Position Versus Momentum Projections for Constrained Hamiltonian Systems *

Werner M. Seiler

Lehrstuhl I für Mathematik,
Universität Mannheim, 68131 Mannheim, Germany
E-mail: seiler@euler.math.uni-mannheim.de

We study the effect of position and momentum projections in the numerical integration of constrained Hamiltonian systems. We show theoretically and numerically that momentum projections are better and more efficient. They lead to smaller error growth rates and affect the energy error much less, as they define a canonical transformation. As a concrete example, the planar pendulum is treated.

Keywords: Constrained dynamics, Hamiltonian system, numerical integration, orthogonal projection, canonical transformation, Dirac bracket, perturbed state space form

AMS Subject classification: Primary: 70H05, 65L05; Secondary: 70F20, 70H15

1. Constrained Hamiltonian Systems

Let the Hamiltonian $H(\mathbf{q},\mathbf{p})$ (with $\frac{\partial^2 H}{\partial \mathbf{p}\partial \mathbf{p}}$ nonsingular) be defined in a $2N$-dimensional phase space and impose K (irreducible) *position constraints* $\phi_\alpha(\mathbf{q}) = 0$. The arising dynamics are usually described by the following differential algebraic equation[1]

$$\dot{\mathbf{q}} = \nabla_{\mathbf{p}} H\,, \quad \dot{\mathbf{p}} = -\nabla_{\mathbf{q}} H - \mu^\alpha \nabla_{\mathbf{q}}\phi_\alpha\,, \quad \phi_\alpha(\mathbf{q}) = 0 \tag{1}$$

with multipliers $\boldsymbol{\mu}$. Differentiating $\boldsymbol{\phi}$ leads to the *momentum constraints*

$$\psi_\alpha(\mathbf{q},\mathbf{p}) = \dot{\phi}_\alpha(\mathbf{q}) = \nabla_{\mathbf{q}}\phi_\alpha \cdot \nabla_{\mathbf{p}} H = \{\phi_\alpha, H\} \tag{2}$$

with the canonical *Poisson bracket* $\{F,G\} = \nabla_{\mathbf{q}} F \cdot \nabla_{\mathbf{p}} G - \nabla_{\mathbf{p}} F \cdot \nabla_{\mathbf{q}} G$ [8, Chapt. VIII]. Differentiating $\boldsymbol{\psi}$ yields an algebraic system for $\boldsymbol{\mu}$

$$\{\psi_\alpha, \phi_\beta\}\mu^\beta = \{H, \psi_\alpha\}\,. \tag{3}$$

One more differentiation gives differential equations for $\boldsymbol{\mu}$; hence (1) has index 3.

Analytically, (1) can be treated as follows: solve (3) for $\boldsymbol{\mu}$, enter the result into (1), choose initial values satisfying *all* constraints and integrate the differential part of (1). Any such computed solution stays on the constraint manifold. The constraints are considered only when choosing the initial data. Numerically, this approach has two disadvantages. Firstly, the *underlying differential equation* obtained by entering the multipliers

* This work was supported by Deutsche Forschungsgemeinschaft.

[1] We use the Einstein convention that a summation over repeated indices is always implied, i. e. we write shortly $\mu^\alpha\phi_\alpha$ instead of $\sum_{\alpha=1}^{K} \mu^\alpha\phi_\alpha$.

into (1) is not Hamiltonian. This excludes for example the use of symplectic integrators. Secondly, in general the numerical solution drifts off the constraint manifold.

Projection methods are a popular cure against the drift [7]. In the simplest form, following a step with a conventional numerical scheme, one projects the computed point $(\tilde{\mathbf{q}}^n, \tilde{\mathbf{p}}^n)$ onto the constraint manifold to obtain the final approximation $(\mathbf{q}^n, \mathbf{p}^n)$. We may distinguish *position projections* where only $\tilde{\mathbf{q}}^n$ is modified so that $\boldsymbol{\phi}(\mathbf{q}^n) = 0$ and *momentum projections* where only $\tilde{\mathbf{p}}^n$ is changed so that $\boldsymbol{\psi}(\mathbf{q}^n, \mathbf{p}^n) = 0$ (with $\mathbf{q}^n = \tilde{\mathbf{q}}^n$).

As $(\mathbf{q}^n, \mathbf{p}^n)$ should be close to $(\tilde{\mathbf{q}}^n, \tilde{\mathbf{p}}^n)$, one prefers *orthogonal* projections. Thus a scalar product must be chosen. Systems where the kinetic energy is a quadratic form in $\mathbf{p}$ are called *natural*, i.e. for them $H(\mathbf{q}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^t M^{-1}(\mathbf{q})\mathbf{p} + V(\mathbf{q})$ with a symmetric and positive definite mass matrix $M(\mathbf{q})$. This form induces the scalar product

$$\langle \mathbf{p}_1 \,|\, \mathbf{p}_2 \rangle_M = \mathbf{p}_1^t M^{-1}(\mathbf{q})\mathbf{p}_2 \,. \tag{4}$$

Two vectors for which (4) vanishes are called *mass-orthogonal*.

Our main result is that momentum projections are not only cheaper than position projections, as they require only the solution of a *linear* system, but that they yield better results, too. The absolute values and the growth rates of all relevant errors (energy and constraint residuals) are smaller. Especially the energy error is much less affected by momentum projections. This will be demonstrated analytically for general (natural) systems and numerically for the specific example of the planar pendulum. Such considerations are of considerable interest for applications like molecular dynamics where computational efficiency is more important than accuracy.

