Mathematisches Forschungsinstitut Oberwolfach

Report No. 17/2021

DOI: 10.4171/OWR/2021/17

Geometric Numerical Integration (hybrid meeting)

Organized by
Marlis Hochbruck, Karlsruhe
Christian Lubich, Tübingen
Robert McLachlan, Palmerston North
Jesús María Sanz-Serna, Madrid

28 March – 3 April 2021

Abstract. The topics of the workshop included interactions between geometric numerical integration and numerical partial differential equations; geometric aspects of stochastic differential equations; interaction with optimisation and machine learning; new applications of geometric integration in physics; problems of discrete geometry, integrability, and algebraic aspects.

Mathematics Subject Classification (2010): Primary: 65; Secondary: 34, 35.

Introduction by the Organizers

Geometric Numerical Integration is an area that has developed vividly over the last decades, starting from the seminal work on symplectic integrators for Hamiltonian ordinary differential equations by Feng Kang and the characterisation of symplectic Runge–Kutta methods simultaneously and independently by Sanz-Serna, Suris, and Lasagni in the late 1980s. Since then, it has been realised that geometric aspects play a fundamental role in the design and analysis of structure-preserving numerical methods for a wide variety of classes of ordinary or partial, deterministic or stochastic differential equations. This has led, on the one hand, to new algorithms that are successfully applied in various fields of science, such as particle physics, celestial mechanics, classical and quantum molecular dynamics, fluid mechanics, to name but a few. On the other hand, rigorous mathematical insight has been obtained into the working mechanisms of structure-preserving numerical integrators, which explains, for example, why preserving geometric properties, such as symplecticity, leads to improved dynamics, such as near-preservation of energy

over long times. The interplay of rigorous mathematical theory, which draws from different fields of mathematics, and the construction of practical algorithms for concrete applications outside mathematics, has fruitfully gone back and forth in the area of Geometric Numerical Integration.

There have been previous Oberwolfach workshops on Geometric Numerical Integration in the years 2006, 2011, and 2016. The field has continued to be very active, gaining substantial new theoretical insight as well as opening up new application areas adapting techniques developed here previously in different contexts.

Directions followed since 2016 and represented in this workshop include

- interactions between geometric integration and several areas of numerical partial differential equations;
- continuing growth of interest in geometric aspects of stochastic ordinary and partial differential equations;
- interaction with optimisation and machine learning;
- new applications of geometric integration in physics;
- problems of discrete geometry, integrability, algebraic aspects.

Some of the items belong to the core area of Geometric Numerical Integration, others use and export techniques developed in this field to other scientific fields outside mathematics and within.

Due to the covid pandemic, the workshop took place as a hybrid workshop, with only seven participants (including two organizers) present at Oberwolfach, all from Germany, and with 50 online participants from Europe, the USA, China, Japan, Australia and New Zealand. To cope with the online format, there were 20 half-hour talks during the workshop spread between 9 am and 11 pm to accommodate participants from different time zones, and there was a junior researcher session with eight contributions. In addition, there were six discussion groups on special topics, with short reports to the general audience on the last day of the workshop, and there were virtual coffee rooms for informal exchange among the participants. From the scientific perspective, the workshop was very useful, as it highlighted current developments in the field and brought together participants with different perspectives in lively interaction.

It must be said, however, that despite the excellent technical support from the MFO, despite the strong efforts by organizers and participants to put up a varied and interesting scientific programme and despite the very active participation, the hybrid workshop could not match a real in-person in-situ Oberwolfach workshop. Aside from irreplaceable personal encounters, it makes a tremendous difference of whether a participant is in Oberwolfach, secluded from everyday life, concentrating in leisure (the Latin *otium*) on the workshop, on mathematical problems and interaction with participants, or follows just some talks in yet another online workshop amidst a heap of everyday duties.

Only four invitations have been canceled, all four by women, and in addition one female online participant did not show up. It appears that the strains of the pandemic are not gender-neutral, at least not among researchers in this field.

As a consequence, there was only a reduced female participation by seven female researchers at the workshop.

Workshop (hybrid meeting): Geometric Numerical Integration

Table of Contents

Abstracts

Methods to compute optimal reparametrizations in shape analysis.

Elena Celledoni
(joint work with Jørgen Riseth and Alexander Schmeding)

Finding the optimal reparametrization in shape analysis of curves or surfaces is a computationally demanding task. The problem can be phrased as an optimisation problem on the infinite dimensional group of orientation preserving diffeomorphisms $\mathrm{Diff}^+(\Omega)$, where Ω is the domain where the curves or surfaces are defined. In the case of curves, one robust approach to compute optimal reparametrizations is based on dynamic programming [3], but this method seems difficult to generalize to surfaces.

We consider here a method inspired by a "Riemannian" gradient descent approach which we illustrate in the case of curves. Suppose $\Omega = [0,1]$ and consider the space $\mathrm{Imm}([0,1],\mathbb{R}^n)$ of C^∞ curves on Ω with nonvanishing derivative, define the corresponding shapes to be equivalence classes under the right action of $\mathrm{Diff}^+(\Omega)$. Let

$$Q : \mathrm{Imm}(I,\mathbb{R}^n) \to C^\infty(I,\mathbb{R}^n), \qquad c(\cdot) \to \sqrt{\|\dot{c}(\cdot)\|}\, c(\cdot)$$

denote the Q-transform, [2]. The distance between two shapes $[c_0]$ and $[c_1]$ is obtained by first considering the Q-transform of the curves $q_0 = Q(c_0)$ and $q_1 = Q(c_1)$ and then solving of the optimization problem

$$\inf_{\varphi\in\mathrm{Diff}^+([0,1])} E(\varphi), \qquad E(\varphi) := \|q_0 - \sqrt{\dot{\varphi}}\, q_1 \circ \varphi\|_{L_2},$$

[2]. Here the subscript L_2 denotes L_2-norm. This is an optimisation problem on an infinite dimensional group, which we approximate by replacing $\mathrm{Diff}^+(\Omega)$ with a finite dimensional space.

A gradient descent approach is obtained by representing the gradient $\mathrm{grad}\, E$ in terms of an othonormal basis of $T_{\mathrm{id}}\mathrm{Diff}^+(\Omega)$ and projecting $\mathrm{grad}\, E$ on a finite dimensional subspace. This approach is equivalent to the algorithm proposed in [2]. We compute the approximation of the optimal reparametrization φ iteratively by the following update rule:

$$\varphi^{(n+1)} = \varphi^{(n)} \circ \left(\mathrm{id} - \eta \sum_{j=1}^{M} \lambda_j v_j\right), \quad n = 0,1,\dots$$

with $\varphi_0 = \mathrm{id}$, where $v_j \in T_{\mathrm{id}}\mathrm{Diff}^+(\Omega)$, $\{\lambda_j\}_{j=1}^M$ are coefficients determined at each iteration n, η is a scalar parameter optimised so to guarantee the invertibility of $\varphi^{(n+1)}$, and $\sum_{j=1}^M \lambda_j v_j$ is the approximation of the gradient.

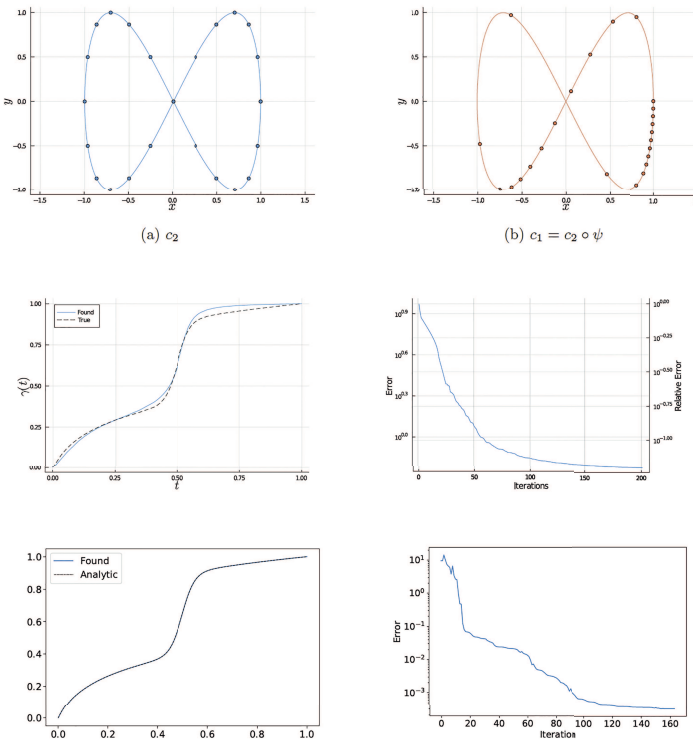

FIGURE 1. Top a curve c_2 (left) and the same curve with a different parametrization $c_1 = c_2 \circ \psi$ (right), in this experiment we aim at reconstructing ψ. Center reconstruction of ψ by the gradient descent approach: left approximation of ψ, right convergence history of the gradient descent iteration. Bottom reconstruction of ψ by a multi-layer approach: left approximation of ψ, right convergence history of the gradient descent iteration optimising simultaneously for M basis elements and for the composition of L diffeomorphisms.

To improve the performance of the method, we approximate the optimal reparametrisation by composing L diffeomorphisms of the form

$$\mathrm{id} + \sum_{j=1}^{M} \lambda_j^\ell v_j, \qquad \ell = 1, \dots, L, \tag{1}$$

where M and L are fixed, and we optimise simultaneously on all the parameters $\{\lambda_j^\ell\}$ for $j = 1, \dots M$ and $\ell = 1, \dots, L$. The comparison of the two algorithms is reported in figure 1 for a simple test example. The algorithm is similar to a deep neural network and its implementation has been carried out using PyTorch.

The method is conceptually the same in the case of surfaces and the details and preliminary experiments are reported in [4].

This algorithm is motivated by results in [1], about the controllability of the group of diffeomorphisms, via the composition of a finite number of elementary diffeomorphisms, each of them here represented with (1). For a compact manifold Ω, any bracket generating family of vector fields on Ω, which is invariant under multiplication by smooth functions, generates the connected component of the identity of the group of diffeomorphisms of Ω, [1].

References

[1] A.A. Agrachev and M. Caponigro, *Controllability on the group of diffeomorphisms*, Ann. I. H. Poincaré, 2009.

[2] S Kurtek, E Klassen, JC Gore, Z Ding, A Sivrastava, *Elastic geodesic paths in shape space of parameterized surfaces*, IEEE transactions on pattern analysis and machine intelligence, 2012.

[3] TB Sebastian, PN Klein, BB Kimia, *On aligning of curves*, IEEE transactions on pattern analysis and machine intelligence, 2003.

[4] JN Riseth, *Gradient-based algorithms in shape analysis for reparametrization of parametric curves and surfaces*, Feb. 2021, Master thesis, NTNU.

Towards Lie–Butcher series for geodesic flows

Hans Munthe-Kaas

(joint work with Kurusch Ebrahimi-Fard, Dominique Manchon)

Summary. The talk presented at this meeting outlined new results on series developments similar to Butcher's B-series for geodesic flows on manifolds equipped with a general affine connection. This includes the Levi–Civita connection on Riemannian metric spaces as a special case. The novel theory paths new ways for analysing numerical integration schemes based on geodesic flows. We also see possible applications in the study of rough paths on affine geometries, problems in control theory as well as open questions in information geometry.

Background. The development and analysis of numerical Lie group integrators [8] has motivated generalisations of Butcher's celebrated B-series [1]. *Lie–Butcher series*, combining B-series with Lie series on manifolds, were introduced and developed in a series of papers [13, 14] in the second half of the 1990s. A geometric understanding of such series is based on affine connections. On the space $\mathcal{X}\mathcal{M}$ of vector fields on a manifold $\mathcal{M}$, we denote the connection as a binary product, $f \triangleright g := \nabla_f g$. In previous works [6, 9, 10, 11, 15], we have investigated particular algebras of *invariant connections*, defined as connections where the curvature and torsion tensors are constant (parallel), so that their covariant derivatives vanish, i.e., $\nabla R = \nabla T = 0$.

In the case where $R = T = 0$, the connection defines a *pre-Lie* algebra [12], i.e., an algebra $(\mathcal{X}\mathcal{M}, \triangleright)$ such that the associator is left symmetric, $\mathrm{a}_\triangleright(f, g, h) = \mathrm{a}_\triangleright(g, f, h)$, where $\mathrm{a}_\triangleright(f, g, h) := f \triangleright (g \triangleright h) - (f \triangleright g) \triangleright h$. The free pre-Lie algebra is the algebra of non-planar rooted trees with grafting [2], yielding classical B-series.

We remark that the algebra of torsion free connections with constant curvature ($\nabla R = 0$, $T = 0$), called a *Lie admissible triple algebra*, will not be discussed here.

Flat connections with constant torsion ($R = 0$, $\nabla T = 0$) yield *post-Lie* algebras, $(\mathcal{XM}, \triangleright, [\_,\_])$. Here, $[f,g] = -T(f,g)$ is a Lie bracket and the connection satisfy

$$f \triangleright [g,h] = [f \triangleright g, h] + [g, f \triangleright h]$$
$$[f,g] \triangleright h = \mathrm{a}_{\triangleright}(f,g,h) - \mathrm{a}_{\triangleright}(g,f,h).$$

See [3] and references in there. The free post-Lie algebra PL in one generator is given in terms of the free Lie algebra over planar rooted trees with left grafting, and the corresponding series are *Lie–Butcher series*. The enveloping algebra $U(PL)$ is a D-algebra, having two associative products; the 'frozen' product $\cdot$ and the Grossman-Larson (GL) product $*$. The so-called frozen exponential $\exp^{\cdot}(tf)$ defines flow along geodesics, such as the action of a Lie group on a homogeneous space. The GL exponential $\exp^{*}(tf)$ defines the exact flow corresponding to $\dot{y} = f(y)$. The basic problem of numerical integration, to express $\exp^{*}(tf)$ in terms of $\exp^{\cdot}(tf)$ is achieved by Runge–Kutta methods in the pre-Lie case and Lie group integrators in the post-Lie case.

The geometric understanding of invariant connections on the tangent bundle was developed by Nomizu [16]. A manifold with a tangent bundle connection being flat and torsion free, $R = T = 0$, is locally a Euclidean space (abelian Lie group). If the tangent bundle admits a flat, constant torsion connection with $\nabla T = 0$ and $R = 0$, the manifold is locally a Lie group. And, if the connection is torsion free with constant curvature ($T = 0$, $\nabla R = 0$), the space is a locally symmetric space.

A major goal has been to understand algebras of general connections. This goal has eluded a systematic study, since the central algebraic structures of pre- and post-Lie algebras rely on an extension of the binary product $\triangleright$ to the enveloping algebra $U(\mathcal{XM})$ such that $f \triangleright (g \triangleright h) = (f * g) \triangleright h$. This is only possible if $R = 0$.

Lie–Butcher series for general connections. A new insight in this work bypasses the obstacle of extending the connection to the enveloping algebra in the general case $R \neq 0$. Indeed, rather than lifting the connection to the enveloping algebra, we extend it to the tensor algebra over $\mathcal{XM}$. Let $D = T(\mathcal{XM}, \triangleright)$, the tensor algebra where the connection is extended to $\triangleright\colon D \times D \to D$ according to

$$f \triangleright (\omega \cdot \tilde{\omega}) = (f \triangleright \omega) \cdot \tilde{\omega} + \omega \cdot (f \triangleright \tilde{\omega})$$
$$(f \cdot \omega) \triangleright \tilde{\omega} = f \triangleright (\omega \triangleright \tilde{\omega}) - (f \triangleright \omega) \triangleright \tilde{\omega}$$

for $f \in \mathcal{XM}$ and $\omega, \tilde{\omega} \in D$. Here $f \cdot \omega$ is the tensor product (concatenation), which plays the role of the frozen product. D has a Hopf algebra structure with the product $\cdot$ and the classical de-shuffle coproduct $\Delta_{\sqcup\!\sqcup}(\alpha) = \sum \alpha_{(1)} \otimes \alpha_{(2)}$ (employing Sweedler's notation). The primitive elements of D

$$\mathfrak{g} = \mathrm{prim}(D) = \{\, \alpha \in D \mid \Delta_{\sqcup\!\sqcup}(\alpha) = \alpha \otimes \mathbb{I} + \mathbb{I} \otimes \alpha \,\}$$

yield a post-Lie algebra such that $U(\mathfrak{g}) = D$. Hence, the algebraic machinery developed in [4] can be employed also in this case. In particular we have the

associative GL product $*\colon D \times D \to D$

$$\alpha * \beta = \sum \alpha_{(1)} \cdot (\alpha_{(2)} \triangleright \beta),$$

satisfying the desired identity

$$f \triangleright (g \triangleright h) = (f * g) \triangleright h.$$

It should be noted that the products are defined on D, not on the enveloping algebra $U(\mathcal{XM})$. In particular the GL commutator $[f, g]_* = f * g - g * f$ on $\mathfrak{g}$ maps to the Jacobi bracket on $\mathcal{XM}$ only when $R = 0$.

Still, the basic properties of the exponential hold, $\exp^{\cdot}(tf)$ defines flow along geodesics, and $\exp^{*}(tf)$ the exact flow, so we are able to employ the full machinery of post-Lie algebras and Lie–Butcher series also for the study of geodesics of general connections.

Interestingly, the geometry of the connection appears in the mapping down from $\mathfrak{g} \subset D$ to $\mathcal{XM}$. Inspired from the series developments of the double exponential in [7], we can prove a general result:

Theorem 0.1. [5] *Any Lie polynomial in* $\mathfrak{g}$ *maps down to a linear combination of the torsion and curvature tensors,* T *respectively* R*, and their covariant derivatives on the space* $\mathcal{XM}$ *of vector fields on a manifold* $\mathcal{M}$*.*

Thus Lie group integrators can be generalised to this setting, where commutators are computed in terms of the torsion and curvature tensors.

