MATHEMATICS OF COMPUTATION
VOLUME 00, NUMBER 00
Xxxx 19xx, PAGES 000–000

NUMERICAL INTEGRATION OF CONSTRAINED HAMILTONIAN SYSTEMS USING DIRAC BRACKETS

WERNER M. SEILER

ABSTRACT. We study the numerical properties of the equations of motion of constrained systems derived with Dirac brackets. This formulation is compared with one based on the extended Hamiltonian. As concrete examples a pendulum in Cartesian coordinates and a chain molecule are treated.

1. INTRODUCTION

The fundamental problem in the numerical integration of a constrained Hamiltonian system (or more generally of any differential algebraic equation [5]) is the drift off the constraint manifold. Geometrically seen, all dynamics happens on this manifold. Only it has a physical meaning; the ambient space is an artifact of the modeling. The dynamics is not well-defined outside the constraint manifold and can be modified, as long as it remains unchanged on the manifold.

Exact solutions are not affected by such modifications. But for numerical solutions any change can make a considerable difference. Due to the discretization error they typically leave the constraint manifold and their stability depends decisively on the properties of the equations in the neighborhood of this manifold.

For Hamiltonian systems Dirac [8, 9] proposed modifications of the dynamics, although for other reasons. He introduced the total and the extended Hamiltonian, respectively, differing from the canonical one by a linear combination of constraint functions. On the constraint manifold both coincide and generate the same dynamics. But the extended Hamiltonian yields more stable equations of motion [15].

We study in this article the Hamilton-Dirac equations of motion [8, 9]. This approach uses a modification of the symplectic structure of the phase space, namely the so-called Dirac bracket, rather than of the Hamiltonian. We will show that it is equivalent to a simplification of the equations of motion derived with the extended Hamiltonian already mentioned in [15].

The basic idea behind the Dirac bracket is the construction of an unconstrained Hamiltonian system (or underlying ordinary differential equation) which has the constraint functions as first integrals. For the special case of a regular system with imposed constraints the impetus-striction formalism [7, 18] achieves the same. In contrast, most index reduction techniques for general differential algebraic equations do not preserve the Hamiltonian structure of the system.

1991 *Mathematics Subject Classification.* Primary 65L05, 70H05; Secondary 70-08.

Key words and phrases. Constrained Hamiltonian system, Dirac bracket, Hamilton-Dirac equations of motion, extended Hamiltonian, numerical integration.

This work was supported by Deutsche Forschungsgemeinschaft.

In order to make this article as self-contained as possible we give in the next two sections a brief review of the Dirac theory and the Hamilton-Dirac equations. In Section 4 we consider the extended Hamiltonian and its relationship to the Dirac theory. After discussing the stability of the constraint manifold for three different approaches we demonstrate the ideas developed so far at a simple toy model. Section 7 specializes the theory to regular systems with imposed constraints. The following two sections contain numerical results for two test problems. Finally, we give some conclusions.

2. The Dirac Theory

Let q^i be coordinates in an N-dimensional configuration space Q. We restrict our presentation to autonomous systems, as explicit time dependencies can always be treated by considering the time as additional coordinate in an extended configuration space. The dynamics of a mechanical system described by a Lagrangian[1] $L(q,\dot{q})$ is given by the Euler-Lagrange equations [11]

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0\,, \qquad i = 1,\dots,N\,. \tag{1}$$

If the Hessian $\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}$ is singular, some equations in (1) are not of second order and the system is *constrained.*

Introduction of the canonically conjugate momenta

$$p_i = \frac{\partial L}{\partial \dot{q}^i}(q,\dot{q}) \tag{2}$$

leads to the Hamiltonian formalism. For a constrained system (2) cannot be solved for all velocities $\dot{q}^i$. Instead one obtains by elimination some *primary constraints*

$$\phi_\alpha(q,p) = 0\,, \qquad \alpha = 1,\dots,A \leq N\,. \tag{3}$$

The *canonical Hamiltonian* of the system is given by

$$H_c(q,p) = p^t\dot{q} - L(q,\dot{q})\,. \tag{4}$$

For an unconstrained system it is obvious that H_c can be considered as a function of (q,p) only, since $\dot{q}$ can be eliminated using (2). Due to the special form of the right hand side of (4), this is also possible in a constrained system, but the resulting H_c is uniquely defined only *on* the constraint manifold. Thus the formalism remains unchanged, if we add an arbitrary linear combination[2] of the constraint functions ϕ [13]. This leads to the *total Hamiltonian*[3]

$$H_t(q,p) = H_c + u^\alpha\phi_\alpha \tag{5}$$

where the multipliers u are a priori arbitrary functions of (q,p).

The standard Hamiltonian formalism is based on the canonical *Poisson bracket* of two phase space functions $F(q,p), G(q,p)$

$$\{F,G\} = \frac{\partial F}{\partial q}\frac{\partial G}{\partial p} - \frac{\partial G}{\partial q}\frac{\partial F}{\partial p}\,. \tag{6}$$

[1] For simplicity we mostly suppress indices, thus q, $\dot{q}$, etc. should be read as vectors.

[2] Here and in the sequel the coefficients of "linear combinations" are allowed to be arbitrary functions of the phase space variables (q,p).

[3] We use the Einstein convention that a summation over repeated indices is always implied.

This bracket is linear in its arguments, skew-symmetric $\{F,G\} = -\{G,F\}$ and satisfies the Jacobi identity $\{F,\{G,H\}\} + \{G,\{H,F\}\} + \{H,\{F,G\}\} = 0$. It gives the phase space the structure of a symplectic manifold. Coordinate transformations $(q,p) \mapsto (Q,P)$ that preserve this structure are called *canonical*.

Using constrained variational calculus one can show that the Euler-Lagrange equations (1) are equivalent to the following first order system [13]

$$\dot{q} = \frac{\partial H_c}{\partial p} + u^\alpha \frac{\partial \phi_\alpha}{\partial p}\,, \qquad \dot{p} = -\frac{\partial H_c}{\partial q} - u^\alpha \frac{\partial \phi_\alpha}{\partial q}\,, \qquad \phi_\alpha = 0\,. \tag{7}$$

The differential part of (7) is not Hamiltonian. However, we can use the following Hamiltonian system

$$\dot{q} = \{q, H_t\}\,, \qquad \dot{p} = \{p, H_t\}\,, \qquad \phi_\alpha = 0\,, \tag{8}$$

as the right hand sides of the differential equations in (7) and (8) differ only by linear combinations of the constraint functions. More generally, the time evolution of any phase space function $F(q,p)$ can be written as

$$\dot{F} = \{F, H_t\}\;. \tag{9}$$

In a consistent theory the constraints $\phi_\alpha = 0$ must be preserved by the evolution of the system. This leads to the conditions

$$\dot{\phi}_\alpha = \{\phi_\alpha, H_t\} \approx 0\,. \tag{10}$$

The $\approx$ signals a *weak equality*; it may hold only after taking the constraints into account. By a standard argument in differential geometry [13] this implies that the Poisson bracket in (10) must be a linear combination of the constraint functions. There are three possibilities: (i) it yields modulo the constraints an equation of the form $1 = 0$; (ii) it becomes $0 = 0$; (iii) we obtain a new equation $\psi(q,p) = 0$.

(i) implies inconsistent equations of motion; they do not possess any solution. (ii) is the desired outcome. (iii) splits into two sub-cases. If ψ depends on some of the multipliers u, we consider it as an equation determining one of them.[4] Otherwise we have a *secondary constraint*. We must then check whether all secondary constraints are preserved by repeating the procedure until we either encounter case (i) or all constraints lead to case (ii). This is the *Dirac algorithm* [8, 9].

The Dirac algorithm is sometimes surprisingly subtle [13]. We consider here only a trivial example with the Lagrangian $L = \frac{1}{2}(\dot{q}^1)^2 - V(q^1,q^2)$. The momenta are $p_1 = \dot{q}^1$ and $p_2 = 0$. Thus there is one primary constraint function $\phi_1 = p_2$. The total Hamiltonian is $H_t = \frac{1}{2}p_1^2 + V(q^1,q^2) + up_2$ with a multiplier u. (10) leads to the secondary constraint function $\phi_2 = \{\phi_1, H_t\} = -V_{q^2}$. Applying again (10) yields $\{\phi_2, H_t\} = -V_{q^1q^2}p_1 - V_{q^2q^2}u = 0$. If we assume that $V_{q^2q^2}$ does not vanish, the Dirac algorithm stops here, as this condition determines the multiplier u.

From the point of view of differential equations, the Dirac theory is a special case of the general problem of *completing* a system of differential equations [28]. This problem is also closely related to the concept of an *index* of a differential algebraic equation. Essentially, the (differential) index corresponds to the number of iterations needed in the Dirac algorithm [27].

[4]Note that as these are weak equations they determine the multipliers only up to linear combinations of the constraint functions.

