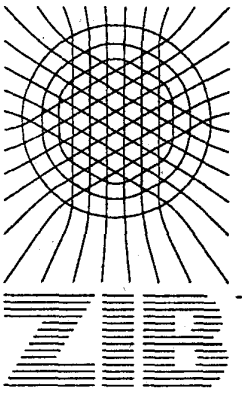
Konrad-Zuse-Zentrum für Informationstechnik Berlin



B. Leimkuhler S. Reich

# The Numerical Solution of Constrained Hamiltonian Systems

Preprint SC 92-16 (August 1992)

# Contents

B. Leimkuhler* S. Reich**

# The Numerical Solution of Constrained Hamiltonian Systems

**Abstract.** A Hamiltonian system subject to smooth constraints can typically be viewed as a Hamiltonian system on a manifold. Numerical computations, however, must be performed in $\mathbf{R}^n$. In this paper, canonical transformations from "Hamiltonian differential–algebraic equations" to ODEs in Euclidean space are considered.

In §2, canonical parameterizations or local charts are developed and it is shown how these can be computed in a practical framework. In §3 we consider the construction of unconstrained Hamiltonian ODE systems in the space in which the constraint manifold is embedded which preserve the constraint manifold as an integral invariant and whose flow reduces to the flow of the constrained system along the manifold. It is shown that certain of these unconstrained Hamiltonian systems force Lyapunov stability of the constraint–invariants, while others lead to an unstable invariant. In §4, we compare various projection techniques which might be incorporated to better insure preservation of the constraint–invariants in the context of numerical discretization. Numerical experiments illustrate the degree to which the constraint and symplectic invariants are maintained under discretization of various formulations.

**Keywords:** differential–algebraic equations, Hamiltonian systems, canonical discretization schemes

**AMS(MOS):** subject classification 65L05

---

* Dept. of Mathematics, University of Kansas, Lawrence, KS 66045, USA
Guest of the Konrad–Zuse–Zentrum Berlin (ZIB), June/July 1992

** Institute of Applied Mathematics and Statistics, O–1086 Berlin,
Visiting Simon Fraser University, Burnaby, B.C., CANADA

# 1. Introduction

Consider a simplified Hamiltonian system of the form:

$$\dot{q} = p \tag{1.1}$$
$$\dot{p} = -\nabla F(q) \tag{1.2}$$

where $q, p \in \mathbf{R}^n$ and $F : \mathbf{R}^n \to \mathbf{R}$ is $C^2$. In places we consider more general Hamiltonians, but a gradient system illuminates the basic questions. The system (1.1)–(1.2) can be accurately solved over long time intervals by a canonical discretization scheme [20] which maintains the symplectic structure of the flow.

A natural question is what happens when (1.1)–(1.2) is constrained by algebraic equations on $q$ and/or $p$. In this paper, we primarily restrict ourselves to the case when the constraints are dependent on $q$ only (as in many mechanical systems), in which case, starting from a Lagrangian variational principle, one arrives at a system of differential–algebraic equations of the form:

$$\dot{q} = p \tag{1.3}$$
$$\dot{p} = -\nabla F(q) - G(q)^t\lambda \tag{1.4}$$
$$0 = g(q) \tag{1.5}$$

where $g : \mathbf{R}^n \to \mathbf{R}^m$ and $G(q) = g'(q) \in \mathbf{R}^{m\times n}$. This system generates a flow on the $(2n - 2m)$–dimensional manifold $\mathcal{M} = \{(q,p) : g(q) = 0, Gp = 0\}$.[1] Although numerous discretization schemes exist for solving these equations directly (see, e.g. [5], [12]), none of the existing methods can be expected to maintain the symplectic structure of (1.3)–(1.5) with any degree of accuracy.

Associated with (1.3)–(1.5) are two important families of ODEs in Euclidean space: the *underlying* and *state–space* form ODEs. An example of an underlying ODE is obtained by first differentiating the constraint $g(q) = 0$ and using (1.3):

$$G\dot{q} = 0 = Gp$$

Then differentiating again yields

$$G\dot{p} + G_q(p,p) = 0$$

(We use the notation $G_q(p,w)$ to denote the derivative of $G$–the tensor second derivative of $g$– operating on vectors $p$ and $w$.) Next we substitute (1.4) and solve the resulting equations for $\lambda$ in terms of $q$ and $p$:

$$\lambda = \Lambda(q,p) = (GG^t)^{-1}(-G\nabla F(q) + G_q(p,p))$$

[1] For notational convenience we frequently supress arguments when these are readily apparent, thus $G$ means $G(q)$, etc.

which, upon reintroduction in (1.4) gives

$$\dot{p} = -(I - \mathcal{H})\nabla F(q) - G^t(GG^t)^{-1}G_q(p,p) \tag{1.6}$$

where $\mathcal{H} = G^t(GG^t)^{-1}G$ is a projector onto the complement of the null space of $G$. We term the ODE comprising $\dot{q} = p$ together with (1.6) the *standard underlying ODE*; it has the feature that the flow it generates reduces to the flow of (1.3)–(1.5) along the constraint manifold $\mathcal{M}$. On the other hand, without enforcing the constraint, (1.3), (1.6) actually defines a flow in $\mathbf{R}^{2n}$. Numerical methods applied directly to this underlying ODE typically drift from the constraint manifold into $\mathbf{R}^{2n}$ during the course of integration, but a popular approach to short time interval computations incorporates numerical discretization of (1.6) and frequent projection onto the constraint [10], [3].

While (1.6) defines one underlying ODE there is a entire family of ODEs whose dynamics reduce to those of the constrained system along $\mathcal{M}$. While (1.6) is not a Hamiltonian system away from $\mathcal{M}$, Hamiltonian ODE systems can be found in the family of underlying ODEs; such systems are developed in §3 using the Poisson bracket formalism of Dirac [9] for constrained Hamiltonian systems.

The second family of ODEs associated with the DAE (1.3)–(1.5) is constructed via a parameterization of the constraint (1.5). Supposing that there is a function $\phi : \mathbf{R}^{n-m} \to \mathbf{R}^n$ with a full rank Jacobian satisfying, for all $\delta \in \mathbf{R}^{n-m}$,

$$g(\phi(\delta)) \doteq 0$$

then with $\delta, \theta \in \mathbf{R}^{n-m}$ the equations

$$\begin{aligned} q &= \phi(\delta) \\ p &= \phi'(\delta)\theta \end{aligned}$$

define an invertible map from $\mathcal{M}$ to $\mathbf{R}^{2n-2m}$. This results in equations in the new variables of the form

$$\begin{aligned} \phi'(\delta)\dot{\delta} &= \phi'(\delta)\theta \\ \phi'(\delta)\dot{\theta} + \frac{\partial \phi'(\delta)\theta}{\partial \delta}\dot{\delta} &= -\nabla F(\phi(\delta)) + G^t\lambda \end{aligned}$$

Now multiplying both equations on the left by $(\phi'^t\phi')^{-1}\phi'^t$ results in

$$\begin{aligned} \dot{\delta} &= \theta \\ \dot{\theta} &= -(\phi'^t\phi')^{-1}\phi'^t(\nabla F(\phi(\delta)) - \phi''(\theta,\theta)) \end{aligned}$$

A *state space form* constructed along these lines will rarely be Hamiltonian. On the other hand, by searching among all parameterization of $\mathcal{M}$ (which do

not necessarily maintain the relation $\dot{\delta} = \theta$), one can find a family of *canonical state space forms* for the constrained problem. This is the approach taken in §2, and a relatively efficient algorithm for solving the discretized equations is presented.

## 2. Hamiltonian State Space Forms

The following theorem shows that there is a family of canonical state space forms based on parameterizations of the constraints. Througout this section we are concerned with a Hamiltonian of the form $H = F(q) + p^t p/2$.

**Proposition 2.1** *If $\phi$ is a parameterization of $g(q) = 0$, then the mapping*

$$\begin{aligned} q &= \phi(\delta) \\ \phi'^t p &= \theta \end{aligned}$$

*defines a canonical map between $\mathcal{M}$ and $\mathbf{R}^{2n-2m}$. The Hamiltonian in the new coordinates is*

$$\hat{H}(\delta, \theta) = \frac{1}{2}\theta^t(\phi'^t\phi')^{-1}\theta + (F \circ \phi)(\delta)$$

*Proof:* This follows from available results in Abraham and Marsden [1] on canonical maps between manifolds. □

In general such a state–space form is computationally impractical because of the need to automatically obtain and then twice differentiate the function $\phi$ defining the parameterization. Here we show how a careful choice of parameterization can lead to a more computable formulation.

