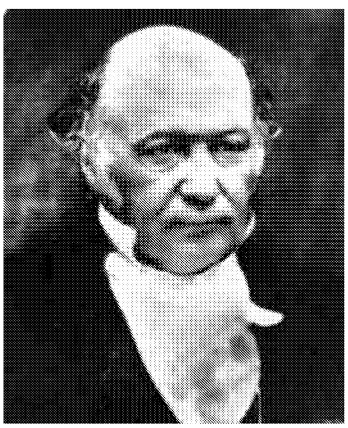
Chapter IV

Symplectic Integration

Hamiltonian systems form the most important class of ordinary differential equations in the context of 'Numerical Geometric Integration' (see the examples of Chapter I). In this chapter we start by discussing the origin of such systems and by studying their geometric properties such as symplecticity. We then turn our attention to numerical integrators which preserve the symplectic structure.

IV.1 Hamiltonian Systems



Sir William Rowan Hamilton[1]

Consider a mechanical system with $q = (q_1, \ldots, q_d)^T$ as *generalized coordinates*, and denote by $T = T(q, \dot q) = \frac{1}{2}\dot q^T M(q)\dot q$ its kinetic energy ($M(q)$ is assumed to be symmetric and positive definite) and by $U = U(q)$ its potential energy. The movement of such a system is described by the solution of the variational problem

$$\int L\Big(q(t), \dot q(t)\Big)\,dt \to \min, \tag{1.1}$$

where $L = T - U$ is the Lagrangian of the system. From the fundamental work of Euler (1744) and Lagrange (1755) at the age of 19 (see [HNW93, p. 8] for some historical remarks) we know that the solutions of (1.1) are determined by the second order differential equation

$$\frac{\partial L}{\partial q} - \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot q}\Big) = 0, \tag{1.2}$$

which constitute the so-called *Euler-Lagrange equations.*

Example 1.1 (Pendulum) We consider the mathematical pendulum (see Sect. I.1) and we take the angle α as generalized coordinate. The kinetic and potential energies are given by $T = m(\dot x^2 + \dot y^2)/2 = m\ell^2\dot\alpha^2/2$ and $U = mgy = -mg\ell\cos\alpha$, respectively, so that the Euler-Lagrange equations become $-mg\ell\sin\alpha - m\ell^2\ddot\alpha = 0$ or equivalently $\ddot\alpha + \frac{g}{\ell}\sin\alpha = 0$.

[1] William Rowan Hamilton, born: 4 August 1805 in Dublin (Ireland), died: 2 September 1865. Picture, copied from `http://www-history.mcs.st-and.ac.uk/history/Mathematicians/Hamilton.html`, where one can also find a short biography.

With the aim of simplifying the structure of the Euler-Lagrange equations and of making them more symmetric, Hamilton [Ha1834] had the idea

- of introducing the new variables

$$p_k = \frac{\partial L}{\partial \dot q_k} \qquad \text{for} \quad k = 1, \ldots, d, \tag{1.3}$$

  the so-called conjugated *generalized momenta*. Observe that for a fixed q we have $p = M(q)\dot q$, so that there is a bijection between $p = (p_1, \ldots, p_d)^T$ and $\dot q$, if $M(q)$ is invertible;

- of considering the *Hamiltonian*

$$H := p^T \dot q - L(q, \dot q) \tag{1.4}$$

  as a function of p and q, i.e., $H(p, q)$.

Theorem 1.2 *Let $M(q)$ and $U(q)$ be continuously differentiable functions. Then, the Euler-Lagrange equations (1.2) are equivalent to the Hamiltonian system*

$$\dot p_k = -\frac{\partial H}{\partial q_k}(p, q), \qquad \dot q_k = \frac{\partial H}{\partial p_k}(p, q), \qquad k = 1, \ldots, d. \tag{1.5}$$

Proof. The definitions (1.3) and (1.4) for the generalized momenta p and for the Hamiltonian function H imply that

$$\begin{aligned} \frac{\partial H}{\partial p} &= \dot q^T + p^T \frac{\partial \dot q}{\partial p} - \frac{\partial L}{\partial \dot q}\frac{\partial \dot q}{\partial p} = \dot q^T, \\ \frac{\partial H}{\partial q} &= p^T \frac{\partial \dot q}{\partial q} - \frac{\partial L}{\partial q} - \frac{\partial L}{\partial \dot q}\frac{\partial \dot q}{\partial q} = -\frac{\partial L}{\partial q}. \end{aligned}$$

The Euler-Lagrange equations (1.2) are therefore equivalent to (1.5). □

If we replace the variable $\dot q$ by $M(q)^{-1}p$ in the definition (1.4) of $H(p, q)$, we obtain

$$\begin{aligned} H(p, q) &= p^T M(q)^{-1} p - L(q, M(q)^{-1}p) = p^T M(q)^{-1} p - \frac{1}{2} p^T M(q)^{-1} p + U(q) \\ &= \frac{1}{2} p^T M(q)^{-1} p + U(q) \end{aligned}$$

and the Hamiltonian is $H = T + U$, which is the *total energy* of the mechanical system.

In the following we consider Hamiltonian systems (1.5), where the Hamiltonian function $H(p, q)$ is arbitrary (not necessarily related to a mechanical problem).

IV.2 Symplectic Transformations

We have already seen in Example 1.2 of Sect. III.1 that the Hamiltonian function $H(p, q)$ is a first integral of the system (1.5). In this section we shall study another important property of Hamiltonian systems – the *symplecticity* of its flow.

For two vectors

$$\xi = \begin{pmatrix} \xi^p \\ \xi^q \end{pmatrix}, \qquad \eta = \begin{pmatrix} \eta^p \\ \eta^q \end{pmatrix}$$

in the (p,q) space (ξ^p, ξ^q, η^p, η^q are in $\mathbb{R}^d$), we consider the parallelogram

$$L = \{t\xi + s\eta \mid 0 \le t \le 1,\ 0 \le s \le 1\}.$$

We then consider its projection $L_i = \{(t\xi_i^p + s\eta_i^p, t\xi_i^q + s\eta_i^q)^T \mid 0 \le t \le 1,\ 0 \le s \le 1\}$ onto the (p_i, q_i) coordinate plane, and we let

$$(dp_i \wedge dq_i)(\xi, \eta) := \text{or.area}\,(L_i) = \det \begin{pmatrix} \xi_i^p & \eta_i^p \\ \xi_i^q & \eta_i^q \end{pmatrix} = \xi_i^p \eta_i^q - \xi_i^q \eta_i^p \tag{2.1}$$

be the oriented area of this projection. Here, dp_i and dq_i select the coordinates of the vectors ξ and η. In an analogous way, we can also define $dp_i \wedge dq_j$, $dp_i \wedge dp_j$ or $dq_i \wedge dq_j$. This *exterior product* is a bilinear map acting on vectors of $\mathbb{R}^{2d}$. It satisfies Grassmann's rules for exterior multiplication

$$dp_i \wedge dp_j = -dp_j \wedge dp_i, \qquad dp_i \wedge dp_i = 0. \tag{2.2}$$

We further consider the differential 2-form

$$\omega^2 := \sum_{i=1}^{d} dp_i \wedge dq_i, \tag{2.3}$$

which will play a central role for Hamiltonian systems. This is again a bilinear mapping. In matrix notation it is given by

$$\omega^2(\xi, \eta) = \xi^T J \eta \qquad \text{with} \qquad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{2.4}$$

where I is the identity matrix of dimension d.