3. Hamilton-Dirac Equations

Let χ_α ($\alpha = 1, \dots, K$) denote all constraint functions, primary ones and those obtained with the Dirac algorithm. They can be divided into two classes by studying the $K \times K$ matrix of their Poisson brackets

$$C_{\alpha\beta} = \{\chi_\alpha, \chi_\beta\} \; . \tag{11}$$

As C is skew-symmetric, its rank M is even. Let us assume for simplicity that after a simple relabeling of the χ_α the top left $M \times M$ sub-matrix of C is regular (in general we must redefine the constraint functions by taking linear combinations to achieve this). Then we call the constraint functions $\chi_1, \dots, \chi_M$ *second class.*

The Poisson bracket of a *first class* constraint function ψ with any other constraint function χ (primary or higher) vanishes weakly

$$\forall \chi : \quad \{\psi, \chi\} \approx 0 \, . \tag{12}$$

In our case the constraint functions $\chi_{M+1}, \dots, \chi_K$ are first class (again we may have to redefine them by taking linear combinations). Obviously this classification can be performed only after *all* constraints have been found.

First class constraints generate gauge symmetries [13]. One example is the following system which came up in a study of Chern-Simons quantum mechanics [10]

$$L = \frac{1}{2}\left(\dot{q}^1 - q^3 q^2\right)^2 + \frac{1}{2}\left(\dot{q}^2 + q^3 q^1\right)^2 \; . \tag{13}$$

It describes a charged particle moving in a plane under the influence of a perpendicular constant magnetic field. There is one primary constraint function $\phi_1 = p_3$ generating one secondary constraint function $\phi_2 = q^2 p_1 - q^1 p_2$. Both are first class and essentially generate the rotational symmetry of the system.

First class constraints lead to arbitrary functions in the general solution of the equations of motion; these are under-determined [28]. In the example described by the Lagrangian (13) q^3 remains arbitrary. In the sequel we will always assume that no first class constraints are present. This is no real restriction, as they appear very rarely in finite-dimensional systems. Furthermore they can always be transformed into second class constraints by a gauge fixing, i.e. by adding further constraints removing the under-determinacy.

Second class constraints signal the presence of unphysical or redundant degrees of freedom; as mentioned above their number M is always even. A trivial example is $q^1 = p_1 = 0$. If there are no first class constraints, the matrix C defined by (11) is regular (otherwise we take the sub-matrix of C corresponding to the second class constraint functions) and we can introduce the *Dirac bracket* [8] of two phase space functions F, G by

$$\{F, G\}^* = \{F, G\} - \{F, \chi_\alpha\} \left(C^{-1}\right)^{\alpha\beta} \{\chi_\beta, G\} \; . \tag{14}$$

In the case of our trivial example this means that we simply omit in (6) the differentiations with respect to q^1, p_1.

The Dirac bracket possesses exactly the same algebraic properties as the canonical Poisson bracket (6): it is linear, skew-symmetric and satisfies the Jacobi identity. Hence it can be used instead of (6) to define a symplectic structure on the phase space. We will show now that restricted to the constraint manifold both brackets generate the same dynamics.

Consider for any function $F(q,p)$ the dynamics defined by

$$\dot{F} = \{F, H_c\}^* \,. \tag{15}$$

We prove in two steps that for initial data on the constraint manifold this dynamics is equivalent to the original one defined by (9). It suffices to show that the right hand sides of the respective equations of motions (9) and (15) are weakly equal, as for such initial data the trajectories never leave the constraint manifold.

As first step we show that the evolution (15) is weakly equal to the one generated by the total Hamiltonian H_t using Dirac brackets:

$$\begin{aligned}\{F, H_t\}^* &= \{F, H_t\} - \{F, \chi_\alpha\}(C^{-1})^{\alpha\beta}\{\chi_\beta, H_t\} \\ &\approx \{F, H_c\} - \{F, \chi_\alpha\}(C^{-1})^{\alpha\beta}\{\chi_\beta, H_c\} + \\ &\qquad u^\gamma\Big(\{F, \chi_\gamma\} - \{F, \chi_\alpha\}(C^{-1})^{\alpha\beta}\{\chi_\beta, \chi_\gamma\}\Big) \\ &= \{F, H_c\}^* \,.\end{aligned} \tag{16}$$

Here we used in the second line that all Poisson brackets involving the multipliers u are multiplied by constraint functions and in the last line the definition (11) of C.

As second step we note that on the constraint manifold Dirac and Poisson bracket generate the same dynamics with H_t:

$$\{F, H_t\}^* = \{F, H_t\} - \{F, \chi_\alpha\}(C^{-1})^{\alpha\beta}\{\chi_\beta, H_t\} \approx \{F, H_t\} \,, \tag{17}$$

as after completion of the Dirac algorithm $\{\chi_\beta, H_t\}$ is for all β a linear combination of constraint functions. We are thus lead to the *Hamilton-Dirac equations*

$$\begin{aligned}\dot{q} = \{q, H_c\}^* &= \frac{\partial H_c}{\partial p} - \frac{\partial \chi_\alpha}{\partial p}(C^{-1})^{\alpha\beta}\{\chi_\beta, H_c\} \,, \\ \dot{p} = \{p, H_c\}^* &= -\frac{\partial H_c}{\partial q} + \frac{\partial \chi_\alpha}{\partial q}(C^{-1})^{\alpha\beta}\{\chi_\beta, H_c\} \,.\end{aligned} \tag{18}$$

For historical correctness one should remark that Dirac did not consider (18). He used the total Hamiltonian H_t instead of the canonical one H_c. But we proved above that the corresponding equations of motion are weakly equal. Computationally the use of H_c is more efficient, as it leads to simpler equations.

The Dirac bracket effectively eliminates the second class constraints, as they become *distinguished* or *Casimir functions:* the Dirac bracket of any phase space function F with a second class constraint function vanishes strongly, i. e. everywhere in phase space, as again by the definition (11) of C

$$\{F, \chi_\gamma\}^* = \{F, \chi_\gamma\} - \{F, \chi_\alpha\}(C^{-1})^{\alpha\beta}\{\chi_\beta, \chi_\gamma\} = 0 \,. \tag{19}$$

4. The Extended Hamiltonian

The distinction into first and second class constraints is an intrinsic one, i. e. it has a geometric meaning. In contrast, the distinction into primary and secondary (or higher) constraints is to some extent artificial and depends on the precise form of the Lagrangian L. There might exist an equivalent Lagrangian, i. e. one describing the same system, yielding different primary constraints.

Furthermore, if one looks at the argument for introducing the total Hamiltonian, one sees that one could also apply it to secondary constraints. These considerations lead to the *extended Hamiltonian* H_e which is the canonical Hamiltonian H_c plus a linear combination of *all* constraint functions and not just the primary ones.

This approach was used by Leimkuhler and Reich [15] for the symplectic integration of constrained Hamiltonian systems. By calling H_e extended Hamiltonian we slightly abuse Dirac's terminology. He added only the first class class constraint functions based on symmetry and not on stability considerations.

Assuming that all constraint functions χ are second class we make the ansatz

$$H_e = H_c + v^\alpha \chi_\alpha . \tag{20}$$

Recall that the v should not be considered as new variables but as so far unknown functions of (q,p)! Demanding $\{\chi_\alpha, H_e\} \approx 0$ yields the condition

$$\{\chi_\alpha, H_e\} = \{\chi_\alpha, H_c\} + \{\chi_\alpha, v^\beta\}\chi_\beta + \{\chi_\alpha, \chi_\beta\}v^\beta \approx 0 , \tag{21}$$

If we discard the Poisson brackets with v, since they are multiplied by constraint functions, (21) becomes a system of linear equations with the particular solution

$$v^\alpha = -(C^{-1})^{\alpha\beta}\{\chi_\beta, H_c\} \tag{22}$$

with C given by (11). Further solutions of the weak equation (21) are obtained by adding an arbitrary linear combination of constraint functions to each of the v^α.

This suggests the following equations of motion

$$\dot q = \{q, H_e\} , \quad \dot p = \{p, H_e\} . \tag{23}$$

We will see below that they yield the correct dynamics, as (23) is weakly equal to the Hamilton-Dirac equations (18).

The extended Hamiltonian leads to considerably more involved equations of motion than the Hamilton-Dirac approach. The multipliers and thus H_e depend on the matrix C^{-1} also appearing in the Dirac bracket (14). In the equations of motion (23) there arise terms from the Poisson brackets of the dynamical variables with the entries of C^{-1} and these terms are typically rather complicated.

Leimkuhler and Reich [15] considered a simplification which they called the "weakly Hamiltonian Dirac formulation". It arises by discarding the terms containing the Poisson brackets with the multipliers. This is allowed, since they vanish weakly. Using the solution (22) for the multipliers we obtain as equations of motion

$$\begin{aligned} \dot q &= \{q, H_e\} \approx \{q, H_c\} - \{q, \chi_\alpha\}(C^{-1})^{\alpha\beta}\{\chi_\beta, H_c\} , \\ \dot p &= \{p, H_e\} \approx \{p, H_c\} - \{p, \chi_\alpha\}(C^{-1})^{\alpha\beta}\{\chi_\beta, H_c\} . \end{aligned} \tag{24}$$

Thus we recover the Hamilton-Dirac equations (18)! Leimkuhler and Reich claimed that they were not Hamiltonian. We can now correct this statement. Although (18) is not Hamiltonian with respect to the canonical Poisson bracket, it is with respect to the Dirac bracket.

