Dynamics, numerical analysis, and some geometry

Ludwig Gauckler * Ernst Hairer † Christian Lubich ‡

March 3, 2018

Abstract

Geometric aspects play an important role in the construction and analysis of structure-preserving numerical methods for a wide variety of ordinary and partial differential equations. Here we review the development and theory of symplectic integrators for Hamiltonian ordinary and partial differential equations, of dynamical low-rank approximation of time-dependent large matrices and tensors, and its use in numerical integrators for Hamiltonian tensor network approximations in quantum dynamics.

1 Introduction

It has become a commonplace notion in all of numerical analysis (which here is understood as comprising the construction and the mathematical analysis of numerical algorithms) that a good algorithm should "respect the structure of the problem" — and in many cases the "structure" is of *geometric* nature. This immediately leads to two basic questions, which need to be answered specifically for each problem:

- How can numerical methods be constructed that "respect the geometry" of the problem at hand?
- What are benefts from using a structure-preserving algorithm for this problem, and how do they come about?

In this note we present results in the numerical analysis of *dynamic* (evolutionary, time-dependent) ordinary and partial differential equations for which geometric aspects play an important role. These results belong to the area that has

*Institut für Mathematik, FU Berlin, Arnimallee 9, D-14195 Berlin, Germany (gauckler@math.fu-berlin.de).

†Section de mathématiques, 2-4 rue du Lièvre, Université de Genève, CH-1211 Genève 4, Switzerland (Ernst.Hairer@unige.ch).

‡Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle 10, D-72076 Tübingen, Germany (lubich@na.uni-tuebingen.de).

become known as *Geometric Numerical Integration*, which has developed vividly in the past quarter-century, with substantial contributions by researchers with very different mathematical backgrounds. We just refer to the books (in chronological order) [SSC94, HLW02, Sur03, LR04, HLW06, Lub08, FQ10, Fao12, WYW13, BC16] and to the Acta Numerica review articles [SS92, IMKNZ00, MW01, MQ02, HLW03, DDE05, BL07, Chu08, HO10, Wan10, CMKO11, AEEVE12, DE13]. In this note we restrict ourselves to some selected topics to which we have contributed.

In Section 2 we begin with reviewing numerical methods for approximately solving *Hamiltonian systems* of ordinary differential equations, which are ubiquitous in many areas of physics. Such systems are characterized by the *symplecticity* of the flow, a geometric property that one would like to transfer to the numerical discretization, which is then called a *symplectic integrator*. Here, the two questions above become the following:

- How are symplectic integrators constructed?
- What are favourable long-time properties of symplectic integrators, and how can they be explained?

The first question relates numerical methods with the theories of Hamilton and Jacobi from the mid-19th century, and the latter question connects numerical methods with the analytical techniques of Hamiltonian perturbation theory, a subject developed from the late 19th throughout the 20th century, from Lindstedt and Poincaré and Birkhoff to Siegel and Kolmogorov, Arnold and Moser (KAM theory), to Nekhoroshev and further eminent mathematicians. This connection comes about via *backward error analysis*, which is a concept that first appeared in numerical linear algebra [Wil60]. The viewpoint is to interpret the numerical approximation as the exact (or almost exact) solution of a modified equation. In the case of a symplectic integrator applied to a Hamiltonian differential equation, the modified differential equation turns out to be again Hamiltonian, with a Hamiltonian that is a small perturbation to the original one. This brings Hamiltonian perturbation theory into play for the long-time analysis of symplectic integrators. Beyond the purely mathematical aspects, it should be kept in mind that symplectic integrators are first and foremost an important tool in computational physics. In fact such numerical methods appeared first in the physics literature [dV56, Ver67, Rut83], in such areas as nuclear physics and molecular dynamics, and slightly later [WH91] in celestial mechanics, which has been the original motivation in the development of Hamiltonian perturbation theory [Poi92, SM71]. It was not least with the use of symplectic integrators that the centuries-old question about the stability of the solar system was finally answered negatively in the last decade by Laskar; see [Las13] and compare also with [Mos78].

In Section 3 we consider numerical methods for finite-dimensional *Hamiltonian systems with multiple time scales* where, in the words of Fermi, Pasta & Ulam [FPU55], "the non-linearity is introduced as a perturbation to a primarily linear problem. The behavior of the systems is to be studied for times which are

long compared to the characteristic periods of the corresponding linear problem." The two basic questions above are reconsidered for such systems. Except for unrealistically small time steps, the backward error analysis of Section 2 does not work for such systems, and a different technique of analysis is required. *Modulated Fourier expansions* in time were originally developed (since 2000) for studying numerical methods for such systems and were subsequently also recognized as a powerful analytical technique for proving new results for continuous systems of this type, including the original Fermi–Pasta–Ulam system. While the canonical transformations of Hamiltonian perturbation theory transform the system into a normal form from which long-time behaviour can be read off, modulated Fourier expansions embed the system into a high-dimensional system that has a Lagrangian structure with invariance properties that enable us to infer long-time properties of the original system. Modulated Fourier expansions do not use nonlinear coordinate transformations, which is one reason for their suitability for studying numerical methods, which are most often not invariant under nonlinear transformations.

In Section 4 we present long-time results for suitable numerical discretizations of *Hamiltonian partial differential equations* such as nonlinear wave equations and nonlinear Schrödinger equations. A number of important results on this topic have been obtained in the last decade, linking the numerical analysis of such equations to recent advances in their mathematical analysis. The viewpoint we take here is to consider the Hamiltonian partial differential equation as an infinite-dimensional system of the oscillatory type of Section 3 with infinitely many frequencies, and we present results on the long-time behaviour of the numerical and the exact solutions that have been obtained with modulated Fourier expansions or with techniques from infinite-dimensional Hamiltonian perturbation theory. We mention, however, that there exist other viewpoints on the equations considered, with different geometric concepts such as multisymplecticity [Bri97, MPS98]. While multisymplectic integrators, which preserve this geometric structure, have been constructed and favourably tested in numerical experiments [BR01, AM04] (and many works thereafter), as of now there appear to be no proven results on the long-time behaviour of such methods.

In Section 5 we consider *dynamical low-rank approximation*, which leads to a different class of dynamical problems with different geometric aspects. The problem here is to approximate large (or rather too large, huge) time-dependent matrices, which may be given explicitly or are the unknown solution to a matrix differential equation, by matrices of a prescribed rank, typically much smaller than the matrix dimension so that a data-compressed approximation is obtained. Such problems of data and/or model reduction arise in a wide variety of applications ranging from information retrieval to quantum dynamics. On projecting the time derivative of the matrices to the tangent space of the manifold of low-rank matrices at the current approximation, this problem leads to a differential equation on the low-rank manifold, which needs to be solved numerically. We present answers to the two basic questions formulated at the beginning of this introduction, for this particular problem. The proposed "geometric" numerical integrator, which is based on splitting the orthogonal projection onto the tan-

gent space, is robust to the (ubiquitous) presence of small singular values in the approximations. This numerically important robustness property relies on a geometric property: The low-rank manifold is a ruled manifold (like a hyperboloid). It contains fat subspaces along which one can pass between any two points on the manifold, and the numerical integrator does just that. In this way the high curvature of the low-rank manifold at matrices with small singular values does not become harmful. Finally, we address the nontrivial extension to tensors of various formats (Tucker tensors, tensor trains, hierarchical tensors), which is of interest in time-dependent problems with several spatial dimensions.

Section 6 on tensor and tensor network approximations in *quantum dynamics* combines the worlds of the previous two sections and connects them with recent developments in computational quantum physics. The reduction of the time-dependent many-particle Schrödinger equation to a low-rank tensor manifold by the Dirac–Frenkel time-dependent variational principle uses a tangent-space projection that is both orthogonal and symplectic. It results in a (non-canonical) Hamiltonian differential equation on the tensor manifold that can be discretized in time by the projector-splitting integrator of Section 5, which is robust to small singular values and preserves both the norm and the energy of the wavefunction.

2 Hamiltonian systems of ordinary differential equations

2.1 Hamiltonian systems

Differential equations of the form (with $\dot{} = \mathrm{d}/\mathrm{d}t$)

$$\dot{p} = - \quad {}_q H(p,q), \quad \dot{q} = + \quad {}_p H(p,q) \tag{2.1}$$

are fundamental to many branches of physics. The real-valued Hamilton function H, defned on a domain of $\mathbb{R}^{d+d}$ (the phase space), represents the total energy and $q(t) \in \mathbb{R}^d$ and $p(t) \in \mathbb{R}^d$ represent the positions and momenta, respectively, of a conservative system at time t. The total energy is conserved:

$$H(p(t),q(t)) = H(p(0),q(0))$$

along every solution $(p(t),q(t))$ of the Hamiltonian differential equations.

Numerical example: We consider four variants of the Euler method, which for a given (small) step size $h > 0$ compute approximations $p_n \approx p(nh)$, $q_n \approx q(nh)$ via

$$p_{n+1} = p_n - h \quad {}_q H(p_{n+\alpha}, q_{n+\beta}), \qquad q_{n+1} = q_n + h \quad {}_p H(p_{n+\alpha}, q_{n+\beta}),$$

with $\alpha, \beta \in 0,1$. For $\alpha = \beta = 0$ this is the explicit Euler method, for $\alpha = \beta = 1$ it is the implicit Euler method. The partitioned methods with $\alpha = \beta$ are known as the *symplectic Euler methods*. All four methods are of order $r = 1$, that is, the error after one step of the method is (h^{r+1}) with $r = 1$.