References

[1] J.C. Butcher, B-Series – Algebraic Analysis of Numerical Methods, Springer Series in Computational Mathematics **55**, Springer International Publishing, 2021

[2] F. Chapoton and M. Livernet, *Pre-Lie algebras and the rooted trees operad*, International Mathematics Research Notices **2001** (2001), 395–408.

[3] C. Curry, K. Ebrahimi-Fard and H.Z. Munthe-Kaas, *What Is a Post-Lie Algebra and Why Is It Useful in Geometric Integration*. In Numerical Mathematics and Advanced Applications (ENUMATH 2017). Lecture Notes in Computational Science and Engineering **126**. Springer, Cham., (2019).

[4] K. Ebrahimi-Fard, A. Lundervold and H.Z. Munthe-Kaas, *On the Lie enveloping algebra of a post-Lie algebra*, Journal of Lie Theory **25** (2015), 1139–1165.

[5] K. Ebrahimi-Fard, D. Manchon, and H.Z. Munthe-Kaas, *Algebras of general connections and Lie–Butcher series for geodesic flows*, preprint 2021.

[6] G. Floystad, H. Munthe-Kaas *Pre-and post-Lie algebras: the algebro-geometric view*, Springer Abel Symposia **13**, Proceedings of the Abel Symposium 2016.

[7] A. V. Gavrilov, *Algebraic properties of the covariant derivative and composition of exponential maps*, Siberian Advances in Mathematics **16** (2006), 54–70.

[8] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett and A. Zanna, *Lie-group methods*, Acta Numerica **9** (2000), 215–365.

[9] A. Lundervold, and H.Z. Munthe-Kaas, *Hopf algebras of formal diffeomorphisms and numerical integration on manifolds*, Contemporary Mathematics **539** (2011), 295–324.

[10] A. Lundervold and H.Z. Munthe-Kaas, *Backward error analysis and the substitution law for Lie group integrators*, Foundations of Computational Mathematics, **13** (2013), 161–186.

[11] A. Lundervold and H.Z. Munthe-Kaas, *On post-Lie algebras, Lie–Butcher series and moving frames*, Foundations of Computational Mathematics **13** (2013), 583–613.

[12] D. Manchon, *A short survey on pre-Lie algebras*, E. Schrödinger Institute Lectures in Mathematical Physics, European Mathematical Society, A. Carey Ed., 2011.
[13] H.Z. Munthe-Kaas, *Lie–Butcher theory for Runge–Kutta methods*, BIT Numerical Mathematics **35** (1995), 572–587.
[14] H.Z. Munthe-Kaas, *Runge–Kutta methods on Lie groups*, BIT Numerical Mathematics **38** (1998), 92–111.
[15] H.Z. Munthe-Kaas and W.M. Wright, *On the Hopf algebraic structure of Lie group integrators*, Foundations of Computational Mathematics **8** (2008), 227–257.
[16] K. Nomizu, *Invariant affine connections on homogeneous spaces*, American Journal of Mathematics **76** (1954), 33–65.

Low regularity integrators for the Gross-Pitaevskii equation

Katharina Schratz
(joint work with Yvonne Bronsard Alama and Yvain Bruned)

A general low regularity framework for nonlinear evolution equations

$$\partial_t u - \mathcal{L}u = f(u, \overline{u}) \qquad (t, x) \in \mathbb{R} \times \Omega \tag{1}$$

was recently introduced in [5] given $\Omega \subset \mathbb{R}^d$ and an initial condition

$$u_{/t=0} = u_0. \tag{2}$$

When $\partial\Omega \neq \emptyset$ the problem shall be equipped with some appropriate homogeneous boundary conditions. The main assumptions on (1) are the following. The linear operator $\mathcal{L}$ is defined on a Hilbert space X of complex valued functions $u \in \mathbb{C}$ with norm denoted by $\|\cdot\|$ and domain $\mathcal{D}(\mathcal{L})$. The complex conjugation $u \mapsto \overline{u}$ is assumed to be an isometry on X. When $\partial\Omega \neq \emptyset$, the boundary conditions will be encoded in the choice of the domain of the operator $\mathcal{L}$. The operator $\overline{\mathcal{L}}$ is defined by $\overline{\mathcal{L}}u = \overline{\mathcal{L}\overline{u}}$. In addition, it is assumed that the nonlinearity f is tensorized under the form

$$f(v, w) = \mathcal{B}\left(F(v) \cdot G(w)\right), \quad F, G : \mathbb{C} \to \mathbb{C}^J,$$

where we use the notation $X \cdot Y = \sum_k X_k Y_k, \quad X, Y \in \mathbb{C}^J$, and $\mathcal{B}$ is a linear operator. The main assumptions on $\mathcal{L}$ are the following

i) $\mathcal{L}$ generates a strongly continuous semigroup $\{e^{t\mathcal{L}}\}_{t\geq 0}$ of contractions on X;
ii) $\mathcal{A} = -\mathcal{L} + \overline{\mathcal{L}}$ generates a group $\{e^{t\mathcal{A}}\}_{t\in\mathbb{R}}$ of unitary operators on X;
iii) $\mathcal{L}$ and $\overline{\mathcal{L}}$ commute: $[\mathcal{L}, \overline{\mathcal{L}}] = 0$.

Under the above assumptions the first-order low regularity Duhamel integrator for (1) takes the form

$$u^{\ell+1} = e^{\tau\mathcal{L}}\Big(u^\ell + \tau\mathcal{B}\left(F(u^\ell) \cdot \varphi_1\big(\tau\mathcal{A}\big)G(\overline{u^\ell})\right)\Big) \quad \text{with} \quad \varphi_1(z) = \frac{e^z - 1}{z} \tag{3}$$

and where we set $u^0 = u_0$. The new scheme (3) in general allows for convergence in a more general setting (for rougher data) than classical methods, such as splitting or exponential integrator methods.

With the framework introduced in [5] we can deal in a unified framework with parabolic, hyperbolic, dispersive as well as mixed equations. However, an important class of equations is missing, namely equations with a potential (or noise).

We close this gap and propose a novel low regularity integrator for the Gross-Pitaevskii (GP) equation

$$i\partial_t u(t,x) = -\Delta u(t,x) + V(x)u(t,x) + |u(t,x)|^2 u(t,x) \quad (t,x) \in \mathbb{R} \times \Omega$$

with non-smooth potential $V(x)$. At first order our low regularity integrator for GP will take the form

$$u^{n+1} = \Phi^{\tau}_{\mathrm{GP}}(u^n) = e^{i\tau\Delta}\Big[u^n - i\tau\left(u^n\varphi_1(-i\tau\Delta)V + (u^n)^2\,\varphi_1(-2i\tau\Delta)\overline{u}^n\right)\Big]. \tag{4}$$

In contrast to classical approximation techniques, such as splitting or exponential integrator methods (see, e.g., [2, 3, 4]), the local error of the low regularity GP integrator (4) will only require the boundedness of first instead of second order spatial derivatives of the potential V and solution u.

Details on the construction and error analysis as well as higher order extensions, following new techniques based on decorated trees series analysis inspired by singular SPDEs (cf. [1]), will be given elsewhere.

References

[1] Y. Bruned, K. Schratz, *Resonance based schemes for dispersive equations via decorated trees*, preprint (2020) http://arxiv.org/abs/2005.01649

[2] D. Cohen, L. Gauckler, *One-stage exponential integrators for nonlinear Schrödinger equations over long times.* BIT 52:877–903 (2012).

[3] T. Jahnke, C Lubich, *Error bounds for exponential operator splittings.* BIT 40:735–744 (2000).

[4] C. Lubich, *On splitting methods for Schrödinger–Poisson and cubic nonlinear Schrödinger equations.* Math. Comp. 77:2141–2153 (2008).

[5] F. Rousset, K. Schratz, *A general framework of low regularity integrators*, to appear in SIAM J. Numer. Anal., http://arxiv.org/abs/2010.01640

Learning effective models from multiscale data: filtering and Bayesian inference

Assyr Abdulle

(joint work with Giacomo Garegnani, Grigorios A. Pavliotis, Andrew M. Stuart, Andrea Zanoni)

1. Introduction

The problem of estimating a stochastic model from time series is important in many disciplines (e.g., chemistry, atmosphere-ocean science [5], econometrics [3], ...). Classes of models used for such inference problems are often based on stochastic differential equations (SDE) of the form

$$\mathrm{d}X_t = f(X_t)\,\mathrm{d}t + g(X_t)\,\mathrm{d}W_t.$$

Inferring the drift $f(x)$ vector and the diffusion tensor $\sigma(x) = g(x)g(x)^T$ from time series is in general challenging. A major issue for such problems is that of model misspecification, when the data is not consistent with the chosen class of models. In this report we describe a new approach to learn coarse grained-models (dynamics at slow time scales) from multiscale data, based on filtering techniques. We show that robust parameter estimation can be derived and that for a class of fast/slow SDEs the theory of homogenization enables a rigorous study of the inference problem.

2. Fast slow SDEs and homogenization

We assume that the given data arise from the following class of overdamped multiscale Langevin SDEs

$$\mathrm{d}X_t^\varepsilon = -\alpha \cdot V'(X_t^\varepsilon)\,\mathrm{d}t - \frac{1}{\varepsilon}p'\left(\frac{X_t^\varepsilon}{\varepsilon}\right)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t, \tag{1}$$

that model the motion of particles in a confining potential which has slow variations $V(x)$ with rapid oscillations superimposed $p(x/\varepsilon)$. Here $\varepsilon > 0$ represents a characteristic size of the small scales in the problem and W_t is a standard one-dimensional Brownian motion. For the rest of the paper we will assume that the fast scale is periodic. We also will assume that $\sigma > 0$, $\alpha \in \mathbb{R}^N$, $V\colon \mathbb{R} \to \mathbb{R}^N$, $V(x) = (V_1(x), V_2(x), \ldots, V_N(x))^\top$, $p\colon \mathbb{R} \to \mathbb{R}$, L-periodic, with $p, V_i \in C^\infty(\mathbb{R})$. For the slow scale potential we further assume that its components V_i and $V_i^{'}$ are polynomially bounded, that V' is Lipschitz continuous and that there exist $a, b > 0$ such that $-a + bx^2 \leq \alpha \cdot V'(x)x$.

Coarse-grained models. The class of models to be fitted to multiscale data are the following "homogenized models"

$$\mathrm{d}X_t = -A \cdot V'(X_t)\,\mathrm{d}t + \sqrt{2\Sigma}\,\mathrm{d}W_t. \tag{2}$$

Under the assumptions above it is possible to show, via homogenization theory, that $X_t^\varepsilon \to X_t$ in law for $\varepsilon \to 0$ [4, Chapter 3]. As mentioned in the introduction, the goal is then to infer the drift coefficient A and the diffusion coefficients Σ from the multiscale data $X^\varepsilon = (X_t^\varepsilon, 0 \leq t \leq T)$.

3. Parameter inference, maximum likelihood estimator.

A classical way to approximate effective drift coefficients A from the coarse-grained observations X (2) is via path-space likelihood expressing the probability of a model X given a drift coefficient $\tilde{A}$

$$L(X \mid \tilde{A}) = \exp\left(-\frac{I(X \mid \tilde{A})}{2\Sigma}\right). \tag{3}$$

Maximizing the functional $L(X \mid \tilde{A})$ with respect to $\tilde{A}$ gives the maximum likelihood estimator (MLE) $\widehat{A}(X,T)$ of A defined by

$$\arg\min_{\tilde{A}\in\mathbb{R}^N} I(X \mid \tilde{A}) = -\left(\int_0^T V'(X_t)\otimes V'(X_t)\,\mathrm{d}t\right)^{-1}\int_0^T V'(X_t)\,\mathrm{d}X_t. \tag{4}$$

The above procedure is well understood. Our goal is however different: estimate $A \in \mathbb{R}^N$ from the multiscale observations $X^\varepsilon = (X_t^\varepsilon, 0 \le t \le T)$. As $X_t^\varepsilon \to X_t$ for $\varepsilon \to 0$ it seems reasonable to define $\widehat{A}(X^\varepsilon, T)$ for the MLE of A with multiscale data. But this turn out to be a wrong approach, indeed, under the assumptions of Section 2 this approach is shown to be biased [8, Thm. 3.4]

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty}\widehat{A}(X^\varepsilon,T) = \alpha. \tag{5}$$

Parameter inference based on subsampling. The following MLE is introduced in [8] (written here for $N = 1$ for simplicity) based on subsampling the data with step δ

$$\widehat{A}_\delta(X^\varepsilon,T) = -\frac{\sum_{i=0}^{M-1} V'(X_{i\delta}^\varepsilon)\left(X_{(i+1)\delta}^\varepsilon - X_{i\delta}^\varepsilon\right)}{\delta\sum_{i=0}^{M-1} V'(X_{i\delta}^\varepsilon)^2}, \qquad M\delta = T. \tag{6}$$

It is shown in [8], again under the assumptions of Section 2, that (6) is an asymptotically unbiased estimator of A in the limit for $\varepsilon \to 0$: if $\delta = \varepsilon^\zeta$, $0 < \zeta < 1$ and $M = \lceil \varepsilon^{-\gamma} \rceil$ with $\gamma > \zeta$, then

$$\lim_{\varepsilon\to 0}\widehat{A}_\delta(X^\varepsilon,T) = A, \quad \text{in probability.} \tag{7}$$

One of the main drawbacks of this approach is its lack of robustness. Indeed for a given T and ε the error depends on the choice of ζ, and it is unknown how to quantify its optimal value (see Figure 1 and [1, Section 5.1.2]). We note that other approaches based on martingale property [7], operator eigenpairs [6] have been developed (we refer to [1] for a more comprehensive literature overview).

Parameter inference based on filtering. We note that subsampling data is a "smoothing" process, so why not directly smoothing the data ? We therefore introduce a filtered process

$$Z_t^\varepsilon = \int_0^t k(t-s)X_s^\varepsilon\,\mathrm{d}s, \tag{8}$$

where the filter $k(r)$ is given by

$$k(r) = C_\beta\delta^{-1/\beta}e^{-\frac{r^\beta}{\delta}}, \quad C_\beta = \beta\,\Gamma(1/\beta)^{-1}, \quad \delta,\beta > 0. \tag{9}$$

For the rest of the paper we assume $\delta > 0, \beta = 1$. In this case the filter has the simple expression $k(r) = \frac{1}{\delta}e^{-\frac{r}{\delta}}$ and we can derive a coupled system of SDEs

$$\begin{aligned}\mathrm{d}X_t^\varepsilon &= -\alpha\cdot V'(X_t^\varepsilon)\,\mathrm{d}t - \frac{1}{\varepsilon}p'\left(\frac{X_t^\varepsilon}{\varepsilon}\right)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t,\\ \mathrm{d}Z_t^\varepsilon &= \frac{1}{\delta}\left(X_t^\varepsilon - Z_t^\varepsilon\right)\,\mathrm{d}t.\end{aligned}$$

It can then be shown that $(X_t^\varepsilon, Z_t^\varepsilon)^\top$ is geometrically ergodic with smooth invariant density $\mu^\varepsilon(\mathrm{d}x, \mathrm{d}z) = \rho^\varepsilon(x,z)\,\mathrm{d}x\,\mathrm{d}z$ that is the solution of an explicit Fokker Planck equation. We then define the filtered MLE by

$$\widehat{A}_k(X^\varepsilon, T) = -\left(\int_0^T V'(Z_t^\varepsilon) \otimes V'(X_t^\varepsilon)\,\mathrm{d}t\right)^{-1} \int_0^T V'(Z_t^\varepsilon)\,\mathrm{d}X_t^\varepsilon. \tag{10}$$

We note (see [1] for a comprehensive explanation)

- $\widehat{A}_k(X^\varepsilon, T)$ is well defined if $\det\left(\int_0^T V'(Z_t^\varepsilon) \otimes V'(X_t^\varepsilon)\,\mathrm{d}t\right) \neq 0$;
- it is essential to keep $\mathrm{d}X_t^\varepsilon$ and $V'(X_t^\varepsilon)$ to prove unbiasedness;
- $\widehat{A}_k(X^\varepsilon, T)$ has to be thought as a perturbation of $\widehat{A}(X^\varepsilon, T)$ at the level of the estimator (i.e., after the maximization process of MLE);
- $\tilde{A}_k(X^\varepsilon, T) = -\left(\int_0^T V'(X_t^\varepsilon) \otimes V'(X_t^\varepsilon)\,\mathrm{d}t\right)^{-1} \int_0^T V'(Z_t^\varepsilon)\,\mathrm{d}X_t^\varepsilon$ is also a valid estimator in the non-homogenized regime (when δ depends on ε).

For this estimator, under the assumptions of Section 2, we can prove [1]

Theorem 3.1 (homogenization regime)**.** *If δ is independent of ε*

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty} \widehat{A}_k(X^\varepsilon, T) = A, \quad \textit{a.s.}$$

Theorem 3.2 (multiscale regime)**.** *If $\delta = \varepsilon^\zeta$, $\quad \zeta \in (0,2)$*

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty} \widehat{A}_k(X^\varepsilon, T) = A, \quad \textit{in probability.}$$

The value $\zeta = 2$ is critical, indeed

Theorem 3.3 (switch to biasedness)**.** *If $\delta = \varepsilon^\zeta$, $\quad \zeta > 2$*

$$\lim_{\varepsilon\to 0}\lim_{T\to\infty} \widehat{A}_k(X^\varepsilon, T) = \alpha, \quad \textit{in probability.}$$

For the diffusion coefficient, the estimator $\widehat{\Sigma}_k(X^\varepsilon, T) := \frac{1}{\delta T}\int_0^T (X_t^\varepsilon - Z_t^\varepsilon)^2\,\mathrm{d}t$, for Σ based on filtering can be employed and proved to be unbiased.

Discussion. MLE based on filtering are robust in practice with respect to the parameter of the filter in contrast to estimators based on subsampling (see e.g. Figure 1, where the subsampling size is not optimal, but hard to find for this example). The MLE based on filtering has also been extended to the Bayesian setting to allow for a probability distribution for the effective drift A and uncertainty quantification. Finally, we note that in many applications only discrete measurements of the diffusion process are available. Recently, using the filtering approach developed in this paper and martingale estimating functions a new estimator for learning homogenised SDEs from noisy discrete data has been introduced [2].

