Found Comput Math
DOI 10.1007/s10208-013-9163-y

FOUNDATIONS OF COMPUTATIONAL MATHEMATICS
The Journal of the Society for the Foundations of Computational Mathematics

Geometric Generalisations of SHAKE and RATTLE

Robert I McLachlan · Klas Modin · Olivier Verdier · Matt Wilkins

Received: 15 July 2012 / Revised: 15 March 2013 / Accepted: 26 April 2013

Abstract A geometric analysis of the SHAKE and RATTLE methods for constrained Hamiltonian problems is carried out. The study reveals the underlying differential geometric foundation of the two methods, and the exact relation between them. In addition, the geometric insight naturally generalises SHAKE and RATTLE to allow for a strictly larger class of constrained Hamiltonian systems than in the classical setting.

In order for SHAKE and RATTLE to be well defined, two basic assumptions are needed. First, a nondegeneracy assumption, which is a condition on the Hamiltonian, i.e., on the dynamics of the system. Second, a coisotropy assumption, which is a condition on the geometry of the constrained phase space. Non-trivial examples of systems fulfilling, and failing to fulfill, these assumptions are given.

Communicated by Arieh Iserles.

R. I McLachlan (✉) · M. Wilkins
Institute of Fundamental Sciences, Massey University, Palmersston North, New Zealand
e-mail: r.mclachlan@massey.ac.nz

M. Wilkins
e-mail: m.c.wilkins@massey.ac.nz

K. Modin
Department of Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden
e-mail: klas.modin@chalmers.se

O. Verdier
Department of Mathematics, NTNU, Trondheim, Norway
e-mail: olivier.verdier@math.ntnu.no

Published online: 25 June 2013

Keywords Symplectic integrators · Constrained Hamiltonian systems · Coisotropic submanifolds · Differential algebraic equations

Mathematics Subject Classification (2010) 37M15 · 65P10 · 70H45 · 65L80

1 Introduction

SHAKE and RATTLE are two commonly used numerical integration methods for Hamiltonian problems subject to holonomic constraints [1, 9, 11, 15, 16]. The difference between the two methods is that RATTLE preserves "hidden" constraints, whereas SHAKE does not. For details and a historical account, see the monographs [10, § 7.2] and [8, § VII.1.4].

In this paper we give a rigorous geometric analysis of the SHAKE and RATTLE methods. Our approach is based on the observation by Reich [15] that SHAKE and RATTLE may be expressed using flow maps. The analysis sheds light on the underlying "geometric foundation" of the two methods, and also on the exact relation between them. In addition, the geometric insight allows us to integrate a larger class of constrained problems than before. Indeed, the geometric versions of SHAKE and RATTLE work for *coisotropic constraints*. This class of constraints is strictly larger than the class of holonomic constraints. In particular, they may depend on both position and momentum.

Throughout the paper we utilise the language of differential geometry. The main reason for doing so is *not* to generalise SHAKE and RATTLE from $\mathbb{R}^{2d}$ to manifolds, but rather because this notation makes the geometric structures more transparent. However, in order to link to the standard literature on SHAKE and RATTLE, we give many key results also in the classical $\mathbb{R}^{2d}$ setting as examples.

Our notation closely follows that of Marsden and Ratiu [13]. In particular, if $\mathcal{M}$ and $\mathcal{N}$ are two manifolds and $f \in \mathcal{C}^\infty(\mathcal{M}, \mathcal{N})$, then $\mathsf{T}f : \mathsf{T}\mathcal{M} \to \mathsf{T}\mathcal{N}$ denotes the tangent derivative. If $\mathcal{N} \subset \mathcal{M}$ is a submanifold, then $\imath_{\mathcal{N}} : \mathcal{N} \to \mathcal{M}$ denotes the inclusion, and the pull-back of differential forms from $\mathcal{M}$ to $\mathcal{N}$ is denoted $\imath_{\mathcal{N}}^*$. If $(\mathcal{P}, \omega)$ is a symplectic manifold, and $H \in \mathcal{C}^\infty(\mathcal{P})$, then X_H denotes the corresponding Hamiltonian vector field, and the Poisson bracket is denoted $\{\cdot, \cdot\}$. The contraction between a vector X and a differential form α is denoted $\mathrm{i}_X \alpha$.

We continue this section with an outline of the paper and the main results.

Problem Formulation Let $(\mathcal{P}, \omega)$ be a symplectic manifold, $H \in \mathcal{C}^\infty(\mathcal{P})$ a smooth function, and $\mathcal{M} \subset \mathcal{P}$ a submanifold. Given $(\mathcal{P}, \omega, H, \mathcal{M})$, the problem is to find a smooth curve $t \mapsto \gamma(t)$ such that

$$\imath_{\mathcal{M}}^*(\mathrm{i}_{\dot\gamma}\omega - \mathrm{d}H) = 0, \qquad \gamma(t) \in \mathcal{M}. \tag{1a}$$

Equation (1a) looks like a Hamiltonian system on $\mathcal{P}$, but constrained to stay on the submanifold $\mathcal{M}$, called the *constraint manifold*. It can be rewritten as

$$\mathrm{i}_{\dot\gamma}\nu - \mathrm{d}H_{\mathcal{M}} = 0, \tag{1b}$$

where $\nu = \imath_{\mathcal{M}}^*\omega$ and $H_{\mathcal{M}} = \imath_{\mathcal{M}}^* H$. From this formulation it is clear that Eqs. (1a), (1b) is intrinsic to $\mathcal{M}$, i.e., it only depends on objects defined on $\mathcal{M}$.

Example 1.1 Let $\mathcal{P} = \mathbb{R}^{2d}$ with canonical coordinates $z = (q, p)$, and let $\mathcal{M} = g^{-1}(\{0\})$, where $g = (g_1, \ldots, g_m) \in \mathcal{C}^\infty(\mathbb{R}^{2d}, \mathbb{R}^m)$. Equations (1a), (1b) then takes the form

$$\begin{aligned} \dot{q} &= H_p(q, p) + g_p(q, p)^\mathsf{T}\lambda, \\ \dot{p} &= -H_q(q, p) - g_q(q, p)^\mathsf{T}\lambda, \\ 0 &= g(q, p), \end{aligned} \tag{2}$$

where g_q and g_p are the partial Jacobian matrices, and $\lambda = (\lambda_1, \ldots, \lambda_m)$ are Lagrange multipliers. Notice that if g does not depend on p, then this reduces to a canonical Hamiltonian system with holonomic constraints.

Owing to the Hamiltonian structure of the equations, the reduction of the differential–algebraic equation (DAE) (2) to an ordinary differential equation (ODE) takes a particular form. This was already noticed by Dirac [2, § 1], and later perfected in [5].

Existence and Uniqueness Since Eqs. (1a), (1b) is intrinsic to $\mathcal{M}$, it is clear that any condition or assumption for existence and uniqueness should also be intrinsic, so it is enough to investigate existence and uniqueness intrinsically on $\mathcal{M}$.

The 2-form ν is closed, but in general degenerate, so $(\mathcal{M}, \nu)$ is not, in general, a symplectic manifold. Instead, it is a *presymplectic* manifold.

The kernel of ν defines a distribution on $\mathcal{M}$ denoted $\ker \nu$. As detailed in Sect. 2.1, the kernel distribution is *integrable*. Thus, for each $z \in \mathcal{M}$, there is a submanifold $[z] \subset \mathcal{M}$ such that $\mathsf{T}_x[z] = \ker \nu_x$ for each $x \in [z]$.

If $\gamma(t)$ is a solution to (1b), then $\mathrm{d}H_{\mathcal{M}}(\gamma(t)) \in \nu(\mathsf{T}_{\gamma(t)}\mathcal{M})$. Thus, solutions stay on the *hidden constraint set*, given by

$$\mathcal{M}^H = \left\{ z \in \mathcal{M} : \mathrm{d}H_{\mathcal{M}}(z) \in \nu(\mathsf{T}_z\mathcal{M}) \right\}. \tag{3}$$

In particular, a necessary condition for Eq. (1b) to have a solution is that the initial data fulfils $\gamma(0) \in \mathcal{M}^H$, which is assumed from here on.

As is further explained in Sect. 2, a sufficient condition for (local) existence and uniqueness of solutions of Eq. (1b), and hence Eqs. (1a), (1b), is the following.

Assumption 1 (Nondegeneracy) For any $z \in \mathcal{M}$, the critical points of the function $H_{[z]} = \imath_{[z]}^* H_{\mathcal{M}}$ are nondegenerate.

Example 1.2 For the classical setting in Example 1.1,

$$\mathcal{M}^H = \left\{ z \in \mathbb{R}^{2d} : g_i(z) = 0,\, X_H(z) \cdot \nabla g_i(z) = 0, i = 1, \ldots, m \right\}$$

and Assumption 1 means that the matrix $g_z(z)^\mathsf{T} H_{zz}(z) g_z(z)$ is invertible for $z \in \mathcal{M}^H$. If g does not depend on p, then this is slightly weaker than the classical assumption that $g_q H_{pp} g_q$ is invertible (see Sect. 5.1.1).

Geometric SHAKE *and* RATTLE We now define SHAKE and RATTLE geometrically (see Fig. 2 for an illustration of the geometrical setting).

Definition 1.3 (Geometric SHAKE) Let φ_h be a method approximating $\exp(hX_H)$ for a given time step h. The SHAKE mapping $S_h\colon \mathcal{M} \ni z_0 \mapsto z_1^- \in \mathcal{M}$ is defined implicitly by

$$\varphi_h(y) \in \mathcal{M}, \quad y \in [z_0] \cap O, \qquad z_1^- = \varphi_h(y),$$

where $O \subset \mathcal{M}$ is a suitable open subset containing $\mathcal{M}^H$.

Definition 1.4 (Geometric RATTLE) Let φ_h be a method approximating $\exp(hX_H)$ for a given time step h. The RATTLE mapping $R_h\colon \mathcal{M}^H \ni z_0 \mapsto z_1 \in \mathcal{M}^H$ is defined implicitly by

$$z_1 \in \left[z_1^-\right] \cap \mathcal{M}^H, \quad z_1^- = S_h(z_0).$$

These are abstract definitions of SHAKE and RATTLE. In order to practically be able to implement them, an implicit definition of $\mathcal{M}$ in terms of constraint functions, and a parameterisation of $[z_0]$, is needed. This issue is discussed in Sect. 4, and is related to Assumption 2 introduced below.

In the holonomic case it is already known that SHAKE and RATTLE essentially yield the same method, since the projection step at the end of RATTLE is "neutralised" by the projection step in SHAKE. This observation is made geometrically precise in Sect. 4, where we show that SHAKE and RATTLE are two different representations of the same fibre mapping.

Example 1.5 It may be illuminating to understand the effect of both methods in a familiar example, that of a planar pendulum, realised as a constrained mechanical system. The ambient space is $\mathbb{R}^4$, and the constraint, which is holonomic, is given by $g(\mathbf{q}, \mathbf{p}) := \|\mathbf{q}\|^1 - 1$. The constraint manifold $\mathcal{M}$ is thus a three-dimensional submanifold of $\mathbb{R}^4$. The Hamiltonian is that of an unconstrained mass in a constant gravity field, $H(\mathbf{q}, \mathbf{p}) := \frac{1}{2}\|\mathbf{p}\|^2 + q_2$. The hidden constraint manifold is the submanifold of $\mathcal{M}$ which consists of the points where the velocity is tangent to $\mathcal{M}$, that is

$$\mathcal{M}^H = \left\{(\mathbf{q}, \mathbf{p}) \in \mathcal{M} : \mathbf{q} \cdot \mathbf{p} = 0\right\}. \tag{4}$$

For this particular case, the SHAKE and RATTLE algorithms are illustrated on Fig. 1.

Well-posedness of SHAKE *and* RATTLE The algebraic equations defining SHAKE and RATTLE can be thought of as discretisations of the original Eqs. (1a), (1b). However, contrary to the continuous case, the discretised equations are *not* intrinsic to $\mathcal{M}$. Thus, well-posedness of SHAKE and RATTLE depends on how $\mathcal{M}$ is embedded in $\mathcal{P}$.

Let $\mathsf{T}\mathcal{M}^\perp$ denote the orthogonal complement of $\mathsf{T}\mathcal{M}$ with respect to the symplectic form ω, i.e., $u \in \mathsf{T}_x\mathcal{M}^\perp$ if and only if $\omega(u, v) = 0$ for all $v \in \mathsf{T}_x\mathcal{M}$.

Definition 1.6 A submanifold $\mathcal{M}$ of $\mathcal{P}$ is called *coisotropic* if $\mathsf{T}\mathcal{M}^\perp \subset \mathsf{T}\mathcal{M}$.

As is explained carefully in Sect. 3, the natural assumption in order for SHAKE and RATTLE to be well-posed is the following, which is a completely extrinsic condition, i.e., it only has to do with how $\mathcal{M}$ is embedded in $\mathcal{P}$.