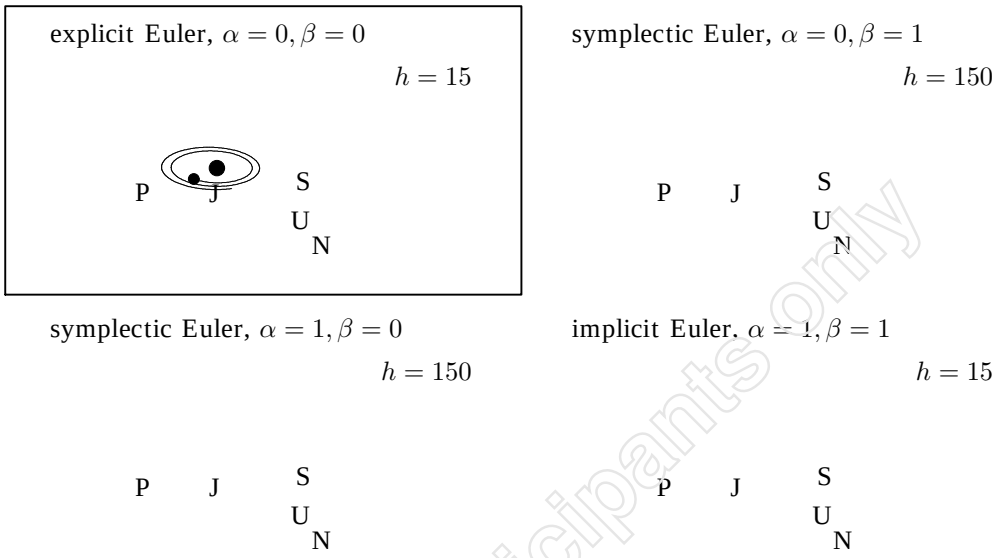

Figure 2.1: Numerical simulation of the outer solar system.

We apply these methods to the outer solar system, which is an N-body problem with Hamiltonian

$$H(p,q) = \frac{1}{2}\sum_{i=0}^{N} \frac{1}{m_i} \; p^i{}^2 - G\sum_{i=1}^{N}\sum_{j=0}^{i-1} \frac{m_i m_j}{q^i - q^j},$$

where $p = (p^0, \ldots, p^N)$, $q = (q^0, \ldots, q^N)$ and $\cdot$ denotes the Euclidean norm, and the constants are taken from [HLW06, Section I.2.4]. The positions $q^i \in \mathbb{R}^3$ and momenta $p^i \in \mathbb{R}^3$ are those of the sun and the five outer planets (including Pluto). Figure 2.1 shows the numerical solution obtained by the four versions of the Euler method on a time interval of 200 000 earth days. For the explicit Euler method the planets spiral outwards, for the implicit Euler method they spiral inwards, fall into the sun and finally are ejected. Both symplectic Euler methods show a qualitatively correct behaviour, even with a step size (in days) that is much larger than the one used for the explicit and implicit Euler methods. Figure 2.2 shows the relative error of the Hamiltonian, $\big(H(p_n,q_n) - H(p_0,q_0)\big)\big/\, H(p_0,q_0)$, along the numerical solution of the four versions of Euler's method on the time interval $0 \le nh \le 200\,000$. Whereas the size of the error increases for the explicit and implicit Euler methods, it remains bounded and small, of a size proportional to the step size h, for both symplectic Euler methods.

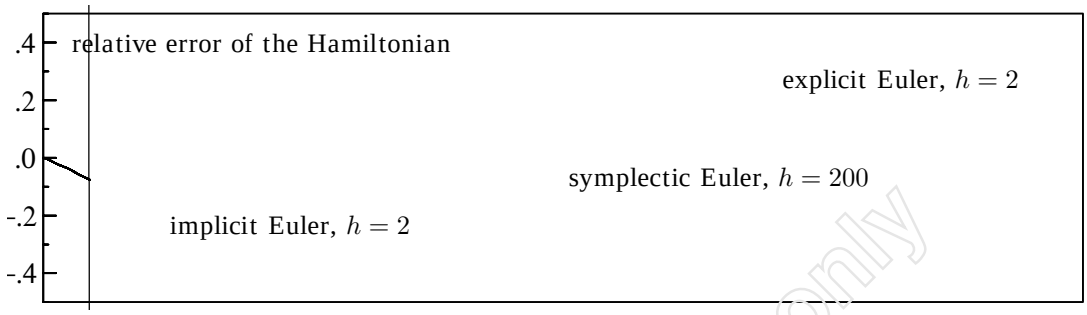

Figure 2.2: Relative error of the Hamiltonian on the interval $0 \le t \le 200\,000$.

2.2 Symplecticity of the flow and symplectic integrators

The time-t *flow* of a differential equation $\dot{y} = f(y)$ is the map φ_t that associates with an initial value y_0 at time 0 the solution value at time t: $\varphi_t(y_0) = y(t)$. Consider now the Hamiltonian system (2.1) or equivalently, for $y = (p, q)$,

$$\dot{y} = J^{-1} \quad H(y) \quad \text{with} \quad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}.$$

The fow φ_t of the Hamiltonian system is *symplectic* (or *canonical*), that is, the derivative matrix $D\varphi_t$ with respect to the initial value satisfes

$$D\varphi_t(y)^\top \, J \, D\varphi_t(y) = J$$

for all y and t for which $\varphi_t(y)$ exists. This quadratic relation is formally similar to orthogonality, with J in place of the identity matrix I, but it is related to the preservation of areas rather than lengths in phase space.

There is also a local converse: If the fow of some differential equation is symplectic, then there exists locally a Hamilton function for which the corresponding Hamiltonian system coincides with this differential equation.

A numerical one-step method $y_{n+1} = \Phi_h(y_n)$ (with step size h) is called symplectic if the numerical fow Φ_h is a symplectic map:

$$D\Phi_h(y)^\top \, J \, D\Phi_h(y) = J.$$

Such methods exist: the "symplectic Euler methods" of the previous subsection are indeed symplectic. This was frst noted, or considered noteworthy, in an unpublished report by de Vogelaere [dV56]. The symplecticity can be readily verifed by direct calculation or by observing that the symplectic Euler methods are symplectic maps with the h-scaled Hamilton function taken as the generating function of a canonical transformation in Hamilton and Jacobi's theory. More than 25 years later, Ruth [Rut83] and Feng Kang [Fen85, Fen86] independently constructed higher-order symplectic integrators using generating functions of Hamilton–Jacobi theory. These symplectic methods require, however, higher derivatives of the Hamilton function. Symplectic integrators began

to find widespread interest in numerical analysis when in 1988 Lasagni, Sanz-Serna and Suris [Las88, SS88, Sur88] independently characterized symplectic Runge–Kutta methods by a quadratic relation of the method coefficients. This relation was already known to be satisfied by the class of Gauss–Butcher methods (the order-preserving extension of Gaussian quadrature formulae to differential equations), which include methods of arbitrary order. Like the Euler methods, Runge–Kutta methods only require evaluations of the vector field, but no higher derivatives.

The standard integrator of molecular dynamics, introduced to the field by Verlet in 1967 [Ver67] and used ever since, is also symplectic. For a Hamiltonian $H(p,q) = \frac{1}{2}p^\top M^{-1}p + V(q)$ with a symmetric positive definite mass matrix M, the method is explicit and given by the formulas

$$\begin{aligned} p_{n+1/2} &= p_n - \tfrac{h}{2} V(q_n) \\ q_{n+1} &= q_n + hM^{-1}p_{n+1/2} \\ p_{n+1} &= p_{n+1/2} - \tfrac{h}{2} V(q_{n+1}). \end{aligned}$$

Such a method was also formulated by the astronomer Störmer in 1907, and can even be traced back to Newton's *Principia* from 1687, where it was used as a theoretical tool in the proof of the preservation of angular momentum in the two-body problem (Kepler's second law), which is preserved by this method (cf. [Wan10]). The above method is referred to as the Störmer–Verlet method, Verlet method or leapfrog method in different communities. The symplecticity of this method can be understood in various ways by relating the method to classes of methods that have proven useful in a variety of applications (cf. [HLW03]): as a *composition method* (it is a composition of the two symplectic Euler methods with half step size), as a *splitting method* (it solves in an alternating way the Hamiltonian differential equations corresponding to the kinetic energy $\frac{1}{2}p^\top M^{-1}p$ and the potential energy $V(q)$), • and as a *variational integrator*: it minimizes the discrete action functional that results from approximating the action integral

$$\int_{t_0}^{t_N} L(q(t), \dot{q}(t))\,\mathrm{d}t \quad \text{with} \quad L(q,\dot{q}) = \frac{1}{2}\dot{q}^\top M\dot{q} - V(q)$$

by the trapezoidal rule and using piecewise linear approximation to $q(t)$. The Störmer–Verlet method can thus be interpreted as resulting from a discretization of the Hamilton variational principle. Such an interpretation can in fact be given for every symplectic method. Conversely, symplectic methods can be *constructed* by minimizing a discrete action integral. In particular, approximating the action integral by a quadrature formula and the positions $q(t)$ by a piecewise polynomial leads to a symplectic partitioned Runge–Kutta method. With Gauss quadrature, this gives a reinterpretation of the Gauss–Butcher methods (cf. [Sur90, MW01, HLW06]).

2.3 Backward error analysis

The numerical example of Section 2.1, and many more examples in the literature, show that symplectic integrators behave much better over long times than their non-symplectic counterparts. How can this be explained, or put differently: How does the geometry lead to favourable dynamics? There is a caveat: As was noted early on [GDC91, CSS92], all the benefts of symplectic integrators are lost when they are used with variable step sizes as obtained by standard step size control. So it is not just about preserving symplecticity.