Above derivation of the extended Hamiltonian H_e is a special case of a more general construction [31]. With any phase space function A we can associate a function $A^* \approx A$ such that $\{A^*, \chi\} \approx 0$ for all constraint functions χ:

$$A^* = A - \chi_\alpha (C^{-1})^{\alpha\beta}\{\chi_\beta, A\} . \tag{25}$$

Using (22) for the multipliers we find $H_e = H_c^*$. The Dirac bracket of two functions A, B is weakly equal to the Poisson bracket of their associated quantities

$$\{A, B\}^* \approx \{A^*, B^*\} . \tag{26}$$

5. Constraint Stability

We discuss the stability of the constraint manifold for three formulations of the equations of motion: (i) the classical one (8) based on the total Hamiltonian H_t, (ii) the Hamilton-Dirac equations (18), and (iii) the equations of motion (23) for the extended Hamiltonian H_e. We assume that we are given the canonical Hamiltonian H_c on a $2N$-dimensional phase space and that after completion of the Dirac algorithm there are $K = 2k$ second class constraints $\chi_\alpha(q,p) = 0$.

Let $\Gamma : (q,p) \mapsto (Q,P)$ be a canonical transformation such that in the new coordinates the constraints are given by $Q^a = P_a = 0$ for $a = 1, \ldots, k$.[5] At least locally such a transformation always exists [13]. Independent of which formulation is used the transformed equations of motion can be split into two subsystems

$$(27)\qquad \begin{pmatrix} \dot{Q}^a \\ \dot{P}_b \end{pmatrix} = \begin{pmatrix} U^a_c(Q^i,P_i)\,Q^c + V^{ac}(Q^i,P_i)\,P_c \\ W_{bc}(Q^i,P_i)\,Q^c + Z^c_b(Q^i,P_i)\,P_c \end{pmatrix}, \qquad \begin{pmatrix} \dot{Q}^r \\ \dot{P}_s \end{pmatrix} = \begin{pmatrix} F^r(Q^i,P_i) \\ G_s(Q^i,P_i) \end{pmatrix}.$$

The first part of (27) reflects that for all consistent formulations of the equations of motion the time derivative of any constraint function must vanish weakly and can thus be written as a linear combination of the constraint functions

$$(28)\qquad \dot{\chi}_\alpha = M^\beta_\alpha(q^i,p_i)\,\chi_\beta\,.$$

Considering the variables (Q^r, P_r) as parameters the origin is a fixed point of the first subsystem of (27) and its stability properties can be used as measure for the stability of the constraint manifold.

Entering the constraints $Q^a = P_a = 0$ into the second part of (27) yields a *Hamiltonian state space form*. If we do not use the constraints, we obtain a perturbed state space form allowing for a perturbation theoretic analysis of the stability of the constraint manifold. However, we will not pursue this approach here but concentrate on the stability of the origin in the first subsystem of (27).

The stability analysis of a Hamiltonian system differs in several aspects from that of a general dynamical system. Its linearization yields a Hamiltonian matrix. If λ is an eigenvalue of such a matrix, $\bar{\lambda}$, $-\lambda$ and $-\bar{\lambda}$ are also eigenvalues [2]. Hence a fixed point can be linearly stable, if and only if all eigenvalues are zero or purely imaginary and their algebraic and geometric multiplicities are equal [17].

Furthermore no asymptotically stable fixed points exist. A stable fixed point is never hyperbolic and the Hartman-Grobman theorem cannot be applied. Actually, linear stability is neither sufficient nor necessary for nonlinear stability. The only simple criterion for nonlinear stability is the theorem of Dirichlet [29]: if the eigenvalues are as above and the Hessian of the Hamiltonian with respect to all canonical variables is definite at the fixed point, then it is nonlinearly stable. Otherwise the stability can be only established with a normal form computation [2, 29].

For the Hamilton-Dirac equations (18) the stability analysis is easy. According to (19) the constraint functions χ_α are distinguished functions

$$(29)\qquad \dot{\chi}_\alpha = \{\chi_\alpha, H_c\}^* = 0\,.$$

Hence the constraint functions are first integrals of the flow generated by (18) and in (27) the matrices U, V, W and Z vanish. This implies that the origin is stable. Obviously, there is no need to distinguish between linear and nonlinear stability.

[5] For the remainder of this section we adopt the following convention: indices a, b, c are always running from 1 to k, indices r, s from $k+1$ to N and indices i, j from 1 to N.

This result has the following geometric meaning. The constraint functions χ foliate the phase space into disjoint submanifolds $\mathcal{M}_\epsilon$ defined by $\chi_\alpha(q,p) = \epsilon_\alpha$ with constants ϵ. Exact solutions of the Hamilton-Dirac equations (18) lie completely on the submanifold $\mathcal{M}_\epsilon$ determined by the initial data. The equations do not "see" the values ϵ; especially $\epsilon = 0$ is not distinguished.[6] Numerical errors are neither damped nor amplified by the dynamics. They lead to different values $\bar\epsilon$ and without further errors the trajectory would stay on the submanifold $\mathcal{M}_{\bar\epsilon}$.

For many constraint functions this result implies that the constraint manifold $\mathcal{M}_0$ is *orbitally stable*[7] though not attractive. For example, if the submanifolds $\mathcal{M}_\epsilon$ are compact, there exists obviously a constant upper bound (depending only on ϵ) for $\mathrm{dist}(X, \mathcal{M}_0)$ with $X \in \mathcal{M}_\epsilon$. The same holds in the for applications important case that the constraint functions χ are quadratic.

In order to study the equations of motion derived with the total and extended Hamiltonian, respectively, we denote by $\tilde{H}_c$, $\tilde{H}_t$ and $\tilde{H}_e$ the with Γ transformed Hamiltonians, by A_c, A_t and A_e their Hessians with respect to the variables (Q^a, P_a) evaluated at the origin and by $J = \begin{pmatrix} 0 & I_k \\ -I_k & 0 \end{pmatrix}$ the $K \times K$ symplectic matrix. For the classical equations of motion (8) it is not possible to make any general statements. Linear stability is decided by the eigenvalues of $B_t = JA_t$. However, (numerical) experience shows that usually the origin is unstable. Otherwise the drift off the constraint manifold would not be a serious problem.

In the approach based on the extended Hamiltonian the precise form of the relevant matrices depends crucially on the chosen solution of the linear system (21) for the multipliers v. Leimkuhler and Reich [15] showed for the special case of the pendulum (see Section 8) that the origin is a center, if the v are determined using (22), whereas for another choice of v it becomes a saddle point. In principle, one could use the stability analysis as a guideline for choosing the precise form of the multipliers. But this seems hardly feasible in practice.

Using (22) for the multipliers we find

$$M_\alpha^\beta = -\left\{\chi_\alpha, (C^{-1})^{\beta\gamma}\{\chi_\gamma, H_c\}\right\} \tag{30}$$

(as above $C_{\alpha\beta} = \{\chi_\alpha, \chi_\beta\}$) and in the transformed coordinates (Q,P)

$$U_b^a = -\frac{\partial^2 \tilde{H}_c}{\partial Q^a \partial P_b}\,, \quad V^{ab} = -\frac{\partial^2 \tilde{H}_c}{\partial Q^a \partial Q^b}\,, \quad W_{ab} = \frac{\partial^2 \tilde{H}_c}{\partial P_a \partial P_b}\,, \quad Z = -U^t\,. \tag{31}$$

Evaluated at the origin the block matrix $\begin{pmatrix} U & V \\ W & Z \end{pmatrix}$ becomes $B_c = JA_c$.

Thus the stability depends not only on the choice of the multipliers v but also on the precise form of the canonical Hamiltonian H_c. Recall from Section 2 that H_c is uniquely defined only on the constraint manifold. We can add arbitrary linear combinations of the primary constraint functions to it and such modifications change the matrix B_c. The Hamilton-Dirac equations are less sensitive to such changes; they affect only the second and not the first subsystem of (27) and thus only the perturbed state space form but not the stability of the constraint manifold.

[6] This is also evident from the fact that the Dirac bracket depends only on the derivatives of the constraint functions and not on the functions themselves.

[7] A manifold $\mathcal{M}$ is called orbitally stable for a dynamical system $\dot z = f(z)$, if for every $\epsilon > 0$ there exists a $\delta > 0$ such that for any solution $z(t)$ satisfying $\mathrm{dist}(z(0), \mathcal{M}) < \delta$ the inequality $\mathrm{dist}(z(t), \mathcal{M}) < \epsilon$ holds [12].