Definition 2.1 A linear mapping $A : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is called *symplectic* (a name suggested by H. Weyl, 1939), if

$$\omega^2(A\xi, A\eta) = \omega^2(\xi, \eta) \qquad \text{for all} \;\; \xi, \eta \in \mathbb{R}^{2d},$$

or, equivalently, if $A^T J A = J$.

In the case $d = 1$, the expression $\omega^2(\xi, \eta) = (dp_1 \wedge dq_1)(\xi, \eta)$ represents the area of the parallelogram spanned by the 2-dimensional vectors ξ and η. Symplecticity of a linear mapping A is therefore equivalent to area preservation. In the general case ($d > 1$), symplecticity means that the sum over the oriented areas of the projections L_i is the same as that for the transformed parallelograms $A(L)_i$.

We now turn our attention to nonlinear mappings. Differentiable functions can locally be approximated by linear mappings. This justifies the following definition.

Definition 2.2 A differentiable function $g : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is called *symplectic* at $(p,q) \in \mathbb{R}^{2d}$, if the Jacobian matrix $g'(p,q)$ is symplectic, i.e., if

$$\omega^2(g'(p,q)\xi, g'(p,q)\eta) = \omega^2(\xi, \eta) \qquad \text{or} \qquad g'(p,q)^T J\, g'(p,q) = J.$$

We next give a geometric interpretation of symplecticity for nonlinear mappings. Consider a 2-dimensional manifold M in the $2d$-dimensional phase space, and suppose that it is given as the image $M = \varphi(K)$ of a compact set $K \subset I\!R^2$, where $\varphi(s,t)$ is a continuously differentiable function. The manifold M can then be considered as the limit of a union of small parallelograms spanned by the vectors

$$\frac{\partial \varphi}{\partial s}(s,t)\, ds \qquad \text{and} \qquad \frac{\partial \varphi}{\partial t}(s,t)\, dt.$$

For one such parallelogram we consider (as above) the sum over the oriented areas of its projections onto the (p_i, q_i) plane. We then sum over all parallelograms of the manifold. In the limit this gives the expression

$$\Omega(M) = \iint_K \omega^2\Big(\frac{\partial \varphi}{\partial s}(s,t), \frac{\partial \varphi}{\partial t}(s,t)\Big)\, ds\, dt. \tag{2.5}$$

Lemma 2.3 *If the mapping $g : I\!R^{2d} \to I\!R^{2d}$ is symplectic for all $(p,q) \in I\!R^{2d}$, then it preserves the expression $\Omega(M)$, i.e.,*

$$\Omega(g(M)) = \Omega(M)$$

holds for all 2-dimensional manifolds M that can be represented as the image of a continuously differentiable function φ.

Proof. The manifold $g(M)$ is parametrized by $g \circ \varphi$. The transformation formula for double integrals therefore implies

$$\Omega(g(M)) = \iint_K \omega^2\Big(\frac{\partial (g\circ\varphi)}{\partial s}(s,t), \frac{\partial (g\circ\varphi)}{\partial t}(s,t)\Big)\, ds\, dt = \Omega(M),$$

because $(g\circ\varphi)'(s,t) = g'(\varphi(s,t))\varphi'(s,t)$ and g is a symplectic transformation. □

For $d = 1$, M is already a subset of $I\!R^2$ and we can take the identity map for φ, so that $M = K$. In this case, $\Omega(M) = \iint_M ds\, dt$ represents the area of M. Hence, Lemma 2.3 states that (for $d = 1$) symplectic mappings are *area preserving*.

We are now able to prove the main result of this section. We use the notation $y = (p,q)$, and we write the Hamiltonian system (1.5) in the form

$$y' = J^{-1}\nabla H(y), \tag{2.6}$$

where J is the matrix of (2.4) and $\nabla H(y) = \operatorname{grad} H(y)^T$.

Recall that the flow $\varphi_t : I\!R^{2d} \to I\!R^{2d}$ of a Hamiltonian system is the mapping that advances the solution by time t, i.e., $\varphi_t(p_0, q_0) = (p(t,p_0,q_0), q(t,p_0,q_0))$, where $p(t,p_0,q_0)$, $q(t,p_0,q_0)$ is the solution of the system corresponding to initial values $p(0) = p_0$, $q(0) = q_0$.

Theorem 2.4 (Poincaré [Po1899]) *Let $H(p,q)$ be a twice continuously differentiable function. Then, the flow φ_t is everywhere a symplectic transformation.*

Proof. The derivative $\partial\varphi_t/\partial y_0$ (with $y_0 = (p_0, q_0)$) is a solution of the variational equation which, for the Hamiltonian system (2.6), is given by $\Psi' = J^{-1}H''(\varphi_t(y_0))\Psi$, where $H''(p,q)$ is the Hessian matrix of $H(p,q)$ ($H''(p,q)$ is symmetric). We therefore obtain

$$\begin{aligned}\frac{d}{dt}\Big(\Big(\frac{\partial\varphi_t}{\partial y_0}\Big)^T J \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)\Big) &= \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)'^T J \Big(\frac{\partial\varphi_t}{\partial y_0}\Big) + \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)^T J \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)' \\ &= \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)^T H''(\varphi_t(y_0)) J^{-T} J \Big(\frac{\partial\varphi_t}{\partial y_0}\Big) + \Big(\frac{\partial\varphi_t}{\partial y_0}\Big)^T H''(\varphi_t(y_0)) \Big(\frac{\partial\varphi_t}{\partial y_0}\Big) = 0,\end{aligned}$$

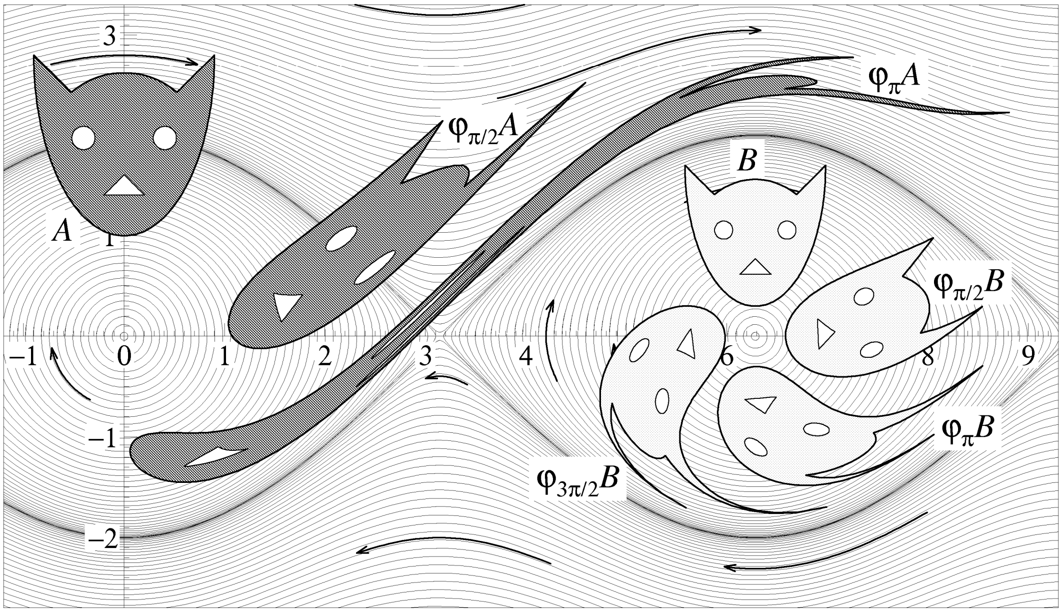


FIG. 2.1: Area preservation of the flow of Hamiltonian systems

because $J^T = -J$ and $J^{-T}J = -I$. Since the relation

$$\Big(\frac{\partial \varphi_t}{\partial y_0}\Big)^T J \Big(\frac{\partial \varphi_t}{\partial y_0}\Big) = J \tag{2.7}$$

is satisfied for $t = 0$ (φ_0 is the identity map), it is satisfied for all t and all (p_0, q_0), as long as the solution remains in the domain of definition of H. □