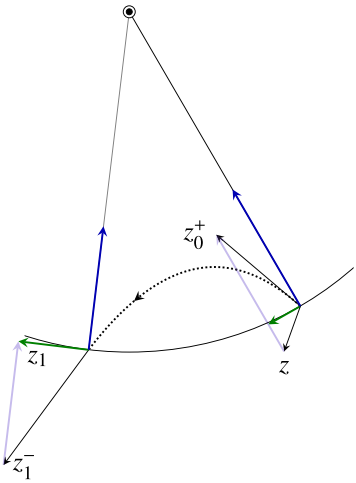


Fig. 1 An illustration of SHAKE and RATTLE in the familiar case of the planar pendulum. The reader is encouraged to compare that situation with the general geometrical description on Fig. 2. We start from a point z, which is on $\mathcal{M}$, but not necessarily on $\mathcal{M}^H$, which means that the velocity is not assumed to be tangent to the constraint circle. The *arrow* in the direction of the rod represents a kick applied to the mass, which has the effect of changing its normal velocity, resulting in a point z_0^+ in the phase space. The next step is to use a symplectic integrator to simulate the unconstrained flow, until the mass reaches the constraint manifold, that is, until the mass is at distance one from the origin; in this case we use the exact, unconstrained solution, which is a parabola. One has to adjust the initial kick in order to reach the constraint manifold exactly after one time-step. The mass reaches the constraint manifold $\mathcal{M}$ at the phase point z_1^- with a non-tangential velocity, which means that z_1^- does not belong to $\mathcal{M}^H$. The effect of RATTLE is now simply to correct the normal velocity by applying an appropriate kick in the rod direction in order to obtain a point $z_1 \in \mathcal{M}^H$

Assumption 2 (Coisotropy) $\mathcal{M}$ is a coisotropic submanifold of $\mathcal{P}$.

Example 1.7 For the setting in Example 1.1, let $g_1, \ldots, g_m$ be the components of the vector valued constraint function g. Then Assumption 2 means that

$$\{g_i, g_j\}(z) = 0, \quad \forall z \in \mathcal{M}.$$

An equivalent interpretation of Assumption 2 is that none of the Lagrange multipliers in Eq. (2) are resolved by differentiating the constraint condition once. From a DAE point of view, the nondegeneracy and coisotropy assumptions together asserts that Eq. (2) has index 3. An important particular case is obtained when g does not depend on p. In that case, Assumption 2 *always holds* (see Sect. 5.1). As shown in Sect. 3.2, Assumption 2 also implies that $[z]$ is parameterised by

$$(\lambda_1, \ldots, \lambda_m) \longmapsto \exp\left(\sum_{i=1}^{m} \lambda_i X_{g_i}\right)(z).$$

In turn, this means that the geometric SHAKE method is given by

$$S_h = \varphi_h \circ \exp\left(\sum_{i=1}^{m} \lambda_i X_{g_i}\right),$$

where $\lambda_1, \ldots, \lambda_m$ are determined implicitly by the conditions $g_i \circ S_h = 0$. Likewise, the geometric RATTLE method is given by

$$R_h = \exp\left(\sum_{i=1}^{m} \sigma_i X_{g_i}\right) \circ S_h,$$

where $\sigma_1, \ldots, \sigma_m$ are determined implicitly by $(X_H \cdot \nabla g_i) \circ R_h = 0$.

It is easy to find instances where the coisotropic and/or the nondegeneracy assumptions do not hold, and where the SHAKE and RATTLE methods are not well-defined. For example, if we take as constraint $H = \text{const}$, then the nondegeneracy assumption does not hold, and it is easy to see that SHAKE and RATTLE are not well-defined. This is expected, since the result by Ge and Marsden [3] asserts that it is not (in general) possible to construct symplectic and energy preserving methods. In Sect. 5.2 we give further examples of failing assumptions. Lastly, in Sect. 5.4 we also give a numerical example of a Hamiltonian problem with mixed position and momentum constraints, where we use the geometric SHAKE and RATTLE methods.

Main Results The main results in the paper can be summarised as follows.

1. Under Assumption 1, the set $\mathcal{M}^H$ is a symplectic submanifold with symplectic form $\varpi = \iota^*_{\mathcal{M}^H} \nu$, and Eqs. (1a), (1b) is well-posed for initial data in $\mathcal{M}^H$ (Theorem 2.7).
2. Under Assumption 1 and Assumption 2, there exists an open set $O \subset \mathcal{M}$, containing $\mathcal{M}^H$, such that the SHAKE map $S_h \colon O \to O$ is well defined and presymplectic, i.e., $S_h^* \nu = \nu$. Further, it is convergent of order at least 1 (Theorem 4.1 and Proposition 4.4).
3. Under Assumption 1 and Assumption 2, the RATTLE map $R_h \colon \mathcal{M}^H \to \mathcal{M}^H$ is well defined and symplectic, i.e., $R_h^* \varpi = \varpi$. Further, it is convergent of order at least 1 (Theorem 4.1 and Proposition 4.4).

2 Hamiltonian Systems on Presymplectic Manifolds

In this section we investigate the geometric structures of Eqs. (1a), (1b) from the intrinsic viewpoint, i.e., without "looking outside" of $\mathcal{M}$.

In general, a *presymplectic manifold* is a pair $(\mathcal{M}, \nu)$, where $\mathcal{M}$ is a smooth manifold, and ν is a closed 2-form on $\mathcal{M}$ called a *presymplectic form*. The difference from a symplectic form is that ν need not be nondegenerate. Thus, a symplectic manifold is a special case of a presymplectic manifold. We review some geometric concepts of presymplectic manifolds that are essential in the remainder. For a thorough treatment, we refer to the book by Libermann and Marle [12].

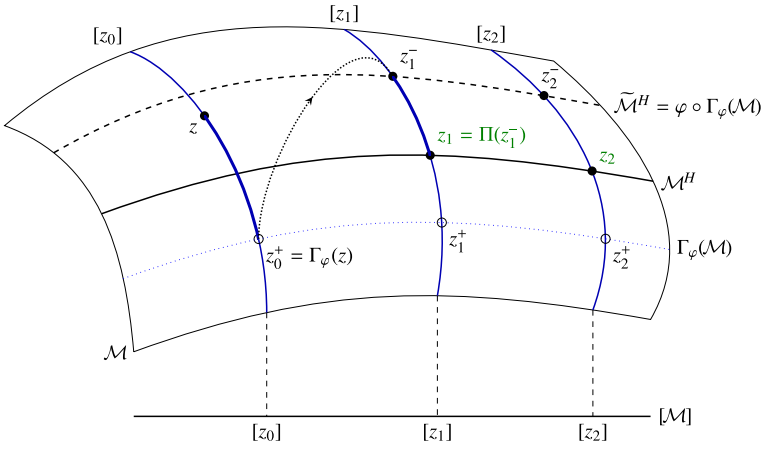


Fig. 2 An illustration of the SHAKE and RATTLE method with underlying symplectic method φ. See also Fig. 1 for an equivalent illustration in the familiar case of the pendulum. The first point z must lie in the vicinity of $\mathcal{M}^H$. The subsequent points $z_1^-, z_2^-, \ldots$ produced by SHAKE will all lie in the modified hidden constraint manifold $\widetilde{\mathcal{M}}^H$. The points $z_1, z_2, \ldots$ produced by RATTLE will lie on $\mathcal{M}^H$. Notice how the points z_k^-, z_k and z_k^+ always lie on the same fibre. Moreover, the precise location of a point on the fibre (for instance z on the figure) is irrelevant for both SHAKE and RATTLE

Given a function $H_{\mathcal{M}} \in \mathcal{C}^\infty(\mathcal{M})$, Eq. (1b) constitutes a Hamiltonian system on $(\mathcal{M}, \nu)$. Since ν might be degenerate, this equation is not, in general, an ordinary differential equation, but instead a DAE. We show in Sect. 2.3 that under Assumption 1 it is an index 1 DAE on $\mathcal{M}$. (In Sect. 3 we take the complementary extrinsic viewpoint, and we show that under Assumption 1 and Assumption 2 Eqs. (1a), (1b) can be interpreted as an index 3 problem on $\mathcal{P}$.)

2.1 Foliation

Throughout the paper we make the following "blanket assumption":

The dimension of the kernel distribution $\ker \nu$ is constant.

One important consequence is that the distribution $\ker \nu$ (now assumed to be *regular*) is *integrable* (cf. [6, Th. 25.2]). That is, at each point $x \in \mathcal{M}$ there is a submanifold $[x] \subset \mathcal{M}$ passing through x whose tangent spaces coincides with the distribution. The submanifolds $[x]$ are called *leaves*, and the collection $[\mathcal{M}]$ of all leaves is called a *foliation*. See Fig. 2 for an illustration of the foliation $[\mathcal{M}]$ of $\mathcal{M}$.

Remark 2.1 The foliation defines an equivalence class by $y \equiv x$ if $y \in [x]$. We denote the set of all such equivalence classes by $[\mathcal{M}]$. The projection is given by

$$\pi : \mathcal{M} \ni x \to [x] \in [\mathcal{M}].$$

The set $[\mathcal{M}]$ may or may not be a smooth manifold. When it is, the presymplectic form ν descends to a symplectic form $\bar{\nu}$ on $[\mathcal{M}]$, and π is a symplectic submersion.

The projection map π being a submersion means that we have a *fibration* of $\mathcal{M}$. Locally, every foliation is a fibration, but not necessarily globally. Throughout the remainder of the paper, we use the word "fibre" instead of "leaf", although the fibration may only be local.

A *presymplectic mapping* is a mapping $\varphi\colon \mathcal{M} \to \mathcal{M}$ that preserves the presymplectic form ν, i.e., for which

$$\nu(u, v) = \varphi^* \nu(u, v) := \nu(\mathsf{T}\varphi \cdot u, \mathsf{T}\varphi \cdot v) \quad \forall u, v \in \mathsf{T}_x \mathcal{M}.$$

There is a certain class of mappings that are trivial in the sense that they reduce to the identity mapping in the quotient manifold $[\mathcal{M}]$.

Definition 2.2 A smooth mapping $\varphi\colon \mathcal{M} \to \mathcal{M}$ is called *trivially presymplectic* if it preserves each fibre, i.e., if

$$\varphi(x) \in [x] \quad \text{for all } x \in \mathcal{M}.$$

The following result is clear.

Proposition 2.3 *If φ is trivially presymplectic, then it is presymplectic.*

2.2 Hidden Constraints

The fact that ν does not have full rank reflects that, in general, the possible solutions of Eq. (1b) do not fill the whole manifold $\mathcal{M}$. Indeed, if a curve $\gamma(t)$ is a solution of Eq. (1b), then $\mathrm{d}H(\gamma(t))$ must be in the set $\nu(\mathsf{T}_{\gamma(t)}\mathcal{M})$. As already seen in Sect. 1, the set of points at which this is fulfilled defines the hidden constraint set $\mathcal{M}^H \subset \mathcal{M}$, given by (3).

Remark 2.4 In general, this set is defined as the locus of m functions, where m is the dimension of $[x]$. However, if the differential of those functions are not independent at the locus points, the set $\mathcal{M}^H$ need not be a submanifold, and if it is a submanifold, it need not be of codimension m. For instance, we may have $\mathcal{M}^H = \mathcal{M}$ if the Hamiltonian H is constant along each fibre $[x]$. This is in particular the case if ν is nondegenerate.

Remark 2.5 The subset $\mathcal{M}^H$ is, strictly speaking, not a set of *hidden* constraints, but rather *implicit* constraints as a consequence of Eq. (1b).

2.3 Nondegeneracy Assumption

In this section we show that the nondegeneracy assumption, Assumption 1, ensures that:

(i) $\mathcal{M}^H$ is a submanifold of $\mathcal{M}$;
(ii) $\varpi = \imath^*_{\mathcal{M}^H} \nu$ is a symplectic form; and

(iii) the initial value problem (1b) is a Hamiltonian problem on $(\mathcal{M}^H, \varpi)$. As a consequence, problem (1b) has unique solutions for initial data in $\mathcal{M}^H$.

From a DAE point of view, Assumption 1 ensures that the DAE (1b) on $\mathcal{M}$ has index 1. As it turns out (see Sect. 4 below), the nondegeneracy assumption, together with Assumption 2, also asserts that the geometrically defined SHAKE and RATTLE methods are well defined.

We start with the observation that $\mathcal{M}^H$ is the set of critical points of $H_{[x]}$.

Proposition 2.6 *For $x \in \mathcal{M}$, let $H_{[x]} = \iota_{[x]}^* H_{\mathcal{M}}$. Then*

$$\mathcal{M}^H = \big\{ y \in \mathcal{M} : \mathrm{d}H_{[x]}(y) = 0 \big\}.$$

Proof If $y \in \mathcal{M}^H \cap [x]$ then $\mathrm{d}H_{[x]}(y) = 0$ since $\mathrm{d}H(y) \in \nu(\mathsf{T}_y\mathcal{M})$. Thus, the set $\mathcal{M}^H \cap [x]$ consists of critical points of the function $H_{[x]}$. □

Theorem 2.7 *Under Assumption 1, the following holds.*

1. *The set $\mathcal{M}^H$ is a submanifold of $\mathcal{M}$.*
2. *At a point $x \in \mathcal{M}^H$ we have*

$$\mathsf{T}_x\mathcal{M} = \mathsf{T}_x\mathcal{M}^H \oplus \ker \nu.$$

   In particular, the presymplectic form ν restricted to $\mathcal{M}^H$ is a symplectic form. Thus, $\mathcal{M}^H$ is a symplectic manifold.
3. *Equation (1b) has unique solutions for initial data in $\mathcal{M}^H$. These solutions are given by the solutions of the Hamiltonian problem on $\mathcal{M}^H$ obtained by restricting H to $\mathcal{M}^H$.*

Proof Each statement is proved, respectively, as follows.

