Larry M. Bates; James M. Nester
On D'Alembert's Principle

Communications in Mathematics, Vol. 19 (2011), No. 1, 57--72

Persistent URL: http://dml.cz/dmlcz/141679

Communications in Mathematics 19 (2011) 57–72

On D'Alembert's Principle

Larry M. Bates, James M. Nester

Abstract. A formulation of the D'Alembert principle as the orthogonal projection of the acceleration onto an affine plane determined by nonlinear nonholonomic constraints is given. Consequences of this formulation for the equations of motion are discussed in the context of several examples, together with the attendant singular reduction theory.

1 D'Alembert's principle

Let us suppose that we have a Lagrangian or Hamiltonian mechanical system and we wish to impose a constraint. The system has an n-dimensional configuration space Q with local coordinates $\{q^a\}$, velocity phase space TQ with the natural chart $\{q^a, v^a\}$, and momentum phase space $P = T^*Q$ with local coordinates (q^a, p_a). To start, suppose we have a Lagrangian of the classical form kinetic energy minus potential energy,

$$l = \frac{1}{2}g(v, v) - u(q).$$

We want to write down the equations of motion if we impose a (possibly time dependent and nonholonomic) constraint of the form

$$c(q, v, t) = 0.$$

Later on we will discuss what happens if we have Lagrangians not of this simple form, or more constraints, but for now it suffices to just to consider this case.

Differentiating the constraint with respect to the time t gives

$$\frac{d}{dt}c = c_a v^a + c_{\dot a}\dot v^a + c_t = 0$$

where $c_a = \partial c/\partial q^a$, $c_{\dot a} = \partial c/\partial v^a$, and $c_t = \partial c/\partial t$. The acceleration a is given by

$$a^b = \dot v^b + \Gamma^b_{kl} v^k v^l$$

2010 MSC: 70H33, 70H45, 37J60, 70F25
Key words: Nonholonomic constraints, d'Alembert's principle.

where the Γ^b_{kl} are the Christoffel symbols of the Levi-Civita connection of the metric. Substituting the acceleration into the differentiated constraint yields

$$c_{\dot{k}} a^k + (c_k v^k - c_{\dot{a}} \Gamma^a_{kl} v^k v^l + c_t) = 0.$$

The important point here is that for fixed (q, v, t) this equation represents an *affine relation* for the accelerations in the tangent space $T_q Q$, where we have used the connection to identify the vertical space $V_{(q,v)}TQ$ with $T_q Q$.

Given the acceleration a of the unconstrained problem, it must be modified so that it lies in the affine plane. This is done by subtracting the component $a_\perp$ of a orthogonal to the plane, so the acceleration of the constrained problem is the difference

$$a_{\text{constrained}} = a - a_\perp$$

and lies in the affine plane. Observe that the vector $a_\perp$ has components

$$a_\perp{}^k = -\lambda g^{kr} c_{\dot{r}}$$

for some real number λ because of the form of the affine equation. Since we already have a Lagrangian description of the unconstrained problem, the force covector for the constrained problem may be written in the form

$$\frac{d}{dt}\left(\frac{\partial l}{\partial v^a}\right) - \frac{\partial l}{\partial q^a} = \lambda c_{\dot{a}}.$$

Observe that the specific form of the Lagrangian was not essential, one can equally well work with the velocity Hessian $g_{ab} := l_{\dot{a}\dot{b}}$ as long as the Lagrangian is regular, i.e., the velocity Hessian defines an invertible metric. Furthermore, the argument generalizes to the case of more than one constraint function, say $c^1 = 0$, $\ldots, c^K = 0$, yielding the constrained Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial l}{\partial v^a}\right) - \frac{\partial l}{\partial q^a} = \lambda_A c^A_{\dot{a}}$$

involving the Lagrange multipliers $\lambda_1, \ldots, \lambda_K$.

2 Other formulations

Assuming the regularity of the Lagrangian, we may push everything over to the cotangent bundle and give a Hamiltonian description as well. This may be written in a coordinate free manner for the vector field

$$X = \dot{q}^a \partial_{q^a} + \dot{p}_b \partial^{p_b}$$

by using ϑ_0, the canonical one-form on T^*Q, the symplectic form $\omega = -d\vartheta_0$, the Legendre transform $\mathscr{L}$, as well as $\vartheta^A := Fc^{A^*}\vartheta_0$, where Fc^A is the fiber derivative of c^A. Set $\phi^A = \mathscr{L}_* \vartheta^A$. Then the constrained Hamilton's equations may be written in the form

$$X \mathbin{\lrcorner} \omega = dh + \lambda_A \phi^A$$

The time derivative of a function f (with Hamiltonian vector field X_f) along an integral curve of the constrained vector field X is

$$\frac{df}{dt} = \langle df, X \rangle = \omega(X_f, X) = -X_f \,\lrcorner\, (X \,\lrcorner\, \omega) = -X_f \,\lrcorner\, (dh + \lambda_A \phi^A).$$

Setting $\psi := dh + \lambda_A \phi^A$, we may write this derivative in Poisson bracket form as

$$\dot{f} = \{f, \psi\}$$

where we are taking the Poisson bracket of the function f and the one form ψ to be $\Lambda(df, \psi)$, where $\Lambda = \omega^{-1}$ is the structure tensor of the Poisson bracket.

