The Geometry of the Gibbs-Appell Equations and Gauss' Principle of Least Constraint*

Andrew D. Lewis

Department of Applied Mechanics
California Institute of Technology
Pasadena, CA 91125

1 June, 1995

Abstract

We present a generalisation of the Gibbs-Appell equations which is valid for general Lagrangians. The general form of the Gibbs-Appell equations is shown to be valid in the case when constraints and external forces are present. In the case when the Lagrangian is the kinetic energy with respect to a Riemannian metric, the Gibbs function is shown to be related to the kinetic energy on the tangent bundle of the configuration manifold with respect to the Sasaki metric. We also make a connection with the Gibbs-Appell equations and Gauss' principle of least constraint in the general case.

1. Introduction

In this paper we study the so-called *Gibbs-Appell* equations and relate these equations to Gauss' principle of least constraint. The basic structure of the Gibbs-Appell equations was introduced by Gibbs [1879] and was developed further by Appell [1900b], Appell [1900a]. The equations have held their appeal mainly because of their very simple form. However, these equations have yet to be placed in a general geometric framework since the presentation typically relies on the mechanical system being a collection of point masses and rigid bodies. In this paper we present the Gibbs-Appell equations for regular Lagrangians and show that, like the equations in their less general form, they provide a very simple way of writing the equations of motion for systems with constraints. The classical reference for Gauss' principle of least constraint is Gauss [1829].

Interest in the Gibbs-Appell equations has carried on through the present in the work of various researchers. A few recent references are Desloge [1988], Townsend [1991], Sharf, D'eleuterio, and Hughes [1992]. Some comments are made in these papers regarding similarities between the Gibbs-Appell equations and the equations of motion presented in Kane [1983]. We shall not comment on this here. We should point out that the above references

*Submitted to *Reports on Mathematical Physics*.

all deal with mechanical systems formed by collections of particles and rigid bodies. A general presentation of mechanics, including constraints, is given in Giachetta [1992]. In fact, in that paper, the general Gibbs function is almost constructed, but is not realised as such. The presentation of the Gibbs-Appell equations in Giachetta [1992] is also quite different from ours. A statement of Gauss' principle of least constraint may also be found in that paper. Other work related to Gauss' principle is that of Udwadia and Kalaba [1992], Kalaba and Udwadia [1993]. However, that work is limited by restrictions on the Lagrangian and on the structure of the configuration manifold for the system (it is assumed to be $\mathbb{R}^n$). Another recent work on Gauss' principle of least constraint is that of Ray [1992].

In Section 2 we present a simple example which uses the classical Gibbs-Appell equations to arrive at the equations of motion. We show "by hand" that these equations are equivalent to Lagrange's equations. In the general formulation we rely heavily on constructions using jet bundles. The necessary background in jet bundles is presented in Section 3. The reader familiar with jets should be able to skip this section except to refer to it for notation. The basic elements of Lagrangian mechanics, the Lagrangian, external forces, and constraints, are formulated quickly in Section 4. In Section 5 we present the Gibbs-Appell formulation of the equations of motion in our general framework. First we construct the Gibbs function. Then, with the Gibbs function in hand, it is an easy matter to give the Gibbs-Appell equations in the absence of constraints and show that these equations are equivalent to Lagrange's equations. This is done in Section 5.2. To present the equations in the presence of constraints takes a bit more development, and this is done in Section 5.3. The interesting special case when the Lagrangian is the kinetic energy with respect to a Riemannian metric is presented in Section 6. In this case the Gibbs function is related to the Sasaki metric on the tangent bundle of the configuration manifold. With our development of the Gibbs-Appell equations, it is a comparatively simple matter to make connections with a general version of Gauss' principle of least constraint which we do in Section 7.

Acknowledgements

The author would like to thank Gabor Stepan for his introduction to the Gibbs-Appell equations. Discussions with Richard Murray and Jim Ostrowski have also been helpful. Jerry Marsden pointed out the link with the Sasaki metric discussed in Section 6.

2. A Simple Example

In this section we present a very simple example as a means of demonstrating how the Gibbs-Appell equations may be applied in their classical form. This serves to motivate our more general discussion. The presentation is somewhat formal in this section.

