Energy-diminishing integration of gradient systems

Ernst Hairer*

Section de mathématiques, Université de Genève, 2-4 rue du Lièvre, CH-1211 Genève 4, Switzerland
*Corresponding author: ernst.hairer@unige.ch

AND

Christian Lubich

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

For gradient systems in Euclidean space or on a Riemannian manifold the energy decreases monotonically along solutions. Algebraically stable Runge–Kutta methods are shown to also reduce the energy in each step under a mild step-size restriction. In particular, Radau IIA methods can combine energy monotonicity and damping in stiff gradient systems. Discrete-gradient methods and averaged vector field collocation methods are unconditionally energy-diminishing, but cannot achieve damping for very stiff gradient systems. The methods are discussed when they are applied to gradient systems in local coordinates as well as for manifolds given by constraints.

Keywords: gradient flow; energy dissipation; implicit Runge–Kutta method; algebraic stability; L-stability; discrete-gradient method; averaged vector field collocation.

1. Introduction

Given a potential function $U : \mathbf{R}^n \to \mathbf{R}$, the associated gradient system is the differential equation (see Hirsch & Smale, 1974; Hale, 1988)

$$\dot{y} = -\nabla U(y). \tag{1.1}$$

More generally, for a potential U defined on a Riemannian manifold $\mathcal{M}$ with metric g, the corresponding gradient system takes the form

$$g_y(\dot{y}, v) = -U'(y)v \quad \text{for all } v \in T_y\mathcal{M}, \tag{1.2}$$

where $T_y\mathcal{M}$ denotes the tangent space of $\mathcal{M}$ at $y \in \mathcal{M}$. Gradient systems arise in a variety of applications, both in a finite-dimensional and infinite-dimensional setting. They describe dissipative physical systems that evolve into a state of minimal energy. Examples include phase-field models in material science (Chen, 2002; Penzler *et al.*, 2012) such as the Allen–Cahn and the Cahn–Hilliard partial differential equations (Chafee & Infante, 1974; Elliott & Zheng, 1986), the porous medium equation (Otto, 2001), and models in image processing (Strzodka *et al.*, 2004; Michailovich *et al.*, 2007), in differential geometry (Droske & Rumpf, 2004; Bobenko & Schröder, 2005) and in quantum systems (Bao & Du, 2004).

A key property of the gradient flow is that the potential $U(y(t))$ decreases monotonically with t:

$$U(y(t)) \leqslant U(y(s)) \quad \text{for } t > s,$$

with strict inequality except at stationary points of U.

In the numerical treatment we would like to preserve the monotonicity, so that, after one step of the method starting from y_0 with a time step h, we would obtain

$$U(y_1) \leqslant U(y_0).$$

One approach to achieve this objective is to determine $y_1 \in \mathcal{M}$ from the condition (Stuart & Humphries, 1996, Section 4.4.6; Otto, 1998)

$$\frac{1}{2h} \operatorname{dist}(y_1, y_0)^2 + U(y_1) \to \min. \tag{1.3}$$

When the manifold is the Euclidean space $\mathbf{R}^n$, then this method becomes the implicit Euler method. When the manifold is embedded by constraints $c(y) = 0$, then it becomes the implicit Euler method constrained to the manifold,

$$y_1 = y_0 - h\nabla U(y_1) - h\nabla c(y_1)\lambda_1, \quad c(y_1) = 0.$$

The implicit Euler method only yields a first-order approximation. Our interest in this note is to explore higher-order methods that ensure a reduction of the potential.

In Section 2 we study the energy reduction in standard implicit Runge–Kutta methods, when applied to (1.1) or to (1.2) if the manifold is given by a local parametrization. It turns out that algebraically stable methods, such as Gauss and Radau IIA methods, behave favourably under a mild time-step restriction that involves a local Lipschitz bound of the Hessian matrix of U. In particular, the Radau IIA methods combine good energy reduction and favourable damping properties in stiff gradient systems. Gauss methods, on the other hand, have no damping property on very stiff gradient systems.