3 Easy consequences

The following is a partial list of easy consequences of the nonlinear formulation so that one may see the similarities and differences with the affine theory. Some of these are known and may be found in earlier work by de Leon et al [5].

3.1 Conservation of energy

Suppose that we have a time independent Lagrangian and impose the constraint of constant energy. Then the Lagrange multiplier is zero, and the problem reduces to Hamilton's equations on a constant energy surface. We may view this as a consistency check for the nonlinear constraint theory.

3.2 Nonconservation of energy

Suppose that we have a time independent Lagrangian and impose a time independent constraint c. Then, letting the energy e be $e = pv - l = l_v v - l$ as usual, we find

$$\dot{e} = \left(\frac{d}{dt} \left(\frac{\partial l}{\partial v^a} \right) - \frac{\partial l}{\partial q^a} \right) v^a,$$

and so by the equation of motion for the constrained problem

$$\dot{e} = \lambda_A \frac{\partial c^A}{\partial v^a} v^a.$$

In general we do not expect this term to vanish, so unlike the case of linear nonholonomic constraints, we do not have energy conservation even in the time independent situation. Since this is not what one would expect from the usual Noether theory, it only goes to show that such problems are really not variational problems in the usual way, even though we have a Lagrangian. However, if the Lagrange multipliers are not zero, there is an important case where this term will vanish, and that is when the constraint functions are each homogeneous of some degree in the velocities. For then, by Euler's theorem on homogeneous functions,

$$\frac{\partial c^A}{\partial v^a} v^a \propto c^A = 0$$

by the constraint equation. Note that this is the case for linear constraints.

3.3 The Lagrange multiplier

Suppose we have just one constraint. The Lagrange multiplier λ is chosen so that the constrained vector field X is tangent to the constraint surface: $\langle dc, X\rangle = 0$. Since the constrained vector field X satisfies

$$X \lrcorner\, \omega = dh + \lambda\phi,$$

we have by symplectic inversion

$$\langle dc, X_h + \lambda\phi^{\#}\rangle = 0,$$

so that

$$\lambda = -\frac{\langle dc, X_h\rangle}{\langle dc, \phi^{\#}\rangle} = -\frac{\{c, h\}}{\langle dc, \phi^{\#}\rangle}$$

where X_h is the Hamiltonian vector field of h, and $\{c, h\}$ is the Poisson bracket of c and h. Note that the solvability of the Lagrange multiplier assumes the independence condition $dc \wedge \phi \neq 0$.

An immediate corollary is that if the constraint function is a first integral of the Hamiltonian, then imposing the integral as a constraint is really no constraint at all, since the multiplier is zero.

If we have multiple constraints, say $c^1, c^2, \ldots, c^K = 0$, then the equations of motion take the form

$$X \lrcorner\, \omega = dh + \lambda_A\phi^A.$$

The Lagrange multipliers λ_A may be found from the K equations

$$\langle dc^A, X_h\rangle + \lambda_B\langle dc^A, \phi^{B\#}\rangle = 0.$$

An evaluation of $\langle dc^A, \phi^{B\#}\rangle$ gives $-M^{AB} = -g^{ab}c^A_{\dot{a}}c^B_{\dot{b}}$, a contravariant metric on the subspace spanned by the one forms $c^A_{\dot{a}}$. As long as this metric is invertible one can uniquely find the λ_A. If g_{ab} has a Euclidean signature this will be the case as long as the one forms $c^A_{\dot{a}}$ are linearly independent. This is not sufficient, however, for Lorentz signature metrics, such as those appearing in our relativistic particle examples below.

4 Examples (1)

4.1 The brachystochrone

A good place to begin is with the brachystochrone. The brachystochrone is the problem where a bead slides down a frictionless wire from rest at $(x, y) = (0, 0)$ to the point $(x, y) = (a, b)$. Here we take the positive y direction to point vertically downwards. The problem is to determine the shape of the wire so as to have a minimum time of descent. For a falling body, from conservation of energy one obtains $v^2 = 2gy$. This can be viewed as a nonholomic constraint. In section 3.1 above we noted that such a constraint does no work—the associated constraint force vanishes. But we can exploit this nonholonomic constraint in another way. Using $dl^2 = dx^2 + dy^2$ we may write the time T of descent to be

$$T = \int dt = \int \frac{dl}{v} = \int \frac{1}{\sqrt{2gy}}\sqrt{\left(\frac{dx}{d\sigma}\right)^2 + \left(\frac{dy}{d\sigma}\right)^2}\, d\sigma.$$

where σ is any path parameter. The parameter invariance is connected with the fact that this Lagrangian is degenerate. Consequently a proper Hamiltonian analysis requires the Dirac algorithm. We will develop that shortly and include a description of the results it yields when applied to this problem. But first let us note that for this problem one can avoid the degeneracy of the Legendre transformation. Observe that our Lagrangian can be viewed as the arc length due to the metric

$$ds^2 = \frac{1}{2gy}(dx^2 + dy^2).$$

And so the objective is just to find a certain geodesic path of this metric. Now it well known that there is an "equivalent" alternative to the arc length Lagrangian for geodesics, namely its square:

$$L = \frac{1}{2gy}\left[\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2\right].$$