1. Let $X_1, \ldots, X_m$ be linearly independent vector fields on M that span the distribution $\ker \nu$. Define the functions $\rho_i(x) := \langle \mathrm{d}H(x), X_i(x) \rangle$. Then, using Proposition 2.6,

$$\mathcal{M}^H = \big\{ x \in \mathcal{M} : \rho_i(x) = 0, i = 1, \ldots, m \big\}.$$

   $\mathcal{M}^H$ is a submanifold if $\mathrm{d}\rho_1(x), \ldots, \mathrm{d}\rho_k(x)$ are linearly independent for every $x \in \mathcal{M}^H$. An equivalent conditions is that the matrix

$$m_{ij} := \big\langle \mathrm{d}\rho_i(x), X_j(x) \big\rangle, \quad i, j = 1, \ldots, m$$

   be invertible for every $x \in \mathcal{M}^H$. Using that $\rho_i(x) = 0$, we get in local coordinates x_j that

$$m_{ij} = X_i^\alpha(x) \frac{\partial^2 H(x)}{\partial x^\alpha \partial x^\beta} X_j^\beta(x),$$

   where $X_i = X_i^\alpha \frac{\partial}{\partial x^\alpha}$. Since $X_1(x), \ldots, X_k(x)$ is a linearly independent basis of $(\ker \nu)_x = \mathsf{T}_x[x]$, Assumption 1 means exactly that this matrix is invertible, which thus proves the first assertion.

2. For the second assertion, using that the codimension of $\mathcal{M}^H$ is m, it suffices to prove that $(\ker \nu)_x \cap \mathsf{T}_x\mathcal{M}^H = 0$ for every $x \in \mathcal{M}^H$. Let $u \in \mathsf{T}_x\mathcal{M}^H$. Then $\langle \mathrm{d}\rho_i(x), u\rangle = 0$. Next, assume that $U \in (\ker \nu)_x$. Then U can be expanded as $U = u^i X_i(x)$. We now get

$$0 = \sum_{i=1}^{k} u^i \big\langle \mathrm{d}\rho_j, X_i(x)\big\rangle = m_{ij}u^i.$$

Under Assumption 1 we know that m_{ij} is invertible, which implies that $U = 0$. Thus, $(\ker \nu)_x \cap \mathsf{T}_x\mathcal{M}^H = 0$, which proves that ν restricted to $\mathcal{M}^H$ is nondegenerate.

3. For the final assertion, it is enough to show that $\gamma(t)$ is a solution to Eq. (1b) if and only if it is a solution to the Hamiltonian problem

$$\nu\big(\dot\gamma(t), U\big) = \big\langle \mathrm{d}H\big(\gamma(t)\big), U\big\rangle, \quad \forall U \in \mathsf{T}\mathcal{M}^H$$

on the symplectic manifold $\mathcal{M}^H$ (for which existence and uniqueness follows from standard ODE theory). As we have seen, under Assumption 1 every $X \in \mathsf{T}_{\gamma(t)}\mathcal{M}$ can be written as $X = U + W$ with $U \in \mathsf{T}\mathcal{M}^H$ and $W \in \ker \nu$. Now,

$$\nu\big(\dot\gamma(t), X\big) = \nu\big(\dot\gamma(t), U\big) \quad \text{and} \quad \big\langle \mathrm{d}H\big(\gamma(t)\big), X\big\rangle = \big\langle \mathrm{d}H\big(\gamma(t)\big), U\big\rangle$$

where the first and second equality follows, respectively, since $W \in \ker \nu$ and

$$\big\langle \mathrm{d}H\big(\gamma(t)\big), W\big\rangle = \Big\langle \mathrm{d}H\big(\gamma(t)\big), \sum_i w^i X_i\big(\gamma(t)\big)\Big\rangle = \sum_i w^i \rho_i\big(\gamma(t)\big) = 0.$$

This ends the proof. □

Remark 2.8 If the prescribed initial condition does not lie in the set $\mathcal{M}^H$, there cannot be any solution curve passing through this point. On the other hand, if $\mathcal{M}^H$ is a submanifold, and if it intersects the fibres of $[\mathcal{M}]$ cleanly, i.e., if the dimension of the intersection is constant, and if that dimension is larger than zero, then the equation may have infinitely many solutions. This is what happens if H is constant on the fibres of $[\mathcal{M}]$.

The following result will be useful in Sect. 4, when we analyse SHAKE and RATTLE.

Corollary 2.9 *Under Assumption 1, there exists an open set $O \subset \mathcal{M}$ containing $\mathcal{M}^H$ such that the equation $y \in [x] \cap \mathcal{M}^H$ has a unique solution for every $x \in O$. The corresponding trivially presymplectic projection map $\Pi\colon O \to \mathcal{M}^H$, defined by $\Pi(x) = y$, is a submersion.*

Proof This follows from Theorem 2.7 item 2, namely that for $x \in \mathcal{M}^H$, $\mathsf{T}_x\mathcal{M} = \mathsf{T}_x\mathcal{M}^H \oplus \ker \nu$. □

3 Coisotropic Constraints

In this section we study the geometry of problem (1a), (1b) from the extrinsic viewpoint. That is, we study properties of $\mathcal{M}$ as a submanifold of the symplectic manifold $(\mathcal{P}, \omega)$. Notice that $\nu := \iota_{\mathcal{M}}^* \omega$ is a presymplectic form on $\mathcal{M}$, since $\mathrm{d}\nu = \mathrm{d}\iota_{\mathcal{M}}^* \omega = \iota_{\mathcal{M}}^* \mathrm{d}\omega = 0$. Thus, any submanifold of a symplectic manifold is automatically a presymplectic manifold.

3.1 Lagrange Multipliers

Typically, a constraint manifold is defined in terms of a number of constraint functions. To this extent, let V be a vector space of dimension m, and denote by V^* its dual. Let $\mathsf{J}\colon \mathcal{P} \to V^*$ be a smooth function such that the constraint submanifold $\mathcal{M}$ is given by

$$\mathcal{M} = \mathsf{J}^{-1}(0) = \big\{ z \in \mathcal{P} : \mathsf{J}(z) = 0 \big\}. \tag{5}$$

If 0 is a regular value for J, i.e., if $\mathsf{T}\mathsf{J}(z)$ has full rank for all $z \in \mathcal{P}$ such that $\mathsf{J}(z) = 0$, then $\mathcal{M}$ is indeed a regular submanifold of $\mathcal{P}$. The dimension m of V is the number of constraints, i.e., the codimension of $\mathcal{M}$.

The problem (1a), (1b) may now be reformulated as finding a smooth curve

$$t \mapsto \big(z(t), \Lambda(t)\big) \in \mathcal{P} \times V$$

such that

$$\begin{aligned} \omega(\dot{z}) &= \mathrm{d}H(z) + \mathrm{d}\langle \mathsf{J}, \Lambda \rangle(z), \\ 0 &= \mathsf{J}(z). \end{aligned} \tag{6a}$$

Here, the notation $\langle \mathsf{J}, \Lambda \rangle$ means the smooth function $z \mapsto \langle \mathsf{J}(z), \Lambda \rangle$, depending on the parameter Λ. The equation can equivalently be written as

$$\begin{aligned} \dot{z} &= X_H(z) + X_{\langle \mathsf{J}, \Lambda \rangle}(z), \\ 0 &= \mathsf{J}(z). \end{aligned} \tag{6b}$$

We sometimes single out a basis $\{e_i\}_{i=1,\dots,m}$ of V and define the functions g_i by

$$g_i(z) := \big\langle \mathsf{J}(z), e_i \big\rangle. \tag{7}$$

Notice that in the case $\mathcal{P} = \mathbb{R}^{2d}$, Eqs. (6a), (6b) coincides with Eq. (2) in Example 1.1 above, with $\lambda = (\lambda_1, \dots, \lambda_m)$ being the coordinate vector of Λ, i.e., $\Lambda = \sum_{i=1}^m \lambda_i e_i$.

The system (6a), (6b) is again a DAE. Under Assumption 1, it follows from Theorem 2.7 above that this DAE has unique solutions for initial data in $\mathcal{M}^H$. From a DAE point of view, Assumption 1 asserts that system (6a), (6b) has index 3.

3.2 Coisotropy Assumption

Due to the solvability result imposed by Assumption 1, the Lagrange multipliers may be resolved as functions of z, which turns Eqs. (6a), (6b) into

$$\dot{z} = X_H(z) + \sum_{i=1}^{m} \lambda_i(z) X_{g_i}(z) =: X(z). \tag{8}$$

Notice that $X(z)$ is only defined for $z \in \mathcal{M}^H$ and also that $X(\mathcal{M}^H) \in \mathsf{T}\mathcal{M}^H$, so $X(z)$ defines an ODE on the hidden constraint manifold $\mathcal{M}^H$. From Theorem 2.7 it follows that its flow is symplectic. However, the individual vector fields $\lambda_i(z) X_{g_i}(z)$ are *not* Hamiltonian vector fields on $\mathcal{P}$ (assuming that $\lambda_i(z)$ is defined also outside of $\mathcal{M}^H$). In this section we present an assumption on the embedding $\iota_{\mathcal{M}} \colon \mathcal{M} \to \mathcal{P}$ which ensures that vector fields of the form $f(z) X_{g_i}(z)$ are trivially presymplectic vector fields on $\mathcal{M}$. As we will see in Sect. 4, this is essential in order to ensure presymplecticity and symplecticity of SHAKE and RATTLE.

Recall from Definition 1.6 that $\mathcal{M}$ is a *coisotropic* submanifold of $\mathcal{P}$ if $\mathsf{T}\mathcal{M}^{\perp} \subset \mathsf{T}\mathcal{M}$. Also recall Assumption 2 above (the coisotropy assumption), which states that $\mathcal{M}$ is a coisotropic submanifold. We continue with some consequences of Assumption 2, which are later used in the geometric analysis of SHAKE and RATTLE.

Remark 3.1 It is straightforward to verify that $\mathcal{M}$ being coisotropic is equivalent to $\mathsf{T}\mathcal{M}^{\perp}$ being *isotropic*, i.e., such that ω restricted to $\mathsf{T}\mathcal{M}^{\perp}$ is zero.

Remark 3.2 A coisotropic submanifold is such that the symplectic form becomes as degenerate as possible (given a fixed number of constraints) when restricted on the submanifold. More precisely, a coisotropic submanifold is such that the dimension of the distribution $\ker \nu$, i.e., dimension of the fibres of $[\mathcal{M}]$, is equal to the number of constraints m.

Remark 3.3 From a theoretical point of view, Assumption 2 is not a restriction on the presymplectic manifold $\mathcal{M}$, since every presymplectic submanifold may be coisotropically embedded in a symplectic manifold [4].

Remark 3.4 In practice as shown in Proposition 3.5, a sufficient condition for the manifold $\mathcal{M}$ defined by the equations $g_i(z) = 0$ for $1 \leq i \leq m$ to be coisotropic is simply that

$$\{g_i, g_j\}(z) = 0, \quad i, j = 1, \ldots, m, \forall z \in \mathcal{M}.$$

In particular, if the manifold $\mathcal{M}$ is defined by one constraint, i.e. if $m = 1$, then it is automatically a coisotropic submanifold.

The following result gives alternative characterisations of coisotropic submanifolds.

Proposition 3.5 *Suppose that $\mathcal{M}$ is a submanifold of $\mathcal{P}$. Then the following conditions are equivalent.*

1. $\mathcal{M}$ *is a coisotropic submanifold, i.e.,* $\mathsf{T}\mathcal{M}^{\perp} \subset \mathsf{T}\mathcal{M}$.
2. $\ker \nu = \mathsf{T}\mathcal{M}^{\perp}$.

Further, if $\mathcal{M}$ *is defined implicitly by* (5)*, then the conditions are also equivalent to*

3. *For any* $\alpha, \beta \in V$*, the functions* $\langle \mathsf{J}, \alpha \rangle$ *and* $\langle \mathsf{J}, \beta \rangle$ *are in involution on* $\mathcal{M}$*, i.e.,*

$$\big\{\langle \mathsf{J}, \alpha \rangle, \langle \mathsf{J}, \beta \rangle\big\}(z) = 0 \quad \text{for } z \in \mathcal{M}.$$

4. *For any* $\alpha \in V$*, the Hamiltonian vector field* $X_{\langle \mathsf{J}, \alpha \rangle}$ *is tangent to* $\mathcal{M}$.

In order to prove this, let us start with a lemma concerning the span of the Hamiltonian vector fields $X_{\langle \mathsf{J}, \alpha \rangle}$.

Lemma 3.6 *Define the distribution*

$$\mathcal{O} = \operatorname{span}\big\{X_{\langle \mathsf{J}, \alpha \rangle}(\mathcal{M}) : \alpha \in V\big\}.$$

Then $\mathsf{T}\mathcal{M}^{\perp} = \mathcal{O}$.