In Section 3 we discuss methods that reduce the energy without any restriction of the step size. It is known from McLachlan *et al.* (1999) that discrete-gradient methods have this property. A particularly interesting example is the average vector field integrator which, applied to (1.1), reads

$$y_1 = y_0 - h \int_0^1 \nabla U(y_0 + \tau(y_1 - y_0)) \, d\tau.$$

Generalizations of this method to higher order are the modified collocation methods of Hairer (2010), which were constructed such that they conserve the energy in Hamiltonian systems. When applied to gradient systems, also these methods diminish the energy without any step-size restriction. Like the Gauss methods, these integrators have no damping property on very stiff gradient systems.

In Section 4 we consider the differential equation (1.2) where the manifold is given by constraints $c(y) = 0$, so that we are concerned with a differential-algebraic equation. We study the energy reduction of implicit Runge–Kutta methods and of suitable extensions to differential-algebraic equations of the discrete-gradient methods and the averaged vector field collocation methods.

2. Implicit Runge–Kutta methods for gradient flows

We consider the gradient system (1.2) in coordinates,

$$G(y)\dot{y} = -\nabla U(y), \tag{2.1}$$

with a potential U that is twice continuously differentiable. The symmetric matrix $G(y)$ is assumed to satisfy, for all y and for all vectors $v \neq 0$,

$$v^{\mathrm{T}}G(y)v \geqslant v^{\mathrm{T}}\Gamma v > 0, \tag{2.2}$$

where Γ is a fixed positive-definite matrix. Along every exact solution we have

$$\frac{\mathrm{d}}{\mathrm{d}t}U(y(t)) = \nabla U(y(t))^{\mathrm{T}}\dot{y}(t) = -\dot{y}(t)^{\mathrm{T}}G(y(t))\dot{y}(t) \leqslant 0, \tag{2.3}$$

implying that $U(y(t))$ is monotonically decreasing.

For the numerical solution of (2.1) we consider implicit Runge–Kutta methods, given by

$$\begin{aligned} Y_i &= y_0 + h\sum_{j=1}^{s} a_{ij}\dot{Y}_j, \quad i = 1, \ldots, s, \\ y_1 &= y_0 + h\sum_{i=1}^{s} b_i\dot{Y}_i, \end{aligned} \tag{2.4}$$

with

$$G(Y_i)\dot{Y}_i = -\nabla U(Y_i). \tag{2.5}$$

Associated to the Runge–Kutta coefficients we consider the diagonal matrix $B = \mathrm{diag}(b_1, \ldots, b_s)$ and the symmetric $s \times s$ matrix M with entries

$$m_{ij} = b_i a_{ij} + b_j a_{ji} - b_i b_j.$$

We denote by δ a bound on the deviation from the starting value,

$$\|y_1 - y_0\| \leqslant \delta, \quad \|Y_i - y_0\| \leqslant \delta, \quad i = 1, \ldots, s. \tag{2.6}$$

Our aim is to prove a discrete analogue of (2.3).

THEOREM 2.1 We consider the gradient system (2.1) with a twice differentiable function $U(y)$ and we assume that the Hessian matrix $\nabla^2 U(y)$ is Lipschitz continuous in the ball of radius δ centred at y_0 with Lipschitz constant L. If the Runge–Kutta weights b_i are non-negative, if the matrix

$$B \otimes \Gamma + \frac{h}{2}M \otimes \nabla^2 U(y_0)$$

is positive definite and if $hL\delta$ is bounded by a sufficiently small constant, then we have

$$U(y_1) \leqslant U(y_0).$$

Proof. We follow the ideas of the proof for algebraic stability of implicit Runge–Kutta methods; see, for example, Hairer & Wanner, 1996, Section IV.12. Taylor series expansion of $U(y_1)$ around y_0 yields

$$U(y_1) = U(y_0) + h \sum_{i=1}^{s} b_i \nabla U(y_0)^{\mathrm{T}} \dot{Y}_i + \frac{h^2}{2} \sum_{i,j=1}^{s} b_i b_j \dot{Y}_i^{\mathrm{T}} \nabla^2 U(\hat{y}) \dot{Y}_j,$$