Note that these results essentially apply also to *non*-Hamiltonian formulations. Usually, the equations of motion are derived as Euler-Lagrange equations. For the numerical integration they are rewritten as a first order system by introducing the velocities $\mathbf{v} = \dot{\mathbf{q}}$. But for a natural system $\mathbf{p} = M\mathbf{v}$ and this linear transformation should not drastically alter the behaviour of the errors. We mainly use the Hamiltonian formalism, as it offers a wider range of techniques like e.g. canonical transformations. For this reason we consider only conventional numerical methods and no symplectic ones.

The basic idea behind our analysis is the construction of two underlying Hamiltonian systems: one differing from (1) only by terms proportional to $\boldsymbol{\phi}$; the other one by terms proportional to $\boldsymbol{\psi}$. If we apply now the corresponding projection, it makes no difference[2] whether we integrate numerically (1) or the respective Hamiltonian system. In order to compare the two kind of projections we can thus study the stability of the constraint manifold for the two underlying systems.

One possibility for an underlying Hamiltonian system are the equations of motion for the *total Hamiltonian* $H_t = H + \mu^\alpha \phi_\alpha$ with $\boldsymbol{\mu}$ the solution of (3):

$$\dot{\mathbf{q}} = \nabla_{\mathbf{p}} H + (\nabla_{\mathbf{p}} \mu^\alpha)\phi_\alpha \,, \quad \dot{\mathbf{p}} = -\nabla_{\mathbf{q}} H - \nabla_{\mathbf{q}}(\mu^\alpha \phi_\alpha) \,. \tag{5}$$

They differ from (1) only by terms proportional to the position constraints $\boldsymbol{\phi}$.

Another possibility consists of using the *Dirac bracket* instead of the canonical Poisson bracket [15]. Denoting all constraints jointly by χ_a, it is defined by

$$\{F, G\}^* = \{F, G\} - \{F, \chi_a\}(C^{-1})^{ab}\{\chi_b, G\} \tag{6}$$

[2] This holds strictly only, if we project each time before we evaluate the equations of motion which is usually not true. But as our numerical example shows, we can neglect this small error.

where $C_{ab} = \{\chi_a, \chi_b\}$ is a skew-symmetric matrix. As equations of motion we take now the *Hamilton-Dirac equations*

$$\dot{\mathbf{q}} = \{\mathbf{q}, H\}^* , \quad \dot{\mathbf{p}} = \{\mathbf{p}, H\}^* . \tag{7}$$

Evaluating the brackets, it is straightforward to show that they differ from (1) only by terms proportional to the momentum constraints $\boldsymbol{\psi}$.

2. Hamiltonian Perturbation Analysis

A coordinate transformation $(\mathbf{q}, \mathbf{p}) \leftrightarrow (\mathbf{Q}, \mathbf{P})$ is called *canonical*, if it preserves the Poisson bracket. The transformation implicitly defined by

$$\mathbf{q} = \nabla_{\mathbf{p}} S(\mathbf{Q}, \mathbf{p}) , \quad \mathbf{P} = \nabla_{\mathbf{Q}} S(\mathbf{Q}, \mathbf{p}) \tag{8}$$

is canonical for any function S with regular Hessian. S is called the *generating function* [8, Chapt. VIII] and (assuming that $\mathbf{Q}$ and $\mathbf{p}$ may be considered as independent variables) every canonical transformation can be derived in this way [6, §§97–99].

Proposition 1. The equations $\phi_\alpha(\mathbf{q}) = \zeta_\alpha$, $\psi_\alpha(\mathbf{q}, \mathbf{p}) = \rho_\alpha$ define for fixed but arbitrary values $\boldsymbol{\zeta}, \boldsymbol{\rho}$ a $2(N-K)$-dimensional submanifold $\mathcal{M}_{\boldsymbol{\zeta},\boldsymbol{\rho}}$ of the full phase space. Let $f^i(\boldsymbol{\xi}, \boldsymbol{\zeta})$ be N functions such that $\phi_\alpha\big(\mathbf{f}(\boldsymbol{\xi}, \boldsymbol{\zeta})\big) = \zeta_\alpha$ and that the matrix $\begin{pmatrix} \nabla_{\boldsymbol{\xi}} \mathbf{f} \\ (\nabla_{\mathbf{q}} \boldsymbol{\phi})^t \end{pmatrix}$ is regular. Then the equations

$$q^i = f^i(\boldsymbol{\xi}, \boldsymbol{\zeta}) , \quad \frac{\partial f^i}{\partial \xi^a} p_i = \pi_a , \quad \frac{\partial \phi_\alpha}{\partial q^i} p_i = \rho_\alpha \tag{9}$$

implicitly define coordinates $(\boldsymbol{\xi}, \boldsymbol{\pi})$ on $\mathcal{M}_{\boldsymbol{\zeta},\boldsymbol{\rho}}$. The restriction of the canonical two-form of the full phase space to $\mathcal{M}_{\boldsymbol{\zeta},\boldsymbol{\rho}}$ is the canonical two-form in the coordinates $(\boldsymbol{\xi}, \boldsymbol{\pi})$.