Proof We show that $\mathcal{O}^{\perp} = \mathsf{T}\mathcal{M}$, which is equivalent to the claim.

$$X \in \mathsf{T}\mathcal{M} \iff \big\langle \mathrm{d}\langle \mathsf{J}, \alpha \rangle, X \big\rangle = 0 \quad \forall \alpha \in V \iff \omega(X_{\langle \mathsf{J}, \alpha \rangle}, X) = 0 \quad \forall \alpha \in V. \quad \square$$

Proof of Proposition 3.5 We do it step by step.

$1 \leftrightarrow 2$ In general,

$$\ker \nu = \mathsf{T}\mathcal{M} \cap \mathsf{T}\mathcal{M}^{\perp},$$

so $\ker \nu = \mathsf{T}\mathcal{M}^{\perp} \iff \mathsf{T}\mathcal{M}^{\perp} \subset \mathsf{T}\mathcal{M}$, and that is the definition of coisotropicity of $\mathcal{M}$.

$1 \leftrightarrow 3$ First, for $x \in \mathcal{M}$

$$\{g_i, g_j\}(x) = 0 \iff \omega\big(X_{g_i}(x), X_{g_j}(x)\big) = 0,$$

so the functions g_i are in involution on $\mathcal{M}$ if and only if $\mathcal{O}$ (defined in Lemma 3.6) is isotropic, which is equivalent to $\mathcal{M}$ being coisotropic.

$3 \leftrightarrow 4$ Finally, it suffices to observe that for a point $x \in \mathcal{M}$,

$$\begin{aligned} X_{\langle \mathsf{J}, \alpha \rangle}(x) \in \mathsf{T}_x\mathcal{M} &\iff \big\langle \mathrm{d}\langle \mathsf{J}, \beta \rangle, X_{\langle \mathsf{J}, \alpha \rangle} \big\rangle(x) = 0 \quad \forall \beta \in V \\ &\iff \big\{\langle \mathsf{J}, \alpha \rangle, \langle \mathsf{J}, \beta \rangle\big\}(x) = 0 \quad \forall \beta \in V. \end{aligned} \quad \square$$

The following results follows directly from Lemma 3.6 and Proposition 3.5.

Corollary 3.7 *Let* $x \in \mathcal{M}$*. Then, under Assumption* 2*, the map*

$$V \ni \alpha \mapsto \exp(X_{\langle \mathsf{J}, \alpha \rangle})(x) \in [x]$$

is a local diffeomorphism. The fibre $[x]$ *is thus locally parametrised by* V.

Corollary 3.8 *Let $f \in \mathcal{C}^\infty(\mathcal{P})$ and $\alpha \in V$. Under Assumption 2 the vector field*

$$X(z) := f(z) X_{\langle \mathsf{J}, \alpha \rangle}(z)$$

is tangent to $\mathcal{M}$, and presymplectic when restricted to $\mathcal{M}$.

3.3 Relation Between X_H and $\mathcal{M}^H$

As a subset of $\mathcal{P}$, the hidden constraint set $\mathcal{M}^H$ is given by the points on $\mathcal{M}$ where the Hamiltonian vector field X_H is tangential to $\mathcal{M}$. Let Π denote the projection onto $\mathcal{M}^H$ defined in Corollary 2.9.

Proposition 3.9 *Under Assumption 2, the hidden constraint set $\mathcal{M}^H$ is*

$$\mathcal{M}^H = \left\{ z \in \mathcal{M} : X_H(z) \in \mathsf{T}_z \mathcal{M} \right\}.$$

Moreover, the differential equation (8) on $\mathcal{M}^H$ can be written

$$\dot{z} = \mathsf{T}_z \Pi \cdot X_H(z).$$

Proof 1. If $z \in \mathcal{M}^H$, then by definition there exists $Y \in \mathsf{T}_z \mathcal{M}$ such that

$$\langle \mathrm{d}H, X \rangle = \omega(Y, X) \quad \forall X \in \mathsf{T}_z \mathcal{M}.$$

Since $\omega(X_H) = \mathrm{d}H$, that is equivalent to

$$X_H - Y \in \mathsf{T}\mathcal{M}^\perp.$$

Noticing that Assumption 2 means that $\mathsf{T}\mathcal{M}^\perp \subset \mathsf{T}\mathcal{M}$, and using $Y \in \mathsf{T}\mathcal{M}$ yields $X_H \in \mathsf{T}_z \mathcal{M}$.

2. The differential equation on $\mathcal{M}^H$ is such that

$$\omega(\dot{z}, X) = \langle \mathrm{d}H, X \rangle \quad \forall X \in \mathsf{T}_z \mathcal{M}$$

so we obtain

$$\dot{z} - Y \in \mathsf{T}_z \mathcal{M}^\perp$$

and $\dot{z} = \mathsf{T}_z \Pi \cdot X_H(z)$. □

Remark 3.10 There are now several ways to compute $\mathcal{M}^H$. First, without any assumption, one can use the definition (3), and its immediate consequence Proposition 2.6. Under Assumption 2, one can also use Proposition 3.9. If the constraint manifold $\mathcal{M}$ is defined as in (5), a further useful description of $\mathcal{M}^H$ is

$$\mathcal{M}^H = \left\{ z \in \mathcal{M} : \{g_i, H\}(z) = 0, i = 1, \ldots, m \right\}.$$

This follows from the observation that

$$X_H \in \mathsf{T}_z \mathcal{M} \iff \langle \mathrm{d}g_i, X_H \rangle = 0, \quad i = 1, \ldots, m \tag{9}$$

and $\langle \mathrm{d}g_i, X_H \rangle = \{g_i, H\}$.

Based on Theorem 2.7, Proposition 3.5 and Lemma 3.6, we can say much more on the behaviour of X_H in a neighbourhood of $\mathcal{M}^H$. Indeed, we have the following result, which is a key ingredient in the well-posedness of SHAKE and RATTLE, as will be explained in Sect. 4.

Lemma 3.11 *Let $y \in \mathcal{M}^H$ and define the function*

$$\begin{aligned} F\colon [y] &\longrightarrow V^* \\ x &\longmapsto \mathrm{d}\mathsf{J}(x) \cdot X_H(x). \end{aligned}$$

Then, under Assumption 1 and Assumption 2, the differential of F at y, i.e., the linear mapping

$$\mathrm{d}F(y) : \mathsf{T}_y[y] \longrightarrow V^*$$

is invertible.

Proof In terms of the previously introduced basis $\{e_i\}_{i=1,\dots,m}$, the function F is given by

$$F(x) = \sum_{i=1}^{m} \big\langle \mathrm{d}g_i(x), X_H(x) \big\rangle e_i = -\sum_{i=1}^{m} \big\langle \mathrm{d}H(x), X_{g_i}(x) \big\rangle e_i.$$

Under Assumption 2, it follows from Proposition 3.5 and Lemma 3.6 that $X_{g_1}(y), \dots, X_{g_m}(y)$ is a basis for $\mathsf{T}_y[y]$. Relative to this basis, and the basis $\{e_i\}_{i=1,\dots,m}$ of V, the Jacobian matrix of $\mathrm{d}F(y)$ is given by

$$m_{ij} := \big\langle \mathrm{d}\big\langle \mathrm{d}H(y), X_{g_i}(y) \big\rangle, X_{g_j}(y) \big\rangle = \big\{ \{H, g_i\}, g_j \big\}(y).$$

Define $\rho_i(x) := \langle \mathrm{d}H(x), X_{g_i}(x) \rangle$. Then $m_{ij} = \langle \mathrm{d}\rho_i(y), X_{g_j}(y) \rangle$. Since $y \in \mathcal{M}^H$ we have $\rho_i(y) = 0$. Now, under Assumption 1 and the exact same argument as in the proof of Theorem 2.7, it follows that m_{ij} is invertible. This concludes the proof. □

4 Geometry of SHAKE and RATTLE

Geometrically, the basic principle of SHAKE, defined in Definition 1.3, with φ_h as an underlying method, can be described as follows. For some initial data $z \in \mathcal{M}$, slide along the fibre with $z^+ \in [z]$ such that $z^- = \varphi_h(z^+)$ "lands" again on the submanifold $\mathcal{M}$. The RATTLE method, defined in Definition 1.4, is then a post-processed version of SHAKE, which is described geometrically as follows. For some initial data $z_0 \in \mathcal{M}^H$, take one step with SHAKE landing on $z_1^- \in \mathcal{M}$, then slide along the fibre $[z_1^-]$ to end up on $z_1 \in \mathcal{M}^H \cap [z_1^-]$. This process is visualised in Fig. 2.

If we assume that the SHAKE map $S_h\colon O \to O$ is well defined for some open subset $O \subset \mathcal{M}$ containing $\mathcal{M}^H$, then we may define a sliding map $\Gamma_{\varphi_h} = \varphi_h^{-1} \circ S_h$. Consequently, it follows from Definition 1.3 that the sliding map $O \ni z \mapsto \Gamma_{\varphi_h}(z) \in O$ is defined implicitly by the equation

$$\varphi_h\big(z^+\big) \in \mathcal{M}, \quad z^+ \in [z] \cap O, \qquad \Gamma_{\varphi_h}(z) = z^+. \tag{10}$$

Notice that:

(i) Γ_{φ_h} is fibre preserving, i.e., trivially presymplectic, and
(ii) Γ_{φ_h} is a projection, i.e., $\Gamma_{\varphi_h} \circ \Gamma_{\varphi_h} = \Gamma_{\varphi_h}$.

Since $S_h = \varphi_h \circ \Gamma_{\varphi_h}$ it follows that SHAKE, if it is well-defined, is a *fibre mapping*, i.e., it maps fibres to fibres. It is, in fact, a little bit more than that, since it maps the whole fibre $[z] \cap O$ to the same point $S_h(z)$. Hence, when using SHAKE it is not important where on the initial fibre $[z]$ one starts (as long as it is close enough to $\mathcal{M}^H$ so that SHAKE is well-defined). Furthermore, regardless of where on the fibre one starts, after one step SHAKE remains on the *modified* hidden constraint set, given by

$$\widetilde{\mathcal{M}}^H := \varphi_h \circ \Gamma_{\varphi_h}(O).$$

Since Γ_{φ_h} is a projection, $\widetilde{\mathcal{M}}^H$ is strictly smaller than O. If SHAKE is well-defined, $\widetilde{\mathcal{M}}^H$ is in fact a symplectic submanifold of $\mathcal{M}$ (see Proposition 4.7 below).

Let Π be the projection on $\mathcal{M}^H$ given in Corollary 2.9. Assume that Π is well-defined on O. Then RATTLE is given by $R_h = \Pi \circ S_h$. Notice that SHAKE and RATTLE define exactly the same fibre mapping. In particular,

$$\Pi \circ (S_h)^k = \Pi \circ (R_h)^k. \tag{11}$$

Thus, RATTLE is only a cosmetic improvement of SHAKE, and has no influence on the numerical scheme except at the last step.

We now give explicit conditions under which SHAKE and RATTLE are well defined and can be computed. More precisely:

1. When is a method φ_h such that the corresponding SHAKE and RATTLE methods are well defined?
2. How can we parameterise Γ_{φ_h} (so that S_h is computable)?
3. Will SHAKE and RATTLE converge to the solution of Eqs. (1a), (1b) as $h \to 0$?
4. Are S_h and R_h presymplectic as mappings $O \to O$?
5. Is R_h symplectic as a mapping $\mathcal{M}^H \to \mathcal{M}^H$, and S_h symplectic as a mapping $\widetilde{\mathcal{M}}^H \to \widetilde{\mathcal{M}}^H$?
6. Can S_h or R_h be reversible?

These questions are addressed in the remainder of this section.

4.1 Well-Posedness

In order for SHAKE and RATTLE to be well defined, we need the "sliding process" to have a locally unique solution. Whether so or not depends on the map φ_h.

Theorem 4.1 *Suppose that Assumption 1 and Assumption 2 hold. Consider a (smooth) method φ_h, consistent with $\dot{z} = X_H(z)$. Then for h small enough and for $z \in \mathcal{M}$ in a neighbourhood of $\mathcal{M}^H$ the equation*

$$\mathsf{J}\big(\varphi_h\big(z^+\big)\big) = 0, \quad z^+ \in [z] \tag{12}$$

has a unique solution.

Proof Let $z \in \mathcal{M}^H$. We define the function $F_h\colon [z] \to V^*$ by

$$F_h\big(z^+\big) := \int_0^1 \frac{\partial}{\partial h}\big(\mathsf{J}\big(\varphi_h\big(z^+\big)\big)\big)(h\tau)\,\mathrm{d}\tau.$$

We see that F_h depends smoothly on h. Notice that for $h \neq 0$,

$$F_h\big(z^+\big) = \frac{\mathsf{J}(\varphi_h(z^+)) - \mathsf{J}(\varphi_0(z^+))}{h} = \frac{\mathsf{J}(\varphi_h(z^+))}{h},$$

because $\varphi_0(z^+) = z^+$ and $z^+ \in \mathcal{M} \iff \mathsf{J}(z^+) = 0$.

Consider now the case $h = 0$. The method φ_h is consistent, so $(\mathrm{d}\varphi_h/\mathrm{d}h)|_{h=0}(z^+) = X_H(z^+)$, which shows that

$$F_0\big(z^+\big) = \mathrm{d}\mathsf{J}\big(z^+\big) \cdot X_H\big(z^+\big). \tag{13}$$

We want to find a neighbourhood $O_h \subset [z]$ of z such that the equation $F_h(z^+) = 0$ with $z^+ \in O_h$ has a unique solution for small enough h. This will prove the claim. The strategy is to show that $\mathrm{d}F_h(z)\colon \mathsf{T}_z[z] \to V^*$ is non-singular. We start with the case $h = 0$.