where $\hat{y}$ lies on the segment connecting y_0 with y_1. From the relation

$$\nabla U(Y_i) = \nabla U(y_0) + h \sum_{j=1}^{s} a_{ij} \int_0^1 \nabla^2 U(y_0 + s(Y_i - y_0)) \, \mathrm{d}s \dot{Y}_j,$$

we extract $\nabla U(y_0)$ and insert it into the previous relation. Using (2.5), this gives

$$U(y_1) = U(y_0) - h \sum_{i=1}^{s} b_i \dot{Y}_i^{\mathrm{T}} G(Y_i) \dot{Y}_i - \frac{h^2}{2} \sum_{i,j=1}^{s} \dot{Y}_i^{\mathrm{T}} (m_{ij} \nabla^2 U(y_0) + R_{ij}) \dot{Y}_j,$$

with remainder terms R_{ij} bounded by $cL\delta$, where c depends only on the coefficients of the Runge–Kutta method. With the relation (2.2) this yields

$$U(y_1) \leqslant U(y_0) - h\dot{Y}^{\mathrm{T}} \left(B \otimes \Gamma + \frac{h}{2} M \otimes \nabla^2 U(y_0) + \frac{h}{2} R \right) \dot{Y},$$

where $\dot{Y} = (\dot{Y}_i)_{i=1}^{s}$ and $R = (R_{ij})_{i,j=1}^{s}$. This proves the statement of the theorem. □

REMARK 2.2 For the implicit Euler method the condition of Theorem 2.1 requires that (for $G(y) = I$)

$$I + \frac{h}{2} \nabla^2 U(y_0) \text{ is positive definite.}$$

On the other hand, the condition for a minimum of (1.3) requires $I + h\nabla^2 U(y_1)$ to be positive semidefinite, which can give a stronger step-size restriction when the Hessian matrix $\nabla^2 U(y_1)$ has negative eigenvalues. To our knowledge, energy decay for the implicit Euler method under such a condition was first studied by Elliott & Stuart (1993, Proposition 5.5).

REMARK 2.3 Humphries & Stuart (1994, Theorem 4.6) showed, for $G(y) = I$, that, for *every* Runge–Kutta method, the energy decreases if the product of the step size h and the Lipschitz constant of ∇U (not that of the Hessian of U) is sufficiently small. In contrast, as we shall see next, Theorem 2.1 shows that, for *suitable* Runge–Kutta methods, the energy is diminished without any restriction of the step size by the Lipschitz constant of ∇U, which in the case of a predominantly quadratic potential can be much larger than the Lipschitz constant of the Hessian $\nabla^2 U$.

Theorem 2.1 shows that favourable Runge–Kutta methods are those with positive weights b_i and a symmetric positive-semidefinite matrix M. Such methods are known as *algebraically stable* methods and play an important role in the numerical treatment of stiff differential equations; see Burrage & Butcher (1979), Crouzeix (1979) and Hairer & Wanner (1996, Section IV.12). Well-known algebraically stable Runge–Kutta methods are the Gauss methods of order $2s$, which have $M = 0$, and the Radau IIA methods of order $2s - 1$.

We now consider potentials of the form

$$U(y) = \tfrac{1}{2}y^{\mathrm{T}}Ay + V(y), \tag{2.7}$$

with a symmetric positive-semidefinite matrix A of arbitrarily large norm. Such problems arise from the spatial discretization of Allen–Cahn and Cahn–Hilliard partial differential equations (Barrett & Blowey, 2002; Feng & Prohl, 2004). We have the following direct consequence to Theorem 2.1.

THEOREM 2.4 Suppose that an algebraically stable Runge–Kutta method is applied to a gradient system with potential (2.7), and assume that the Hessian matrix $\nabla^2 V(y)$ is Lipschitz continuous in the ball of radius δ centred at y_0 with Lipschitz constant L. Then, independently of the norm of A, if the matrix

$$B \otimes \Gamma + \frac{h}{2}M \otimes \nabla^2 V(y_0) \tag{2.8}$$

is positive definite, and if $hL\delta$ is bounded by a sufficiently small constant, we have

$$U(y_1) \leqslant U(y_0).$$

Proof. This result follows from the fact that $M \otimes A$ is positive semidefinite. $\square$

Note that, for Gauss methods, we have $M = 0$, and hence the matrix (2.8) reduces to $B \otimes \Gamma$, which is always positive definite.