Proof. This is a generalisation of a similar proposition in [11]. The existence of $\mathbf{f}$ follows from the implicit function theorem. Let the symplectic two-form be $\Omega = \omega_{ij} dz^i \wedge dz^j$ in some coordinate system z^i on the phase space. If a submanifold M is described parametrically by $z^i = z^i(y^a)$, Ω induces on M the two-form $\tilde{\Omega} = \tilde{\omega}_{ab} dy^a \wedge dy^b$ defined by $\tilde{\omega}_{ab} = \omega_{ij} \frac{\partial z^i}{\partial y^a} \frac{\partial z^j}{\partial y^b}$ [10]. In our case $\Omega = dq^i \wedge dp_i$ and thus $\tilde{\Omega} = \big[\frac{\partial q^i}{\partial \xi^a} \frac{\partial p_i}{\partial \pi_b} - \frac{\partial p_i}{\partial \xi^a} \frac{\partial q^i}{\partial \pi_b}\big] d\xi^a \wedge d\pi_b$. Entering (9) and noting that $\frac{\partial f^i}{\partial \xi^a} \frac{\partial p_i}{\partial \pi_b} = \delta_a^b$ we obtain $\tilde{\Omega} = d\xi^a \wedge d\pi_a$. □

On $\mathcal{M}_{\boldsymbol{\zeta},\boldsymbol{\rho}}$ we may consider (9) as implicitly defining a coordinate transformation $,_{\boldsymbol{\zeta},\boldsymbol{\rho}} : (\mathbf{q}, \mathbf{p}) \mapsto (\boldsymbol{\xi}, \boldsymbol{\pi})$. According to Proposition 1 $,_{\boldsymbol{\zeta},\boldsymbol{\rho}}$ is canonical and it is easy to see that its generating function is $S(\boldsymbol{\xi}, \mathbf{p}) = \mathbf{f}(\boldsymbol{\xi}, \boldsymbol{\zeta})\mathbf{p}$. A function $F(\mathbf{q}, \mathbf{p})$ on $\mathcal{M}_{\boldsymbol{\zeta},\boldsymbol{\rho}}$ can be transformed into a function $\tilde{F}(\boldsymbol{\xi}, \boldsymbol{\pi})$ satisfying $F = \tilde{F} \circ ,_{\boldsymbol{\zeta},\boldsymbol{\rho}}$. The transformed Hamiltonian $\tilde{H}$ for a natural system with $M = I$ is given by

$$\tilde{H}(\boldsymbol{\xi}, \boldsymbol{\pi}) = \frac{1}{2} \boldsymbol{\pi}^t \left[(\nabla_{\boldsymbol{\xi}} \mathbf{f})^t \nabla_{\boldsymbol{\xi}} \mathbf{f}\right]^{-1} \boldsymbol{\pi} + \frac{1}{2} \boldsymbol{\rho}^t \left[\nabla_{\mathbf{q}} \boldsymbol{\phi} (\nabla_{\mathbf{q}} \boldsymbol{\phi})^t\right]^{-1} \boldsymbol{\rho} + V \circ \mathbf{f}, \tag{10}$$

as $(\nabla_{\mathbf{q}} \boldsymbol{\phi})(\nabla_{\boldsymbol{\xi}} \mathbf{f}) = 0$ by definition of $\mathbf{f}$ and $\mathbf{p} = \begin{pmatrix} \nabla_{\boldsymbol{\xi}} \mathbf{f} \\ (\nabla_{\mathbf{q}} \boldsymbol{\phi})^t \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\pi} \\ \boldsymbol{\rho} \end{pmatrix}$ by (9).

We can now introduce *perturbed Hamiltonian state space forms* where $\boldsymbol{\zeta}, \boldsymbol{\rho}$ model the constraint residuals. For the Hamilton-Dirac equations (7) we get

$$\dot{\boldsymbol{\xi}} = \{\boldsymbol{\xi}, \tilde{H}\}\,, \quad \dot{\boldsymbol{\pi}} = \{\boldsymbol{\pi}, \tilde{H}\}\,. \tag{11}$$

For (5) we must use $\tilde{H}_t = \tilde{H} + \tilde{\mu}^\alpha \zeta_\alpha$ leading to a different perturbed state space form. But the *unperturbed* state space form obtained by setting $\boldsymbol{\zeta} = \boldsymbol{\rho} = 0$ is identical.

The position constraint residuals $\boldsymbol{\zeta}$ appear in $\tilde{H}$ only via the functions $\mathbf{f}$; in $\tilde{H}_t$ we get an extra term $\tilde{\boldsymbol{\mu}} \cdot \boldsymbol{\zeta}$. The momentum constraint residuals $\boldsymbol{\rho}$ appear in (10) also in form of an extra term, a quadratic form. Extra terms in the Hamiltonian lead to extra terms in the equations of motion which may change their qualitative properties profoundly. But if momentum projections are used, i. e. $\boldsymbol{\rho} = 0$, no extra terms appear.

We can refine the perturbation analysis by considering $\boldsymbol{\zeta}, \boldsymbol{\rho}$ as time-dependent. This does not change the canonical transformation $\Gamma_{\boldsymbol{\zeta},\boldsymbol{\rho}}$, but we must subtract from $\tilde{H}$ the time derivative of the generating function [8]

$$\frac{\partial S}{\partial t} = \dot{\boldsymbol{\zeta}}^t (\nabla_{\boldsymbol{\zeta}} \mathbf{f}) \mathbf{p}\,. \tag{12}$$

Applying $\Gamma_{\boldsymbol{\zeta},\boldsymbol{\rho}}$ with time-dependent residuals $\boldsymbol{\zeta}, \boldsymbol{\rho}$ to (5) or (7) yields differential equations for $\boldsymbol{\xi}, \boldsymbol{\pi}, \boldsymbol{\zeta}, \boldsymbol{\rho}$. Those for $\dot{\boldsymbol{\xi}}$ and $\dot{\boldsymbol{\pi}}$ are the corresponding state space form; those for $\dot{\boldsymbol{\zeta}}, \dot{\boldsymbol{\rho}}$ are in general not Hamiltonian. The origin $\boldsymbol{\zeta} = \boldsymbol{\rho} = 0$ is a fixed point for the latter equations and its stability determines the drift off the constraint manifold.