The example is a particle of mass m moving in $\mathbb{R}^3$ subject to the constraint

$$\dot{z} - y\dot{x} = 0$$

and to an arbitrary force $F = (F_x, F_y, F_z)$. The constraint specifies a distribution D on $\mathbb{R}^3$ and we may choose

$$\left\{ \frac{\partial}{\partial y}, \frac{\partial}{\partial z} + y\frac{\partial}{\partial x} \right\}$$

as a basis for D. Thus, as coordinates for the admissible states we can use

$$(x, y, z, \nu^1 \triangleq \dot{y}, \nu^2 \triangleq \dot{z} + y\dot{x}).$$

This gives the three relations

$$\begin{aligned} \nu^1 &= \dot{y} \\ \nu^2 &= \dot{z} + y\dot{x} \\ 0 &= \dot{z} - y\dot{x}. \end{aligned} \tag{2.1}$$

The equations (2.1) may be inverted to give

$$\begin{aligned} \dot{x} &= \frac{1}{2y}\nu^2 \\ \dot{y} &= \nu^1 \\ \dot{z} &= \frac{1}{2}\nu^2. \end{aligned} \tag{2.2}$$

Thus we compute

$$\begin{aligned} \ddot{x} &= \frac{1}{2y}\dot{\nu}^2 - \frac{1}{2y^2}\nu^1\nu^2 \\ \ddot{y} &= \dot{\nu}^1 \\ \ddot{z} &= \frac{1}{2}\dot{\nu}^2. \end{aligned} \tag{2.3}$$

The Gibbs function for this system is given by

$$G = \frac{1}{2}m(\ddot{x}^2 + \ddot{y}^2 + \ddot{z}^2).$$

If we substitute the relations (2.3) into the Gibbs function we obtain

$$G = \frac{1}{2}m\left(\frac{1}{4y^2}(\dot{\nu}^2)^2 + \frac{1}{4y^4}(\nu^1)^2(\nu^2)^2 - \frac{1}{2y^3}\dot{\nu}^2\nu^1\nu^2 + (\dot{\nu}^1)^2 + \frac{1}{4}(\dot{\nu}^2)^2\right).$$

The forces which appear in the Gibbs-Appell equations are defined by

$$F_1 = \frac{\partial\dot{x}}{\partial\nu^1}F_x + \frac{\partial\dot{y}}{\partial\nu^1}F_y + \frac{\partial\dot{z}}{\partial\nu^1}F_z = F_y$$
$$F_2 = \frac{\partial\dot{x}}{\partial\nu^2}F_x + \frac{\partial\dot{y}}{\partial\nu^2}F_y + \frac{\partial\dot{z}}{\partial\nu^2}F_z = \frac{1}{2y}F_x + \frac{1}{2}F_z$$

The Gibbs-Appell equations are

$$\frac{\partial G}{\partial\dot{\nu}^a} = F_a, \qquad a = 1, 2.$$

Doing the calculations gives

$$m\dot{\nu}^1 = F_y \tag{2.5a}$$
$$\left(\frac{m}{4y^2} + \frac{m}{4}\right)\dot{\nu}^2 - \frac{m}{4y^3}\nu^1\nu^2 = \frac{1}{2y}F_x + \frac{1}{2}F_z. \tag{2.5b}$$

If we append to these equations the equations (2.2), we obtain the correct number of equations describing the motion of the system with the given constraints.

It remains to be shown, however, that these equations are the same as Lagrange's equations. Let us quickly go through this formalism for this example to verify that the two approaches are in fact equivalent. Lagrange's equations are given by

$$m\ddot{x} = -\lambda y + F_x \tag{2.6a}$$
$$m\ddot{y} = F_y \tag{2.6b}$$
$$m\ddot{z} = \lambda + F_z \tag{2.6c}$$

where λ is a Lagrange multiplier. Substituting (2.6c) into (2.6a) to eliminate λ gives

$$m\ddot{x} + my\ddot{z} = F_x + yF_z$$

which may easily be seen to be equivalent to (2.5b) using (2.3). Also, it is clear that (2.6b) is equivalent to (2.5a). This shows that, for this example, the Gibbs-Appell equations are equivalent to Lagrange's equations.

3. Jet Bundles

We shall need to know some fairly detailed structure of bundles of jets from $\mathbb{R}$ to the configuration manifold Q. In particular we need the structure of the jet bundles as affine bundles. Therefore, we start with a discussion of affine bundles in Section 3.1. With this we discuss the necessary facts about jet bundles in Sections 3.2, 3.3, and 3.4. So-called *special vector fields* are discussed in Section 3.5. Finally, we introduce the acceleration derivative in Section 3.6.