For a quadratic potential (i.e., $V(y) = 0$ in (2.7)) and for $G(y) = I$ the numerical solution of an implicit Runge–Kutta method is given by

$$y_1 = R(-hA)y_0,$$

where $R(z)$ is the stability function; see Hairer & Wanner (1996, Section IV.3). Every algebraically stable Runge–Kutta method is A-stable, so that $\|R(-hA)\| = \max_i |R(-h\lambda_i)| \leqslant 1$, where λ_i are the eigenvalues of the matrix A. Radau IIA methods have $R(\infty) = 0$ and hence eigencomponents corresponding to very large eigenvalues of A are strongly damped. This is not the case for Gauss methods, for which $|R(\infty)| = 1$. The importance of the *damping property* $|R(\infty)| < 1$ for the approximation properties of a Runge–Kutta method applied to semilinear parabolic equations is well understood; see Lubich & Ostermann (1993, 1996). For the role of the condition $R(\infty) = 0$ in the approximation of stiff differential equations of singular perturbation type see Hairer & Wanner (1996, Chapter VI).

3. Unconditionally energy-diminishing methods

3.1 *Discrete-gradient methods*

Discrete-gradient methods, applied to (1.1), are of the form

$$\bar{G}(y_1, y_0)(y_1 - y_0) = -h\bar{\nabla}U(y_1, y_0), \tag{3.1}$$

where $\bar{G}(y_1, y_0)$ is a symmetric positive-definite matrix with $\bar{G}(y, y) = G(y)$, and the discrete gradient satisfies

$$\bar{\nabla}U(y, z)^{\mathrm{T}}(y - z) = U(y) - U(z), \quad \bar{\nabla}U(y, y) = \nabla U(y). \tag{3.2}$$

Well-known examples are the midpoint discrete gradient (Gonzalez, 1996)

$$\bar{\nabla} U(y, z) = \nabla U \left(\frac{y + z}{2} \right) + \frac{U(y) - U(z) - \nabla U((y + z)/2)^{\mathrm{T}}(y - z)}{\|y - z\|^2} (y - z),$$

and the average vector field (McLachlan *et al.*, 1999)

$$\bar{\nabla} U(y, z) = \int_0^1 \nabla U(z + \tau(y - z)) \, \mathrm{d}\tau.$$

With $\bar{G}(y, z) = G((y + z)/2)$ these methods are of order 2. Higher-order discrete-gradient methods are derived in Dahlby *et al.* (2011) and Norton & Quispel (2012).

As was noted in McLachlan *et al.* (1999), multiplying (3.1) by $(y_1 - y_0)^{\mathrm{T}}$ and using (3.2) immediately yields $U(y_1) \leqslant U(y_0)$. This does not require any restriction on the step size other than what is needed for the existence of the numerical solution.

In the case of a quadratic potential, both discrete-gradient methods given above reduce to the implicit midpoint rule. Therefore, these methods show no damping for very stiff gradient systems.

3.2 *Averaged vector field collocation methods*

The following class of methods was introduced in Hairer (2010) and Cohen & Hairer (2011), where they are called energy-preserving collocation methods in view of their energy conservation for Hamiltonian systems. For given distinct nodes $c_1, \ldots, c_s$ (usually $0 \leqslant c_i \leqslant 1$) we use the notation

$$\ell_i(\tau) = \prod_{j=1, j \neq i}^{s} \frac{\tau - c_j}{c_i - c_j}, \quad b_i = \int_0^1 \ell_i(\tau) \, \mathrm{d}\tau$$

for the Lagrange polynomials. Assuming $b_i \neq 0$ for all i, we consider the polynomial $u(t)$ of degree s satisfying $u(t_0) = y_0$ and, for $i = 1, \ldots, s$,

$$G(u(t_0 + c_i h))\dot{u}(t_0 + c_i h) = -\frac{1}{b_i} \int_0^1 \ell_i(\tau) \nabla U(u(t_0 + \tau h)) \, \mathrm{d}\tau, \qquad (3.3)$$

and define $y_1 = u(t_0 + h)$. This method is of order $2s$ if the nodes are those of the Gaussian quadrature.