For the Hamilton-Dirac equations $\dot{\boldsymbol{\zeta}} = \dot{\boldsymbol{\rho}} = 0$ [15]. Based on this result we can analyse the use of momentum projections in the numerical integration of (1). The dynamics themselves do not lead to any growth of the constraint residuals. Following Alishenas [1] we use a continuous model for the error propagation and assume that because of numerical errors $\dot{\mathbf{q}} = \{\mathbf{q}, H\}^* + \boldsymbol{\epsilon}(t)$ with $||\boldsymbol{\epsilon}(t)|| < \hat{\epsilon}$ in the integration interval. Then $\dot{\boldsymbol{\zeta}} = \boldsymbol{\epsilon}(t)$ and the position constraint residual can grow *at most linearly.*

For (5) one cannot make such general statements. But the following simple argument for $K = 1$ shows that we must expect a worse behaviour. In the coordinates $(\mathbf{q}, \mathbf{p})$ the growth of the constraint residuals is determined by $\dot{\phi} = \{\phi, H_t\} = \{\phi, \mu\}\phi + \psi$ and $\dot{\psi} = \{\psi, H_t\} = \{\psi, \mu\}\phi$. Linearising at the origin yields $\dot{\phi} = 2a\phi + \psi$, $\dot{\psi} = b\phi$ with some time-dependent coefficients a, b. The eigenvalues of this system are $a \pm \sqrt{a^2 + b}$. Thus in general the origin is (linearly) unstable. Assuming that because of position projections $\phi(t) \approx \phi_0 \ll 1$, we still find from the equation for $\dot{\psi}$ that already the dynamics lead to an at least linear growth of the momentum constraint residual.

3. The Planar Pendulum

The planar pendulum represents a simple example of a constrained natural system. It is described by the Hamiltonian $H = \frac{1}{2}(p_x^2 + p_y^2) + y$. The position constraint is $\phi(x, y) = \frac{1}{2}(x^2 + y^2 - 1) = 0$ and the momentum constraint $\psi(x, y, p_x, p_y) = xp_x + yp_y = 0$. For the multiplier one obtains $\mu = \frac{p_x^2 + p_y^2 - y}{x^2 + y^2}$. The classical equations of motion (1) are

$$\dot{x} = p_x\,, \quad \dot{y} = p_y\,, \quad \dot{p}_x = -\mu x\,, \quad \dot{p}_y = -\mu y - 1\,. \tag{13}$$

The underlying Hamiltonian system (5) defined by $H_t = H + \mu\phi$ is

$$\begin{aligned} \dot{x} &= p_x + \frac{2p_x}{x^2+y^2}\,\phi\,, & \dot{y} &= p_y + \frac{2p_y}{x^2+y^2}\,\phi\,, \\ \dot{p}_x &= -\mu x + \frac{2x\mu}{x^2+y^2}\,\phi\,, & \dot{p}_y &= -\mu y - 1 + \frac{2y\mu+1}{x^2+y^2}\,\phi\,. \end{aligned} \tag{14}$$

The Dirac bracket is $\{F,G\}^* = \{F,G\} - \frac{1}{x^2+y^2}\big(\{F,\phi\}\{\psi,G\} - \{F,\psi\}\{\phi,G\}\big)$ and the Hamilton-Dirac equations (7) yield the underlying system

$$\begin{aligned} \dot{x} &= p_x - \frac{x}{x^2+y^2}\,\psi\,, & \dot{y} &= p_y - \frac{y}{x^2+y^2}\,\psi\,, \\ \dot{p}_x &= -\mu x + \frac{p_x}{x^2+y^2}\,\psi\,, & \dot{p}_y &= -\mu y - 1 + \frac{p_y}{x^2+y^2}\,\psi\,. \end{aligned} \tag{15}$$

Obviously, on the constraint manifold (13), (14) and (15) are identical.

The canonical transformation (9) can be written explicitly as

$$\,_{\zeta,\rho}\;:\;\begin{cases} x = \sqrt{2\zeta+1}\sin\xi\,, & y = \sqrt{2\zeta+1}\cos\xi\,, \\ p_x = \dfrac{\pi\cos\xi + \rho\sin\xi}{\sqrt{2\zeta+1}}\,, & p_y = \dfrac{-\pi\sin\xi + \rho\cos\xi}{\sqrt{2\zeta+1}}\,. \end{cases} \tag{16}$$

This yields for the transformed Hamiltonian

$$\tilde{H}(\xi,\pi) = \frac{1}{2}\frac{\pi^2+\rho^2}{2\zeta+1} + \sqrt{2\zeta+1}\cos\xi - \frac{\rho\dot{\zeta}}{2\zeta+1} \tag{17}$$

where the last term is the time derivative (12) of the generating function.