6. A Theoretical Example

We demonstrate our theoretical results at a simple system defined by the Lagrangian $L = \frac{1}{2}(\dot{q}^1)^2 - V(q^1 - q^2)$ where we assume that the function V satisfies $V(0) = V'(0) = 0$ and $V''(0) \neq 0$. Later we will further specialize to $V(x) = \frac{1}{2}x^2$. We showed already in Section 2 how the Dirac algorithm works for such a system: there is one primary constraint function $\phi_1 = p_2$ and one secondary $\phi_2 = V'(q^1 - q^2)$. The total Hamiltonian is

$$H_t = \frac{1}{2}p_1^2 + p_1 p_2 + V(q^1 - q^2)\,. \tag{32}$$

Since $\{\phi_1, \phi_2\} = V''(q^1 - q^2)$, the Dirac bracket is defined by

$$\{F, G\}^* = \{F, G\} + \frac{\{F, \phi_1\}\{\phi_2, G\} - \{F, \phi_2\}\{\phi_1, G\}}{V''(q^1 - q^2)}\,. \tag{33}$$

It is also straightforward to determine the extended Hamiltonian

$$H_e = \tfrac{1}{2}p_1^2 + p_1 p_2 + V(q^1 - q^2) - \frac{\left[V'(q^1 - q^2)\right]^2}{V''(q^1 - q^2)}\,. \tag{34}$$

The canonical transformation Γ may be taken as

$$Q^1 = V'(q^1 - q^2)\,, \quad Q^2 = q^1\,, \quad P_1 = -\frac{p_2}{V''(q^1 - q^2)}\,, \quad P_2 = p_1 + p_2\,. \tag{35}$$

In order to decide the stability of the origin for the first subsystem of (27) we need the Hessian of the total and the extended Hamiltonian, respectively, with respect to Q^1 and P_1 evaluated at the origin. After some straightforward applications of the chain rule one finds

$$A_t = \begin{pmatrix} \left[V''(0)\right]^{-1} & 0 \\ 0 & -V''(0) \end{pmatrix}, \qquad A_e = \begin{pmatrix} -\left[V''(0)\right]^{-1} & 0 \\ 0 & -V''(0) \end{pmatrix}. \tag{36}$$

A_t is indefinite, thus we cannot apply the criterion of Dirichlet and obtain no information about the nonlinear stability for the classical formulation. In contrast, A_e is definite (independent of the sign of $V''(0)$) and thus for the extended Hamiltonian formulation the origin is even nonlinearly stable.

Although it is possible to solve exactly the equations of motion for arbitrary functions V, we specialize now to $V(x) = \frac{1}{2}x^2$. The classical equations of motion are computed using $\dot{F} = \{F, H_t\}$

$$\dot{q}^1 = p_1 + p_2\,, \quad \dot{q}^2 = p_1\,, \quad \dot{p}_1 = -\dot{p}_2 = q^2 - q^1\,. \tag{37}$$

We solve this system for some initial point $(q_0^1, q_0^2, p_{10}, p_{20})$. Denoting the constraint residuals at this point by $\rho = p_{20}$ and $\zeta = q_0^1 - q_0^2$, we can write the solution as

$$\begin{aligned} q^1(t) &= q_0^1 + (p_{10} + \rho)t\,, \\ q^2(t) &= q_0^1 + (p_{10} + \rho)t + \tfrac{1}{2}(\rho - \zeta)e^{-t} - \tfrac{1}{2}(\rho + \zeta)e^t\,, \\ p_1(t) &= p_{10} + \rho - \tfrac{1}{2}(\rho - \zeta)e^{-t} - \tfrac{1}{2}(\rho + \zeta)e^t\,, \\ p_2(t) &= \tfrac{1}{2}(\rho - \zeta)e^{-t} - \tfrac{1}{2}(\rho + \zeta)e^t\,. \end{aligned} \tag{38}$$

This implies that for $\rho + \zeta \neq 0$ the constraint residuals grow exponentially and the constraint manifold is unstable. For general V we can conclude that the constraint manifold is linearly unstable. This also follows from the fact that the eigenvalues of B_t are ± 1.

The Hamilton-Dirac equations are very simple for this system. They are computed using $\dot F = \{F, H_c\}^*$ with the Dirac bracket (33)

$$\dot q^1 = \dot q^2 = p_1\,, \qquad \dot p_1 = \dot p_2 = 0\,. \tag{39}$$

Taking the same initial point as above we find the solution

$$\begin{aligned} q^1(t) &= q_0^1 + p_{10}t\,, & p_1(t) &= p_{10}\,,\\ q^2(t) &= q_0^1 + \zeta + p_{10}t\,, & p_2(t) &= \rho\,. \end{aligned} \tag{40}$$

The constraint residuals remain constant and the constraint manifold is orbitally stable but not attractive in accordance with our results above.

The equations of motion for the dynamics $\dot F = \{F, H_e\}$ defined with the extended Hamiltonian (34) are identical with the classical ones (37) except for the sign of the forces. But this sign drastically changes the solution

$$\begin{aligned} q^1(t) &= q_0^1 + (p_{10}+\rho)t\,, & p_1(t) &= p_{10} + \rho + \zeta\sin t - \rho\cos t\,,\\ q^2(t) &= q_0^1 + (p_{10}+\rho)t - \rho\sin t - \zeta\cos t\,, & p_2(t) &= -\zeta\sin t + \rho\cos t\,. \end{aligned} \tag{41}$$

The constraint residuals do not grow but oscillate. The origin is a center of the first part of (27); the eigenvalues of $B_e = JA_e$ are $\pm i$. Again this implies that the constraint manifold is orbitally stable though not attractive. As mentioned above this holds even for general V.

If we consider the full solution instead of only the constraint residuals, we see that the difference between the position coordinates in (41) for different values of ρ grows linearly. In the Hamilton-Dirac equations even this difference remains constant. They represent the most stable formulation for a numerical integration among the three considered.

This example also demonstrates well the effect of a redefinition of H_c. Instead of using $H_c = \frac{1}{2}p_1^2 + V(q^1 - q^2)$ we take $\bar H_c = \frac{1}{2}(p_1 + p_2)^2 + V(q^1 - q^2)$. Obviously, H_c and $\bar H_c$ differ only by a multiple of the primary constraint function $\phi_1 = p_2$. The multiplier u must now satisfy $u \approx 0$; thus we set $\bar H_t = \bar H_c$. For $V(x) = \frac{1}{2}x^2$ the classical equations of motion are

$$\dot q^1 = \dot q^2 = p_1 + p_2\,, \quad \dot p_1 = -\dot p_2 = q^2 - q^1\,. \tag{42}$$

Although the constraint manifold is still unstable for this dynamics, the residuals grow now only linearly and no longer exponentially, as the general solution is

$$\begin{aligned} q^1(t) &= q_0^1 + (p_{10}+\rho)t\,, & p_1(t) &= p_{10} + \zeta t\,,\\ q^2(t) &= q_0^1 + \zeta + (p_{10}+\rho)t\,, & p_2(t) &= \rho - \zeta t\,. \end{aligned} \tag{43}$$

Using the standard choice (22) for the multipliers we obtain for the extended Hamiltonian $\bar H_e = \frac{1}{2}(p_1 + p_2)^2 + V(q^1 - q^2) - \big[V'(q^1 - q^2)\big]^2/V''(q^1 - q^2)$. In the quadratic case the general solution is the same as for the total Hamiltonian $\bar H_t$; only the signs of the terms ζt in the momenta changes. Hence, the constraint manifold is now linearly unstable for the formulation based on the extended Hamiltonian.

7. Regular Systems with Imposed Constraints

For applications the most important case of a constrained system is described by a regular Lagrangian L_0 and subject to k externally imposed holonomic constraints $\phi_\alpha(q) = 0$. In principle, this situation *cannot* be treated within the Dirac formalism, as it covers only singular Lagrangians. Therefore one introduces Lagrange multipliers λ^α and considers the Lagrangian $L = L_0 + \lambda^\alpha\phi_\alpha$. In contrast to the multipliers u in the Dirac theory, the λ must be considered as additional dynamical variables and not as undetermined functions. Now L is obviously singular, as it does not depend on the "velocities" $\dot{\lambda}$.

For the Hamiltonian formalism we must introduce canonically conjugate momenta π_α for the λ^α. The primary constraints are simply given by $\pi = 0$. If we denote by H_0 the Hamiltonian for the regular system, the canonical Hamiltonian of the constrained system is $H_c = H_0 - \lambda^\alpha \phi_\alpha$; the total one is $H_t = H_c + u^\alpha \pi_\alpha$. The Dirac algorithm yields the secondary constraints $\phi_\alpha = 0$ and the tertiary constraints $\psi_\alpha = \{\phi_\alpha, H_0\} = 0$. The next step determines λ

$$\{\psi_\alpha, H_0\} - \lambda^\beta \{\psi_\alpha, \phi_\beta\} = 0\,. \tag{44}$$

The fifth and last step yields $u = 0$.