Following [18] we define $q = \phi(\delta)$ as the solution of the nonlinear system

$$Aq = \delta \tag{2.1}$$

$$g(q) = 0 \tag{2.2}$$

where $A$ has been chosen so that

$$R = \begin{bmatrix} A \\ G \end{bmatrix}$$

is a nonsingular matrix. Previous authors have used the induced state space form obtained by setting $\delta' = \theta$ to solve multibody dynamics problems, but instead, we here choose $\theta$ to insure a canonical map,

$$p = (I - \mathcal{H})A^t\theta \tag{2.3}$$

with $\mathcal{H} = G^t(GG^t)^{-1}G$. Thus

$$\delta' = Aq' = Ap = A(I - \mathcal{H})A^t\theta \tag{2.4}$$

and, finally, since $A(I - \mathcal{H})A^t$ is nonsingular,

$$-A(I - \mathcal{H})\nabla F(q) = A(I - \mathcal{H})A^t\theta' - A(I - \mathcal{H})\mathcal{H}_q(p, A^t\theta) \tag{2.5}$$

gives an equation solvable for $\theta'$.

The above choice of canonical SSF requires determination of the derivative $\mathcal{H}_q$ of $\mathcal{H}$. Because $\mathcal{H}$ is expressed in terms of $G$, $\mathcal{H}_q$ requires knowledge of $G_q$. Although this could be somewhat difficult to compute in general cases, for certain practical applications, the second derivatives of the constraint functions can be obtained systematically at the same time the constraints are being generated. It is also possible to employ the new automatic differentiation techniques of [11] to compute the derivatives "on the fly."

A more challenging problem is probably determination of the matrix $A$. This matrix is typically held constant over several steps and must be updated from time to time as integration progresses. One possibility is to compute the matrix $A$ at the beginning of each step using a technique such as singular value decomposition, but this may be relatively expensive, especially for large systems.

In [14], a particular case of the canonical parameterization was defined via a *coordinate partitioning* (in the sense of [19]); here $A$ would be a matrix of 1's and 0's chosen to "pick out" a minimal subset of the coordinates which can be used to define a local chart.

## 2.1 Computation of the Canonical SSF

It must be pointed out that although we began this section treating a problem with a separable Hamiltonian (i.e. $H(q,p) = T(p) + V(q)$), the Hamiltonian of the canonical state space form ODE is not separable. What this means that e.g. the efficient explicit Runge–Kutta Nyström methods of Okunbor and Skeel [16] will not be canonical on the Hamiltonian state space form. On the other hand, we here show that the mixed set of equations (2.1)–(2.5) in $q$, $p$, $\delta$ and $\theta$ can be solved relatively efficiently by using implicit (e.g. Gauss–Legendre) Runge–Kutta discretization methods, by a scheme based on functional iteration. (2.1)–(2.5) can be viewed as defining differential equations $\delta' = f(\delta,\theta)$, $\theta' = g(\delta,\theta)$ for $\delta$ and $\theta$. Consider the $s$–stage Runge–Kutta method which computes an approximation $(\delta_{n+1},\theta_{n+1}) \approx (\delta(t_0+(n+1)h),\theta(t_0+(n+1)h))$ in terms of $(\delta_n,\theta_n)$ via

$$\Delta_i = \delta_n + h\sum_{j=1}^{s} a_{i,j}\dot{\Delta}_j \tag{2.6}$$

$$\Theta_i = \theta_n + h\sum_{j=1}^{s} a_{i,j}\dot{\Theta}_j \tag{2.7}$$

$$\delta_{n+1} = \delta_n + h\sum_{j=1}^{s} b_j\dot{\Delta}_j \tag{2.8}$$

$$\theta_{n+1} = \theta_n + h\sum_{j=1}^{s} b_j\dot{\Theta}_j \tag{2.9}$$

where $\dot{\Delta}_j = f(\Delta_j, \Theta_j)$, $\dot{\Theta}_j = f(\Delta_j, \Theta_j)$. For properly chosen $a_{ij}$ and $b_j$, this method generates a canonical map [20]. The computation of (2.6)–(2.9) is a little unusual for (2.1)–(2.5) because the state–space form is given implicitly. One approach is to lump all of the equations at each step together and solve them via a Newton iteration, but this is unnatural for nonstiff problems like the ones we would expect to encounter in a Hamiltonian framework [8]. A much better approach to computing a step is summarized in Algorithm 2.1.

**Algorithm 2.1**

*Given $\delta_n$, $\theta_n$, tolerance $\epsilon$*

*1. Predict values for the stage variables $\Delta_i^{(0)}$, $\Theta_i^{(0)}$, $i = 1, \ldots, s$.*

*2.* **repeat** *for $k = 0, 1, \ldots$.*

*2a. Solve the systems*

$$\begin{aligned} \Delta_i^{(k)} &= AQ_i^{(k)} \\ 0 &= g(Q_i^{(k)}) \end{aligned}$$

*for $Q_i^{(k)}$, $i = 1, \ldots, s$.*

*2b. Compute:*

$$\begin{aligned} \mathcal{H}_i^{(k)} &= \mathcal{H}(Q_i^{(k)}) \\ P_i^{(k)} &= (I - \mathcal{H}_i^{(k)})A^t\Theta_i^{(k)} \\ \mathcal{H}_{q,i}^{(k)} &= \mathcal{H}_q(P_i^{(k)}, A^t\Theta_i^{(k)}) \\ f_i^{(k)} &= -\nabla F(Q_i^{(k)}) \end{aligned}$$

*for $i = 1, \ldots, s$.*

*2c. Solve the equations*

$$\begin{aligned} A(I - \mathcal{H}_i^{(k)})A^t\dot{\Theta}_i^{(k)} &= A(I - \mathcal{H}_i^{(k)})f_i^{(k)} + A(I - \mathcal{H}_i^{(k)})\mathcal{H}_{q,i}^{(k)} \\ \dot{\Delta}_i^{(k)} &= A(I - \mathcal{H}_i^{(k)})A^t\Theta_i^{(k)} \ . \end{aligned}$$

*for $\dot{\Delta}_i^{(k)}$ and $\dot{\Theta}_i^{(k)}$, $i = 1, \ldots, s$.*

*2d. Compute $\Delta_i^{(k+1)} = \delta_n + h\sum_{j=1}^{s} a_{ij}\dot{\Delta}_j^{(k)}$ and $\Theta_i^{(k+1)} = \theta_n + h\sum_{j=1}^{s} a_{ij}\dot{\Theta}_j^{(k)}$.*

*2e.* **until** $\|\Delta_i^{(k+1)} - \Delta_i^{(k)}\| < \epsilon$ *and* $\|\Theta_i^{(k+1)} - \Theta_i^{(k)}\| < \epsilon$.

*3. Compute $\delta_{n+1} = \delta_n + h\sum_{j=1}^{s} b_j\dot{\Delta}_j^{(k+1)}$ and $\theta_{n+1} = \theta_n + h\sum_{j=1}^{s} b_j\dot{\Theta}_j^{(k+1)}$.*

Algorithm 2.1 is really nothing more than functional iteration for the equations in the reduced variables, so it always converges for sufficiently small $h$. Also, observe that the iterations in steps 2a, 2b, 2c, and 2d could be trivially parallelized by using "parallel do loops" across the stage index i. The functional iteration must be iterated to convergence (i.e to within rounding error) to insure that it defines a symplectic map.

# 3. Hamiltonian Underlying ODEs

We now examine the possibility of obtaining Hamiltonian underlying ODEs as an alternative to the computation of the state space form. In case the constraint is linear,

$$Gq = 0$$

with $G$ constant, the standard underlying ODE (1.1)–(1.2) reduces to

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -(I - \mathcal{H})\nabla F(q) \end{aligned}$$

This ODE system is not Hamiltonian because the projection of $\nabla F$ is not necessarily the gradient of any function, however it is easy to construct an underlying ODE which is Hamiltonian: we simply note that if $q$ lies on $Gq = 0$, then $(I - \mathcal{H})q = q$ so that

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -(I - \mathcal{H})\nabla F((I - \mathcal{H})q) \end{aligned}$$

is also an underlying ODE–and this one *is* a Hamiltonian system.

For the nonlinearly constrained case, we make use of Dirac's theory of constrained Hamiltonian systems [9].

## 3.1 Nonlinearly Constrained Hamiltonians

For a Hamiltonian function $H = H(q,p)$ and a scalar–valued function $\phi = \phi(q,p)$, the condition for $\phi$ to be an invariant under the flow of the Hamiltonian system derived from $H$ is just that the *Poisson bracket* of $\phi$ with $H$ vanishes, i.e.

$$\sum_{i=1}^{n} \frac{\partial \phi}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial \phi}{\partial p_i}\frac{\partial H}{\partial q_i} =: \{\phi, H\} = 0$$

Here a distinction must be made between two types of invariants. $\phi(q,p) = 0$ is said to be a *strong* invariant of the flow derived from $H$ in case $\{\phi, H\}$ vanishes identically. A weak invariant is one that satisfies $\{\phi, H\} = 0$ *only when* $\phi(q,p) = 0$. We make use of the following elementary properties of Poisson brackets for functions $\phi, \psi, \omega : \mathbf{R}^{2n} \to \mathbf{R}$ and real constants $\alpha_1, \alpha_2$:

(i) $\{\phi, \psi\} = -\{\psi, \phi\}$

(ii) $\{\phi, \phi\} = 0$

(iii) $\{\alpha_1\phi_1 + \alpha_2\phi_2, \psi\} = \alpha_1\{\phi_1, \psi\} + \alpha_2\{\phi_2, \psi\}$;
$\{\psi, \alpha_1\phi_1 + \alpha_2\phi_2\} = \alpha_1\{\psi, \phi_1\} + \alpha_2\{\psi, \phi_2\}$

(iv) $\{\phi, \psi\omega\} = \{\phi, \psi\}\omega + \{\phi, \omega\}\psi$

For arbitrary $H$ and $\phi$ consider the adjusted (constrained) Hamiltonian function

$$H^{(1)} = H + \mu\phi$$

Reasoning from physical principles, $\mu$ arises as a generalized Lagrange multiplier in attempting to minimize a certain functional subject to the constraint $\phi = 0$ [9].

