Sanz-Serna, J. M. (2016). Symplectic Runge--Kutta Schemes for Adjoint Equations, Automatic Differentiation, Optimal Control, and More. *SIAM Review*, 58(1), 3–33.

DOI: 10.1137/151002769

Symplectic Runge-Kutta schemes for adjoint equations, automatic differentiation, optimal control and more

J.M. Sanz-Serna*

September 8, 2014

Abstract

It is well known that symplectic Runge-Kutta and Partitioned Runge-Kutta methods exactly preserve *quadratic* first integrals (invariants of motion) of the system being integrated. While this property is often seen as a mere curiosity (it does not hold for arbitrary first integrals), it plays an important role in the computation of numerical sensitivities, optimal control theory and Lagrangian mechanics, as described in this paper, which, together with some new material, presents in a unified way a number of results now scattered or implicit in the literature. Some widely used procedures, such as the direct method in optimal control theory and the computation of sensitivities via reverse accumulation imply 'hidden' integrations with symplectic Partitioned Runge-Kutta schemes.

AMS numbers: 34H05, 49A10, 65L06, 65K10, 65P10, 70H25

Keywords: Runge-Kutta methods, Partitioned Runge-Kutta methods, symplectic integration, Hamiltonian systems, variational equations, adjoint equations, computation of sensitivities, Lagrange multipliers, automatic differentiation, optimal control, Lagrangian mechanics, reflected and transposed Runge-Kutta schemes, differential-algebraic problems, constrained controls

1 Introduction

Symplectic Runge-Kutta (RK) [22], [28], [36] and Partitioned Runge-Kutta (PRK) [1], [37] formulae were introduced to integrate Hamiltonian systems in long time intervals. They are defined in terms of a purely geometric property and provided the first widely studied instance of what it was later termed *geometric integration* [29]. While it is well known that symplectic RK methods exactly preserve all *quadratic* first integrals

*Departamento de Matemáticas, Universidad Carlos III de Madrid, Spain. jmsanzserna@gmail.com

(invariants of motion) of the system being integrated, such a property is often seen as a curiosity: it does not hold for arbitrary first integrals. The aim of this paper is to emphasize that the conservation of quadratic invariants plays an important role in the computation of numerical sensitivities, in optimal control theory and in classical mechanics. Actually, some widely used procedures, such as the direct method in optimal control theory and the computation of sensitivities via reverse accumulation imply 'hidden' integrations with symplectic PRK schemes; therefore the theory of symplectic PRK integration should be helpful in understanding such procedures.

The paper presents, in what we hope is a coherent way, some new results together with results that are already available in the literature of different communities. It is not always the case that such known results appear in their sources in the unifying language used here. In order to cater for a variety of possible readers, this paper is written without assuming much background. We hope it will help researchers in optimal control to better understand RK schemes and, similarly, encourage RK experts to consider sensitivities and optimal control problems.

Section 2 provides background on numerical integrators. We introduce the necessary notation and recall a number of properties of *symplectic* RK and related schemes. In particular, we quote some results (Theorems 1, 3) that ensure the exact preservation by the integrator of quadratic conservation laws; such a preservation is the linchpin of the paper.

Section 3 is devoted to the integration of the adjoint variational equations used to perform sensitivity analysis. While it is well known that an RK method $\mathcal{M}$ applied to the variational equations of a system $\mathcal{S}$ automatically produces the variational equations for the discretisation of $\mathcal{S}$ by means of $\mathcal{M}$ (Theorem 5), the situation for the adjoints is more complicated, cf. [34]. There are three cases of increasing complexity:

- $\mathcal{S}$ is integrated with a *symplectic* RK scheme $\mathcal{M}$. Then the application of $\mathcal{M}$ to the adjoint equations of $\mathcal{S}$ produces the adjoint equations for discretisation of $\mathcal{S}$ by means of $\mathcal{M}$ (Theorem 6).

- $\mathcal{S}$ is integrated with a *non-symplectic* RK scheme $\mathcal{M}$ *with non-vanishing weights*. Then, the adjoint equations for the discretisation are obtained by integrating the adjoint equations of $\mathcal{S}$ with a *different* set of RK coefficients, so that the overall procedure is a symplectic PRK method (Theorem 7). The recipe for the adjoint coefficients is given in formula (41) below. The method used for the adjoint equations will in general be of lower order than the RK scheme $\mathcal{M}$ used for the main integration and will also have different stability properties. For these reasons non-symplectic methods $\mathcal{M}$ should be used with care. The computation of sensitivities of the discrete solution via *automatic differentiation with reverse accummulation* implicitly provides the *symplectic PRK integration* of the adjoint equations with coefficients (41) (Theorem 8).

- $\mathcal{S}$ is integrated with a *non-symplectic* RK scheme $\mathcal{M}$ having one or more null weights. Then, to obtain the adjoint equations of the discretisation, the continuous adjoint equations have to be integrated with a fancy integrator outside the RK class (Appendix A). Again an order reduction is likely to take place and

again the fancy integration is implicitly performed whenever differentiation with reverse accumulation is used.

Section 4 deals with the Mayer optimal control problem in the case of unconstrained controls. The situation there is quite similar to that in the preceding section (the case of vanishing weights is discussed in Appendix A):

- For a symplectic RK method, *commutation* [26] takes place : the discretisation of the continuous first order conditions necessary for optimality provides the first order necessary conditions for the discrete solution (Theorem 10).

- When the equations for the states are discretised with a *non-symplectic* RK scheme *with non-vanishing weights*, to achieve commutation the costate equations have to be integrated by means of a clever set of coefficients that does not coincide with the set used for the states (Theorem 10). With this clever set, the overall integration (states+costates) is performed with a symplectic PRK method. In general, an order reduction will take place for states, costates and controls. As first noted by Hager [15], the required set of coefficients is alternatively defined, not by imposing symplecticness of the integration, but by using the *direct* approach, i.e. by minimising the cost in the discrete realm with the help of Lagrange multipliers (Theorem 11).

For a *symplectic* RK or PRK integration of the system for states and costates, the direct and indirect approach are mathematically equivalent. When a non-symplectic PRK is used in the indirect approach, the discrete solution *cannot* be reached via the direct approach, which always implies a symplectic integration of the states+costates system.

Some extensions are presented in Section 5. Section 6 is devoted to classical mechanics. Hamilton's variational principle may of course be viewed as a control problem: it is a matter of minimising a functional subject to differential constraints. As it is well known, the application of the theory of optimal control to this situation replicates the standard procedure to obtain Hamilton's canonical equations from Hamilton's principle. In the discrete realm, this process provides the variational derivation of symplectic PRK integrators, originally due to Suris [37].

There are two appendices with more technical material. The first deals with the problem of how to 'supplement' a given non-symplectic RK method with vanishing weights so as to have a symplectic algorithm for partitioned systems. The second relates the contents of the paper and the theory of reflection and transposition of RK coefficients introduced by Schrerer and Türke [32].

In order not to clutter the exposition with unwanted details, I shall not be concerned with technical issues such as existence of solutions of implicit integrators, smoothness requirements and so on. These may be very important in some circumstances (e.g. smoothness poses difficulties if the controls are constrained, see [9]).

2 Numerical integrators

In this section we review some results on RK and related methods. For more details the reader is referred to [31], [5], [17], [19], [20].

2.1 Runge-Kutta schemes

An RK method with s stages is specified by $s^2 + 2s$ numbers

$$a_{ij}, \quad i,j = 1,\dots,s, \qquad b_i,\ c_i, \quad i = 1,\dots,s. \tag{1}$$

Given a D-dimensional differential system, $F : \mathbb{R}^D \times \mathbb{R} \to \mathbb{R}^D$,

$$\frac{d}{dt}y = F(y,t), \tag{2}$$

to be studied in an interval, $t_0 \le t \le t_0 + T$, and an initial condition

$$y(t_0) = A \in \mathbb{R}^D, \tag{3}$$

the method (1) finds approximations y_n to the values $y(t_n)$, $n = 0, 1, \dots, N$, of the solution of (2)–(3), $t_0 < t_1 < \cdots < t_N = t_0 + T$, by setting $y_0 = A$ and, recursively,

$$y_{n+1} = y_n + h_n \sum_{i=1}^{s} b_i K_{n,i}, \qquad n = 0, 1 \dots, N-1. \tag{4}$$

Here $h_n = t_{n+1} - t_n$ denotes the step-length and $K_{n,i}$, $i = 1,\dots,s$, are the 'slopes'

$$K_{n,i} = F(Y_{n,i}, t_n + c_i h_n) \tag{5}$$

at the so-called internal stages $Y_{n,i}$. The vectors $Y_{n,1}$,..., $Y_{n,s}$ are in turn defined by the relations

$$Y_{n,i} = y_n + h_n \sum_{j=1}^{s} a_{ij} K_{n,j}, \quad i = 1,\dots,s. \tag{6}$$

In the particular case where the matrix (a_{ij}) is, perhaps after renumbering the stages, lower triangular (explicit RK methods), the stages are computed recursively from (5)–(6). In the general case, (5)–(6) provides, for each n, a system of coupled equations to be solved for the stages.

The internal stages should not be confused with the values y_n output by the integrator and may merely be regarded as auxiliary variables. Alternatively, the vector $Y_{n,i}$ is sometimes viewed as an approximation to the off-step value $y(t_n + c_i h_n)$. It is important to emphasise that the differences $y(t_n + c_i h_n) - Y_{n,i}$ are typically much larger than the differences $y(t_n) - y_n$.

When the system (2) is autonomous, i.e. $F = F(y)$, the c_i play no role. At the other end of the spectrum, if F is independent of y, the RK discretisation amounts to the use in the interval $t_0 \le t \le t_0 + T$ of the composite quadrature rule based on the *abscissas* c_i and the *weights* b_i.

An RK scheme is said to possess order ρ if, for $t_0 \le t_n \le t_0 + T$ and smooth problems, $|y_n - y(t_n)| = \mathcal{O}(h^\rho)$, where $h = \max_n h_n$. The expansion of the local truncation error in powers of the step-length h_n includes, for each power h_n^k, $k = 1, 2, \ldots$, one or several elementary differentials of F; an integrator has order $\ge \rho$ if and only if, in that expansion, the coefficients of the elementary differentials of orders $k = 1, \ldots, \rho$ vanish. For instance, the relations (order conditions)

$$\sum_{i=1}^{s} b_i = 1, \quad \sum_{i,j=1}^{s} b_i a_{ij} = \frac{1}{2}, \quad \sum_{i,j,k=1}^{s} b_i a_{ij} a_{jk} = \frac{1}{6}, \quad \sum_{i,j,k=1}^{s} b_i a_{ij} a_{ik} = \frac{1}{3}, \tag{7}$$

ensure order at least 3 for autonomous problems. They correspond to the elementary differentials F (of order 1), $(\partial_y F)F$ (of order 2) and $(\partial_y F)(\partial_y F)F$, $(\partial^2_{yy} F)[F, F]$ (both of order 3) ($\partial_y F$ is the Jacobian matrix and $\partial^2_{yy} F$ the tensor of second derivatives). Since the work of Butcher in the early 1960's, order conditions and elementary differentials are studied with the help of graphs. To impose order $\ge \rho$ for autonomous problems, there is an independent order condition for each rooted tree with ρ or fewer vertices. Most, but not all, useful RK schemes satisfy $c_i = \sum_j a_{ij}$ for each i; for them order ρ for autonomous problems implies order ρ for all problems.

The present paper is based on the following 1987 result of Cooper [8]. It ensures that *some* RK methods automatically inherit the quadratic conservation laws possessed by the system being integrated.

Theorem 1 *Assume that the system (2) possesses a quadratic first integral I, i.e. $I(\cdot, \cdot)$ is a real-valued bilinear mapping in $\mathbb{R}^D \times \mathbb{R}^D$ such that, for each A and t_0, the solution $y(t)$ of (2)–(3) satisfies $(d/dt)I(y(t), y(t)) \equiv 0$. The relations*

$$b_i a_{ij} + b_j a_{ji} - b_i b_j = 0, \qquad i, j = 1, \ldots, s, \tag{8}$$

guarantee that, for each RK trajectory $\{y_n\}$ satisfying (4)–(6), $I(y_n, y_n)$ is independent of n.

We shall not reproduce here the proof of this result; it is similar to that of Theorem 3 below. The relations (8) are essentially necessary for an RK scheme to conserve *each* quadratic first integral of *each* differential system [17, Chapter VI, Theorems 7.6, 7.10].