This rather long derivation can be shortened by *not* introducing the total Hamiltonian H_t and the momenta π. Starting with H_c and imposing $\phi = 0$ as primary constraints leads to equivalent results, as in the end $\pi = u = 0$. The standard approach is to take the Hamiltonian equations of motion for H_c and augment them by the constraints to get the following differential algebraic equation (see (7))

$$\dot{q} = \frac{\partial H_0}{\partial p}\,, \qquad \dot{p} = -\frac{\partial H_0}{\partial q} + \lambda^\alpha \frac{\partial \phi_\alpha}{\partial q}\,, \qquad \phi_\alpha = 0\,. \tag{45}$$

By differentiating twice the last equation in (45) one can derive exactly the same equation (44) for λ as in the Dirac theory. With $Q_{\alpha\beta} = \{\phi_\alpha, \psi_\beta\}$ it has the solution

$$\lambda^\alpha \approx (Q^{-1})^{\alpha\beta}\{\psi_\beta, H_0\}\,. \tag{46}$$

The main problem in using Dirac brackets is the inversion of the matrix C of the Poisson brackets of the constraint functions. For a larger number K of constraints one can no longer do this symbolically. Thus one must numerically invert a $K \times K$ matrix at each evaluation of the equations of motion. In our special case $K = 2k$ and C can be partitioned into four $k \times k$ sub-matrices

$$C = \begin{pmatrix} 0 & Q \\ -Q^t & S \end{pmatrix} \tag{47}$$

where Q is as above and $S_{\alpha\beta} = \{\psi_\alpha, \psi_\beta\}$. The inversion of such a matrix can be reduced to the inversion of one $k \times k$ matrix plus two matrix multiplications, as

$$C^{-1} = \begin{pmatrix} Q^{-t}SQ^{-1} & -Q^{-t} \\ Q^{-1} & 0 \end{pmatrix}\,. \tag{48}$$

The Hamilton-Dirac equations take now the following form

$$\begin{aligned} \dot{q} =& \quad \frac{\partial H_0}{\partial p} - (Q^{-1})^{\alpha\beta}\frac{\partial \psi_\alpha}{\partial p}\psi_\beta\,, \\ \dot{p} =& -\frac{\partial H_0}{\partial q} - (Q^{-1})^{\alpha\beta}\frac{\partial \phi_\alpha}{\partial q}\{\psi_\beta, H_0\} \\ &\qquad + \left[(Q^{-t}SQ^{-1})^{\alpha\beta}\frac{\partial \phi_\alpha}{\partial q} + (Q^{-1})^{\alpha\beta}\frac{\partial \psi_\alpha}{\partial q}\right]\psi_\beta\,. \end{aligned} \tag{49}$$

Taking (46) into account we see that they differ from (45) only by some terms multiplied by ψ. Thus both formulations are weakly equal. Note that the position constraint functions ϕ do not appear explicitly!

We cannot apply here the results of Ascher e. a. [3] on the stabilization of differential algebraic equations. They subtract the constraint functions multiplied by some matrix from an underlying differential equation. If the product of this matrix and the Jacobian of the constraints is positive definite, the constraint manifold is asymptotically stable for the obtained flow. The Hamilton-Dirac equations can be understood within this general scheme, but the result does not hold, as the matrix product is not positive definite. We got weaker stability results. But this was to be expected, as we still have a Hamiltonian system, whereas the approach of Ascher e. a. destroys this property.

For the extended Hamiltonian we make the following ansatz

$$H_e = H_0 - \lambda^\alpha \phi_\alpha + \mu^\alpha \psi_\alpha \,. \tag{50}$$

For λ we recover the result (46); for μ we obtain

$$\mu^\alpha = (Q^{-1})^{\alpha\beta}\psi_\beta \,. \tag{51}$$

Thus μ vanishes weakly und could be taken as zero. But then $H_e = H_t$ and we get as equations of motion the classical ones (45) plus terms $(\partial\lambda^\alpha/\partial q)\phi_\alpha$ and $(\partial\lambda^\alpha/\partial p)\phi_\alpha$, respectively (compare (7) and (8)).

The multipliers λ depend on derivatives of the momentum constraint functions ψ. Since they occur in the extended Hamiltonian H_e, we need three differentiations of the original constraint functions ϕ to set up the equations of motion opposed to the Hamilton-Dirac equations where two differentiations suffice.

8. Numerical Example I: The Pendulum

A classical example of a constrained system is the planar pendulum in Cartesian coordinates. For simplicity, all constants like length, mass, etc. are set to 1. The Lagrangian of the underlying regular system is $L_0 = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) - y$. We add the constraint function $\phi = \frac{1}{2}(x^2 + y^2 - 1)$ with a multiplier λ to get the Lagrangian for the pendulum $L = L_0 + \frac{1}{2}\lambda(x^2 + y^2 - 1)$. The canonically conjugate momenta are just the velocities: $p_x = \dot{x}$, $p_y = \dot{y}$. Checking whether the evolution generated by the Hamiltonian $H_t = \frac{1}{2}(p_x^2 + p_y^2) + y - \frac{1}{2}\lambda(x^2 + y^2 - 1)$ preserves the primary constraint $\phi = 0$ yields a secondary constraint $\psi = \{\phi, H_t\} = xp_x + yp_y = 0$. Then the Dirac algorithm stops, as the next step only determines λ. Since $\{\phi,\psi\} = x^2+y^2$ the Dirac bracket takes the form

$$\{F,G\}^* = \{F,G\} + \frac{1}{x^2+y^2}\Big(\{F,\phi\}\{\psi,G\} - \{F,\psi\}\{\phi,G\}\Big)\,. \tag{52}$$

By taking Dirac brackets with the Hamiltonian $H_0 = (p_x^2 + p_y^2)/2 + y$ of the underlying regular system we finally arrive at the Hamilton-Dirac equations,

$$\dot{x} = p_x + \mu x\,, \quad \dot{y} = p_y + \mu y\,, \quad \dot{p}_x = \lambda x - \mu p_x\,, \quad \dot{p}_y = \lambda y - \mu p_y - 1\,, \tag{53}$$

where λ, μ are given by

$$\lambda = -\frac{p_x^2 + p_y^2 - y}{x^2 + y^2}\,, \qquad \mu = -\frac{xp_x + yp_y}{x^2+y^2}\,. \tag{54}$$

We compare this formulation with the one based on the extended Hamiltonian $H_e = H_0 - \lambda\phi + \mu\psi$ with the multipliers λ, μ again given by (54). This corresponds

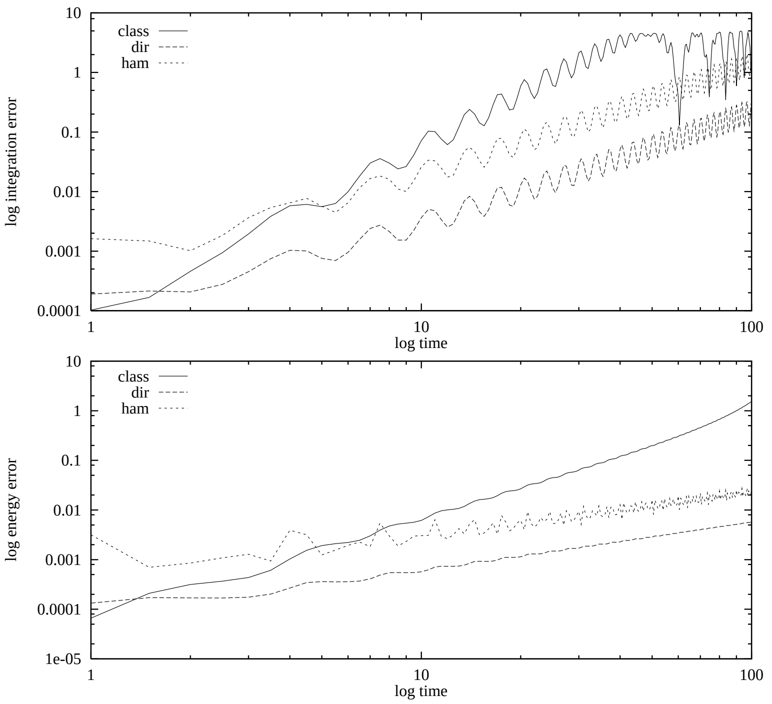

FIGURE 1. Integration and energy error (pendulum)

to the choice (22) for the solution of (21). The equations of motion are

$$\begin{gathered} \dot{x} = \left(2 - \frac{1}{x^2+y^2}\right) p_x + 2\mu x\,, \quad \dot{y} = \left(2 - \frac{1}{x^2+y^2}\right) p_y + 2\mu y\,, \\ \dot{p}_x = \lambda \frac{x}{x^2+y^2} - 2\mu p_x - 2\mu^2 x\,, \\ \dot{p}_y = \lambda \frac{y}{x^2+y^2} - 2\mu p_y - 2\mu^2 y - \frac{1}{2}\left(1 + \frac{1}{x^2+y^2}\right). \end{gathered} \tag{55}$$

They differ from (53) only by linear combinations of constraint functions. As expected, they are more complicated and thus more expensive to evaluate.