This will give the same path, but with a uniform speed parameterization. Moreover, for our problem it is apparent that we specifically want the unit speed geodesics of this metric. Then the parameter is actually the physical time. A virtue of this new Lagrangian is that it is non-degenerate, so there is no complication in passing to the Hamiltonian. The Hamiltonian equations of motion for the associated Hamiltonian $h = gy(p_x^2 + p_y^2)/2$ are (upon setting $g = 1$)

$$\dot{x} = yp_x, \qquad \dot{p}_x = 0, \qquad \dot{y} = yp_y, \qquad \dot{p}_y = -\frac{h}{y}.$$

The translational invariance of the metric in the x direction implies the conservation of p_x, so we may reduce at $\mu = p_x$ along with $h = 1$ to get the reduced equation

$$\dot{y} = yp_y = \sqrt{y - \mu^2 y^2}.$$

This separates and integrates to $s = \frac{1}{\mu}\arccos(2\mu^2 y - 1)$. Inverting yields $y = \frac{1}{2\mu^2}(1 - \cos\mu s)$, and hence $x = \frac{1}{2\mu}(s - \frac{1}{\mu}\sin\mu s)$, from which we recognize the familiar cycloidal solutions. In this example we have finessed the degeneracy of the Legendre transformation. Further on we will reconsider that issue. Other aspects of this problem are discussed in [6].

4.2 Dirac constraints

To deal with the proper dynamical formulation for relativistic particles, which also involves finding an appropriate parametrization (the proper time) we first sketch the Dirac Hamiltonian theory. In order to pass from the Lagrange equations together with the undetermined constraint forces $\lambda_A \partial c^A/\partial v^k$, to the Hamiltonian description, one uses the Legendre transformation. If the Lagrangian is not regular, so the momenta are not independent, then they satisfy some primary constraints $\Phi_\alpha(q,p) = 0$. Following Dirac, one includes these constraints in the Hamiltonian with Lagrange multipliers, so the total Hamiltonian takes the form $h = h_0 + u^\alpha \Phi_\alpha$.

The Hamiltonian evolution equations, including the velocity constraint forces, take the form

$$\begin{aligned}\frac{dq^k}{dt} &= \frac{\partial h_0}{\partial p_k} + u^\alpha \frac{\partial \Phi_\alpha}{\partial p_k},\\ \frac{dp_k}{dt} &= -\frac{\partial h_0}{\partial p_k} - u^\alpha \frac{\partial \Phi_\alpha}{\partial p_k} + \lambda_A \phi^A_k.\end{aligned}$$

These differential equations are to be considered along with the two sets of constraint equations

$$\Phi_\alpha = 0, \qquad c^A = 0.$$

The first possible obstruction is whether the constraint functions c^A can be chosen so that there exist one-forms on phase space ϕ^A_k which are related by the degenerate Legendre transform to $c^A_{\dot{k}}$. Since the dynamical equations are required to preserve the constraints, this leads to additional conditions which may yield new constraints or fix the multipliers. In general, in attempting to determine the unknown multipliers one can expect similar outcomes to the usual Dirac procedure:

1. there may be no solution,
2. there may be additional constraints,
3. the solution may not be unique.

Observe that the constraints $\Phi_\alpha = 0$ are well defined on phase space and their time derivatives are linear in the unknown multipliers. In this case we know from Dirac how to proceed. However, the constraints $c^A = 0$ are velocity constraints, they are not defined on phase space. In general there are no phase space functions which are related to c^A. Fortunately, from the Hamiltonian perspective there is a natural "inverse" to the Legendre transformation, given by the first half of the Hamiltonian evolution equations. Hence the velocity constraint function can be given in terms of the phase space variables as

$$c^A(q^k, v^k) = c^A\left(q^k, \frac{\partial h_0}{\partial p_k} + u^\alpha \frac{\partial \Phi_\alpha}{\partial p_k}\right) = 0.$$

If c^A is linear in v this expression will be linear in the multipliers, and there is no insurmountable difficulty. In the general case the velocity constraints could be *nonlinear* functions of the unknown multipliers. Moreover, preserving these constraints could lead to expressions involving the derivatives of the multipliers. This is the second obstruction.

In the relativistic examples below the velocity constraints are linear in an expression which is homogeneous of degree one in velocity, so they turn out to be linear in the multiplier, so there is no difficulty.