Notice that for this special system the extra terms in $\tilde{H}$ do not depend on the dynamical variables (ξ,π). Thus they affect only the energy error but do not lead to extra terms in the perturbed state space forms. These are for (14)

$$\dot{\xi} = \frac{(4\zeta+1)}{(2\zeta+1)^2}\pi\,, \quad \dot{\pi} = \frac{\zeta+1}{\sqrt{2\zeta+1}}\sin\xi \tag{18}$$

and for the Hamilton-Dirac equations (15)

$$\dot{\xi} = \frac{1}{2\zeta+1}\pi\,, \quad \dot{\pi} = \sqrt{2\zeta+1}\sin\xi\,. \tag{19}$$

Now we must analyse the growth of the constraint residuals ζ,ρ. Applying the transformation (16) to (14) yields the following system for them

$$\dot{\zeta} = \frac{4\zeta+1}{2\zeta+1}\rho\,, \quad \dot{\rho} = 4\frac{\zeta(\rho^2+\pi^2)}{(2\zeta+1)^2} - \frac{\zeta\cos\xi}{\sqrt{2\zeta+1}}\,. \tag{20}$$

The eigenvalues of the linearised system are $\pm\sigma$ with $\sigma = \sqrt{4\pi^2 - \cos\xi}$. Thus whenever σ is real, the origin is unstable for (20).

If we assume that due to position projections $\zeta \approx \zeta_0 \ll 1$, we obtain for ρ the Riccati equation $\dot{\rho} = \zeta_0(\sigma^2 + 4\rho^2)$. For the initial data $\rho(0) = \rho_0$ it has the solution $\frac{\sigma}{2}\tan\big[2\sigma\zeta_0 t + \arctan(\frac{2\rho_0}{\sigma})\big] \approx \rho_0 + \zeta_0(\sigma^2 + 4\rho_0^2)t + O(\zeta_0^2 t^2)$. Thus for small t already the dynamics yield an almost linear growth of the momentum constraint residual and it even blows up after a finite time $t_\infty \sim \zeta_0^{-1}$. Note that this still holds, if $\rho_0 = 0$.

4. Projections and Canonical Transformations

Proposition 2. A mass-orthogonal momentum projection is a canonical transformation for natural systems.

Proof. For natural systems the momentum constraint manifold is the hyperplane defined by $\langle\nabla_{\mathbf{q}}\phi_\alpha\,|\,\mathbf{p}\rangle_M = 0$. Thus $\nabla_{\mathbf{q}}\phi_\alpha$ are normal vectors and a mass-orthogonal projection has the form $\mathbf{P} = \mathbf{p} + (\nabla_{\mathbf{q}}\boldsymbol{\phi})^t\boldsymbol{\nu}$. The multipliers $\boldsymbol{\nu}$ are determined by the condition that for a given point $\mathbf{p}^*$ the transformed point $\mathbf{P}^*$ lies on the constraint manifold. This yields a linear system for $\boldsymbol{\nu}$: $\langle\nabla_{\mathbf{q}}\phi_\alpha\,|\,\mathbf{p}^* + \nu^\beta\nabla_{\mathbf{q}}\phi_\beta\rangle_M = 0$. Its coefficient matrix $R = (\nabla_{\mathbf{q}}\boldsymbol{\phi})^t M^{-1}(\nabla_{\mathbf{q}}\boldsymbol{\phi})$ is symmetric and positive definite; thus a unique solution exists.

Consider the family of generating functions $S_\lambda(\mathbf{Q},\mathbf{p}) = \mathbf{Q}\mathbf{p} + \lambda^\alpha\phi_\alpha(\mathbf{Q})$ with arbitrary constants $\boldsymbol{\lambda}$. It defines the canonical transformation $\mathbf{q} = \mathbf{Q}$, $\mathbf{P} = \mathbf{p} + (\nabla_{\mathbf{q}}\boldsymbol{\phi})^t\boldsymbol{\lambda}$. Choosing $\boldsymbol{\lambda} = \boldsymbol{\nu}$ we recover the mass-orthogonal momentum projection. □

Proposition 2 holds only for natural systems where a natural scalar product exists. But one can generalise it with a more operational point of view. The projection requires for general Hamiltonians the solution of a non-linear system. Similarly, we must solve a non-linear system in $\boldsymbol{\lambda}$ for the construction of the canonical transformation mapping $\mathbf{p}^*$ on the momentum constraint manifold. If both systems are treated in a certain way by Newton's method, the results are identical at each iteration step [14].

Proposition 3. A non-trivial position projection is never canonical.

Proof. One can show that all possible extensions of a point transformation $\mathbf{q} = \mathbf{F}(\mathbf{Q})$ to a canonical one can be described by generating functions of the form $S = \mathbf{F}(\mathbf{Q})\mathbf{p} + G(\mathbf{Q})$ with some scalar function G [6, §§105–108]. Hence the momenta transform according to $\mathbf{p} = (\nabla_{\mathbf{Q}}\mathbf{F})\mathbf{P} + \nabla_{\mathbf{Q}}G$. Since they remain unchanged in a position projection, $\nabla_{\mathbf{Q}}\mathbf{F}$ must be the identity matrix and G must vanish. □

For an integration method of order r, the (local) constraint residuals are of order $O(h^{r+1})$ [9]. Thus we may expect a projection to change the energy also in this order. However, for momentum projections the situation is much more favourable.

Proposition 4. If an integration method of order r is used, a mass-orthogonal momentum projection changes the energy in $O(h^{2r+2})$.