3.1. Affine Bundles In this section we give a review of affine bundles as discussed in Goldschmidt [1967]. We begin with affine spaces.

3.1 DEFINITION: An *affine space* over a field k modeled on a vector space E is a set A which is a homogeneous space for the additive group of E, and upon which the action of E is free. If $e \in E$ and $a \in A$ we denote by $e + a$ the action of e on a. □

If $a, b \in A$ we denote by $a - b$ the unique element of E defined by $(a - b) + b = a$. Also, for each $a \in A$ we have a natural identification of A with E given by $b \mapsto b - a \in E$. This identification induces a vector space structure on A where the zero vector is a. A non-empty subset U of A is an *affine subspace* of A if U is a subspace of A considered as a vector space with the vector space structure induced by some element $a \in U$. An affine subspace U may be thought of as an affine space modeled on the subspace of E given by all elements $x - y$ where $x, y \in U$. Clearly then, if F is a subspace of E and if $a \in A$, the set

$$\{e + a \mid \; e \in F\}$$

is an affine subspace of A.

Now we discuss affine bundles.

3.2 DEFINITION: Let $\pi\colon E \to B$ be a vector bundle. An *affine bundle* over B modeled on E is a fibre bundle $\tau\colon A \to B$ and a map $\phi\colon E \times_B A \to A$ such that the following diagram commutes

$$\begin{array}{ccc} E \times_B A & \xrightarrow{\phi} & A \\ {\scriptstyle pr_1}\downarrow & & \downarrow{\scriptstyle \tau} \\ E & \xrightarrow[\pi]{} & B \end{array}$$

and such that $e + a \triangleq \phi(e,a)$ makes A_b into an affine space modeled on E_b for each $b \in B$. Here pr_1 is projection onto the first factor. Let F be a vector subbundle of E. A *subbundle* of A modeled on F is a subbundle U of the fibre bundle $\tau\colon A \to B$ such that U_b is an affine subspace of A_b associated with F_b. □

Finally, we give some obvious notation to an object which we will have occasion to use. Let M be a differentiable manifold, let $\mathcal{S}$ be a sequence of fibre bundles

$$A_0 \xleftarrow{\tau_1} A_1 \longleftarrow \cdots \longrightarrow A_k \xleftarrow{\tau_k} A_{k+1} \longrightarrow \cdots$$

and suppose that we have fibre bundles $\pi_k\colon A_k \to M$ for $k = 0, 1, \ldots$ Suppose that $\pi\colon E \to M$ is a vector bundle. We may then define a family of vector bundles by pull-back,

$$\mathcal{F} = \{\pi_k^*\pi\colon \pi_k^* E \to A_k\}_{k \in \mathbb{Z}^+}.$$

We shall say that $\mathcal{S}$ is a *tower of affine bundles modeled on* $\mathcal{F}$ if, for each $k \in \mathbb{Z}^+$, $\tau_k\colon A_k \to A_{k-1}$ is an affine bundle modeled on $\pi_{k-1}^*\pi\colon \pi_{k-1}^* E \to A_{k-1}$.

3.2. The Bundle of Jets from $\mathbb{R}$ to Q The notation in this section follows that of Golubitsky and Guillemin [1973].

We first need to say what we mean when two curves have the same derivative up to some order at a point. Let $c_1 : \mathbb{R} \to Q$ and $c_2 : \mathbb{R} \to Q$ be two curves on Q so that $c_1(t) = c_2(t) = q$ for some $t \in \mathbb{R}$. Let $(q^1, \ldots, q^n)$ be a coordinate chart around q. We shall say that c_1 and c_2 *agree at order* k *at* q if the kth time derivatives of the components $(q^1(s), \ldots, q^n(s))$ agree at $s = t$. It may be seen that this definition of equivalence is independent of coordinate chart. If c_1 and c_2 agree at order k at q we shall write

$$c_1^{(k)}(t) = c_2^{(k)}(t).$$

3.3 DEFINITION: Let Q be a differentiable manifold, let $t \in \mathbb{R}$, and let $c_1, c_2\colon \mathbb{R} \to Q$ be curves on Q such that $c_1(t) = c_2(t) = q$. We say that c_1 and c_2 are *equivalent to order* m *at* t if