Example 2.5 Consider the pendulum problem (Example 1.1) with the normalization $m = \ell = g = 1$. We then have $q = \alpha$, $p = \dot\alpha$, and the Hamiltonian is given by

$$H(p, q) = p^2/2 - \cos q.$$

Fig. 2.1 shows level curves of this function, and it also illustrates the area preservation of the flow φ_t. Indeed, by Theorem 2.4 and Lemma 2.3 the area of A and $\varphi_t(A)$ as well as those of B and $\varphi_t(B)$ are the same, although their appearance is completely different.

We next show that symplecticity of a flow is a characteristic property for Hamiltonian systems.

Theorem 2.6 *Let $f : I\!\!R^{2d} \to I\!\!R^{2d}$ be continuously differentiable. Then, $y' = f(y)$ is a Hamiltonian system, if and only if its flow $\varphi_t(y)$ is symplectic for all $y \in I\!\!R^{2d}$ and for all sufficiently small t.*

Proof. The necessity follows from Theorem 2.4. We therefore assume that the flow φ_t is symplectic, and we have to prove the existence of a function $H(y)$ such that $f(y) = J^{-1}\nabla H(y)$. Differentiating (2.7) and using the fact that $\partial\varphi_t/\partial y_0$ is solution of the variational equation $\Psi' = f'(\varphi_t(y_0))\Psi$, we obtain

$$\frac{d}{dt}\Big(\Big(\frac{\partial \varphi_t}{\partial y_0}\Big)^T J \Big(\frac{\partial \varphi_t}{\partial y_0}\Big)\Big) = \Big(\frac{\partial \varphi_t}{\partial y_0}\Big)\Big(f'(\varphi_t(y_0))^T J + J f'(\varphi_t(y_0))\Big)\Big(\frac{\partial \varphi_t}{\partial y_0}\Big) = 0.$$

Putting $t = 0$, it follows from $J = -J^T$ that $Jf'(y_0)$ is a symmetric matrix for all y_0. Lemma 2.7 below shows that $Jf(y)$ can be written as the gradient of a function $H(y)$. □

Lemma 2.7 *Let $f : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ be continuously differentiable, and assume that the Jacobian $f'(y)$ is symmetric for all y. Then, there exists a function $H : \mathbb{R}^{2d} \to \mathbb{R}$ such that $f(y) = \nabla H(y)$, i.e., the vector field $f(y)$ possesses a potential $H(y)$.*

Proof. Since f is defined on the whole space, we can define

$$H(y) = \int_0^1 y^T f(ty)\, dt + Const.$$

Differentiation with respect to y_k, and using the symmetry assumption $\partial f_i/\partial y_k = \partial f_k/\partial y_i$ yields

$$\frac{\partial H}{\partial y_k}(y) = \int_0^1 \Big(f_k(ty) + y^T \frac{\partial f}{\partial y_k}(ty)t \Big)\, dt = \int_0^1 \frac{d}{dt}\Big(t f_k(ty) \Big)\, dt = f_k(y),$$

which proves the statement. □

Lemma 2.7 and Theorem 2.6 remain valid for functions $f : U \to \mathbb{R}^{2d}$ with $U \subset \mathbb{R}^{2d}$, if U is star-shaped or, more generally, if U is a simply connected domain. A counter-example, which shows that the statement of Theorem 2.6 is not true for general U, is given in Exercise 8.

IV.3 Symplectic Runge-Kutta Methods

Since the property of symplecticity is characteristic of Hamiltonian systems (Theorem 2.6), it is natural to search for numerical methods that share this property. After some pioneering work of de Vogelaere [Vo56], Ruth [Ru83] and Feng Kang [FeK85], the systematic study of symplectic methods started around 1988. A characterization of symplectic Runge-Kutta methods (Theorem 3.4 below) has been found independently by Lasagni [La88], Sanz-Serna [SS88] and Suris [Su89].

Definition 3.1 A numerical one-step method $y_1 = \Phi_h(y_0)$ is called *symplectic* if, when applied to a smooth Hamiltonian system, the mapping Φ_h is everywhere a symplectic transformation.

Example 3.2 We consider the harmonic oscillator

$$H(p, q) = (p^2 + q^2)/2,$$

so that the Hamiltonian system becomes $\dot p = -q$, $\dot q = p$. We apply six different numerical methods to this problem: the explicit Euler method (I.1.4), the symplectic Euler method (I.1.8), and the implicit Euler method (I.1.5), as well as the second order method of Runge

$$k_1 = f(y_0), \qquad k_2 = f(y_0 + hk_1/2), \qquad y_1 = y_0 + hk_2, \tag{3.1}$$

the Verlet scheme (I.3.6), and the implicit midpoint rule (I.1.6). For a set of initial values (p_0, q_0) (the dark set in Fig. 3.1) we compute 16 steps with step size $h = \pi/8$ for the first order methods, and 8 steps with $h = \pi/4$ for the second order methods. Since the exact solution is periodic with period 2π, the numerical result of the last step approximates the set of initial values. One clearly observes that the explicit Euler, the implicit Euler and the second order explicit method of Runge are not symplectic (not area preserving).