$\mu$ must be chosen to insure $\phi = 0$ along solutions, which certainly hold if $\phi$ is a weak invariant of the flow of $H^{(1)}$. For this to happen, we need that

$$\begin{aligned} 0 \approx \{\phi, H^{(1)}\} &= \{\phi, H\} + \{\phi, \mu\phi\} \\ &= \{\phi, H\} + \{\phi, \phi\}\mu + \{\phi, \mu\}\phi \end{aligned}$$

Taking $\phi = 0$ in the above and noting that $\{\phi, \phi\} = 0$, we must have $\{\phi, H\} \approx 0$. When this does not happen, we treat the equation $\psi \equiv \{\phi, H\} = 0$ as a new constraint and consider the revised Hamiltonian

$$H^{(2)} = H + \mu_1\phi + \mu_2\psi$$

If we now attempt to insure that both $\phi \approx 0$ and $\psi \approx 0$ for the flow of $H^{(2)}$, we find that the key issue concerns the invertibility of the matrix of Poisson brackets

$$R = \begin{bmatrix} \{\phi, \phi\} & \{\phi, \psi\} \\ \{\psi, \phi\} & \{\psi, \psi\} \end{bmatrix}$$

When $R$ is nonsingular, we can solve for $(\mu_1, \mu_2)$ so that both $\phi$ and $\psi$ are invariant for $H^{(2)}$.

We now turn to the case of a vector–valued constraint function. The main thing to bear in mind here is that, in the end, the constraints must be treated all at once, not one at a time. Given a vector of constraints $\phi = 0$, one must first augment these constraints by all of the "hidden" constraints which arise by taking Poisson brackets with the augmented Hamiltonians, i.e. through the recursive differentiation of the constraints and substitution of the differential equations derived from the Hamiltonian. This approach is taken in [14] in deriving control laws for constrained systems, where it is shown that two steps of the reduction process are sufficient if the constraints are independent and holonomic, i.e. essentially only dependent on $q$.

As an example, if we follow the reduction for independent constraints of the form $g(q) = 0$, we obtain the hidden constraints $G(q)p = 0$. The next step is construction of the constrained Hamiltonian $H_T$ from $H$ and the constraints, thus we set

$$H_T(q,p) := H(q,p) + \mu(q,p)^t g(q) + \eta(q,p)^t G(q)p$$

Equations for $\mu$ and $\eta$ can be derived directly by insuring that $g(q) = 0$ and $G(q)p = 0$ are either weak or strong invariants of the flow derived from $H_T$. A slight generalization of the Poisson bracket notation to handle multiple constraints makes this straightforward.

**Definition 3.1** *Given vector valued functions $\phi : \mathbf{R}^{2n} \to \mathbf{R}^l$ and $\psi : \mathbf{R}^{2n} \to \mathbf{R}^m$, the Poisson bracket of $\phi$ and $\psi$ is the $l \times m$ matrix whose $(i,j)$–component is defined by*

$$(\{\phi, \psi\})_{i,j} = \{\phi_i, \psi_j\}$$

The following proposition shows how the generalized Poisson bracket can be evaluated in terms of the Jacobians of the vector functions.

**Proposition 3.1** *Given vector valued functions $\phi : \mathbf{R}^{2n} \to \mathbf{R}^l$ and $\psi : \mathbf{R}^{2n} \to \mathbf{R}^m$, let $\phi_q, \phi_p \in \mathbf{R}^{l \times n}$, $\psi_q, \psi_p \in \mathbf{R}^{m \times n}$, and denote the Jacobian matrices of the indicated function with respect to the indicated variables. Then*

$$\{\phi, \psi\} = \phi_q \psi_p^t - \phi_p \psi_q^t$$

Using Proposition 3.1, it is easy to see that $\{\phi, \psi\} = -\{\psi, \phi\}^t$. Proposition 3.2 is also useful in calculations:

**Proposition 3.2** *If $\phi$ and $\psi$ are as in Proposition 3.1, and $\lambda : \mathbf{R}^{2n} \to \mathbf{R}^m$, then*

$$\{\phi, \lambda^t \psi\} = \{\phi, \psi\}\lambda + \{\phi, \lambda\}\psi$$

*Proof:*

$$\begin{aligned} \{\phi, \lambda^t\psi\} &= \phi_q(\lambda^t\psi)_p^t - \phi_p(\lambda^t\psi)_q^t \\ &= \phi_q(\psi_p^t\lambda + \psi^t\lambda_p) - \phi_p(\psi_q^t\lambda + \psi^t\lambda_q) \\ &= \{\phi, \psi\}\lambda + \{\phi, \lambda\}\psi \end{aligned}$$

The generalized Poisson bracket described here is purely a computational device and not technically a Poisson bracket in the classical sense (see e.g. [17]). In particular, the Poisson bracket of a vector function with itself is a skew symmetric matrix; moreover, the development of a Jacobi identity for this new bracket would require that the concept be further generalized to allow one to take the Poisson bracket of a matrix–valued function with a vector–valued function.

To get an invariant, we require

$$\begin{aligned} \{g, H_T\} = \{g, H\} + \{g, \mu^t g\} + \{g, \eta^t G p\} &= 0 \\ \{Gp, H_T\} = \{Gp, H\} + \{Gp, \mu^t g\} + \{Gp, \eta^t G p\} &= 0 \end{aligned}$$

Working out the Poisson brackets in the first equation, we get

$$-\{g,H\} = \{g,g\}\mu + \{g,\mu\}g + \{g,Gp\}\eta + \{g,\eta\}Gp \tag{3.1}$$

If we do not take the constraints to be satisfied and seek $\mu$ and $\eta$ so that e.g. $\{g,H_T\} \equiv 0$, then we need to solve a system of partial differential equations which actually becomes singular along the constraints, thus it seems to be too much to ask for strong invariance of the constraints.

On the other hand, for a weak invariant, we may assume that $g = Gp = 0$. Next, note that $\{g,g\}$ vanishes because $g$ is a function of $q$ only. Moreover, $\{g,Gp\} = GG^t$, thus

$$-\{g,H\} = GG^t\eta$$

This can be solved for $\eta$ provided $G$ has full rank.

The second equation can be reduced to

$$-\{Gp,H\} = -GG^t\mu + \{Gp,\mu\}g + [(Gp)_qG^t - G(Gp)_q^t]\eta - \{Gp,\eta\}Gp$$

Again, for weak invariance, the terms multiplied by $g$ and $Gp$ drop out and we are left with equations which uniquely determine $\mu$. Once $\mu$ and $\eta$ are known, the Hamiltonian function $H_T$ is determined and the unconstrained equations of motion can be found by differentiating $H_T$:

$$\dot{q} = \nabla_p H + \mu_p^t g + \eta_p^t Gp + G^t\eta \tag{3.2}$$

$$\dot{p} = -\nabla_q H - \mu_q^t g - G^t\mu - \eta_q^t Gp - (Gp)_q^t\eta \tag{3.3}$$

From a computational point of view, it may be quite involved to formulate the system in this manner. In particular, we now need to to compute *third* derivatives of $g$ and second derivatives of $H$. Below we will consider some simplifications in the hopes of improving the computational efficacy of Hamiltonian formulation.