In many applications the system (2) is Hamiltonian. This means that D is even and, after writing $y = [q^\mathsf{T}, p^\mathsf{T}]^\mathsf{T}$, $F = [f^\mathsf{T}, g^\mathsf{T}]^\mathsf{T}$, with $q, p, f, g \in \mathbb{R}^d$, $d = D/2$, there exists a real-valued function $H(p, q, t)$ (the Hamiltonian) such that $f^r = \partial H/\partial p^r$, $g^r = -\partial H/\partial q^r$, $r = 1, \ldots, d$ (superscripts indicate components). Hamiltonian systems are characterised geometrically by the symplectic property of the corresponding solution flow [2]. When $d = 1$, simplecticness means conservation of oriented area; in higher dimensions a similar but more complicated interpretation exists. It is often advisable [31], [17], [23] to integrate Hamiltonian problems by means of so-called symplectic algorithms, i.e. algorithms such that the transformation $y_n \mapsto y_{n+1}$ in $\mathbb{R}^{2d}$ is symplectic; those algorithms are particularly advisable in integrations where the interval $t_0 \le t \le t_0 + T$ is long (for a recent reference in that connection, see [11], which is part of a project to integrate the solar system over a 60 million year interval). Using

the method of modified equations [14], each numerical solution may (approximately) be interpreted as a true solution of a nearby differential system called the modified system. For symplectic methods applied to Hamiltonian systems, the modified system is Hamiltonian; for non-symplectic discretisations, the modified system, while close to the system being integrated, is not Hamiltonian and this fact is likely to imply a substantial distortion of the long-time dynamics [31], [17].

The first symplectic integrators were constructed in an *ad hoc* way; it was later discovered (independently by Lasagni [22], Suris [36] and the present author [28]) that the class of RK methods contains many symplectic schemes:

Theorem 2 *Assume that the system (2) is Hamiltonian. The relations (8) guarantee that the mapping $y_n \mapsto y_{n+1}$ defined in (4)–(6) is symplectic.*

The proof of Theorem 2, not included here, is very similar to the proof of Theorem 1. Just as for the conservation of quadratic first integrals, it turns out, see [31], Section 6.5, that the relations (8) are essentially necessary for $y_n \mapsto y_{n+1}$ to be symplectic for each Hamiltonian system.

The set of relations (8) thus ensures *two* different properties: quadratic conservation and symplecticness. These two properties are not unrelated: symplecticness may be viewed a consequence of the quadratic conservation because, as noted in [3], the preservation of the symplectic structure by a Hamiltonian solution flow may be interpreted as a bilinear first integral of the solution flow of the associated variational system.

The symplectic character of RK schemes satisfying (8) has attracted much attention in view of the importance of Hamiltonian systems in the applications. On the other hand, it is fair to say that quadratic conservation has been largely unnoticed or even disregarded as a mere curiosity. For this reason, while schemes satisfying (8) could have been called conservative, the following terminology is standard:

Definition 1 *The RK scheme (1) is called* symplectic *(or canonical) if (8) holds.*

Our focus in this paper is on symplectic schemes in as far as they conserve quadratic invariants, as these are actually crucial in several applications. The discussion of any possible benefits derived from the symplectic character of the map $y_n \mapsto y_{n+1}$, including the existence of modified Hamiltonian systems, are out of our scope here. The paper [7] is, in this sense, complementary to the present work.

It was proved in [30] that the relations (8) act as simplifying assumptions *vis-à-vis* the order conditions: once these relations are imposed, the order conditions corresponding to the different elementary differentials/rooted trees are no longer independent. For instance, it is a simple exercise to show that, when (8) holds, the second order condition in (7) is a consequence of the first and therefore symplectic RK schemes of order ≥ 1 automatically possess order ≥ 2. Similarly the last order condition in (7) is a consequence of the first three. In this way, for a general RK methods to have order ≥ 3 for autonomous problems, there are 4 order conditions; for symplectic methods the number is only 2. For a symplectic RK method to have order $\geq \rho$ for autonomous problems there is an order condition for each so-called non-superfluous free tree with $\leq \rho$ vertices.

There are many symplectic RK methods [31] including the Gauss methods (of maximal order $2s$ and positive weights) as first shown in [28]; however no symplectic RK scheme is explicit.

2.2 Partitioned Runge-Kutta schemes

In some applications the components of the vector y in (2) appear partitioned into two blocks: $y = [q^\mathsf{T}, p^\mathsf{T}]^\mathsf{T}$, $q \in \mathbb{R}^{D-d}$, $p \in \mathbb{R}^d$. Hamiltonian problems, where $d = D/2$, provide an example, as we have just seen. In those cases it may make sense to use a set of coefficients (1) for the integration of the block q and a second set

$$A_{ij}, \quad i,j = 1,\dots,s, \qquad B_i,\, C_i, \quad i = 1,\dots,s, \tag{9}$$

for the integration of the block p. (There is no loss of generality in assuming that the number of stages s in (9) coincides with that in (1): see [31] Remark 3.2.) The overall method is called a PRK scheme. PRK methods are not a mathematical nicety: the Verlet algorithm, the method of choice in molecular dynamics [33] is one of them. A more precise description follows.

Denote by $F = [f^\mathsf{T}, g^\mathsf{T}]^\mathsf{T}$, $f \in \mathbb{R}^{D-d}$, $g \in \mathbb{R}^d$ the partitioning of F induced by the partitioning of y, so that (2) reads

$$\frac{d}{dt}q = f(q,p,t), \qquad \frac{d}{dt}p = g(q,p,t); \tag{10}$$

then the equations for the step $n \to n+1$ of the PRK method (1), (9) are

$$q_{n+1} = q_n + h_n \sum_{i=1}^{s} b_i k_{n,i}, \quad p_{n+1} = p_n + h_n \sum_{i=1}^{s} B_i \ell_{n,i}, \quad n = 0,\dots,N-1, \tag{11}$$

where

$$k_{n,i} = f(Q_{n,i}, P_{n,i}, t_n + c_i h_n), \qquad \ell_{n,i} = g(Q_{n,i}, P_{n,i}, t_n + C_i h_n), \tag{12}$$

and the internal stages $Q_{n,i}$, $P_{n,i}$, $i = 1,\dots,s$, are defined by the relations

$$Q_{n,i} = q_n + h_n \sum_{i=1}^{s} a_{ij} k_{n,j}, \qquad P_{n,i} = p_n + h_n \sum_{j=1}^{s} A_{ij} \ell_{n,j}. \tag{13}$$

Clearly an RK scheme may be regarded as a particular instance of a PRK method where the two sets (1), (9) happen to coincide. For PRK methods to possess order $\geq \rho$ for autonomous problems, there is an order condition associated with each bicolour rooted tree with ρ or less vertices (see e.g. [17, Chapter III]). For order ≥ 2 the order conditions are:

$$\textstyle\sum_i b_i = 1, \quad \sum_i B_i = 1, \tag{14}$$

$$\textstyle\sum_{ij} b_i a_{ij} = \frac{1}{2}, \quad \sum_{ij} b_i A_{ij} = \frac{1}{2}, \quad \sum_{ij} B_i a_{ij} = \frac{1}{2}, \quad \sum_{ij} B_i A_{ij} = \frac{1}{2}; \tag{15}$$

they correspond to the elementary differentials f, g, $(\partial_x f)f$, $(\partial_x f)g$, $(\partial_x g)f$, $(\partial_x g)g$ respectively. It will be important later to note that, if the PRK (1), (9) has order ρ,

then the RK scheme with coefficients (1) and the RK scheme with coefficients (9) have both order ρ. The converse it is not true: if (1) and (9) are the coefficients of two RK schemes of order ρ, then the combined PRK scheme may have order $< \rho$. This is plain in (15), where the second and third relations are necessary for the PRK to have order ≥ 2 but are obviously not required for (1) and (9) to be the coefficients of two different RK schemes of order ≥ 2.

For PRK methods, the result corresponding to Theorem 1 is (cf. [17, Chapter IV, Theorem 2.4], where only the autonomous case is envisaged):

Theorem 3 *Assume that $S(\cdot,\cdot)$ is a real-valued bilinear map in $\mathbb{R}^d \times \mathbb{R}^{D-d}$ such that, for each t_0 and A, the solution $y(t) = [q(t)^{\mathsf{T}}, p(t)^{\mathsf{T}}]^{\mathsf{T}}$ of (3), (10), satisfies*

$$\frac{d}{dt} S(q(t), p(t)) \equiv 0.$$

The relations

$$b_i = B_i, \quad i = 1, \ldots, s, \quad b_i A_{ij} + B_j a_{ji} - b_i B_j = 0, \quad i, j = 1, \ldots, s, \tag{16}$$

and

$$c_i = C_i, \quad i = 1, \ldots, s, \tag{17}$$

guarantee that, for each PRK trajectory satisfying (11)–(13), $S(q_n, p_n)$ is independent of n.

As in the case of RK methods, the condition in the theorem is necessary for conservation to hold for all S and all partitioned differential systems, see [17, Chapter VI, Theorems 7.6, 7.10]. In the particular case of autonomous problems the abscissas play no role. Thus, to achieve conservation, it is not necessary to impose the condition (17) whenever f and g are independent of t. Note that the theorem only applies to a quadratic function of the form $S(q,p)$ which is not the most general possible; for instance the inner product $q^{\mathsf{T}} q$ is not included in that format.

Before proving the theorem we present a simple algebraic auxiliary result that will be used repeatedly later in other contexts.

Lemma 1 *Let q_n, p_n, Q_i, P_i, $k_{n,i}$, $\ell_{n,i}$ be arbitrary vectors satisfying (11) and (13). If S is bilinear and (16) holds, then*

$$S(q_{n+1}, p_{n+1}) - S(q_n, p_n) = h_n \sum_i b_i \Big(S(k_{n,i}, P_{n,i}) + S(Q_{n,i}, \ell_{n,i}) \Big). \tag{18}$$

Proof: Since S is bilinear, we may write from (11)

$$\begin{aligned} S(q_{n+1}, p_{n+1}) - S(q_n, p_n) \quad &= \quad h_n \sum_i b_i S(k_{n,i}, p_n) + h_n \sum_j B_j S(q_n, \ell_{n,j}) \\ &\quad + h_n^2 \sum_{ij} b_i B_j S(k_{n,i}, \ell_{n,j}). \end{aligned}$$

Now use (13) to eliminate q_n and p_n from the right-hand side:

$$\begin{aligned} S(q_{n+1},p_{n+1}) - S(q_n,p_n) &= h_n \sum_i b_i S(k_{n,i}, P_{n,i} - h_n \sum_j A_{ij}\ell_{n,j}) \\ &\quad + h_n \sum_j B_j S(Q_{n,j} - \sum_i a_{ji} k_{n,i}, \ell_{n,j}) \\ &\quad + h_n^2 \sum_{ij} b_i B_j S(k_{n,i}, \ell_{n,j}). \end{aligned}$$

In view of the bilinearity and (16), the proof is complete. □

Proof of the theorem: Conservation of S implies that $S(f(q,p,t),p)+S(q,g(q,p,t)) \equiv 0$, because, along each solution $q(t)$, $p(t)$,

$$S\big(\frac{d}{dt}q(t),p(t)\big) + S\big(q(t),\frac{d}{dt}p(t)\big) = \frac{d}{dt}S(q(t),p(t)) = 0.$$

Therefore (12) and (17) entail that the right-hand side of (18) vanishes. □

For the preservation of the symplectic structure, the result (derived in [37] and [1] independently) is:

Theorem 4 *Assume that the system (10) is Hamiltonian. The relations (16)–(17) guarantee that the mapping $(q_n,p_n) \mapsto (q_{n+1},p_{n+1})$ defined in (11)–(13) is symplectic.*

The conditions (16)–(17) are essentially necessary for symplecticness [31] and hence the following definition:

Definition 2 *The PRK scheme (1), (9) is called* symplectic *if (16)–(17) hold.*

If the PRK is symplectic, there is a reduction in the number of independent order conditions; the classes of equivalent order conditions were first described by Hairer [16]. An alternative treatment (see [25]) based on so-called H-trees was given by Murua in his 1995 thesis, cf. [4]. For instance, for a symplectic PRK method to have order ≥ 4 it is necessary to impose 13 order conditions: for general PRK methods that number is 36.

3 Variational systems and their adjoints

We now explore the role of symplectic RK schemes when integrating adjoint variational systems. A comprehensive discussion of the use of adjoints to determine sensitivities is not within our scope here. The paper [12] provides a general introduction, together with applications to aerodynamics. Applications of adjoints to atmospheric models are discussed in [27]. Of course the idea of an adjoint problem is not restricted to differential equations; see [6] for an early paper describing a very general framework.