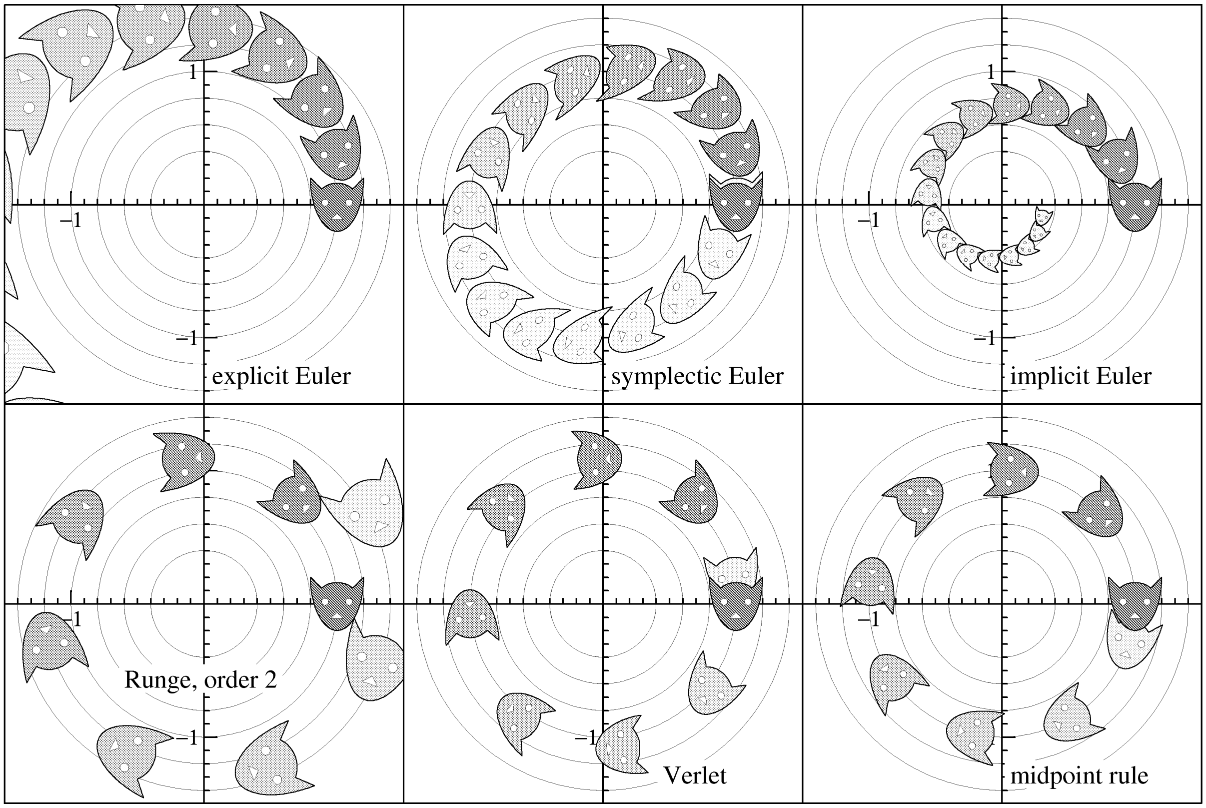


FIG. 3.1: Area preservation of numerical methods for the harmonic oscillator

The other methods are symplectic (see Theorem 3.4 and Theorem 3.5), although the approximation at the end of the integration may be quite different from the initial set. Only the implicit midpoint rule preserves exactly the quadratic invariant $(p^2 + q^2)/2$.

For the study of symplecticity of numerical integrators we follow the approach of [BoS94], which is based on the following lemma.

Lemma 3.3 *For Runge-Kutta methods and for partitioned Runge-Kutta methods the following diagram commutes:*

$$\begin{array}{ccc} y' = f(y), \quad y(0) = y_0 & \longrightarrow & \begin{array}{c} y' = f(y), \quad y(0) = y_0 \\ \Psi' = f'(y)\Psi, \quad \Psi(0) = I \end{array} \\ \Big\downarrow \textit{method} & & \Big\downarrow \textit{method} \\ \{y_n\} & \longrightarrow & \{y_n, \Psi_n\} \end{array}$$

(horizontal fleches mean 'differentiation'). Therefore, the numerical result y_n, Ψ_n, obtained from applying the method to the problem augmented by its variational equation, is equal to the numerical solution for $y' = f(y)$ augmented by its derivative $\Psi_n = \partial y_n / \partial y_0$.

Proof. This result is very important when the derivative of the numerical solution with respect to the initial value is needed. It is proved by implicit differentiation. Let us illustrate this for Euler's method

$$y_{n+1} = y_n + hf(y_n).$$

We consider y_n and y_{n+1} as functions of y_0, and we differentiate the equation, defining the numerical method, with respect to y_0. For Euler's method this gives

$$\frac{\partial y_{n+1}}{\partial y_0} = \frac{\partial y_n}{\partial y_0} + hf'(y_n)\,\frac{\partial y_n}{\partial y_0},$$

which is exactly the same relation that we get, when we apply the method to the variational equation. Since $\partial y_0/\partial y_0 = I$, we have $\partial y_n/\partial y_0 = \Psi_n$ for all n. □

The main observation is now that the symplecticity condition (2.7) is a quadratic first integral of the variational equation. The following characterization of symplectic methods is therefore not surprising.

Theorem 3.4 *If the coefficients of a Runge-Kutta method satisfy*

$$b_i a_{ij} + b_j a_{ji} = b_i b_j \qquad \textit{for all} \;\; i,j = 1,\ldots,s, \tag{3.2}$$

then it is symplectic.[2]

Proof. We write the Hamiltonian system together with its variational equation as

$$y' = J^{-1}\nabla H(y), \qquad \Psi' = J^{-1}H''(y)\Psi. \tag{3.3}$$

It follows from

$$(J^{-1}H''(y)\Psi)^T J\Psi + \Psi^T J(J^{-1}H''(y)\Psi) = 0$$

(see also the proof of Theorem 2.4) that $\Psi^T J\Psi$ is a first integral of the augmented system (3.3). Since this first integral is quadratic, it is exactly preserved by Runge-Kutta methods satisfying (3.2) (see Theorem III.3.2). Hence, $\Psi_1^T J\Psi_1 = \Psi_0^T J\Psi_0$ holds. The symplecticity of the Runge-Kutta method Φ_h then follows from Lemma 3.3, because for $\Psi_0 = I$ we have $\Psi_1 = \Phi_h'(y_0)$. □

Theorem 3.5 *If the coefficients of a partitioned Runge-Kutta method (II.4.2) satisfy*

$$b_i\widehat{a}_{ij} + \widehat{b}_j a_{ji} \;=\; b_i\widehat{b}_j \qquad \textit{for} \;\; i,j = 1,\ldots,s, \tag{3.4}$$

$$b_i \;=\; \widehat{b}_i \qquad \textit{for} \;\; i = 1,\ldots,s, \tag{3.5}$$

then it is symplectic.

If the Hamiltonian is of the form $H(p,q) = T(p) + U(q)$, i.e., it is separable, then the condition (3.4) alone implies the symplecticity of the numerical flow.

Proof. We write the solution Ψ of the variational equation as

$$\Psi = \begin{pmatrix} \Psi^p \\ \Psi^q \end{pmatrix}.$$

Then, the Hamiltonian system together with its variational equation (3.3) is a partitioned system with variables (p, Ψ^p) and (q, Ψ^q). Every component of

$$\Psi^T J\Psi = (\Psi^p)^T\Psi^q - (\Psi^q)^T\Psi^p$$

is of the form (III.3.5), so that Theorem 3.3 can be applied. □

[2]For irreducible Runge-Kutta methods the condition (3.2) is also necessary for symplecticity.