4.3 The relativistic particle

We apply the above procedure to the following relativistic particle Lagrangians with their associated *proper time-constant magnitude* velocity constraints (see for example Gràcia [8] or Krupková [10])

$$l_1 = -mc\sqrt{-g_{\mu\nu}v^\mu v^\nu} - V(x) + qv^\mu A_\mu, \qquad C_1 = c - \sqrt{-g_{\mu\nu}v^\mu v^\nu},$$
$$l_2 = \frac{1}{2}mg_{\mu\nu}v^\mu v^\nu - V(x) + qv^\mu A_\mu, \qquad C_2 = \frac{1}{2c}\big[g_{\mu\nu}v^\mu v^\nu = c^2\big].$$

Here $v^\mu = dx^\mu/d\sigma$, where σ is an *a priori* arbitrary time parameter. Just for this example we have changed our notation for the constraints in order to avoid any possible confusion with the speed of light c. In both cases the dynamical equation of motion can be rearranged to be in the form

$$\frac{d}{d\sigma}(mg_{\mu\nu}u^\nu) = qv^\nu(\partial_\mu A_\nu - \partial_\nu A_\mu) - \partial_\mu V + \lambda g_{\mu\nu}u^\nu,$$

with $u^\mu = dx^\mu/d\tau$ with τ being the proper time. Contracting the equation of motion with the 4-velocity gives the value of the multiplier: $-u^\mu\partial_\mu V - \lambda c^2 = 0$. The final form of the equation of motion is then

$$\frac{d}{d\tau}(mg_{\mu\nu}u^\nu) = qF_{\mu\nu}u^\nu - (\delta^\nu_\mu + c^{-2}u^\nu g_{\mu\gamma}u^\gamma)\partial_\nu V,$$

where, as usual, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the Maxwell field. The authors know of no textbook that gives a proper treatment of the relativistic velocity constraint. Typically, the equations of motion are obtained and then the constraint is imposed (see for example [7]) without any consideration of the force of constraint. This happens to work if the the dynamics is compatible with the constraint, as is the case of a charged particle interacting with the Maxwell field. Then the velocity constraint does no work, so the multiplier λ vanishes. This is not the case for the scalar potential, where the force must be Lorentz orthogonal to the 4-velocity. There does not seem to be any way to get the proper relativistic force due to a scalar potential directly from a Lagrangian. Nonholonomic constraints play an essential role.

It is also of interest to give the Hamiltonian formulation of these examples. For the Lagrangian l_1, the Legendre transformation is degenerate, and so there is a primary constraint. From

$$p_\mu = \frac{\partial l_1}{\partial v^\mu} = mc\frac{g_{\mu\nu}v^\nu}{\sqrt{-g_{\alpha\beta}v^\alpha v^\beta}} + qA_\mu,$$

set $\mathbf{p}_\mu := p_\mu - qA_\mu$ so the primary constraint has the form

$$g^{\mu\nu}\mathbf{p}_\mu\mathbf{p}_\nu = -(mc)^2.$$

The Hamiltonian $h_0 = p_\mu v^\mu - l_1 = V(x)$ and so

$$h = V + \frac{N}{2m}\big[g^{\mu\nu}\mathbf{p}_\mu\mathbf{p}_\nu + (mc)^2\big],$$

where the primary Dirac constraint Lagrange multiplier N is called the lapse. Hamilton's equations take the form

$$\frac{dx^\mu}{d\sigma} = \frac{N}{m} g^{\mu\nu} \mathbf{p}_\nu,$$
$$\frac{dp_\mu}{d\sigma} = -\partial_\mu V + \frac{N}{m} \mathbf{p}_\alpha g^{\alpha\nu} (q\partial_\mu A_\nu) + \lambda \frac{\mathbf{p}_\mu}{mc}.$$

The velocity constraint $0 = c - \sqrt{-\frac{N^2}{m^2}\mathbf{p}_\mu \mathbf{p}^\mu}$ determines that the Dirac multiplier $N = 1$. The preservation of the primary Dirac constraint determines the velocity constraint multiplier since $0 = -\mathbf{p}^\mu \partial_\mu V - \lambda \mathbf{p}_\mu \mathbf{p}^\mu/(mc)$ implies that $\lambda = (mc)^{-1}\mathbf{p}^\mu \partial_\mu V$.

4.4 The brachystochrone as a constrained system

Now we can briefly return to the brachystochrone as a constrained system. The Lagrangian has a general curve parameter σ and thus a gauge freedom type degeneracy, which leads to a primary constraint

$$2gy(p_x^2 + p_y^2) - 1 = 0.$$

The Lagrangian is homogeneous of degree one in the velocities. Consequently by Euler's theorem the energy function vanishes. Then the Dirac Hamiltonian is just given by a Lagrange multiplier multiple of the primary constraint:

$$H = \frac{u}{4}\big[2gy(p_x^2 + p_y^2) - 1\big].$$

The primary constraint is preserved, and it is first class; the multiplier is an undetermined gauge parameter. The simple choice $u = 1$ gives the same Hamiltonian equations found earlier. One might also consider including the gauge fixing condition $\dot{x}^2 + \dot{y}^2 = 2gy$ as a non-holonomic constraint along with its attendant constraint force. This is essentially just imposing the constant energy as a non-holonomic constraint, and, as mentioned, such a constraint does no work and has vanishing constraint force. This is representative of what happens for other time parameter gauge invariant actions (e.g., Jacobi), and their gauge fixing options.