3.1 The continuous problem: quadratic conservation

In this section we consider a d-dimensional differential system

$$\frac{d}{dt}x = f(x,t). \tag{19}$$

We denote by $\alpha \in \mathbb{R}^d$ the corresponding initial value and by $\bar{x}(t)$ the solution that arises from the perturbed initial condition $\bar{x}(t_0) = \alpha + \eta$. Linearisation of (19) around $x(t)$ shows that, as $|\eta| \to 0$, $\bar{x}(t) = x(t) + \delta(t) + o(|\eta|)$, where δ solves the (linear) *variational system* (see e.g. [19] Section I.14)

$$\frac{d}{dt}\delta = \partial_x f(x(t),t)\,\delta, \tag{20}$$

($\partial_x f$ is the Jacobian matrix of f with respect to x). Thus, when $x(t)$ is known, solving for $\delta(t_0+T)$ the initial-value problem given by (20) and $\delta(t_0) = \eta$ yields an estimate for the change in solution $\bar{x}(t) - x(t)$.

The *adjoint* system of (20) is given by

$$\frac{d}{dt}\lambda = -\partial_x f(x(t),t)^{\mathsf{T}}\,\lambda. \tag{21}$$

(To avoid confusion, variables in this paper are always *column vectors*; from a mathematical point of view it would have been better to write sensitivities, Lagrange multipliers and momenta as row vectors, as they belong to the dual space of the space of states.) The right-hand side in (21) has been chosen in such a way that the the following proposition is valid. More precisely, it is best to think that *the adjoint is the system for which (23) below holds*.

Proposition 1 *For each x, δ, $\lambda \in \mathbb{R}^d$ and real t:*

$$\big(-\partial_x f(x,t)^{\mathsf{T}}\,\lambda\big)^{\mathsf{T}}\delta + \lambda^{\mathsf{T}}\partial_x f(x,t)\delta = 0.$$

Therefore if $\delta(t)$ and $\lambda(t)$ are arbitrary solutions of (20), (21) respectively, then

$$\frac{d}{dt}\lambda(t)^{\mathsf{T}}\delta(t) = \Big(\frac{d}{dt}\lambda(t)\Big)^{\mathsf{T}}\delta(t) + \lambda(t)^{\mathsf{T}}\Big(\frac{d}{dt}\delta(t)\Big) \equiv 0 \tag{22}$$

and accordingly

$$\lambda(t_0+T)^{\mathsf{T}}\delta(t_0+T) = \lambda(t_0)^{\mathsf{T}}\delta(t_0). \tag{23}$$

Why is the adjoint system useful? Regard η as a parameter and assume that we are interested in finding $\omega^{\mathsf{T}}\delta(t_0+T)$ for fixed $\omega \in \mathbb{R}^d$, i.e. in estimating, at the final time t_0+T, the change along the direction of ω of the solution of (19) induced by the initial perturbation $\alpha \mapsto \alpha + \eta$. (For instance choosing ω equal to the r-th co-ordinate vector would correspond to estimating the change in the r-th component of the solution.) When $x(t)$ is known, we solve (21) with the *final* condition $\lambda(t_0+T) = \omega$ and note that the quantity we seek coincides with $\lambda(t_0)^{\mathsf{T}}\eta$ because, from the proposition,

$$\omega^{\mathsf{T}}\delta(t_0+T) = \lambda(t_0+T)^{\mathsf{T}}\delta(t_0+T) = \lambda(t_0)^{\mathsf{T}}\delta(t_0) = \lambda(t_0)^{\mathsf{T}}\eta.$$

The advantage of this procedure is that, as η varies, the computation of $\lambda(t_0)^{\mathsf{T}}\eta$ requires only *one* integration of (21); the computation of $\omega^{\mathsf{T}}\delta(t_0+T)$ via (20) would need a fresh integration for each new choice of η.

As an application, consider the task of computing the gradient, $\nabla_\alpha \mathcal{C}(x(t_0+T))$, of a real-valued function $\mathcal{C}$ with respect to the initial data α. We set $\omega = \nabla_x \mathcal{C}(x(t_0+T))$ in the preceding construction and successively let the r-th coordinate vector, $r = 1,\dots,d$, play the role of η to conclude that the gradient sought has the value $\lambda(t_0)$ where $\lambda(t)$ is the solution of the adjoint system with final condition $\lambda(t_0+T) = \nabla_x \mathcal{C}(x(t_0+T))$. Only one integration is required to find d derivatives $\partial/\partial\alpha^r$. The adjoint system (21) 'pulls back' gradients with respect to $x(t_0+T)$ into gradients with respect to $x(t_0)$.

3.2 The continuous problem: Lagrange multipliers

We shall also need an alternative derivation of the recipe $\nabla_\alpha \mathcal{C}(x(t_0+T)) = \lambda(t_0)$ just found. Since the use of Lagrange multipliers (see e.g. [12, Section 2.5]) in this connection (as distinct from their use in minimisation) may not be known to some readers, we give full details (see e.g. [12, Section 2.5]). Define the Lagrangian functional $\mathcal{L} = \mathcal{L}(\hat\alpha, \hat x, \hat\lambda_0, \hat\lambda)$

$$\mathcal{L} = \mathcal{C}(\hat x(t_0+T)) - \lambda_0^{\mathsf{T}}\big(\hat x(t_0) - \hat\alpha\big) - \int_{t_0}^{t_0+T} \hat\lambda(t)^{\mathsf{T}}\Big(\frac{d}{dt}\hat x(t) - f(\hat x(t),t)\Big)\,dt,$$

where, $\hat\alpha$, $\hat\lambda_0$ are arbitrary vectors, $\hat x$, $\hat\lambda$ arbitrary functions. Whenever $\hat x$ is a solution of (19) and $\hat x(t_0) = \hat\alpha$, the value of $\mathcal{L}(\hat\alpha, \hat x, \hat\lambda_0, \hat\lambda)$ coincides with $\mathcal{C}(\hat x(t_0+T))$.

If η and δ are the variations in α and $\hat x$ respectively, the variation $\delta\mathcal{L}$ of the functional is

$$\begin{aligned}\delta\mathcal{L} \;=\; & \nabla_x \mathcal{C}(\hat x(t_0+T))^{\mathsf{T}}\delta(t_0+T) - \hat\lambda_0^{\mathsf{T}}\big(\delta(t_0)-\eta\big) \\ & - \int_{t_0}^{t_0+T} \hat\lambda(t)^{\mathsf{T}}\Big(\frac{d}{dt}\delta(t) - \partial_x f(\hat x(t),t)\delta(t)\Big)\,dt;\end{aligned}$$

so that, after integration by parts,

$$\begin{aligned}\delta\mathcal{L} \;=\; & \big(\nabla_x \mathcal{C}(\hat x(t_0+T)) - \hat\lambda(t_0+T)\big)^{\mathsf{T}}\delta(t_0+T) + \hat\lambda(t_0)^{\mathsf{T}}\eta \\ & + \big(\hat\lambda(t_0) - \hat\lambda_0\big)^{\mathsf{T}}\delta(t_0) \\ & + \int_{t_0}^{t_0+T} \Big(\frac{d}{dt}\hat\lambda(t)^T\delta(t) + \hat\lambda(t)^T\partial_x f(\hat x(t),t)\delta(t)\Big)\,dt.\end{aligned}$$

We now make specific choices λ_0, λ for the (so far arbitrary) multipliers; we impose the requirements (21), $\lambda_0 = \lambda(t_0)$, and $\lambda(t_0+T) = \nabla_x \mathcal{C}(x(t_0+T))$. This choice ensures that, at x, α, the *intermediate* variation $\delta(t)$ does not contribute to $\delta\mathcal{L}$; we then have, at x, α, $\delta\mathcal{L} = \lambda(t_0)^{\mathsf{T}}\eta$ and therefore $\lambda(t_0) = \nabla_\alpha \mathcal{C}(x(t_0+T))$. The original system (19) and the initial condition may also be retrieved from the Lagrangian by making zero the variations with respect to $\hat\lambda$ and $\hat\lambda_0$ respectively.

The same approach may also be used if we wish to make things more involved and introduce the velocities $(d/dx)\hat{x} = \hat{k}$ as new arguments in the Lagrangian. Dropping the hat to simplify the notation, the Lagrangian becomes

$$\begin{aligned}\mathcal{L} &= \mathcal{C}(x(t_0+T)) - \lambda_0^{\mathsf{T}}\big(x(t_0)-\alpha\big) \\ &\quad - \int_{t_0}^{t_0+T} \lambda(t)^{\mathsf{T}}\Big(\frac{d}{dt}x(t) - k(t)\Big)\,dt \\ &\quad - \int_{t_0}^{t_0+T} \Lambda(t)^{\mathsf{T}}\Big(k(t) - f(x(t),t)\Big)\,dt. \end{aligned} \tag{24}$$

(Note that while, for extra clarity, different symbols were used before for the argument $\hat{x}$ in the Lagrangian and the solution x of the initial value problem that features in $\mathcal{C}(x(t_0+T))$, we shall not be so careful hereafter.) Taking variations and choosing the multipliers to cancel the undesired contributions to $\delta\mathcal{L}$, leads to the relations $\lambda(t_0) = \nabla_\alpha \mathcal{C}(x(t_0+T))$, $\lambda(t_0+T) = \nabla_x \mathcal{C}(x(t_0+T))$, $\lambda_0 = \lambda(t_0)$ found above and, additionally, to $\Lambda(t) \equiv \lambda(t)$ (as expected).

3.3 The discrete problem: RK integration

Let us suppose that (19) has been discretised by means of the RK scheme (1) to get, $n = 0, \ldots, N-1$,

$$x_{n+1} = x_n + h_n \sum_{i=1}^{s} b_i k_{n,i}, \tag{25}$$

$$k_{n,i} = f(X_{n,i}, t_n + c_i h_n), \quad i = 1, \ldots, s, \tag{26}$$

$$X_{n,i} = x_n + h_n \sum_{j=1}^{s} a_{ij} k_{n,j}, \quad i = 1, \ldots, s, \tag{27}$$

and that, in analogy with the preceding material, we wish to estimate the impact on x_N of a perturbation of the initial condition $x_0 = \alpha$. Linearisation of the RK equations (25)–(27) around x_n, $X_{n,i}$ shows that the perturbed RK solution $\bar{x}_n$, $n = 0, \ldots, N$, satisfies $\bar{x}_n = x_n + \delta_n + o(|\eta|)$ with

$$\delta_{n+1} = \delta_n + h_n \sum_{i=1}^{s} b_i d_{n,i}, \tag{28}$$

$$d_{n,i} = \partial_x f(X_{n,i}, t_n + c_i h_n)\Delta_{n,i}, \quad i = 1, \ldots, s, \tag{29}$$

$$\Delta_{n,i} = \delta_n + h_n \sum_{j=1}^{s} a_{ij} d_{n,j}, \quad i = 1, \ldots, s \tag{30}$$

(the vectors $d_{n,i}$ and $\Delta_{n,i}$ are the variations in the slopes $k_{n,i}$ and stages $X_{n,i}$ respectively).

On the other hand, if we regard the given differential equations (19) together with the variational equations (20) as a $2d$-dimensional system for the vector $y = [x^{\mathsf{T}}, \delta^{\mathsf{T}}]^{\mathsf{T}}$ and apply the RK scheme as in (4)–(6), we also arrive at (25)–(30). We have thus proved, as in, say, [17, Chapter VI, Lemma 4.1]:

Theorem 5 *The processes of RK discretisation and forming variational equations commute: the RK discretisation of the continuous variational equations (19)–(20) yields the variational equations (25)–(30) for the RK discretisation.*

The situation for the adjoint equations is not quite as neat (cf. [34]). In order to find the discrete sensitivity $\omega^{\mathsf{T}}\delta_N$ we would like to numerically integrate (21) with final condition $\lambda_N = \omega$ in such a way that (cf. (23))

$$\lambda_N^{\mathsf{T}}\delta_N = \lambda_0^{\mathsf{T}}\delta_0. \tag{31}$$

Although in actual computation the approximations λ_n are to be found without using the equations (28)–(30) for δ_n (this is the whole point behind the use of adjoints), let us consider for a moment the $3d$-dimensional system (19)–(21) for the extended vector $y = [x^{\mathsf{T}}, \delta^{\mathsf{T}}, \lambda^{\mathsf{T}}]^{\mathsf{T}}$. Then the condition (31) demands that we integrate this large system in such a way as to *exactly* preserve the invariant $I(y(t), y(t)) = \lambda(t)^{\mathsf{T}}\delta(t)$ in (22). According to Theorem 1, we may achieve this goal by using the RK scheme (1) *provided that it is symplectic*. This results in the relations (25)–(30) in tandem with $(n = 0, \ldots, N-1)$:

$$\lambda_{n+1} = \lambda_n + h_n \sum_{i=1}^{s} b_i \ell_{n,i}, \tag{32}$$

$$\ell_{n,i} = -\partial_x f(X_{n,i}, t_n + c_i h_n)^{\mathsf{T}} \Lambda_{n,i}, \quad i = 1, \ldots, s, \tag{33}$$

$$\Lambda_{n,i} = \lambda_n + h_n \sum_{j=1}^{s} a_{ij} \ell_{n,j}, \quad i = 1, \ldots, s. \tag{34}$$

Let us summarise the preceding discussion:

Theorem 6 *Assume that the $3d$-dimensional system (19)–(21) is discretised by a* symplectic *RK scheme (1). Then for any RK solution (31) holds. In particular, for the RK solution specified by the initial condition $x_0 = \alpha$, $\delta_0 = \eta$ together with the final condition $\lambda_N = \omega$,*

$$\omega^{\mathsf{T}}\delta_N = \lambda_0^{\mathsf{T}}\eta.$$

For a non-symplectic RK scheme of order ρ, $\omega^{\mathsf{T}}\delta_N$ and $\lambda_0^{\mathsf{T}}\eta$ are approximations of order ρ to their continuous counterparts $\omega^{\mathsf{T}}\delta(t_0 + T)$ and $\lambda(t_0)^{\mathsf{T}}\eta$ respectively and therefore $\lambda_0^{\mathsf{T}}\eta$ will be a $\mathcal{O}(h^\rho)$ approximation to the true sensitivity $\omega^{\mathsf{T}}\delta_N$ of the discrete solution.