THEOREM 3.1 We assume that the nodes c_i are such that $b_i > 0$ for all i. When applied to the gradient system (2.1), the averaged vector field collocation method then reduces the energy so that $U(y_1) \leqslant U(y_0)$.

Proof. From the fundamental theorem of calculus we have

$$U(u(t_0 + h)) - U(u(t_0)) = h \int_0^1 \dot{u}(t_0 + \tau h)^{\mathrm{T}} \nabla U(u(t_0 + \tau h)) \, \mathrm{d}\tau.$$

Substituting

$$\dot{u}(t_0 + \tau h) = \sum_{i=1}^{s} \ell_i(\tau)\dot{u}(t_0 + c_i h)$$

in this relation and using (3.3) yields

$$U(y_1) - U(y_0) = -h \sum_{i=1}^{s} b_i \dot{Y}_i^{\mathrm{T}} G(Y_i) \dot{Y}_i \leqslant 0,$$

where $Y_i = u(t_0 + c_i h)$ and $\dot{Y}_i = \dot{u}(t_0 + c_i h)$.   □

In the case of a quadratic potential, the averaged vector field collocation method reduces to the Gauss collocation method. Therefore, also these methods show no damping for very stiff gradient systems.

4. Constrained gradient systems

On a manifold $\mathcal{M}$ given by constraints $c(y) = 0$ with a continuously differentiable $c : \mathbf{R}^n \to \mathbf{R}^m$, the gradient system becomes

$$\begin{aligned} G(y)\dot{y} &= -\nabla U(y) - \nabla c(y)\lambda, \\ 0 &= c(y), \end{aligned} \tag{4.1}$$

where $\lambda \in \mathbf{R}^m$ is the vector of Lagrange multipliers and $\nabla c(y)$ denotes the transpose of the Jacobian matrix of $c(y)$.

4.1 Runge–Kutta methods

A Runge–Kutta method applied to (4.1) is given by (2.4), where

$$G(Y_i)\dot{Y}_i = -\nabla U(Y_i) - \nabla c(Y_i)\Lambda_i, \tag{4.2}$$

and Λ_i is defined either by

$$0 = c(Y_i) \quad \text{(index-2 formulation) or}$$

$$0 = \nabla c(Y_i)^{\mathrm{T}} \dot{Y}_i \quad \text{(index-1 formulation).}$$

In general, the approximation y_1 does not lie on the manifold, and we therefore add an energy-preserving projection after every step:

$$\begin{aligned} \hat{y}_1 &= y_1 - \nabla c(y_1)\lambda_1 - \nabla U(y_1)\mu, \\ 0 &= c(\hat{y}_1), \\ 0 &= U(\hat{y}_1) - U(y_1). \end{aligned}$$

The next step would then start with $\hat{y}_1$, which is an approximation of the full order of the Runge–Kutta method (Hairer & Wanner, 1996, Section VI.1). Theorem 2.1 extends directly to the index-1 formulation.

THEOREM 4.1 We consider the constrained gradient system (4.1) and apply an implicit Runge–Kutta method in its index-1 formulation. Then, the statement of Theorem 2.1 remains valid without any change.

Proof. The proof is nearly the same as that of Theorem 2.1. Instead of using (2.5) we now use (4.2) and we note that $\dot{Y}_i^{\mathrm{T}} \nabla c(Y_i) = 0$ for the index-1 formulation of the method.   □

4.2 *Discrete-gradient methods*

A natural extension of discrete-gradient methods to systems of the form (4.1) is

$$\bar{G}(y_1, y_0)(y_1 - y_0) = -h\bar{\nabla}U(y_1, y_0) - h\bar{\nabla}c(y_1, y_0)\lambda,$$
$$0 = c(y_1). \tag{4.3}$$

Starting from a consistent initial value, $c(y_0) = 0$, and using the defining property of the discrete gradient, we note that

$$\bar{\nabla}c(y_1, y_0)^{\mathrm{T}}(y_1 - y_0) = c(y_1) - c(y_0) = 0.$$

Therefore, multiplying the first equation of (4.3) by $(y_1 - y_0)^{\mathrm{T}}$ yields again $U(y_1) \leqslant U(y_0)$.