5 The distributional splitting

In this section we derive the key distributional splitting and look at the special case of homogeneous constraints. A consequence of homogeneity is that there is a distributional formulation of the constrained Hamiltonian equations. The proofs of these results are little changed from the earlier work of Bates and Śniatycki [2] who treated the linear case, but we reproduce them here for the sake of completeness.

Let M be the manifold given by the common zeroes of the constraint functions $\{c^1, \dots, c^K\}$, and let F be the distribution consisting of vectors in the kernel of the forms $\{\phi^1, \dots, \phi^K\}$,

$$F = \{v \in TP \mid \langle \phi^A, v\rangle = 0,\, a = 1, \dots, K\}.$$

Set H to be the distribution formed by the intersection

$$H = F \cap TM.$$

So far the only assumption needed on the metric was that it was nondegenerate. However, in order to progress with the theory for indefinite metrics, we need an additional assumption.

Definition 1. The constraint manifold M is said to be g-nondegenerate if the restriction of the metric g to the distribution $\pi_* TM^\omega$ is nondegenerate. Here $\pi : P \to Q$ is the cotangent bundle projection.[1]

It is easy to check that this if M is given locally by the common zeroes of constraint functions $c^1, \dots, c^K$, then the condition is equivalent to the nondegeneracy of the matrix M^{AB} of section 3.3 whose AB component is $g(\pi_* X_{c^A}, \pi_* X_{c^B})$.

Theorem 1. *On a g-nondegenerate constraint manifold M, the restriction of ω to the distribution H, denoted ω_H, is nondegenerate.*

Proof. Since the forms ϕ^A are assumed independent and semi-basic (since they annihilate the vertical space VTP), we may assert the existence of $n-K$ additional independent semi-basic one-forms $\phi^{K+1}, \dots, \phi^n$. In the local chart $\{q^1, \dots, p_n\}$, $\phi^a = \phi^a_i dq^i$. Let ϕ be the matrix with ab component ϕ^a_b. Our assumption implies that the matrix ϕ is invertible. Define forms χ_a by

$$\chi_a = (\phi^{-1})^j_a dp_j, \quad a = 1, \dots, n.$$

It then follows that $\{\phi^1, \dots, \phi^n, \chi_1, \dots, \chi_n\}$ is a symplectic coframe as

$$\phi^a \wedge \chi_a = \phi^a_i (\phi^{-1})^j_a dq^i \wedge dp_j = dq^a \wedge dp_a = \omega.$$

Since the restriction of ω to F is

$$\omega|_F = \phi^{K+1} \wedge \chi_{K+1} + \dots + \phi^n \wedge \chi_n,$$

it follows that the symplectic perpendicular F^ω is

$$F^\omega = \ker\{\phi^1, \dots, \phi^n, \chi_{K+1}, \dots, \chi_n\},$$

and this implies that F is coisotropic. Since M is defined by the common zeroes of $c^1, \dots, c^K$, tangent vectors to M are defined by the kernel of the forms

$$\psi^A = dc^A = c^A_{,m} dq^m + c^{A,r} dp_r \quad A = 1, \dots, K.$$

It follows that the intersection of F^ω and TM is given by

$$F^\omega \cap TM = \ker\{\phi^1, \dots, \phi^n, \chi_{K+1}, \dots, \chi_n, \psi^1, \dots, \psi^K\}.$$

[1] That F^ω could intersect the constraint manifold M nontransversally was overlooked in the original proof for linear constraints found in [2], where I thought that the determinental multiplier was always 1.

The conclusion will follow if we can show that

$$\phi^1 \wedge \cdots \wedge \phi^n \wedge \chi_{K+1} \wedge \cdots \wedge \chi_n \wedge \psi^1 \wedge \cdots \wedge \psi^K$$

is a volume, since then it follows that $F^\omega \cap TM = 0$, and so $F^\omega \cap H = 0$. Since $F^\omega \oplus H = F$ by dimension count, and F is coisotropic, we may conclude that ω_H is nondegenerate.

To actually show that the $2n$ form is a volume, first observe that since the result is a pointwise result, we may choose, for a fixed point z, a local symplectic chart such that

$$\phi^1(z) = dq^1, \ldots, \phi^n(z) = dq^n, \chi_1(z) = dp_1, \ldots, \chi_n(z) = dp_n.$$

In other words, $\phi^a_b(z) = \delta^a_b$. This means that the wedge product

$$\phi^1 \wedge \cdots \wedge \phi^n \wedge \chi_{K+1} \wedge \cdots \wedge \chi_n \wedge \psi^1 \wedge \cdots \wedge \psi^K$$

will equal $\det(g_K) dq^1 \wedge \cdots \wedge dq^n \wedge dp_1 \wedge \cdots \wedge dp_n$ where $\det(g_K)$ is the determinant of the upper left $K \times K$ block of the metric g^{ab} in this frame (this is just the earlier defined M^{AB}). This is immediate once one realizes that the only part of the forms ψ^a that survive the wedge product with all of the ϕ^a are the terms $\phi^a_r g^{rs} dp_s$, and all the terms involving $dp_{K+1}, \ldots, dp_n$ are annihilated by being wedged with $\chi_{K+1} \wedge \cdots \wedge \chi_n$. Observe that the inequality $\det(g_K) \neq 0$ is exactly the condition of g-nondegeneracy in our special frame. □