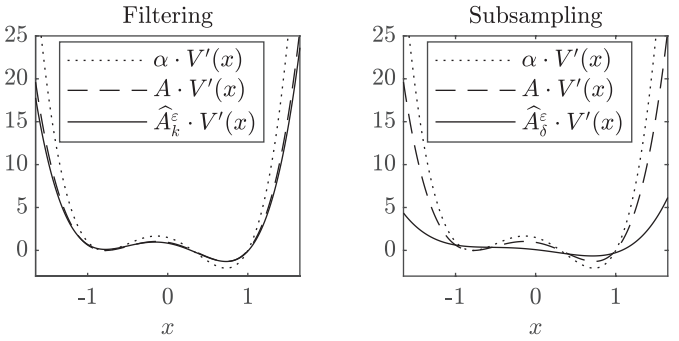

FIGURE 1. Multi-dimensional problem $N = 4$, $V_i(x) = T_i(x)$ i-th Chebyshev polynomials, $p^{\epsilon}(x) = \cos(x/\epsilon)$, $\alpha = (-1, -1/2, 1, 1/2)$, $\varepsilon = 0.05$, $T = 10^3$, subsampling $\delta = \varepsilon^{2/3}$, taken from [1].

REFERENCES

[1] A. Abdulle, G. Garegnani, G. A. Pavliotis, A. M. Stuart, and A. Zanoni, *Drift estimation of multiscale diffusions based on filtered data* (2020), arXiv preprint arXiv:2009.13457.

[2] A. Abdulle, G. A. Pavliotis, and A. Zanoni, *Eigenfunction martingale estimating functions and filtered data for drift estimation of discretely observed multiscale diffusions* (2021), arXiv preprint arXiv:2104.10587.

[3] Y. Aït-Sahalia and J. Jacod, *High-frequency financial econometrics*, Princeton University Press (2014).

[4] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic analysis for periodic structures*, North-Holland Publishing Co., Amsterdam (1978).

[5] C. J. Cotter and G. A. Pavliotis, *Estimating eddy diffusivities from noisy Lagrangian observations*, Commun. Math. Sci. **7** (2009), 805–838.

[6] D. Crommelin and E. Vanden-Eijnden, *Diffusion estimation from multiscale data by operator eigenpairs*, Multiscale Model. Simul. **9** (2011), 1588–1623.

[7] S. Kalliadasis, S. Krumscheid, and G. A. Pavliotis, *A new framework for extracting coarse-grained models from time series with multiscale structure*, J. Comput. Phys. **296** (2015), 314–328.

[8] G. A. Pavliotis and A. M. Stuart, *Parameter estimation for multiscale diffusions*, J. Stat. Phys. **127** (2007), 741–781.

Constraint-based regularization of neural networks

BENEDICT LEIMKUHLER

(joint work with Timotheé Pouchon, Amos Storkey, Tiffany Vlaar)

View a neural network as defining a map $x \in \mathbb{R}^m, \theta \in \mathbb{R}^d \mapsto \Phi(\theta, x) \in \mathbb{R}^n$ and suppose that the goal is to use this form to characterize an unknown abstract functional relationship $f : \mathbb{R}^m \to \mathbb{R}^n$. To this end, one selects a training data set D–a set of input-output pairs $(x_\alpha \in \mathbb{R}^m, y_\alpha = f(x_\alpha) \in \mathbb{R}^n)$–and trains the model to represent this data, i.e., so that $y_\alpha \approx \Phi(\theta, x_\alpha)$. Neural network training can be

viewed as solving an optimization problem:

$$\theta = \text{argmin}_\theta L(\theta, D),$$

where $L(\theta, D)$ is a loss function which quantifies the defect in the neural network input-output relationship for the given training data D. There are often so many parameters compared to the data that it is possible to train the network (find parameters) so that on the training data set $y_\alpha \equiv \Phi(\theta, x_\alpha)$, i.e to zero loss. In fact there are likely to be many choices of parameters for which this relationship is exact or very accurate for all training data. The challenge is then to identify the parameters so that the learned functional relationship generalizes to a large family of unseen ("test") data; this is precisely why the choice of training procedure is important.

This formulation is ill-posed until we explain what we mean by "generalizes" and describe carefully the training data set. If all our training data lies on a low-dimensional submanifold, that is $x_\alpha \in M$ where $\dim M < m$. Then it should be clear that we cannot hope to represent the functional relationship for data points far from M. One way to proceed is to assume that there exists an underlying data distribution ϱ (for example, the distribution of all handwritten digits represented by 28×28 grayscale pixel arrays, as in the MNIST data set) and that the training data is a representative sampling of ϱ. Once this is assumed the goal becomes that of representing other unseen samples drawn from ϱ. Alternatively, one may avoid issues of explicitly characterizing the dataset and adopt a "Bayesian perspective" by describing the posterior parameter distribution as the product of a prior π_0 and a likelihood function $\pi((x, y)|\theta)$ (based on the neural network model); this is the approach we have taken in our work. Bayes' formula gives

$$\pi(\theta|D) \propto \pi_0(\theta) \prod_{\alpha=1}^{|D|} \pi((x_\alpha, y_\alpha)|\theta).$$

In typical practice the training is associated to maximizing π (i.e. finding its modes), which can be seen to be equivalent to minimizing the loss function L defined by $L(\theta, D) \equiv -\ln \pi(\theta|D)$.

Neural networks have discontinuous derivatives, something which limits and complicates analysis. Despite this, neural networks are typically trained using a gradient descent procedure. Due to the large data sets involved, the full gradient is replaced by an approximation. If $G(\theta) \equiv \nabla_\theta L(\theta, D)$, then we replace this by $\tilde{G} \approx G$, typically defined by restricting the gradient loss calculation via stochastic subsampling of the training data, and take Euler steps on the flow

$$\frac{d\theta}{dt} = -\tilde{G}.$$

This method is surprisingly effective, due probably the restriction to certain functional forms in defining the nodes of the network: neural networks are normally Lipschitz continuous and have piecewise smooth derivatives. Since neural network training is not just optimization, but optimization which "generalizes well", one may choose modified forms of the loss function (or perturbations thereof) in order

to enhance the latter property. These modifications are often described broadly as "regularization". Good neural network training relies on finding both a good neural network model and also a suitable regularization in order to drive the optimizer towards highly generalizable optima, that is ones that give an accurate representation of the function f on unseen (but in some sense nearby) data to that used in training. Examples of regularizations include: weight decay (adding an L_2 penalty term in the weights to the loss function), batch normalization and other renormalization schemes (which rescale and/or shift the parameters in each layer of a neural network), dropout and weight thinning (which introduce a stochastic sparsification of the network), and early stopping (whereby the training is automatically interrupted before the optimum is reached).

Since some neural network training procedures can be viewed as either introducing a normalizing step or penalty term it is natural to consider removing these in favor of introducing constraints, for example changing the stochastic gradient descent dynamics to

$$\dot{\theta} = -\tilde{G}(\theta) - h'(\theta)^T \lambda, \tag{1}$$

$$0 = h(\theta). \tag{2}$$

Note that the form here masks some complexity since the stochastic gradient that appears here does not have a trivial expression in terms of Gaussian noise perturbation of the true gradient (the stochastic gradient noise is spatially anisotropic). The constraints can be designed in such a way as to limit the Lipschitz bound for the neural network model.

In our work we have been exploring alternative approaches to neural network training which further replace the stochastic gradient descent method itself by a discretized stochastic differential equation which incorporates small additive noise in all directions (overdamped Langevin dynamics) as in the stochastic gradient Langevin dynamics method of Welling and Teh [4], in all force components (underdamped Langevin dynamics), or using a thermostat control like those popular in molecular dynamics (e.g. Nosé-Hoover/Adaptive Langevin as in [5]). This can be shown to have a regularizing effect on the model parameterization [1] (Figs. 1 and 2). It also helps to overcome barriers in the loss landscape that may limit exploration at early stages. (Although it is likely true that in the heavily overparameterized regime the barriers diminish and one gradually convexifies the problem–see e.g. [2] for a theoretical discussion in a simplified setting—some problems can be handled well short of this heavily overparameterized regime with consequent benefits due to smaller weight spaces, tightly wound spiral data for example–see the figures).

The framework of additive noise coupled with stochastic gradient descent replaces the constrained system (1)-(2) by a system of constrained SDEs [3]. The properties of these SDEs are a little difficult to study due to the stochastic gradient terms, but we can study them in the limit of an accurate gradient to explore their diffusive (hypocoercive/hypoelliptic) properties. In recent years, we have made significant strides in the area of numerical methods for such SDEs which allow

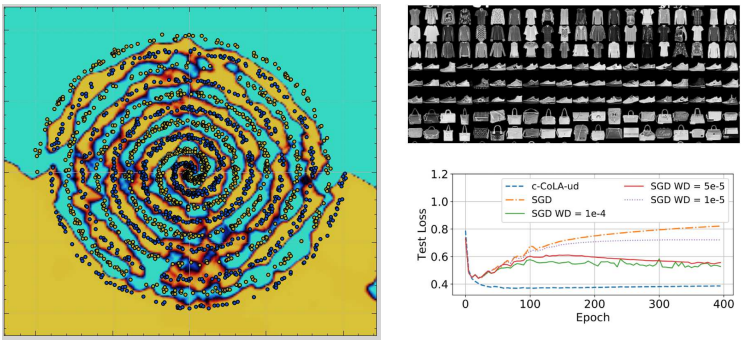

FIGURE 1. Illustrations. Left: (unpublished) tightly wound spiral data treated using a single hidden layer perceptron with 500 hidden nodes (far below the overparameterized limit). Colored domains represent classification regions generated by the neural network. Training was performed to 97% test accuracy using a constrained variant of the AdLaLa method of [1] with large interval (circle) constraints on the output weights (radius = 80) and much milder interval constraints on input weights (radius =1) and biases (radius=2). Right: results for the "fashion MNIST" training dataset for classification consisting of labelled raster images of clothing items shows clear signs of overfitting in an unconstrained SGD scheme (not mitigated by weight decay) whereas our underdamped Langevin algorithm with constraints (c-CoLA-ud) demonstrates monotonicity in the loss curve–details in [3].

the calculation of reliable long term (equilibrium) averages by looking at both unconstrained and constrained SDEs and their discretization by splitting methods, see e.g. [6, 7]. As an illustration of the potential benefits of using additive noise in conjunction with constraints, see Fig 2 where we applied stochastic gradient Langevin dynamics (SGLD) with constraints to spiral data classification with significant improvement on SGD with constraints. We are currently exploring the use of such SDEs and discretization schemes in connection with constraint manifolds that are designed to provide relatively simple but effective control of the weights during training. An elementary (but useful) example is interval constraints of the form $|\theta_i| \leq \bar{\theta}$ which can be introduced elegantly by adding a slack variable ξ_i and the constraint $\theta_i^2 + \xi_i^2 = \bar{\theta}^2$. We have also examined more complicated constraints, for example orthogonality constraints $W_i^T W_i = I$ based on weight matrices W_i which appear naturally at the ith layer of a deep network. In some cases these constraints can dramatically improve the generalization performance of neural networks. We are currently improving these algorithms and studying their application in a wide range of settings, including for problems in image analysis

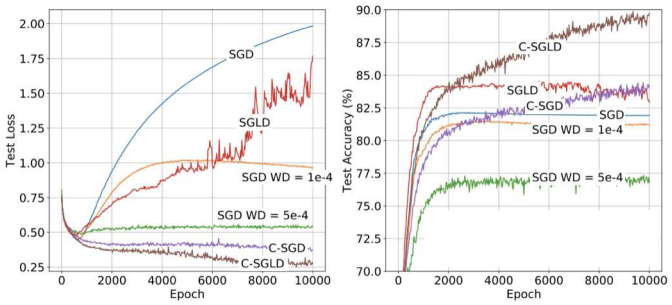

FIGURE 2. The figure shows the test loss curve at left and the test accuracy on the right for spiral data classification using several different algorithms to train a two-layer perceptron, including methods that incorporate additive noise (SGLD, see [4] and its constrained variant C-SGLD). C-SGLD also significantly improves on the use of "weight-decay" (WD). See [3] for more detail.

(fashion MNIST, CIFAR-10), discrete datasets (e.g. Chess endgames and poker hand classification), molecular structure analysis and natural language processing.

REFERENCES

[1] B. Leimkuhler, C. Matthews, and T. Vlaar, Partitioned integrators for thermodynamic parameterization of neural networks, *Foundations of data science* **1**(4): 457-489, 2019.

[2] S. Arora, N. Cohen, and E. Hazan, On the Optimization of Deep Networks: Implicit Acceleration by Overparameterization, Proceeedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, PMLR 80, 2018.

[3] B. Leimkuhler,T. Pouchon, T. Vlaar, and A. Storkey, Constraint-based regularization of neural networks, Procedings of Neurips OPT2020: 12th Annual Workshop on Optimization for Machine Learning, 2020.

[4] M. Welling and Y.-W. Teh, Bayesian Learning via Stochastic Gradient Langevin Dynamics, Proceeedings of the 35th International Conference on Machine Learning, 2011.

[5] A. Jones and B. Leimkuhler, Adaptive stochastic methods for sampling driven molecular systems, *Journal of Chemical Physics* **135**, 84125, 2011.

[6] B. Leimkuhler, C. Matthews and G. Stoltz, The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics, *IMA Journal on Numerical Analysis* **36**, 13–79,2016.

[7] B. Leimkuhler and C. Matthews, Efficient molecular dynamics using geodesic integration and solvent-solute splitting. *Proceedings of the Royal Society* **A 472**,20160138, 2016.

Recent advances in structure preserving dynamical low-rank algorithms

Lukas Einkemmer

(joint work with Ilon Joseph, Jingwei Hu, Lexing Ying, Christian Lubich)

Kinetic equations are used extensively in a range of applications from plasma physics to radiative heat transfer. However, because they are formulated in an up to six-dimensional phase space their numerical solution is extremely challenging. A direct discretization suffers from the fact that the required degrees of freedom scale as $O(n^6)$, where n is the number of grid points in each direction. This is called the curse of dimensionality. While many improvements, both in designing better numerical methods and implementing them efficiently on modern computer hardware (see, e.g., [17, 6, 4, 1, 12]) have been made, this remains a fundamental limitation for the widespread use of kinetic simulations.

Complexity reduction techniques offer the possibility to significantly reduce this computational burden. However, for kinetic equations many classic methods (such as sparse grids) have only met with limited success [14]. Because of this, particle/Monte Carlo methods are still widely used. However, this approach suffers from numerical noise that only decreases slowly as the number of particles are increased and are not able to resolve regions of phase space with low density.

Recently, dynamical low-rank algorithms based on a projector splitting approach have been identified as an extremely promising complexity reduction technique for kinetic problems. In the simplest case, this takes the form of a singular value decomposition (SVD). The dynamical low-rank approach, which we consider here, formulates the equations of motions in terms of the low-rank factors. The low-rank factors are lower dimensional functions that only depend on a subset of the phase space variables.

For example, for the Vlasov equation, a typical kinetic problem,

$$\partial_t f(t,x,v) + v \cdot \nabla_x f(t,x,v) - E(t,x) \cdot \nabla_v f(t,x,v) = 0$$

a dynamical low-rank approximation employs $X_i(t,x)$, which depend only on the physical space variables x, and $V_j(t,v)$, which depend only on velocity space variables v. A numerical solution is then sought in the following form

$$f(t,x,v) \approx \sum_{i=1}^{r} \sum_{j=1}^{r} X_i(t,x) S_{ij}(t) V_j(t,v),$$

where r is the rank of the approximation and we note that $S(t) \in \mathbb{R}^{r \times r}$ does neither depend on physical space nor on velocity space variables. A set of partial differential equations is then derived that describe the dynamics of the low-rank factors constrained to the approximation space (a manifold of low-rank functions with rank r). That is, we obtain partial differential equations for the low-rank factors X_i, V_i, and S_{ij}. The main computational saving is due to the fact that these functions are (at most) three dimensional. We also mention that robust integrators for the resulting equations have been proposed [16, 15, 2].

It has been demonstrated that such dynamical low-rank approximations can be very effective for kinetic equations [9, 7, 19]. Both for problems in plasma physics as well as for transport problems/radiative transfer. Recently, some advances have been made to understand why and in which situations low-rank approximations are particularly effective. In this context we mention that the solution of the linearized Vlasov equation is low-rank [11] and that, in many case, diffusive and fluid limits of collisional kinetic problems have a low-rank structure that can be captured by well designed numerical methods [3, 5, 7].

Nevertheless, a well known property of the low-rank approach, particularly relevant to the present workshop, is that it destroys the physical structure of the underlying equation. In the context of kinetic equations this means, e.g., that mass, momentum, and energy are no longer conserved if a classic low-rank approximation is applied. This situation is in stark contrast to direct solvers or particle methods, where usually at least mass and momentum are conserved up to machine precision.

While some proposed remedies, such as [10, 19, 18], can alleviate this issue to some extend, until very recently there has not been a low-rank approximation that for nonlinear problems naturally conserves the invariants of the solution. The key insight to achieve this is to recognize that if certain fixed basis functions are appended to the approximation space then conservation of mass, momentum, and energy can be achieved. However, in the analytical formulation neither of these functions belong to the usually used L^2 space. This is remedied by using a suitably modified function space. The approach requires the derivation of new evolution equations for the low-rank factors. To be more precise, the test functions in the variational problem have to be chosen in such a way as to be compatible with the fixed basis functions and modified function spaces. This approach yields a mass, momentum, and energy conservative dynamical low-rank scheme which can then be combined with appropriate conservative time and space discretizations. It also exactly satisfies the underlying conservation laws that give rise to these invariants. In the PDE setting this is even more important for accurate and stable long time integration than just conserving the invariants. For more information we refer the reader to [8].