In practice, the variational equations (20) do not need to be integrated. We successively find $x_0, x_1, \ldots, x_N$ via (25)–(27) and, once these are available, we set $\lambda_N = \omega$, and compute $\lambda_{N-1}, \ldots, \lambda_0$ from (32)–(34) taken in the order $n = N-1, N-2, \ldots, 0$. For this reason, it may be advisable to rewrite (32)–(34) in the following 'reflected' form (see Appendix B) that emphasises that the approximation λ_n at t_n is to be found

from the approximation λ_{n+1} at t_{n+1}:

$$\lambda_n = \lambda_{n+1} + (-h_n)\sum_{i=1}^{s} b_i \ell_{n,i}, \tag{35}$$

$$\ell_{n,i} = -\partial_x f(X_{n,i}, t_{n+1} + (1-c_i)(-h_n))^{\mathsf{T}} \Lambda_{n,i}, \quad i = 1, \ldots, s, \tag{36}$$

$$\Lambda_{n,i} = \lambda_{n+1} + (-h_n)\sum_{j=1}^{s} (b_j - a_{ij})\ell_{n,j}, \quad i = 1, \ldots, s. \tag{37}$$

In analogy to the continuous case, for a symplectic RK discretisation, the gradient $\nabla_\alpha \mathcal{C}(x_N)$ may be computed by finding λ_0 from the recursion (32)–(34) (or (35)–(37)) with $\lambda_N = \nabla_x \mathcal{C}(x_N)$.

3.4 The discrete problem: PRK integration

Theorem 6 may be generalised easily with the help of Theorem 3. Hereafter it is understood that when using the PRK scheme the x, δ equations are integrated with the set of coefficients (1) (so that the δ_n are exactly the variations in x_n) and the λ equations with the set of coefficients (9). In other words, the system is partitioned as $q = [x^{\mathsf{T}}, \delta^{\mathsf{T}}]^{\mathsf{T}}$, $p = \lambda$. This approach leads to (25)–(30) supplemented by the relations obtained by replacing the lower case coefficients a_{ij}, b_i, c_i in (32)–(34) by their upper case counterparts:

$$\lambda_{n+1} = \lambda_n + h_n \sum_{i=1}^{s} B_i \ell_{n,i}, \tag{38}$$

$$\ell_{n,i} = -\partial_x f(X_{n,i}, t_n + C_i h_n)^{\mathsf{T}} \Lambda_{n,i}, \quad i = 1, \ldots, s, \tag{39}$$

$$\Lambda_{n,i} = \lambda_n + h_n \sum_{j=1}^{s} A_{ij} \ell_{n,j}, \quad i = 1, \ldots, s. \tag{40}$$

The generalisation of Theorem 6 is:

Theorem 7 *Assume that the* $3d$*-dimensional system (19)–(21) is discretised by a* symplectic *PRK scheme (1), (9). Then (31) holds for any PRK solution. In particular, for the PRK solution specified by the initial condition* $x_0 = \alpha$, $\delta_0 = \eta$ *together with the final condition* $\lambda_N = \omega$,

$$\omega^{\mathsf{T}} \delta_N = \lambda_0^{\mathsf{T}} \eta.$$

Once more, for a symplectic PRK discretisation, the gradient $\nabla_\alpha \mathcal{C}(x_N)$ coincides with λ_0 if $\lambda_N = \nabla_x \mathcal{C}(x_N)$. For a non-symplectic discretisation of the adjoint equations, λ_0 is a only an approximation to $\nabla_\alpha \mathcal{C}(x_N)$. For this reason non-symplectic PRK discretisations cannot be implied by the direct differentiation procedure described in Section 3.2.

How do we compute *exactly* (i.e. up to round-off) the sensitivity $\omega^{\mathsf{T}} \delta_N$ with the help of the adjoint system when the x integration has been performed with a non-symplectic RK scheme (1) and Theorem 6 cannot be invoked? Theorem 7 suggests the way. For

simplicity we only look at the case where in (1) none of the weights b_i, $i = 1, \ldots, s$, vanishes (for the general situation see Appendix A). From the coefficients in (1) we compute a new set

$$A_{ji} = b_i - b_i a_{ij}/b_j, \quad i, j = 1, \ldots, s, \quad B_i = b_i, \quad C_i = c_i \quad i = 1, \ldots, s. \quad (41)$$

In view of (16)–(17), we now have a PRK scheme for the discretisation of (19)–(21) and Theorem 7 applies. If (1) is explicit, the computations required to descend from λ_N to λ_0 are also explicit. Here is the simplest example. Assume that the x equations are integrated with the explicit Euler rule ($s = 1$, $a_{11} = 0$, $b_1 = 1$, $c_1 = 0$). With that choice, $X_{n,1} = x_n$ and

$$x_{n+1} = x_n + h_n f(x_n, t_n).$$

The trick just described yields $A_{11} = 1$, $B_1 = 1$, $C_1 = 0$. Accordingly, the stage $\Lambda_{n,1}$ coincides with λ_{n+1} and using (11) we see that the required λ integrator is:

$$\lambda_{n+1} = \lambda_n - h_n \partial_x f(x_n, t_n)^{\mathsf{T}} \lambda_{n+1}.$$

Obviously this is *not* the explicit Euler rule, because λ in the right-hand side appears at time t_{n+1}. And, unless the problem is autonomous, it is not the implicit Euler rule either because t is evaluated at the retarded time t_n. (For RK enthusiasts only: the coefficients $A_{11} = 1$, $B_1 = 1$, $C_1 = 0$ correspond to the Radau IA method of one stage introduced by Ehle, [20, Section IV.5].)

In the particular situation where the x integration has been performed by a symplectic RK method (symplectic RK methods possess non-vanishing weights [31], Section 8.2), the recipe (41) will lead to $A_{ij} = a_{ij}$ and the resulting PRK method will coincide with the original RK method. In the general case, for (31) to hold, *the adjoint equations for λ have to be integrated with coefficients different from those used for the original equations for x.*

There are hidden difficulties with the use of this recipe. When stability is an issue, as in stiff problems or time-discretisations of partial differential equations, it is necessary to investigate carefully the stability behaviour of the λ integration [34]. On the other hand, and as noted before, the order of accuracy of the overall PRK, x, λ, integrator may be lower than the order of the RK method (1) for x we started with. When investigating the order of the overall PRK method we have to take into account that the right-hand side of (19) is independent of λ and the right-hand side of (21) is linear in λ. These features imply that many elementary differentials vanish and that accordingly it is not necessary to impose the order conditions associated with them. Furthermore we have to take into account the reduction in the number of independent order conditions implied by symplecticness.

3.5 The discrete problem: automatic differentiation

According to the preceding discussion, for any RK integration of (19) with nonzero weights, it is possible to find $\nabla_\alpha \mathcal{C}(x_N)$ by means of an integration of the adjoint equations with the coefficients (41). It is however clear that it is also perfectly possible to compute $\nabla_\alpha \mathcal{C}(x_N)$ by repeatedly using the chain rule in (25)–(27), something that

we shall perform presently. Since $\mathcal{C}$ is scalar and $\alpha \in \mathbb{R}^d$, where d is possibly large, reverse accumulation [13] is to be preferred and this may be performed with the help of Lagrange multipliers as in Section 3.2.

We shall need the following auxiliary result:

Lemma 2 *Suppose that the mapping $\Omega : \mathbb{R}^{d+d'} \to \mathbb{R}^{d'}$ is such that the Jacobian matrix $\partial_\gamma \Omega$ is invertible at a point $(\alpha_0, \gamma_0) \in \mathbb{R}^d \times \mathbb{R}^{d'}$, so that in the neighborhood of α_0, the equation $\Omega(\alpha, \gamma) = 0$ defines γ as a function of α. Consider a real-valued function in $\mathbb{R}^d$ of the form $\psi(\alpha) = \Psi(\alpha, \gamma(\alpha))$, for some $\Psi : \mathbb{R}^{d+d'} \to \mathbb{R}$. There exists a unique vector $\lambda_0 \in \mathbb{R}^{d'}$ such that (superscripts denote components):*

$$\begin{aligned} \nabla_\alpha \psi|_{\alpha_0} &= \nabla_\alpha \Psi|_{(\alpha_0,\gamma_0)} + \sum_{r=1}^{d'} \lambda_0^r \nabla_\alpha \Omega^r|_{(\alpha_0,\gamma_0)}, \\ 0 &= \nabla_\gamma \Psi|_{(\alpha_0,\gamma_0)} + \sum_{r=1}^{d'} \lambda_0^r \nabla_\gamma \Omega^r|_{(\alpha_0,\gamma_0)}. \end{aligned}$$

Proof: The second requirement may be rewritten as

$$(\partial_\gamma \Omega)^{\mathsf{T}} \lambda_0 = -\nabla_\gamma \Psi, \tag{42}$$

with the matrix and right-hand side evaluated at α_0, γ_0. This is a linear system that uniquely defines λ_0. To check that the vector λ_0 we have just found satisfies the first requirement, we use the chain rule

$$\partial_\alpha \psi|_\alpha = \partial_\alpha \Psi|_{(\alpha,\gamma(\alpha))} + \partial_\gamma \Psi|_{(\alpha,\gamma(\alpha))} \partial_\alpha \gamma|_\alpha,$$

differentiate $\Omega(\alpha, \gamma(\alpha)) = 0$ to get

$$\partial_\alpha \Omega|_{(\alpha,\gamma(\alpha))} + \partial_\gamma \Omega|_{(\alpha,\gamma(\alpha))} \partial_\alpha \gamma|_\alpha = 0,$$

evaluate at α_0, and eliminate $\partial_\alpha \gamma|_{\alpha_0}$. □

It is useful to rephrase the lemma by introducing the Lagrangian

$$\mathcal{L}(\alpha, \gamma, \lambda) = \Psi(\alpha, \gamma) + \lambda^T \Omega(\alpha, \gamma).$$

so that the relation $\Omega(\alpha_0, \gamma_0) = 0$ and the equation (42) that defines de multiplier are respectively

$$\nabla_\lambda \mathcal{L}(\alpha, \gamma, \lambda)|_{(\alpha_0,\gamma_0,\lambda_0)} = 0, \qquad \nabla_\gamma \mathcal{L}(\alpha, \gamma, \lambda)|_{(\alpha_0,\gamma_0,\lambda_0)} = 0,$$

while the gradient we seek is computed as

$$\nabla_\alpha \psi|_{\alpha_0} = \nabla_\alpha \mathcal{L}(\alpha, \gamma, \lambda)|_{(\alpha_0,\gamma_0,\lambda_0)}.$$

Note that these developments mimic the material in Section 3.2, with γ playing the part of $\hat{x}$, γ_0 the part of x, etc.

In numerical differentiation, ψ is the function whose gradient is to be evaluated, the components of α are the independent variables, and the components of γ represent

intermediate stages towards the computation of ψ. (For instance, in the simple case ($d = 1$) where $\psi(\alpha) = \alpha\sqrt{1 + \alpha\exp(\alpha)\cos(\exp(\alpha))}$, we may set the constraints $\Omega^1 = \gamma^1 - \exp(\alpha) = 0$, $\Omega^2 = \gamma^2 - \cos(\gamma^1) = 0$, $\Omega^3 = \gamma^3 - \alpha\gamma^1\gamma^2 = 0$, $\Omega^4 = \gamma^4 - \sqrt{1 + \gamma^3}$, $\psi = \alpha\gamma^4$.) The interpretation of the γ^r as successive stages implies that, in practice, Ω will possess a lower triangular structure: Ω^r will only involve $\gamma^1,\ldots,\gamma^r$. The evaluation of ψ successively finds the numerical values of $\gamma^1,\ldots,\gamma^{d'}$ in a forward fashion. The numerical values of the components λ_0^r, are then found by *backward* substitution in the upper-triangular linear system (42) and finally the lemma yields the required value of the gradient. If Ψ and Ω have been judiciously chosen, then the mappings $\nabla_\alpha\Psi$, $\nabla_\gamma\Psi$, $\nabla_\alpha\Omega^r$, $\nabla_\alpha\Omega^r$ required to compute the gradient will have simple analytic expressions, easily derived by a human or by a computer programme.