$$c_1^{(k)}(t) = c_2^{(k)}(t)$$

for $k = 1, \ldots, m$. We shall denote the equivalence class by $[c_1]_m$. We denote the set of all such equivalence classes by $J^m(\mathbb{R}, Q)_{t,q}$. The set

$$J^m(\mathbb{R}, Q) \triangleq \mathop{\mathring{\bigcup}}_{(t,q) \in \mathbb{R} \times Q} J^m(\mathbb{R}, Q)_{t,q}$$

is called the set of *m-jets from $\mathbb{R}$ to Q*. By definition we take $J^0(\mathbb{R}, Q) = \mathbb{R} \times Q$. For $[c]_m \in J^m(\mathbb{R}, Q)_{t,q}$ we call t the *source* of $[c]_m$ and q the *target* of $[c]_m$. □

We will be interested in the sets of 1-jets and 2-jets. If $(q^1, \ldots, q^n)$ is a coordinate chart for Q, we have natural coordinates for $J^1(\mathbb{R}, Q)$ given by

$$(t, q^1, \ldots, q^n, v^1, \ldots, v^n).$$

In a similar manner we have natural coordinates

$$(t, q^1, \ldots, q^n, v^1, \ldots, v^n, a^1, \ldots, a^n)$$

for $J^2(\mathbb{R}, Q)$. We shall write a typical element of $J^1(\mathbb{R}, Q)$ as v and a typical element of $J^2(\mathbb{R}, Q)$ as a. Elements of $J^1(\mathbb{R}, Q)$ and $J^2(\mathbb{R}, Q)$ transform in natural coordinates in specific ways according to the change of coordinates on Q. To be specific, if $(Q^1, \ldots, Q^n)$ are coordinates for Q different than $(q^1, \ldots, q^n)$, we have, with the obvious notation,

$$\begin{aligned} V^i &= \frac{\partial Q^i}{\partial q^j} v^j \\ A^i &= \frac{\partial Q^i}{\partial q^j} a^j + \frac{\partial^2 Q^i}{\partial q^j \partial q^k} v^j v^k. \end{aligned}$$

We now define a family of projections from "higher" jet bundles to "lower" jet bundles. If $l < m$ there is a canonical projection $\tau_{m,l} \colon J^m(\mathbb{R}, Q) \to J^l(\mathbb{R}, Q)$ which "forgets" the higher order of equivalence. We also define projections $\rho_m \colon J^m(\mathbb{R}, Q) \to Q$ by $\rho_m \triangleq pr_2 \circ \tau_{m,0}$ where $pr_2 \colon \mathbb{R} \times Q \to Q$ is the projection onto the second factor. Note that in natural coordinates for $J^1(\mathbb{R}, Q)$ we have

$$\rho_1(t, q^1, \ldots, q^n, v^1, \ldots, v^n) = (q^1, \ldots, q^n)$$

and in natural coordinates for $J^2(\mathbb{R}, Q)$ we have

$$\rho_2(t, q^1, \ldots, q^n, v^1, \ldots, v^n, a^1, \ldots, a^n) = (q^1, \ldots, q^n).$$

If $c \colon \mathbb{R} \to Q$ is a map, $j^m c \colon \mathbb{R} \to J^m(\mathbb{R}, Q)$ will denote the map which assigns to t the equivalence class $[c]_m \in J^m(\mathbb{R}, Q)_{t,c(t)}$. If the curve c is given by

$$s \mapsto (q^1(s), \ldots, q^n(s)),$$

then the map $j^1 c$ is given by

$$j^1 c(t) = \left(t, q^1(t), \ldots, q^n(t), \frac{dq^1}{ds}(t), \ldots, \frac{dq^n}{ds}(t)\right)$$

and the map $j^2 c$ is given by

$$j^2 c(t) = \left(t, q^1(t), \ldots, q^n(t), \frac{dq^1}{ds}(t), \ldots, \frac{dq^n}{ds}(t), \frac{d^2 q^1}{ds^2}(t), \ldots, \frac{d^2 q^n}{ds^2}(t)\right).$$

For each $t \in \mathbb{R}$ and $q \in Q$ we have a canonical identification of $T_q Q$ with $J^1(\mathbb{R}, Q)_{t,q}$. We will implicitly utilise this identification at times.