Another setting in which dynamical low-rank approximations play an important role are kinetic systems that are close to either the diffusive or fluid limit [13]. For many problems such as linear transport or the Boltzmann equation in the weakly compressible regime, the corresponding limit is low-rank and, as has been shown in the literature, asymptotic preserving schemes can be constructed that preserve this structure numerically [3, 5, 7]. In this situation dynamical low-rank algorithms are particularly effective because it is known that the solution can be represented with a small rank.

However, until very recently the arguably most interesting case, namely the Navier–Stokes limit of the Boltzmann equation, could not be treated in that way. The reason for this is that the limit, given by a Maxwellian, is not low-rank. In [8] this problem has been solved by performing the low-rank approximation in

a very particular form. More specifically, a multiplicative deviation g from the equilibrium M is introduced. The density function is then written as $f = Mg$, where the low-rank approximation is only applied to g. The difficulty in that approach is that certain integrals required for the dynamical low-rank algorithm can not be easily separated into x and v dependent components. To avoid this problem, the integrals are written as convolutions and FFT based techniques are used to efficiently evaluate them. This results in a numerical scheme that is rank 1 in the limit of large collisionality and exactly captures the fluid limits described by the Navier–Stokes/Euler equations.

References

[1] J. Bigot, V. Grandgirard, G. Latu, C. Passeron, F. Rozar, and O. Thomine. Scaling GYSELA code beyond 32K-cores on Blue Gene/Q. In *ESAIM: Proceedings*, volume 43, pages 117–135, 2013.

[2] G. Ceruti and C. Lubich. An unconventional robust integrator for dynamical low-rank approximation. *arXiv:2010.02022*, 2020.

[3] Z. Ding, L. Einkemmer, and Q. Li. Error analysis of an asymptotic preserving dynamical low-rank integrator for the multi-scale radiative transfer equation. *arXiv:1907.04247*, 2019.

[4] L. Einkemmer. A mixed precision semi-Lagrangian algorithm and its performance on accelerators. In *High Performance Computing & Simulation (HPCS), 2016 International Conference on*, pages 74–80, 2016.

[5] L. Einkemmer. A low-rank algorithm for weakly compressible flow. *SIAM J. Sci. Comput.*, 41(5):A2795–A2814, 2019.

[6] L. Einkemmer. A performance comparison of semi-Lagrangian discontinuous Galerkin and spline based Vlasov solvers in four dimensions. *J. Comput. Phys.*, 376:937–951, 2019.

[7] L. Einkemmer, J. Hu, and Y. Wang. An asymptotic-preserving dynamical low-rank method for the multi-scale multi-dimensional linear transport equation. *arXiv:2005.06571*, 2020.

[8] L. Einkemmer and I. Joseph. A mass, momentum, and energy conservative dynamical low-rank scheme for the Vlasov equation. *arXiv:2101.12571*, 2021.

[9] L. Einkemmer and C. Lubich. A low-rank projector-splitting integrator for the Vlasov–Poisson equation. *SIAM J. Sci. Comput.*, 40:B1330–B1360, 2018.

[10] L. Einkemmer and C. Lubich. A quasi-conservative dynamical low-rank algorithm for the Vlasov equation. *SIAM J. Sci. Comput.*, 41(5):B1061–B1081, 2019.

[11] L. Einkemmer, A. Ostermann, and C. Piazzola. A low-rank projector-splitting integrator for the Vlasov–Maxwell equations with divergence correction. *J. Comput. Phys.*, 403:109063, 2020.

[12] F. Filbet and E. Sonnendrücker. Comparison of Eulerian Vlasov solvers. *Comput. Phys. Commun.*, 150(3):247–266, 2003.

[13] J. Hu, S. Jin, and Q. Li. Asymptotic-preserving schemes for multiscale hyperbolic and kinetic equations. In *Handbook of Numerical Analysis*, volume 18, pages 103–129. 2017.

[14] K. Kormann and E. Sonnendrücker. Sparse grids for the Vlasov–Poisson equation. *Sparse Grids Appl.*, 109:163–190, 2014.

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

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

[17] M. Palmroth, U. Ganse, Y. Pfau-Kempf, M. Battarbee, L. Turc, T. Brito, M. Grandin, S. Hoilijoki, A. Sandroos, and S. von Alfthan. Vlasov methods in space physics and astrophysics. *Living Rev. Comput. Astrophys*, 4(1):1, 2018.

[18] Z. Peng and R.G. McClarren. A high-order/low-order (HOLO) algorithm for preserving conservation in time-dependent low-rank transport calculations. *arXiv:2011.06072*, 2020.
[19] Z. Peng, R.G. McClarren, and M. Frank. A low-rank method for two-dimensional time-dependent radiation transport calculations. *J. Comput. Phys.*, page 109735, 2020.

Geometric Integration of Degenerate Lagrangian Systems

Michael Kraus

The Geometric Numerical Integration literature describes numerous structure-preserving algorithms for canonical Hamiltonian and regular Lagrangian systems. Noncanonical Hamiltonian and degenerate Lagrangian systems, on the other hand, are rarely discussed. Such systems play an important role in reduced charged particle dynamics like the guiding centre model, population dynamics like the Lotka–Volterra model, or nonlinear oscillators. The following is a short overview of the issues that arise when discretising such systems and a discussion of possible strategies for their structure-preserving integration.

The most general form of the dynamical equations of a Hamiltonian system, also referred to as Poisson system, with state vector $q \in \mathbb{R}^m$ is

$$\dot{q} = \mathcal{P}(q)\,\nabla H(q). \tag{1}$$

Here, $\mathcal{P}(q)$ is an anti-symmetric matrix, possibly degenerate, satisfying

$$\sum_{l=1}^{m} \left[\frac{\partial \mathcal{P}^{ij}(q)}{\partial q^l}\,\mathcal{P}^{lk}(q) + \frac{\partial \mathcal{P}^{jk}(q)}{\partial q^l}\,\mathcal{P}^{li}(q) + \frac{\partial \mathcal{P}^{ki}(q)}{\partial q^l}\,\mathcal{P}^{lj}(q) \right] = 0 \quad \text{for } 1 \le i,j,k \le m.$$

In the following, we are concerned with a special case of such systems, namely noncanonical symplectic systems. In that case $\mathcal{P}(q)$ is even-dimensional ($m = 2d$), non-degenerate and thus invertible, so that we can write

$$\mathcal{P}(q) = \Omega^{-1}(q), \tag{2}$$

with Ω a noncanonical symplectic form whose components are in general nonlinear functions of the state variables q. If Ω is constant and takes the values

$$\Omega_c = \begin{pmatrix} \mathbb{0} & -\mathbb{1} \\ \mathbb{1} & \mathbb{0} \end{pmatrix},$$

the system is said to be of canonical symplectic form.

Noncanonical symplectic systems are tightly linked to a special class of degenerate Lagrangian systems whose Lagrangian is linear in velocities and given by

$$L(q,\dot{q}) = \vartheta(q)\cdot\dot{q} - H(q). \tag{3}$$

The corresponding Euler–Lagrange equations,

$$\frac{\partial L}{\partial q}(q,\dot{q}) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}(q,\dot{q})\right) = 0, \tag{4}$$

are equivalent to (1) in the case of (2). Here, the symplectic potential ϑ is a general, possibly nonlinear function $\mathbb{R}^{2d} \to \mathbb{R}^{2d}$, such that $\Omega = \mathsf{d}\vartheta$, with d denoting the

exterior derivative. Note that by mild abuse of notation we do not distinguish between differential forms ϑ and Ω and their components in some local basis.

This connection suggests that variational integrators [2] should provide suitable means for the construction of structure-preserving integrators for noncanonical symplectic systems. Let us denote the discrete trajectory by $q_d = \{q_n\}_{n=0}^N$. A discrete Lagrangian is constructed as a function of two consecutive points (q_n, q_{n+1}) along that trajectory, by approximating the continuous Lagrangian L via finite difference, Runge–Kutta or other suitable methods. For example, a simple midpoint approximation is given by

$$L_d(q_n, q_{n+1}) = h\, L\left(\frac{q_n + q_{n+1}}{2}, \frac{q_{n+1} - q_n}{h}\right). \tag{5}$$

In a similar fashion as in the continuous case, Hamilton's principle of stationary action leads to the discrete Euler–Lagrange equations,

$$D_2 L_d(q_{n-1}, q_n) + D_1 L_d(q_n, q_{n+1}) = 0 \quad \text{for all} \quad n, \tag{6}$$

where D_i denotes thederivative w.r.t. the i-th argument. In the specific case of degenerate Lagrangians, for which the continuous Euler–Lagrange equations are first-order differential equations, the discrete Euler–Lagrange equations constitute a multi-step method. This implies that these integrators are susceptible to parasitic modes and require two sets of initial data even though the continuous Euler–Lagrange equations are of first order.

A remedy for the initialisation issue can be found by using the discrete fibre derivative to rewrite the discrete Euler–Lagrange equations (6) in the so-called position-momentum (PM) form

$$\begin{aligned} p_n &= -D_1 L_d(q_n, q_{n+1}), \\ p_{n+1} &= D_2 L_d(q_n, q_{n+1}). \end{aligned} \tag{7}$$

Given (q_n, p_n), this system can be solved for (q_{n+1}, p_{n+1}) if the discrete Lagrangian L_d is non-degenerate, i.e.,

$$\det \left| \frac{\partial^2 L_d}{\partial q_n^i \, \partial q_{n+1}^j} \right| \neq 0.$$

Interestingly, most discrete Lagrangians like (5) are non-degenerate even if the corresponding continuous Lagrangian L is degenerate. With the PM-form (7), the continuous fibre derivative can be used to obtain a second initial condition p_0 as function of q_0,

$$p_0 = \frac{\partial L}{\partial \dot{q}}(q_0) = \vartheta(q_0).$$

While this solves the initialisation problem, it does not solve the problem of parasitic modes. Nonetheless, the PM-formulation provides an interesting geometric picture of the parasitic modes developing in solutions of the variational integrator. Introducing the PM-form implicitly amounted to rewriting the equations of motion as an index-two differential-algebraic system,

$$\dot{z} = \Omega_c^{-1}\big(\nabla H(z) + \nabla \phi^T(z)\, v\big), \quad 0 = \phi(z) = p - \vartheta(q), \quad z = (q, p). \tag{8}$$

The problem that arises, is that the variational integrator does not preserve the constraint $\phi(z) = 0$, so that the numerical solution drifts away from the constraint submanifold, i.e., $p_n \neq \vartheta(q_n)$ for $n \geq 1$, even though $p_0 = \vartheta(q_0)$.

This can be remedied by combining the PM-form (7) with projection methods [3]. In particular, symmetric projection methods [1] show very good long-time stability and energy conservation properties, although the resulting projected variational integrators are usually not symplectic. In a similar but more general fashion, applying the SPARK methodology [6] to (8) leads to a large variety of integrators, many of which exhibit good properties although not being symplectic.

For certain degenerate Lagrangians, it is also possible to construct symplectic Runge–Kutta methods as well as one-step variational integrators. This applies to Lagrangians (3) with $q \in \mathbb{R}^{2d}$ and ϑ such that d of its components vanish,

$$\vartheta_\mu = 0 \qquad \text{for} \qquad \mu = d+1, \, ..., 2d.$$

It was already pointed out that, although the continuous Lagrangian is degenerate, the corresponding discrete Lagrangians typically are not. Thus a decisive structural property of the continuous system is lost in discretisation: it's degeneracy. When the degeneracy is retained in the discrete Lagrangian, the discrete Euler–Lagrange equations (6) constitute one-step methods referred to as Degenerate Variational Integrators (DVIs) [4, 5]. For simplicity, consider the Lagrangian

$$L(q^1, q^2, \dot{q}^1, \dot{q}^2) = \vartheta_1(q^1, q^2) \cdot \dot{q}^1 - H(q^1, q^2),$$

which is discretised as

$$L_d(q_n^1, q_n^2, q_{n+1}^1, q_{n+1}^2) = h \left[\vartheta_1(q_n^1, q_n^2) \cdot \frac{q_{n+1}^1 - q_n^1}{h} - H(q_n^1, q_n^2) \right].$$

The discrete Euler–Lagrange equations (6) can be written as

$$\begin{aligned}
\vartheta_1(q_n^1, q_n^2) &= \vartheta_1(q_{n-1}^1, q_{n-1}^2) + h \nabla_1 \vartheta_1(q_n^1, q_n^2) \cdot v_n^1 \\
&\qquad - h \nabla_1 H(q_n^1, q_n^2), && \text{for } n = 1, ..., N-1, \\
v_n^1 &= \big(\nabla_2 \vartheta_1(q_n^1, q_n^2) \big)^{-1} \nabla_2 H(q_n^1, q_n^2), && \text{for } n = 0, ..., N-1, \\
q_{n+1}^1 &= q_n^1 + h v_n^1, && \text{for } n = 0, ..., N-1.
\end{aligned}$$

Note that this system is underdetermined, as it lacks equations that determine x_N^2. Motivated by the discrete symplecticity condition following from the boundary values in the action principle [2], the system is closed by adding the equations

$$\begin{aligned}
\vartheta_1(q_N^1, q_N^2) &= \vartheta_1(q_{N-1}^1, q_{N-1}^2) + h \nabla_1 \vartheta_1(q_N^1, q_N^2) \cdot v_N^1 - h \nabla_1 H(q_N^1, q_N^2), \\
v_N^1 &= \big(\nabla_2 \vartheta_1(q_N^1, q_N^2) \big)^{-1} \nabla_2 H(q_N^1, q_N^2).
\end{aligned}$$

With this closure, the full set of discrete Euler–Lagrange equations preserves the discrete symplectic form $\Omega_d = \mathsf{d}\vartheta_d$ with potential

$$\vartheta_d(q^1, q^2) = \big[\vartheta_1(q^1, q^2) - h \nabla_1 \vartheta_1(q^1, q^2) \cdot v^1 + h \nabla_1 H(q^1, q^2) \big] \, \mathsf{d}q^1.$$

The DVI thus obtained is of first-order [4]. While second-order generalisations also exist [5], higher-order DVIs are currently not known. Moreover this approach is

limited to the aforementioned special form of the degenerate Lagrangian and thus not always applicable. This and the previous discussion raise several questions:

(1) Can DVIs be generalised to arbitrary order, using e.g. Galerkin or Runge–Kutta discretisations of the Lagrangian?
(2) Can DVIs [4, 5] and symplectic Runge–Kutta methods [7] be obtained for general degenerate Lagrangians of the form (3), without assuming that any of the components of the symplectic potential ϑ vanish?
(3) Do symplectic Runge–Kutta methods [7] follow from a discrete variational principle?

Addressing these questions is of great practical relevance as geometric integrators for degenerate Lagrangian systems and in particular for the guiding centre model are much-needed in application fields such as astro and fusion plasma physics.

References

[1] E. Hairer, *Symmetric projection methods for differential equations on manifolds*, BIT Numerical Mathematics **40** (2000), 726–734.
[2] J. E. Marsden, M. West, *Discrete mechanics and variational integrators*, Acta numerica **10** (2001), 357–514.
[3] M. Kraus, *Projected Variational Integrators for Degenerate Lagrangian Systems*, Preprint (2017), arXiv:1708.07356.
[4] C. L. Ellison, J. M. Finn, J. W. Burby, M. Kraus, H. Qin, W. M. Tang, *Degenerate variational integrators for magnetic field line flow and guiding center trajectories*, Physics of Plasmas **25** (2018), 052502.
[5] J. W. Burby, J. M. Finn, C. L. Ellison, *Improved accuracy in degenerate variational integrators for guiding center and magnetic field line flow*, Preprint (2021), arXiv:2103.05566.
[6] M. Kraus, *SPARK Methods for Degenerate Lagrangian Systems*, in preparation.
[7] M. Kraus, *Symplectic Runge–Kutta Methods for Degenerate Lagrangian Systems*, in preparation.

Majorants series for the N-body problem

Philippe Chartier
(joint work with M. Antoñana and A. Murua)

This talk is concerned with the solution of the N-body problem: considering N masses m_i, $i = 1, \dots, N$, moving in a three-dimensional space under the influence of gravitational forces, Newton's law describe the evolution of their *positions* q_i and *velocities* v_i for $i = 1, \dots, N$ through the equations

$$m_i \frac{d^2 q_i}{dt^2} = \sum_{j \neq i} \frac{G m_i m_j}{\|q_j - q_i\|^3} (q_j - q_i), \tag{1}$$

where G is the gravitational constant and $\|q_j - q_i\|$ is the distance between q_i and q_j in the euclidean norm of $\mathbb{R}^3$.

The N-body equations are of great importance in physics and celestial mechanics in particular and the question of representing its solution in the form of a convergent series has been a long-standing problem. For an introduction on the historical and practical aspects of the subject, we refer to [1]. To make a long

story short, let us just recall that the problem was solved for $N = 3$ by Karl Fritiof Sundman [2] and Qiu-Dong Wang for the $N \geq 3$-case [6] in the 1990s.

As an intermediate step to obtain a representation of each solution of (1) in terms of a series expansion convergent for all t in its maximal interval of existence, both authors rewrite the N-body equations (1) in terms of a new independent variable τ related to the physical time t by

$$\frac{d\tau}{dt} = s(q(t))^{-1}, \quad \tau(0) = 0,$$

for an appropriate *time-renormalization function* $s(q)$ depending on the positions $q = (q_1, \ldots, q_N)$. In [1], we proposed new time-renormalization functions for the purpose of simulating the N-body problem with constant stepsizes without degrading the accuracy of the computed trajectories. In contrast with previously known functions [2, 6], ours depend not only on positions but also on velocities (up to our knowledge, for the first time). Noticeably, these global time-renormalizations were shown to be *uniform* in the sense that the solution of the time-renormalized equations in the fictitious time τ can be extended *analytically* to the strip

$$\{\tau \in \mathbb{C} \ : \ |\mathrm{Im}(\tau)| \leq \beta\}$$

for some $\beta > 0$ independent of the initial conditions $(q^0, v^0) \in \mathbb{R}^{6N}$ (provided that $q_i^0 \neq q_j^0$ for all $i \neq j$) and the masses m_i, $1 \leq i \leq N$.