Define the distribution K by $K = TM \cap H^\omega$. Since $TM = H \oplus K$, the constrained vector field X may be decomposed as $X = X^H + X^K$. Two extreme cases of this are when the constraint is the Hamiltonian itself, so the Lagrange multiplier vanishes and $X = X^K$, and when the constraints are homogeneous, and then $X = X^H$. To see this, observe that for $A = 1, \ldots, K$, the pairing

$$\langle \phi^A, X \rangle = \langle \phi^A, \dot{q} \rangle = 0$$

by homogeneity, which implies that X is in H. Evaluating the constrained equation of motion

$$X \lrcorner\, \omega = dh + \lambda_A \phi^A$$

on the distribution H annihilates the terms involving the Lagrange multipliers, and we obtain

$$X \lrcorner\, \omega_H = dh_H.$$

The expression dh_H denotes the restriction of dh to the distribution H, and we may think of the constrained Hamiltonian equations as being in distributional form.

6 Conservation laws

In Hamiltonian mechanics symmetry (and the closely related reduction theory) are usually studied together with conservation laws because of their equivalence, which is the content of the first Noether theorem. In nonholonomic systems, this equivalence is in general broken, so not all symmetries yield conservation laws, and not all conservation laws yield symmetries (see [2] for a simple example.) Looking ahead to the reduction theory, we will say that a vector is *horizontal* if it lies in the distribution H.

Suppose that we have a Lie group G acting in a Hamiltonian way on the phase space P such that it possesses a momentum map $j : P \to \mathfrak{g}^*$. In other words, for each $\zeta \in \mathfrak{g}^*$, the momentum j_ζ corresponding to ζ

$$j_\zeta = \langle j, \zeta \rangle$$

has a Hamiltonian vector field X_ζ satisfying

$$X_\zeta \lrcorner\, \omega = dj_\zeta.$$

If we further assume that G leaves the constraint manifold M, the constraint forms ϕ^a and the Hamiltonian h invariant, then it also leaves the Lagrange multipliers invariant, and thus the structure of the equations of motion

$$X \lrcorner\, \omega = dh + \lambda_A \phi^A$$

is invariant as well. The vector field X_ζ is called an infinitesimal symmetry.

Lemma 1. *The momentum j_ζ associated to ζ is conserved if the vector field X_ζ is horizontal.*

Proof. This is a simple calculation:

$$\langle dj_\zeta, X \rangle = X \lrcorner\, (X_\zeta \lrcorner\, \omega\rangle = -X_\zeta \lrcorner\, (dh + \lambda_A \phi^A) = -X_\zeta \lrcorner\, \lambda_A \phi^A = 0. \qquad \square$$

Observe that the invariance of the constraint manifold M was never used in the proof of the lemma. This implies that the following more general theorem is true.

Theorem 2. *Let f be a function with Hamiltonian vector field X_f with the property that it preserves the Hamiltonian h and lies in the distribution F consisting of the kernel of the constraint forms ϕ^A. Then f is a constant of motion.*

7 Symmetry and reduction

So far the nonlinear constraint theory looks virtually identical to the linear theory. It is in the reduction by symmetry that the nonlinear case differs, and this is because the constrained vector field does not have to lie in the distribution H.

Recall that the time derivative of a function f along an integral curve of the constrained vector field X is

$$\frac{df}{dt} = \langle df, X \rangle = -X_f \lrcorner\, (dh + \lambda_A \phi^A).$$

Set $\psi := dh + \lambda_A \phi^A$, write this derivative in Poisson bracket form as

$$\dot{f} = \{f, \psi\}$$

where we are taking the Poisson bracket of the function f and the one form ψ to be $\Lambda(df, \psi)$, where $\Lambda = \omega^{-1}$ is the structure tensor of the Poisson bracket.

In specific problems it is often more convenient to solve for the Lagrange multipliers explicitly and then use a Dirac bracket-like formulation in which the multipliers are eliminated. Since our bracket convention implies

$$\{f, dh + \lambda_A \phi^A\} = \{f, h\} + \lambda_A \{f, \phi^A\}$$

and the preservation of the constraints by the dynamics implies

$$0 = \dot{c}^A = X_{c^A} \lrcorner (X \lrcorner \omega) = \omega(X_h, X_{c^A}) - \lambda_B \{c^A, \phi^B\},$$

and we already have the matrix $M^{BD} := -\{c^B, \phi^D\}$ with assumed inverse $M_{EF}M^{FD} = \delta_E^D$ from which the Lagrange multiplier λ_A may be found as

$$\lambda_A = -M_{AB}\{h, c^B\}.$$

This implies that the bracket formulation of the equations of motion may be given as

$$\dot{f} = \{f, h\} + \{f, \phi^A\} M_{AB} \{c^B, h\}.$$

It is important to note that this dynamical equation is only defined on the image of the constraint manifold under the Legendre transformation, and not on the entire phase space. In practice however, keeping track of this is not an issue, and we will continue to ignore it as it makes no essential conceptual difference in what follows.