IV.4 Symplecticity for Linear Problems

For quadratic Hamiltonians $H(y) = \frac{1}{2} y^T C y$ (where C is a symmetric real matrix) the corresponding system (2.6) is linear,

$$y' = J^{-1} C y. \tag{4.1}$$

Lemma 4.1 *A Runge-Kutta method, applied with step size h to a linear system $y' = Ly$, is equivalent to*

$$y_1 = R(hL) y_0, \tag{4.2}$$

where the rational function $R(z)$ is given by

$$R(z) = 1 + z b^T (I - zA)^{-1} \mathbb{1}, \tag{4.3}$$

$A = (a_{ij})$, $b^T = (b_1, \ldots, b_s)$, *and* $\mathbb{1}^T = (1, \ldots, 1)$. *The function $R(z)$ is called the* stability function *of the method.*

Proof. The Runge-Kutta method (Definition II.1.1), applied to $y' = Ly$, reads

$$hk_i = hL\Big(y_0 + \sum_{j=1}^{s} a_{ij} h k_j\Big)$$

or, using the the supervector $K = (k_1^T, \ldots, k_s^T)^T$,

$$(I - A \otimes hL) hK = \mathbb{1} \otimes hL y_0$$

(here, $A \otimes B = (a_{ij} B)$ denotes the tensor product of two matrices or vectors). Computing hK from this relation, and inserting it into $y_1 = y_0 + \sum_{i=1}^{s} b_i h k_i = y_0 + (b^T \otimes I)(hK)$ proves the statement. □

For the explicit Euler method, the implicit Euler method and the implicit midpoint rule, the stability function $R(z)$ is given by

$$1 + z, \qquad \frac{1}{1-z}, \qquad \frac{1 + z/2}{1 - z/2}.$$

Theorem 4.2 *For Runge-Kutta methods the following statements are equivalent:*

- *the method is symmetric for linear problems $y' = Ly$;*
- *the method is symplectic for problems (4.1) with symmetric C;*
- *the stability function satisfies $R(-z)R(z) = 1$ for all complex z.*

Proof. The method $y_1 = R(hL) y_0$ is symmetric, if and only if $y_0 = R(-hL) y_1$ holds for all initial values y_0. But this is equivalent to $R(-hL)R(hL) = I$.

Since $\Phi_h'(y_0) = R(hL)$, symplecticity of the method for the problem (4.1) is defined by $R(hJ^{-1}C)^T J R(hJ^{-1}C) = J$. For $R(z) = P(z)/Q(z)$ this is equivalent to

$$P(hJ^{-1}C)^T J P(hJ^{-1}C) = Q(hJ^{-1}C)^T J Q(hJ^{-1}C). \tag{4.4}$$

By the symmetry of C, the matrix $L := J^{-1}C$ satisfies $L^T J = -JL$ and hence also $(L^k)^T J = J(-L)^k$ for $k = 0, 1, 2, \ldots$. Consequently, (4.4) is equivalent to

$$P(-hJ^{-1}C) P(hJ^{-1}C) = Q(-hJ^{-1}C) Q(hJ^{-1}C),$$

which is nothing else than $R(-hJ^{-1}C) R(hJ^{-1}C) = I$. □

We remark that symmetry and symplecticity are equivalent properties of Runge-Kutta methods only for linear problems. For general nonlinear problems, there exist symmetric methods that are not symplectic, and there exist symplectic methods that are not symmetric. For example, the *trapezoidal rule*

$$y_1 = y_0 + \frac{h}{2}\Big(f(y_0) + f(y_1)\Big) \tag{4.5}$$

is symmetric, but it does not satisfy the condition (3.2) for symplecticity. In fact, this is true for all Lobatto IIIA methods (see Example II.4.6). On the other hand, the method of Table 4.1 satisfies the symplecticity condition (3.2), but it is clearly not symmetric (the weights do not satisfy $b_{s+1-i} = b_i$).

TABLE 4.1: A symplectic Radau method of order 5 [Su93]

$$\begin{array}{c|ccc}
\frac{4-\sqrt{6}}{10} & \frac{16-\sqrt{6}}{72} & \frac{328-167\sqrt{6}}{1800} & \frac{-2+3\sqrt{6}}{450} \\
\frac{4+\sqrt{6}}{10} & \frac{328+167\sqrt{6}}{1800} & \frac{16+\sqrt{6}}{72} & \frac{-2-3\sqrt{6}}{450} \\
1 & \frac{85-10\sqrt{6}}{180} & \frac{85+10\sqrt{6}}{180} & \frac{1}{18} \\
\hline
 & \frac{16-\sqrt{6}}{36} & \frac{16+\sqrt{6}}{36} & \frac{1}{9}
\end{array}$$

IV.5 Campbell-Baker-Hausdorff Formula

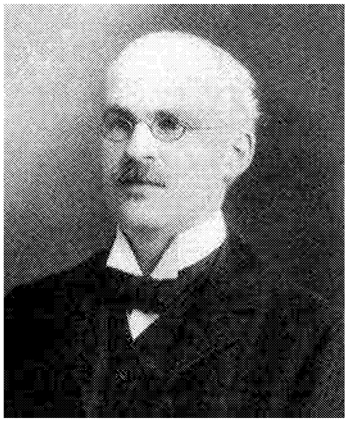
This section is devoted to the derivation of the Campbell-Baker-Hausdorff (short CBH or BCH) formula. It was claimed in 1898 by J.E. Campbell and proved independently by Baker [Ba05] and Hausdorff [Hau06]. This formula will be the essential ingredient for the discussion of splitting methods (Sect. IV.6).

Let A and B be two non-commuting matrices (or operators, for which the compositions $A^k B^l$ make sense). The problem is to find a matrix $C(t)$, such that

$$\exp(tA)\exp(tB) = \exp C(t). \tag{5.1}$$

As long as we do not need the explicit form of $C(t)$, this is a simple task: the expression $\exp(tA)\exp(tB)$ is a series of the form $I + t(A+B) + \mathcal{O}(t^2)$ and, assuming that $C(0) = 0$, $\exp C(t)$ is also close to the identity for small t. Therefore, we can apply the logarithm to (5.1) and we get $C(t)$. Using the series expansion $\log(1+x) = x - x^2/2 + \ldots$, this yields $C(t)$ as a series in powers of t. It starts with $C(t) = t(A+B) + \mathcal{O}(t^2)$ and it has a positive radius of convergence, because it is obtained by elementary operations of convergent series. Consequently, the obtained series for $C(t)$ will converge for bounded A and B, if t is sufficiently small.

The main problem of the derivation of the BCH formula is to get explicit formulas for the coefficients of the series of $C(t)$. With the help of the following lemma, recurrence relations for these coefficients will be obtained, which allow for an easy computation of the first terms.



John Edward Campbell[3]

Henry Frederick Baker[4]

Felix Hausdorff[5]

Lemma 5.1 *Let A and B be (non-commuting) matrices. Then, (5.1) holds, where $C(t)$ is the solution of the differential equation*

$$C' = A + B + \frac{1}{2}[A-B, C] + \sum_{k\ge 2} \frac{B_k}{k!} \operatorname{ad}_C^k(A+B) \tag{5.2}$$

with initial value $C(0)=0$. Recall that $\operatorname{ad}_C A = [C,A] = CA - AC$, and that B_k denote the Bernoulli numbers as in Lemma III.6.2.