Our main contribution is to explicitly construct *majorants* for the expansions of the solution of the N-body problem as a series in powers of either t or τ. We furthermore construct majorants for the expansion of discrete solutions of the time-renormalized equations in series of powers of the step-size τ .

Generally speaking, a power series $f = \sum_{k\geq 0} f_k\, t^k \in \mathbb{R}^n[[t]]$ is said to be *majored* by $\bar{f} = \sum_{k\geq 0} \bar{f}_k\, t^k \in \mathbb{R}_+[[t]]$ if, for all $k \in \mathbb{N}$,

$$\|f_k\| \leq \bar{f}_k$$

where $\|\cdot\|$ is the euclidean norm in $\mathbb{R}^n$ and we then write

$$f \trianglelefteq \bar{f}.$$

The application of the technique of majorant equations goes back to the proof of Cauchy-Kovalevskaya theorem [5] (see also [4] for a specific application to ordinary differential equations) and has several advantages in our context:

- it allows for an easy estimate of the radius of convergence of the series f;
- simple rules on majorant series apply to the usual operations on power series, such as addition, multiplication, derivation and integration;
- when used for the time-renormalized N-body equations, it leads to improved estimates of the value of β;
- it can be used to analyze numerical discretizations of the time-renormalized equations, and in particular, to obtain bounds for their local errors.

References

[1] M. Antoñana, P. Chartier, J. Makazaga, A. Murua , *Global time-regularisation of the gravitational N-body problem*, SIAM J. Appl. Dyn. Syst., 19(4): 2658-2681, 2020.
[2] K. F. Sundman, *Mémoire sur le problème des trois corps*, Acta Mathematica. 36: 105–179, 1912.
[3] M. Antoñana, P. Chartier, A. Murua, *Majorants series for the N-body problem*, https://arxiv.org/pdf/2103.12839.pdf, 2021, submitted.
[4] *Majorants for formal power series*, Joris van der Hoeven, unpublished notes.
[5] S. von Kowalevsky, *Zur theorie der partiellen differentialgleichungen*, vol. 1875, no. 80, 1875, pp. 1-32.
[6] Q.-D. Wang, *The global solution of the n-body problem*, Celestial Mechanics and Dynamical Astronomy, 50 (1): 73–88, 1991.

Fractalization and quantization in linear and nonlinear dispersive partial differential equations and Fermi-Pasta-Ulam-Tsingou lattices

Peter J. Olver

The evolution, through spatially periodic linear dispersion, of rough initial data produces fractal, non-differentiable profiles at irrational times and, for asymptotically polynomial large wave dispersion relations, quantized structures at rational times. Such phenomena have been observed in models that arise in fluid mechanics, optics, and quantum mechanics, including integrable equations such as the Korteweg–deVries, nonlinear Schrödinger equation, and Benjamin–Ono equations, and their non-integrable counterparts with higher degree nonlinearities. Ramifications and recent progress on the analysis and numerics of these remarkable phenomena were presented.

The second part of the talk concentrated on new results on the dispersive fractalization and quantization of solutions to periodic linear Fermi–Pasta–Ulam–Tsingou systems, which produce integrable partial differential equation models in the continuum limit. For nonlinear periodic FPUT systems, our numerical results suggest a somewhat similar behavior in the presence of small nonlinearities, which disappears as the nonlinear force increases in magnitude. However, even with the high-order splitting methods we employed, these numerical investigations are so far limited to nonlinear FPUT chains with a smaller number of masses than would be needed to resolve the issue unambiguously.

References

[1] Berry, M.V., Klein, S., *Integer, fractional and fractal Talbot effects*, J. Mod. Optics **43** (1996), 2139–2164.
[2] Berry, M.V., Marzoli, I., Schleich, W., *Quantum carpets, carpets of light*, Physics World 14 (6) (2001), 39–44.
[3] Bourgain, J., *Periodic Korteweg de Vries equation with measures as initial data*, Selecta Math., **3** (1997), 115–159.
[4] Boulton, L., Olver, P.J., Pelloni, B., Smith, D.A., *New revival phenomena for linear integro-differential equations*, Stud. Appl. Math., to appear.
[5] Chen, G., Olver, P.J., *Dispersion of discontinuous periodic waves*, Proc. Roy. Soc. London A **469** (2013), 20120407.

[6] Chen, G., Olver, P.J., *Numerical simulation of nonlinear dispersive quantization*, Discrete Cont. Dyn. Syst. A **34** (2013), 991–1008.
[7] Chousionis, V., Erdoğan, M.B., Tzirakis, N., *Fractal solutions of linear and nonlinear dispersive partial differential equations*, Proc. London Math. Soc. **110** (2015), 543–564.
[8] Dauxois, T., *Fermi, Pasta, Ulam, and a mysterious lady*, Physics Today **61** (1) (2008), 55–57.
[9] Erdoğan, M.B., Shakan, G., *Fractal solutions of dispersive partial differential equations on the torus*, Selecta Math., **25** (2019), 11.
[10] Fermi, E., Pasta, J., Ulam, S., *Studies of nonlinear problems*. I., Los Alamos Report LA 1940, 1955, in: *Nonlinear Wave Motion*, A.C. Newell, ed., Lectures in Applied Math., vol. 15, American Math. Soc., Providence, R.I., 1974, pp. 143–156.
[11] Holden, H., Karlsen, K.H., Risebro, N.H., Tao, T., *Operator splitting for the KdV equation*, Math. Comp. **80** (2011), 821–846.
[12] Kapitanski, L., Rodnianski, I., *Does a quantum particle know the time?*, in *Emerging Applications of Number Theory*, D. Hejhal, J. Friedman, M.C. Gutzwiller and A.M. Odlyzko, eds., IMA Volumes in Mathematics and its Applications, vol. 109, Springer Verlag, New York, 1999, pp. 355–371
[13] Olver, P.J., *Dispersive quantization*, Amer. Math. Monthly **117** (2010), 599–610.
[14] Olver, P.J., Sheils, N.E., *Dispersive Lamb systems*, J. Geom. Mech. **11** (2019), 239–254.
[15] Olver, P.J., Sheils, N.E., Smith, D.A., *Revivals and fractalisation in the linear free space Schrödinger equation*, Quart. Appl. Math. **78** (2020), 161–192,
[16] Olver, P.J., Stern, A., *Dispersive fractalization in Fermi-Pasta-Ulam-Tsingou lattices*, European J. Appl. Math., to appear.
[17] Olver, P.J., Tsatis, E., *Points of constancy of the periodic linearized Korteweg–deVries equation*, Proc. Roy. Soc. London A **474** (2018), 20180160.
[18] Oskolkov, K.I., *A class of I.M. Vinogradov's series and its applications in harmonic analysis*, in: *Progress in Approximation Theory*, Springer Ser. Comput. Math., 19, Springer, New York, 1992, pp. 353–402.
[19] Talbot, H.F., *Facts related to optical science*. No. IV, Philos. Mag. **9** (1836), 401–407.
[20] Thaller, B., *Visual Quantum Mechanics*, Springer–Verlag, New York, 2000.
Zabusky, N.J., *Computational synergetics and mathematical innovation*, J. Comp. Phys. **43** (1981) 195–249.
[21] Zabusky, N.J., Kruskal, M.D., *Interaction of "solitons" in a collisionless plasma and the recurrence of initial states*, Phys. Rev. Lett. **15** (1965) 240–243.

Variational Structures in Cochain Projection Based Variational Discretizations of Lagrangian PDEs

Melvin Leok
(joint work with Brian Tran)

Compatible discretizations, such as finite element exterior calculus, provide a discretization framework that respect the cohomological structure of the de Rham complex, which can be used to systematically construct stable mixed finite element methods. Multisymplectic variational integrators are a class of geometric numerical integrators for Lagrangian and Hamiltonian field theories, and they yield methods that preserve the multisymplectic structure and momentum-conservation properties of the continuous system. We investigate the synthesis of these two approaches, by constructing discretization of the variational principle for Lagrangian

field theories utilizing structure-preserving finite element projections. In our investigation, compatible discretization by cochain projections plays a pivotal role in the preservation of the variational structure at the discrete level, allowing the discrete variational structure to essentially be the restriction of the continuum variational structure to a finite-dimensional subspace. The preservation of the variational structure at the discrete level will allow us to construct a discrete Cartan form, which encodes the variational structure of the discrete theory, and subsequently, we utilize the discrete Cartan form to naturally state discrete analogues of Noether's theorem and multisymplecticity, which generalize those introduced in the discrete Lagrangian variational framework by Marsden et al. [6]. We study both covariant spacetime discretization and canonical spatial semi-discretization, and subsequently relate the two in the case of spacetime tensor product finite element spaces.

The problem of structure-preservation in numerical discretizations of partial differential equations has primarily been studied in two disjoint stages, the first involving the semidiscretization of the spatial degrees of freedom, and the second having to do with the time-integration of the resulting coupled system of ordinary differential equations. Implicit in such an approach is the use of tensor product meshes in spacetime. In the context of spatial semidiscretization, the notion of structure-preservation is focused on compatible discretizations, that preserve in some manner the functional and geometric relationships between the different function spaces that arise in the partial differential equation, and in the context of time-integration, geometric numerical integrators aim to preserve geometric invariants like the symplectic or Poisson structure, energy, momentum, and the nonlinear manifold structure of the configuration spaces, like its Lie group, homogeneous space, or Riemannian structure.

Lagrangian partial differential equations are an important class of partial differential equations that exhibit geometric structure that can benefit from numerical discretizations that preserve such structure. This can either be viewed as an infinite-dimensional Lagrangian system with time as the independent variable, or a finite-dimensional Lagrangian multisymplectic field theory with space and time as independent variables. Lagrangian variational integrators are a popular method for systematically constructing symplectic integrators of arbitrarily high-order, and satisfy a discrete Noether's theorem that relates group-invariance with momentum conservation. A group-invariant (and hence momentum-preserving) variational integrator can be constructed from group-equivariant interpolation spaces.

We study the variational finite element discretization of Lagrangian field theories from two perspectives; we begin by investigating directly discretizing the full variational principle over the full spacetime domain, which we refer to as the "covariant" approach, and subsequently study semi-discretization of the instantaneous variational principle on a globally hyperbolic spacetime, which we refer to as the "canonical" approach. This paper can be considered a discrete analogue to the program initiated in Gotay et al. [2, 3], which lays the foundation for relating the covariant and canonical formulations of Lagrangian field theories

through their (multi)symplectic structures and momentum maps. One of the goals of understanding the relation between these two different formulations is to systematically relate the covariant gauge symmetries of a gauge field theory to its initial value constraints. This is seen, for example, in general relativity, where the diffeomorphism gauge invariance gives rise to the Einstein constraint equations over the initial data hypersurface (see, for example, Gourgoulhon [4]). When one semi-discretizes such gauge field theories, the discrete initial data must satisfy an associated discrete constraint. We aim to make sense of the discrete geometric structures in the covariant and canonical discretization approaches as a foundation for understanding the discretization of gauge field theories.

We begin by formulating a discrete variational principle in the covariant approach, utilizing the finite element construction to appropriately restrict the variational principle. We show that a cochain projection from the underlying de Rham complex onto the finite element spaces yields a natural discrete variational principle that is compatible with the holonomic jet structure of a Lagrangian field theory. We then show that discretizing by cochain projections leads to a naturality relation between the continuous variational problem and the discrete variational problem; this naturality then implies that discretization and the variational principle commute and also, that discretizing at the level of the configuration bundle or at the level of the jet bundle are equivalent. Subsequently, by decomposing the finite element spaces into boundary and interior components, we define a discrete Cartan form in analogy with the continuum Cartan form which will, in a sense, encode the discrete variational structure. With particular choices of finite element spaces, this discrete Cartan form recovers the notion of the discrete Cartan form introduced by Marsden et al. [6]; however, we note that our notion of a discrete Cartan form is more general and furthermore, since our discrete variational problem is naturally related to the continuum variational problem, we are able to explicitly discuss in what sense the discrete Cartan form converges to the continuum Cartan form. Using this discrete Cartan form, we state and prove discrete analogues of the multisymplectic form formula and Noether's theorem. Finally, we reinterpret and concisely summarize the preceding sections by interpreting the discrete variational structures as elements of a discrete variational complex.

We study the semi-discretization of the canonical formulation of a Lagrangian field theory on a globally hyperbolic spacetime. We discretize the instantaneous variational principle utilizing cochain projections onto finite element spaces over a Cauchy surface, which gives rise to a semi-discrete Euler–Lagrange equation. We relate this semi-discrete Euler–Lagrange equation to a Hamiltonian flow on a symplectic semi-discrete phase space. We will discuss in what sense the symplectic structure on the semi-discrete phase space arises from a symplectic structure on the continuum phase space. Subsequently, we will investigate the energy-momentum map structure associated to the semi-discrete phase space and discuss how, under appropriate equivariance conditions on the projection, the energy-momentum map structure on the semi-discrete phase space arises as the pullback of the energy-momentum map structure on the continuum phase space. This lays a foundation

for understanding initial value constraints when discretizing field theories with gauge symmetries. Finally, we relate the covariant and canonical discretization approaches in the case of tensor product finite element spaces.

The underlying theme of this work is that, when one discretizes the variational principle utilizing compatible discretization techniques, the associated (covariant or canonical) discretization inherits discrete variational structures which can be viewed as pullbacks or projections of the associated continuum variational structures. These discrete variational structures allow one to investigate structure-preservation under discretization of important physical properties, such as momentum conservation, symplecticity, and (gauge) symmetries.

References

[1] E. S. Gawlik and M. Leok. Interpolation on symmetric spaces via the generalized polar decomposition. *Found. Comput. Math.*, 18(3):757–788, 2018.
[2] M. J. Gotay, J. Isenberg, J. E. Marsden, and R. Montgomery. Momentum maps and classical relativistic fields. Part I: Covariant field theory. (preprint, `arXiv:physics/9801019 [math-ph]`), 1998.
[3] M. J. Gotay, J. Isenberg, J. E. Marsden, and R. Montgomery. Momentum maps and classical relativistic fields. Part II: Canonical analysis of field theories. (preprint, `arXiv:physics/0411032[math-ph]`), 2004.
[4] E. Gourgoulhon. *3+1 Formalism in General Relativity*, volume 846. Springer-Verlag, Berlin, 2012.
[5] M. Leok. Variational discretizations of gauge field theories using group-equivariant interpolation. *Found. Comput. Math.*, 19(5):965–989, 2019.
[6] J. E. Marsden, G. W. Patrick, and S. Shkoller. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Commun. Math. Phys.*, 199(2):351–395, 1998.
[7] B. Tran and M. Leok. Multisymplectic Hamiltonian variational integrators. (preprint, `arXiv:2101.07536 [math.NA]`), 2021.
[8] B. Tran and M. Leok. Variational Structures in Cochain Projection Based Variational Discretizations of Lagrangian PDEs. (preprint, `arXiv:2103.09338 [math.NA]`), 2021.

Geometric variational finite element discretization in fluid dynamics

François Gay-Balmaz
(joint work with Evan S. Gawlik)

1. Introduction

We present a class of finite element variational integrators for fluid models. These integrators are derived by discretizing, in a structure preserving way, the Lie group formulation of fluid dynamics on diffeomorphism groups and the associated variational principles. A main step in the approach is the identification of a subspace of the discrete Lie algebra with a Raviart-Thomas finite element space. The resulting scheme has remarkable conservative properties.

2. Variational formulation for fluids

The regular solutions of Euler type fluid flows in a bounded domain $\Omega \subset \mathbb{R}^n$ with smooth boundary can be formally regarded as curves $\varphi : [0,T] \to \mathrm{Diff}(\Omega)$ that are critical for the Hamilton principle

$$\delta \int_0^T L(\varphi, \partial_t \varphi) \mathrm{d}t = 0 \tag{1}$$

with respect to variations $\delta\varphi$ vanishing at the endpoints. Here $\mathrm{Diff}(\Omega)$ is the group of diffeomorphisms of Ω and $\varphi(t) : \Omega \to \Omega$ is the map sending the position X of a fluid particle at time 0 to its position $x = \varphi(t, X)$ at time t. The function $L : T\,\mathrm{Diff}(\Omega) \to \mathbb{R}$ in (1) is the Lagrangian of the fluid. For incompressible fluids, one uses the subgroup $\mathrm{Diff}_{\mathrm{vol}}(\Omega)$ of volume preserving diffeomorphisms of Ω.

The Hamilton principle (1) induces a variational formulation in the Eulerian frame, called the Euler-Poincaré principle, which, for barotropic fluids, reads

$$\delta \int_0^T \ell(u, \rho)\, \mathrm{d}t = 0, \tag{2}$$

with respect to variations δu and $\delta\rho$ of the form

$$\delta u = \partial_t v + \pounds_u v, \quad \delta\rho = -\operatorname{div}(\rho v), \tag{3}$$

with $v : [0,T] \to \mathfrak{X}(\Omega)$ and $v(0) = v(T) = 0$. Here $\mathfrak{X}(\Omega)$ denotes the Lie algebra of $\mathrm{Diff}(\Omega)$, which consists of vector fields on Ω, tangent to $\partial\Omega$.

3. Discrete variational formulation for fluids

Discrete diffeomorphisms. The method is based on the use of a finite dimensional Lie group approximation of $\mathrm{Diff}(\Omega)$, given by

$$G_h^r = \{q \in GL(V_h^r) \mid q\mathbf{1} = \mathbf{1}\}, \tag{4}$$

where $V_h^r = \{f \in L^2(\Omega) \mid f|_K \in P_r(K),\ \forall K \in \mathcal{T}_h\}$ is the discontinuous Galerkin space of order r associated to a triangulation $\mathcal{T}_h$ of Ω, and where $\mathbf{1}$ is a discrete representative of the constant function 1. The condition $q\mathbf{1} = \mathbf{1}$ encodes the fact that constant functions are preserved by the action of a diffeomorphism. Elements in the Lie algebra

$$\mathfrak{g}_h^r = \{A \in L(V_h^r, V_h^r) \mid A\mathbf{1} = 0\} \tag{5}$$

are potential candidates to be discrete vector fields. As linear maps in $\mathfrak{g}_h^r$ these discrete vector fields act as discrete derivations on V_h^r. It is thus natural to choose them as distributional directional derivatives.