Let G be a symmetry group of the dynamical system, by which we mean that the group G acts symplectically on the phase space, preserves the constraint manifold and constraint forms, and leaves the Hamiltonian invariant. These assumptions imply that the group action also preserves the Poisson bracket, the constraint distributions H and K as well as the Lagrange multipliers.

Now, if f is a function invariant under the action of the symmetry group G, $f \in C^\infty(P)^G$, then the Poisson bracket $\{f, \psi\}$ is invariant as well, since both Λ and ψ are. This implies that the map

$$\{\cdot, \psi\} : C^\infty(P)^G \to C^\infty(P)^G : f \to \{f, \psi\}$$

is an outer Poisson derivation[2] (it is an outer derivation since it involves an invariant one-form and not an invariant function) on the ring of invariant functions, and so if $f \in C^\infty(P)^G$, $\dot{f} = \{f, \psi\}$ may be viewed as a differential equation on the reduced space P/G. In this way a vector field $\bar{X}$ is defined on the reduced space, and this is the projection of X by the quotient map $P \to P/G$ when restricted to the quotient of the constraint manifold M/G. A key benefit to considering reduction in this formulation is that the construction is well-defined even when the quotient space is not a manifold, as long as we assume that the group action is proper, for then the quotient space is a subcartesian differential space, and the invariant functions still separate points on the quotient (a good reference for this material may be found in Śniatycki [13]). In this sense we may view the Poisson bracket-like formulation as providing the singular reduction of nonlinear nonholonomic constraints simply by restriction to the invariant functions.

[2]It is important to note here that there is no statement that the dynamical flow preserves the Poisson bracket, even on the constraint manifold.

8 Examples (2)

8.1 The central force problem with a speed constraint

Consider the motion of a particle in space subject to the action of a central force and and the constraint of constant speed. Then the Hamiltonian may be written as

$$h = \frac{1}{2}|p|^2 + V(r)$$

with $r = |q|$. Take the constraint to be $c = \frac{1}{2}|p|^2$. The equations of motion are

$$\begin{aligned}\dot{q}^a &= p_a \\ \dot{p}_a &= -\frac{q^a}{r}V'(r) - \lambda p_a\end{aligned}$$

Solving for the Lagrange multiplier yields

$$\lambda = -\frac{1}{2c}\frac{q \cdot p}{r}V'(r).$$

Since the problem is rotationally invariant, we introduce the three independent rotation invariants $\sigma_1 = r^2 = |q|^2$, $\sigma_2 = q \cdot p$, and $\sigma_3 = |p|^2 = 2c$. The singularly reduced problem in σ_1-σ_2 space (with the semi-algebraic constraint $\sigma_1 \geq 0$) has equation of motion

$$\begin{aligned}\dot{\sigma}_1 &= 2\sigma_2, \\ \dot{\sigma}_2 &= 2\left(c - \left(\sigma_1 - \frac{1}{2c}\sigma_2^2\right)\frac{dV}{d\sigma_1}\right).\end{aligned}$$

The dynamical interest here is in the destabilization of the circular orbits in Keplerian like potentials with the imposition of the constant speed constraint. If this seems counterintuitive, one may think of it as the imposition of the constant speed constraint, even though it is rotationally invariant, means that the angular momentum is no longer conserved.

8.2 A classical particle with spin

The reduction by symmetry required that the bracket map the invariant functions to invariant functions:

$$\{\cdot, \psi\} : C^\infty(P)^G \to C^\infty(P)^G.$$

However, and this is the crucial point, there is no requirement that the form ψ actually be invariant as well in order to have reduced dynamics[3]. Thus, it is possible to have reduction to invariant functions without the full problem being invariant. From a distributional point of view, this says that reduction exists in the case when the constrained vector field varies under the group action, but only in directions that are parallel to the tangents to the group action. A nontrivial example of this

[3]There *is* such a requirement if one wants to have the reduced dynamics still in Poisson like form, with a corresponding reduced form $\bar{\psi}$.

behaviour may be found in the reduction of a classical spinning rigid body to the classical particle with spin. Consider a uniformly charged symmetric rigid body with moment of inertia I, total charge q, mass m and gyromagnetic ratio g in the presence of a magnetic field B. The equations of motion are

$$\begin{aligned}\dot{x} &= v,\\ \dot{v} &= \frac{q}{m}(v \times B) + \frac{gI}{m} DB^*(x)(k(\operatorname{Ad}_A X))\\ \dot{A} &= A(X - g \operatorname{Ad}_{A^{-1}} B(x)),\\ \dot{X} &= 0.\end{aligned}$$