We now apply this technique to find $\nabla_\alpha\mathcal{C}(x_N)$. In (25)–(27) we think that (the components of) x_n, $n = 0, ..., N$, and $k_{n,i}$, $n = 0, \ldots, N-1$, $i = 1, \ldots, s$, play the role of (the components of) γ and introduce the Lagrangian

$$\begin{aligned}\mathcal{C}(x_N) - \lambda_0^{\mathsf{T}}(x_0 - \alpha) - \textstyle\sum_{n=0}^{N-1} h_n\lambda_{n+1}^{\mathsf{T}}\Big[\frac{1}{h_n}(x_{n+1} - x_n) - \sum_{i=1}^{s} b_i k_{n,i}\Big] \\ - \textstyle\sum_{n=0}^{N-1} h_n \sum_{i=1}^{s} b_i\Lambda_{n,i}^{\mathsf{T}}\Big[k_{n,i} - f(X_{n,i}, t_n + c_i h_n)\Big],\end{aligned} \tag{43}$$

where we understand that the stage vectors $X_{n,i}$ have been expressed in terms of the x_n and $k_{n,i}$ by means of (27). Clearly this discrete Lagrangian is the natural RK approximation to (24).

A straightforward application of Lemma 2 now directly yields the following result, where we note that the hypothesis $b_i \neq 0$, $i = 1, \ldots, s$, is natural because, when, say, $b_1 = 0$, the Lagrangian (43) does not incorporate the constraint $k_{n,1} = f(X_{n,1}, t_n + c_1 h_n)$. (The case of zero weights is considered in Appendix A.)

Theorem 8 *Consider the RK equations (25)–(27), with $b_i \neq 0$, $i = 1, \ldots, s$. The computation of $\nabla_\alpha\mathcal{C}(x_N)$ based on the use of Lemma 2 with Lagrangian (43) leads to the relations (38)–(40), with the coefficients A_{ij}, B_i, C_i given by (41), together with $\nabla_x\mathcal{C}(x_N) = \lambda_N$, $\nabla_\alpha\mathcal{C}(x_N) = \lambda_0$.*

Note that, in the situation of the theorem, λ_N, λ_{N-1}, λ_{N-2}, ... successively yield $\nabla_{x_N}\mathcal{C}(x_N)$, $\nabla_{x_{N-1}}\mathcal{C}(x_N)$, $\nabla_{x_{N-2}}\mathcal{C}(x_N)$, ... It is well known that the reverse mode of differentiation implies an integration of the adjoint equations. The theorem shows additionally that, for an RK computation of x, the implied adjoint equation integration is such that the x, λ system is discretised with a *symplectic* PRK method. In this way the chain rule provided us with symplectic integration *before the latter was invented.*

A further remark: the use of the chain rule with forward accumulation implies an RK integration of the variational equations (20) with the original RK coefficients (1). In agreement with a previous discussion, the forward mode is more expensive; each partial derivative $\partial/\partial\alpha^r$, $r = 1, \ldots, d$, in the gradient requires a separate integration.

4 A simple optimal control problem

We explore next the role of symplectic methods when integrating the differential equations that arise in some optimal control problems [35], [38], [39]. In this section we

look at the simplest case, where the developments are very similar to those just considered; more general problems are treated in the next.

4.1 The continuous problem

Consider now the d-dimensional system

$$\frac{d}{dt}x = f(x,u,t), \tag{44}$$

where x is the state vector and u a ν-dimensional vector of controls. Our aim is to find functions $x(t)$ and $u(t)$, subject to (44) and the initial condition $x(t_0) = \alpha \in \mathbb{R}^d$, so as to minimise a given cost function $\mathcal{C}(x(t_0+T))$.

The variational equation is (cf. (20))

$$\frac{d}{dt}\delta = \partial_x f(x(t),u(t),t)\,\delta + \partial_u f(x(t),u(t),t)\,\zeta, \tag{45}$$

where ∂_u is the Jacobian matrix of f with respect to u and ζ denotes the variation in u, see e.g. [35, Section 2.8], [38, Section 5.1]. Now $\delta(t_0) = 0$, as $x(t_0)$ remains nailed down at α.

An adjoint system (cf. (21))

$$\frac{d}{dt}\lambda = -\partial_x f(x(t),u(t),t)^{\mathsf{T}}\,\lambda, \tag{46}$$

and constraints

$$\partial_u f(x(t),u(t),t)^{\mathsf{T}}\lambda(t) = 0, \tag{47}$$

are introduced, see e.g. [35, Section 9.2], in such a way that the following proposition holds (cf. Proposition 1). Again the key is to ensure the validity of the conservation property (23).

Proposition 2 *For each choice of vectors x, u, δ, ζ, λ and real t:*

$$\Big(-\partial_x f(x,u,t)^{\mathsf{T}}\,\lambda\Big)^{\mathsf{T}}\delta + \lambda^{\mathsf{T}}\Big(\partial_x f(x,u,t)\delta + \partial_u f(x,u)\zeta\Big) = 0. \tag{48}$$

Therefore if $\delta(t)$, $\lambda(t)$, $\zeta(t)$ satisfy (45)–(47), then (22)–(23) hold.

The use of the proposition is as follows. We solve the two-point boundary problem given by the states+costates system (44), (46)–(47) with initial/final conditions

$$x(t_0) = \alpha, \qquad \lambda(t_0+T) = \nabla\mathcal{C}(x(t_0+T)). \tag{49}$$

Then, the variation $\delta(t_0+T)$ at the end of the interval is orthogonal to the gradient of the cost since, from (23),

$$\nabla\mathcal{C}(x(t_0+T))^{\mathsf{T}}\delta(t_0+T) = \lambda(t_0+T)^{\mathsf{T}}\delta(t_0+T) = \lambda(t_0)^{\mathsf{T}}\delta(t_0) = 0. \tag{50}$$

This of course means that any solution $[x(t)^\mathsf{T}, \lambda(t)^\mathsf{T}, u(t)^\mathsf{T}]^\mathsf{T}$ of the boundary-value problem satisfies the first-order necessary condition for $\mathcal{C}$ to attain a minimum. As in sensitivity analyses, the costates λ may be interpreted as *Lagrange multipliers*.

It is customary to introduce the function $H(x, \lambda, u, t) = \lambda^\mathsf{T} f(x, u, t)$ (pseudo-Hamiltonian) so that (44), (46)–(47) take the very symmetric form

$$\frac{d}{dt}x = \nabla_\lambda H, \quad \frac{d}{dt}\lambda = -\nabla_x H, \quad \nabla_u H = 0. \tag{51}$$

4.2 The discrete problem: indirect approach

In the indirect approach, approximations to the optimal states, costates and controls are obtained by discretisation of the boundary value problem (44), (46)–(47), (49). Note that we have to tackle a *differential-algebraic* system [20, Chapter VI.1], with the controls being algebraic variables as $(d/dt)u$ does not feature in any of the equations (44), (46)–(47). Under suitable technical assumptions (invertibility of the second derivative of H with respect to u), the system is of *index one*. This means that the constraints (47) may be used to express, locally around the solution of interest, the algebraic variables as functions of the differential variables, $u = \Phi(x, \lambda, t)$. (When applying the implicit function theorem, the relevant Jacobian matrix is the Hessian $\partial_{uu}H$ and this will generically be positive definite, if Pontryagin's principle [38, Section 7.2] holds so that $H(x, \lambda, \cdot, t)$ is minimised by $\Phi(x, \lambda, t)$.) For a system of index one we may think that the right-hand sides of (44) and (46) have been written as functions of x, λ and t by setting $u = \Phi(x, \lambda, t)$, thus transforming the differential-algebraic system into a system of ordinary differential equations. In fact the transformed system is the canonical Hamiltonian system with Hamiltonian function $\mathcal{H}(x, \lambda, t) = H(x, \lambda, \Phi(x, \lambda, t), t)$, because the chain rule and $\nabla_u H = 0$ imply that, in (51), $\nabla_x H(x, \lambda, u, t) = \nabla_x \mathcal{H}(x, \lambda, t)$ and $\nabla_x H(x, \lambda, u, t) = \nabla_x \mathcal{H}(x, \lambda, t)$. This Hamiltonian system may be discretised with the PRK scheme (1), (9). (Recall that RK schemes are included as particular cases where both sets of coefficients just coincide.) The discrete equations are solved to find the approximations x_n and λ_n to $x(t_n)$, $\lambda(t_n)$ and finally the approximations to the controls are retrieved as $u_n = \Phi(x_n, \lambda_n, t_n)$.

The analytic expression of the implicit function Φ will in general not be available, so that it will not be possible to find $\mathcal{H}$ explicitly. This is not a hindrance: the approximations x_n, λ_n, u_n that one would get by a PRK integration of the Hamiltonian system may be found in practice as solutions of the set of equations (52)–(59) below, obtained by direct discretisation of the differential-algebraic format (44), (46)–(47). The equivalence between the two approaches, differential and differential-algebraic is seen by eliminating the controls from (52)–(59), see [20, Chapter VI.1].

The discrete equations are ($n = 0, \ldots, N-1$):

$$x_{n+1} = x_n + h_n \sum_{i=1}^{s} b_i k_{n,i}, \tag{52}$$

$$k_{n,i} = f(X_{n,i}, U_{n,i}, t_n + c_i h_n), \quad i = 1, \ldots, s, \tag{53}$$

$$X_{n,i} = x_n + h_n \sum_{j=1}^{s} a_{ij} k_{n,j}, \quad i = 1, \ldots, s, \tag{54}$$

$$\lambda_{n+1} = \lambda_n + h_n \sum_{i=1}^{s} B_i \ell_{n,i}, \tag{55}$$

$$\ell_{n,i} = -\partial_x f(X_{n,i}, U_{n,i}, t_n + C_i h_n)^{\mathsf{T}} \Lambda_{n,i}, \quad i = 1, \ldots, s, \tag{56}$$

$$\Lambda_{n,i} = \lambda_n + h_n \sum_{j=1}^{s} A_{ij} \ell_{n,j}, \quad i = 1, \ldots, s, \tag{57}$$

$$\partial_u f(X_{n,i}, U_{n,i}, t_n + C_i h_n)^{\mathsf{T}} \Lambda_{n,i} = 0, \quad i = 1, \ldots, s, \tag{58}$$

together with ($n = 0, \ldots, N$)

$$\partial_u f(x_n, u_n, t_n)^{\mathsf{T}} \lambda_n = 0, \tag{59}$$

and the boundary conditions $x_0 = \alpha$, $\lambda_N = \nabla C(x_N)$ from (49).

What is the accuracy of this technique? We encounter the same difficulty we found in the preceding section: relevant here is the order of the overall PRK scheme rather than the (possibly higher) order of the RK coefficients (1) used for the state variables. In the preceding section the approximations x_n are found independently of the λ_n and, accordingly, the possible order reduction does not affect them. In the optimal control problem, states and costates are coupled and any order reduction will harm both of them. This was first noted by Hager who also provided relevant counterexamples, see [15, Table 3]. Hager (Proposition 6.1) also shows that there is no order reduction for explicit, forth order RK schemes with positive weights.

The obvious analogue of Theorem 5 holds: the variations δ_n in the discrete solution x_n satisfy the equations that result from discretising (45) with the coeffients (1). These equations are (28) and (30) where now

$$k_{n,i} = \partial_x f(X_{n,i}, U_{n,i}, t_n + c_i h_n)\, \Delta_{n,i} + \partial_u f(X_{n,i}, U_{n,i}, t_n + c_i h_n)\, Z_{n,i}, \tag{60}$$

($\Delta_{n,i}$, $Z_{n,i}$ are the stages associated with the variables δ and ζ).

Assume next that the PRK is *symplectic*. Recall that symplecticness may be the result of choosing the RK coefficients (1) ($b_i \neq 0$, $i = 1, \ldots, s$) for the state variables and retrieving from (41) the coefficients (9) for the integration of the adjoint system. The symplecticness of the integrator makes it possible to formulate a discrete analogue of Proposition 2.

Theorem 9 *Assume that x_n, λ_n, u_n, $n = 0, \ldots, N$, satisfy the equations (52)–(59) arising from the application of a* symplectic *PRK method and that, furthermore, δ_n, $n = 0, \ldots, N$, $\delta_0 = 0$, are the variations in x_n. Then, for $n = 0, \ldots, N-1$,*

$$\lambda_{n+1}^{\mathsf{T}} \delta_{n+1} = \lambda_n^{\mathsf{T}} \delta_n.$$

The PRK scheme may be a symplectic RK scheme or the result of choosing freely the RK coefficients (1), $b_i \neq 0$, $i = 1, \ldots, s$, for the states and then using (41) to determine the coefficients for the integration of the costates.