Proof. We follow [Var74, Sect. 2.15] and we consider a matrix function $Z(s,t)$ such that

$$\exp(sA)\exp(tB) = \exp Z(s,t). \tag{5.3}$$

Using Lemma III.6.1, the derivative of (5.3) with respect to s is

$$A\,\exp(sA)\exp(tB) = d\exp_{Z(s,t)}\Big(\frac{\partial Z}{\partial s}(s,t)\Big)\exp Z(s,t),$$

so that

$$\frac{\partial Z}{\partial s} = d\exp_Z^{-1}(A) = A - \frac{1}{2}[Z,A] + \sum_{k\ge 2}\frac{B_k}{k!}\operatorname{ad}_Z^k(A). \tag{5.4}$$

We next take the inverse of (5.3)

$$\exp(-tB)\exp(-sA) = \exp(-Z(s,t)),$$

and differentiate this relation with respect to t. As above we get

$$\frac{\partial Z}{\partial t} = d\exp_{-Z}^{-1}(B) = B + \frac{1}{2}[Z,B] + \sum_{k\ge 2}\frac{B_k}{k!}\operatorname{ad}_Z^k(B), \tag{5.5}$$

[3]John Edward Campbell, born: 27 May 1862 in Lisburn, Co Antrim (Ireland), died: 1 October 1924.
[4]Henry Frederick Baker, born: 3 July 1866 in Cambridge (England), died: 17 March 1956.
[5]Felix Hausdorff, born: 8 November 1869 in Breslau (Germany), died: 26 January 1942. All three pictures are copied from `http://www-history.mcs.st-and.ac.uk/~history/Mathematicians`, where one can also find short biographies.

because $\mathrm{ad}_{-Z}^{k}(B) = (-1)^k \mathrm{ad}_Z^k(B)$ and the Bernoulli numbers satisfy $B_k = 0$ for odd $k \ge 2$. A comparison of (5.1) with (5.3) gives $C(t) = Z(t,t)$. The stated differential equation for $C(t)$ therefore follows from $C'(t) = \frac{\partial Z}{\partial s}(t,t) + \frac{\partial Z}{\partial t}(t,t)$, and from adding the relations (5.4) and (5.5). □

Using Lemma 5.1 we can compute the first coefficients of the series

$$C(t) = tC_1 + t^2 C_2 + t^3 C_3 + t^4 C_4 + \dots \,. \tag{5.6}$$

Inserting this ansatz into (5.2) and comparing like powers of t gives

$$\begin{aligned} C_1 &= A + B \\ 2\,C_2 &= \frac{1}{2}[A-B, A+B] = [A,B] \\ 3\,C_3 &= \frac{1}{2}\Big[A-B, \frac{1}{2}[A,B]\Big] \;=\; \frac{1}{4}\Big[A,[A,B]\Big] + \frac{1}{4}\Big[B,[B,A]\Big] \\ 4\,C_4 &= \;\dots\; = \;\frac{1}{6}\Big[A,[B,[B,A]]\Big]. \end{aligned} \tag{5.7}$$

For the simplification of the expression for C_4 we have made use of the Jacobi identity (III.2.2). The next coefficient C_5 contains already 6 independent terms, and for higher order the expressions become soon very complicated.

For later use (construction of splitting methods) we also need a formula for the symmetric composition

$$\exp(tA)\exp(tB)\exp(tA) = \exp D(t). \tag{5.8}$$

Taking the inverse of (5.8), we see that $\exp(-D(t)) = \exp(D(-t))$, so that $D(-t) = -D(t)$, and the expansion of $D(t)$ is in odd powers of t:

$$D(t) = tD_1 + t^3 D_3 + t^5 D_5 + \dots \,. \tag{5.9}$$

By repeated application of the BCH formula (5.1) with coefficients given by (5.7) we find that

$$\begin{aligned} D_1 &= 2A + B \\ 3\,D_3 &= \frac{1}{2}\Big[B,[B,A]\Big] - \frac{1}{2}\Big[A,[A,B]\Big]. \end{aligned} \tag{5.10}$$

Remark 5.2 If A and B are bounded operators, the series (5.6) and (5.9) converge for sufficiently small t. We are, however, also interested in the situation, where A and B are unbounded differential operators. In this case, we still have a formal identity. This means that if we expand both sides of the identities (5.1) or (5.8) into powers of t, then the corresponding coefficients are equal. Truncation of these series therefore introduces a defect of size $\mathcal{O}(t^N)$, where N can be made arbitrarily large.

IV.6 Splitting Methods

For a motivation of splitting methods, let us consider a Hamiltonian system with separable Hamiltonian $H(p,q) = T(p) + V(q)$. It is the sum of two Hamiltonians, which depend either only on p or only on q. The corresponding Hamiltonian systems

$$\begin{aligned} \dot p &= 0 \\ \dot q &= T_p(p) \end{aligned} \qquad \text{and} \qquad \begin{aligned} \dot p &= -V_q(q) \\ \dot q &= 0 \end{aligned} \tag{6.1}$$

can be solved exactly and yield

$$\begin{aligned} p(t) &= p_0 \\ q(t) &= q_0 + t\,T_p(p_0) \end{aligned} \qquad \text{and} \qquad \begin{aligned} p(t) &= p_0 - t\,V_q(q_0) \\ q(t) &= q_0, \end{aligned} \tag{6.2}$$

respectively. Denoting the flows of these two systems by φ_t^T and φ_t^V, one can check that the symplectic Euler method (I.1.8) is nothing other than the composition $\varphi_h^T \circ \varphi_h^V$. Since φ_h^T and φ_h^V are both symplectic transformations, and since the composition of symplectic maps is again symplectic, this gives an elegant proof of the symplecticity of the symplectic Euler method. Furthermore, the adjoint of the symplectic Euler method can be written as $\varphi_h^V \circ \varphi_h^T$, and by (I.3.7) the Verlet scheme becomes $\varphi_{h/2}^V \circ \varphi_h^T \circ \varphi_{h/2}^V$.

The idea of splitting can be applied to very general situations. We consider an arbitrary (not necessarily Hamiltonian) system $y' = f(y)$ in $\mathbb{R}^n$, which can be split as

$$y' = f_1(y) + f_2(y). \tag{6.3}$$

We further assume that the flows $\varphi_t^{[1]}$ and $\varphi_t^{[2]}$ of the systems $y' = f_1(y)$ and $y' = f_2(y)$ can be calculated explicitly (later in Chapter V we shall study splitting methods, where $\varphi_t^{[1]}$ and $\varphi_t^{[2]}$ are replaced with some numerical approximations). An extension of the symplectic Euler method to the new situation is

$$\varphi_h^{[1]} \circ \varphi_h^{[2]}, \tag{6.4}$$

which is often called the *Lie-Trotter formula* [Tr59]. By Taylor expansion we find that $(\varphi_h^{[1]} \circ \varphi_h^{[2]})(y_0) = \varphi_h(y_0) + \mathcal{O}(h^2)$, so that (6.4) gives an approximation of order 1 to the solution of (6.3). The analogue of the Verlet scheme is

$$\varphi_{h/2}^{[1]} \circ \varphi_h^{[2]} \circ \varphi_{h/2}^{[1]}, \tag{6.5}$$

which is the so-called *Strang splitting*[6] [Str68]. Due to its symmetry it is a method of order 2. The order can be still further increased by suitably composing the flows $\varphi_t^{[1]}$ and $\varphi_t^{[2]}$. According to [McL95] we distinguish the following cases:

- **Non-symmetric**. Such methods are of the form

$$\varphi_{b_m h}^{[2]} \circ \varphi_{a_m h}^{[1]} \circ \varphi_{b_{m-1} h}^{[2]} \circ \ldots \circ \varphi_{a_2 h}^{[1]} \circ \varphi_{b_1 h}^{[2]} \circ \varphi_{a_1 h}^{[1]} \tag{6.6}$$

  (a_1 or b_m or both of them are allowed to be zero).

- **Symmetric**. Symmetric methods are obtained by a composition of the form

$$\varphi_{a_m h}^{[1]} \circ \varphi_{b_m h}^{[2]} \circ \ldots \circ \varphi_{a_1 h}^{[1]} \circ \varphi_{b_1 h}^{[2]} \circ \varphi_{a_1 h}^{[1]} \circ \ldots \circ \varphi_{b_m h}^{[2]} \circ \varphi_{a_m h}^{[1]} \tag{6.7}$$

  (here, b_1 or a_m or both are allowed to be zero).