Proof. The projection may be considered as the flow of the Hamiltonian $H_\lambda = \lambda^\alpha\phi_\alpha(\mathbf{q})$ given by $\mathbf{q}(t) = \mathbf{q}_0$, $\mathbf{p}(t) = \mathbf{p}_0 - t[\nabla_{\mathbf{q}}\boldsymbol{\phi}(\mathbf{q}_0)]^t\boldsymbol{\lambda}$. The energy error is determined by the change of H along an integral curve of H_λ. At $t = 1$ in first order $\Delta E \approx \dot{H} = \{H, H_\lambda\} = -\boldsymbol{\lambda}\cdot\boldsymbol{\psi}$. In the proof of Proposition 2 we saw that $\boldsymbol{\lambda} = -R^{-1}\boldsymbol{\psi}$ and thus $\Delta E \approx \boldsymbol{\psi}^t R^{-1}\boldsymbol{\psi}$. As $\boldsymbol{\psi} = O(h^{r+1})$, the projection changes the energy in $O(h^{2r+2})$. □

5. An Analytical Example

We apply one step of the Euler method $z^{n+1} = z^n + hf(t, z^n)$ to the classical equations of motion (13) of the pendulum. Let the initial point be $A = (x, y, p_x, p_y)$, the result

$B = (\hat{x}, \hat{y}, \hat{p}_x, \hat{p}_y)$. Projecting B on the position constraint manifold leads to the point $B^{(p)} = (\hat{x}^{(p)}, \hat{y}^{(p)}, \hat{p}_x, \hat{p}_y)$; momentum projection to the point $B^{(m)} = (\hat{x}, \hat{y}, \hat{p}_x^{(m)}, \hat{p}_y^{(m)})$. We compare the constraint residuals and the energy error at B, $B^{(p)}$, and $B^{(m)}$.

If $\Phi = \phi(x, y)$ and $\Psi = \psi(x, y, p_x, p_y)$ denote the constraint residuals at A, we obtain at B (abbreviating unimportant but lengthy coefficients by dots)

$$\phi(\hat{x}, \hat{y}) = \Phi + \Psi h + \tfrac{1}{2}(p_x^2 + p_y^2)h^2\,, \tag{21a}$$

$$\psi(\hat{x}, \hat{y}, \hat{p}_x, \hat{p}_y) = \Psi - [(\ldots)\Psi + p_y]\,h^2\,. \tag{21b}$$

Obviously, the zeroth order terms are just the residuals at A and we obtain residuals of $O(h^2)$, if A satisfies both constraints, i.e. $\Phi = \Psi = 0$. Note that in the case of the momentum constraint residual $\Psi = 0$ suffices to obtain this order.

The difference ΔE of the energies at A and B is

$$\Delta E = (\ldots)\Psi h + (\ldots)h^2\,. \tag{21c}$$

Thus $\Delta E = O(h)$ for a general point A and $O(h^2)$, if A satisfies the momentum constraint. The position constraint residual Φ does not appear here.

The *position projection* requires the solution of a non-linear system. We assume that the step size h is so small that one Newton step suffices to get the position constraint residual below some prescribed tolerance. Then the projection has the simple form $\hat{x}^{(p)} = (1-\nu)\hat{x}$ and $\hat{y}^{(p)} = (1-\nu)\hat{y}$ where the multiplier is given by $\nu = \phi(\hat{x}, \hat{y})/(\hat{x}^2 + \hat{y}^2)$.

We expand the constraint residuals and the energy error in series in h.

$$\phi\big(\hat{x}^{(p)}, \hat{y}^{(p)}\big) = (\ldots)\Phi^2 + (\ldots)\Phi\Psi h + \big[(\ldots)\Phi + (\ldots)\Psi^2\big]h^2 + O(h^3)\,, \tag{22a}$$

$$\psi\big(\hat{x}^{(p)}, \hat{y}^{(p)}, \hat{p}_x, \hat{p}_y\big) = (\ldots)\Psi + (\ldots)\Psi^2 h + \big[(\ldots)\Psi - (\ldots)\big]h^2 + O(h^3)\,, \tag{22b}$$

$$\Delta E^{(p)} = (\ldots)\Phi + \big[(\ldots)\Psi + (\ldots)\Phi\big]h + O(h^2)\,. \tag{22c}$$

The energy error has now a zeroth order term for a general point A. If we assume that A satisfies both constraints, these expressions simplify considerably and we obtain

$$\phi\big(\hat{x}^{(p)}, \hat{y}^{(p)}\big) = \tfrac{1}{8}(p_x^2 + p_y^2)^2 h^4 + O(h^6)\,, \tag{23a}$$

$$\psi\big(\hat{x}^{(p)}, \hat{y}^{(p)}, \hat{p}_x, \hat{p}_y\big) = -p_y h^2 + \tfrac{1}{2}p_y(p_x^2 + p_y^2)h^4 + O(h^6)\,, \tag{23b}$$

$$\Delta E^{(p)} = \tfrac{1}{2}\Big[p_x(xp_y - yp_x) + x^2 + (p_x^2 + p_y^2)^2\Big]h^2 + O(h^3)\,. \tag{23c}$$

The vanishing of both, position and momentum, constraint residuals at A is necessary to achieve the improved orders. This can be seen from the last terms in (22a) and (22c).