## 3.2 Calculation of the Wedge Product

The standard we take for judging the methods presented for the numerical integration of (1.3)–(1.5) is invariance of the wedge product $dq \wedge dp$ of differentials along the numerical solution. The procedure for computing these products is worth setting out. Consider the ODE

$$\dot{w} = h(w) \tag{3.4}$$

If (3.4) is first discretized by an $s$–stage one–step (Runge–Kutta) method:

$$W_i = w_n + h\sum_{j=1}^{s} a_{ij}h(W_j) \tag{3.5}$$

$$w_{n+1} = w_n + h\sum_{j=1}^{s} b_j h(W_j) \tag{3.6}$$

then the differentials satisfy equations

$$\begin{aligned} dW_i &= dw_n + h\sum_{j=1}^{s} a_{ij}h'(W_j)dW_j \\ dw_{n+1} &= dw_n + h\sum_{j=1}^{s} b_j h'(W_j)dW_j \end{aligned}$$

Hence we can compute the behavior of differentials under the discrete version of the flow in tandem with the computation of each time step. (The differentials should be computed after the stage values $W_j$ are known). If we are interested in solutions of (3.4) that lie on an invariant manifold $\mathcal{M}$, it is important that we choose our initial values and differentials for computations to be in the cotangent bundle $T^*\mathcal{M}$. However, for computational purposes, the differentials can just be identified with vectors in the tangent space. For (1.3)–(1.5), this means that $(q_0, p_0, dq_0, dp_0)$ should satisfy the consistency conditions

$$\begin{aligned} 0 &= g(q_0) \\ 0 &= G(q_0)p_0 \\ 0 &= G(q_0)dq_0 \\ 0 &= G_q(p_0, dq_0) + Gdp_0 \end{aligned}$$

where $G_q$ in the latter equation is to be evaluated at $q = q_0$. Now to see how the wedge product fares under discretization, we first select $(q_0, p_0) \in \mathcal{M}$ and two independent vectors $(dq_0^1, dp_0^1)$ and $(dq_0^2, dp_0^2)$ for which the tangency conditions are satisfied at $(q_0, p_0)$. We compute a step of the numerical solution starting from $(q_0, p_0)$ and calculate the corresqonding change in the differentials. This process is iterated to obtain differentials at every step. The wedge product at the $n$th step is calculated as the value of the quadratic form

$$\begin{bmatrix} dq_n^1 \\ dp_n^1 \end{bmatrix}^t \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} dq_n^2 \\ dp_n^2 \end{bmatrix}$$

Although the wedge product as defined will be maintained in the symplectic discretization of Hamiltonian underlying ODEs, this does not mean that the solution and the differentials stay in the tangent bundle $T\mathcal{M}$ at each step.

In Figure 3.1, a numerical experiment with the Hamiltonian underlying ODE for the pendulum in cartesian coordinates is summarized. Here we start with the Hamiltonian $H(q,p) = (p_1^2 + p_2^2)/2 + q_2$, so that gravity and mass are normalized to one. We introduce the position constraint $g(q_1, q_2) = q_1^2 + q_2^2 - 1 = 0$ and the associated constraint on $p$: $q_1p_1 + q_2p_2 = 0$. We computed $\mu$ and $\eta$ as described above. Starting from the initial configuration $(q_1, q_2, p_1, p_2) = (1, 0, 0, -2)$, the resulting Hamiltonian underlying ODE

(3.2)–(3.3) was solved using the implicit midpoint method. Sample differentials of the solution and the wedge product were simultaneously computed using the approach discussed above. The leftmost figures in 3.1 show the extent to which the constraint residuals are maintained during integration; the right figure demonstrates that the wedge product is indeed invariant under the discrete version of the flow.

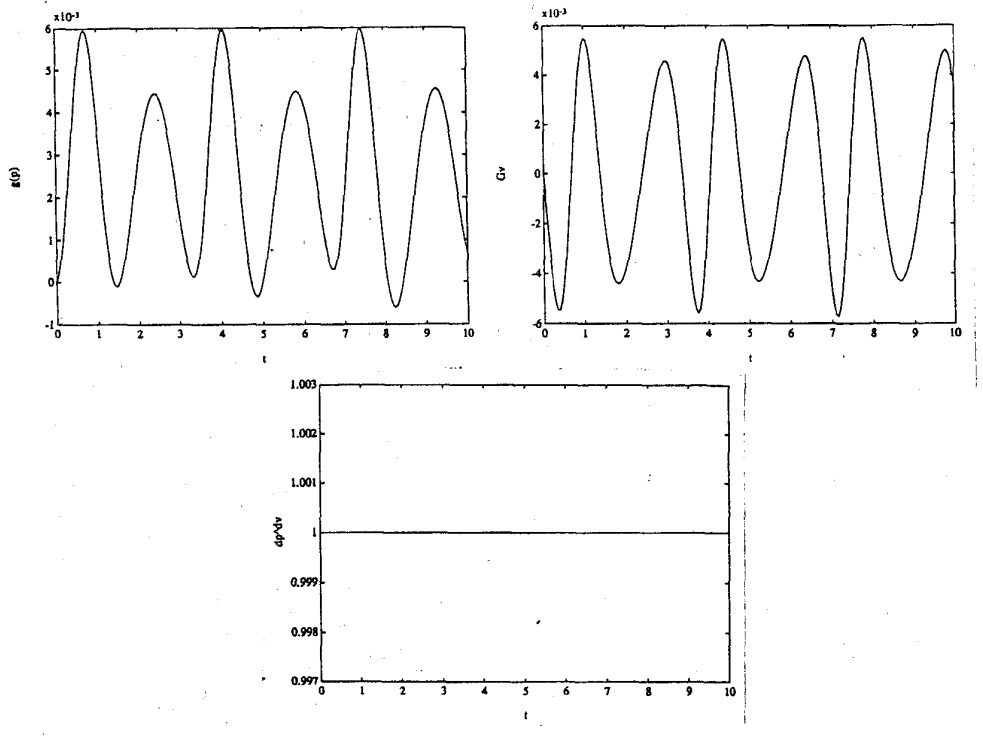


Figure 3.1: The Hamiltonian underlying ODE of Dirac for the pendulum; constraints and wedge product.

## 3.3 Weakly Hamiltonian Underlying ODE

Dirac's process requires the differentiation of the constraint multipliers $\mu$ and $\eta$; since $\mu$ and $\eta$ depend on second derivatives of $g$ and first derivatives of $H$, construction of a Hamiltonian underlying ODE along the lines of Dirac's theory in general requires third derivatives of $g$ and second derivatives of $H$. However, along the constraint manifold $\mathcal{M}$, which is an invariant under the flow of (3.2)–(3.3), the terms multiplying the partial derivatives of $\mu$ and $\eta$ vanish, and we are left with a simplified system:

$$\dot{q} = \nabla_p H + G^t \eta \tag{3.7}$$

$$\dot{p} = -\nabla_q H - G^t \mu - (Gp)^t_q \eta \tag{3.8}$$

This system (referred to below as the "Weakly Hamiltonian Dirac formulation") behaves like a Hamiltonian system for initial values chosen on the constraint manifold; in fact, *any* underlying ODE is a Hamiltonian system along the constraint manifold. But under numerical discretization we cannot in general expect the constraints to be maintained exactly, so that a canonical ODE discretization scheme applied to (3.7)–(3.8) would not result in a canonical step–to–step map. On the other hand, (3.7)–(3.8) requires only the computation of second derivatives of $g$ and first derivatives of $H$, hence it may be much more easily computed for certain problems. This formulation has been treated in the literature regarding Lagrangian formulations of the equations of motion (see e.g. [10]). Its behavior with respect to the wedge product will be discussed in more detail in §4.

## 3.4 A Simplification

The process outlined above for obtaining a weak invariant is not completely well–defined. For example, the determination of $\mu$ from

$$-\{g, H\} = GG^t\eta$$

can be done in any number of ways if we are allowed to freely use the relation $g = 0$ or $Gp = 0$. For example, taking

$$-\{g, H\} + \alpha g = GG^t\eta$$

does not cause $g = 0$ to cease to be an invariant of the flow ultimately obtained, but it may change characteristics of that flow for points near $\mathcal{M}$ where $g \neq 0$. We may also note that if $H(q,p) = F(q) + \frac{p^tp}{2}$ and $g = g(q)$, then $\{g, H\} = Gp$ which is weakly zero when we are constraining with respect to both $g(q) = 0$ and $Gp = 0$. This means that we have the obvious choice of taking $\{g, H\} = 0$ which leads to $\eta = 0$, or to follow the derivation as outlined above which would lead to $\eta \neq 0$ away from $\mathcal{M}$. There is no obvious, a priori reason to favor one of these formulations over the other.

If we take $\eta = 0$, we get

$$H_T = H + \mu^t g$$

so that, after insuring that $\mathcal{M}$ is invariant, we arrive at

$$\dot{q} = p + \mu_p^t g \tag{3.9}$$
$$\dot{p} = -\nabla F - \mu_q^t g - G^t\mu \tag{3.10}$$

where $\mu = (GG^t)^{-1}(G\nabla F - G_q(p,p))$. This system requires the computation of third derivatives of $g$ and second derivatives of $H$ as before.

Besides providing a simplified Hamiltonian formulation, (3.9)–(3.10) has the immediate and natural consequence of showing that along the constraint ($g = 0$), the standard underlying ODE generates a Hamiltonian flow. However, the formulation (3.9)–(3.10) contains a somewhat surprising instability which can be observed in computations whenever numerical discretization induces a perturbation of the constraint. In Figure 3.2, the implicit midpoint method (a canonical discretization scheme) has been applied to solve (3.9)–(3.10) for the cartesian pendulum discussed above with fixed stepsize $h = .01$ from $t = 0$ to $t = 1$ with the same initial conditions as for Figure 3.1. Although the wedge product is maintained in this case, the constraint residuals are very rapidly growing in time.

Figure 3.2 appears to contradict a result in Cooper [6] that says that quadratic integral invariants are precisely maintained by one step methods (like the implicit midpoint method) which are "marginally algebraically stable." However, the integral invariants in [6] are always taken to be *strong* invariants which implies that the invariant manifold is in a certain sense locally stable; as we see below, this is not the case for a weak invariant.