- **Symmetric, composed of symmetric steps**. We let $\Phi_h = \varphi_{h/2}^{[1]} \circ \varphi_h^{[2]} \circ \varphi_{h/2}^{[1]}$ or $\Phi_h = \varphi_{h/2}^{[2]} \circ \varphi_h^{[1]} \circ \varphi_{h/2}^{[2]}$, and we consider the composition

$$\Phi_{b_m h} \circ \Phi_{b_{m-1} h} \circ \ldots \circ \Phi_{b_1 h} \circ \Phi_{b_0 h} \circ \Phi_{b_1 h} \circ \ldots \circ \Phi_{b_{m-1} h} \circ \Phi_{b_m h}. \tag{6.8}$$

[6] The article [Str68] deals with spatial discretizations of partial differential equations such as $u_t = Au_x + Bu_y$. There, the functions f_i typically contain differences in only one spatial direction.

An early contribution to this subject is the article of Ruth [Ru83], where, for the special case (6.1), a non-symmetric method (6.6) of order 3 with $m = 3$ is constructed. A systematic study of such methods has started with the articles of Suzuki [Su90, Su92] and Yoshida [Yo90].

In all three situations the problem is the same: *what are the conditions on the parameters a_i, b_i, such that the compositions (6.6), (6.7) or (6.8) approximate the flow φ_h of (6.3) to a given order p?*

In order to compare the expressions (6.6), (6.7) and (6.8) with the flow φ_h of (6.3), it is convenient to introduce the differential operators D_i (Lie derivative) which, for differentiable functions $F : \mathbb{R}^n \to \mathbb{R}^n$, are defined by

$$D_i F(y) = F'(y) f_i(y), \tag{6.9}$$

where $f_i(y)$ is the function of (6.3). This means that, if $y(t)$ is a solution of $y' = f_i(y)$, then

$$\frac{d}{dt} F(y(t)) \;=\; (D_i F)(y(t)). \tag{6.10}$$

Applying iteratively this operator to the identity map $F(y) = y$, we obtain for the solution $y(t)$ of $y' = f_i(y)$ that $y'(t) = D_i y(t)$, $y''(t) = D_i^2 y(t)$, etc. Consequently, for analytic functions, the solution $\varphi_t^{[i]}(y_0)$ is given by

$$\varphi_t^{[i]}(y_0) \;=\; \sum_{k\geq 0} \frac{t^k}{k!} D_i^k y_0 \;=\; \exp(tD_i) y_0. \tag{6.11}$$

Lemma 6.1 *Let $\varphi_t^{[1]}$ and $\varphi_t^{[2]}$ be the flows of the differential equations $y' = f_1(y)$ and $y' = f_2(y)$, respectively. For their composition we then have*

$$\Big(\varphi_s^{[1]} \circ \varphi_t^{[2]}\Big)(y_0) = \exp(tD_2)\, \exp(sD_1)\, y_0$$

(observe the reversed order of the operators).

Proof. For an arbitrary smooth function $F(y)$, it follows from an iterated application of (6.10) that

$$\frac{d^k}{dt^k} F\Big(\varphi_t^{[2]}(y_0)\Big) = D_2^k F\Big(\varphi_t^{[2]}(y_0)\Big),$$

so that by Taylor series expansion $F(\varphi_t^{[2]}(y_0)) = \sum_{k\geq 0} \frac{t^k}{k!} D_2^k F(y_0)$. Putting $F(y) := \varphi_s^{[1]}(y)$ and using (6.11) gives

$$\Big(\varphi_s^{[1]} \circ \varphi_t^{[2]}\Big)(y_0) = \Big(\sum_{k\geq 0} \frac{t^k}{k!} D_2^k\Big)\Big(\sum_{l\geq 0} \frac{s^l}{l!} D_1^l\Big) y_0 \tag{6.12}$$

which proves the statement. □

In general, the two operators D_1 and D_2 do not commute, so that the composition $\exp(tD_2)\exp(tD_1)y_0$ is different from $\exp(t(D_1 + D_2))y_0$, which represents the solution $\varphi_t(y_0)$ of $y' = f(y) = f_1(y) + f_2(y)$.

Order Conditions The derivation of the order conditions for splitting methods can be done as follows: with the use of Lemma 6.1 we write the method as a product of exponentials, then we apply the Campbell-Baker-Hausdorff formula to get one exponential of a series in powers of h. Finally, we compare this series with $h(D_1 + D_2)$, which corresponds to the exact solution of (6.3).

Let us illustrate this procedure with the methods of type (6.8) (see [Yo90]). Using Lemma 6.1 and formulas (5.8), (5.10), the second order integrator $\Phi_h = \varphi^{[1]}_{h/2} \circ \varphi^{[2]}_h \circ \varphi^{[1]}_{h/2}$ can be written as

$$\begin{aligned} \Phi_h &= \exp(hD_1/2)\exp(hD_2)\exp(hD_1/2) \\ &= \exp(h\alpha_1 + h^3\alpha_3 + h^5\alpha_5 + \ldots), \end{aligned} \tag{6.13}$$

where $\alpha_1 = D_1 + D_2$, $\alpha_3 = \frac{1}{12}[D_2, [D_2, D_1]] - \frac{1}{24}[D_1, [D_1, D_2]]$. The Lie-bracket for differential operators is defined in the usual way, i.e., $[D_1, D_2] = D_1D_2 - D_2D_1$. We next define $\Psi^{(j)}$ recursively by

$$\Psi^{(0)} = \Phi_{b_0 h}, \qquad \Psi^{(j)} = \Phi_{b_j h} \circ \Psi^{(j-1)} \circ \Phi_{b_j h}, \tag{6.14}$$

so that $\Psi^{(m)}$ is equal to the method (6.8).

Lemma 6.2 *The operators $\Psi^{(j)}$, defined by (6.14) and (6.13), satisfy*

$$\Psi^{(j)}(y_0) = \exp\Big(A_{1,j}h\alpha_1 + A_{3,j}h^3\alpha_3 + A_{5,j}h^5\alpha_5 + B_{5,j}h^5[\alpha_1, [\alpha_1, \alpha_3]] + \mathcal{O}(h^7)\Big)y_0 \tag{6.15}$$

where

$$A_{1,0} = b_0, \qquad A_{3,0} = b_0^3, \qquad A_{5,0} = b_0^5, \qquad B_{5,0} = 0$$

and

$$\begin{aligned} A_{1,j} &= A_{1,j-1} + 2b_j \\ A_{3,j} &= A_{3,j-1} + 2b_j^3 \\ A_{5,j} &= A_{5,j-1} + 2b_j^5 \\ B_{5,j} &= B_{5,j-1} + \frac{1}{6}\Big(A_{1,j-1}^2 b_j^3 - A_{1,j-1}A_{3,j-1}b_j - A_{3,j-1}b_j^2 + A_{1,j-1}b_j^4\Big). \end{aligned}$$