Much insight into this question is obtained from the viewpoint of *backward analysis*, where the result of one step of a numerical method for a differential equation $\dot{y} = f(y)$ is interpreted as the solution to a modifed differential equation (or more precisely formal solution, having the same expansion in powers of the step size h)

$$\dot{\widetilde{y}} = f(\widetilde{y}) + hf_1(\widetilde{y}) + h^2 f_2(\widetilde{y}) + h^3 f_3(\widetilde{y}) + \dots \ .$$

The question then is how geometric properties of the numerical method, such as symplecticity, are refected in the modifed differential equation. It turns out that in the case of a *symplectic* integrator applied to a Hamiltonian differential equation, each of the perturbation terms is a *Hamiltonian* vector feld, $f_j(y) = J^{-1} \ \ H_j(y)$ (at least locally, on simply connected domains). The formal construction was frst given by Moser [Mos68], where the problem of interpolating a near-identity symplectic map by a Hamiltonian fow was considered. For the important class of symplectic partitioned Runge–Kutta methods (which includes all the examples mentioned in Section 2.2), a different construction in [Hai94], using the theory of P-series and their associated trees, showed that the perturbation Hamiltonians H_j are indeed *global*, defned on the same domain on which the Hamilton function H is defned and smooth. Alternatively, this can also be shown using the explicit generating functions for symplectic partitioned Runge–Kutta methods as derived by Lasagni; see [HLW06, Sect. IX.3]. This global result is in particular important for studying the behaviour of symplectic integrators for near-integrable Hamiltonian systems, which are considered in neighbourhoods of tori. It allows us to bring the rich arsenal of Hamiltonian perturbation theory to bear on the long-time analysis of symplectic integrators.

The step from a formal theory (with the three dots at the end of the line) to rigorous estimates was taken by Benettin & Giorgilli [BG94] (see also [HL97, Rei99] and [HLW06, Chapter IX] for related later work), who showed that in the case of an analytic vector feld f, the result $y_1 = \Phi_h(y_0)$ of one step of the numerical method and the time-h fow $\widetilde{\varphi}_h(y_0)$ of the corresponding modifed differential equation, suitably truncated after $N \sim 1/h$ terms, difer by a term that is exponentially small in $1/h$:

$$\Phi_h(y_0) - \widetilde{\varphi}_h(y_0) \ \ \leq Ch\,\mathrm{e}^{-c/h},$$

uniformly for y_0 varying in a compact set. The constants C and c can be given explicitly. It turns out that c is inversely proportional to a local Lipschitz

2. o tegr and near- tegr Hamiltonian

ymplectic integrators en- e prxpert- whn ar in-

which dz noue e ar nou conju-
a symplectic d) rgumen a

the long-time behavious of numerical w sitlation where the producg of with step size is not so small con to
condition

constant L of f, and hence the estimate is meaningful only under the condition $hL \ll 1$. We note that in an oscillatory Hamiltonian system, L corresponds to the highest frequency in the system.

A different approach to constructing a modified Hamiltonian whose flow is exponentially close to the near-identity symplectic map is outlined by Neishtadt [Nei84], who exactly embeds the symplectic map into the flow of a non-autonomous Hamiltonian system with rapid oscillations and then uses averaging to obtain an autonomous modified Hamiltonian.

2.4 Long-time near-conservation of energy

The above results immediately explain the observed near-preservation of the total energy by symplectic integrators used with constant step size: Over each time step, and as long as the numerical solution stays in a fixed compact set, the Hamilton function $\widetilde{H}$ of the optimally truncated modified differential equation is almost conserved up to errors of size $(he^{-c/h})$. On writing $\widetilde{H}(y_n) - \widetilde{H}(y_0)$ as a telescoping sum and adding up the errors, we thus obtain

$$\widetilde{H}(y_n) - \widetilde{H}(y_0) = \quad (e^{-c/2h}) \quad \text{for} \quad nh \le e^{c/2h}.$$

For a symplectic method of order r, the modified Hamilton function $\widetilde{H}$ is (h^r) close to the original Hamilton function H, uniformly on compact sets, and so we have *near-conservation of energy over exponentially long times*:

$$H(y_n) - H(y_0) = \quad (h^r) \quad \text{for} \quad nh \le e^{c/2h}.$$

Symplecticity is, however, not necessary for good energy behaviour of a numerical method. First, the assumption can clearly be weakened to conjugate symplecticity, that is, the one-step method $y_{n+1} = \Phi_h(y_n)$ is such that $\Phi_h = \chi_h^{-1} \circ \Psi_h \circ \chi_h$ where the map Ψ_h is symplectic. But then, for some methods such as the Störmer–Verlet method, long-time near-conservation of energy can be proved by an argument that does not use symplecticity, but just the time-symmetry $\Phi_{-h} \circ \Phi_h = \mathrm{id}$ of the method [HLW03]. That proof is similar in spirit to proving the conservation of the energy $\frac{1}{2}p^\top M^{-1}p + V(q) = \frac{1}{2}\dot q^\top M \dot q + V(q)$ for the second-order differential equation $M\ddot q + \quad V(q) = 0$ by taking the inner product with $\dot q$ and noting that there results a total differential: $\frac{\mathrm{d}}{\mathrm{d}t}(\frac{1}{2}\dot q^\top M \dot q + V(q)) = 0$. This kind of argument can be extended to proving long-time near-conservation of energy and momentum for symmetric multiV b y wn g-tit

backward error analysis and the perturbation theory of integrable systems, a rich mathematical theory originally developed for problems of celestial mechanics [Poi92, SM71, AKN97].

A Hamiltonian system with the (real-analytic) Hamilton function $H(p,q)$ is called *integrable* if there exists a symplectic transformation $(p,q) = \psi(a,\theta)$ to *action-angle variables* (a,θ), defned for actions $a = (a_1,\ldots,a_d)$ in some open set of $\mathbb{R}^d$ and for angles θ on the d-dimensional torus $\mathbb{T}^d = (\theta_1,\ldots,\theta_d)$; $\theta_i \in \mathbb{R} \bmod 2\pi$, such that the Hamiltonian in these variables depends only on the actions:

$$H(p,q) = H(\psi(a,\theta)) = K(a).$$

In the action-angle variables, the equations of motion are simply $\dot{a} = 0, \dot{\theta} = \omega(a)$ with the *frequencies* $\omega = (\omega_1,\ldots,\omega_d)^T = {}_aK$. For every a, the torus $(a,\theta) : \theta \in \mathbb{T}^d$ is thus invariant under the fow. We express the actions and angles in terms of the original variables (p,q) via the inverse transform as

$$(a,\theta) = (I(p,q), \Theta(p,q))$$

and note that the components of $I = (I_1,\ldots,I_d)$ are frst integrals (conserved quantities) of the integrable system.

The efect of a small perturbation of an integrable system is well under control in subsets of the phase space where the frequencies ω satisfy Siegel's *diophantine condition*:

$$k \cdot \omega \geq \gamma \, k^{-\nu} \quad \text{for all } k \in \mathbb{Z}^d,\ k = 0,$$

$HKd\ell v$ I M / - rß • ÐM2M p + 7
for some positive constants γ and ν, with $k = \sum_i k_i$. For $\nu > d-1$, almost all frequencies (in the sense of Lebesgue meass ub5 Ans [

initial values and a Cantor set of step sizes this holds even perpetually, as the existence of invariant tori of the numerical integrator close to the invariant tori of the integrable system was shown by Shang [Sha99, Sha00].

The linear error growth persists when the symplectic integrator is applied to a *perturbed integrable system* $H(p,q) + \varepsilon G(p,q)$ with a perturbation parameter of size $\varepsilon = \quad (h^\alpha)$ for some positive exponent α. Perturbed integrable systems have KAM tori, i.e., deformations of the invariant tori of the integrable system corresponding to diophantine frequencies ω, which are invariant under the fow of the perturbed system. If the method is applied to such a perturbed integrable system, then the numerical method has almost-invariant tori over exponentially long times [HL97]. For a Cantor set of non-resonant step sizes there are even truly invariant tori on which the numerical one-step map reduces to rotation by $h\omega$ in suitable coordinates [HLW02, Sect. X.6.2].

In a very diferent line of research, one asks for integrable discretizations of integrable systems; see the monumental treatise by Suris [Sur03].

2.6 Hamiltonian systems on manifolds

In a more general setting, a Hamiltonian system is considered on a symplectic manifold, which is a manifold with a closed, non-degenerate alternating two-form ω, called the symplectic form. Given a smooth Hamilton function $H : \quad \mathbb{R}$, the corresponding Hamiltonian diferential equation is to fnd $u : [0,T] \quad$ such that

$$\omega_{u(t)}(\dot{u}(t), v) = \mathrm{d}H(u(t))[v] \qquad \text{for all } v \_ T_{u(t)} \quad ,$$

where $T_u \quad$ denotes the tangent space at u of , for a given initial value $u(0) = u_0 \_ \quad$. On inserting $v = \dot{u}(t)$ it is seen that the total energy $H(u(t))$ is constant in time. We write again $u(t) = \varphi_t(u_0)$ to indicate the dependence on the initial value. The fow map φ_t is *symplectic* in the sense that the symplectic form ω is preserved along the fow: for all t and u_0 where $\varphi_t(u_0)$ exists,

$$\omega_{\varphi_t(u_0)}(\mathrm{d}\varphi_t(u_0)[\xi], \mathrm{d}\varphi_t(u_0)[\eta]) = \omega_{u_0}(\xi,\eta) \quad \text{for all } \xi,\eta \_ T_{u_0} \quad ; \quad \text{or } \varphi_t^*\omega = \omega.$$

Contrary to the canonical Hamiltonian systems considered before, no general prescription is known how to construct a symplectic numerical integrator for a general Hamiltonian system on a general symplectic manifold.

However, for the important class of Hamiltonian systems with holonomic constraints, there exist symplectic extensions of the Störmer–Verlet method [And83, LR94] and of higher-order partitioned Runge–Kutta methods [Jay96]. Here the symplectic manifold is the submanifold of $\mathbb{R}^{2d}$ given by constraints $g(q) = 0$, which constrain only the positions, together with the implied constraints for the momenta, $Dg(q) \quad {}_p H(p,q) = 0$.