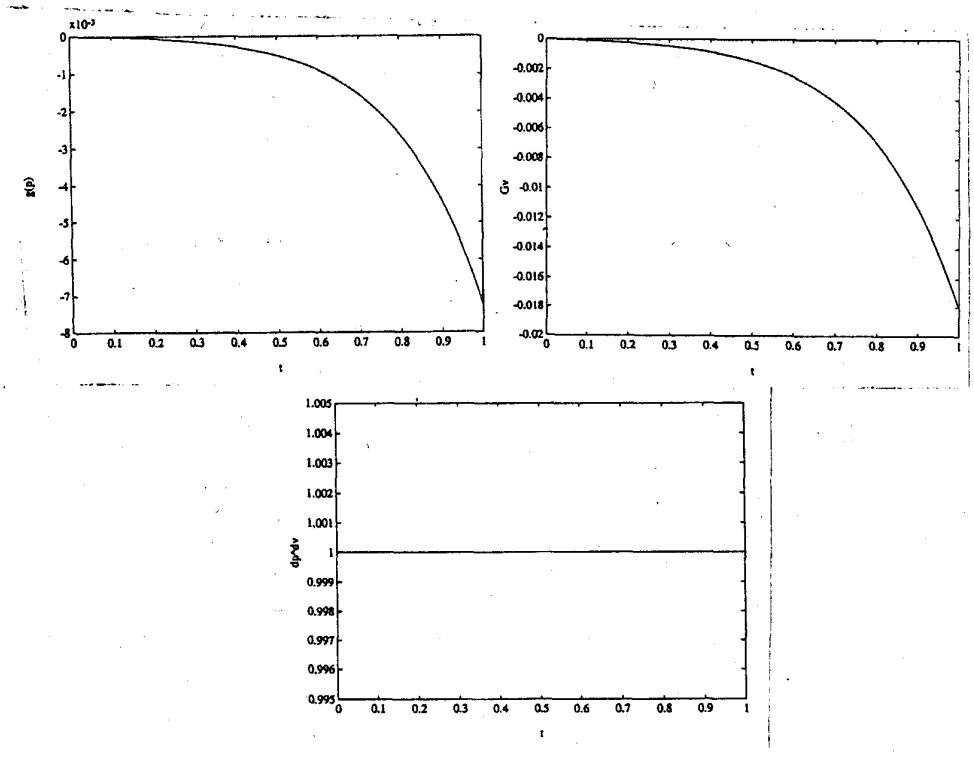


Figure 3.2: Constraint residuals and wedge product for the simplified gradient system formulation.

## 3.5 Stability of the Constraint–Invariants

Let us begin with the case of a linearly constrained quadratic Hamiltonian with constraint $\phi = Gq$ with $G$ constant. Here we find that the simplified Hamiltonian system based on $\dot{H}_T = H + \mu^t Gq$ is

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -(I - \mathcal{H})\nabla F + D^2 F \mathcal{H} q \end{aligned}$$

where $D^2F$ is the Hessian matrix of $F$. Multiplying both equations by the (here constant) projector $\mathcal{H}$, we obtain

$$\begin{aligned} \mathcal{H}\dot{q} &= \mathcal{H}p \\ \mathcal{H}\dot{p} &= \mathcal{H}D^2 F \mathcal{H} q \end{aligned}$$

Since $\mathcal{H} = \mathcal{H}^2$, we can change to variables $r = \mathcal{H}q$, $s = \mathcal{H}p$ and write

$$\dot{r} = s \tag{3.11}$$

$$\dot{s} = Br \tag{3.12}$$

where $B = \mathcal{H}D^2F\mathcal{H}$. Invariance of the constraints translates to $r = s = 0$. If the Hessian is constant and positive definite, as is often the case, $B$ is positive semidefinite, and the equilibrium position in (3.11)–(3.12) will be a saddle point. In this situation, one can expect an instability under perturbation of the constraint–invariants introduced via discretization.

If we perform a similar analysis starting from the Hamiltonian $H_T = \frac{p^t p}{2} + F(q) + \mu^t Gq + \eta^t Gp$ of Dirac, we arrive at equations:

$$\begin{aligned} \dot{q} &= (I - 2\mathcal{H})p \\ \dot{p} &= -(I - \mathcal{H})\nabla F + D^2 F \mathcal{H} q \end{aligned}$$

Multiplying the equations by $\mathcal{H}$, we get

$$\begin{aligned} \mathcal{H}\dot{q} &= \mathcal{H}(I - 2\mathcal{H})p = -\mathcal{H}p \\ \mathcal{H}\dot{p} &= \mathcal{H}D^2 F \mathcal{H} q \end{aligned}$$

Hence the corresponding system of differential equations for the constraint residuals is

$$\dot{r} = -s \tag{3.13}$$

$$\dot{s} = Br \tag{3.14}$$

Now the equilibrium position $r = s = 0$ has become a stable center under the assumption that $D^2F$ is positive definite.

It is interesting in this context to compare also the restricted formulations (3.7)–(3.8) and the standard underlying ODE. In the case of linear constraints, (3.7)–(3.8) leads to:

$$\begin{aligned} \dot{q} &= (I - \mathcal{H})p \\ \dot{p} &= -(I - \mathcal{H})\nabla F \end{aligned}$$

and hence to

$$\begin{aligned} \dot{r} &= 0 \\ \dot{s} &= 0 \end{aligned}$$

for the projections $r = \mathcal{H}q$, $s = \mathcal{H}p$. The standard underlying ODE, on the other hand, is nothing other than

$$\begin{aligned} \dot{r} &= s \\ \dot{s} &= 0 \end{aligned}$$

(This case is well studied, see, e.g., [15].)

## 3.6 Nonlinear Constraints

We begin the discussion by writing the equations of motion for both formulations in case $\{\phi, \phi\} = 0$. Beginning with the constraint $\phi = 0$, we derive a constraint of the form $\psi = \{\phi, H\}$. Assuming $\{\phi, \psi\}$ is nonsingular, we arrive at the Hamiltonian $H_T = H + \mu^t\phi + \eta^t\psi$. The conditions on $\mu$ and $\eta$ reduce to

$$\begin{aligned} \{\phi, H_T\} &\approx 0 \quad \Rightarrow \quad \{\phi, H\} + \{\phi, \psi\}\eta = 0 \\ \{\psi, H_T\} &\approx 0 \quad \Rightarrow \quad \{\psi, H\} + \{\psi, \phi\}\mu + \{\psi, \psi\}\eta = 0 \end{aligned}$$

Assuming $\{\psi, \phi\}$ is an invertible matrix, and using $\{\phi, H\} = \psi$, we have

$$\begin{aligned} \eta &= -\{\phi, \psi\}^{-1}\psi \\ \mu &= -\{\psi, \phi\}^{-1}\left[\{\psi, H\} - \{\psi, \psi\}\{\phi, \psi\}^{-1}\psi\right] \end{aligned}$$

Next, we write differential equations for the constraint residuals, thus

$$\begin{aligned} \dot{\phi} &= \{\phi, H_T\} = \psi + \{\phi, -\{\psi, \phi\}^{-1}\left[\{\psi, H\} - \{\psi, \psi\}\{\phi, \psi\}^{-1}\psi\right]\}\phi \\ &\quad + \{\phi, -\{\phi, \psi\}^{-1}\psi\}\psi + \{\phi, \psi\}\left[-\{\phi, \psi\}^{-1}\psi\right] \\ \dot{\psi} &= \{\psi, H_T\} = \{\psi, H\} + \{\psi, -\{\psi, \phi\}^{-1}\left[\{\psi, H\} - \{\psi, \psi\}\{\phi, \psi\}^{-1}\psi\right]\}\phi \\ &\quad + \{\psi, \phi\}\left[-\{\psi, \phi\}^{-1}(\{\psi, H\} - \{\psi, \psi\}\{\phi, \psi\}^{-1}\psi)\right] \\ &\quad + \{\psi, -\{\phi, \psi\}^{-1}\psi\}\psi + \{\psi, \psi\}[-\{\phi, \psi\}^{-1}\psi] \end{aligned}$$

This simplifies to the system

$$\begin{aligned}\dot{\phi} &= -\{\phi,\{\psi,\phi\}^{-1}\{\psi,H\}\}\phi + \{\phi,\{\psi,\phi\}^{-1}\{\psi,\psi\}\{\phi,\psi\}^{-1}\psi\}\phi \\ &\quad -\{\phi,\{\phi,\psi\}^{-1}\psi\}\psi \qquad (3.15)\\ \dot{\psi} &= -\{\psi,\{\psi,\phi\}^{-1}\{\psi,H\}\}\phi + \{\psi,\{\psi,\phi\}^{-1}\{\psi,\psi\}\{\phi,\psi\}^{-1}\psi\}\phi \\ &\quad -\{\psi,\{\phi,\psi\}^{-1}\psi\}\psi \qquad (3.16)\end{aligned}$$