In order to show the necessity of stabilizing the constraint manifold we compare these two formulations with the classical equations of motion (45)

$$\dot{x} = p_x\,, \quad \dot{y} = p_y\,, \quad \dot{p}_x = \lambda x\,, \quad \dot{p}_y = \lambda y - 1 \tag{56}$$

with λ again given by (54).

We integrated numerically all three formulations with the classical fourth order Runge-Kutta method for the following initial data: $x^0 = 1$, $y^0 = 0$, $p_x^0 = 0$, $p_y^0 = -2$. For these values the pendulum rotates clockwise with a period of $T \approx 3.31$. We integrated over the interval $t \in [0, 100]$, i. e. roughly over 30 periods, with a constant step size of $h = 0.1 \approx T/33$. Fig. 1 contains logarithmic plots of the integration and the energy error; Fig. 2 shows the position and momentum constraint residuals. `class`, `dir` and `ham` label the curves for the classical equations of motion (56), for the Hamilton-Dirac equations (53) and for the equations (55) derived with the extended

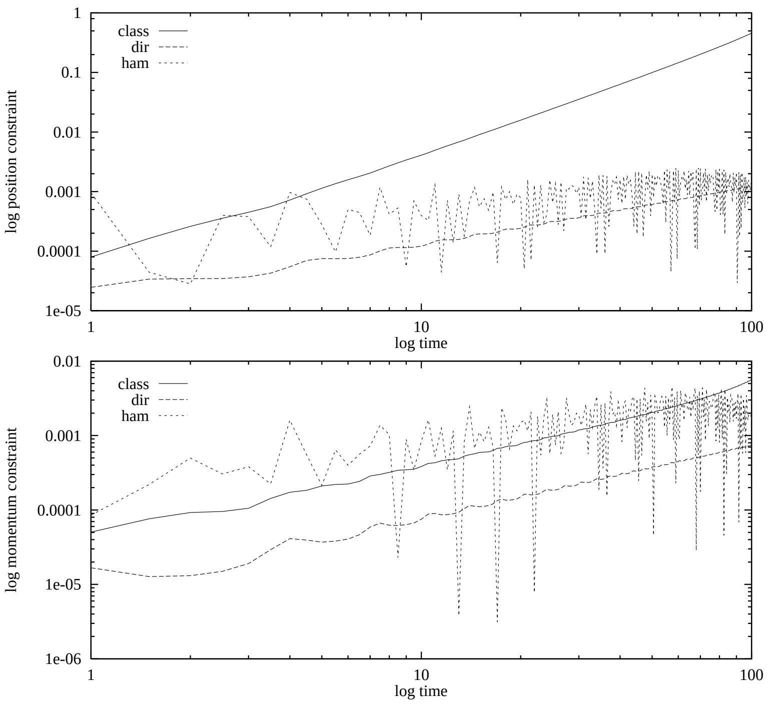

FIGURE 2. Position and momentum constraint residual (pendulum)

Hamiltonian, respectively. The integration error was estimated by comparing with the solution of the state space form $\ddot{\varphi} = -\sin\varphi$ computed with $h/10$.

Since the amplitude of our pendulum is 1 and its maximal momentum about 2.45, the computed values can surely be considered as useless, if the integration error exceeds 1. Thus the Hamilton-Dirac equations are the only formulation where the numerical integration does not clearly break down before the end of the considered interval. With the extended Hamiltonian one obtains reasonable results until approximately $t = 70$; with the classical formulation until about $t = 30$.

The stabilizing effect of the extended Hamiltonian and of the Dirac bracket, respectively, shows not only in the lower absolute values of the errors but also in their growth. Both formulations show a quadratic growth of the integration error and a linear growth of the energy error. Taking only the time into account where the classical formulation yields reasonable results, its integration error grows cubically and its energy error quadratically. For the Hamilton-Dirac equations the constraint residuals grow linearly, for the extended Hamiltonian even less. In the classical formulation the position constraint residual shows a quadratic growth, whereas the momentum constraint residual behaves also linearly.

As already mentioned in Section 5 in the case of the pendulum with our choice for the multipliers λ, μ in the extended Hamiltonian the origin is a center of (28). This can be observed in the figures, as the curves are always very "shaky" for this formulation. The different stability properties of the various equations of motion can also be clearly seen, if one chooses inconsistent initial data. The Hamilton-Dirac equations do not notice the inconsistency but produce a solution with the same

	$y(2T)$	$y(4T)$	$y(10T)$	$y(20T)$	$y(30T)$
`class`	$9.66 \cdot 10^{-3}$	$7.28 \cdot 10^{-2}$	$9.22 \cdot 10^{-1}$	$2.16 \cdot 10^{-1}$	$3.05 \cdot 10^{-1}$
`ham`	$5.75 \cdot 10^{-3}$	$1.77 \cdot 10^{-2}$	$8.98 \cdot 10^{-2}$	$3.17 \cdot 10^{-1}$	$6.21 \cdot 10^{-1}$
`dir`	$8.28 \cdot 10^{-4}$	$2.63 \cdot 10^{-3}$	$1.39 \cdot 10^{-2}$	$5.18 \cdot 10^{-2}$	$1.13 \cdot 10^{-1}$

TABLE 1. Phase error (pendulum)

behavior as in the figures, however now with respect to the manifold $\phi(x, y) = \zeta$ and $\psi(x, y, p_x, p_y) = \rho$ defined by the initial data. In contrast, the solution of the equations of motion for the extended Hamiltonian tries to reach the true constraint manifold but effectively oscillates about it.

Another important aspect is how much of the periodicity of the solution is maintained during the numerical integration. Table 1 contains the numerical values of y after several revolutions. The correct value would be zero for our initial data. Phase portraits of the numerical solutions (not shown here) also clearly demonstrate that the classical formulation leads only for a rather short time to an acceptable approximation of the true solution.

In a comparison one must also take the computational costs into account. Using the Hamilton-Dirac equations requires only about 5% more computing time than the classical formulation, whereas the extended Hamiltonian needs almost 65% more time. The difference in computational efficiency becomes even larger with a variable step size. Using a fifth order Runge-Kutta-Fehlberg method the integration of the equations of motion derived with the extended Hamiltonian needs between 50% and 100% more evaluations of the equations for the same prescribed precision than the Dirac bracket approach.

9. NUMERICAL EXAMPLE II: A CHAIN MOLECULE

As a larger example we consider a problem in molecular dynamics already used by Leimkuhler and Skeel [16] in the context of constrained dynamics. It consists of a planar chain molecule with $N = 7$ atoms. The bonds between them are assumed to have a fixed length. This condition yields the constraints. The interaction of the atoms is described by a Lennard-Jones potential

$$V = 0.1 \sum_{j>i} (r_{ij}^{-12} - 2r_{ij}^{-6}) \tag{57}$$

where r_{ij} denotes the distance between atom i and atom j.

One global energy minima of the molecule is the hexagonal structure shown in the left part of Fig. 3. We took this as initial configuration in our computations. At the ends of the chain we started with initial velocities of equal amplitude ($v_0 = 0.25$) but opposite direction; the remaining atoms are initially at rest. The emerging dynamics can be split into a rigid body rotation of the whole chain and small vibrations of each atom around its equilibrium position. The right part of Fig. 3 shows the motion of an end atom of the chain.

Integration methods for differential algebraic systems are often based on backward differentiation formulae. As Leimkuhler and Skeel [16] reported, this approach leads to physically unacceptable solutions. Such methods were originally developed

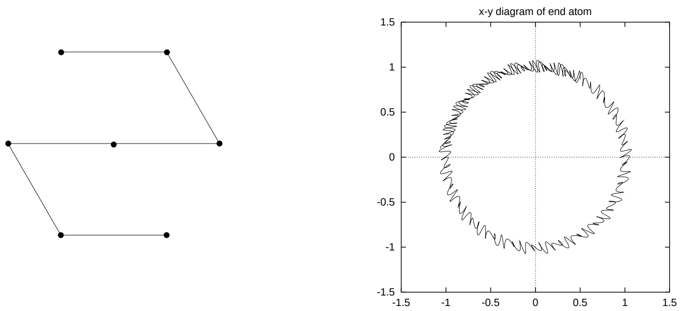

FIGURE 3. Global energy minimum and motion of end atom

for stiff systems. After a short time they completely eliminate the vibrational degrees of freedom of the system and yield a pure rigid body rotation. This implies a significant violation of energy conservation.

We have chosen this model in order to demonstrate that the Dirac bracket approach can be reasonably applied even for larger systems. Actually in this example it is still easily possible to perform all necessary calculations by hand based on our results in Section 7. We did not try to do this for the method of the extended Hamiltonian, as it would lead to very complex equations of motion.