Apart from holonomic mechanical systems, there exist specially tailored symplectic integrators for particular problem classes of non-canonical Hamiltonian systems. These are often splitting methods, as for example, for rigid body dynamics [DLM97, BCF01], for Gaussian wavepackets in quantum dynamics [FL06], and for post-Newtonian equations in general relativity [LWB10].

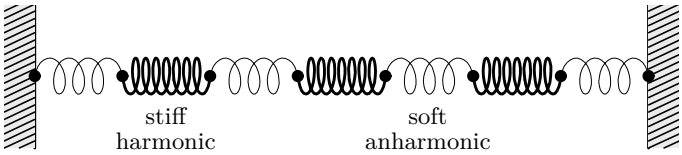

Figure 3.1: Chain of alternating stif harmonic and soft anharmonic springs.

3 Hamiltonian systems with multiple time scales

3.1 Oscillatory Hamiltonian systems

The numerical experiment by Fermi, Pasta and Ulam in 1955, which showed unexpected recurrent behaviour instead of relaxation to equipartition of energy in a chain of weakly nonlinearly coupled particles, has spurred a wealth of research in both mathematics and physics; see, e.g., [Gal08, BI05, For92, Wei97]. Even today, there are only few rigorous mathematical results for large particle numbers in the FPU problem over long times [BP06, HL12], and rigorous theory is lagging behind the insight obtained from carefully conducted numerical experiments [BCP13].

Here we consider a related class of oscillatory Hamiltonian systems for which the long-time behaviour is by now quite well understood analytically both for the continuous problem and its numerical discretizations, and which show interesting behaviour on several time scales. The considered multiscale Hamiltonian systems couple high-frequency harmonic oscillators with a Hamiltonian of slow motion. An illustrative example of such a Hamiltonian is provided by a Fermi–Pasta–Ulam type system of point masses interconnected by stif harmonic and soft anharmonic springs, as shown in Figure 3.1; see [GGMV92] and [HLW06, Section I.5]. The general setting is as follows: For positions $q = (q_0, q_1, \dots, q_m)$ and momenta $p = (p_0, p_1, \dots, p_m)$ with $p_j, q_j \in \mathbb{R}^{d_j}$, let the Hamilton function be given by

$$H(p,q) = H_\omega(p,q) + H_{\text{slow}}(p,q),$$

where the oscillatory and slow-motion energies are given by

$$H_\omega(p,q) = \sum_{j=1}^{m} \frac{1}{2}\Big(p_j{}^2 + \omega_j^2 \, q_j{}^2 \Big), \qquad H_{\text{slow}}(p,q) = \frac{1}{2}\, p_0{}^2 + U(q)$$

with high frequencies

$$\omega_j \geq \varepsilon^{-1}, \qquad 0 < \varepsilon \ll 1.$$

The coupling potential U is assumed smooth with derivatives bounded independently of the small parameter ε. On eliminating the momenta $p_j = \dot{q}_j$, the Hamilton equations become the system of second-order diferential equations

$$\ddot{q}_j + \omega_j^2 q_j = - \quad {}_j U(q), \qquad j = 0, \dots, m,$$

where $\quad {}_j$ denotes the gradient with respect to q_j, and where we set $\omega_0 = 0$. We are interested in the behaviour of the system for initial values with an oscillatory

energy that is bounded independently of ε:

$$H_\omega(p(0), q(0)) \leq \text{Const.}$$

This system shows different behaviour on different time scales:

(i) almost-harmonic motion of the fast variables (p_j, q_j) $(j = 0)$ on time scale ε;

(ii) motion of the slow variables (p_0, q_0) on the time scale ε^0;

(iii) energy exchange between the harmonic oscillators with the same frequency on the time scale ε^{-1};

The relationship between the two techniques of proof, (H) and (F), is not clear at present. The proofs look very different in the basic arguments, in the geometric content and in the technical details, yet lead to very similar results about the long-time behaviour of the continuous problem.

3.2 Modulated Fourier expansion

Modulated Fourier expansions in time have proven useful in the long-time analysis of differential equations where the nonlinearity appears as a perturbation to a primarily linear problem (as laid out in the programme of [FPU55] cited in the introduction). This encompasses important classes of Hamiltonian ordinary and partial differential equations. The approach can be successfully used for the analysis of the continuous problems as well as for their numerical discretizations, as is amply shown in the corresponding references in this and the next section. In particular for the analysis of numerical methods, it offers the advantage that it does not require nonlinear coordinate transforms. Instead, it embeds the original system in a high-dimensional system of modulation equations that has a Lagrangian / Hamiltonian structure with invariance properties. In addition to the use of modulated Fourier expansions as an analytical technique, they have been used also as a numerical approximation method in [HLW02, Chapter XIII] and [Coh04, CDI09, CDI10, FS14, BCZ14, Zha17].

We now describe the basic steps how, for the problem of the previous subsection, a simple ansatz for the solution over a short time interval leads to long-time near-conservation results for the oscillatory energies $E_j = \frac{1}{2}(\dot{q}_j{}^2 + \omega_j^2 \, q_j{}^2)$. We approximate the solution q_j of the second-order differential equation of the previous section as a *modulated Fourier expansion*,

$$q_j(t) \approx \sum_k z_j^k(t)\, \mathrm{e}^{\mathrm{i}(k\cdot\omega)t} \qquad \text{for short times } 0 \le t \le 1,$$

with modulation functions z_j^k, all derivatives of which are required to be bounded independently of ε. The sum is taken over a finite set of multi-indices $k = (k_1, \dots, k_m) \in \mathbb{Z}^m$, and $k \cdot \omega = \sum k_j \omega_j$. The slowly changing modulation functions are multiplied with the highly oscillatory exponentials $\mathrm{e}^{\mathrm{i}(k\cdot\omega)t} = \prod_{j=1}^m \big(\mathrm{e}^{\mathrm{i}\omega_j t}\big)^{k_j}$, which are products of solutions to the linear equations $\ddot{x}_j + \omega_j^2 x_j = 0$. Such products can be expected to be introduced into the solution q_j by the nonlinearity.

Similar multiscale expansions have appeared on various occasions in the literature. The distinguishing feature here is that such a short-time expansion is used to derive long-time properties of the Hamiltonian system.

3.2.1 Modulation system and non-resonance condition

When we insert this ansatz into the differential equation and collect the coefficients to the same exponential $\mathrm{e}^{\mathrm{i}(k\cdot\omega)t}$, we obtain the infinite system of modu-

lation equations for $\mathbf{z} = (z_j^k)$

$$\left(\omega_j^2 - (k \cdot \omega)^2\right) z_j^k + 2\mathrm{i}(k \cdot \omega)\dot{z}_j^k + \ddot{z}_j^k = -\frac{\partial}{\partial z_j^{-k}}(\mathbf{z}).$$

The left-hand side results from the linear part $\ddot{q}_j + \omega_j^2 q_j$ of the differential equation. The right-hand side results from the nonlinearity and turns out to have a ge

under a continuous group action (a geometric property) yields the existence of conserved quantities of the motion (a dynamic property). By Noether's theorem, the modulation equations thus conserve

$$\ell(\mathbf{y}, \dot{\mathbf{y}}) = -\mathrm{i} \sum_j \sum_k k_\ell \omega_\ell \, y_j^{-k} \, \dot{y}_j^k .$$

Since the modulation equations are solved only up to a defect (ε^N) in the construction of the modulated Fourier expansion, the functions $_\ell$ are almost-conserved quantities with (ε^{N+1}) deviations over intervals of length (1). They turn out to be (ε) close to the oscillatory energies E_ℓ. By patching together many short time intervals, the drift in the almost-invariants $_\ell$ is controlled to remain bounded by $Ct\varepsilon^{N+1} \leq C\varepsilon$ over long times $t \leq \varepsilon^{-N}$, and hence also the deviation in the oscillatory energies E_ℓ is only (ε) over such long times. We thus obtain long-time near-conservation of the oscillatory energies E_ℓ.

3.3 Long-time results for numerical integrators

Modulated Fourier expansions were first developed in [HL00b] and further in [HLW02, Chapter XIII] to understand the observed long-term near-conservation of energy by some numerical methods for step sizes for which the smallness condition $hL \ll 1$ of the backward error analysis of Section 2.3 is not fulfilled. For the numerical solution of the differential equation of Section 3.1, we are interested in using numerical integrators that allow large step sizes h such that $h/\varepsilon \geq c_0 > 0$. In this situation, the one-step map of a numerical integrator is no longer a near-identity map, as was the case in Section 2.

For a class of time-symmetric *trigonometric integrators*, which are exact for the uncoupled harmonic oscillator equations $\ddot{x}_j + \omega_j^2 x_j = 0$ and reduce to the Störmer–Verlet method for $\omega_j = 0$, the following results are proved for step sizes h that satisfy a *numerical non-resonance condition*:

$$h\omega_j \text{ is bounded away (by } \overline{h}) \text{ from a multiple of } \pi.$$

Under just this condition it is shown in [CGHL15], using modulated Fourier expansions, that the slow energy H_{slow} is nearly preserved along the numerical solution for very long times $t \leq h^{-N}$ for arbitrary $N \geq 1$ provided the total energy remains bounded along the numerical solution. If in addition,

$$\begin{array}{c} \text{sums of } \pm h\omega_j \text{ with at most } N+1 \text{ terms are} \\ \text{bounded away from non-zero multiples of } 2\pi, \end{array}$$

then also the total and oscillatory energies H and H_ω are nearly preserved along the numerical solution for $t \leq h^{-N}$ for the symplectic methods among the considered symmetric trigonometric integrators. Modified total and oscillatory energies are nearly preserved by the non-symplectic methods in this class. These results yield the numerical version of property (vi) above. A numerical version of property (v) was shown in [CHL05]. The single-frequency case was previously

studied in [HL00b]. For the *Störmer–Verlet method*, which can be interpreted as a trigonometric integrator with modified frequencies, related long-time results are given in [HL00a, CGHL15].