For determining the local stability of the constraints, we would like to expand the right–hand sides here and then eliminate all second–order terms in the constraint residuals. Unfortunately, this is slightly complicated since Proposition 3.2, for example, only applies to products of two vector quantities, and, generally speaking, we haven't yet defined the Poisson bracket of a vector quantity with a matrix–valued quantity like $\{\phi,\psi\}^{-1}$. On the other hand, by looking at Definition 3.1, it is straightforward to see that, for example, if $f_1, f_2, f_3$ are vector functions in $q$ and $p$, and $A$ is a matrix valued function of $q$ and $p$, then

$$\{f_1, A^t f_2\} f_3 = \{f_1, f_2\} A + Q(f_2, f_3)$$

where $Q$ is some tensor evaluated at $f_2$, $f_3$; importantly, $Q(f_2, f_3)$ is a quadratic term in $f_2$ and $f_3$. With this observation, we can neglect the second order terms in (3.15)–(3.16) to obtain

$$\begin{aligned}\dot{\phi} &= -\psi - \{\phi,\{\psi,\phi\}^{-1}\{\psi,H\}\}\phi \qquad (3.17)\\ \dot{\psi} &= -\{\psi,\{\psi,\phi\}^{-1}\{\psi,H\}\}\phi \qquad (3.18)\end{aligned}$$

On the other hand, if we start with $\hat{H}_T \doteq H + \mu^t\phi$ then we obtain via the same sort of calculations

$$\begin{aligned}\dot{\phi} &= \psi - \{\phi,\{\psi,\phi\}^{-1}\{\psi,H\}\}\phi \qquad (3.19)\\ \dot{\psi} &= -\{\psi,\{\psi,\phi\}^{-1}\{\psi,H\}\}\phi \qquad (3.20)\end{aligned}$$

Note that the only difference between (3.17)–(3.18) and (3.19)–(3.20) is the sign that appears with $\psi$ in the first equation of each system.

Lets turn to an example. For the cartesian pendulum, if we neglect the potential entirely and take $H(q,p) = p^t p/2$ and a single constraint, $(q^t q - 1)/2 = 0$, then the Dirac Hamiltonian becomes

$$H_T(q,p) = \frac{p^t p}{2} + \frac{1}{2}\frac{p^t p}{q^t q}(q^t q - 1) - \frac{(q^t p)^2}{q^t q}$$

Differentiating gives

$$\begin{aligned}\dot{q} &= \left(\frac{2q^tq-1}{q^tq}I - \frac{2}{q^tq}qq^t\right)p \\ \dot{p} &= -\frac{p^tp}{(q^tq)^2}q + 2\frac{q^tp}{q^tq}p - 2\left(\frac{q^tp}{q^tq}\right)^2 q\end{aligned}$$

Now setting $\rho = (q^tq-1)/2$, $\sigma = q^tp$, we can write equations for $\dot{\rho}$ and $\dot{\sigma}$:

$$\begin{aligned}\dot{\rho} &= q^t\dot{q} = q^tp - 2q^tp + \frac{q^tq-1}{q^tq}q^tp = -\sigma + \frac{\sigma\rho}{\rho+1/2} \\ \dot{\sigma} &= q^t\dot{p} + p^t\dot{q} = -\frac{p^tp}{q^tq} + p^tp - 2\frac{(q^tp)^2}{q^tq} + p^tp\frac{q^tq-1}{q^tq} \\ &= 2p^tp\frac{\rho}{\rho+1/2} - \frac{\sigma^2}{\rho+1/2}\end{aligned}$$

The term $p^tp$ is a nuisance. If we treat it as a time dependent coefficient, linearizing at $\rho = \sigma = 0$, we get

$$\begin{aligned}\dot{\rho} &= -\sigma \\ \dot{\sigma} &= 4p^tp\rho\end{aligned}$$

which makes the origin a center. Under these conditions, we would expect behavior as in part (a) of Figure 3.3; this agrees with the numerical experiment shown in Figure 3.1

By contrast, if we had only made use of constraints on $q$ in formulating the system, we would have had after following the above analysis and linearizing,

$$\begin{aligned}\dot{\rho} &= \sigma \\ \dot{\sigma} &= 4p^tp\rho\end{aligned}$$

meaning that the origin has become a saddle point and the behavior over time would be as indicated in part (b) of Figure 3.3; this is exactly what we have seen in Figure 3.2.

Although it appears that the general nonlinear case can be quite complicated, some generalization of the comparative analysis for linear systems of the first part of this section is possible via linearization of nonlinear constraints if we bear in mind that a potential energy function always has a positive definite Hessian at least in the neighborhood of a dynamic equilibrium [7].

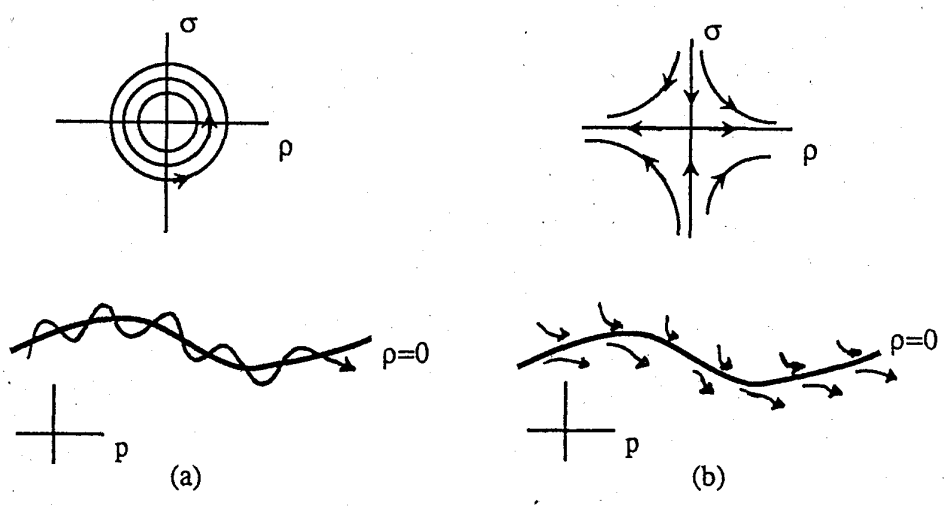



Figure 3.3: Stability of Constraints (a) for Dirac formulation, and (b) for reduced formulation

# 4. Projection Methods

Since linear constraint–invariants are locally stable for (3.7)–(3.8) and only mildly unstable for the standard underlying ODE, and since each of those systems is the restriction of a Hamiltonian system to the constraint manifold, it is natural to ask whether we could not get away with solving one of these two simplified systems using a canonical integration method over moderate time intervals without too much variation in the wedge product. Numerical experiments with, respectively, the standard underlying ODE and (3.7)–(3.8) formulations are summarized in Figures 4.1 and 4.2. Note that nonlinearities in the equations and discretization have made the constraint residuals oscillate in the case of the underlying ODE and introduced some growth in the case of the weakly Hamiltonian Dirac formulation. Perhaps more curious is the apparent periodic oscillation of the wedge product in the latter case.

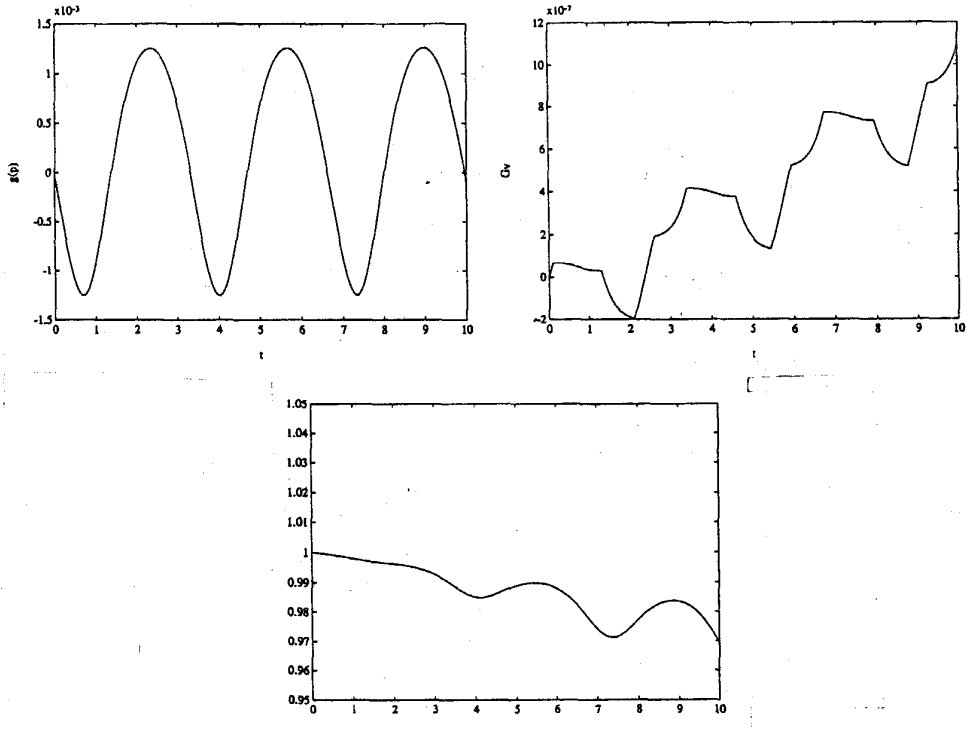


Figure 4.1: Standard underlying ODE formulation of the Pendulum.