Here $x \in \mathbb{R}^3$ is the position of the centre of mass, $v \in \mathbb{R}^3$ is the velocity of the centre of mass, $A \in SO(3)$ is the orientation of the rigid body, and $X \in so(3)$ is the angular velocity. k is the Killing metric on the rotation group. We have also employed the identification of antisymmetric matrices and 3-vectors as appropriate. The nonlinear nonholonomic constraint of constant length of angular momentum is applied, and the problem is reduced with respect to the action of the rotation group to get Souriau's model of a classical particle with spin

$$\begin{aligned}\dot{x} &= v,\\ \dot{v} &= \frac{q}{m}(v \times B) + \frac{g}{m} DB^*(x)(S)\\ \dot{S} &= g[S, B]\end{aligned}$$

The details of this calculation, but not this point of view, may be found in [4].

9 Notes

1. In the mechanics literature there are differing definitions of just what constitutes D'Alembert's principle, the virtual work of perfect constraints, etc. It seems to us that the D'Alembert principle is at heart the choice to pick the *orthogonal* projection of the acceleration of the unconstrained problem onto the affine plane. One could pick a different projection, but it would in general result in different constrained equations of motion. In coming to this understanding we profited greatly from the thoughtful discussions in Marle [11] and Rosenberg [12], as well as many insightful comments from J. Śniatycki.

2. One may observe a superficial analogy between the constructions of various authors computing a nonholonomic bracket (van der Schaft and Maschke [14], Bates [1], de Leon [5], Koon and Marsden [9] etc) and this paper. However, there is a fundamental difference in that we are using the standard Poisson bracket and putting all of the nonholonomic information into the one-form that drives the dynamics.

3. In various examples one may of course consider what happens to the image of the various distributions H, K under reduction. This is of interest in the special case where the reduced distribution $\bar{K} = 0$, so the invariant part of the dynamics satisfies the distributional Hamiltonian equations. Our results

would indicate that this is not the most fundamental way to view these problems, and this is why we have stressed that the fundamental structure that enables reduction by symmetry is the action of the outer Poisson derivation on the invariant functions. Compare the discussion in Cantrijn et al [3].

Acknowledgement

JMN was supported by the National Science Council of the R.O.C. under the grant NSC-99-2112-M-008-004 and in part by the National Center of Theoretical Sciences (NCTS).

References

[1] L. Bates: Examples of singular nonholonomic reduction. *Reports on Mathematical Physics* 42 (1998) 231–247.

[2] L. Bates, J. Śniatycki: Nonholonomic reduction. *Reports on Mathematical Physics* 32 (1993) 99–115.

[3] F. Cantrijn, M. de Leon, J. Marrero, D. de Diego: Reduction of constrained systems with symmetries. *Journal of mathematical physics* 40 (1999) 795–820.

[4] R. Cushman, D. Kempainen, J. Śniatycki: A classical particle with spin realized by reduction of a nonlinear nonholonomic constraint. *Reports on mathematical physics* 41 (1) (1998) 133–142.

[5] M. de León, D. de Diego: Mechanical systems with nonlinear constraints. *International journal of theoretical physics* 36 (4) (1997) 979–995.

[6] M. Giaquinta, S. Hildebrandt: *Calculus of Variations I*. Number 310 in Grundlehren der mathematischen Wissenschaften. Springer-Verlag (1996).

[7] H. Goldstein, C. Poole, J. Safko: *Classical mechanics*. Addison-Wesley, third edition (2002).

[8] X. Grácia, R. Martin: Regularity and symmetries of nonholonomic systems. *J. Phys. A: Math. Gen.* 38 (2005) 1071–1087.

[9] W. Koon, J. Marsden: Poisson reduction for nonholonomic mechanical systems with symmetry. *Reports on mathematical physics* 42 (1998) 101–134.

[10] O. Krupková, J. Musilová: The relativistic particle as a mechanical system with non-holonomic constraints. *J. Phys. A: Math. Gen.* 34 (2001) 3859–3875.

[11] C.-M. Marle: Various approaches to conservative and nonconservative nonholonomic systems. *Reports on mathematical physics* 42 (1998) 211–229.

[12] R. Rosenberg: *Analytical dynamics of discrete systems*. Plenum press (1997).

[13] J. Śniatycki: Orbits of families of vector fields on subcartesian spaces. *Annales de L'Institut Fourier* 53 (2003) 2257–2296.

[14] A. van der Schaft, B. Maschke: On the hamiltonian formulation of nonholonomic mechanical systems. *Reports on mathematical physics* 34 (2) (1994) 225–233.

Authors' addresses:
Larry M. Bates, Department of Mathematics, University of Calgary, Calgary, Alberta, Canada T2N 1N4

E-mail: bates@ucalgary.ca

James M. Nester, Department of Physics, Graduate Institute of Astronomy, and Center for Mathematics and Theoretical Physics, National Central University, Chungli 320, Taiwan

E-mail: `nester@phy.ncu.edu.tw`

Received: 12 August, 2011
Accepted for publication: 20 August, 2011
Communicated by: Olga Krupková