Proof. We use the formulas (5.8), (5.9), (5.10) with tA replaced with $b_jh\alpha_1 + (b_jh)^3\alpha_3 + \ldots$, and tB replaced with $A_{1,j-1}h\alpha_1 + A_{3,j-1}h^3\alpha_3 + \ldots$. This gives $\Psi^{(j)}(y_0) = \exp(D(h))y_0$ with

$$\begin{aligned} D(h) &= (2b_j + A_{1,j-1})h\alpha_1 + (2b_j^3 + A_{3,j-1})h^3\alpha_3 + (2b_j^5 + A_{5,j-1})h^5\alpha_5 + B_{5,j-1}h^5[\alpha_1, [\alpha_1, \alpha_3]] \\ &+ \frac{1}{6}\Big[A_{1,j-1}h\alpha_1, \Big[A_{1,j-1}h\alpha_1 + A_{3,j-1}h^3\alpha_3\,,\, b_jh\alpha_1 + b_j^3h^3\alpha_3\Big]\Big] \\ &- \frac{1}{6}\Big[b_jh\alpha_1, \Big[b_jh\alpha_1 + b_j^3h^3\alpha_3\,,\, A_{1,j-1}h\alpha_1 + A_{3,j-1}h^3\alpha_3\Big]\Big] + \mathcal{O}(h^7). \end{aligned}$$

A comparison of $D(h)$ with the argument in (6.15) proves the statement. □

Theorem 6.3 *The order conditions for the splitting method (6.8) are:*

- *order 2:* $A_{1,m} = 1,$
- *order 4:* $A_{1,m} = 1, \quad A_{3,m} = 0,$
- *order 6:* $A_{1,m} = 1, \quad A_{3,m} = 0, \quad A_{5,m} = 0, \quad B_{5,m} = 0.$

The coefficients $A_{i,m}$ *and* $B_{5,m}$ *are those defined in Lemma 6.2.*

Proof. This is an immediate consequence of Lemma 6.2, because the conditions of order p imply that the Taylor series expansion of $\Psi^{(m)}(y_0)$ coincides with that of the solution $\varphi_h(y_0) = \exp(h(D_1 + D_2))y_0$ up to terms of size $\mathcal{O}(h^p)$. □

It is interesting to note that the order conditions of Theorem 6.3 do not depend on the special form of α_3 and α_5 in (6.13). We also remark that the composition methods of Sect. II.6 are a special case of the splitting method (6.8). Theorem 6.3 therefore explains the order conditions (II.6.3) and (II.6.4), which were mysterious with the techniques of Chapter II.

[Yo90] solves the order conditions for order 6 with $m = 3$ (four equations for the four parameters b_0, b_1, b_2, b_3). He finds three solutions, one of which is given in the end of Sect. II.6. [Yo90] also presents some methods of order 8. A careful investigation of symmetric splitting methods of orders 2 to 8 can be found in [McL95]. There, several new methods with small error constants are presented.

Remark 6.4 We emphasize that splitting methods are an important tool for the construction of symplectic integrators. If we split a Hamiltonian as $H(y) = H_1(y) + H_2(y)$, and if we consider the vector fields $f_i(y) = J^{-1}\nabla H_i(y)$, then the flows $\varphi_t^{[i]}$ are symplectic, and therefore all splitting methods are automatically symplectic.

IV.7 Volume Preservation

IV.8 Generating Functions

IV.9 Variational Approach

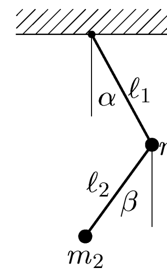
Marsden, etc

IV.10 Symplectic Integrators on Manifolds

IV.11 Exercises

1. Prove that a linear transformation $A : \mathbb{R}^2 \to \mathbb{R}^2$ is symplectic, if and only if $\det A = 1$.
2. Prove that the flow of a Hamiltonian system satisfies $\det \varphi_t'(y) = 1$ for all y and all t. Deduce from this result that the flow is *volume preserving*, i.e., for $B \subset \mathbb{R}^{2d}$ it holds that $\mathrm{vol}\,(\varphi_t(B)) = \mathrm{vol}\,(B)$ for all t.

3. Consider the Hamiltonian system $y' = J^{-1}\nabla H(y)$ and a variable transformation $y = \varphi(z)$. Prove that, for a symplectic transformation $\varphi(z)$, the system in the z-coordinates is again Hamiltonian with $\widetilde{H}(z) = H(\varphi(z))$.
4. Consider a Hamiltonian system with $H(p,q) = \frac{1}{2}p^T p + V(q)$. Let $q = \chi(Q)$ be a change of position coordinates. How has one to define the variable P (as a function of p and q) so that the system in the new variables (P,Q) is again Hamiltonian?
Result. $P = \chi'(Q)^T p$.
5. Let α and β be the generalized coordinates of the double pendulum, whose kinetic and potential energies are



$$\begin{aligned} T &= \frac{m_1}{2}(\dot{x}_1^2 + \dot{y}_1^2) + \frac{m_2}{2}(\dot{x}_2^2 + \dot{y}_2^2) \\ U &= m_1 g y_1 + m_2 g y_2. \end{aligned}$$

Determine the generalized momenta of the corresponding Hamiltonian system.
6. Consider the transformation $(r,\varphi) \mapsto (p,q)$, defined by
$$p = \psi(r)\cos\varphi, \qquad q = \psi(r)\sin\varphi.$$
For which function $\psi(r)$ is it a symplectic transformation?
7. Write Kepler's problem with Hamiltonian
$$H(p,q) = \frac{1}{2}(p_1^2 + p_2^2) - \frac{1}{\sqrt{q_1^2 + q_2^2}}$$
in polar coordinates $q_1 = r\cos\varphi$, $q_2 = r\sin\varphi$. What are the conjugated generalized momenta p_r, p_φ? What is the Hamiltonian in the new coordinates.
8. On the set $U = \{(p,q)\,;\, p^2 + q^2 > 0\}$ consider the differential equation
$$\begin{pmatrix} \dot{p} \\ \dot{q} \end{pmatrix} = \frac{1}{p^2 + q^2}\begin{pmatrix} p \\ q \end{pmatrix}. \tag{11.1}$$
a) Prove that its flow is symplectic everywhere on U.
b) On every simply-connected subset of U the vector field (11.1) is Hamiltonian (with $H(p,q) = \mathrm{Im}\,\log(p + iq) + Const$).
c) It is not possible to find a differentiable function $H : U \to \mathbb{R}$ such that (11.1) is equal to $J^{-1}\nabla H(p,q)$ for all $(p,q) \in U$.
Remark. The vector field (11.1) is called *locally Hamiltonian.*
9. Prove that the definition (2.5) of $\Omega(M)$ does not depend on the parametrization φ, i.e., the parametrization $\psi = \varphi \circ \alpha$, where α is a diffeomorphism between suitable domains of $\mathbb{R}^2$, leads to the same result.
10. Prove that the coefficient C_4 in the series (5.6) of the Campbell-Baker-Hausdorff formula is given by $[A,[B,[B,A]]]/6$.
11. Deduce the BCH formula from the Magnus expansion (III.6.9).
Hint. For constant matrices A and B consider the matrix function $A(t)$, defined by $A(t) = B$ for $0 \le t \le 1$ and $A(t) = A$ for $1 \le t \le 2$.
12. Prove that the series (5.6) of the BCH formula converges for $|t| < \ln 2/(\|A\| + \|B\|)$.
13. What are the conditions on the parameters a_i and b_i, such that the splitting method (6.6) is of order 2, of order 3?
14. How many order conditions have to be satisfied by a symmetric splitting method (6.7) to get order 4? The result is 4.