The numerical version of the energy transfer of property (iii) was studied in [HLW02, Section XIII.4] and in [CHL05, MS14]. Getting the energy transfer qualitatively correct by the numerical method turns out to put more restrictions on the choice of methods than long-time energy conservation.

While we concentrated here on long-time results, it should be mentioned that fixed-time convergence results of numerical methods for the multiscale problem as $h \to 0$ and $\varepsilon \to 0$ with $h/\varepsilon \geq c_0 > 0$ also pose many challenges; see, e.g., [GASSS99, HL99, GH06, BGG+17] for systems with constant high frequencies and also [LW14, HL16] for systems with state-dependent high frequencies, where near-preservation of adiabatic invariants is essential. We also refer to [HLW06, Chapters XIII and XIV] and to the review [CJLL06].

4 Hamiltonian partial differential equations

There is a vast literature on the long-time behaviour of nonlinear wave equations, nonlinear Schrödinger equations and other Hamiltonian partial differential equations; see, e.g., the monographs [Kuk93, Bou99, Cra00, Kuk00, KP03, GK14] where infinite-dimensional versions of Hamiltonian perturbation theory are developed. Here we consider a few analytical results that have recently been transfered also to numerical discretizations.

4.1 Long-time regularity preservation

We consider the nonlinear wave equation (or nonlinear Klein–Gordon equation)

$$\partial_t^2 u = \partial_x^2 u - \rho u + g(u), \qquad u = u(x,t) \in \mathbb{R}$$

with 2π-periodic boundary condition in one space dimension, a positive mass parameter ρ and a smooth nonlinearity $g = G'$ with $g(0) = g'(0) = 0$. This equation is a Hamiltonian partial differential equation $\partial_t v = -\nabla_u H(u,v)$, $\partial_t u = \nabla_v H(u,v)$ (where $v = \partial_t u$) with Hamilton function

$$H(u,v) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\frac{1}{2} \Big(v^2 + (\partial_x u)^2 + \rho u^2 \Big) - G(u) \right) \mathrm{d}x$$

on the Sobolev space H^1 of 2π-periodic functions.

Written in terms of the Fourier coefficients u_j of $u(x,t) = \sum_{j \in \mathbb{Z}} u_j(t) \mathrm{e}^{\mathrm{i}jx}$, the nonlinear wave equation takes the form of the oscillatory second-order differential equation of Section 3.1, but the system is now infinite-dimensional:

$$\ddot{u}_j + \omega_j^2 u_j = \gamma_j g(u), \qquad j \in \mathbb{Z},$$

where γ_j gives the jth Fourier coefficient and $\omega_j = \sqrt{j^2 + \rho}$ are the frequencies.

The following result is proved, using infinite-dimensional Hamiltonian perturbation theory (Birkhof normal forms), by Bambusi [Bam03], for arbitrary $N \geq 1$: Under a non-resonance condition on the frequencies ω_j, which is satisfed for almost all values of the parameter ρ, and for initial data (u^0, v^0) that are ε-small in a Sobolev space $H^{s+1} \times H^s$ with sufficiently large $s = s(N)$, the harmonic energies $E_j = \frac{1}{2}(\, \dot{u}_j\,^2 + \omega_j^2\, u_j\,^2)$ are nearly preserved over the time scale $t \leq \varepsilon^{-N}$, and so is the $H^{s+1} \times H^s$ norm of the solution $(u(t), v(t))$.

An alternative proof using modulated Fourier expansions was given in [CHL08b] with the view towards transfering the result to numerical discretizations with trigonometric integrators as done in [CHL08a], for which in addition also a numerical non-resonance condition is required.

Related long-time near-conservation results are proved for other classes of Hamiltonian differential equations, in particular for nonlinear Schrödinger equations with a resonance-removing convolution potential, in [Bou96, BG06, Gré07] using Birkhof normal forms and in [GL10a] using modulated Fourier expansions. These results are transfered to numerical discretization by Fourier collocation in space and splitting methods in time in [FGP10a, FGP10b, GL10b

are

ɜ mo0ulated Fot ier expt sions.

The above hurdles are overcome in [CHL08a] for the nonlinearly perturbed wave equation of the previous subsection discretized by Fourier collocation in space and symplectic trigonometric integrators in time. Here, high regularity of the numerical solution and near-conservation of energy are proved simultaneously using modulated Fourier expansions. In [GL10b], this technique and the energy conservation results are taken further to a class of nonlinear Schrödinger equations with a resonance-removing convolution potential (in arbitrary space dimension) discretized by Fourier collocation in space and a splitting method in time.

Long-time near-conservation of energy for symplectic splitting methods applied to the nonlinear Schrödinger equation in one space dimension (without a resonance-removing convolution potential) is shown in [FG11, Fao12] with a backward error analysis adapted to partial differential equations and, under weaker step size restrictions, in [Gau16] with modulated Fourier expansions. In

on Birkhof normal forms and the other one on modulated Fourier expansions. The latter technique is used in [FGL14] to transfer the result to numerical discretization using Fourier collocation in space and a splitting method for time discretization. The long-time orbital stability under smooth perturbations is in contrast to the instability under rough perturbations shown in [Han14].

5 Dynamical low-rank approximation

Low-rank approximation of too large matrices and tensors is a fundamental approach to data compression and model reduction in a wide range of application areas. Given a matrix $A \in \mathbb{R}^{m\times n}$, the best rank-$r$ approximation to A with respect to the distance given by the Frobenius norm (that is, the Euclidean norm of the vector of entries of a matrix) is known to be obtained by a truncated singular value decomposition: $A \approx \sum_{i=1}^{r} \sigma_i u_i v_i^\top$, where $\sigma_1, \dots, \sigma_r$ are the r largest singular values of A, and $u_i \in \mathbb{R}^m$ and $v_i \in \mathbb{R}^n$ are the corresponding left and right singular vectors, which form an orthonormal basis of the range and corange, respectively, of the best approximation. Hence, only r vectors of both length m and n need to be stored. If $r \ll \min(m,n)$, then the requirements for storing and handling the data are significantly reduced.

When $A(t) \in \mathbb{R}^{m\times n}$, $0 \le t \le T$, is a time-dependent family of large matrices, computing the best rank-r approximation would require singular value decompositions of $A(t)$ for every time instance t of interest, which is often not computationally feasible. Moreover, when $A(t)$ is not given explicitly but is the unknown solution to a matrix differential equation $\dot{A}(t) = F(t, A(t))$, then computing the best rank-r approximation would require to first solve the differential equation on $\mathbb{R}^{m\times n}$, which may not be feasible for large m and n, and then to compute the singular value decompositions at all times of interest, which may again not be feasible.

5.1 Dynamical low-rank approximation of matrices

An alternative — and often computationally feasible — approach can be traced back to Dirac [Dir30] in a particular context of quantum dynamics (see also the next section). Its abstract version can be viewed as a nonlinear Galerkin method on the tangent bundle of an approximation manifold and reads as follows: Consider a differential equation $\dot{A}(t) = F(t, A(t))$ in a (finite- or infinite-dimensional) Hilbert space , and let be a submanifold of . An approximation $Y(t) \in$ to a solution $A(t)$ (for $0 \le t \le T$) is determined by choosing the time derivative $\dot{Y}(t)$ as the orthogonal projection of the vector field $F(t, Y(t))$ to the tangent space $T_{Y(t)}$ at $Y(t) \in$:

$$\dot{Y}(t) = P_{Y(t)} F(t, Y(t)), \tag{5.1}$$

where P_Y denotes the orthogonal projection onto the tangent space at $Y \in$. Equation (5.1) is a differential equation on the approximation manifold , which is complemented with an initial approximation $Y(0) \in$ to $A(0) \in$.

When is a fat space, then this is the standard Galerkin method, which is a basic approximation method for the spatial discretization of partial differential equations. When is not fat, then the tangent space projection P_Y depends on Y, and (5.1) is a nonlinear differential equation even if F is linear.

For the dynamical low-rank approximation of time-dependent matrices, (5.1) is used with chosen as the manifold of rank-r matrices in the space $= \mathbb{R}^{m\times n}$ equipped with the Frobenius inner product (the Euclidean inner product of the matrix entries). This approach was first proposed and studied in [KL07a]. The rank-r matrices are represented in (non-unique) factorized form as

$$Y = USV^\top,$$

where $U \in \mathbb{R}^{m\times r}$ and $V \in \mathbb{R}^{n\times r}$ have orthonormal columns and $S \in \mathbb{R}^{r\times r}$ is an invertible matrix. The intermediate small matrix S is not assumed diagonal, but it has the same non-zero singular values as Y. Differential equations for the factors U, S, V can be derived from (5.1) (uniquely under the gauge conditions $U^\top \dot U = 0$ and $V^\top \dot V = 0$). They contain the inverse of S as a factor on the right-hand side. It is a typical situation that S has small singular values, because in order to obtain accurate approximability, the discarded singular values need to be small, and then the smallest retained singular values are usually not much larger. Small singular values complicate the analysis of the approximation properties of the dynamical low-rank approximation (5.1), for a geometric reason: the curvature of the rank-r manifold at Y (measured as the local Lipschitz constant of the projection map $Y \mapsto P_Y$) is proportional to the inverse of the smallest singular value of Y. It seems obvious that high curvature of the approximation manifold can impair the approximation properties of (5.1), and for a general manifold this is indeed the case. Nevertheless, for the manifold of rank-r matrices there are numerical and theoretical results in [KL07a] that show good approximation properties also in the presence of arbitrarily small singular values.