The analysis of the *momentum projection* is simpler, as ψ is linear. One easily finds $\hat{p}_x^{(m)} = \hat{p}_x - \lambda\hat{x}$ and $\hat{p}_y^{(m)} = \hat{p}_y - \lambda\hat{y}$ with $\lambda = \psi(\hat{x}, \hat{y}, \hat{p}_x, \hat{p}_y)/(\hat{x}^2 + \hat{y}^2)$. The position constraint residual is again given by (21a). The momentum constraint residual always vanishes, as we perform an exact projection. The energy error is

$$\Delta E^{(m)} = (\ldots)\Psi^2 + (\ldots)\Psi h + \big[(\ldots)\Psi + (\ldots)\big]h^2 + O(h^3)\,. \tag{24}$$

If A satisfies the momentum constraint, then $\Delta E - \Delta E^{(m)} = O(h^4)$. This was to be expected by Proposition 4, as the Euler method is of order $r = 1$. In contrast, $\Delta E - \Delta E^{(p)} = \frac{1}{2}y(p_x^2 + p_y^2)h^2 + O(h^3)$ for A on the constraint manifold. Whenever $y < 0$ position projections *enlarge* the energy error in leading order in h.

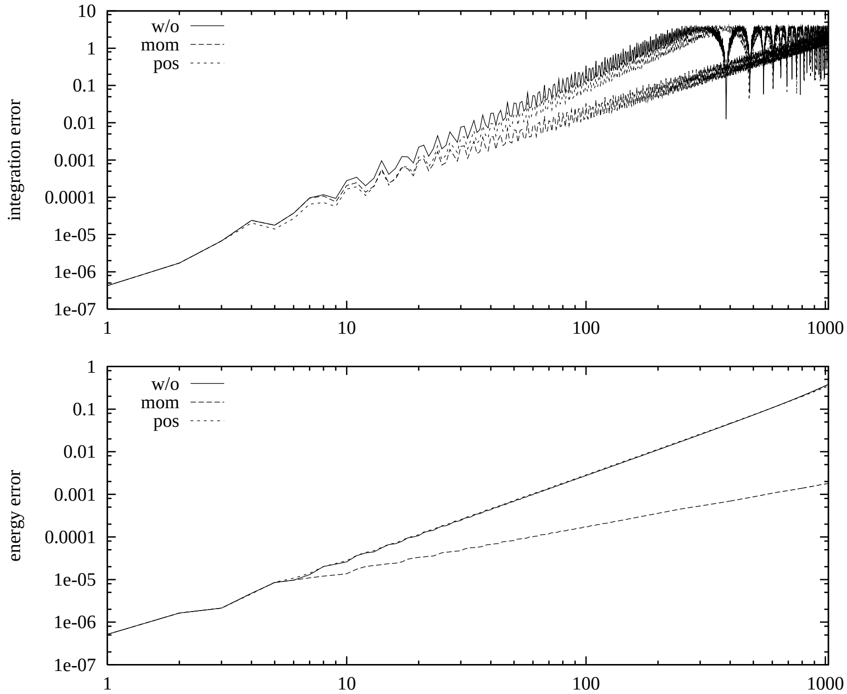

Figure 1. Integration and energy error for the planar pendulum

One can do the same calculations for the second order Runge-Kutta method $z^{n+1/2} = z^n + hf(t, z^n)/2$, $z^{n+1} = z^n + hf(t + h/2, z^{n+1/2})$. The results do not differ much. The residuals and the energy error are of higher order, but the qualitative picture remains the same. We find $\Delta E - \Delta E^{(m)} = O(h^6)$ in agreement with Proposition 4; but $\Delta E - \Delta E^{(p)} = O(h^3)$. Position projections enlarge the energy error whenever $yp_y > 0$, i.e. whenever the pendulum approaches an equilibrium.

6. A Numerical Example

We integrated the equations of motion (13) of the pendulum for the initial data $(x^0, y^0, p_x^0, p_y^0) = (1, 0, 0, -2)$ until $t = 1023$ with the classical fourth-order Runge-Kutta method using the step size $h = 0.025$. For these data the pendulum rotates clockwise with the period $T = 3.31$. We projected, when a residual exceeded $\epsilon = 10^{-6}$. Figure 1 (upper part) shows the integration error (estimated by comparing with an integration of the state-space form with step size $h/10$) without (`w/o`), with momentum (`mom`) and with position projections (`pos`).

Position projections hardly improve the results. They yield the same energy error as without, whereas momentum projections significantly reduce it (Figure 1, lower part). Position projections have no effect on the momentum constraint residual. In contrast, momentum projections improve the position constraint residual by more than two orders of magnitude compared to without projections. In the end it is $5 \cdot 10^{-4}$.

Momentum projections also yield smaller error *growth rates*. Without projections the integration error grows cubically, the energy error and the position constraint residual quadratically and the momentum constraint residual linearly. These rates are not changed by position projections. Momentum projections lead to a quadratic growth of the integration error and a linear growth of energy error and position constraint residual.

This can be partially explained by our results in Sections 2 and 3 where e. g. the linear growth of the remaining residual after projection on one constraint was predicted. The growth rates of the energy errors come from the perturbed Hamiltonian $\tilde{H}$. The momentum constraint residual ρ grows linearly after position projections. As it appears quadratically in (17), we can expect an at least quadratic growth of the energy error. For the error after momentum projections the dependency of $\tilde{H}$ on the position constraint residual ζ is decisive. The series expansion contains a linear term that dominates the higher order terms because of the smallness of ζ in our integration interval. As ζ grows linearly, so does the energy error.

In order to check the periodicity of the solutions we computed a *periodogram* from the values at $t = 0, 1, 2, \dots, 1023$. With momentum projections it hardly differs from the one obtained from the state space form and consists essentially of one spike at $f = 0.302$ with amplitude 0.39. Since $f = 1/T$, the periodicity is very well maintained. Without projections the spike is smeared over the range $0.3 - 0.37$ with a maximal amplitude of 0.05. Position projections yield only a small improvement.