If we denote the coordinates of atom i by (x^i, y^i) and its momenta by p_x^i, p_y^i, the underlying regular Hamiltonian is

$$H_0 = \frac{1}{2}\sum_{i=1}^{N}\Big[(p_x^i)^2 + (p_y^i)^2\Big] + V(x,y) \tag{58}$$

with V given by (57) and $r_{ij}^2 = (x^i - x^j)^2 + (y^i - y^j)^2$. The constraints are

$$\begin{aligned} \phi_\alpha &= \frac{1}{2}\big[(\Delta x^\alpha)^2 + (\Delta y^\alpha)^2 - L^2\big] = 0\,, \\ \psi_\alpha &= \Delta x^\alpha \Delta p_x^\alpha + \Delta y^\alpha \Delta p_y^\alpha = 0\,, \end{aligned} \qquad \alpha = 1, \dots, N-1\,, \tag{59}$$

where L stands for the length of the bonds and where we have introduced the short hand $\Delta x^\alpha = x^\alpha - x^{\alpha+1}$ and so on.

Computing the entries of the matrices Q, S defined in Section 7 we obtain for Q

$$\begin{aligned} \{\phi_\alpha, \psi_\beta\} = 2\delta_{\alpha\beta}\big[(\Delta x^\alpha)^2 + (\Delta y^\alpha)^2\big] - {}& \\ \delta_{\alpha+1,\beta}\big[\Delta x^\alpha \Delta x^\beta + \Delta y^\alpha \Delta y^\beta\big] - {}& \\ \delta_{\alpha,\beta+1}\big[\Delta x^\beta \Delta x^\alpha + \Delta y^\beta \Delta y^\alpha\big]& \end{aligned} \tag{60}$$

and for S, respectively,

$$\begin{aligned} \{\psi_\alpha, \psi_\beta\} = \delta_{\alpha+1,\beta}\big[\Delta x^\alpha \Delta p_x^\beta - \Delta x^\beta \Delta p_x^\alpha + \Delta y^\alpha \Delta p_y^\beta - \Delta y^\beta \Delta p_y^\alpha\big] - {}& \\ \delta_{\alpha,\beta+1}\big[\Delta x^\beta \Delta p_x^\alpha - \Delta x^\alpha \Delta p_x^\beta + \Delta y^\beta \Delta P_y^\alpha - \Delta y^\alpha \Delta p_y^\beta\big]\,.& \end{aligned} \tag{61}$$

Both matrices are tridiagonal, as we have a chain structure or "nearest neighbors constraints:" ϕ_α, ψ_α involve only data of the atoms α and $\alpha + 1$. The inversion of such matrices has a linear complexity and can thus be done very fast.

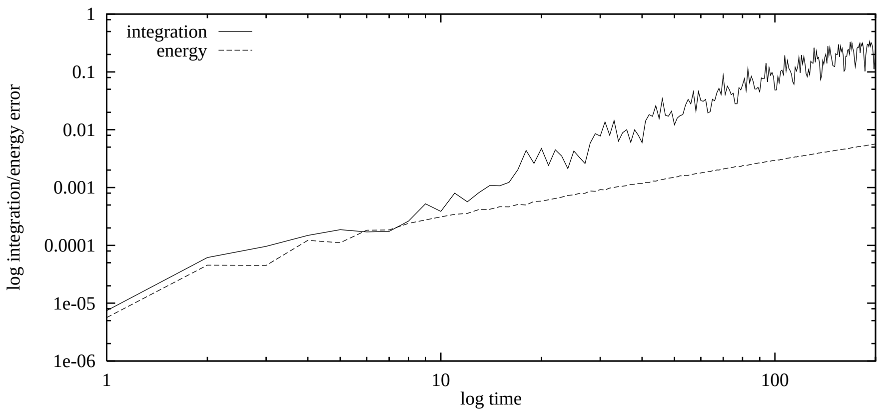

FIGURE 4. Integration and energy error (chain molecule)

The Poisson brackets of the coordinates with the constraint functions are

$$\{x^i, \phi_\alpha\} = 0\,, \qquad \{p_x^i, \psi_\alpha\} = (\delta^i_{\alpha+1} - \delta^i_\alpha)\Delta p_x^\alpha\,, \\ \{p_x^i, \phi_\alpha\} = -\{x^i, \psi_\alpha\} = (\delta^i_{\alpha+1} - \delta^i_\alpha)\Delta x^\alpha \tag{62}$$

and corresponding expressions for y, p_y. Finally, we calculate

$$\{\psi_\alpha, H_0\} = (\Delta p_x^\alpha)^2 + (\Delta p_y^\alpha)^2 - \\ \Delta x^\alpha \left(\frac{\partial V}{\partial x^\alpha} - \frac{\partial V}{\partial x^{\alpha+1}}\right) - \Delta y^\alpha \left(\frac{\partial V}{\partial y^\alpha} - \frac{\partial V}{\partial y^{\alpha+1}}\right)\,. \tag{63}$$

(60–63) contain all expressions needed to set up the equations of motion (49). The evaluation of the potential and the two matrix multiplications have a complexity quadratic in the number N of atoms. All other operations are linear in N. Thus the Dirac bracket could be applied without problems to much larger molecules.

We integrated the system for the initial conditions described above with the classical fourth order Runge-Kutta method for the interval $t \in [0, 200]$. As one can see from the right part of Fig. 3 this corresponds roughly to 5/4 periods of the rigid body rotation of the molecule. Figs. 4 and 5 show the results for the constant step size $h = 0.1$. As for the Hamilton-Dirac formulation of the pendulum, the integration error grows quadratically, all others errors about linearly.

Comparing with Leimkuhler and Skeel [16], we find that at least regarding energy conservation their approach using the RATTLE algorithm [1] leads to better results. Their energy error remains more or less constant over the full integration interval $t \in [0, 200]$. The explanation is simple: RATTLE is a *symplectic integrator* [24].

It is well-known that such methods often perform superior in long time integrations, especially with respect to energy conservation. Since almost all known symplectic integrators preserve only the canonical Poisson bracket, it may appear that they are not applicable in the case of a modified bracket structure like the Dirac bracket used in the Hamilton-Dirac equations.

We hope to discuss this problem in more detail in the future, but we want to comment briefly on some preliminary experiments with a *canonical* symplectic integrator: the implicit midpoint rule. The arising nonlinear equations were solved with a simple functional iteration to a tolerance of 10^{-5}.

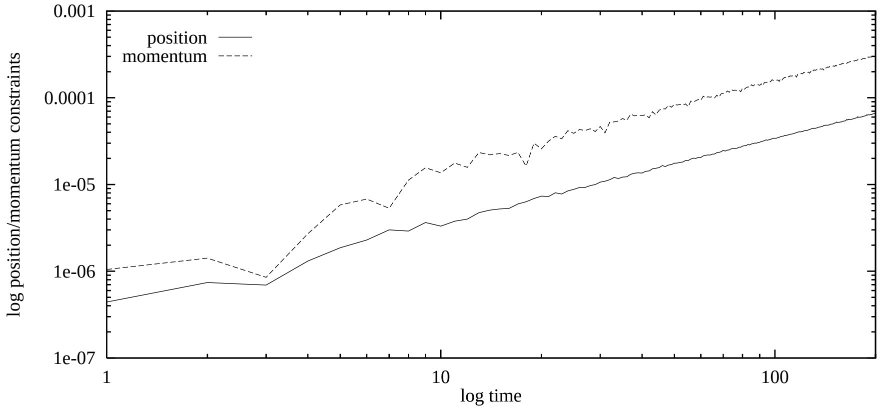

FIGURE 5. Position and momentum constraint residuals (chain molecule)

Although the implicit midpoint rule is only second order opposed to the fourth order scheme used so far, it conserved the energy for the same step size better by almost an order of magnitude. The error growth is less than linear; the constraints residuals improve by more than an order of magnitude. If the step size is halved, the energy error becomes smaller by more than an order of magnitude and remains almost constant about 10^{-4} over the full integration interval.

A partial explanation might be given as follows. On the constraint manifold the Dirac bracket represents the symplectic structure induced by the canonical Poisson bracket [13, 30]. For small constraint residuals a canonical symplectic integrator thus defines in good approximation a symplectic mapping for the Dirac bracket, too. But for other methods for the symplectic integration of constrained systems the situation does not differ much, as they require the solution of nonlinear equations [14, 22]. In a numerical computation they are also only approximations of symplectic mappings, if these equations are not solved exactly.

The implicit midpoint rule preserves quadratic first integrals, if the arising nonlinear equations are solved exactly [6]. In our example the constraints and the energy are defined by quadratic functions. The constraints are first integrals for the Hamilton-Dirac equations. Thus it is not surprising that we find small errors.

We may expect that the higher the precision with which the nonlinear equations are solved, the more the implicit midpoint rule behaves like a true symplectic integrator for the Dirac bracket. First numerical tests seem to confirm this conjecture. In the case of the pendulum one observes for example much smaller phase errors compared with Table 1.