5.2 Projector-splitting integrator

The numerical solution of the differential equations for U, S, V encounters difficulties with standard time integration methods (such as explicit or implicit Runge–Kutta methods) when S has small singular values, since the inverse of S appears as a factor on the right-hand side of the system of differential equations.

A numerical integration method for these differential equations with remarkable properties is given in [LO14]. It is based on splitting the tangent space projection, which at $Y = USV^\top$ is an alternating sum of three subprojections:

$$P_Y Z = ZVV^\top - UU^\top ZVV^\top + UU^\top Z.$$

Starting from a factorization $Y_n = U_n S_n V_n^\top$ at time t_n, the corresponding splitting integrator updates the factorization of the rank-r approximation to $Y_{n+1} = U_{n+1} S_{n+1} V_{n+1}^\top$ at time t_{n+1}. It alternates between solving (approximately if need be) matrix differential equations of dimensions $m \times r$ (for US),

$r \times r$ (for S), $n \times r$ (for $VS^\top$) and doing orthogonal decompositions of matrices of these dimensions. The inverse of S does not show up in these computations.

The projector-splitting integrator has a surprising exactness property: if the given matrix $A(t)$ is already of rank r for all t, then the integrator reproduces $A(t)$ exactly [LO14]. More importantly, the projector-splitting integrator is robust to the presence of small singular values: it admits convergent error bounds that are independent of the singular values [KLW16]. The proof uses the above exactness property and a geometric peculiarity: in each substep of the algorithm, the approximation moves along a fat subspace of the manifold of rank-r matrices. In this way, the high curvature due to small singular values does no harm.

5.3 Dynamical low-rank approximation of tensors

The dynamical low-rank approximation and the projector-splitting integrator have been extended from matrices to tensors $A(t) \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ such that the favourable approximation and robustness properties are retained; see [KL10, LRSV13, AJ14, LOV15, LVW17]. The dynamical low-rank approximation can be done in various tensor formats that allow for a notion of rank, such as Tucker tensors, tensor trains, hierarchical tensors, and general tensor tree networks; see [Hac12, UV13] for these concepts and for some of their geometric properties.

6 Quantum dynamics

6.1 The time-dependent variational approximation principle

The time-dependent Schrödinger equation for the N-particle wavefunction $\psi = \psi(x_1, \ldots, x_N, t)$,

$$\mathrm{i}\partial_t \psi = H\psi,$$

posed as an evolution equation on the complex Hilbert space $\mathcal{H} = L^2((\mathbb{R}^3)^N, \mathbb{C})$ with a self-adjoint Hamiltonian operator H, is not accessible to direct numerical treatment in the case of several, let alone many particles. "One must therefore resort to approximate methods", as Dirac [Dir30] noted already in the early days of quantum mechanics. For a particular approximation scheme, which is nowadays known as the time-dependent Hartree–Fock method, he used the tangent space projection (5.1) for the Schrödinger equation. Only later was this recognized as a general approximation approach, which is now known as the (Dirac–Frenkel) time-dependent variational principle in the physical and chemical literature: Given a submanifold $\mathcal{M}$ of $\mathcal{H}$, an approximation $u(t) \in \mathcal{M}$ to the wavefunction $\psi(\cdot, t) \in \mathcal{H}$ is determined by the condition that

$$\dot{u} \text{ is chosen as that } w \in T_u\mathcal{M} \text{ for which } \|\mathrm{i}w - Hu\| \text{ is minimal.}$$

This is precisely (5.1) in the context of the Schrödinger equation: $\dot{u} = P_u \frac{1}{\mathrm{i}} H u$. If we assume that the approximation manifold is such that for all $u \in$,

$$T_u \quad \text{is a complex vector space,}$$

(so that with $v \in T_u$, also $\mathrm{i}v \in T_u$), then the orthogonal projection P_u turns out to be also a *symplectic* projection with respect to the canonical symplectic two-form on given by $\omega(\xi, \eta) = 2 \operatorname{Im} \langle \xi, \eta \rangle$ for $\xi, \eta \in$, and is a symplectic manifold. With the Hamilton function $H(u) = \langle u, Hu \rangle$, the differential equation for u can then be rewritten as

$$\omega(\dot{u}, v) = \mathrm{d}H(u)[v] \qquad \text{for all } v \in T_u ,$$

which is a Hamiltonian system on the symplectic manifold ; cf. Section 2.6. The total energy $H(u)$ is therefore conserved along solutions, and the flow is symplectic on . The norm is conserved if contains rays, i.e., with $u \in$ also $\alpha u \in$ for all $\alpha > 0$. We refer to the books [KS81, Lub08] for geometric, dynamic and approximation aspects of the time-dependent variational approximation principle.

6.2 Tensor and tensor network approximations

In an approach that builds on the time-honoured idea of separation of variables, the multi-configuration time-dependent Hartree method (MCTDH) [MMC90, MGW09] uses the time-dependent variational principle to determine an approximation to the multivariate wavefunction that is a linear combination of products of univariate functions:

$$u(x_1, \dots, x_N, t) = \sum_{i_1=1}^{r_1} \cdots \sum_{i_N=1}^{r_N} c_{i_1,\dots,i_N}(t)\, \varphi_{i_1}^{(1)}(x_1, t) \dots \varphi_{i_N}^{(N)}(x_N, t).$$

The time-dependent variational principle yields a coupled system of ordinary differential equations for the coefficient tensor $\big(c_{i_1,\dots,i_N}(t)\big)$ of full multilinear rank and low-dimensional nonlinear Schrödinger equations for the single-particle functions $\varphi_{i_n}^{(n)}(x_n, t)$, which are assumed orthonormal for each $n = 1, \dots, N$. Well-posedness and regularity for this nonlinear system of evolution equations is studied in [KL07b], and an asymptotic error analysis of the MCTDH approximation for growing ranks r_n is given in [CL10].

The projector-splitting integrator of Section 5.2 is extended to MCTDH in [Lub15]. The nonlinear MCTDH equations are thus split into a chain of linear single-particle differential equations, alternating with orthogonal matrix decompositions. The integrator conserves the L^2 norm and the total energy and, as is proved in [LVW17], it is robust to the presence of small singular values in matricizations of the coefficient tensor.

In the last decade, tensor network approximations, and in particular matrix product states, have increasingly come into use for the description of strongly interacting quantum many-body systems; see, e.g., [VMC08, CV09, SPM+15].

Matrix product states (known as tensor trains in the mathematical literature [Ose11]) approximate the wavefunction by

$$u(x_1, \ldots, x_N, t) = G_1(x_1, t) \cdot \ldots \cdot G_N(x_N, t)$$

with matrices $G_n(x_n, t)$ of compatible (low) dimensions. This approach can be viewed as a non-commutative separation of variables. Its memory requirements grow only linearly with the number of particles N, which makes the approach computationally attractive for many-body systems. The approximability of the wavefunction or derived quantities by this approach is a different issue, with some excellent computational results but hardly any rigorous mathematical theory so far.

For the numerical integration of the equations of motion that result from the time-dependent variational approximation principle, the projector-splitting integrator has recently been extended to matrix product states in [LOV15, HLO$^+$16], with favourable properties like the MCTDH integrator. The important robustness to the presence of small singular values is proved in [KLW16], again using the property that the integrator moves along fat subspaces within the tensor manifold.

Acknowledgement

We thank Balázs Kovács, Frank Loose, Hanna Walach, and Gerhard Wanner for helpful comments.

References

[AEEVE12] A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numerica*, 21:1–87, 2012.

[AJ14] A. Arnold and T. Jahnke. On the approximation of high-dimensional differential equations in the hierarchical Tucker format. *BIT Numer. Math.*, 54(2):305–341, 2014.

[AKN97] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt. *Mathematical Aspects of Classical and Celestial Mechanics*. Springer, Berlin, 1997.

[AM04] U. M. Ascher and R. I. McLachlan. Multisymplectic box schemes and the Korteweg-de Vries equation. *Appl. Numer. Math.*, 48(3-4):255–269, 2004. Workshop on Innovative Time Integrators for PDEs.

[And83] H. C. Andersen. Rattle: a "velocity" version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.*, 52:24–34, 1983.

[Bam03] D. Bambusi. Birkhof normal form for some nonlinear PDEs. *Comm. Math. Phys.*, 234:253–285, 2003.

[BC16] S. Blanes and F. Casas. *A Concise Introduction to Geometric Numerical Integration*. Monographs and Research Notes in Mathematics. CRC Press, Boca Raton, FL, 2016.

[BCF01] G. Benettin, A. M. Cherubini, and F. Fassò. A changing-chart symplectic algorithm for rigid bodies and other Hamiltonian systems on manifolds. *SIAM J. Sci. Comput.*, 23:1189–1203, 2001.

[BCP13] G. Benettin, H. Christodoulidi, and A. Ponno. The Fermi–Pasta–Ulam problem and its underlying integrable dynamics. *J. Stat. Phys.*, 152(2):195–212, 2013.

[BCZ14] W. Bao, Y. Cai, and X. Zhao. A uniformly accurate multiscale time integrator pseudospectral method for the klein–gordon equation in the nonrelativistic limit regime. *SIAM J. Numer. Anal.*, 52:2488–2511, 2014.

[BFG13] D. Bambusi, E. Faou, and B. Grébert. Existence and stability of ground states for fully discrete approximations of the nonlinear Schrödinger equation. *Numer. Math.*, 123(3):461–492, 2013.

[BG94] G. Benettin and A. Giorgilli. On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms. *J. Statist. Phys.*, 74:1117–1143, 1994.

[BG06] D. Bambusi and B. Grébert. Birkhof normal form for partial differential equations with tame modulus. *Duke Math. J.*, 135(3):507–567, 2006.

[BGG87] G. Benettin, L. Galgani, and A. Giorgilli. Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. Part I. *Comm. Math. Phys.*, 113:87–103, 1987.