Popular approaches to overcoming drift that have been used in multibody simulation are Baumgarte stabilization [4] and projection of the numerical solution onto the constraint manifold. Baumgarte's method corresponds to adding a term to the underlying ODE to cause the constraint to become

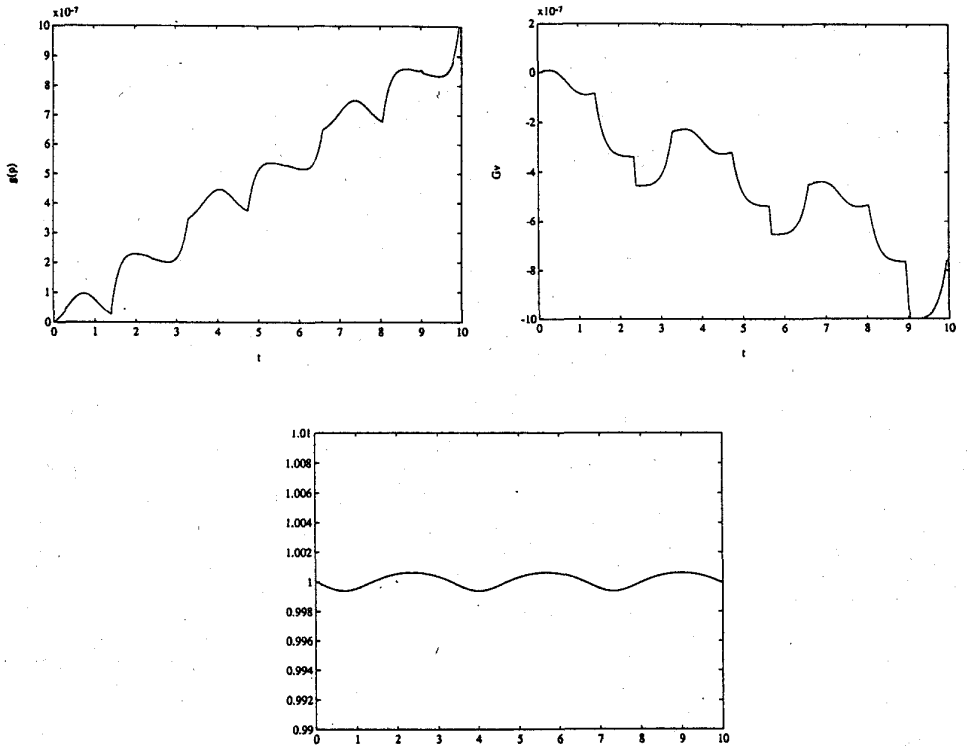


Figure 4.2: The weakly Hamiltonian Dirac formulation for the pendulum.

asymptotically stable. A simple variant in the case of a linearly constrained gradient system would be to solve

$$\begin{aligned} \dot{q} &= p - \mathcal{H}q \\ \dot{p} &= -(I - \mathcal{H})\nabla F - \mathcal{H}p \end{aligned}$$

Then the ODE system for the residuals becomes

$$\begin{aligned} \dot{r} &= s - r \\ \dot{s} &= -s \end{aligned}$$

which makes 0 asymptotically stable. Although these sorts of stabilizations certainly have a place in short time interval integrations, they destroy the Hamiltonian structure, and we do not consider them further here.

Among projection methods we consider the orthogonal projections which have been used by a number of authors (see e.g. [13], [10], [21]). We posed the following question: what happens to the wedge product if we project the numerical solution (obtained via a canonical discretization scheme) of a Hamiltonian system onto an invariant at each step? It is enough to describe

various approaches in terms of the general ODE with invariant

$$\begin{aligned} \dot{w} &= h(w) \\ q(w) &= 0 \end{aligned}$$

There are numerous constraint projections one might consider for this purpose. Here we restrict ourselves to two schemes. One idea is to take a step using some numerical discretization of $\dot{w} = h(w)$ and then to project the resulting value back onto the constraint manifold. For example, we might compute

$$W_i = w_n + h\sum_{j=1}^{s} a_{ij}h(W_j) \tag{4.1}$$

$$\tilde{w}_{n+1} = w_n + h\sum_{j=1}^{s} b_j h(W_j) \tag{4.2}$$

and then solve

$$\begin{aligned} w_{n+1} &= \tilde{w}_{n+1} - q'(w_{n+1})^t\lambda_{n+1} \\ 0 &= q(w_{n+1}) \end{aligned}$$

$d\tilde{w}_{n+1}$ can be obtained by differentiation of (4.1)–(4.2). Differentiating the projection step, we get

$$\begin{aligned} dw_{n+1} &= d\tilde{w}_{n+1} - q'(w_{n+1})^t d\lambda_{n+1} - q''(w_{n+1})^t(dw_{n+1}, \lambda_{n+1}) \\ 0 &= q'(w_{n+1})dw_{n+1} \end{aligned}$$

assuming $q'$ has full rank, we will be able to solve for $dw_{n+1}$ and $d\lambda_{n+1}$ at each step. We term this approach *endpoint projection*. Another approach is based on projection of the intermediate (stage) values: we solve

$$\begin{aligned} W_i &= w_n + h\sum_{j=1}^{s} a_{ij}\dot{W}_j - q'(W_i)^t\Lambda_i \\ 0 &= q(W_i) \end{aligned}$$

for $i = 1, \ldots, s$, then compute

$$w_{n+1} = w_n + h\sum_{j=1}^{s} b_j h(W_j)$$

Here, too, one obtains equations for the differentials in a straightforward way. In the case of the implicit midpoint method, we refer to this scheme as midpoint projection.

The stability of these projection schemes has been analyzed by Ascher and Petzold [3]. The first scheme is numerically stable while the second is known to have an instability. On the other hand, the discussion of [3] was not directly concerned with Hamiltonian systems, so it is interesting to ask whether the additional structure present in Hamiltonian systems might not enable the use of midpoint projection.

In Figures 4.3 and 4.4, we show the effects of, respectively, endpoint and midpoint projection on the numerical solution of the true Hamiltonian formulation of Dirac for the cartesian pendulum.

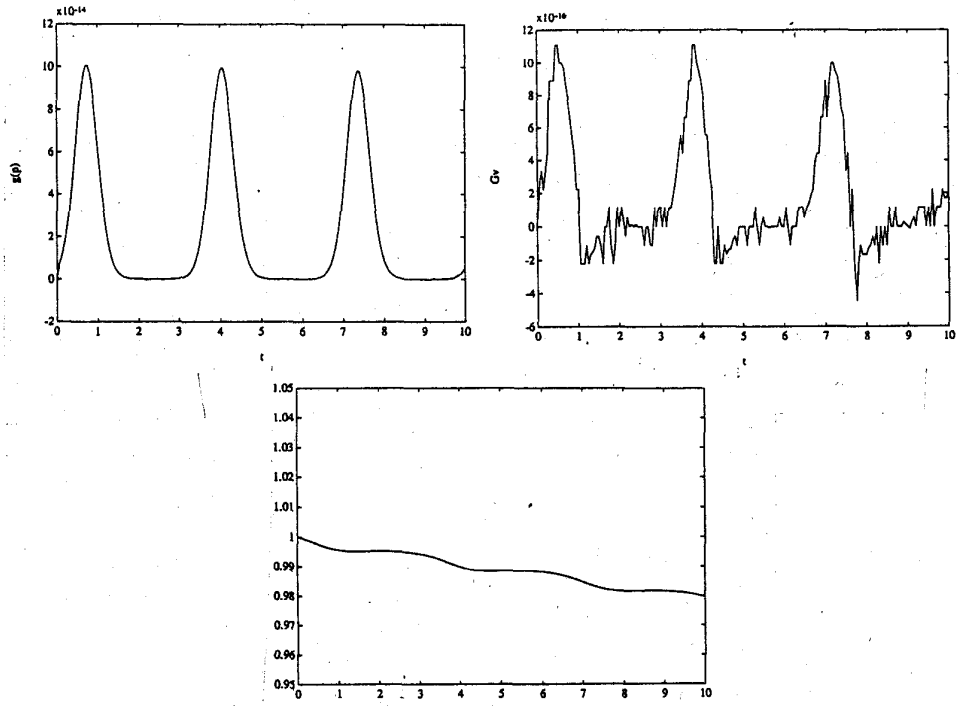


Figure 4.3: Endpoint projection for the Hamiltonian formulation of Dirac.

Although the evident instability of the midpoint projection method makes it appear to be unsuitable for long term computations, it should be pointed out that both projection schemes would probably only be employed occasionally during the integration time interval. In a final experiment, the weakly Hamiltonian Dirac formulation was solved on $[0, 10]$ with infrequent endpoint projection every 20 steps; the results are summarized in Figure 4.5.

Figure 4.5 points out a side effect of infrequent projection: in the likely event that the solution exhibits periodic or quasiperiodic regularity, the projections are unlikely to be performed in phase with the dynamical behavior and this leads to a loss of regular structure in the solutions that result.