Note on Notation: It is common to see natural coordinates for $J^1(\mathbb{R}, Q)$ written as

$$(t, q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n).$$

We will stick to using v's instead of $\dot{q}$'s unless there is a specific curve on Q we are considering and so we wish to think of v^i as dq^i/dt. In this case we will use $\dot{q}^i$. Similar remarks hold for using a^i as opposed to using $\ddot{q}^i$. □

3.3. The Structure of the Tangent Bundle of a Jet Bundle There will be a couple of instances where it will be useful to realise the special structure of the tangent bundle of a jet bundle. We state this as a lemma.

3.4 Lemma: *$J^m(\mathbb{R}, Q)$ may be canonically embedded as a submanifold of $TJ^{m-1}(\mathbb{R}, Q)$.*

Proof: Let $[c]_m \in J^m(\mathbb{R}, Q)_{t,q}$. Recall that $j^{m-1}c$ is a curve on $J^{m-1}(\mathbb{R}, Q)$. Therefore $(j^{m-1}c)'(t) \in TJ^{m-1}(\mathbb{R}, Q)$. In this manner we obtain a 1–1 correspondence between $J^m(\mathbb{R}, Q)$ and a subset of $TJ^{m-1}(\mathbb{R}, Q)$. In natural coordinates it is easy to verify that this inclusion is an embedding. ■

We may write the above embeddings in the case where $m = 1, 2$. For $m = 1$ we have

$$(t, q^i, v^j) \hookrightarrow (t, q^i, 1, v^j)$$

and for $m = 2$ we have

$$(t, q^i, v^j, a^k) \hookrightarrow (t, q^i, v^j, 1, v^k, a^l).$$

We shall also be interested in the vertical bundle corresponding to the fibration $\sigma_1 \colon J^1(\mathbb{R}, Q) \to J^0(\mathbb{R}, Q)$. We shall denote this subbundle of $TJ^1(\mathbb{R}, Q)$ by $VJ^1(\mathbb{R}, Q)$. Note that we may naturally identify $VJ^1(\mathbb{R}, Q)$ with the pull-back bundle $\rho_1^* TQ$. In natural coordinates this identification has the form

$$((t, q^i, v^j), u^k) \mapsto (t, q^i, v^j, 0, 0, u^k). \tag{3.1}$$

3.4. The Affine Structure of $\boldsymbol{J^m(\mathbb{R}, Q)}$ Now we examine the affine structure of $J^k(\mathbb{R}, Q)$ over $J^{k-1}(\mathbb{R}, Q)$. This point of view is taken in Goldschmidt and Sternberg [1973] and Hermann [1982] with both references drawing from Goldschmidt [1967].

We first need to generate a family of vector bundles which will serve as a model for the tower of affine bundles given by $\tau_{k,k-1} \colon J^k(\mathbb{R}, Q) \to J^{k-1}(\mathbb{R}, Q)$. We may construct a family of vector bundles by pull-back as follows. For $k > 1$ we define the pull-back bundle

$$\rho_k^* \tau_Q \colon \rho_k^* TQ \to J^k(\mathbb{R}, Q)$$

where $\tau_Q \colon TQ \to Q$ is the tangent bundle projection. We shall define $\tau_k = \rho_k^* \tau_Q$. This defines a family of vector bundles

$$\mathfrak{J}_Q \triangleq \{\tau_k \colon \rho_k^* TQ \to J^k(\mathbb{R}, Q)\}_{k \in \mathbb{Z}^+}.$$

Now we may state the result.

3.5 PROPOSITION: *Let $\sigma_k = \tau_{k,k-1}$ for $k \in \mathbb{Z}^+$. Then the sequence*

$$J^0(\mathbb{R}, Q) \xleftarrow{\sigma_1} J^1(\mathbb{R}, Q) \longleftarrow \cdots \longrightarrow J^k(\mathbb{R}, Q) \xleftarrow{\sigma_k} J^{k+1}(\mathbb{R}, Q) \longrightarrow \cdots$$

is a tower of affine bundles modeled on $\mathfrak{J}_Q$.