[BGG89] G. Benettin, L. Galgani, and A. Giorgilli. Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. Part II. *Comm. Math. Phys.*, 121:557–601, 1989.

[BGG+17] S. Buchholz, L. Gauckler, V. Grimm, M. Hochbruck, and T. Jahnke. Closing the gap between trigonometric integrators and splitting methods for highly oscillatory differential equations. *IMA J. Numer. Anal.*, page drx007, 2017.

[BGPP13] D. Bambusi, A. Giorgilli, S. Paleari, and T. Penati. Normal form and energy conservation of high frequency subsystems without non-resonance conditions. *Istituto Lombardo (Rend. Scienze)*, 147:1–17, 2013.

[BI05] G.P. Berman and F.M. Izrailev. The Fermi–Pasta–Ulam problem: ffty years of progress. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 15:015104, 2005.

[BL07] S.D. Bond and B.J. Leimkuhler. Molecular dynamics and the accuracy of numerically computed averages. *Acta Numerica*, 16:1–65, 2007.

[Bou96] J. Bourgain. Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations. *Geom. Funct. Anal.*, 6:201–230, 1996.

[Bou99] J. Bourgain. *Global solutions of nonlinear Schrödinger equations*, volume 46 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1999.

[BP06] D. Bambusi and A. Ponno. On metastability in FPU.

[CHL08a] D. Cohen, E. Hairer, and C. Lubich. Conservation of energy, momentum and actions in numerical discretizations of nonlinear wave equations. *Numer. Math.*, 110:113–143, 2008.

[CHL08b] D. Cohen, E. Hairer, and C. Lubich. Long-time analysis of nonlinearly perturbed wave equations via modulated Fourier expansions. *Arch. Ration. Mech. Anal.*, 187:341–368, 2008.

[Chu08] M. T. Chu. Linear algebra algorithms as dynamical systems. *Acta Numerica*, 17:1–86, 2008.

[CJLL06] D. Cohen, T. Jahnke, K. Lorenz, and C. Lubich. Numerical integrators for highly oscillatory Hamiltonian systems: a review. In *Analysis, modeling and simulation of multiscale problems*, pages 553–576. Springer, Berlin, 2006.

[CL10] D. Conte and C. Lubich. An error analysis of the multi-configuration time-dependent Hartree method of quantum dynamics. *ESAIM: Mathematical Modelling and Numerical Analysis*, 44(4):759–780, 2010.

[CMKO11] S. H. Christiansen, H. Z. Munthe-Kaas, and B. Owren. Topics in structure-preserving discretization. *Acta Numerica*, 20:1–119, 2011.

[Coh04] D. Cohen. *Analysis and numerical treatment of highly oscillatory differential equations*. PhD thesis, Univ. Genève, 2004.

[Cra00] W. Craig. *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, volume 9 of *Panoramas et Synthèses*. Société Mathématique de France, Paris, 2000.

[CSS92] M. P. Calvo and J. M. Sanz-Serna. Variable steps for symplectic integrators. In *Numerical Analysis 1991*, Res. Notes Math. Ser. 260, pages 34–48, Dundee, 1992. Pitman.

[CV09] J. I. Cirac and F. Verstraete. Renormalization and tensor product states in spin chains and lattices. *J. Physics A: Mathematical and Theoretical*, 42(50):504004, 2009.

[DDE05] K. Deckelnick, G. Dziuk, and C. M. Elliott. Computation of geometric partial differential equations and mean curvature flow. *Acta Numerica*, 14:139–232, 2005.

[DE13] G. Dziuk and C. M. Elliott. Finite element methods for surface PDEs. *Acta Numerica*, 22:289–396, 2013.

[Dir30] P. A. M. Dirac. Note on exchange phenomena in the Thomas atom. *Math. Proc. Cambridge Phil. Soc.*, 26(3):376–385, 1930.

[DLM97] A. Dullweber, B. Leimkuhler, and R. McLachlan. Symplectic splitting methods for rigid body molecular dynamics. *J. Chem. Phys. 107 No.*, 15:5840–5851, 1997.

[dV56] R. de Vogelaere. Methods of integration which preserve the contact transformation property of the Hamiltonian equations. Technical report, Dept. Math. Univ. of Notre Dame, Notre Dame, Ind , 1956.

[Fao12] E. Faou. *Geometric Numerical Integration and Schrödinger Equations*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2012.

[Fen85] K. Feng. On difference schemes and symplectic geometry. In *Proceedings of the 5-th Intern. Symposium on differential geometry & differential equations 1984*, pages 42–58, Beijing, August 1985.

[Fen86] K. Feng. Difference schemes for Hamiltonian formalism and symplectic geometry. *J. Comp. Math.*, 4:279–289, 1986.

[FG11] E. Faou and B. Grébert. Hamiltonian interpolation of splitting approximations for nonlinear PDEs. *Found. Comput. Math.*, 11(4):381–415, 2011.

[FGL13] E. Faou, L. Gauckler, and C. Lubich. Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus. *Comm. Partial Differential Equations*, pages 1123–1140, 2013.

[FGL14] E. Faou, L. Gauckler, and C. Lubich. Plane wave stability of the split-step Fourier method for the nonlinear Schrödinger equation. *Forum of Mathematics, Sigma*, 2:e5, 2014.

[FGP10a] E. Faou, B. Grébert, and E. Paturel. Birkhof normal form for splitting methods applied to semilinear Hamiltonian PDEs. part I. Finite-dimensional discretization. *Numer. Math.*, 114(3):429–458, 2010.

[FGP10b] E. Faou, B. Grébert, and E. Paturel. Birkhof normal form for splitting methods applied to semilinear Hamiltonian PDEs. part II. Abstract splitting. *Numer. Math.*, 114(3):459–490, 2010.

[FL06] E. Faou and C. Lubich. A Poisson integrator for Gaussian wavepacket dynamics. *Computing and Visualization in Science*, 9(2):45–55, 2006.

[For92] J. Ford. The Fermi–Pasta–Ulam problem: paradox turns discovery. *Physics Reports*, 213:271–310, 1992.

[FPU55] E. Fermi, J. Pasta, and S. Ulam. Studies of non linear problems. Technical Report LA-1940, Los Alamos, 1955. Later published in E. Fermi: Collected Papers, Chicago, 1965 and reprinted in G. Gallavotti (ed.), The Fermi–Pasta–Ulam Problem. A Status Report. Springer Lect. Notes Phys. 728, 2008.

[FQ10] K. Feng and M.-Z. Qin. *Symplectic Geometric Algorithms for Hamiltonian Systems*. Zhejiang Science and Technology Publishing House, Hangzhou, 2010. Translated and revised from the Chinese original. With a foreword by Feng Duan.

[FS14] E. Faou and K. Schratz. Asymptotic preserving schemes for the klein–gordon equation in the non-relativistic limit regime. *Numer. Math.*, 126(3):441–469, 2014.

[Gal08] G. Gallavotti, editor. *The Fermi–Pasta–Ulam problem*, volume 728 of *Lecture Notes in Physics*. Springer, Berlin, 2008. A status report.

[GL10a] L. Gauckler and C. Lubich. Nonlinear Schrödinger equations and their spectral semi-discretizations over long times. *Found. Comput. Math.*, 10:141–169, 2010.

[GL10b] L. Gauckler and C. Lubich. Splitting integrators for nonlinear Schrödinger equations over long times. *Found. Comput. Math.*, 10:275–302, 2010.

[Gré07] B. Grébert. Birkhof normal form and Hamiltonian PDEs. In *Partial differential equations and applications*, volume 15 of *Sémin. Congr.*, pages 1–46. Soc. Math. France, Paris, 2007.

[GW17] L. Gauckler and D. Weiss. Metastable energy strata in numerical discretizations of weakly nonlinear wave equations. *Discrete Contin. Dyn. Syst.*, 37(7):3721–3747, 2017.

[Hac12] W. Hackbusch. *Tensor spaces and numerical tensor calculus*. Springer Science & Business Media, 2012.

[Hai94] E. Hairer. Backward analysis of numerical integrators and symplectic methods. *Annals of Numerical Mathematics*, 1:107–132, 1994.

[Han14] Z. Hani. Long-time instability and unbounded Sobolev orbits for some periodic nonlinear Schrödinger equations. *Arch. Ration. Mech. Anal.*, 211(3):929–964, 2014.

[HL97] E. Hairer and C. Lubich. The life-span of backward error analysis for numerical integrators. *Numer. Math.*, 76:441–462, 1997. Erratum: http://www.unige.ch/math/folks/hairer/.

[HL99] M. Hochbruck and C. Lubich. A Gautschi-type method for oscillatory second-order diferential equations. *Numer. Math.*, 83:403–426, 1999.

[HL00a] E. Hairer and C. Lubich. Energy conservation by Störmer-type numerical integrators. In D. F. Grifths G. A. Watson, editor, *Numerical Analysis 1999*, pages 169–190. CRC Press LLC, 2000.

[HL00b] E. Hairer and C. Lubich. Long-time energy conservation of numerical methods for oscillatory diferential equations. *SIAM J. Numer. Anal.*, 38:414–441, 2000.

[HL04] E. Hairer and C. Lubich. Symmetric multistep methods over long times. *Numer. Math.*, 97:699–723, 2004.

[HL12] E. Hairer and C. Lubich. On the energy distribution in Fermi–Pasta–Ulam lattices. *Arch. Ration. Mech. Anal.*, 205(3):993–1029, 2012.

[HL16] E. Hairer and C. Lubich. Long-term analysis of the Störmer–Verlet method for Hamiltonian systems with a solution-dependent high frequency. *Numer. Math.*, 34:119–138, 2016.

[HL17] E. Hairer and C. Lubich. Symmetric multistep methods for charged-particle dynamics. *SMAI J. Comput. Math., to appear*, 2017.