Space of distributional derivatives. It is shown in [1] that the subspace $S_h^r \subset \mathfrak{g}_h^r$ of distributional derivatives is isomorphic to a Raviart-Thomas space:

$$S_h^r \longleftrightarrow RT_{2r}(\mathcal{T}_h) = \{u \in H_0(\mathrm{div}, \Omega) \mid u|_K \in (P_{2r}(K))^n + xP_{2r}(K), \forall K \in \mathcal{T}_h\}\,. \tag{6}$$

This result relates the discrete Lie group formulation for fluids to finite element methods.

Discrete Lagrangians. In order to define the semidiscrete Lagrangian on the Lie algebra $\mathfrak{g}_h^r$, we introduce the *Lie algebra-to-vector field map* $\widehat{\ }: \mathfrak{g}_h^r \subset L(V_h^r, V_h^r) \to [V_h^r]^n$ given by

$$\widehat{A} := \sum_{k=1}^{n} A(I_h^r(x^k))e_k, \tag{7}$$

where $I_h^r : L^2(\Omega) \to V_h^r$ is the L^2-orthogonal projector. Given a continuous Lagrangian, for instance an expression $\ell = \ell(u, \rho)$ for barotropic fluids, the discrete Lagrangian $\ell_d : \mathfrak{g}_h^r \times V_h^r \to \mathbb{R}$ is defined by $\ell_d(A_h, \rho_h) = \ell(\widehat{A_h}, \rho_h)$. Given a discrete Lagrangian, one needs to restrict its domain of definition to a subspace $\Delta_h^R \subset S_h^r$ on which the discrete Lagrangian is nondegenerate. The diagram below illustrates the situation that we consider and the relation between the various spaces, [1].

$$\begin{array}{ccccccc} RT_{2r}(\mathcal{T}_h) & \xleftrightarrow[(6)]{} & S_h^r & \hookrightarrow & \mathfrak{g}_h^r & \xrightarrow{\widehat{\ }} & [V_h^r]^n \\ \uparrow & & \uparrow & & & & \\ R_h & \longleftrightarrow & \Delta_h^R & & & & \end{array}$$

Semidiscrete Euler-Poincaré principle. In this setting, given $\ell_d : \mathfrak{g}_h^r \times V_h^r \to \mathbb{R}$ and $\Delta_R^h \subset S_h^r$ on which ℓ_d is nondegenerate, the semidiscrete version of (2)–(3) reads

$$\delta \int_0^T \ell_d(A_h, \rho_h)\mathrm{d}t = 0,$$

with respect to variations of the form

$$\delta A_h = \partial_t B_h + [B_h, A_h] \quad \text{and} \quad \delta\rho_h = -\rho_h \cdot B_h,$$

for all $B_h(t) \in \Delta_h^R$ with $B_h(0) = B_h(T) = 0$. The critical conditions for this principle yields the finite element semidiscrete fluid equations, [1, 2].

Temporal discretization. The variational character of the fluid equations can be exploited also at the temporal level, by deriving the temporal scheme via a time discretization of the Euler-Poincaré variational principle. Alternatively, it also admits a time discretization that exactly preserves the total energy, [1, 2, 3].

4. Conservative properties and illustration

When applied to the incompressible fluid with variable density, our finite element scheme preserves the following quantities at the full discrete level: *total mass*, *total squared density*, *total energy*, and the *incompressibility constraint*, [2].

When applied to incompressible MHD with variable density, our scheme can be adapted to additionally preserve the *magnetic helicity*, the *cross-helicity*, and the *solenoidal constraint*, [3].

The figure below illustrates the Rayleigh-Taylor instability test for a compressible fluid with Lagrangian

$$\ell(u, \rho, s) = \int_\Omega \Big[\frac{1}{2}\rho|u|^2 - \rho e(\rho, \eta) - \rho\Phi\Big]\mathrm{d}x, \tag{8}$$

with s the entropy density, $\eta = \frac{s}{\rho}$, $e(\rho, \eta) = e^{\eta/C_v}\rho^{\gamma-1}$, $\gamma = 5/3$, $C_v = 1$, and $\Phi = -y$. We take $\Delta t = 0.01$, $R_h = RT_0(\mathcal{T}_h)$ and V_h^1 on a uniform triangulation $\mathcal{T}_h$ of $\Omega = (0, 1/4) \times (0, 1)$ with $h = 2^{-8}$, [1]. We incorporate upwinding by using the strategy detailed in [1], which retains the scheme's energy-preserving property.

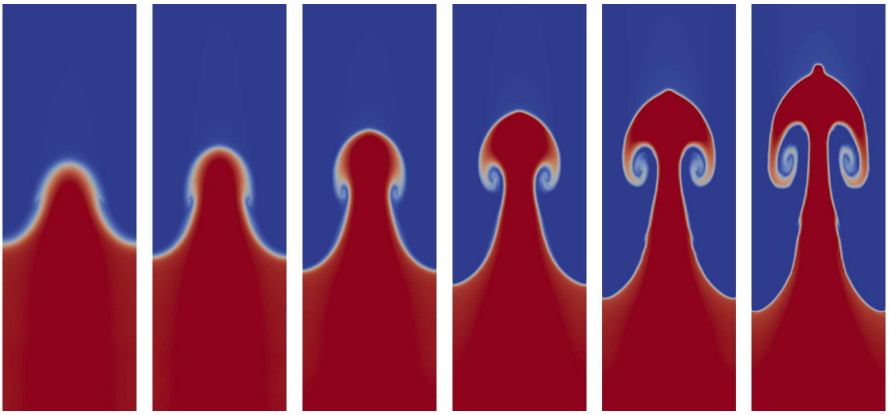

FIGURE 1. Contours of the mass density at $t = 1.0, 1.2, 1.4, 1.6, 1.8, 2.0$.

REFERENCES

[1] E. S. Gawlik and F. Gay-Balmaz, *A variational finite element discretization of compressible flow*, Found. Comp. Math. 2020, accepted.

[2] E. S. Gawlik and F. Gay-Balmaz, *A conservative finite element method for the incompressible Euler equations with variable density*, J. Comp. Phys., 412(2020) 109439.

[3] E. S. Gawlik and F. Gay-Balmaz, *A finite element method for MHD that preserves energy, cross-helicity, magnetic helicity, incompressibility, and* div $B = 0$., *arXiv: 2012.04122.*

Energy-stable parametric finite element methods (ES-PFEM) for geometric PDEs and applications

WEIZHU BAO

In this talk, I begin with a review of different geometric flows (PDEs) including mean curvature (curve shortening) flow, surface diffusion flow, Willmore flow, etc., which arise from materials science, interface dynamics in multi-phase flows, biology membrane, computer graphics, geometry, etc. Different mathematical formulations and numerical methods for mean curvature flow are then discussed [1, 5, 6], including marker-particle method, θ-L method, finite element method (FEM), and parametric finite element method (PFEM), etc. In particular, a semi-implicit energy-stable parametric finite element method (ES-PFEM) is presented in details [4, 8]. The ES-PFEM enjoys a few advantages for discretizing geometric flows: (i) it is semi-implicit and only a linear system is to be solved at each

time step, (ii) it preserves energy dissipation and thus it is unconditionally energy stable, and (iii) it guarantees asymptotic mesh equal-distribution and thus no re-meshing is needed during time evolution. Then the ES-PFEM is extended to surface diffusion flow and anistropic surface diffusion flow [7, 3]. Finally, sharp interface models and their PFEM and ES-PFEM approximations are presented for solid-state dewetting in materials science [2, 8, 9]. This talk is based on joint works with Wei Jiang, Yifei Li, Yan Wang and Quan Zhao.

References

[1] W. Bao, W. Jiang, Y. Wang and Q. Zhao, *A parametric finite element method for solid-state dewetting problems with anisotropic surface energies*, J. Comput. Phys., **330** (2017), 380–400.

[2] W. Bao and Q. Zhao, *An energy-stable parametric finite element method for simulating solid-state dewetting problems in three dimensions*, arXiv: 2012.11404.

[3] W. Bao and Q. Zhao, *A structure-preserving parametric finite element method for surface diffusion*, in preparation.

[4] J. W. Barrett, H. Garcke and R. Nürnberg, *A parametric finite element method for fourth order geometric evolution equations*, J. Comput. Phys., **222** (2007), 441–467.

[5] J. W. Barrett, H. Garcke and R. Nürnberg, *On the variational approximation of combined second and fourth order geometric evolution equations*, SIAM J. Sci. Comput., **29** (2007), 1006–1041.

[6] J. W. Barrett, H. Garcke and R. Nürnberg, *Parametric finite element approximations of curvature driven interface evolutions*, Handbook of Numerical Analysis (Andrea Bonito and Ricardo H. Nochetto, eds.), **21** (2020), 275-423, Elsevier, Amsterdam.

[7] Y. Li and W. Bao, *An energy-stable parametric finite element method for anisotropic surface diffusion*, arXiv: 2012.05610.

[8] Q. Zhao, W. Jiang and W. Bao, *An energy-stable parametric finite element method for simulating solid-state dewetting*, IMA J. Numer. Anal., https://doi-org.libproxy1.nus.edu.sg/10.1093/imanum/draa070.

[9] Q. Zhao, W. Jiang and W. Bao, *A parametric finite element method for solid-state dewetting problems in three dimensions*, SIAM J. Sci. Comput., **42** (2020), B327–B352.

Structure preserving deep learning

Carola-Bibiane Schönlieb

(joint work with Elena Celledoni, Matthias J. Ehrhardt, Christian Etmann, Robert I McLachlan, Brynjulf Owren, Ferdia Sherry)

Over the past few years, deep learning has risen to the foreground as a topic of massive interest, mainly as a result of successes obtained in solving large-scale image processing tasks. There are multiple challenging mathematical problems involved in applying deep learning: most deep learning methods require the solution of hard optimisation problems, and a good understanding of the tradeoff between computational effort, amount of data and model complexity is required to successfully design a deep learning approach for a given problem. A large amount of progress made in deep learning has been based on heuristic explorations, but there is a growing effort to mathematically understand the structure in existing

deep learning methods and to systematically design new deep learning methods to preserve certain types of structure in deep learning.

In a recent article [1], we review a number of these directions: some deep neural networks can be understood as discretisations of dynamical systems, neural networks can be designed to have desirable properties such as invertibility or group equivariance, and new algorithmic frameworks based on conformal Hamiltonian systems and Riemannian manifolds to solve the optimisation problems have been proposed.

My talk picked out three topics from the above, namely: neural networks as discretisations of dynamical systems and possible notions of stability of deep learning that can be drawn from the numerical analysis of ODEs; invertible neural networks and their use for computing probability density functions from a finite number of samples; and equivariant neural networks and their favourable performance for image processing tasks. We discuss these topics and raise open problems that we consider to be interesting directions for future research.

References

[1] E. Celledoni, M. J. Ehrhardt, C. Etmann, R. I. McLachlan, B. Owren, C.-B. Schönlieb, F. Sherry, *Structure preserving deep learning*, arXiv:2006.03364, to appear in EJAM 2021.

Mixing Time Guarantees for Unadjusted Hamiltonian Monte Carlo

Nawaf Bou-Rabee
(joint work with Andreas Eberle)

In this talk, we present quantitative upper bounds on the total variation mixing time of the Markov chain corresponding to the unadjusted Hamiltonian Monte Carlo (uHMC) algorithm. For two general classes of models and fixed time discretization step size h, the mixing time is shown to depend only logarithmically on the dimension. Moreover, we provide quantitative upper bounds on the total variation distance between the invariant measure of the uHMC chain and the true target measure. As a consequence, we show that an ε-accurate approximation of the target distribution μ in total variation distance can be achieved by uHMC for a broad class of models with $O\left(d^{3/4}\epsilon^{-1/2}\log(d/\varepsilon)\right)$ gradient evaluations, and for mean field models with weak interactions with $O\left(d^{1/2}\epsilon^{-1/2}\log(d/\varepsilon)\right)$ gradient evaluations. The proofs are based on the construction of successful couplings for uHMC that realize the upper bounds.

References

[1] N. Bou-Rabee and A. Eberle, *Mixing Time Guarantees for Unadjusted Hamiltonian Monte Carlo*, to appear on arXiv:2105.00887 [math.PR]

[2] A. Durmus and A. Eberle, *Asymptotic bias of inexact Markov Chain Monte Carlo methods in high dimensions*, to appear on arXiv: [math.PR]

[3] N. Bou-Rabee and K. Schuh, *Convergence of unadjusted Hamiltonian Monte Carlo for mean-field models*, arXiv:2009.08735 [math.PR]

[4] N. Bou-Rabee and A. Eberle, *Two-Scale Coupling for Hamiltonian Monte Carlo in Infinite Dimensions*, Stochastics and Partial Differential Equations: Analysis and Computations **9** (2021), 207–242.
[5] N. Bou-Rabee, A. Eberle, and R. Zimmer *Coupling and Convergence for Hamiltonian Monte Carlo*, Annals of Applied Probability **30** (2020), 1209–1250.
[6] N. Bou-Rabee and J. M. Sanz-Serna, *Geometric Integrators and the Hamiltonian Monte Carlo method*, Acta Numerica **27** (2018), 113–206.
[7] N. Bou-Rabee and J. M. Sanz-Serna, *Randomized Hamiltonian Monte Carlo*, Annals of Applied Probability **27** (2017), 2159–2194.

On superconvergence features of methods based on the Crank-Nicolson scheme in the context of diffusion PDEs (deterministic and stochastic)

Gilles Vilmart
(joint work with Assyr Abdulle, Ibrahim Almuslimani, Guillaume Bertoli, Christophe Besse, and Charles-Edouard Bréhier)

In this talk,[1] we present two situations where the Crank-Nicolson method is surprisingly more accurate than one could expect:

- **in the context of splitting methods for parabolic PDEs** [3, 4]**:** we show that the Strang splitting method applied to a diffusion-reaction equation with inhomogeneous general oblique boundary conditions is of order two when the diffusion equation is solved with the Crank-Nicolson method, while order reduction occurs in general if using other Runge-Kutta schemes or even the exact flow itself for the diffusion part. We prove [4] these results when the source term only depends on the space variable, an assumption which makes the splitting scheme equivalent to the Crank-Nicolson method itself applied to the whole problem.
- **in the context of ergodic parabolic stochastic PDEs** [5, 1, 2]**:** Although the Crank-Nicolson method can sample exactly the invariant measure of ergodic stochastic differential equations in the Gaussian case, it is only A-stable and lacks the L-stability property which is desirable for a fast convergence to equilibrium. Using the idea of post-processing, we investigate how the L-stability property and the exactness for the invariant measure in the Gaussian case can be achieved simultaneously. We present such schemes applied to nonlinear ergodic problems in the context of implicit Runge-Kutta methods [5] and in the context of explicit stabilized Runge-Kutta methods [1], which can be shown to be strongly convergent [2] for a class of quasilinear parabolic stochastic PDEs, including the quasilinear stochastic heat equation with space-time white noise.

[1] The slides of the talk are available at: `www.unige.ch/~vilmart/talks.html`

References

[1] A. Abdulle, I. Almuslimani, and G. Vilmart, *Optimal explicit stabilized integrator of weak order one for stiff and ergodic stochastic differential equations*, SIAM/ASA J. Uncertain. Quantif. **6** (2018), 937–964.

[2] A. Abdulle, C.-E. Bréhier, and G. Vilmart, *Convergence analysis of explicit stabilized integrators for parabolic semilinear stochastic PDEs*, arXiv:2102.03209, Submitted.

[3] G. Bertoli and G. Vilmart, *Strang splitting method for semilinear parabolic problems with inhomogeneous boundary conditions: a correction based on the flow of the nonlinearity*, SIAM J. Sci. Comput. **42** (2020), A1913–A1934.

[4] G. Bertoli, C. Besse, and G. Vilmart, *Superconvergence of the Strang splitting when using the Crank-Nicolson scheme for parabolic PDEs with oblique boundary conditions*, arXiv:2011.05178, submitted.

[5] C.-E. Bréhier and and G. Vilmart, *High-order integrator for sampling the invariant distribution of a class of parabolic stochastic PDEs with additive space-time noise*, SIAM J. Sci. Comput **38** (2016) A2283–A2306.

A convergent evolving finite element algorithm for Willmore flow of closed surfaces

Balázs Kovács

(joint work with Buyang Li and Christian Lubich)

The elastic bending energy or Willmore energy (named after Thomas Willmore) of a surface $\Gamma := \Gamma[X(\cdot,t)] = \{X(p,t) : p \in \Gamma^0\}$, described by a flow map $X : \Gamma^0 \times [0,T] \to \mathbb{R}^3$, is given as

$$W(\Gamma) = \frac{1}{2}\int_\Gamma H^2,$$

where H is the mean curvature of the surface. *Willmore flow* is the L^2 gradient flow of surfaces for the elastic bending energy. It plays an important role in modelling lipid bilayers, biomembranes, vesicles, regularization of phase-field systems, etc.

The negative L^2 gradient of the Willmore energy W for a two-dimensional surface Γ in $\mathbb{R}^3$ has no tangential contribution and reads

$$v = \Big(\Delta_\Gamma H + H(\tfrac{1}{2}H^2 - 2K)\Big)\,\nu \qquad \text{on } \Gamma, \tag{1}$$

with outward normal vector field ν, mean curvature $H = \kappa_1 + \kappa_2$ (here taken without a factor 1/2) and Gaussian curvature $K = \kappa_1\kappa_2$, where κ_1 and κ_2 are the principal curvatures on Γ. Willmore gives a proof of this result and attributes the formula (in the stationary case) to Thomsen (1924) (who mentions Schadow (1922)) and Blaschke (1926). For curves the study of "Willmore energy" goes back to D. Bernoulli, S. Germain and S.D. Poisson.