Proof: Use Lemma 1 with $S(q, p) = \lambda^{\mathsf{T}} \delta$. This results in

$$\lambda_{n+1}^{\mathsf{T}} \delta_{n+1} - \lambda_n^{\mathsf{T}} \delta_n = h_n \sum_i b_i (\Lambda_{n,i}^{\mathsf{T}} k_{n,i} + \ell_{n,i}^{\mathsf{T}} \Delta_{n,i})$$

where $k_{n,i}$ and $\ell_{n,i}$ come from (60) and (56) respectively. According to (48), each of the terms being summed vanishes. □

When the boundary conditions (49) are imposed,

$$\nabla \mathcal{C}(x_N)^{\mathsf{T}} \delta_N = \lambda_N^{\mathsf{T}} \delta_N = \lambda_0^{\mathsf{T}} \delta_0 = 0,$$

which means that the discrete solution satisfies the first-order necessary conditions for $\mathcal{C}(x_N)$ to achieve a minimum subject to the constrains (52)–(54) and $x_0 = \alpha$. In this way we have proved that *symplectic discretisation* commutes [26] with the process of forming necessary conditions for minimisation:

Theorem 10 *A solution $\{x_n\}$, $\{\lambda_n\}$, $\{u_n\}$ of the equations (52)–(59) arising from discretising with a* symplectic *PRK integrator the necessary conditions for the continuous optimal control problem satisfies the necessary conditions for $\mathcal{C}(x_N)$ to achieve a minimum subject to the discrete constraints (52)–(54) and $x_0 = \alpha$. The PRK scheme may be a symplectic RK scheme or the result of choosing freely the RK coefficients (1), $b_i \neq 0$, $i = 1, \ldots, s$, for the states and then using (41) to determine the coefficients for the integration of the costates.*

When the states+costates system is integrated by means of a non-symplectic PRK, x_N will not satisfy the necessary conditions for $\mathcal{C}$ to be minimised subject to the constraints (52)–(54) and $x_0 = \alpha$. Therefore non-symplectric PRK discretisations *cannot* be obtained via the direct approach considered next.

4.3 The discrete problem: direct approach

The direct approach (see e.g. [38, Chapter 9]) based on RK discretisation begins by applying the scheme (1) to the differential equation (44) to get (52)–(54). Then, these equations and $x_0 = \alpha$ are seen as constraints of a finite-dimensional optimisation problem for the minimisation of $\mathcal{C}(x_N)$.

We use the standard method of Lagrange multipliers based on the Lagrangian in (43), trivially adapted to the present circumstances by letting f depend on the controls. The method leads in a straightforward way to the following result, first proved by Hager [15], see also [4]. (However [15] does not point out that the relations (41) correspond to symplecticness. Furthermore [15] and [4] do not use a discrete Lagrangian obtained by discretisation of the continuous Lagrangian.)

Theorem 11 *The first-order necessary conditions for the minimisation of $\mathcal{C}(x_N)$ subject to $x_0 = \alpha$ and (52)–(54), $b_i \neq 0$, $i = 1, \ldots, s$, are $x_0 = \alpha$, $\nabla \mathcal{C}(x_N) = \lambda_N$ together with (52)–(58), with the coefficients A_{ij}, B_i, C_i given by (41).*

In other words, when the direct approach is used, we arrive at *exactly the same set of equations* for x_n, λ_n, $X_{n,i}$, $\Lambda_{n,i}$, $U_{n,i}$ we obtained, with the help of RK technology, via the indirect approach in Theorem 10. Let us observe that the direct approach does not provide 'natural' approximations u_n to $u(t_n)$. Hager [15] suggests to define u_n by locally minimising $H(x_n, \lambda_n, u, t_n)$ which leads to (59). He also notes ([15], Table 4) that the order of convergence of the control stages $U_{n,i}$ may be lower than that in u_n, something that it is not suprising at all: typically, internal stages are more inaccurate than end-of-step approximations. We remark that, in the direct approach and once the RK method for x has been chosen, the minimisation of $\mathcal{C}$ implicitly provides the 'right' coefficients A_{ij}, B_i, C_i to be used in the integration of the costates in order to ensure symplecticness of the overall PRK integrator. In the indirect approach those coefficients have to be determined by using the relations (16)–(17) and Theorem 3.

While the direct and indirect approaches may be seen as mathematically equivalent here, both have their own interest. The direct approach suggests to solve the discrete PRK equations with the help of optimisation techniques and these may be an efficient choice in practice. On the other hand, the direct approach 'hides' the PRK integration of the costates, a fact that may lead to the wrong impression that the order of accuracy of the overall procedure coincides with the order of the RK scheme used to discretise the differential constraint (44). This was emphasised in [15], where the order of the PRK method (1), (9), (41) is called the order of the RK method (1) *for optimal control problems*. A discussion of the advantages of the direct and indirect approaches is not within our scope here, see e.g. [38, Chapter 9], [10].

5 Some extensions

Let us look at some extensions of the preceding material.

5.1 Generalised conservation

Here are simple generalisations of Theorems 1 and 3. Only Theorem 13 will be proved; the other proof is of course very similar. For a typical application of Theorem 12, take the case where y comprises positions and velocities of a mechanical system, I is the kinetic energy and φ is the work of the forces.

Theorem 12 *Assume that, for the differential system (2), there exist a real-valued bilinear mapping I in $\mathbb{R}^D \times \mathbb{R}^D$ and a real-valued function φ in $\mathbb{R}^D$ such that, for each solution $y(t)$*

$$\frac{d}{dt} I(y(t), y(t)) = \varphi(y(t))$$

and, therefore,

$$I(y(t_0 + T), y(t_0 + T)) - I(y(t_0), y(t_0)) = \int_{t_0}^{t_0+T} \varphi(y(t))\, dt.$$

If the system is integrated by means of a symplectic RK scheme as in (4)–(6), then

$$I(y_N, y_N) - I(y_0, y_0) = \sum_{n=0}^{N-1} h_n \sum_{i=1}^{s} b_i \, \varphi(Y_{n,i}).$$

Note that the last sum, based on the RK quadrature weights b_i and in the approximation $y(t_n + c_i h_n) \approx Y_{n,i}$, is the 'natural' RK discretisation of the corresponding integral.

Theorem 13 *Assume that, for the partitioned system (10), there exist a real-valued bilinear map S in $\mathbb{R}^{D-d} \times \mathbb{R}^d$ and a real-valued function φ in $\mathbb{R}^{D-d} \times \mathbb{R}^d$, such that for each solution*

$$\frac{d}{dt} S(q(t), p(t)) = \varphi(q(t), p(t))$$

and, therefore,

$$S(q(t_0 + T), p(t_0 + T)) - S(q(t_0), p(t_0)) = \int_{t_0}^{t_0+T} \varphi(q(t), p(t)) \, dt.$$

If the system is integrated by means of a symplectic PRK scheme as in (11)–(13), then

$$S(q_N, p_N) - S(q_0, p_0) = \sum_{n=0}^{N-1} h_n \sum_{i=1}^{s} b_i \, \varphi(Q_{n,i}, P_{n,i}).$$

Proof: Use Lemma 1 and note that, under the present hypotheses,

$$S(k_{n,i}, P_{n,i}) + S(Q_{n,i}, \ell_{n,i}) = \varphi(Q_{n,i}, P_{n,i}),$$

because $S(f(q,p,t),p) + S(q, g(q,p,t)) \equiv \phi(q,p)$ (cf. the proof of Theorem 3). □

5.2 Other optimal control problems

Consider first the situation in Section 4, but assume that the value $x(t_0)$ is not prescribed. Then $\delta(t_0)$ is free and for (50) to hold it is necessary to impose the condition $\lambda(t_0) = 0$. This replaces in (49) the initial condition $x(t_0) = \alpha$. The results in Section 4 are valid in this setting after the obvious modifications.

We next look at the case where (44) and $x(0) = \alpha$ are imposed, but the cost function is given by

$$\mathcal{C}(x(t_0 + T)) + \int_{t_0}^{t_0+T} \mathcal{D}(x(t), u(t), t) \, dt \tag{61}$$

(this is sometimes called a Mayer-Lagrange cost [38], as distinct from the Mayer cost $\mathcal{C}(x(t_0 + T))$ envisaged before). The adjoint system and constraints are, respectively,

$$\begin{aligned} &\frac{d}{dt}\lambda = -\partial_x f(x,u,t)^{\mathsf{T}} \, \lambda - \nabla_x \mathcal{D}(x,u,t), \\ &\partial_u f(x,u,t)^{\mathsf{T}} \lambda + \nabla_u \mathcal{D}(x,u,t) = 0. \end{aligned}$$

These are of the form in (51) for the pseudo-Hamiltonian $H = \lambda^T f + \mathcal{D}$.

The conservation property (23) is replaced by the generalised conservation

$$\lambda(t_0+T)^{\mathsf{T}}\delta(t_0+T) - \lambda(t_0)^{\mathsf{T}}\delta(t_0) + \int_{t_0}^{t_0+T} \Big(\nabla_x \mathcal{D}(x(t),u(t),t)^{\mathsf{T}}\delta(t) + \nabla_u \mathcal{D}(x(t),u(t),t)^{\mathsf{T}}\zeta(t)\Big)\, dt = 0,$$

which holds for arbitrary $\delta(t)$, $\lambda(t)$ satisfying the variational equations (45), the adjoint system and the constraints. After setting $\delta(t_0) = 0$ and $\lambda(t_0+T) = \nabla\mathcal{C}(x(t_0+T))$, we recover the first-order necessary conditions for the minimisation of the cost.

For a symplectic PRK discretisation of the algebraic-differential system, Lemma 1 may be used, just as in the proof of Theorem 13, to show (the notation should be clear by now):

$$\lambda_N^{\mathsf{T}}\delta_N - \lambda_0^{\mathsf{T}}\delta_0 + \sum_{n=0}^{N-1} h_n \sum_{i=1}^{s} b_i \Big(\nabla_x \mathcal{D}(X_{n,i},U_{n,i},t_n+c_ih_n)^{\mathsf{T}}\Delta_{n,i} + \nabla_u \mathcal{D}(X_{n,i},U_{n,i},t_n+c_ih_n)^{\mathsf{T}} Z_{n,i}\Big) = 0.$$

By setting $\lambda_N = \nabla\mathcal{C}(x_N)$ and $\delta_0 = 0$, we get the necessary condition for the discrete solution to minimise the discretised cost

$$\mathcal{C}(x_N) + \sum_{n=0}^{N-1} h_n \sum_{i=1}^{s} b_i \mathcal{D}(X_{n,i},U_{n,i}).$$

Therefore also in this case, results corresponding to Theorems 10 and 11 hold for a symplectic PRK discretisation.

It is of course possible to combine the cost (61) with alternative boundary specifications. If $x(t_0)$ is not prescribed, then we have to impose $\lambda(t_0) = 0$, as pointed out above. If both $x(t_0) = \alpha$ and $x(t_0+T) = \beta$ are imposed (in which case the term $\mathcal{C}(x(t_0+T))$ may be dropped from the cost), then $\lambda(t_0)$ and $\lambda(t_0+T)$ are both free.

5.3 Constrained controls

Let us go back once more to the problem in Section 4 and suppose that the controls u are constrained so that, for each t, it is demanded that $u(t) \in U$, where U is a given closed, convex subset of $\mathbb{R}^\nu$. Then (see e.g. [15]), the constraint (47) on λ has to be replaced by

$$u(t) \in U, \qquad -\partial_u f(x(t),u(t),t)^{\mathsf{T}}\lambda(t) \in N_U(u(t)),$$

where $N_U(u)$ is the cone of all vectors $w \in \mathbb{R}^\nu$ such that, for each $v \in U$, $w^{\mathsf{T}}(v-u) \le 0$. Proceeding as in Proposition 2, we see that now $(d/dt)\lambda(t)^{\mathsf{T}}\delta(t) \ge 0$ and therefore

$$\nabla\mathcal{C}(x(t_0+T))^{\mathsf{T}}\delta(t_0+T) \ge 0,$$

which is the necessary condition for a minimum in the continuous problem. For a PRK discretisation of the boundary value for the states+costates system, the relation $(d/dt)\lambda(t)^{\mathsf{T}}\delta(t) \geq 0$ implies

$$k_{n,i}^{\mathsf{T}}\Lambda_{n,i} + \Delta_{n,i}^{\mathsf{T}}\ell_{n,i} \geq 0$$

and therefore we may use Lemma 1 yet again to conclude that for symplectic PRK methods and if *the weights b_i are positive*,

$$\nabla\mathcal{C}(x_N)^{\mathsf{T}}\delta_N \geq 0.$$

Once more, results similar to Theorems 10 and 11 hold. See [9] for order reduction results.