[HLO+16] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete. Unifying time evolution and optimization with matrix product states. *Physical Review B*, 94(16):165116, 2016.

[HLW02] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Series in Computational Mathematics 31. Springer-Verlag, Berlin, 2002.

[HLW03] E. Hairer, C. Lubich, and G. Wanner. Geometric numerical integration illustrated by the Störmer–Verlet method. *Acta Numerica*, 12:399–450, 2003.

[HLW06] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Series in Computational Mathematics 31. Springer-Verlag, Berlin, 2nd edition, 2006.

[HO10] M. Hochbruck and A. Ostermann. Exponential integrators. *Acta Numerica*, 19:209–286, 2010.

[IMKNZ00] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. *Acta Numerica*, pages 215–365, 2000.

[Jay96] L. Jay. Symplectic partitioned Runge–Kutta methods for constrained Hamiltonian systems. *SIAM J. Numer. Anal.*, 33:368–387, 1996.

[KL07a] O. Koch and C. Lubich. Dynamical low-rank approximation. *SIAM J. Matrix Anal. Appl.*, 29(2):434–454, 2007.

[KL07b] O. Koch and C. Lubich. Regularity of the multi-configuration time-dependent Hartree approximation in quantum molecular dynamics. *ESAIM: Mathematical Modelling and Numerical Analysis*, 41(2):315–331, 2007.

[KL10] O. Koch and C. Lubich. Dynamical tensor approximation. *SIAM J. Matrix Anal. Appl.*, 31(5):2360–2375, 2010.

[KLW16] E. Kieri, C. Lubich, and H. Walach. Discretized dynamical low-rank approximation in the presence of small singular values. *SIAM J. Numer. Anal.*, 54(2):1020–1038, 2016.

[KP03] T. Kappeler and J. Pöschel. *KdV & KAM*, volume 45 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*. Springer-Verlag, Berlin, 2003.

[KS81] P. Kramer and M. Saraceno. *Geometry of the time-dependent variational principle in quantum mechanics*, volume 140 of *Lecture Notes in Physics*. Springer, Berlin, 1981.

[Kuk93] S. B. Kuksin. *Nearly integrable infinite-dimensional Hamiltonian systems*, volume 1556 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1993.

[Kuk00] S. B. Kuksin. *Analysis of Hamiltonian PDEs*, volume 19 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2000.

[Las88] F. M. Lasagni. Canonical Runge–Kutta methods. *ZAMP*, 39:952–953, 1988.

[Las13] J. Laskar. Is the solar system stable? In *Chaos*, volume 66 of *Prog. Math. Phys.*, pages 239–270. Birkhäuser/Springer, Basel, 2013.

[LO14] C. Lubich and I. V. Oseledets. A projector-splitting integrator for dynamical low-rank approximation. *BIT Numer. Math.*, 54(1):171–188, 2014.

[LOV15] C. Lubich, I. V. Oseledets, and B. Vandereycken. Time integration of tensor trains. *SIAM J. Numer. Anal.*, 53(2):917–941, 2015.

[LR94] B. Leimkuhler and S. Reich. Symplectic integration of constrained Hamiltonian systems. *Math. Comp.*, 63:589–605, 1994.

[LR04] B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*. Cambridge Monographs on Applied and Computational Mathematics 14. Cambridge University Press, Cambridge, 2004.

[LRSV13] C. Lubich, T. Rohwedder, R. Schneider, and B. Vandereycken. Dynamical approximation by hierarchical Tucker and tensor-train tensors. *SIAM J. Matrix Anal. Appl.*, 34(2):470–494, 2013.

[Lub08] C. Lubich. *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008.

[Lub15] C. Lubich. Time integration in the multiconfiguration time-dependent Hartree method of molecular quantum dynamics. *Applied Mathematics Research eXpress*, 2015(2):311–328, 2015.

[LVW17] C. Lubich, B. Vandereycken, and H. Walach. Time integration of rank-constrained Tucker tensors. *Preprint, arXiv:1709.02594*, 2017.

[LW14] C. Lubich and D. Weiss. Numerical integrators for motion under a strong constraining force. *Multiscale Modeling & Simulation*, 12(4):1592–1606, 2014.

[LWB10] C. Lubich, B. Walther, and B. Brügmann. Symplectic integration of post-Newtonian equations of motion with spin. *Physical Review D*, 81(10):104025, 2010.

[MGW09] H.-D. Meyer, F. Gatti, and G. A. Worth. *Multidimensional quantum dynamics*. John Wiley & Sons, 2009.

[MMC90] H.-D. Meyer, U. Manthe, and L. S. Cederbaum. The multi-confgurational time-dependent Hartree approach. *Chem. Phys. Letters*, 165(1):73–78, 1990.

[Mos68] J. Moser. Lectures on Hamiltonian systems. *Mem. Am. Math. Soc.*, 81:1–60, 1968.

[Mos78] J. Moser. Is the solar system stable? *Mathematical Intelligencer*, 1:65–71, 1978.

[MPS98] J. E. Marsden, G. W. Patrick, and S. Shkoller. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Comm. Math. Phys.*, 199(2):351–395, 1998.

[MQ02] R. I. McLachlan and G. R. W. Quispel. Splitting methods. *Acta Numerica*, 11:341–434, 2002.

[MS14] R. I. McLachlan and A. Stern. Modifed trigonometric integrators. *SIAM J. Numer. Anal.*, 52:1378–1397, 2014.

[MW01] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, 10:1–158, 2001.

[Nei84] A. I. Neishtadt. The separation of motions in systems with rapidly rotating phase. *Journal of Applied Mathematics and Mechanics*, 48(2):133–139, 1984.

[Ose11] I. V. Oseledets. Tensor-train decomposition. *SIAM J. Sci. Comput.*, 33(5):2295–2317, 2011.

[Poi92] H. Poincaré. *Les Méthodes Nouvelles de la Mécanique Céleste Tome I.* Gauthier-Villars, Paris, 1892.

[Rei99] S. Reich. Backward error analysis for numerical integrators. *SIAM J. Numer. Anal.*, 36:1549–1570, 1999.

[Rut83] R. D. Ruth. A canonical integration technique. *IEEE Trans. Nuclear Science*, NS-30:2669–2671, 1983.

[Sha99] Z. Shang. KAM theorem of symplectic algorithms for Hamiltonian systems. *Numer. Math.*, 83:477–496, 1999.

[Sha00] Z. Shang. Resonant and Diophantine step sizes in computing invariant tori of Hamiltonian systems. *Nonlinearity*, 13:299–308, 2000.

[SM71] C. L. Siegel and J. K. Moser. *Vorlesungen über Himmelsmechanik*, volume 85 of *Grundlehren d. math. Wiss.*. Springer-Verlag, 1971. First German edition: C. L. Siegel, *Vorlesungen über Himmelsmechanik*, Grundlehren vol. 85, Springer-Verlag, 1956.

[SPM+15] S. Szalay, M. Pfefer, V. Murg, G. Barcza, F. Verstraete, R. Schneider, and Ö. Legeza. Tensor product methods and entanglement optimization for ab initio quantum chemistry. *Int. J. Quantum Chem.*, 115(19):1342–1391, 2015.

[SS88] J. M. Sanz-Serna. Runge-Kutta schemes for Hamiltonian systems. *BIT*, 28:877–883, 1988.

[SS92] J. M. Sanz-Serna. Symplectic integrators for Hamiltonian problems: an overview. *Acta Numerica*, 1:243–286, 1992.

[SSC94] J. M. Sanz-Serna and M. P. Calvo. *Numerical Hamiltonian problems*, volume 7 of *Applied Mathematics and Mathematical Computation*. Chapman & Hall, London, 1994.

[Sur88] Y. B. Suris. On the conservation of the symplectic structure in the numerical solution of Hamiltonian systems. In S. S. Filippov, editor, *Numerical Solution of Ordinary Differential Equations*, pages 148–160. Keldysh Institute of Applied Mathematics USSR Academy of Sciences, Moscow, 1988. in Russian.

[Sur90] Y. B. Suris. Hamiltonian methods of Runge–Kutta type and their variational interpretation. *in Russian Math. Model.*, 2:78–87, 1990.

[Sur03] Y. B. Suris. *The Problem of Integrable Discretization: Hamiltonian Approach*, volume 219 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2003.

[UV13] A. Uschmajew and B. Vandereycken. The geometry of algorithms using hierarchical tensors. *Linear Algebra Appl.*, 439(1):133–166, 2013.

[Ver67] L. Verlet. Computer "experiments" on classical fuids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review*, 159:98–103, 1967.

[VMC08] F. Verstraete, V. Murg, and J. I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Advances in Physics*, 57(2):143–224, 2008.

[Wan10] G. Wanner. Kepler, Newton and numerical analysis. *Acta Numerica*, 19:561–598, 2010.

[Wei85] M. I. Weinstein. Modulational stability of ground states of nonlinear Schrödinger equations. *SIAM J. Math. Anal.*, 16(3):472–491, 1985.

[Wei97] T.P. Weissert. *The Genesis of Simulation in Dynamics: Pursuing the Fermi–Pasta–Ulam Problem*. Springer-Verlag, New York, 1997.

[WH91] J. Wisdom and M. Holman. Symplectic maps for the n-body problem. *Astron. J.*, 102:1528–1538, 1991.

[Wil60] J. H. Wilkinson. Error analysis of floating-point computation. *Numer. Math.*, 2:319–340, 1960.

[WYW13] X. Wu, X. You, and B. Wang. *Structure-Preserving Algorithms for Oscillatory Differential Equations*. Springer Science & Business Media, 2013.

[Zha17] X. Zhao. Uniformly accurate multiscale time integrators for second order oscillatory differential equations with large initial data. *BIT Numer. Math.*, 57:649–683, 2017.