The most striking result is that these considerable improvements have been achieved with only 155 momentum projections, i.e. on average after 260 integration steps. In contrast, position projections were needed after almost every step. With a tighter error tolerance the results for momentum projections further improve, whereas this makes hardly any difference for position projections. With $\epsilon = 10^{-8}$ one needs on average after 3 integration steps a momentum projection, the maximal value of the integration error is about $3 \cdot 10^{-3}$, of the energy error $3 \cdot 10^{-5}$ and of the position constraint residual 10^{-5}.

Other numerical methods yield similar results. Hairer and Wanner [9, p. 472] applied the Dormand-Prince 5(4) pair to the pendulum and observed that the integration error became even worse, when position projections were used. They also noted that adding position projections to momentum projections hardly changes the results.

7. Conclusions

Alishenas [1,2] showed already that in Lagrangian systems it is better to preserve the velocity constraints than the position constraints. We provided further evidence within the Hamiltonian formalism: momentum constraint residuals yield a quadratic extra term in the perturbed Hamiltonian and momentum projections are a canonical transformation and affect thus the energy error less than position projections.

The importance of the momentum constraints $\boldsymbol{\psi} = 0$ is easily understood geometrically. They represent a tangency condition for the position constraints $\boldsymbol{\phi} = 0$. Their preservation leads thus also to a stabilisation of the position constraints. But the preservation of $\boldsymbol{\phi} = 0$ does not influence the momentum constraints. This difference can be observed clearly in our numerical example.

We considered only simple projection methods, but our results can also be used in other ways. Substituting in the equations of motion (1) the position constraints $\boldsymbol{\phi} = 0$

by the momentum ones $\psi = 0$ reduces the index to 2 and the system can be rather efficiently treated by standard methods [16]. Or one incorporates the momentum projections into a numerical method; examples are the half-explicit Runge-Kutta methods of Brasey and Hairer [3,4]. Or one performs the projections within the equations of motion. Several mass-orthogonal formulations have been derived by Brauchli [5,17]; for Hamiltonian systems this leads to the impetus-striction formalism [14].

Propositions 2 and 3 imply that mass-orthogonal momentum projections do not destroy the symplectic nature of a numerical method in contrast to position projections. However, the naive symplectic integration of underlying Hamiltonian systems becomes rather expensive, as these are usually no longer separable and implicit methods must be used. One the other hand, the canonical transformation used in the proof of Proposition 2 (and thus momentum projection) is the basis of Reich's symplectic composition methods for constrained systems [13] which include the popular RATTLE scheme [12].

References

[1] T. Alishenas. *Zur Numerischen Behandlung, Stabilisierung durch Projektion und Modellierung Mechanischer Systeme mit Nebenbedingungen und Invarianten.* PhD thesis, Dept. of Numerical Analysis and Computer Science, Royal Institute of Technology, Stockholm, Sweden, 1992.

[2] T. Alishenas and Ö. Ólafsson. Modeling and velocity stabilization of constrained mechanical systems. *BIT*, 34:455–483, 1994.

[3] V. Brasey and E. Hairer. Half-explicit Runge-Kutta methods for differential-algebraic systems of index 2. *SIAM J. Num. Anal.*, 30:538–552, 1993.

[4] V. Brasey and E. Hairer. Symmetrized half-explicit methods for constrained mechanical systems. *Appl. Num. Math.*, 13:23–31, 1993.

[5] H. Brauchli. Mass-orthogonal formulation of equations of motion for multibody systems. *ZAMP*, 42:169–182, 1991.

[6] C. Carathéodory. *Variationsrechnung und Partielle Differentialgleichungen erster Ordnung, Band I.* Teubner, Leipzig, 1956.

[7] E. Eich, C. Führer, B. Leimkuhler, and S. Reich. Stabilization and projection methods for multibody dynamics. Research Report A281, Helsinki University of Technology, Institute of Mathematics, 1990.

[8] H. Goldstein. *Classical Mechanics.* Addison-Wesley, New York, 1980.

[9] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II.* Springer Series in Computational Mathematics 14. Springer-Verlag, Berlin, 1996.

[10] M. Henneaux and C. Teitelboim. *Quantization of Gauge Systems.* Princeton University Press, 1992.

[11] B.J. Leimkuhler and S. Reich. Symplectic integration of constrained Hamiltonian systems. *Math. Comp.*, 63:589–605, 1994.

[12] B.J. Leimkuhler and R.D. Skeel. Symplectic numerical integrators in constrained Hamiltonian systems. *J. Comp. Phys.*, 112:117–125, 1994.

[13] S. Reich. Symplectic integration of constrained Hamiltonian systems by composition methods. *SIAM J. Num. Anal.*, 33:475–491, 1996.

[14] W.M. Seiler. Impetus-striction formalism and the numerical integration of constrained Hamiltonian systems. In preparation, 1997.

[15] W.M. Seiler. Numerical integration of constrained Hamiltonian systems using Dirac brackets. *Math. Comp.*, to appear, 1998.

[16] B. Simeon. MBSPACK — numerical integration software for constrained mechanical motion. *Surv. Math. Ind.*, 5:169–202, 1995.

[17] M. Sofer, O. Melliger, and H. Brauchli. Numerical behaviour of different formulations for multibody dynamics. In Ch. Hirsch, O.C. Zienkiewicz, and E. Oñate, editors, *Numerical Methods in Engineering '92*, pages 277–284. Elsevier, Amsterdam, 1992.