6 Lagrangian mechanics

Let us now consider Lagrangian mechanical systems [2]. Denote by $\mathcal{L}(x,u,t)$ the Lagrangian function, where $x \in \mathbb{R}^d$ are the Lagrangian co-ordinates and $u = (d/dt)x$ the corresponding velocities. According to Hamilton's principle, the trajectories $t \mapsto x(t)$ of the system are characterised by the fact that they render stationary (often minimum) the action integral

$$\int_{t_0}^{t_0+T} \mathcal{L}(x(t),u(t),t)\,dt,$$

among all curves $t \mapsto \bar{x}(t)$ with $\bar{x}(t_0) = x(t_0)$ and $\bar{x}(t_0+T) = x(t_0+T)$. This may of course be viewed as a control problem to make stationary (or even maximum) the cost (61) with $\mathcal{C} \equiv 0$ and $\mathcal{D} = -\mathcal{L}$, subject to the constraint $\dot{x} = u$ with fixed end-values $x(t_0)$ and $x(t_0+T)$. The theory in Section 5 applies. The pseudo-Hamiltonian is $H(x,\lambda,u,t) = \lambda^{\mathsf{T}}u - \mathcal{L}(x,u,t)$. The constraint $\nabla_u H = 0$ reads $\lambda = \nabla_u\mathcal{L}(x,u,t)$; thus the control costates coincide with the mechanical momenta. The elimination of the controls with the help of Pontryagin's principle would determine u as a function $\Phi(x,\lambda,t)$ by maximising (recall that we are here trying to maximise the cost!) the function $u \mapsto H(x,\lambda,u,t)$. In mechanics, this exactly corresponds with the theory of the Legendre transformation as presented in [2, Section 14]: that theory shows that, if $\mathcal{L}$ is a strictly convex function of u, then, at given x and t, the velocity vector u that corresponds to a given value of the momentum λ is globally uniquely defined and maximises $\lambda^{\mathsf{T}}u - \mathcal{L}(x,u,t)$. In most mechanical problems $\mathcal{L} = \mathcal{T}(x,u,t) - \mathcal{V}(x,t)$, with $\mathcal{T}$ and $\mathcal{V}$ the kinetic and potential energy respectively, and $\mathcal{T}$ is quadratic, positive-definite in u, thus ensuring the required convexity. In control theory the elimination of the controls u in the pseudo-Hamiltonian H gives rise to the 'control' Hamiltonian $\mathcal{H}$; correspondingly, in mechanics the Hamiltonian is defined as the result of expressing in $\lambda^{\mathsf{T}}u - \mathcal{L}(x,u,t)$ the velocities as functions of the momenta (and x and t). Finally the evolution of the states and costates (mechanical co-ordinates and momenta) obeys Hamilton's canonical equations. Hamiltonian solution flows are symplectic and, in this way, we have travelled all the way from action minimisation to symplecticness.

A similar journey may take place in the discrete realm. Choose any RK scheme (1) with nonzero weights to discretise the differential constraint $(d/dt)x = u$ and minimise the associated discrete action

$$\sum_{n=0}^{N-1} h_n \sum_{i=1}^{s} b_i \, \mathcal{L}(X_{n,i}, U_{n,i}, t_n + c_i h_n).$$

As we know from Theorem 10, this direct approach implies a symplectic PRK integration of the Hamiltonian system for x and λ, where the λ equations are integrated with the coefficients (9). This is nothing more than the variational construction of PRK symplectic integrators, already presented in the early paper [37] by Suris (see [24] for more information on integrators based on the principle of least action, cf. [21]). In this way, Hager's result [15] may be viewed as an extension of Suris work to general control problems.

Acknowledgments. I am extremely indebted to E. Zuazua for providing me with the initial motivation for this research. He, A. Murua E. Trélat kindly provided several useful references. Additionally, A. Murua has to be thanked for sharing some of his many insights; in particular in connection with automatic differentiation. This research is supported by project MTM2010-18246-C03-01 from Ministerio de Ciencia e Innovación, Spain.

References

[1] L. ABIA AND J. M. SANZ-SERNA, *Partitioned Runge-Kutta methods for separable Hamiltonian problems*, Math. Comput. 60 (1993), pp. 617–634.

[2] V. I. ARNOLD, *Mathematical Methods of Classical Mechanics, 2nd ed.*, Springer, New York, 1989.

[3] P. B. BOCHEV AND C. SCOVEL, *On quadratic invariants and symplectic structure*, BIT 34 (1994), pp. 337-345.

[4] J. F. BONNANS AND J. LAURENT-VARIN, *Computation of order conditions for symplectic partitioned Runge-Kutta schemes with application to optimal control*, Numer. Math. 103 (2006), pp. 1-10.

[5] J. C. BUTCHER, *Numerical Methods for Ordinary Differential Equations, 2nd ed.*, J. Wiley, Chichester, 2008.

[6] D. G. CACUCI, *Sensitiviy theory for nonlinear systems. I. Nonlinear functional analysis approach*, J. Math, Phys. 22 (1981), pp. 2794–2802.

[7] M. CHYBA, E. HAIRER, AND G. VILMART, *The role of symplectic methods in optimal control*, Optim. Control Appl. Meth. 30 (2009), pp. 367–382.

[8] G. J. COOPER, *Stability of Runge-Kutta methods for trajectory problems*, IMA J. Numer. Anal. 7 (1987), pp. 1–13.

[9] A. L. DONTCHEV, W. W. HAGER, AND V. M. VELIOV, *Second-order Runge-Kutta approximations in control constrained optimal control*, SIAM J. Numer. Anal. 38 (2000), pp. 202–226.

[10] S. ERVEDOZA AND E. ZUAZUA, *Numerical Approximation of Exact Control Waves*, Springer, New York, 2013.

[11] A. FARRÉS, J. LASKAR, S. BLANES, F. CASAS, J. MAKAZAGA, AND A. MURUA, *High precision Symplectic Integrators for the Solar System*, Celestial Mechanics and Dynamical Astronomy 116 (2013), pp. 141–174.

[12] M. B. GILES AND N. A. PIERCE, *An Introduction to the adjoint approach to design*, Flow, Turbulence and Combustion 65 (2000), pp. 393415.

[13] A. GRIEWANK, *A mathematical view of automatic differentiation*, Acta Numerica 12 (2003), pp. 321–398.

[14] D. F. GRIFFITHS AND J. M. SANZ-SERNA, *On the scope of the method of modified equations*, SIAM J. Sci. Statist. Comput. 7 (1986), pp. 994-1008.

[15] W. W. HAGER, *Runge-Kutta methods in optimal control and the transformed adjoint system*, Numer. Math. 87 (2000), pp. 247–282.

[16] E. HAIRER, *Backward analysis of numerical integrators and symplectic methods*, Annals of Numerical Mathematics 1 (1994), pp. 107–132.

[17] E. HAIRER, CH. LUBICH, AND G. WANNER, *Geometric Numerical Integration*, 2nd ed., Springer, Berlin, 2006.

[18] E. HAIRER, A. MURUA, AND J. M. SANZ-SERNA, *The nonexistence of symplectic multiderivative Runge-Kutta methods*, BIT 34 (1994), pp. 80-87.

[19] E. HAIRER, S. P. NØRSETT AND G. WANNER, *Solving Ordinary Differential Equations I, Nonstiff Problems, 2nd ed.,* Springer, Berlin, 1993.

[20] E. HAIRER, AND G. WANNER, *Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems, 2nd ed.,* Springer, Berlin, 1996.

[21] S. LALL AND M. WEST, *Discrete variational Hamiltonian mechanics*, J. Phys. A: Math. Gen. 39 (2006), pp. 5509–5519.

[22] F. M. LASAGNI, *Canonical Runge-Kutta methods*, ZAMP 39 (1988), pp. 952-953.

[23] B. LEIMKUHLER AND S. REICH, *Simulating Hamiltonian Dynamics,* Cambridge University Press, Cambridge, 2005.

[24] J. E. MARSDEN AND M. WEST, *Discrete mechanics and variational integrators*, Acta Numerica 10 (2001), pp. 357–514.

[25] A. MURUA, *On order conditions for partitioned symplectic methods*, SIAM J. Numer. Anal. 34 (1997), pp. 2204-22011.

[26] I. M. ROSS, *A road for optimal control: the right way to commute*, Annals of the New York Academy of Sciences 1065 (2006), pp. 210–231.

[27] A. SANDU, D. N. DAESCU, G. R. CARMICHAEL, T. CHAI, *Adjoint sensitivity analysis of regional air quality models*, J. of Comput. Phys. 204 (2005), pp. 222–252.

[28] J. M. SANZ-SERNA, *Runge-Kutta schemes for Hamiltonian systems*, BIT 28 (1988), pp. 877–883.

[29] J. M. SANZ-SERNA, *Geometric integration*, In : The State of the Art in Numerical Analysis, I. S. Duff and G. A. Watson eds., Clarendon Press, Oxford 1997, pp. 121-143.

[30] J. M. SANZ-SERNA AND L. ABIA, *Order conditions for canonical Runge-Kutta schemes*, SIAM J. Numer. Anal 28 (1991), pp. 1081–1096.

[31] J. M. SANZ-SERNA AND M. P. CALVO, *Numerical Hamiltonian Problems,* Chapman and Hall, London, 1994.

[32] R. R. SCHERER AND H. TÜRKE, *Reflected and transposed Runge-Kutta methods*, BIT 23 (1983) pp. 262–266.

[33] T. SCHLICK, *Molecular Modelling and Simulation: An Interdisciplinary Guide, 2nd. ed.,* Springer, New York, 2010.

[34] Z. SIRKES AND E. TZIPERMAN, *Finite difference of adjoint or adjoint of finite difference?*, Monthly Weather Rev. 49 (1997), pp. 5–40.

[35] E. T. SONTAG, *Mathematical Control Theory, Deterministic Finite Dimensional Systems, 2nd. ed.*, Springer, New York, 1998.

[36] Y. B. SURIS, *Preservation of symplectic structure in the numerical solution of Hamiltonian systems*, in Numerical Solution of Differential Equations, S. S. Filippov ed., Akad. Nauk. SSSR, Inst. Prikl. Mat., Moscow, 1988, pp. 138–144 (in Russian).

[37] Y. B. SURIS, *Hamiltonian methods of Runge-Kutta type and their variational interpretation*, Math. Model. 2 (1990), pp. 78-87 (in Russian).

[38] E. TRÉLAT, *Contrôle Optimal: Théorie et Applications*, Vuibert, Paris, 2005.

[39] J. ZABCZYK, *Mathematical Control Theory, An Introduction*,Birkhäuser, Boston, 1995.

Appendix A: Schemes with some vanishing weights

If one or more wieghts b_i in (1) vanish, then it is not possible to use the recipe (41) to define the coefficients required to create a combined symplectic PRK method (1), (9). Given the partitioned system (10) and the q coefficients (1), how to integrate the

p equations so as to have a symplectic scheme? The solution to this problem is rather weird and it is best to begin with the simplest example.

Let us study the second-order scheme (due to Runge in his 1895 original paper [19, Section II.1]), $s = 2$,

$$a_{11} = a_{21} = a_{22} = 0,\ a_{12} = 1/2,\ b_1 = 1,\ b_2 = 0,\ c_1 = 1/2,\ c_2 = 0. \tag{62}$$

While it is customary to label the stages so that the abscissas c_i increase with i, we have departed from this practice; if we adopted it, formula (67) below would get a rather disordered appearance.