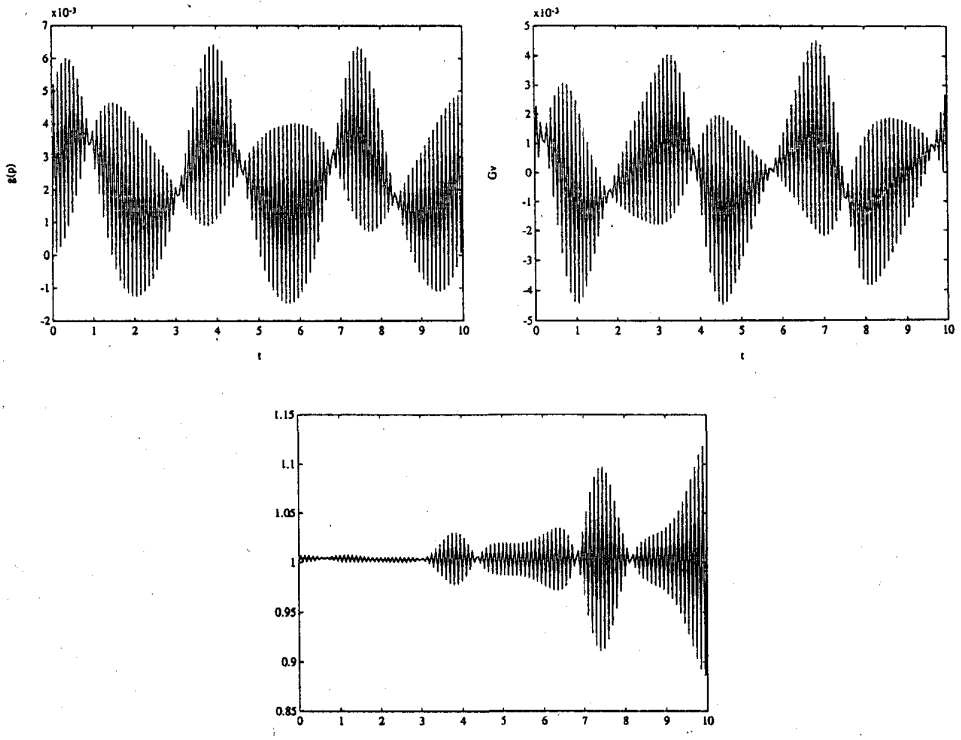


Figure 4.4: Midpoint projection for the Hamiltonian formulation of Dirac.

**Acknowledgement** Linda Petzold pointed us to this problem and guided us in the early stages and we are very thankful for her patience and support. The current work was initiated while the second author was visiting the University of Kansas and completed while the first author was a guest of the Konrad–Zuse–Zentrum in Berlin. We would like to thank Peter Deuflhard and Jacek Szmigielski for their comments and encouragement.

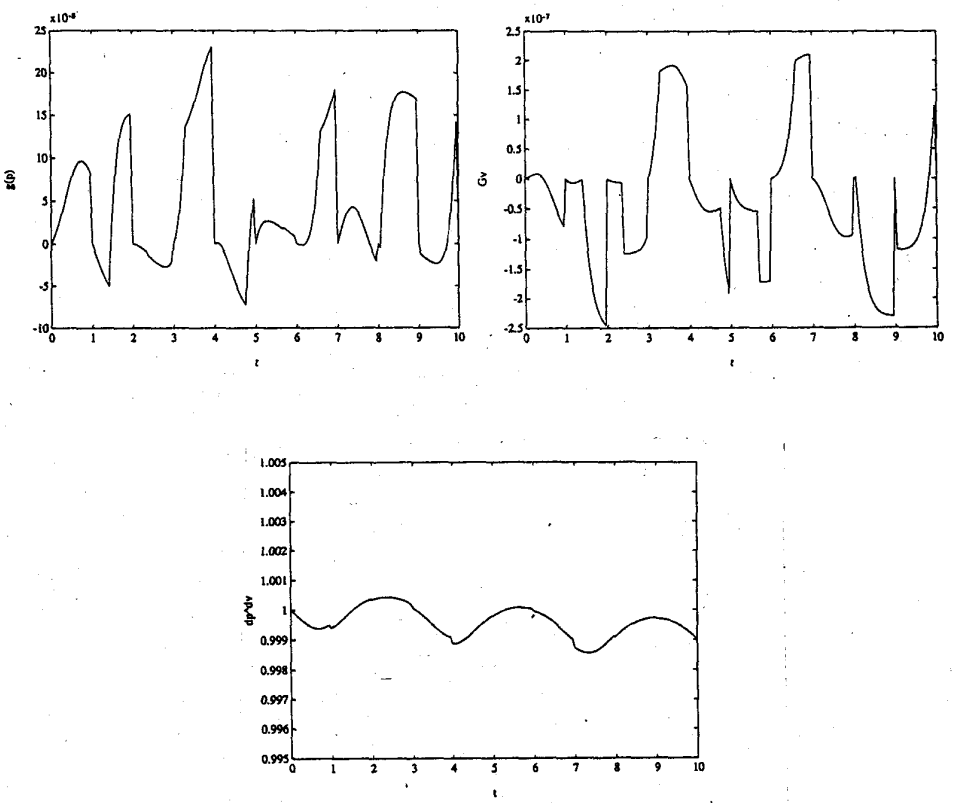



Figure 4.5: Infrequent endpoint projection of the weakly Hamiltonian Dirac formulation.

# Bibliography


[1] Abraham, R., and Marsden, J.E., *Foundations of mechanics*, 2nd Edition, Benjamin Cummings Publishing Co., Inc., Reading, MA., 1978.

[2] Arnold, V.I., *Mathematical methods of classical mechanics*, Springer Graduate Texts in Mathematics No. 60, Springer–Verlag, NY, 1975.

[3] Ascher, U., and Petzold, L.R., "Stability of computational methods for constrained dynamics systems," Report UCRL–JC–107050, Lawrence Livermore National Lab., Livermore, 1991.

[4] Baumgarte, J., "Stabilization of constraints and integrals of motion in dynamical systems," *Comp. Math. Appl. Mech. Eng.* 1 (1976), 1–16.

[5] Brenan, K.E., Campbell, S.L., and Petzold, L.R., *Numerical solution of initial value problems in differential-algebraic equations*, North–Holland, 1989.

[6] Cooper, G.J., Stability of Runge–Kutta methods for trajectory problems, *IMA Journal on Numerical Analysis*, 7, 1–13, 1987.

[7] Courant, R., and Hilbert, D., *Methods of Mathematical Physics*, John Wiley and Sons, NY, 1953.

[8] Deuflhard, P., Uniqueness theorems for stiff ODE initial value problems, Preprint No. SC–87–3, Konrad–Zuse–Zentrum für Informationstechnik Berlin, 1987.

[9] Dirac, P.A.M., *Lectures on Quantum Mechanics*, Belfer Graduate School Monographs, No. 3, Yeshiva University, 1964.

[10] Eich, E., Führer, C., Leimkuhler, B., and Reich, S., Stabilization and projection methods for multibody dynamics," Report A281, Helsinki U. of Tech., Helsinki, 1990.

[11] Griewank, A., On automatic differentiation, *Mathematical Programming, Recent Developments and Applications*, M. Iri and K. Tanake, Eds., Kluwer Academic Pub., Norwell, MA.

[12] Hairer, E., Lubich, Ch., and Roche, M., *The Numerical Solution of Differential–Algebraic Systems by Runge–Kutta Methods*, Lecture Notes in Mathematics, N. 1409, Springer–Verlag, 1989.

[13] Lubich, Ch., $h^2$ extrapolation methods for differential–algebraic systems of index two, *IMPACT*, 1, 260–268, 1989.

[14] McClamroch, N. H. and Bloch, A. M., Control of constrained Hamiltonian systems and applications to control of constrained robots

[15] Mrziglod, T., Zur Theorie und numerischen Realisierung von Lösungmethoden bei Differentialgleichungen mit angekoppelten algebraischen Gleichungen," Diplomarbeit, Math. Inst., Univ. zu Köln, 1987.

[16] Okunbor, D., and Skeel, R.D., An explicit Runge–Kutta–Nyström method is canonical if and only if its adjoint is explicit, *Siam Journal on Numerical Analysis*, 29, No. 2, 1992.

[17] Olver, P., *Applications of Lie Groups to Differential Equations*, Springer–Verlag, 1986.

[18] Potra, F., and Rheinboldt, W., On the numerical solution of the Euler–Lagrange equations, R. Deyo and E. Haug, editors, *NATO Advanced Research Workshop on Real–Time Integration Methods for Mechanical System Simulation*, Springer–Verlag, 1990.

[19] Rheinboldt, W., *Numerical Analysis of Parameterized Nonlinear Equations*, University of Arkansas Lecture Notes in the Mathematical Sciences, 7, John Wiley and Sons, N.Y., 1986.

[20] Sanz–Serna, J.M., Runge–Kutta schemes for Hamiltonian systems, *BIT*, 28, 877–883, 1988.

[21] Simeon, B., Numerical integration of multibody systems by a projection technique, Report TUM–M9201, Math. Inst., TU–München, 1992.