Numerical methods for Willmore flow based on evolving surface finite element methods have been proposed by Rusu (2005), Dziuk (2008), and Barrett, Garcke & Nürnberg (2007,2008) each based on different variational formulations of (1). However, convergence of a surface finite element method for the Willmore flow of *closed surfaces* has remained an open problem.

The objective of [3] *is to construct a convergent evolving surface finite element algorithm for the Willmore flow of closed two-dimensional surfaces in three-dimensional space.*

In this Oberwolfach report we focus on two main aspects of the approach taken in [3], which lead to achieve the above goal:

(i) Our algorithm is based on a coupled system of fourth-order evolution equations for the geometric quantities in (1), inspired by the approach of Huisken (1984) and our previous work on mean curvature flow [2].
(ii) To present a novel stability analysis based on energy estimates, by (essentially) compressing a 13 page proof to a single figure.

(i) The key idea is to use fourth-order parabolic evolution equations for the mean curvature H and the normal vector ν along the Willmore flow:

$$\begin{aligned}\partial^\bullet H = & -\big(\Delta_\Gamma + |A|^2\big)(\Delta_\Gamma H + Q),\\ \partial^\bullet \nu = & \big(-\Delta_\Gamma + (HA - A^2)\big)(\Delta_\Gamma \nu + |A|^2\nu) + |\nabla_\Gamma H|^2 \nu\\ & - 2\big(\nabla_\Gamma \cdot (A\nabla_\Gamma H)\big)\,\nu - A^2\nabla_\Gamma H - \nabla_\Gamma Q,\end{aligned}$$

where $A = \nabla_\Gamma \nu$ and Q denotes the non-linear term in (1). We derive an algorithm based on the weak formulation of the system that couples these evolution equations to the velocity law (1) and to the ordinary differential equation $\partial_t X = v$ for the positions. Here, H and ν are considered to be independently evolving unknowns that are not directly extracted from the surface at any given time. This is different from the previously mentioned approaches.

(ii) Stability, which is here understood as bounding the errors in terms of consistency defects and initial errors. For the velocity law, stability is shown using a stability bound for the interpolation of products of surface finite element functions. The main idea for the stability estimates for the second-order system for the geometric variables is to exploit the *anti-symmetric* structure of the semi-discrete error equations and combine it with *multiple energy estimates*, testing with both the errors and their time derivative. The structure of the energy estimates is sketched in Figure 1. The proof is performed in the matrix–vector formulation of the numerical method, and it uses technical lemmas relating different finite element surfaces that were shown in [1] and [2] (these are vastly *more* important than this short report, or the corresponding talk, could suggest). A key step in the proof is to establish $W^{1,\infty}$-norm error bounds for all variables, which are obtained from the time-uniform H^1-norm error bounds using an inverse estimate.

In Theorem 4.1 of [3] optimal-order H^1-norm semi-discrete convergence estimates are shown for all variables X, v, ν, H. Convergence is shown towards sufficiently regular solutions of Willmore flow, which excludes the formation of singularities (within the considered time interval).

For a perturbed torus the surface evolution (towards a Clifford torus) and Willmore energy ($W(\Gamma) \geq 4\pi^2$) is shown in Figure 2.

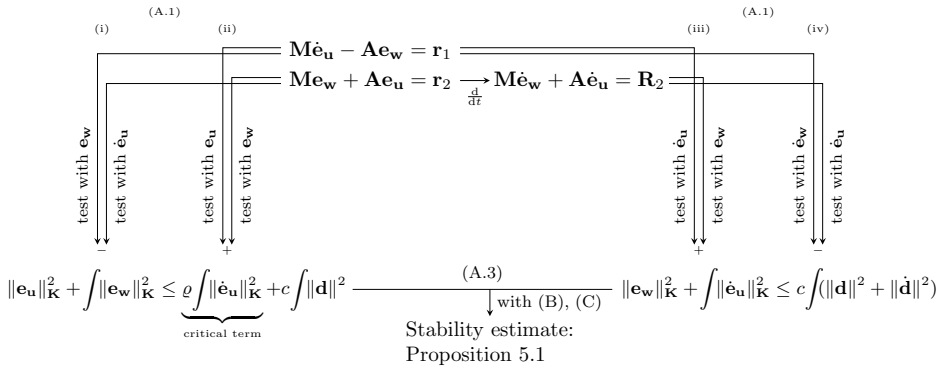

FIGURE 1. Sketch of the structure of the energy estimates.

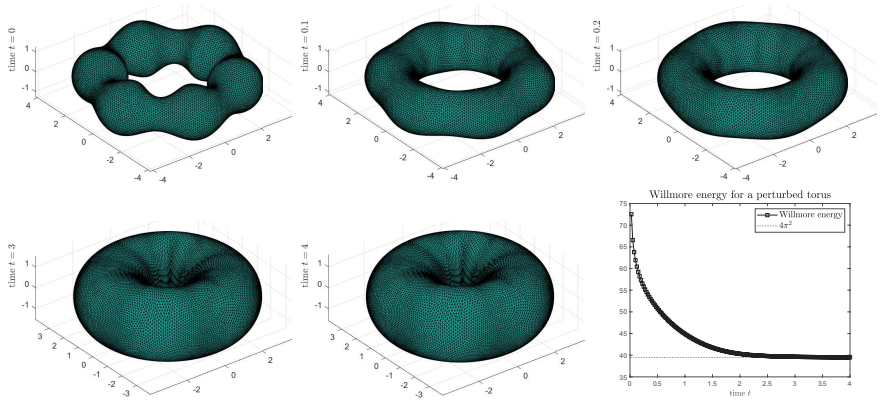

FIGURE 2. Surface evolutions at different times and corresponding Willmore energy for a perturbed torus.

REFERENCES

[1] B. Kovács, B. Li, C. Lubich, and C.A. Power Guerra. Convergence of finite elements on an evolving surface driven by diffusion on the surface. *Numer. Math.*, 137(3):643–689, 2017.

[2] B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for mean curvature flow of closed surfaces. *Numer. Math.*, 143(4):797–853, 2019.

[3] B. Kovács, B. Li, and C. Lubich. A convergent evolving finite element algorithm for Willmore flow of closed surfaces. arXiv:2007.15257, 2020.

Functional equivariance of numerical integrators

Ari Stern
(joint work with Robert I. McLachlan)

In numerical ordinary differential equations, it is well known that all B-series methods (including Runge–Kutta methods) preserve linear invariants, while only certain ones preserve quadratic invariants. Linear invariants arising in physical systems include mass, charge, and linear momentum; quadratic invariants include angular momentum and other momentum maps, as well as the canonical symplectic form for Hamiltonian systems. (See [1] and references therein.)

However, for partial differential equations describing time evolution, it is desirable for a numerical integrator to preserve not only global invariants but also local conservation laws. For instance, the evolution may preserve total mass (a global invariant), but the mass in a particular region may change by flowing through the boundary of the region (a local conservation law). Another example is the canonical multisymplectic conservation law for Hamiltonian PDEs, which is a quadratic local conservation law for the variational equation.

In order to answer the question of which methods preserve local conservation laws such as these, we investigate an even more general question: When does a numerical integrator preserve the evolution of certain classes of observables (e.g., linear, quadratic), even when these observables are not invariant?

This talk presents forthcoming work [2] that answers this question by studying the *functional equivariance* of numerical integrators. Given a dynamical system $\dot{y} = f(y)$, we say that a method is equivariant with respect to a functional F if applying the method to the augmented system

$$\dot{y} = f(y), \qquad \dot{z} = F'(y)f(y),$$

preserves the relation $z = F(y)$. We prove that, when $\mathcal{F}$ is a class of functionals including linear functionals, a method is functionally equivariant with respect to $\mathcal{F}$ if and only if it preserves invariants in $\mathcal{F}$. In particular, this implies that all B-series methods preserve local conservation laws involving linear quantities (like mass and charge), while those preserving quadratic invariants also preserve local conservation laws involving quadratic quantities, including the multisymplectic conservation law.

References

[1] E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration*, vol. 31 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 2006.
[2] R. I. McLachlan and A. Stern, *Functional equivariance and conservation laws in numerical integration*, in preparation.

Implicit symplectic methods for high precision numerical integration of the Solar System

Ander Murua
(joint work with M. Antoñana, J. Makazaga, E. Alberdi)

High precision long-term dynamic simulation of planetary systems requires computationally expensive numerical integrations. Efficient computation of long-term ephemerides of the Solar System can be achieved by exploiting the near Keplerian motion of the planets around the Sun. The main idea is to alternate Keplerian motions with appropriate corrections accounting for the planetary interactions. In this sense, symplectic splitting integrators for high precision integration of the Solar System has been constructed and tested in [2, 3]. Further improvement of the efficiency of numerical simulations demands the development of fast and accurate algorithms that take advantage of parallel computer architectures. Although some attempts have been made in this sense for symplectic splitting integrators for Solar System simulations, the sequential nature of that kind of schemes makes it difficult to get substantial improvements from parallelization strategies.

In this work, we present an optimized implementation of the 16th order symplectic FCIRK method for Hamiltonian systems of the form

$$H(q,p) = H_K(q,p) + H_I(q,p),$$
$$H_K(q,p) = \sum_{i=1}^{n} \left(\frac{\|p_i\|^2}{2\mu_i} - \frac{\mu_i\, k_i}{\|q_i\|} \right),$$

where $q = (q_1, \ldots, q_n)$, $p = (p_1, \ldots, p_n)$, $q_i, p_i \in \mathbb{R}^3$, $i = 1, \ldots, n$. The interaction Hamiltonian $H_I(q,p)$ is assumed to be of smaller magnitude than the Keplerian Hamiltonian $H_K(q,p)$. In particular, the $(n+1)$-body problem corresponding to n planets orbiting a central star can be rewritten in this form when written in appropriate variables (such as Jacobi coordinates or Poincaré's canonical Heliocentric coordinates). Of course, the Hamiltonian of models including natural satellites of some of the planets also can be written in this form.

The propagation of roundoff errors is a key limiting factor for very long term integrations of the Solar System, and sometimes the precision provided by the standard double precision arithmetic is not good enough. A good option may be to use instead the extended precision 80-bit floating point arithmetic (one bit for the sign, 15 bits for the exponent, and 64 bits for the significand) available for most intel processors. Compared to double precision arithmetic (with 53-bit significands) 11 additional binary digits of precision are gained by making all the computations in extended precision arithmetic. To further reduce roundoff errors, one could resort to quadruple precision arithmetic, but doing all the computations in that arithmetic is generally not an option, as it would require at least twenty times more computing time. An intermediate solution is to use a mixed-precision arithmetic: to use quadruple precision for the most critical computations and perform the rest of the computation in a lower precision arithmetic. In the case of explicit splitting methods, a natural mixed-precision strategy is to compute

the Keplerian flows in quadruple precision, and the corrections corresponding to the planetary interactions (which are typically of smaller magnitude) in 80-bit arithmetic.

In [1], a new family of symplectic integration methods, flow-composed implicit Runge-Kutta (FCIRK) methods, is introduced which is also appropriate for the numerical integration of the Solar System, or other systems with near-Keplerian motions. They are similar to symplectic splitting schemes in that Keplerian motions are alternated with corrections of smaller magnitude. But while in symplectic splitting integrators the corrections are computed as the solution operator of the interaction Hamiltonian, in FCIRK methods, such corrections correspond to the application to a transformed ODE system of one step of an implicit Runge-Kutta (IRK) method based on collocation with Gauss-Legendre nodes. As their underlying IRK methods, FCIRK methods are super-convergent: the method based on s Gauss-Legendre nodes is of order $2s$. After some preliminary numerical experiments with high precision Solar System simulations, we have concluded that the method with $s = 8$ nodes (hence, of order $2s = 16$) is a good choice. We refer to that method as FCIRK16.

Compared to symplectic splitting integrators, the implementation of FCIRK methods is more involved as an implicit system of equations has to be solved at each time-step. In return, such methods are better suited than symplectic splitting integrators

- for exploiting parallel computer architectures (most of the computations can be done in s processors in parallel), and
- for the mixed-precision approach mentioned above, as fewer Keplerian motions in quadruple precision have to be computed for the same level of precision and a given integration interval.

Our implicit symplectic integrators are based on the following factorization of the h-flow φ_h of $H = H_K + H_I$

$$\varphi_h = \varphi^K_{h/2} \circ \psi_h \circ \varphi^K_{h/2} \tag{1}$$

where φ^K_h is the h-flow of H_K and

$$\psi_h := \varphi^K_{-h/2} \circ \varphi_h \circ \varphi^K_{-h/2}.$$

We show that for any solution $U(t)$ of

$$\frac{d}{dt}U = F(U, t - h/2), \quad \text{where} \quad F(U,t) := J^{-1}\nabla_U H_I(\varphi^K_t(U)), \tag{2}$$

it holds that

$$\psi_h(U(0)) = U(h).$$

One step of a FCIRK method is obtained by replacing in (1) the symplectic map ψ_h by a high order aproximation obtained by applying a symplectic implicit Runge-Kutta method to (2).

In our preliminary implementation of FCIRK schemes [1], time-steps of constant length were used. However, for long-term simulations of realistic models

of the Solar System, the larger bodies of the asteroid belt (specially Ceres, Pallas, and Vesta) experience close encounters that have a significant impact in the chaotic behavior of the Solar System [4]. Unfortunately, the precision of the numerical integration performed with any integrator with constant time-steps (in physical time) is degraded during close enough encounters, a degradation that may be more pronounced for FCIRK scheme (and the Wisdom-Holman map and its generalizations) due to the temporary loss of the hierarchy $H_I << H_K$ during extreme close encounters. On the other hand, it is known that the advantages of symplectic integrators for the long-term integration of Hamiltonian systems are lost if standard adaptive time-stepping strategies are used. Motivated by that, and under the assumption that close encounters that seriously degrade the accuracy seldom occur, a mild form of time-stepping adaptivity is implemented in our code: an automatic mechanism is used to identify steps with significantly larger local errors (presumably, due to close encounters) which are then resolved with higher precision time-steps; constant time-step length is used otherwise, so that the integration enjoys the advantages of constant time-step symplectic integration during long integration subintervals.

In our numerical experiments, we consider a Newtonian model of the Solar System with 16 point masses: the Sun, the eight planets, Pluto, the Moon as a separated body, and five of the main bodies in the Asteroid Belt: Ceres, Pallas, Vesta, Iris, Bamberga. In our results for that particular model,we conclude that:

- FCIRK16 is very well suited for the mixed precision approach: In this model, the above-mentioned mixed precision approach increases considerably (by a factor of six) the CPU time of symplectic splitting integrators. In contrast, in our FCIRK implementation, the mixed precision approach comes essentially for free.
- The parallel version of FCIRK16 implemented with OpenMP with four threads, a speedup of 2 is obtained. This results in a better efficiency than the best symplectic splitting methods in [2] for precisions higher than standard double precision arithmetic.

In addition, compared to symplectic splitting integrators, that only admit interaction Hamiltonians of the form $H_I(p,q) = H_C(p) + H_D(q)$, FCIR16 admits Hamiltonians of arbitrary form (in particular, those arising from relativistic corrections).

References

[1] M. Antoñana, J. Makazaga, A. Murua, *New Integration Methods for Perturbed ODEs Based on Symplectic Implicit Runge-Kutta Schemes with Application to Solar System Simulations*, J Sci Comput **76** (2018), 630–650.

[2] S. Blanes, F. Casas, A. Farrés, J. Laskar, J. Makazaga, and A. Murua, *New families of symplectic splitting methods for numerical integration in dynamical astronomy*, Appl. Numer. Math. **68** (2013), 58-72.

[3] A. Farrés, J. Laskar, S. Blanes, F. Casas, J. Makazaga, and A. Murua, *High precision Symplectic Integrators for the Solar System*, Cel. Mech. & Dyn. Astron. **116** (2013), 141-174.
[4] J. Laskar, M. Gastineau, J.B Delisle, A. Farrés, and A. Fienga, *Strong chaos induced by close encounters with ceres and vesta*, Astronomy and Astropysics **532** L4 (2011).

Deducing properties of ODEs from their discretization

Reinout Quispel
(joint work with David McLaren, Charalambos Evripidou)

1. Introduction

Consider a polynomial ordinary differential equation (ODE)

$$\frac{dx}{dt} = f(x) \tag{1}$$

In this paper we address the question

Q1: What rational conserved integral(s) and (inverse) polynomial volume form (if any) does the ODE possess?

Since finding rational integrals generally requires solving a nonlinear problem we propose a three step program, that, using a certain ansatz, only requires the solution of linear problems:

Step 1: Discretise the ODE using a "suitable" method. In this paper we will use Kahan's method (but much of the following also holds for certain other birational integration methods given in the references). Compute the Jacobian determinant J of the discretisation, and factorise J.

Step 2: Use the factors of J as candidate discrete cofactors for finding discrete Darboux polynomials (DPs).

Step 3: Take the continuum limits of the discrete cofactors and DPs found in step 2. If possible, use these DPs as building blocks for time-dependent/time-independent first integrals and preserved measure of the ODE, If one is very lucky, it may even be possible to use them to derive the exact solution to the initial value problem fore the ODE.

2. Step 1

Our ongoing example in this paper will be the ODE

$$\begin{aligned} \dot{x} &= 2 - 2x + xz \\ \dot{y} &= -y + yz \\ \dot{z} &= -y - 3z + z^2 \end{aligned} \tag{2}$$

For any quadratic ODE:

$$\frac{dx_i}{dt} = \sum_{j,k} a_{ijk} x_j x_k + \sum_j b_{ij} x_j + c_i \tag{3}$$

Kahan's "unconventional" method is defined by

$$\frac{x_i' - x_i}{h} = \sum_{j,k} a_{ijk} \frac{x_j' x_k + x_j x_k'}{2} + \sum_j b_{ij} \frac{x_j + x_j'}{2} + c_i \tag{4}$$

here $x_i := x_i(nh)$, $x_i' := x_i((n+1)h)$, and h is the timestep.