An extension of discrete-gradient methods to problems with constraints was proposed in Gonzalez (1999) within a Hamiltonian setting, and in Gonzalez (2000) for models in nonlinear elasticity.

4.3 *Averaged vector field collocation methods*

The extension of the averaged vector field collocation method to the constrained system (4.1) replaces (3.3) by

$$G(u(t_0 + c_ih))\dot{u}(t_0 + c_ih) = -\frac{1}{b_i}\int_0^1 \ell_i(\tau)\nabla U(u(t_0 + \tau h))\,\mathrm{d}\tau - \frac{1}{b_i}\int_0^1 \ell_i(\tau)\nabla c(u(t_0 + \tau h))\,\mathrm{d}\tau\,\Lambda_i,$$

where Λ_i is determined such that

$$\int_0^1 \ell_i(\tau)\nabla c(u(t_0 + \tau h))^{\mathrm{T}}\,\mathrm{d}\tau\,\dot{u}(t_0 + c_ih) = 0. \tag{4.4}$$

Summing this relation from $i = 1$ to s yields

$$0 = h\int_0^1 \nabla c(u(t_0 + \tau h))^{\mathrm{T}}\dot{u}(t_0 + \tau h)\,\mathrm{d}\tau = c(y_1) - c(y_0),$$

so that y_1 automatically satisfies the constraint. When the c_i are chosen as the Gaussian nodes, the method has order at least $s + 1$. This follows from approximating the integral in (4.4) by the Gaussian quadrature, which gives an error of size $\mathcal{O}(h^{s+1})$. If the constraints $c(y) = 0$ are linear or quadratic, condition (4.4) is equivalent to $\nabla c(u(t_0 + c_ih))^{\mathrm{T}}\dot{u}(t_0 + c_ih) = 0$ and the analysis of Cohen & Hairer (2011) can be applied to prove that the order equals $2s$.

The same proof as for Theorem 3.1 shows that $U(y_1) \leqslant U(y_0)$. As for the unconstrained case, this does not require any restriction on the step size other than that needed for the existence of the numerical solution.

Deutsche Forschungsgemeinschaft, SFB/TR 71.

References

BAO, W. & DU, Q. (2004) Computing the ground state solution of Bose–Einstein condensates by a normalized gradient flow. *SIAM J. Sci. Comput.*, **25**, 1674–1697.

BARRETT, J. & BLOWEY, J. (2002) Finite element approximation of an Allen–Cahn/Cahn–Hilliard system. *IMA J. Numer. Anal.*, **22**, 11–71.

BOBENKO, A. & SCHRÖDER, P. (2005) Discrete Willmore flow. *Eurographics Symposium on Geometry Processing* (H. P. M. Desbrun ed.). Aire-la-Ville, Switzerland: Eurographics Association, pp. 101–110.

BURRAGE, K. & BUTCHER, J. C. (1979) Stability criteria for implicit Runge–Kutta methods. *SIAM J. Numer. Anal.*, **16**, 46–57.

CHAFEE, N. & INFANTE, E. F. (1974) A bifurcation problem for a nonlinear partial differential equation of parabolic type. *Appl. Anal.*, **4**, 17–37.

CHEN, L. (2002) Phase-field models for microstructure evolution. *Annu. Rev. Mater. Res.*, **32**, 113–140.

COHEN, D. & HAIRER, E. (2011) Linear energy-preserving integrators for Poisson systems. *BIT*, **51**, 91–101.

CROUZEIX, M. (1979) Sur la B-stabilité des méthodes de Runge–Kutta. *Numer. Math.*, **32**, 75–82.

DAHLBY, M., OWREN, B. & YAGUCHI, T. (2011) Preserving multiple first integrals by discrete gradients. *J. Phys. A*, **44**, 305205, 14.