Using (13) and appealing to Lemma 3.11, $\mathrm{d}F_0(z)$ is invertible under Assumption 1 and Assumption 2. Thus, by the inverse function theorem, we can find an open neighbourhood $O_0 \subset [z]$ such that $F_0\colon O_0 \to F_0(O_0)$ is a diffeomorphism. Also, since F_h depends smoothly on h, it follows that $\mathrm{d}F_h(z)$ is invertible for small enough h. Thus, for small enough h, we can find an open neighbourhood $O_h \subset [z]$ such that $F_h\colon O_h \to F_h(O_h)$ is a diffeomorphism.

It remains now to show that $0 \in F_h(O_h)$, so that the equation $F_h(z^+) = 0$ with $z^+ \in O_h$ has a unique solution. Now, if $z \in \mathcal{M}^H$ it follows from Proposition 3.9 that $X_H(z)$ is tangential to $\mathcal{M}$, which means that $\mathrm{d}\mathsf{J}(z) \cdot X_H(z) = 0$, so we get $F_0(z) = 0$. Thus, $0 \in F_0(O_0)$, and it follows by smoothness that $0 \in F_h(O_h)$ for small enough h. □

Remark 4.2 The coisotropy assumption is essential for the result of Theorem 4.1, because it uses Lemma 3.11 which depends on that assumption in an essential manner. On the other hand, if (12) has a unique solution in a neighbourhood of $\mathcal{M}^H$, then Assumption 2 must hold.

Indeed, the theorem implies that $[z]$ is locally diffeomorphic to V^*. This implies that their dimension is the same, so $\ker \nu$ must have the same dimension as V^*.

As a result,

$$\dim \ker \nu = \dim V^* = \operatorname{codim} \mathcal{M} = \dim \mathsf{T}\mathcal{M}^\perp.$$

Now, in general $\ker \nu = \mathsf{T}\mathcal{M} \cap \mathsf{T}\mathcal{M}^\perp$, so we get $\mathsf{T}\mathcal{M}^\perp \subset \mathsf{T}\mathcal{M}$, which implies that $\mathcal{M}$ is a coisotropic submanifold.

The result of Theorem 4.1 allows one to define the sliding map Γ_{φ_h} by (10). The corresponding SHAKE and RATTLE maps, S_h and R_h, are thus defined when h is sufficiently small.

4.2 Fibre Parametrisation

When the manifold $\mathcal{M}$ is defined implicitly as the locus of functions g_i, one can express the effect of Γ_{φ_h} using the flows of the Hamiltonian vector fields X_{g_i}.

Proposition 4.3 *Suppose that the φ_h-SHAKE method is well defined, so that Γ_{φ_h} is well defined. Under Assumption 2 there exist functions $\lambda_1, \ldots, \lambda_m \in \mathcal{C}^\infty(\mathcal{M})$ such that the sliding map Γ_{φ_h} is given by*

$$\Gamma_{\varphi_h}(x) = \exp\left(h \sum_{i=1}^{m} \lambda_i(x) X_{g_i}\right)(x).$$

Proof Follows directly from Corollary 3.7. □

4.3 Convergence

We may now show the convergence of SHAKE and RATTLE. The proof is essentially a standard convergence argument for RATTLE, which is a numerical method on the manifold $\mathcal{M}^H$. The convergence of SHAKE is then obtained using (11).

Proposition 4.4 *Suppose that Assumption 1 and Assumption 2 hold. Let φ_h be a method consistent with $\dot{z} = X_H(z)$. Then the φ_h-SHAKE and φ_h-RATTLE methods are convergent of order at least* 1.

Proof We first turn to RATTLE. The continuous system is an ordinary differential equation on $\mathcal{M}^H$ with vector field $\mathsf{T}\Pi \circ X_H$ (see Proposition 3.9). Since it is an ordinary differential equation, we only need to show that RATTLE (defined in Definition 1.4) is consistent and standard arguments may then be used to show convergence of order 1 (see [7, § II.3]). Using that $R_h = \Pi \circ \varphi_h \circ \Gamma_{\varphi_h}$ and differentiating at $h = 0$ we get

$$\left.\frac{\partial R_h}{\partial h}\right|_{h=0} = \mathsf{T}\Pi \circ \left(\left.\frac{\partial \varphi_h}{\partial h}\right|_{h=0} + \left.\frac{\partial \Gamma_{\varphi_h}}{\partial h}\right|_{h=0}\right),$$

which follows since $\varphi_0 = \mathrm{id}$ and $\Gamma_{\varphi_0} = \mathrm{id}$. The second term is in the kernel of $\mathsf{T}\Pi$, so

$$\left.\frac{\partial R_h}{\partial h}\right|_{h=0} = \mathsf{T}\Pi \circ \left.\frac{\partial \varphi_h}{\partial h}\right|_{h=0}.$$

The assumption that φ_h is consistent with $\dot{z} = X_H(z)$ means that $\frac{\partial \varphi_h}{\partial h}|_{h=0} = X_H$. Consistency of R_h then follows.

Now, using (11) and that $\Gamma_{\varphi_h}(z) = z + \mathcal{O}(h)$ when $h \to 0$, we also see that SHAKE converges of order at least 1. □

Remark 4.5 Interestingly, the local error of SHAKE is only $\mathcal{O}(h)$, so standard arguments do not apply to show convergence. However, due to the fact that SHAKE is the same fibre mapping as RATTLE, this error does not accumulate, and the global error is still $\mathcal{O}(h)$.

4.4 Symplecticity

We examine in which sense SHAKE and RATTLE may be regarded as symplectic methods. The essential result is that both SHAKE and RATTLE are *presymplectic*, i.e., they preserve the presymplectic structure of $\mathcal{M}$.

Proposition 4.6 *Let φ_h be a symplectic method. Then the corresponding* SHAKE *map S_h and* RATTLE *map R_h, regarded as mappings $\mathcal{M} \to \mathcal{M}$, are presymplectic.*

Proof Since Γ_{φ_h} is trivially presymplectic, it is in particular presymplectic, so

$$\nu(u, v) = \nu(\mathsf{T}\Gamma_{\varphi_h} \cdot u, \mathsf{T}\Gamma_{\varphi_h} \cdot v), \quad \forall u, v \in \mathsf{T}\mathcal{M}.$$

Further, since φ_h is symplectic it follows that

$$\begin{aligned}\nu(u, v) &= \nu\big(\mathsf{T}\varphi_h \cdot (\mathsf{T}\Gamma_{\varphi_h} \cdot u), \mathsf{T}\varphi_h \cdot (\mathsf{T}\Gamma_{\varphi_h} \cdot v)\big) \\ &= \nu\big(\mathsf{T}(\varphi_h \circ \Gamma_{\varphi_h}) \cdot u, \mathsf{T}(\varphi_h \circ \Gamma_{\varphi_h}) \cdot v\big) \\ &= \nu(\mathsf{T}S_h \cdot u, \mathsf{T}S_h \cdot v), \quad \forall u, v \in \mathsf{T}\mathcal{M}.\end{aligned}$$

Thus, S_h is presymplectic.

Moreover, since $R_h = \Pi \circ S_h$ and Π is trivially presymplectic, R_h is also presymplectic. □

Proposition 4.7 *Let φ_h be a symplectic method. Under Assumption 1 and Assumption 2, the set $\widetilde{\mathcal{M}}^H$ is a symplectic submanifold of $\mathcal{M}$, with symplectic form $\tilde{\varpi} = \iota^*_{\widetilde{\mathcal{M}}^H}\nu$.*

Proof Under Assumption 1 it follows from Theorem 2.7 that the set $(\mathcal{M}^H, \varpi)$ is a symplectic submanifold of $\mathcal{M}$. We first show that $\widetilde{\mathcal{M}}^H$ is diffeomorphic to $\widetilde{\mathcal{M}}^H$, and thus also a submanifold of $\mathcal{M}$, by constructing a diffeomorphism $\mathcal{M}^H \to \widetilde{\mathcal{M}}^H$. Under Assumption 1 and Assumption 2, the map $\Gamma_{\varphi_h} \colon \mathcal{M}^H \to \mathcal{M}$ is well defined. Let $z \in \mathcal{M}^H$. By construction we have $\ker(\mathsf{T}_z\Gamma_{\varphi_h}) = \mathsf{T}_z[z]$. Thus, $\mathsf{T}_z\Gamma_{\varphi_h} \colon \mathsf{T}_z\mathcal{M}^H \to \mathsf{T}_{\Gamma_{\varphi_h}(z)}\mathcal{M}$ is injective, so $\Gamma_{\varphi_h} \colon \mathcal{M}^H \to \Gamma_{\varphi_h}(\mathcal{M}^H)$ is a diffeomorphism. Next, since $\varphi_h \colon \mathcal{P} \to \mathcal{P}$ is a diffeomorphism, it also holds that $\varphi_h \colon \Gamma_{\varphi_h}(\mathcal{M}^H) \to \varphi_h(\Gamma_{\varphi_h}(\mathcal{M}^H)) = \widetilde{\mathcal{M}}^H$ is a diffeomorphism. Thus, $S_h = \varphi_h \circ \Gamma_{\varphi_h}$ is a diffeomorphism as a map $\mathcal{M}^H \to \widetilde{\mathcal{M}}^H$, so $\mathcal{M}^H$ and $\widetilde{\mathcal{M}}^H$ are diffeomorphic.

Next, we show that the form $\tilde{\varpi} = \iota^*_{\widetilde{\mathcal{M}}^H}\nu$ is nondegenerate. Let $y = S_h(z)$ and $u, v \in \mathsf{T}_y\widetilde{\mathcal{M}}^H$. From Proposition 4.6 it follows that S_h is presymplectic, so

$$\nu(u, v) = \nu\big(\mathsf{T}_y S_h^{-1} \cdot u, \mathsf{T}_y S_h^{-1} \cdot v\big) = \varpi\big(\mathsf{T}_y S_h^{-1} \cdot u, \mathsf{T}_y S_h^{-1} \cdot v\big).$$

Since $\mathsf{T}_y S_h^{-1} \colon \widetilde{\mathcal{M}}^H \to \mathcal{M}^H$ is invertible, and since $(\widetilde{\mathcal{M}}^H, \varpi)$ is a symplectic submanifold, it follows that $\nu(u, v) = 0$ for all $v \in \mathsf{T}_y\widetilde{\mathcal{M}}^H$ only if $u = 0$, which shows that $(\widetilde{\mathcal{M}}^H, \tilde{\varpi})$ is a symplectic manifold. □

The following result follows directly from Theorem 2.7, Proposition 4.6, and Proposition 4.7.

Corollary 4.8 *Let φ_h be a symplectic method and assume that Assumption 1 and Assumption 2 hold. Then*:

- *The* SHAKE *map S_h, regarded a mapping $\widetilde{\mathcal{M}}^H \to \widetilde{\mathcal{M}}^H$, is symplectic.*
- *The* RATTLE *map R_h, regarded as a mapping $\mathcal{M}^H \to \mathcal{M}^H$, is symplectic.*

4.5 Time Reversibility

Just as in the holonomic case, if the underlying method φ_h is *symmetric*, i.e., $\varphi_h^{-1} = \varphi_{-h}$, then RATTLE is also symmetric and thus of second order. Note that SHAKE can never be symmetric, because although S_h preserves $\widetilde{\mathcal{M}}^H$, the reverse method S_{-h} does not.

Proposition 4.9 *If the underlying method φ_h is symmetric, then so is* RATTLE, *considered as a method from $\mathcal{M}^H$ to $\mathcal{M}^H$.*

Proof This follows from the symmetry property of the map Γ_{φ_h}. In general, if $z_0^+ = \Gamma_\varphi(z_0)$, $z_1^- = \varphi(z_0^+)$, and $z_1 = \Pi(z_1^-)$ (see Fig. 2), then

$$z_1^- = \Gamma_{\varphi^{-1}}(z_1). \tag{14}$$

This follows from Theorem 4.1, because z_1^- is a solution of the equation $\mathsf{J}(\varphi^{-1}(z_1)) = 0$, so it must be equal to $\Gamma_{\varphi^{-1}}(z_1)$.

Suppose that we start from a point $z_0 \in \mathcal{M}^H$. The image by the RATTLE map is $z_1 = R_h(z_0)$. Now, since $z_0 = \Pi(z_0^+)$ (as we assumed that $z_0 \in \mathcal{M}^H$), and (14), we obtain $z_0 = \Pi \circ \varphi_h^{-1} \circ \Gamma_{\varphi_h^{-1}}$.

If we assume now that φ_h is symmetric, i.e., $\varphi_h^{-1} = \varphi_{-h}$, we obtain $z_0 = R_{-h}(z_1)$, so RATTLE is symmetric. □

5 Examples

In this section we give examples of constrained problems that can be solved with the geometric SHAKE and RATTLE methods. In Sect. 5.2 we also give non-trivial examples where the nondegeneracy assumption fails to hold.