It is not hard to show that eq(4) can be rearranged as follows:

$$\frac{x' - x}{h} = \left(I - \frac{h}{2} f'(x) \right)^{-1} f(x), \tag{5}$$

This defines the Kahan map $x_{n+1} = \phi(x_n)$ [1].

Next we compute the Jacobian determinant J of ϕ:

$$J(x) = \left| \frac{\partial \phi_i(x)}{\partial x_j} \right|, \tag{6}$$

and use an algebraic manipulation package to factorise J.

For our example

$$J(x) = \frac{K_1 K_2 K_3 K_4}{D_1 D_2^4}, \tag{7}$$

where $K_i, D_j (i = 1, \ldots, 4; j = 1, 2)$ are given in the appendix.

3. Step 2

Given a map $x_{n+1} = \phi(x_n)$, a polynomial $P(x)$ is a (discrete) DP of ϕ if there exists a (non-tautological) rational function C s.t.

$$P(x_{n+1}) = C(x_n) P(x_n), \tag{8}$$

where $P(x_{n+1}) = P(\phi(x_n))$ and C is called the (discrete) cofactor of P [2, 3].

i	C_i	P_i	C_i
1	$\frac{K_1}{D_2}$	$z-y-3$	z
2	$\frac{K_2}{D_2}$	$2z+y$	$z-3$
3	$\frac{K_3}{D_2}$	y	$z-1$
4	$\frac{K_4}{D_1 D_2}$	$x+y+z-1$	$z-2$

TABLE 1

Ansatz: Given a rational map ϕ with Jacobian determinant

$$J(\mathbf{x}) = \frac{\prod_{i=1}^{l} K_i^{b_i}(\mathbf{x})}{\prod_{j=1}^{m} D_j^{m_j}(\mathbf{x})}, \tag{9}$$

we try all cofactors (up to a certain polynomial degree d) of the form

$$C(\mathbf{x}) = \pm\frac{\prod_{i=1}^{l} K_i^{f_i}(\mathbf{x})}{\prod_{j=1}^{m} D_j^{g_j}(\mathbf{x})}, \tag{10}$$

where $f_i, g_j \in \mathbb{N}_0$.

NOTE

(1) There is a finite number of these co-factors up to a certain degree. For each of this finite number of co-factors, we only need to solve a linear problem (up to a chosen degree)!
(2) If $C(\mathbf{x}) = J(\mathbf{x})$, the corresponding Darboux polynomials are inverse densities of preserved measures.

The discrete cofactors C_i and corresponding DPs P_i for our example are given in the first two columns of Table 1:

4. STEP 3

The continuum limits $\bar{P}_i$, $\bar{C}_i$ are given by $lim_{h\to 0} P_i$ resp. $lim_{h\to 0}\frac{C_i-1}{h}$, and satisfy the ODEs

$$\dot{\bar{P}}_i = \bar{C}_i \bar{P}_i \tag{11}$$

A useful property of the cofactor $\bar{C}_i$ is [4]

$$\dot{\bar{P}}_i = \bar{C}_i \bar{P}_i \to \dot{\mathcal{P}} = \mathcal{C}\mathcal{P} \tag{12}$$

where

$$\mathcal{P} := \prod_i \bar{P}_i^{\alpha_i}, \mathcal{C} := \sum_i \alpha_i \bar{C}_i \tag{13}$$

This has the following implications:

(1) $\mathcal{C}(x) = 0 \to \dot{\mathcal{P}} = 0 \to \mathcal{P}$ is a first integral
(2) $\mathcal{C}(x) = C \to \dot{\mathcal{P}} = C\mathcal{P} \to \mathcal{P}(x(t)) = \mathcal{P}(x(0))e^{Ct}$
(3) $\mathcal{C}(x) = div f(x) \to \dot{\mathcal{P}} = \mathcal{C}\mathcal{P} \to f$ preserves the measure $\frac{dx}{\mathcal{P}(x)}$

For our problem, the $\bar{C}_i$ are given in the last column of Table 1. (For affine DPs, $\bar{P}_i = P_i$. For two theorems regarding affine DPs, cf [2]).

Note that

$$\bar{C}_1 - \bar{C}_2 = 3, \bar{C}_1 - \bar{C}_3 = 1, \bar{C}_1 - \bar{C}_4 = 2 \tag{14}$$

Hence

$$\frac{P_1}{P_2} = \frac{z-y-3}{2z+y} = k_2 e^{3t} \tag{15}$$

$$\frac{P_1}{P_3} = \frac{z-y-3}{y} = k_3 e^{t} \tag{16}$$

$$\frac{P_1}{P_4} = \frac{z-y-3}{x+y+z-1} = k_4 e^{2t} \tag{17}$$

and this yields 2 time-independent first integrals:

$$I_1 = \frac{P_3^2}{P_1 P_4} = \frac{y^2}{(z-y-3)(x+y+z-1)} \tag{18}$$

$$I_2 = \frac{P_3 P_4}{P_1 P_2} = \frac{y(x+y+z-1)}{(z-y-3)(2z+y)} \tag{19}$$

Hence f is integrable. Moreover $J = C_1 C_2 C_3 C_4$ implies that f preserves the measure

$$\frac{dxdydz}{P_1 P_2 P_3 P_4} = \frac{dxdydz}{y(2z+y)(z-y-3)(x+y+z-1)} \tag{20}$$

Equations (15), (16), & (17) can be combined to give the explicit solution of ODE (2)

$$x = \frac{6e^{2t}k_4k_2 + 2\left(-3e^{t}k_2 + \left(1 + k_2e^{3t}\right)k_4\right)k_3}{k_4\left(2e^{3t}k_2k_3 + 3e^{2t}k_2 - k_3\right)} \tag{21}$$

$$y = \frac{6k_2e^{2t}}{-2e^{3t}k_2k_3 - 3e^{2t}k_2 + k_3} \tag{22}$$

$$z = \frac{-3k_3 + 3e^{2t}k_2}{2e^{3t}k_2k_3 + 3e^{2t}k_2 - k_3} \tag{23}$$

Acknowledgement: We are grateful to our collaborators E Celledoni, G Gubbiotti, R McLachlan, B Owren and B Tapley for many illuminating discussions.

Appendix

The explicit factors of J(x) in eq (7) are:

$$
\begin{aligned}
K_1 &= -1/4\,h^2x_2 + 3/4\,h^2x_3 - 3/4\,h^2 - 1/2\,hx_3 - h + 1 \\
K_2 &= -1/4\,h^2x_2 - 1/4\,h^2x_3 - 3/4\,h^2 - 1/2\,hx_3 + h + 1 \\
K_3 &= -1/4\,h^2x_2 - 3/4\,h^2x_3 + 3/4\,h^2 - 1/2\,hx_3 + 2\,h + 1 \\
K_4 &= 1/8\,h^3x_2x_3 - 1/8\,h^3{x_3}^2 - 1/4\,h^3x_2 + \frac{7\,h^3x_3}{8} + 1/4\,h^2{x_3}^2 - 3/4\,h^3 \\
&\quad -1/4\,h^2x_2 - 1/4\,h^2x_3 - 5/4\,h^2 - hx_3 + h + 1 \\
D_1 &= -1/2\,hx_3 + h + 1 \\
D_2 &= 1/2\,h^2{x_3}^2 + 1/4\,h^2x_2 - 5/4\,h^2x_3 + 3/4\,h^2 - 3/2\,hx_3 + 2\,h + 1
\end{aligned}
$$

References

[1] Celledoni E, McLachlan RI, Owren B, Quispel, GRW 2013, Geometric properties of Kahan's method *J. Phys. A* **46** 12 025201

[2] Celledoni E, Evripidou C, McLaren DI, Owren B, Quispel GRW, Tapley BK, and van der Kamp P, Using discrete Darboux polynomials to detect and determine preserved measures and integrals of rational maps, *J. Phys. A* **52** (2019) 31 31LT01

[3] Celledoni E, Evripidou C, McLaren DI, Owren B, Quispel GRW, and Tapley BK, Detecting and determining preserved measures and integrals of rational maps, submitted to J.Phys. A (2021)

[4] Goriely A, Integrability and Nonintegrability of Dynamical Systems, World Scientific Publishing, Singapore (2001), Section 2.5

Reporter: Gianluca Ceruti

Participants

Prof. Dr. Assyr Abdulle
Chair of Computational Mathematics
and Numerical Analysis
Institute of Mathematics
École Polytechnique Fédérale de
Lausanne
Station 8
1015 Lausanne
SWITZERLAND

Prof. Dr. Weizhu Bao
Department of Mathematics
National University of Singapore
Lower Kent Ridge Road
Singapore 119 076
SINGAPORE

Dr. Joackim Bernier
Institut de Mathématiques de Toulouse
Université Paul Sabatier
118, route de Narbonne
31062 Toulouse Cedex 9
FRANCE

Prof. Dr. Sergio Blanes
Instituto de Matematica Multidisciplinar
Universitat Politecnica de Valencia
Edificio 8-G, piso 2
Camino de Vera/sn
46022 Valencia
SPAIN

Dr. Geir Bogfjellmo
7491 Trondheim
NORWAY

Dr. Nawaf Bou-Rabee
Department of Mathematical Sciences,
Rutgers
Camden NJ 08102
UNITED STATES

Prof. Dr. Chris J. Budd
Dept. of Mathematical Sciences
University of Bath
Claverton Down
Bath BA2 7AY
UNITED KINGDOM

Prof. Dr. Fernando Casas
Departament de Matematiques
Universitat Jaume I
12071 Castellon
SPAIN

Prof. Dr. Elena Celledoni
Department of Mathematical Sciences
Norwegian University of Science and
Technology
A. Getz vei 1
7491 Trondheim
NORWAY

Gianluca Ceruti
Mathematisches Institut
Universität Tübingen
Auf der Morgenstelle 10
72076 Tübingen
GERMANY

Dr. Philippe Chartier
INRIA
Université de Rennes 1
Campus de Beaulieu
35042 Rennes Cedex
FRANCE

Dr. David Cohen
Mathematical Sciences
Chalmers University of Technology and
University of Gothenburg
SE-412 96 Gothenburg
Sweden
412 96 Göteborg
SWEDEN

Dr. Nicolas Crouseilles
Inria
I.R.M.A.R.
Université de Rennes I
Campus de Beaulieu
35042 Rennes Cedex
FRANCE

Dr. Charles Curry
Department of Mathematical Sciences
NTNU
7491 Trondheim
NORWAY

Benjamin Dörich
Institut für Angewandte und
Numerische Mathematik
Karlsruher Institut für Technologie
(KIT)
Englerstrasse 2
76131 Karlsruhe
GERMANY

Kurusch Ebrahimi-Fard
Department of Mathematical Sciences
Norwegian University of Science and
Technology
Alfred Getz Vei 1
7491 Trondheim
NORWAY

Assoc. Prof. Dr. Lukas Einkemmer
Institut für Mathematik
Universität Innsbruck
Technikerstrasse 13
6020 Innsbruck
AUSTRIA

Alejandro Escorihuela-Tomàs
Departament de Matemàtiques
Universitat Jaume I
12071 Castellon
SPAIN

Prof. Dr. Erwan Faou
INRIA
Département de Mathématiques
ENS Rennes
Campus Ker Lann
Avenue Robert Schumann
35170 Bruz
FRANCE

Prof. Dr. Jason E. Frank
Mathematisch Instituut
Universiteit Utrecht
Budapestlaan 6
3584 CD Utrecht
NETHERLANDS

Dr. Francois Gay-Balmaz
CNRS / Laboratoire de meteorologie
dynamique
École Normale Supérieure
24, rue Lhomond
75005 Paris Cedex
FRANCE

Prof. Dr. Ernst Hairer
Département de Mathématiques
Université de Geneve
C.P. 64
2-4, rue du Lievre
1211 Genève 4
SWITZERLAND

Prof. Dr. Marlis Hochbruck
Institut für Angewandte und
Numerische Mathematik
Karlsruher Institut für Technologie
(KIT)
Englerstrasse 2
76131 Karlsruhe
GERMANY

Prof. Dr. Arieh Iserles
Department of Applied Mathematics and Theoretical Physics
University of Cambridge
Wilberforce Road
Cambridge CB3 0WA
UNITED KINGDOM

Prof. Dr. Tobias Jahnke
Institut für Angewandte und Numerische Mathematik
Karlsruher Institut für Technologie (KIT)
Englerstrasse 2
76131 Karlsruhe
GERMANY

Dr. Balázs Kovács
Mathematisches Institut
Universität Tübingen
Auf der Morgenstelle 10
72076 Tübingen
GERMANY

Dr. Michael Kraus
Max-Planck-Institut für Plasmaphysik
Boltzmannstr. 2
85748 Garching bei München
GERMANY

Prof. Dr. Benedict Leimkuhler
School of Mathematics
University of Edinburgh
James Clerk Maxwell Bldg.
King's Buildings, Mayfield Road
Edinburgh EH9 3JZ
UNITED KINGDOM

Prof. Dr. Melvin Leok
Department of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0112
UNITED STATES

Prof. Dr. Christian Lubich
Mathematisches Institut
Universität Tübingen
Auf der Morgenstelle 10
72076 Tübingen
GERMANY

Prof. Dr. Robert McLachlan
Institute of Fundamental Sciences
Massey University
Private Bag 11-222
Palmerston North 5300
NEW ZEALAND

Dr. Yuto Miyatake
Cybermedia Center
Osaka University
Machikaneyama 1-32, Toyonaka
Osaka 560-0043
JAPAN

Assoc. Prof. Dr. Klas Modin
Department of Mathematical Sciences
Chalmers University of Technology
412 96 Göteborg
SWEDEN

Prof. Dr. Hans Z. Munthe-Kaas
Department of Mathematics
University of Bergen
Postbox 7803
5020 Bergen
NORWAY

Prof. Dr. Ander Murua
Informatika Fakultatea
Universidad del Pais Vasco
Lardizabal pasealehua,1
20018 Donostia/San Sebastian 20018
SPAIN

Dr. Sina Ober-Blöbaum
Institut für Mathematik
Universität Paderborn
Warburger Str. 100
33098 Paderborn
GERMANY

Christian Offen
Institut für Mathematik
Universität Paderborn
Warburger Straße 100
33098 Paderborn
GERMANY

Prof. Dr. Peter J. Olver
School of Mathematics
University of Minnesota
127 Vincent Hall
206 Church Street S. E.
Minneapolis MN 55455-0436
UNITED STATES

Prof. Dr. Alexander Ostermann
Institut für Mathematik
Universität Innsbruck
Technikerstrasse 13
6020 Innsbruck
AUSTRIA

Prof. Dr. Brynjulf Owren
Department of Mathematical Sciences
NTNU
7491 Trondheim
NORWAY

Prof. Dr. Reinout Quispel
Department of Mathematics
La Trobe University
Bundoora 3086
AUSTRALIA

Prof. Dr. Oliver Sander
Institut für Numerische Mathematik
Technische Universität Dresden
Zellescher Weg 12-14
01069 Dresden
GERMANY

Prof. Dr. Jesús Maria Sanz-Serna
Departamento de Matematicas
Universidad Carlos III de Madrid
Avenida de la Universidad, 30
28911 Leganes (Madrid)
SPAIN

Prof. Dr. Carola-Bibiane Schönlieb
Department of Applied Mathematics and
Theoretical Physics (DAMTP)
Centre for Mathematical Sciences
Wilberforce Road
Cambridge CB3 0WA
UNITED KINGDOM

Stefan Schrammer
Institut für Angewandte und
Numerische Mathematik
Karlsruher Institut für Technologie
(KIT)
Englerstrasse 2
76131 Karlsruhe
GERMANY

Prof. Dr. Katharina Schratz
Laboratoire Jacques-Louis Lions
Sorbonne Université
4, place Jussieu
75252 Paris Cedex 05
FRANCE

Ferdia Sherry
Department of Applied Mathematics and
Theoretical Physics (DAMTP)
Centre for Mathematical Sciences
Wilberforce Road
Cambridge CB3 0WA
UNITED KINGDOM

Prof. Dr. Eric Sonnendrücker
Max-Planck-Institut für Plasmaphysik
Boltzmannstraße 2
85748 Garching bei München
GERMANY

Prof. Dr. Ari Stern
Department of Mathematics and Statistics
Washington University in St. Louis
Campus Box 1146
One Brookings Drive
St. Louis, MO 63130-4899
UNITED STATES

Prof. Dr. Yuri B. Suris
Institut für Mathematik
Technische Universität Berlin
Sekr. MA 7-1
Straße des 17. Juni 136
10623 Berlin
GERMANY

Prof. Dr. Bart Vandereycken
Section of Mathematics
University of Geneva
Rue du Lièvre 2-4
1211 Genève 4
SWITZERLAND

Dr. Mats Vermeeren
School of Mathematics,
University of Leeds
LS2 9JT Leeds
UNITED KINGDOM

Dr. Gilles Vilmart
Section de Mathématiques
Université de Geneve
UNI DUFOUR
24, rue du Général Dufour
P.O. Box 64
1211 Genève 4
SWITZERLAND

Milo Viviani
Scuola Normale Superiore di Pisa
Piazza dei Cavalieri 7
56126 Pisa
ITALY

Dr. Anke Wiese
Department of Actuarial Mathematics and Statistics
Heriot-Watt University
Riccarton
Edinburgh EH14 4AS
UNITED KINGDOM

Prof. Dr. Antonella Zanna
Department of Mathematics
University of Bergen
P.B. 7803
Allégaten 41
5020 Bergen
NORWAY

Dr. Konstantinos Zygalakis
School of Mathematics
University of Edinburgh
James Clerk Maxwell Bldg.
Edinburgh EH9 3FD
UNITED KINGDOM