DROSKE, M. & RUMPF, M. (2004) A level set formulation for Willmore flow. *Interfaces Free Bound.*, **6**, 361–378.

ELLIOTT, C. M. & STUART, A. M. (1993) The global dynamics of discrete semilinear parabolic equations. *SIAM J. Numer. Anal.*, **30**, 1622–1663.

ELLIOTT, C. M. & ZHENG, S. (1986) On the Cahn–Hilliard equation. *Arch. Ration. Mech. Anal.*, **96**, 339–357.

FENG, X. & PROHL, A. (2004) Error analysis of a mixed finite element method for the Cahn–Hilliard equation. *Numer. Math.*, **99**, 47–84.

GONZALEZ, O. (1996) Time integration and discrete Hamiltonian systems. *J. Nonlinear Sci.*, **6**, 449–467.

GONZALEZ, O. (1999) Mechanical systems subject to holonomic constraints: differential-algebraic formulations and conservative integration. *Phys. D*, **132**, 165–174.

GONZALEZ, O. (2000) Exact energy and momentum conserving algorithms for general models in nonlinear elasticity. *Comput. Methods Appl. Mech. Engrg.*, **190**, 1763–1783.

HAIRER, E. (2010) Energy-preserving variant of collocation methods. *J. Numer. Anal. Ind. Appl. Math.*, **5**, 73–84.

HAIRER, E. & WANNER, G. (1996) *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*, 2nd edn. Springer Series in Computational Mathematics, vol. 14. Berlin: Springer.

HALE, J. K. (1988) *Asymptotic Behavior of Dissipative Systems*. Mathematical Surveys and Monographs, vol. 25. Providence, RI: American Mathematical Society, pp. x+198.

HIRSCH, M. W. & SMALE, S. (1974) *Differential Equations, Dynamical Systems, and Linear Algebra*. Pure and Applied Mathematics, vol. 60. New York: Academic Press, pp. xi+358.

HUMPHRIES, A. R. & STUART, A. M. (1994) Runge–Kutta methods for dissipative and gradient dynamical systems. *SIAM J. Numer. Anal.*, **31**, 1452–1485.

LUBICH, C. & OSTERMANN, A. (1993) Runge–Kutta methods for parabolic equations and convolution quadrature. *Math. Comp.*, **60**, 105–131.

LUBICH, C. & OSTERMANN, A. (1996) Runge–Kutta time discretization of reaction-diffusion and Navier–Stokes equations: nonsmooth-data error estimates and applications to long-time behaviour. *Appl. Numer. Math.*, **22**, 279–292. Special issue celebrating the centenary of Runge–Kutta methods.

MCLACHLAN, R. I., QUISPEL, G. R. W. & ROBIDOUX, N. (1999) Geometric integration using discrete gradients. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **357**, 1021–1045.

MICHAILOVICH, O., RATHI, Y. & TANNENBAUM, A. (2007) Image segmentation using active contours driven by the Bhattacharyya gradient flow. *IEEE Trans. Image Process.*, **16**, 2787–2801.

NORTON, R. & QUISPEL, G. (2012) Discrete gradient methods for preserving a first integral of an ordinary differential equation. *Discrete Contin. Dyn. Syst.* **34**, 1147–1170.

OTTO, F. (1998) Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory. *Arch. Ration. Mech. Anal.*, **141**, 63–103.

OTTO, F. (2001) The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, **26**, 101–174.

PENZLER, P., RUMPF, M. & WIRTH, B. (2012) A phase-field model for compliance shape optimization in nonlinear elasticity. *ESAIM Control Optim. Calc. Var.*, **18**, 229–258.

STRZODKA, R., DROSKE, M. & RUMPF, M. (2004) Image registration by a regularized gradient flow. A streaming implementation in DX9 graphics hardware. *Computing*, **73**, 373–389.

STUART, A. M. & HUMPHRIES, A. R. (1996) *Dynamical Systems and Numerical Analysis*. Cambridge Monographs on Applied and Computational Mathematics, vol. 2. Cambridge: Cambridge University Press, pp. xxii+685.