We regularise the zero weight and consider the one-parameter family, $\epsilon \neq 0$:

$$a_{11} = a_{21} = a_{22} = 0,\ a_{12} = 1/2,\ b_1 = 1,\ b_2 = \epsilon,\ c_1 = 1/2,\ c_2 = 0. \tag{63}$$

(The regularised scheme is not even consistent, but this does not hinder the argument.) From (41), we set

$$A_{11} = 1,\ A_{12} = A_{22} = \epsilon,\ A_{21} = 1 - 1/(2\epsilon),\ B_1 = 1,\ B_2 = \epsilon,\ C_1 = 1/2,\ C_2 = 0. \tag{64}$$

Thus, the PRK specified by (63)–(64) is symplectic for each ϵ. The idea now is to take limits as $\epsilon \to 0$; the limit integrator, *if it exists*, will preserve quadratic invariants and, when applied to Hamiltonian problems, the symplectic structure. The difficulty is that from the equation that defines $P_{n,2}$

$$P_{n,2} = p_n + h_n \left(1 - \frac{1}{2\epsilon}\right) g(Q_{n,1}, P_{n,1}, t_n + h_n/2) + h_n \epsilon\, g(Q_{n,2}, P_{n,2}, t_n)$$

we may expect that, for fixed q_n, p_n, the stage vector $P_{n,2}$ grows unboundedly as $\epsilon \to 0$ and that, therefore, a limit integrator cannot be defined. However, the stage $P_{n,2}$ only affects $P_{n,1}$ and p_{n+1} through the *small* coefficients $A_{1,2} = B_2 = \epsilon$, and this makes it possible to prove that the limit scheme exists for some particular differential equations. Specifically, we assume in the remainder of this section that in the partitioned differential system (10) being integrated, f and g have the special form

$$f = f(q,t) \qquad g = L(q,t) + M(q,t)p \tag{65}$$

(with $q = x$, $p = \lambda$, this format includes the system (19), (21) in Section 3). When (65) holds, the q integration with coefficients (63) converges, as $\epsilon \to 0$, to the integration with the originally given coefficients (62). The system for the p stages P_1, P_2 (the index n is sometimes dropped to shorten the formulas) may be written as

$$\begin{aligned} P_1 &= p_n + h_n(L_1 + M_1 P_1) + h_n(\epsilon L_2 + h_n M_2 m_2), \\ m_2 &= \frac{\epsilon}{h_n} p_n + \big(\epsilon - \frac{1}{2}\big)(L_1 + M_1 P_1) + \epsilon(\epsilon L_2 + h_n M_2 m_2), \end{aligned}$$

where we have scaled $m_2 = (\epsilon/h_n)P_2$ to avoid blow-up and used the abbreviations

$$\begin{aligned} L_1 &= L(Q_1, t_n + h_n/2), & M_1 &= M(Q_1, t_n + h_n/2), \\ L_2 &= L(Q_2, t_n), & M_2 &= M(Q_2, t_n). \end{aligned}$$

Now take limits as $\epsilon \to 0$, to get

$$\begin{aligned} P_1 &= p_n + h_n(L_1 + M_1 P_1) + h_n^2 M_2 m_2, \\ m_2 &= -\frac{1}{2}(L_1 + M_1 P_1). \end{aligned}$$

Since $B_1 = A_{11}$ and $B_2 = A_{12}$, the end-of-step approximations is given by $p_{n+1} = P_1$.

We write these equations in a way similar to (11)–(13):

$$\begin{aligned} p_{n+1} &= p_n + h_n \ell_1 + h_n^2 M_2 m_2, \\ \ell_1 &= g(Q_1, P_1, t_n + h_n/2), \\ M_2 &= M(Q_2, t_n), \\ P_1 &= p_n + h_n \ell_1 + h_n^2 M_2 m_2, \\ m_2 &= -\frac{1}{2}\ell_1. \end{aligned} \tag{66}$$

The combination of these formulas for p with the scheme (62) for q is a *first-order* integrator that conserves quadratic invariants as in Theorem 3 and, for Hamiltonian problems, preserves the symplectic structure. Of course the integrator is *not* a PRK method; since $M = \partial_p g$, the formula (66) is reminiscent of Runge-Kutta methods that use higher derivatives of the solution [19, Section II.13].(Such high-order derivative methods cannot be symplectic for general problems [18].) Note that, while ℓ_1 is an approximation to the first derivative $(d/dt)p$, the vector $M_2 m_2$ has the dimensions of the second derivative $(d^2/dt^2)p$.

Let us now turn to the general case. Assume that in (1) the first r weights $b_1, \ldots, b_r$ do not vanish, while $b_{r+1} = \cdots = b_s = 0$. The regularisation procedure used for Runge's method leads to the fancy integrator:

$$p_{n+1} = p_n + h_n \sum_{i=1}^{r} b_i \ell_i + h_n^2 \sum_{\alpha=r+1}^{s} M_\alpha m_\alpha. \tag{67}$$

$$\begin{aligned} P_i = p_n + h_n \sum_{j=1}^{r} \big(b_j - \frac{b_j a_{ji}}{b_i}\big)\ell_j \\ + h_n^2 \sum_{\beta=r+1}^{s} \big(1 - \frac{b_j a_{\beta i}}{b_i}\big) M_\beta m_\beta, \qquad i = 1, \ldots, r, \end{aligned} \tag{68}$$

$$m_\alpha = -\sum_{j=1}^{r} b_j a_{j\alpha} \ell_j - h_n \sum_{\beta=r+1}^{s} a_{\beta\alpha} M_\beta m_\beta, \qquad \alpha = r+1, \ldots, s. \tag{69}$$

Here the r vectors ℓ_i are as in (12), so that the method uses r slopes and additionally $s - r$ matrices $M_\alpha = M(Q_\alpha, t_n + c_\alpha h_n)$. From the relations (69) the m_α may be viewed as functions of the ℓ_i.

The following result is a consequence of the construction via regularisation:

Theorem 14 *Consider partitioned systems of the special format (65), where the q equations are integrated with the RK scheme (1), $b_1 \neq 0, \ldots, b_r \neq 0$, $b_{r+1} = \cdots =$*

$b_s = 0$, and the p equations with the formulas in (67)–(69). If $S(q(t), p(t))$ is a conserved quantity as in Theorem 3, then $S(q_n, p_n)$ is independent of n. If the system is Hamiltonian, then the map $(q_n, p_n) \mapsto (q_{n+1}, p_{n+1})$ is symplectic.

Proofs of this theorem that do not rely on taking limits as $\epsilon \to 0$ are of course possible. For such an alternative proof of the conservation of S, we may note that manipulations (not reproduced here) similar to those used to prove Lemma 1 show that for the present method, in lieu of (18), we may write:

$$\begin{aligned} S(q_{n+1}, p_{n+1}) - S(q_n, p_n) &= h_n \sum_{i=1}^{r} b_i \big(S(k_i, P_i) + S(Q_i, \ell_i)\big) \\ &\quad + h_n^2 \sum_{\alpha=r+1}^{s} \big(S(k_\alpha, m_\alpha) + S(Q_\alpha, M_\alpha m_\alpha)\big). \end{aligned}$$

This is an algebraic identity that does not require that the system integrated to be conservative. When S is conserved, the first sum vanishes as in the proof of Theorem 3. For the second sum note that from $S(f(q,t), p) + S(q, L(q,t) + M(q,t)p) \equiv 0$ it follows that $S(f, p) + S(q, Mp) \equiv 0$.

For the adjoint equations in Section 3, the conclusion of Theorem 7 holds if the x equations are integrated with a (nonsymplectic) RK method with one or more vanishing weights and the λ equations are integrated as in (67)–(69). Similarly Theorem 8 holds for a suitable choice of the Lagrangian (details will not be given, but see below).

What is the situation for the control problem in Section 4? Recall that the corresponding system of *differential equations* is given by (44), (46), where, in the right-hand sides, u has been expressed as $u = \Phi(x, \lambda, t)$. That system of differential equations does *not* possess the format (65) for which (67) makes sense and, accordingly, we cannot provide analogues to Theorems 9 and 10.

In order to gain additional insight, let us use the direct approach based on Runge's second order integrator (62). We define the Lagrangian (compare with (43) and note consistency with (24) due to the factor h_n^2):

$$\begin{aligned} \mathcal{C}(x_N) - \lambda_0^{\mathsf{T}}(x_0 - \alpha) - \sum_{n=0}^{N-1} h_n \lambda_{n+1}^{\mathsf{T}} \Big[\frac{1}{h_n}(x_{n+1} - x_n) - k_{n,1}\Big] \\ - \sum_{n=0}^{N-1} h_n \Lambda_n^{\mathsf{T}} \Big[k_{n,1} - f(X_{n,1}, U_{n,1}, t_n + h_n/2)\Big] \\ - \sum_{n=0}^{N-1} h_n^2 \mu_n^{\mathsf{T}} \Big[k_{n,2} - f(X_{n,2}, U_{n,2}, t_n)\Big], \end{aligned}$$

where, as on other occasions, the stages $X_{n,1} = x_n + (h_n/2)k_{n,2}$, $X_{n,2} = x_n$ must be seen as known functions of x_n and $k_{n,2}$. Taking gradients with respect to x_n, $k_{n,1}$,

$k_{n,2}$ leads to the necessary conditions

$$\begin{aligned}\lambda_{n+1} &= \lambda_n - (\partial_x f(X_{n,1}, U_{n,1}, t_n + h_n/2))^{\mathsf{T}} \Lambda_n - h_n^2 (\partial_x f(X_{n,2}, U_{n,2}, t_n))^{\mathsf{T}} \mu_n, \\ \Lambda_n &= \lambda_{n+1}, \\ \mu_n &= \frac{1}{2} (\partial_x f(X_{n,1}, U_{n,1}, t_n + h_n/2))^{\mathsf{T}} \Lambda_n;\end{aligned}$$

which clearly correspond to the integrator (66).(By considering the case where f is independent of u, this shows that Theorem 8 holds in this case.) However, taking gradients with respect to $U_{n,1}$ and $U_{n,2}$ yields

$$(\partial_u f(X_{n,1}, U_{n,1}, t_n + h_n/2))^{\mathsf{T}} \Lambda_n = 0, \qquad (\partial_u f(X_{n,2}, U_{n,2}, t_n))^{\mathsf{T}} \mu_n = 0.$$

The second equation is totally meaningless. It cannot be seen as a discretisation of (47) because μ_n is not an approximation to the costate λ; it does not even possess the right dimensions for that to happen. The values of $U_{n,2}$ retrieved from this constraint will have no relation to the true optimal controls. The paper [15] nicely illustrates this with an example (see also [9]).

Since the trouble arises by the presence of the controls, things may be fixed by tampering with $U_{n,2}$, as pointed out in [15], [9]. However, there is no shortage of RK schemes with nonzero (or even positive) weights, so that, in practice, resorting to such fixes seems ill advised.

Appendix B: Reflecting and transposing RK coefficients

Scherer and Türke [32] associated with the set of RK coefficients (1) two new sets called the reflection and the transposition of the original. The reflected coefficients are given by ($i, j = 1, \ldots, s$)

$$a^r_{ij} = b_j - a_{ij}, \quad b^r_i = b_i, \quad c^r_i = 1 - c_i$$

and the transposed coefficients are defined, only for methods with nonvanishing weights b_i, by

$$a^t_{ij} = b_j a_{ji}/b_i, \quad b^t_i = b_i, \quad c^t_i = 1 - c_i.$$

The operations of reflection and transposition commute: the transposition of the reflection coincides with the reflection of the transposition as both lead to

$$a^{rt}_{ij} = b_j - b_j a_{ji}/b_i, \quad b^{rt}_i = b_i, \quad c^{rt}_i = c_i.$$

Furthermore both operations are involutions: each is its own inverse.

The paper [32] introduces the operations of reflection and transposition as algebraic manipulations that make it possible to interrelate important families of RK methods; no attempt is made there to interpret computationally the meaning of integrating with the reflected or transposed coefficients. What do reflection and composition mean? The interpretation of reflection is well known [31, Section 3.6], [19, Chapter II, Theorem 8.3]: a step of length $-h_n$ with the reflected RK method inverts the transformation

$y_n \mapsto y_{n+1}$ induced by a step of length h_n with the original method. In this paper we have seen this idea at work when moving from (32)–(34) to (35)–(37). The formulas (41) provide meaning to the idea of transposition: to construct a symplectic PRK out of a given RK method with nonvanishing weights the p coefficients are determined by reflecting and transposing the given q coefficients. The transposed of the q coefficients are then those required to integrate backwards the p equations in, say, sensitivity analyses.

As a further illustration of these ideas, consider the linear non-autonomous system

$$\frac{d}{dt}q = M(t)q, \quad \frac{d}{dt}p = -M(t)^{\mathsf{T}}p,$$

integrated with the PRK method (1), (9). Since p and q are uncoupled, this amounts to an RK integration of the q equations with the coefficients (1) together with an RK integration of the p equations with the coefficients (9). The system has the invariant $q^{\mathsf{T}}p$; Theorem 3 ensures that it will be preserved if the p coefficients are the transposition of the reflection of the q coefficients. Both sets of coefficients only coincide if q itself is integrated symplectically. A nonsymplectic integration of q is possible, but then one has to compensate by integrating the p equations in an appropriate way and the order and stability of the p integration has to be investigated separately. Again, if the p equations are integrated backward in time, then, preservation of $q^{\mathsf{T}}p$ requires that such backward integration be performed with the transposition of the coefficients used to propagate q forward.

The material in the preceding section shows that, for systems of the special form (65), the scheme (67) may be viewed as the reflected and transposed of (1) in the case of zero weights.

We conclude this appendix with a remark on terminology. Monographs such as [17] and [31] use the word 'adjoint' to refer to the method with reflected coefficients. Section 3 and the discussion in this section suggest that, in order to proceed as in the differential equation case, it would have been better to keep the word adjoint for the reflected and transposed method. And call reflected to what in [17] or [31] is called adjoint. With that alternative terminology, for RK schemes, symplecticness would simply be self-adjointness.