differential equations. J. Math. Study 47 (2014) 111–150.
- [2] G. Benettin, L. Galgani and A. Giorgilli, The dynamical foundations of classical statistical mechanics and the Boltzmann-Jeans conjecture. In Seminar on Dynamical Systems (St. Petersburg, 1991). Vol. 12 of *Progr. Nonlin. Differ. Eq. Appl.* Birkhäuser, Basel (1994) 3–14.
- [3] F. Castella, P. Chartier and E. Faou, An averaging technique for highly oscillatory Hamiltonian problems. SIAM J. Numer. Anal. 47 (2009) 2808–2837.
- [4] P. Chartier, A. Murua and J.M. Sanz-Serna, Higher-order averaging, formal series and numerical integration I: B-series. Found. Comput. Math. 10 (2010) 695–727.
- [5] D. Cohen, Analysis and Numerical Treatment of Highly Oscillatory Differential Equations. Ph.D thesis, University of Geneva (2004).
- [6] D. Cohen, E. Hairer and Ch. Lubich, Modulated Fourier expansions of highly oscillatory differential equations. Found. Comput. Math. 3 (2003) 327–345.
- [7] D. Cohen, E. Hairer and Ch. Lubich, Numerical energy conservation for multi-frequency oscillatory differential equations. BIT 45 (2005) 287–305.
- [8] D. Cohen, T. Jahnke, K. Lorenz and Ch. Lubich, Numerical integrators for highly oscillatory Hamiltonian systems: a review. In Analysis, modeling and simulation of multiscale problems. Springer, Berlin (2006) 553–576.
- [9] M. Condon, A. Deaño and A. Iserles, On second-order differential equations with highly oscillatory forcing terms. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **466** (2010) 1809–1828.
- [10] M. Condon, A. Deaño and A. Iserles, On systems of differential equations with extrinsic oscillation. Discrete Contin. Dyn. Syst. 28 (2010) 1345–1367.
- [11] M. Condon, A. Deaño and A. Iserles, Asymptotic solvers for oscillatory systems of differential equations. \vec{SeMA} J. (2011) 79–101
- [12] E. Faou and K. Schratz, Asymptotic preserving schemes for the Klein–Gordon equation in the non-relativistic limit regime. Numerische Mathematik 126 (2013) 441–469.
- [13] L. Gauckler, Long-time analysis of Hamiltonian partial differential equations and their discretizations. PhD. thesis, Universität Tübingen (2010). http://nbn-resolving.de/urn:nbn:de:bsz:21-opus-47540.
- [14] V. Grimm and M. Hochbruck, Error analysis of exponential integrators for oscillatory second-order differential equations. J. Phys. A 39 (2006) 5495–5507.
- [15] E. Hairer and Ch. Lubich, Long-time energy conservation of numerical methods for oscillatory differential equations. SIAM J. Numer. Anal. 38 (electronic) (2000) 414–441.
- [16] E. Hairer, Ch. Lubich and G. Wanner, Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations. Vol. 31 of Springer Ser. Comput. Math. Springer, Berlin (2002).
- [17] M. Hochbruck and Ch. Lubich, A Gautschi-type method for oscillatory second-order differential equations. Numer. Math. 83 (1999) 403–426.
- [18] M. Hochbruck and A. Ostermann, Exponential integrators. Acta Numerica 19 (2010) 209-286.

- [19] C. Le Bris and F. Legoll, Integrators for highly oscillatory Hamiltonian systems: an homogenization approach. Discrete Contin. Dyn. Syst. Ser. B 13 (2010) 347–373.
- [20] R.I. McLachlan and A. Stern, Modified trigonometric integrators. SIAM J. Numer. Anal. 52 (2014) 1378–1397.
- [21] A. Stern and E. Grinspun, Implicit-explicit variational integration of highly oscillatory problems. Multiscale Model. Simul. 7 (2009) 1779–1794.
- [22] M. Tao, H. Owhadi and J.E. Marsden, Nonintrusive and structure preserving multiscale integration of stiff ODEs, SDEs, and Hamiltonian systems with hidden slow dynamics via flow averaging. *Multiscale Model. Simul.* 8 (2010) 1269–1324.
- [23] B. Wang and X. Wu, A new high precision energy-preserving integrator for system of oscillatory second-order differential equations. *Phys. Lett. A* **376** (2012) 1185–1190.