10. CONCLUSION

There exist two basic strategies for dealing numerically with differential algebraic equations. One can modify the equations; this leads to stabilization and index reduction techniques. Or one designs special numerical schemes like projection methods. Obviously, these two strategies are complementary and can be combined. We studied in this article the first approach for the special case of constrained Hamiltonian systems.

There have been attempts to stabilize general differential algebraic equations [3]. However, no systematic solution with a solid theoretical foundation has emerged so far. A classical example for the problems encountered is the Baumgarte stabilization [4] where the choice of the parameters is to a large extent still a question of try and error. In physical problems like Hamiltonian mechanics the differential equations possess special properties. In this article we exploited the symplectic structure of the phase space to derive stable equations of motion.

The constraint manifold is always stable for the Hamilton-Dirac equations. In contrast its stability for the classical formulation or for the one based on the extended Hamiltonian depends on choices made for the canonical Hamiltonian or the multipliers. This was demonstrated by the toy model in Section 6. The neutral stability obtained with the Dirac bracket is in some sense the optimum one can achieve in a *Hamiltonian* formulation of the equations of motion. If the constraint manifold was asymptotically stable, the theorem of Liouville on the conservation of phase space volume [2] would imply an instability within the constraint manifold.

The Faddeev-Jackiw formalism [25], a first order approach to constrained dynamics, uses an extended phase space and modifies the symplectic structure there. Although the modified structure coincides with the Dirac bracket on the original phase space, the arising equations of motion differ from the Hamilton-Dirac equations. But in numerical experiments both approaches lead to almost identical results even for long integration times. This clearly indicates that the physical properties of the Dirac bracket are the cause of the observed stability.

For systems with a large number of constraints the efficiency of the Dirac bracket approach depends crucially on the matrix C which must be inverted at each evaluation of the Hamilton-Dirac equations. As the example of the chain molecule demonstrated, this inversion can be significantly simplified by exploiting special constraint structures like “nearest neighbors constraints.” Note that for the Hamilton-Dirac equations it suffices to invert numerically, whereas the extended Hamiltonian approach needs in addition derivatives of C^{-1} to set up the equations of motion.

For regular systems with imposed constraints there exists a comparatively cheap way to exploit much of the stability of the Hamilton-Dirac equations without really using them [26]. We noted already in Section 7 that for this class of systems the classical Hamiltonian equations of motion (45) and the Hamilton-Dirac equations (49) differ only by terms multiplied by the momentum constraint functions ψ. If we perform at each step a *momentum projection* in order to exactly maintain these constraints, the two formulations are equivalent and we can ignore the additional terms which are rather expensive to evaluate. For most systems occurring in applications the constraint functions ψ are linear in the momenta p. Thus the projection requires only the solution of a linear system and is rather cheap.

The constraint functions become first integrals for the Hamilton-Dirac equations; for the equations derived with the extended Hamiltonian they represent only weak invariants. For a higher degree of constraint preservation one may thus use a special scheme for maintaining invariants. Moan [19] constructed recently explicit Runge-Kutta methods preserving quadratic first integrals. Among them is a second order method with three stages that applied to the chain molecule yields with less evaluations better results than the classical fourth order method. For general constraint functions one could use the discrete gradient methods proposed by Quispel e. a. [20, 21], although their construction appears rather expensive.

Finally, we note that Dirac brackets can be generalized to infinite-dimensional systems. Thus this approach could also be useful for problems in electrodynamics, continuum mechanics etc. like the impetus-striction formalism [7, 18]. Salmon [23] showed for example that the semi-geostrophic equations for a rotating fluid often used in meteorology or oceanography can be derived using Dirac brackets.

Acknowledgments

I am indebted to R. Coleman for valuable hints on Hamiltonian stability theory. The comments of an anonymous referee lead to an improved presentation of the material.

References

[1] H.C. Andersen, *Rattle — a velocity version of the Shake algorithm for molecular dynamics calculations*, J. Comp. Phys. **52** (1983), 24–34.
[2] V.I. Arnold, *Mathematical methods of classical mechanics*, Graduate Texts in Mathematics 60, Springer-Verlag, New York, 1978.
[3] U.M. Ascher, H. Chin, and S. Reich, *Stabilization of DAEs and invariant manifolds*, Num. Math. **67** (1994), 131–149.
[4] J. Baumgarte, *Stabilization of constraints and integrals of motion in dynamical systems*, Comp. Meth. Appl. Mech. Eng. **1** (1972), 1–16.
[5] K.E. Brenan, S.L. Campbell, and L.R. Petzold, *Numerical solution of initial-value problems in differential-algebraic equations*, Classics in Applied Mathematics 14, SIAM, Philadelphia, 1996.
[6] G.J. Cooper, *Stability of Runge-Kutta methods for trajectory problems*, IMA J. Num. Anal. **7** (1987), 1–13.
[7] D.J. Dichmann, J.H. Maddocks, and R.L. Pego, *Hamiltonian dynamics of an elastica and the stability of solitary waves*, Arch. Rat. Mech. Anal. **135** (1996), 357–396.
[8] P.A.M. Dirac, *Generalized Hamiltonian dynamics*, Can. J. Math. **2** (1950), 129–148.
[9] ———, *Generalized Hamiltonian dynamics*, Proc. Roy. Soc. A **246** (1958), 326–332.
[10] G.V. Dunne, R. Jackiw, and C.A. Trugenberger, *Topological (Chern-Simons) quantum mechanics*, Phys. Rev. D **41** (1990), 661–666.
[11] H. Goldstein, *Classical mechanics*, Addison-Wesley, New York, 1980.
[12] W. Hahn, *Stability of motion*, Grundlehren der mathematischen Wissenschaften 138, Springer-Verlag, Berlin, 1967.
[13] M. Henneaux and C. Teitelboim, *Quantization of gauge systems*, Princeton University Press, 1992.
[14] L.O. Jay, *Symplectic partitioned Runge-Kutta methods for constrained Hamiltonian systems*, SIAM J. Numer. Anal. **33** (1996), 368–387.
[15] B.J. Leimkuhler and S. Reich, *Symplectic integration of constrained Hamiltonian systems*, Math. Comp. **63** (1994), 589–605.
[16] B.J. Leimkuhler and R.D. Skeel, *Symplectic numerical integrators in constrained Hamiltonian systems*, J. Comp. Phys. **112** (1994), 117–125.
[17] R.S. MacKay, *Stability of equilibria of Hamiltonian systems*, Nonlinear Phenomena and Chaos (S. Sarkar, ed.), Hilger, Bristol, 1986, pp. 254–270.
[18] J.H. Maddocks and R.L. Pego, *An unconstrained Hamiltonian formulation for incompressible fluid flow*, Comm. Math. Phys. **170** (1995), 207–217.
[19] P.C. Moan, *The numerical solution of ordinary differential equations with conservation laws*, Master's thesis, Norwegian Institute of Technology, Trondheim, 1996.
[20] G.R.W. Quispel and H.W. Capel, *Solving ODEs numerically while preserving all first integrals*, Preprint, La Trobe University, Melbourne, 1997.
[21] G.R.W. Quispel and G.S. Turner, *Discrete gradient methods for solving ODEs numerically while preserving a first integral*, J. Phys. A **29** (1996), L341–L349.
[22] S. Reich, *Symplectic integration of constrained Hamiltonian systems by composition methods*, SIAM J. Num. Anal. **33** (1996), 475–491.
[23] R. Salmon, *Semigeostrophic theory as a Dirac-bracket projection*, J. Fluid Mech. **196** (1988), 345–358.

[24] J.M. Sanz-Serna and M.P. Calvo, *Numerical Hamiltonian problems*, Applied Mathematics and Mathematical Computation 7, Chapman & Hall, London, 1994.

[25] W.M. Seiler, *Involution and constrained dynamics II: The Faddeev-Jackiw approach*, J. Phys. A **28** (1995), 7315–7331.

[26] ———, *Momentum versus position projections for constrained Hamiltonian systems*, Preprint, Universität Karlsruhe, 1997.

[27] ———, *Numerical analysis of constrained Hamiltonian systems and the formal theory of differential equations*, Math. Comp. Simul. **to appear** (1997).

[28] W.M. Seiler and R.W. Tucker, *Involution and constrained dynamics I: The Dirac approach*, J. Phys. A **28** (1995), 4431–4451.

[29] C.L. Siegel and J.K. Moser, *Lectures on celestial mechanics*, Grundlehren der mathematischen Wissenschaften 187, Springer-Verlag, Berlin, 1971.

[30] J. Śniatycki, *Dirac brackets in geometric dynamics*, Ann. Inst. Henri Poincaré A **20** (1974), 365–372.

[31] K. Sundermeyer, *Constrained dynamics*, Lecture Notes in Physics 169, Springer-Verlag, New York, 1982.

Institut für Algorithmen und Kognitive Systeme, Universität Karlsruhe, D-76128 Karlsruhe, Germany

E-mail address: `werner.seiler@informatik.uni-karlsruhe.de`