5.1 Holonomic Case

In this section we study the classical, so-called *holonomic* case, where the constraints depend only on the position q, and not on the momentum p.

5.1.1 Classical Assumptions

Consider the classical case of constrained mechanical systems, where the symplectic manifold $\mathcal{P}$ is simply $\mathbb{R}^{2n}$, with coordinates (q^i, p_i), equipped with the canonical symplectic form $\omega = \sum_i \mathrm{d}q^i \wedge \mathrm{d}p_i$. The constraints are given by $g(q) = 0$ for a

function g defined from $\mathcal{P}$ to $V := \mathbb{R}^m$. It should be emphasised that g does not depend on p, i.e., $g_p = 0$. The fibre passing through the point $z = (q, p)$ is then an affine subspace described parametrically by

$$\left\{\left(q, p + g_q(q)^{\mathsf{T}}\lambda\right) : \lambda \in \mathbb{R}^m\right\}.$$

Assumption 2 is automatically fulfilled, because $\frac{\partial g_i}{\partial p_j} = 0$ implies

$$\{g_i, g_j\} = 0.$$

Moreover, the standard assumption placed on the Hamiltonian H is that the matrix

$$g_q(q)H_{pp}(q, p)g_q(q)^{\mathsf{T}} \quad \text{be invertible} \tag{15}$$

(see [8, § VII.1.13], [15]), which implies Assumption 1.

The condition (15) is more stringent than Assumption 1, since the latter only involves critical points of H on a fibre, i.e., on the points lying on the hidden constraint manifold $\mathcal{M}^H$. It is clear that the condition (15) is too restrictive when the fibres are compact, and we will examine such an example in Sect. 5.4.

5.1.2 Integrators on Cotangent Bundles

We explain here how to construct symplectic integrators on cotangent bundles $\mathsf{T}^*\mathcal{Q}$, for a configuration manifold

$$\mathcal{Q} := \left\{q \in \mathbb{R}^n : g(q) = 0\right\}. \tag{16}$$

The constraint manifold $\mathcal{M}$ is then

$$\mathcal{M} := \left\{(q, p) \in \mathsf{T}^*\mathbb{R}^n : g(q) = 0\right\}. \tag{17}$$

The crucial observation is that $\mathsf{T}^*\mathcal{Q}$ is *canonically* symplectomorphic to $[\mathcal{M}]$. We can now use SHAKE to construct a symplectic integrator on $\mathsf{T}^*\mathcal{Q}$ as follows.

Suppose that we are given a Hamiltonian H on $\mathsf{T}^*\mathbb{R}^n$. Starting from a point $\xi_0 \in \mathsf{T}^*\mathcal{Q} \simeq [\mathcal{M}]$, one lifts it to a point $z_0 \in \mathcal{M}$ such that $[z_0] = \xi_0$. SHAKE then produces a new point z_1^-, which by projection gives $\xi_1 := [z_1^-]$. Note that ξ_1 does not depend on the choice of the point z_0 on the fibre. In particular, note that the mapping $\xi_0 \mapsto \xi_1$ is exactly the same if RATTLE is used instead of SHAKE. The mapping $\xi_0 \mapsto \xi_1$ is then a symplectic integrator on $\mathsf{T}^*\mathcal{Q}$ for the Hamiltonian on $[\mathcal{M}]$ given by the projection of $\imath^*_{\mathcal{M}^H} H$. This method has the same order of convergence as RATTLE.

For instance, in Example 1.5 the configuration manifold $\mathcal{Q}$ is a circle, so $\mathcal{M}$ is the vector bundle of planes above the points of the circle. Each such plane is spanned by two vectors: one normal and one tangential to the circle. SHAKE and RATTLE differ only by the normal component. Since a co-vector in $\mathsf{T}^*\mathcal{Q}$ is defined by its scalar product with tangent vectors in $\mathsf{T}\mathcal{Q}$, its normal component has no effect, so SHAKE and RATTLE have identical effect as integrators on $\mathsf{T}^*\mathcal{Q}$.

5.1.3 *Classical* SHAKE *and* RATTLE

The classical SHAKE method is only defined for separable Hamiltonian functions (see [8, § VII.1.23], [11, 15]). However, it is readily extended to general Hamiltonians as the mapping $(q_0, p_0) \mapsto (q_1, p_1^-)$ defined by

$$\begin{aligned} p_0^+ &= p_0 - \frac{h}{2} g_q(q_0)^\mathsf{T} \lambda, \\ p_{1/2} &= p_0^+ - \frac{h}{2} H_q(q_0, p_{1/2}), \\ q_1 &= q_0 + \frac{h}{2}\big(H_p(q_0, p_{1/2}) + H_p(q_1, p_{1/2})\big), \\ p_1^- &= p_{1/2} - \frac{h}{2} H_q(q_1, p_{1/2}), \\ 0 &= g(q_1). \end{aligned}$$

The classical RATTLE method is obtained by adding a projection step onto the manifold $\mathcal{M}^H$, i.e., the next step is instead $(q_1, p_1^-) \mapsto (q_1, p_1)$ defined, on top of SHAKE, as

$$\begin{aligned} p_1 &= p_1^- + g_q(q_1)^\mathsf{T} \mu, \\ 0 &= g_q(q_1) H_p(q_1, p_1). \end{aligned}$$

We see that the mapping $(q_0, p_0^+) \mapsto (q_1, p_1^-)$ is the Störmer–Verlet method, so the classical SHAKE and RATTLE methods are obtained using the Störmer–Verlet method as the underlying unconstrained symplectic method.

Similarly, in [8, § VII.1.3], a first order method is described as

$$\begin{aligned} p_0^+ &= p_0 - h g_q(q_0)^\mathsf{T} \lambda, \\ p_1^- &= p_0^+ - h H_q\big(q_0, p_1^-\big), \\ q_1 &= q_0 + h H_p\big(q_0, p_1^-\big), \\ 0 &= g(q_1). \end{aligned}$$

We see that the mapping $(q_0, p_0^+) \mapsto (q_1, p_1^-)$ is the symplectic Euler method, so this is the SHAKE method obtained using symplectic Euler as the underlying unconstrained symplectic method. When complemented with the projection step

$$\begin{aligned} p_1 &= p_1^- - h g_q(q_1)^\mathsf{T} \mu, \\ 0 &= g_q(q_1) H_p(q_1, p_1), \end{aligned}$$

that method becomes the corresponding RATTLE method.

5.2 Examples of Failing Nondegeneracy Assumption

Let us examine an example where Assumption 1 is not fulfilled. The manifold $\mathcal{P}$ is $\mathbb{R}^4$ with coordinates $q, p, \overline{q}, \overline{p}$, and the symplectic form is $\omega = \mathrm{d}q \wedge \mathrm{d}p + \mathrm{d}\overline{q} \wedge \mathrm{d}\overline{p}$. The constraint function J is

$$\mathsf{J}(q, p, \overline{q}, \overline{p}) = \overline{q}.$$

The corresponding fibres are the lines parametrised by $\overline{p}$, i.e., the fibres are the lines of equation $(q, p, \overline{q}) = (q_0, p_0, \overline{q}_0)$. We choose the Hamiltonian

$$H = \frac{p^2}{2} + \overline{p}q.$$

The hidden constraint manifold is then $\mathcal{M}^H = \{q = 0\}$. The restriction of H on a fibre is a linear function (because H is linear in $\overline{p}$), so its critical points are degenerate, and Assumption 1 is not fulfilled. As a result, there are extra constraints which prevent the problem from being well-posed on $\mathcal{M}^H$. It is readily verified that the system has in fact index five instead of three, the tertiary and quaternary constraints being, respectively, $p = 0$ and $\overline{p} = 0$.

If we had assumed instead that $H = \frac{p^2}{2}$, the hidden constraint set would be $\mathcal{M}^H = \mathcal{M}$, because H is now constant on the fibres (because H does not depend on $\overline{p}$). In that case, there are infinitely many solutions to the problem at hand.

5.3 Index 1 Constraints

Let $\mathcal{P} = \mathbb{R}^{2n+2k}$ with coordinates $q \in \mathbb{R}^n$, $p \in \mathbb{R}^n$, $\alpha \in \mathbb{R}^k$, $\beta \in \mathbb{R}^k$, symplectic form $\omega = \mathrm{d}q \wedge \mathrm{d}p + \mathrm{d}\alpha \wedge \mathrm{d}\beta$, Hamiltonian $H(q, p, \beta)$, and consider the holonomic constraint $\alpha = 0$. Because the constraints are holonomic, the coisotropy assumption is satisfied (see Sect. 5.1). The presymplectic manifold is $\mathcal{M} = \mathbb{R}^{2n+k}$ with coordinates (q, p, β) and presymplectic form $\nu = \mathrm{d}q \wedge \mathrm{d}p$. The hidden constraint submanifold is $\mathcal{M}^H = \{(q, p, \lambda) : H_\beta(q, p, \beta) = 0\}$, giving the presymplectic system

$$\begin{aligned} \dot{q} &= H_p(q, p, \beta), \\ \dot{p} &= -H_q(q, p, \beta), \\ 0 &= H_\beta(q, p, \beta). \end{aligned} \tag{18}$$

The nondegeneracy assumption is that $H_{\beta\beta}$ is non-singular on $\mathcal{M}^H$. When this holds, the secondary constraint can be solved for $\beta = B(q, p)$. This is the case of index one constraints considered in [14]. In this paper, a symplectic integrator is introduced and studied for the index one constrained Hamiltonian system (18), consisting of a direct application of a symplectic Runge–Kutta method to the full system (18). In the case when φ_h is the midpoint rule, we now show that this method is equivalent to the application of SHAKE or RATTLE applied on $\mathcal{P}$. Indeed, Hamilton's equations on $\mathcal{P}$

are given by

$$\begin{aligned}\dot{q} &= H_p(q, p, \beta),\\ \dot{p} &= -H_q(q, p, \beta),\\ \dot{\alpha} &= H_\beta(q, p, \beta),\\ \dot{\beta} &= 0,\end{aligned}$$

for which the midpoint discretisation, defining φ_h, is

$$\begin{aligned}\frac{q_1 - q_0}{h} &= H_p\left(\frac{q_0+q_1}{2}, \frac{p_0+p_1}{2}, \frac{\beta_0+\beta_1}{2}\right),\\ \frac{p_1 - p_0}{h} &= -H_q\left(\frac{q_0+q_1}{2}, \frac{p_0+p_1}{2}, \frac{\beta_0+\beta_1}{2}\right),\\ \frac{\alpha_1 - \alpha_0}{h} &= H_\beta\left(\frac{q_0+q_1}{2}, \frac{p_0+p_1}{2}, \frac{\beta_0+\beta_1}{2}\right),\\ \beta_1 - \beta_0 &= 0,\end{aligned} \tag{19}$$

while the constraint flow is given by $\exp(\lambda X_\alpha)\colon (q, p, \alpha, \beta) \mapsto (q, p, \alpha, \beta - \lambda)$. Initially, the primary constraints are satisfied, i.e. $\alpha_0 = 0$, and the constraint flow determines β_1 so that $\alpha_1 = 0$, i.e., so that

$$0 = H_\beta\left(\frac{q_0+q_1}{2}, \frac{p_0+p_1}{2}, \frac{\beta_0+\beta_1}{2}\right)$$

with solution $\frac{\beta_0+\beta_1}{2} = B(\frac{q_0+q_1}{2}, \frac{p_0+p_1}{2})$. This is exactly the method of [14] in the case of the midpoint rule. The final step of RATTLE chooses β so that $0 = H_\beta(q_1, p_1, \beta)$, but (as we have seen in the general case), this value is irrelevant for the next step.

5.4 Harmonic Constraints

Consider the symplectic manifold $\mathcal{P} = \mathsf{T}^*\mathbb{R}^2$ equipped with the canonical symplectic form

$$\omega(q_0, p_0, q_1, p_1) = \mathrm{d}q_0 \wedge \mathrm{d}p_0 + \mathrm{d}q_1 \wedge \mathrm{d}p_1,$$

where $\mathbf{q} = (q_0, q_1)$ and $\mathbf{p} = (p_0, p_1)$ are canonical coordinates. We often make the identification $\mathsf{T}^*\mathbb{R}^2 \simeq \mathbb{C}^2$ by introducing the complex coordinates $z_0 = q_0 + \mathrm{i}p_0$ and $z_1 = q_1 + \mathrm{i}p_1$.

Consider now the constraint space $V = \mathbb{R}$ and the single constraint function

$$g(\mathbf{q}, \mathbf{p}) := \|\mathbf{q}\|^2 + \|\mathbf{p}\|^2 - 1. \tag{20}$$

Note that the constraint manifold is given by $\mathcal{M} = S^3$, i.e., the unit sphere in $\mathbb{R}^4$. It is automatically a coisotropic submanifold since it has codimension one (see Remark 3.4).