Idea of Proof: The proposition will be proved if we can show that $\sigma_k^{-1}(u)$ is an affine space modeled on T_qQ where $u \in J^{k-1}(\mathbb{R}, Q)_{t,q}$. Let $v \in T_qQ$ and let $[c]_k \in \sigma_k^{-1}(u)$. Suppose that $\tilde{c}$ is a deformation of the curve c in a neighborhood $U$of t. Thus $\tilde{c}\colon U \times (-\epsilon, \epsilon) \to Q$ is such that $\tilde{c}(s, 0) = c(s)$ for all $s \in U$. Further suppose that

$$\left.\frac{d\tilde{c}(t,\tau)}{d\tau}\right|_{\tau=0} = v.$$

Then we may define a curve on Q by

$$s \mapsto \tilde{c}(s, (s-t)^k). \tag{3.2}$$

The k-jet of this curve at $s = t$ will be denoted

$$v + [c]_k. \tag{3.3}$$

We claim that this defines an action of T_qQ on $\sigma_k^{-1}(u)$ which make the latter into an affine space modeled on the former. To show this, we must prove that

i) $v + [c]_k \in \sigma_k^{-1}(u)$,

ii) the map (3.3) defines an action,

iii) the action is free, and

iv) the action is transitive.

i: To see that $v + [c]_k \in \sigma_k^{-1}(u)$ we need only observe that the $(k-1)$-jet of the curve defined by (3.2) is the same as that of c. This may be seen by applying the chain rule to the former curve.

ii: To show that (3.3) does define an action, suppose that $v_1, v_2 \in T_qQ$. By performing successive deformations of c using these two vectors one may show that

$$(v_1 + v_2) + [c]_k = v_1 + (v_2 + [c]_k).$$

This implies that (3.3) does indeed define an action.

iii: Suppose that

$$v + [c]_k = [c]_k.$$

Then the curve defined by (3.2) must agree with c to order k. This is only true if $v = 0$. Thus the action is free.

iv: Let $[c']_k \in \sigma_k^{-1}(u)$. By applying the chain rule, one may see that there is a vector $v \in T_qQ$ so that

$$v + [c]_k = [c']_k.$$

This completes the proof. ∎

We now look at the affine structure of $J^1(\mathbb{R}, Q)$ and $J^2(\mathbb{R}, Q)$ in natural coordinates. Suppose that $v \in T_qQ$ has components $(v^1, \dots, v^n)$ and that $u \in J^1(\mathbb{R}, Q)_{t,q}$ has coordinates $(t, q^1, \dots, q^n, u^1, \dots, u^n)$. Then we have

$$v + u = (t, q^1, \dots, q^n, u^1 + v^1, \dots, u^n + v^n).$$

Similarly, for $a \in J^2(\mathbb{R}, Q)$ we have

$$v + a = (t, q^1, \dots, q^n, u^1, \dots, u^n, a^1 + v^1, \dots, a^n + v^n).$$

Observe that these definitions obey the appropriate transformation laws for jets. Also note that $J^1(\mathbb{R}, Q)$ has more structure than just that of an affine bundle over $J^0(\mathbb{R}, Q)$. It is, in fact, a vector bundle.

3.5. Special Vector Fields Let $X\colon J^1(\mathbb{R}, Q) \to TJ^1(\mathbb{R}, Q)$ be a vector field on $J^1(\mathbb{R}, Q)$. We say that X is *special* if $T\sigma_1 \circ X(v) = v \in TJ^0(\mathbb{R}, Q)$ (recalling that $J^1(\mathbb{R}, Q) \subset TJ^0(\mathbb{R}, Q)$ by Lemma 3.4). It is easy to show that in natural coordinates a special vector field must have the form

$$X(t, q^i, v^j) = (t, q^i, v^j, 1, v^k, X^l(t, q, v)).$$

From Lemma 3.4 with $m = 2$ we see that a special vector field restricts to $J^2(\mathbb{R}, Q) \subset TJ^1(\mathbb{R}, Q)$. Thus we can regard a special vector field as a map from $J^1(\mathbb{R}, Q)$ to $J^2(\mathbb{R}, Q)$ given in natural coordinates by

$$X(t, q^i, v^j) = (t, q^i, v^j, X^k(t, q, v)).$$

In the sequel we shall at times view a special vector field in this manner.

3.6. The Acceleration Derivative In this section we introduce a new derivative which will be useful to us in describing the Gibbs-Appell equations.

Recall from Section 3.1 that if A is an affine space modeled on a vector space E, then there is a canonical identification of A with E for each $a \in A$. $\sigma_2\colon J^2(\mathbb{R}, Q) \to J^1(\mathbb{R}, Q)$ is an affine bundle whose fibre, $\sigma_2^{-1}(v)$, is modeled on the vector space $T_{\rho_1(v)}Q$. Therefore, for each $a \in \sigma_2^{-1}(v)$, we have a canonical identification of $\sigma_2^{-1}(v)$ with $T_{\rho_1(v)}Q$.

Now let f be a function on $J^1(\mathbb{R}, Q)$ and fix $\bar{v} \in J^1(\mathbb{R}, Q)$. Let $\tilde{f}$ denote the restriction of f to $\sigma_2^{-1}(\bar{v})$. For each $\bar{a} \in \sigma_2^{-1}(\bar{v})$ we may regard $\tilde{f}$ as a function on $T_{\rho_1(\bar{v})}Q$ and, therefore, $\boldsymbol{d}\tilde{f}(\bar{a})$ may be identified with an element of $T^*_{\rho_1([c]_1)}Q$. Performing this construction defines a map from $J^2(\mathbb{R}, Q)$ to T^*Q which we call the *acceleration derivative* of f and denote by $\boldsymbol{d_a} f$. Let $\rho_1^*\pi_Q\colon \rho_1^*T^*Q \to J^1(\mathbb{R}, Q)$ denote the pull-back of T^*Q to $J^1(\mathbb{R}, Q)$. Here $\pi_Q\colon T^*Q \to Q$ is the cotangent bundle projection. We can also regard $\boldsymbol{d_a} f$ as taking its values in $\rho_1^*T^*Q$ and we shall use the fact that the following diagram commutes.

$$\begin{tikzcd} J^2(\mathbb{R}, Q) \arrow[rr, "\boldsymbol{d_a} f"] \arrow[dr, "\rho_2"'] & & \rho_1^*T^*Q \arrow[dl, "\pi_Q \circ pr_2"] \\ & Q & \end{tikzcd}$$

Here $pr_2 \colon \rho_1^* T^*Q = J^1(\mathbb{R}, Q) \times_Q T^*Q \to T^*Q$ is the projection onto the second factor of the fibred product.

Let us compute what the acceleration derivative looks like in natural coordinates. Let us fix $\bar{v} = (\bar{t}, \bar{q}^i, \bar{v}^j) \in J^1(\mathbb{R}, Q)$. Also fix $\bar{a} = (\bar{t}, \bar{q}^i, \bar{v}^j, \bar{a}^k) \in J^2(\mathbb{R}, Q)$. The identification of $\sigma_2^{-1}(\bar{v})$ with $T_{\rho_1(\bar{v})}Q$ at $\bar{a}$ is given by

$$(\bar{t}, \bar{q}^1, \ldots, \bar{q}^n, \bar{v}^1, \ldots, \bar{v}^n, a^1, \ldots, a^n) \mapsto (\bar{q}^1, \ldots, \bar{q}^n, a^1 - \bar{a}^1, \ldots, a^n - \bar{a}^n).$$

We can define a function on $T_{\rho_1(v)}Q$ by

$$\tilde{f}(v^1, \ldots, v^n) = f(\bar{t}, \bar{q}^1, \ldots, \bar{q}^n, \bar{v}^1, \ldots, \bar{v}^n, v^1 + \bar{a}^1, \ldots, v^n + \bar{a}^n).$$

We see that $\boldsymbol{d_a} f(\bar{a}) = \boldsymbol{d}\tilde{f}(0)$. Therefore we have

$$\boldsymbol{d_a} f(\bar{a}) = \frac{\partial f}{\partial a^i}(\bar{a}) dq^i.$$

When we deal with constrained systems we shall consider the acceleration derivative restricted to an affine subbundle of $J^2(\mathbb{R}, Q)$. Let $\bar{E}$ be a vector subbundle of $\rho_1^* TQ$. Thus $\bar{E}$ is a specification of a subspace of TQ as a function of $J^1(\mathbb{R}, Q)$. Now suppose that E is an affine subbundle of $\sigma_2 \colon J^2(\mathbb{R}, Q) \to J^1(\mathbb{R}, Q)$ modeled on $\bar{E}$. We may restrict any function f on $J^2(\mathbb{R}, Q)$ to E and follow the above construction to arrive at a map $\boldsymbol{d}_{\boldsymbol{a}}^E(f \mid E) \colon E \to \bar{E}^*$ so that the following diagram commutes.

$$\begin{tikzcd}
E \arrow[rr, "\boldsymbol{d}_{\boldsymbol{a}}^E(f|E)"] \arrow[dr, "\rho_2|E"'] & & \bar{E}^* \arrow[dl] \\
& Q &
\end{tikzcd}$$

Without additional structure we cannot identify $\bar{E}_v^*$ as a subset of $T^*_{\rho_1(v)}Q$.