The flow of X_g expressed in complex coordinates is given by

$$\exp(tX_g)(z_0, z_1) = \left(e^{it} z_0, e^{it} z_1\right).$$

Thus, it follows from Corollary 3.7 that the fibre $[(z_0, z_1)]$ passing through (z_0, z_1) is given by the great circle

$$\left[(z_0, z_1)\right] = \left\{\left(e^{is} z_0, e^{is} z_1\right) \in \mathbb{C}^2 : s \in \mathbb{R}\right\}.$$

These fibres are exactly the fibres in the classical Hopf fibration of the 3-sphere in circles. Recall that the Hopf map is the submersion

$$(z_0, z_1) \mapsto \left(2z_0\bar{z_1}, |z_0|^2 - |z_1|^2\right), \tag{21}$$

which projects $\mathbb{C}^2$ onto $\mathbb{R}^3 \simeq \mathbb{C} \times \mathbb{R}$ and maps 3-spheres onto 2-spheres. The quotient set $[\mathcal{M}]$ is thus a manifold which is diffeomorphic to the 2-sphere.

We now investigate the validity of Assumption 1 for two different choices of the Hamiltonian function.

5.4.1 Hamiltonian without Potential Energy

First consider the Hamiltonian

$$H(\mathbf{q}, \mathbf{p}) = \frac{\|\mathbf{p}\|^2}{2}.$$

The governing equations (6a), (6b) are in this case given by

$$\begin{aligned} \dot{\mathbf{q}} &= (1 + \lambda)\mathbf{p}, \\ \dot{\mathbf{p}} &= -\lambda\mathbf{q}, \\ 1 &= \|\mathbf{q}\|^2 + \|\mathbf{p}\|^2. \end{aligned} \tag{22}$$

Let $(\mathbf{q}, \mathbf{p}) \in \mathcal{M}$ and let

$$\mathcal{H}(\theta) := H\left(\cos(\theta)\mathbf{q} + \sin(\theta)\mathbf{p}, -\sin(\theta)\mathbf{q} + \cos(\theta)\mathbf{p}\right),$$

so that $\mathcal{H}$ is the restriction of H to the fibre containing $(\mathbf{q}, \mathbf{p})$ and parameterised by $\theta \in \mathbb{R}/\mathbb{Z}$. Notice that $\mathcal{H}(0) = H(\mathbf{q}, \mathbf{p})$. After trigonometric simplification we obtain

$$\mathcal{H}(\theta) = \frac{1}{4}\left(\|\mathbf{p}\|^2\left(1 + \cos(2\theta)\right) + \|\mathbf{q}\|^2\left(1 - \cos(2\theta)\right) + 2\mathbf{q} \cdot \mathbf{p} \sin(2\theta)\right).$$

From Proposition 2.6 it follows that $(\mathbf{q}, \mathbf{p}) \in \mathcal{M}^H$ if and only if it is a critical point of H in the direction of the fibre, i.e., $\mathcal{H}'(0) = 0$. By differentiating, one obtains

$$\mathcal{H}'(\theta) = \frac{1}{2}\left(\left(\|\mathbf{q}\|^2 - \|\mathbf{p}\|^2\right)\sin(2\theta) + 2\mathbf{q} \cdot \mathbf{p} \cos(2\theta)\right), \tag{23}$$

which leads to

$$\mathcal{M}^H = \big\{(\mathbf{q}, \mathbf{p}) \in \mathcal{M} : \mathbf{q} \cdot \mathbf{p} = 0\big\}. \tag{24}$$

This could also have been derived by differentiating the constraints in the differential algebraic equation (22), or by using one of the equivalent formulations of Remark 3.10.

The critical point $(\mathbf{q}, \mathbf{p}) \in \mathcal{M}^H$ is nondegenerate if and only if $\mathcal{H}''(\theta) \neq 0$, which, using (23), means that $\|\mathbf{q}\|^2 - \|\mathbf{p}\|^2 \neq 0$. From this we also see that $(\mathbf{q}, \mathbf{p})$ is degenerate critical point if and only if $\mathbf{q} \cdot \mathbf{p} = 0$ and $\|\mathbf{q}\| = \|\mathbf{p}\|$, which, again using (23), implies that the Hamiltonian is constant on the whole fibre. A closer examination shows that there are two such fibres. By the Hopf map these two fibres correspond to the two antipodal points $(\mathrm{i}, 0)$ and $(-\mathrm{i}, 0)$ on the 2-sphere in $\mathbb{C} \times \mathbb{R}$. We call those two points the *exceptional points* and the corresponding fibres the *exceptional fibres*. If we remove the points where $\|\mathbf{q}\| = \|\mathbf{p}\|$ from the constraint manifold, then $\mathcal{M}^H$ becomes a manifold with two connected components (one cannot get from a point where $\|\mathbf{q}\| < \|\mathbf{p}\|$ to a point where $\|\mathbf{q}\| > \|\mathbf{p}\|$ without crossing a point where $\|\mathbf{q}\| = \|\mathbf{p}\|$).

By differentiating the secondary constraints $\mathbf{q} \cdot \mathbf{p} = 0$ with respect to time, and substituting for $\dot{\mathbf{q}}$ and $\dot{\mathbf{p}}$ using (22) we obtain

$$\lambda = \frac{\|\mathbf{p}\|^2}{\|\mathbf{q}\|^2 - \|\mathbf{p}\|^2}. \tag{25}$$

Since both the Hamiltonian and the constraint function are conserved along a solution trajectory, it follows that $\|\mathbf{q}\|$ and $\|\mathbf{p}\|$ are constant on each trajectory, so the Lagrange multiplier λ is constant. Using these facts it is straightforward to show that the exact solution to (22) is given by

$$\begin{aligned} \mathbf{q}(t) &= R\left(\frac{t}{\alpha - 1/\alpha}\right)\mathbf{q}_0, \\ \mathbf{p}(t) &= R\left(\frac{t}{\alpha - 1/\alpha}\right)\mathbf{p}_0, \end{aligned} \tag{26}$$

where $(\mathbf{q}_0, \mathbf{p}_0) \in \mathcal{M}^H$ are the initial conditions, $\alpha = \pm\|\mathbf{p}_0\|/\|\mathbf{q}_0\|$, and

$$R(\theta) := \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

(The sign of α depends on which connected component of the constraint manifold the point $(\mathbf{q}_0, \mathbf{p}_0)$ belong to.) Thus, the exact solutions are given by rotations in the $\mathbf{q}$ and $\mathbf{p}$ planes. These trajectories are great circles of the constraint manifold $\mathcal{M} = S^3$, so both the fibres and the solution trajectories are given by great circles. Notice that the solution trajectory passing through $(\mathbf{q}_0, \mathbf{p}_0)$ crosses the fibre passing through $(\mathbf{q}_0, \mathbf{p}_0)$ twice (see Fig. 3).

We now turn to the solution of (22) by SHAKE and RATTLE. Indeed, SHAKE works as follows: for some $(\mathbf{q}_0, \mathbf{p}_0)$, move along the fibre so that the length of $(\mathbf{q}_1^-, \mathbf{p}_1^-)$

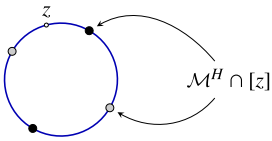


Fig. 3 An illustration of the intersection of a fibre $[z]$ with $\mathcal{M}^H$ for the system (22). Fibres are big circles on the sphere S^3. The two antipodal points of $\mathcal{M}^H$ on each fibre, marked with the same colour on the picture, are part of the same trajectory. Those two trajectories belong to two distinct connected components of $\mathcal{M}^H$

is 1 (to "land" again on $\mathcal{M}$). Now, $(\mathbf{q}_0^+, \mathbf{p}_0^+)$ is related to $(\mathbf{q}_1^-, \mathbf{p}_1^-)$ by an orthogonal reflection in the hyperplane perpendicular to $X_H(\mathbf{q}_0^+, \mathbf{p}_0^+) = (\mathbf{p}_0^+, 0)$. One can check that this reflection sends fibres to fibres. It is also straightforward to check that the reflection leaves solution trajectories invariant. Therefore, the fibre passing through $(\mathbf{q}_1^-, \mathbf{p}_1^-)$ intersects the same solution trajectory as the fibre passing through $(\mathbf{q}_0, \mathbf{p}_0)$, which means that SHAKE jumps between fibres of the same solution trajectory. Hence, RATTLE reproduces the exact flow of (22) up to a time reparametrisation.

Let us gather our findings:

Proposition 5.1 *Consider the symplectic manifold* $\mathcal{P} = \mathsf{T}^*\mathbb{R}^2 \simeq \mathbb{C}^2$, *the constraint submanifold*

$$\mathcal{M} := \left\{(z_0, z_1) \in \mathcal{P} : |z_0|^2 + |z_1|^2 = 1 \text{ and } \left(2z_0\bar{z_1}, |z_0|^2 - |z_1|^2\right) \neq (\pm \mathrm{i}, 0)\right\},$$

and the Hamiltonian $H(z_0, z_1) = (\Im(z_0)^2 + \Im(z_1)^2)/2$. *Then*:

1. *Both Assumption 1 and Assumption 2 hold.*
2. *The hidden constraint manifold* $\mathcal{M}^H$ *is defined by* (24). *It has two connected components.*
3. *The exact solution is given by* (26).
4. *The solution trajectory and fibre corresponding to some initial data in* $\mathcal{M}^H$ *intersect twice.*
5. SHAKE *maps between fibres that intersect the same exact solution trajectory.*
6. *Up to time reparametrisation,* RATTLE *produces the exact solution.*

We now turn to a slightly modified problem, for which RATTLE does not reproduce the exact trajectory.

5.4.2 Hamiltonian with Linear Potential Energy

With the same constraint, we consider now instead the Hamiltonian

$$H(\mathbf{q}, \mathbf{p}) = \frac{\|\mathbf{p}\|^2}{2} - \mathbf{g} \cdot \mathbf{q}, \tag{27}$$

where $\mathbf{g}$ is a fixed given vector.

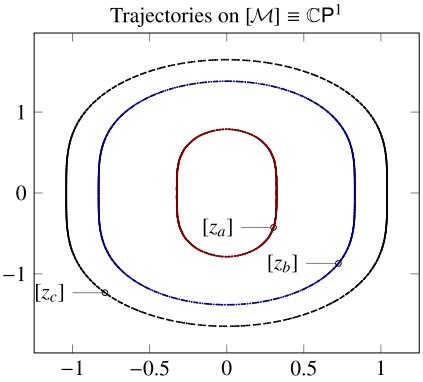


Fig. 4 Simulation results for the Hamiltonian (27) with the vector $\mathbf{g} := (0, -1)$. Both the SHAKE and RATTLE methods use the midpoint rule as underlying symplectic integrator. The time step is $h = 0.1$ in all the examples. We choose the initial conditions in Table 1. The phase trajectories for the three initial conditions z_a, z_b and z_c are displayed by first applying the Hopf map (21) and then using a stereographic projection defined as $\mathbb{C} \times \mathbb{R} \ni (\zeta, \rho) \mapsto \frac{\zeta}{\sqrt{|\zeta|^2+|\rho|^2}-\rho} \in \mathbb{C}$. The sphere $[\mathcal{M}]$ is thus identified to the Riemann sphere $\mathbb{C}\mathsf{P}^1$, and the trajectories are plotted on the complex plane. Note that we plot the projected trajectories on $[\mathcal{M}]$, and SHAKE and RATTLE become the same mapping when considered as a mapping on $[\mathcal{M}]$. The essential observation is that the simulated trajectories are closed, thus reproducing an important property of the exact solution. The innermost trajectory, corresponding to the initial value z_a, is also plotted in Fig. 6 along with the fibres lying above it. The energy and the hidden constraint are also plotted for that trajectory in Fig. 5

Now the structure of the hidden constraint set $\mathcal{M}^H$ is more involved. Let, as above, $\mathcal{H}$ be the restriction of H to the fibre passing through $(\mathbf{q}, \mathbf{p}) \in \mathcal{M}$. Then we get

$$\mathcal{H}'(\theta) = \frac{1}{2}\big(\|\mathbf{q}\|^2 - \|\mathbf{p}\|^2\big)\sin(2\theta) + \mathbf{q}\cdot\mathbf{p}\cos(2\theta) - \mathbf{g}\cdot\mathbf{q}\sin(\theta) + \mathbf{g}\cdot\mathbf{p}\cos(\theta). \quad (28)$$

The intersection between $[(\mathbf{q}, \mathbf{p})]$ and $\mathcal{M}^H$ is now more complicated than for the previous Hamiltonian. We have

$$\big[(\mathbf{q}, \mathbf{p})\big] \cap \mathcal{M}^H = \big\{\big(\mathbf{q}\cos(\theta) + \mathbf{p}\sin(\theta), -\sin(\theta)\mathbf{q} + \cos(\theta)\mathbf{p}\big) : \mathcal{H}'(\theta) = 0\big\}.$$

From the form (28) of $\mathcal{H}'(\theta)$ it follows that $[(\mathbf{q}, \mathbf{p})] \cap \mathcal{M}^H$ contains either:

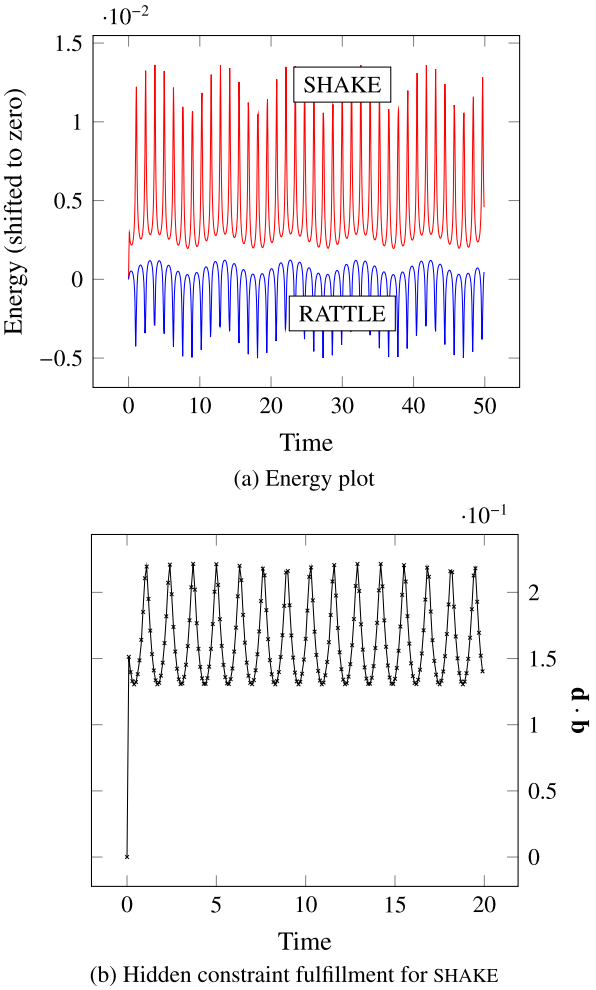
(i) four nondegenerate critical points, or
(ii) two nondegenerate critical points, or
(iii) two nondegenerate critical points and one degenerate critical point.

Which case that occurs depends on the magnitude of the $\cos(2\theta)$, $\sin(2\theta)$ coefficients in (28) in relation to the $\cos(\theta)$, $\sin(\theta)$ coefficients in (28): if the former are larger we get (i), if the latter are larger we get (ii), in the limit situation we get (iii). Thus, for some initial data it might (and will, see Fig. 6) happen that the critical points in $\mathcal{M}^H$ constituting the solution trajectory cease to be nondegenerate, in which case Assumption 1 is not fulfilled and the equation is no longer well posed.

In Fig. 4 we give some simulation results using SHAKE and RATTLE for globally well posed trajectories.

Table 1 Three initial conditions used in Fig. 4. All initial conditions lie on $\mathcal{M}^H$

	z_a	z_b	z_c
q_0	−0.7865261200000000	−0.4963624948824013	0.3477491188213400
q_1	−0.4043988000000000	−0.7319740436366664	−0.8131619010029159
p_0	−0.3880746864163783	−0.4275775933953260	−0.4368285559113795
p_1	0.2173391755798215	0.1225384882604160	−0.0837800227934176



(a) Energy plot

(b) Hidden constraint fulfillment for SHAKE

Fig. 5 Plot (**a**) shows the energy evolution for the SHAKE and RATTLE methods using the initial condition z_a in Table 1, corresponding to the inner most trajectory in Fig. 4. Notice that there is a slight difference between SHAKE and RATTLE, corresponding to the representation of the method either on $\widetilde{\mathcal{M}}^H$ or $\mathcal{M}^H$ as described above. Also notice that there is no drift in energy, in accordance with the general theory of symplectic integrators. Plot (**b**) shows the fulfillment of the hidden constraint by the SHAKE method using the same initial data z_a. The hidden constraint stays away from $\mathcal{M}^H$ at the same order of magnitude as the time step (0.1 in this case). A similar plot for RATTLE would show that the hidden constraint is fulfilled up to machine accuracy. Recall that this does not mean that RATTLE performs better than SHAKE, as RATTLE can be seen as a mere optional post-processing step. Note also that both methods fulfill the primary constraint (20) up to machine accuracy

In Fig. 5 we display the energy evolution of SHAKE and RATTLE and fulfillment of the hidden constraints for the SHAKE method.

In Fig. 6 we plot the magnitude of $\mathcal{H}'$ for the fibres along a numerical solution trajectory computed with SHAKE.

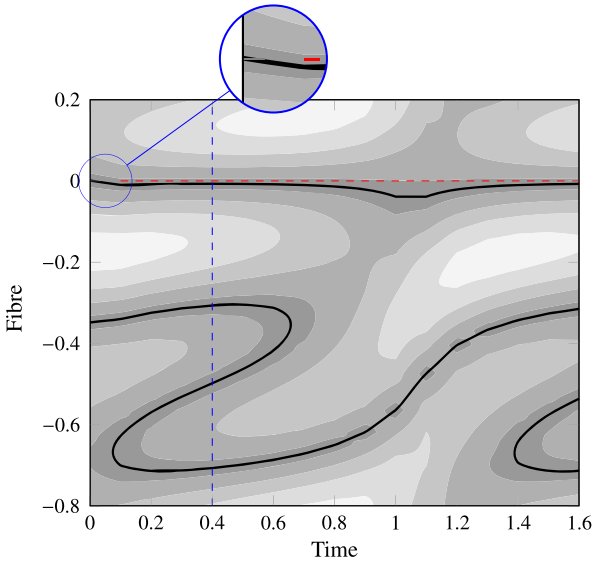


Fig. 6 Level curves showing the magnitude of $\mathcal{H}'$ (as given by (28)) for the fibres along the trajectory computed using SHAKE with initial data z_a given in Table 1. The points where $|\mathcal{H}'| = 0$ are those points lying on $\mathcal{M}^H$. *Black* corresponds to $|\mathcal{H}'| = 0$, so the *black curves* in the graphs give the intersection between $\mathcal{M}^H$ and the fibres above the trajectory. By construction, we are taking a coordinate in $[0, 1)$ on the circle and shifting it along the computed trajectory so that $\widetilde{\mathcal{M}}^H$ is always at coordinate 0 (the *dashed horizontal line* in the figure). The graph is periodic in the vertical direction because the fibres are circles; one such circle is given by the *dashed vertical line*. The graph is also periodic in the time direction, because the solutions are time periodic (see Fig. 4). Notice that, although the initial condition lies on $\mathcal{M}^H$, SHAKE persistently stays off $\mathcal{M}^H$ after one step (see the zoomed-in section). This is because $\widetilde{\mathcal{M}}^H$ is slightly offset from $\mathcal{M}^H$, in accordance with the theory in Sect. 4. The result of RATTLE is obtained by projecting vertically onto $\mathcal{M}^H$. Note that $\mathcal{M}^H$ crosses the fibres either two, three or four times, in accordance with the classification described in Sect. 5.4.2 (also compare to Fig. 3). In particular, there is another component of $\mathcal{M}^H$ crossing the fibre in the lower part of the figure. The inflexion points of that curve are points where Assumption 1 is not fulfilled anymore. The differential algebraic equation is not well defined at those points, and both SHAKE and RATTLE fail if such a point is reached. This failure takes two forms: the Newton iteration may either fail to converge, or it may find a spurious solution by jumping to another component of $\mathcal{M}^H$ which is better behaved

6 Discussion

The geometric version of the Dirac constraint algorithm [5] in the instance used here delivers a chain of submanifolds

$$\mathcal{M}^H \hookrightarrow \mathcal{M} \hookrightarrow \mathcal{P}$$

in which $\mathcal{M}^H$ is symplectic. The geometric RATTLE algorithm delivers symplectic integrators on $\mathcal{M}^H$. However, $\mathcal{P}$, the intervening presymplectic manifold $\mathcal{M}$, its coisotropy, and knowledge of its fibres, are essential to the algorithm. If the fibres cannot be explicitly parametrised, the algorithm is still formally defined, but more computation would be required in practice—for example, by integrating the fibres numerically to roundoff error. This is an extreme version of a situation common in

numerical analysis, in which allowing a wider class of methods (e.g. implicit Runge–Kutta methods, for which implicit equations have to be solved numerically) enables a wider class of properties.

If $\mathcal{M}$ is not coisotropic, then the coisotropic embedding theorem [4] says that there exists a symplectic manifold $\mathcal{P}'$ such that $\mathcal{M}$ is coisotropically embedded in $\mathcal{P}'$ (Remark 3.3). Thus, abstractly at least, one can extend the Hamiltonian on $\mathcal{M}$ arbitrarily to $\mathcal{P}'$ and apply the geometric RATTLE algorithm, for the rest of the required structure is intrinsic to $\mathcal{M}$. In specific examples it may be possible to carry this out by finding a suitable symplectic vector space $\mathcal{P}'$. The same remark holds if the given data is a Hamiltonian on a presymplectic manifold, see Sect. 5.3.

However, there remain many constrained problems which do not fall into the classes considered here. The most fundamental one has data $(\mathcal{P}, \mathcal{M}', H)$ where $\mathcal{M}'$ is a symplectic submanifold of the symplectic manifold $\mathcal{P}$. We do not know of symplectic integrators for this problem. They would provide symplectic integrators for a wide class of symplectic manifolds. A very general situation is that provided by the geometric version of the Dirac constraint algorithm [5], which, from presymplectic data $(\mathcal{M}, \omega, H)$, produces a nested sequence of submanifolds

$$\mathcal{M}_K \hookrightarrow \mathcal{M}_{K-1} \hookrightarrow \cdots \hookrightarrow \mathcal{M}_1 := \mathcal{M}$$

defined by

$$\mathcal{M}_{K+1} := \left\{ x \in \mathcal{M}_K : \mathrm{d}H(x) \in \omega(\mathcal{M}_K) \right\}$$

and a (possibly nonunique) vector field X such that $\iota_{\mathcal{M}_K}(\mathrm{i}_X \omega - \mathrm{d}H) = 0$. One would like to integrate an index-K DAE on $\mathcal{M}$ or an index-$K+1$ DAE on a symplectic embedding of $\mathcal{M}$ so as to preserve the constraints and $\iota_{\mathcal{M}_K}\omega$.

Finally, we mention another class of integrators for the holonomic case, known as SPARK, for Symplectic Partitioned Additive Runge–Kutta [9]. These generalise RATTLE to higher order. They are *partitioned* (use different RK coefficients for the q and p components) and *additive* (use different RK coefficients for the constraint and regular forces). The holonomic assumption is used in two critical steps: first, it means that the flow of the constraint force is given by Euler's method; second, it means that the q-component of the constraint forces vanishes. This allows their RK coefficients to be arbitrary, which means that the RK coefficients of the p-component can be arbitrary, and can be chosen to include stages at the endpoints. Thus, this approach does not immediately give an algorithm for problems of the type $(\mathcal{P}, \mathcal{M}', H)$ mentioned above. The situation is similar to the relationship between splitting methods and RK methods; we do not know if SPARK can be adapted to more general constraints.

Acknowledgements O. Verdier would like to acknowledge the support of the GeNuIn Project, funded by the Research Council of Norway, the Marie Curie International Research Staff Exchange Scheme Fellowship within the European Commission's Seventh Framework Programme as well as the hospitality of the Institute for Fundamental Sciences of Massey University, New Zealand, where some of this research was conducted. K. Modin would like to thank the Marsden Fund in New Zealand, the Department of Mathematics at NTNU in Trondheim, the Royal Swedish Academy of Science and the Swedish Research Council, contract VR-2012-335, for support. We would like to thank the reviewers for helpful suggestions.

References

1. H.C. Andersen, Rattle: a "velocity" version of the shake algorithm for molecular dynamics calculations, *J. Comput. Phys.* **52**, 24–34 (1983).
2. P. Dirac, *Lectures on Quantum Mechanics* (Dover Publications, New York, 2001).
3. Z. Ge, J.E. Marsden, Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators, *Phys. Lett. A* **133**, 134–139 (1988).
4. M.J. Gotay, On coisotropic imbeddings of presymplectic manifolds, *Proc. Am. Math. Soc.* **84**, 111–114 (1982).
5. M. Gotay, J. Nester, G. Hinds, Presymplectic manifolds and the Dirac–Bergmann theory of constraints, *J. Math. Phys.* **19**, 2388 (1978).
6. V. Guillemin, S. Sternberg, *Symplectic Techniques in Physics* (Cambridge University Press, Cambridge, 1990).
7. E. Hairer, S. Nørsett, G. Wanner, *Solving Ordinary Differential Equations: Nonstiff Problems* (Springer, Berlin, 1993).
8. E. Hairer, C. Lubich, G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations* (Springer, Berlin, 2006).
9. L. Jay, Symplectic partitioned Runge-Kutta methods for constrained Hamiltonian systems, *SIAM J. Numer. Anal.* **33**, 368–387 (1996).
10. B. Leimkuhler, S. Reich, *Simulating Hamiltonian Dynamics* (Cambridge University Press, Cambridge, 2004).
11. B. Leimkuhler, R. Skeel, Symplectic numerical integrators in constrained Hamiltonian systems, *J. Comput. Phys.* **112**, 117–125 (1994).
12. P. Libermann, C. Marle, *Symplectic Geometry and Analytical Mechanics* (D. Reidel Publishing Co., Dordrecht, 1987).
13. J. Marsden, T. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems* (Springer, New York, 1999).
14. R. McLachlan, K. Modin, O. Verdier, M. Wilkins, Symplectic integrators for index one constraints (2012). arXiv:1207.4250.
15. S. Reich, Symplectic integration of constrained Hamiltonian systems by composition methods, *SIAM J. Numer. Anal.* **33**, 475–491 (1996).
16. J.-P. Ryckaert, G. Ciccotti, H.J. Berendsen, Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.* **23**, 327–341 (1977).