4. Lagrangian Mechanics Using Jet Bundles

In this section we present the basic elements of Lagrangian mechanics formulated on jet bundles. The basic elements we shall use are the Lagrangian, external forces, and constraints. We tie everything together by saying what is meant by a solution to Lagrange's equations with a given external force and a given set of constraints.

4.1. The Lagrangian A *Lagrangian* on Q is a function on $J^1(\mathbb{R}, Q)$. Define the *fundamental tensor* corresponding to L on $J^1(\mathbb{R}, Q)$ by

$$g = \frac{\partial^2 L}{\partial v^i \partial v^j} dq^i \otimes dq^j.$$

It is easy to verify that the above definition of g is independent of the choice of natural coordinates for $J^1(\mathbb{R}, Q)$. Note that for fixed $v \in J^1(\mathbb{R}, Q)$ we may regard $g(v)$ as a symmetric bilinear form on $T_{\rho_1(v)}Q$. This allows us to define a symmetric bilinear form on the fibres of ρ_1^*TQ and so define a bundle map $\flat\colon \rho_1^*TQ \to \rho_1^*T^*Q$. We say that L is *regular* if this map is nondegenerate on each fibre. In this case the inverse bundle map will be denoted $\sharp\colon \rho_1^*T^*Q \to \rho_1^*TQ$. Note that $\flat$ and $\sharp$ are not quite the usual musical homomorphisms associated with a Riemannian metric. They are, however, the natural generalisation of those notions. In this paper we shall assume that the Lagrangians we deal with are regular. A regular Lagrangian defines a special vector field via Lagrange's equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right) - \frac{\partial L}{\partial q^i} = \frac{\partial^2 L}{\partial v^i \partial v^j}a^j + \frac{\partial^2 L}{\partial v^i \partial q^j}v^j + \frac{\partial^2 L}{\partial v^i \partial t} - \frac{\partial L}{\partial q^i} = 0.$$

The corresponding vector field on $J^1(\mathbb{R}, Q)$ is

$$X_L = \frac{\partial}{\partial t} + v^i \frac{\partial}{\partial q^i} + g^{ij}\left(\frac{\partial L}{\partial q^j} - \frac{\partial^2 L}{\partial v^j \partial q^k}v^k - \frac{\partial^2 L}{\partial v^j \partial t}\right)\frac{\partial}{\partial v^i}.$$

Viewed as taking its values in $J^2(\mathbb{R}, Q)$ as in Section 3.5, X_L has the form

$$(t, q^i, v^j) \mapsto \left(t, q^i, v^j, g^{kl}\left(\frac{\partial L}{\partial q^l} - \frac{\partial^2 L}{\partial v^l \partial v^m}v^m - \frac{\partial^2 L}{\partial v^l \partial t}\right)\right)$$

in natural coordinates. See Lewis [1995] for an example of how Lagrange's equations may be developed intrinsically.

4.2. External Forces We shall define a *force* to be a semi-basic differential one-form on the bundle $\rho_1\colon J^1(\mathbb{R}, Q) \to Q$. In natural coordinates such a one-form is given as

$$F = F_i(t, q, v)dq^i.$$

It is also possible to consider a force to be a section of the bundle $\rho_1^*\pi_Q\colon \rho_1^*T^*Q \to J^1(\mathbb{R}, Q)$. If we think of a force in this manner, using the identification of ρ_1^*TQ with $VJ^1(\mathbb{R}, Q)$ given by (3.1), we may regard $F^\sharp$ as a vertical vector field on the bundle $\sigma_1\colon J^1(\mathbb{R}, Q) \to J^0(\mathbb{R}, Q)$. In fact, in Giachetta [1992] forces are *defined* in this manner. However, we feel it is important to regard forces as differential forms since this is how they arise from first principles. Only by using the Lagrangian (by means of the map $\sharp$) can we consider a force to be a vertical vector field on $\sigma_1\colon J^1(\mathbb{R}, Q) \to J^0(\mathbb{R}, Q)$. One certainly does not want the idea of a force to depend on the Lagrangian, so the notion of a force as a one-form is fundamental.

Note that if X is a special vector field, then we may also think of $X + F^\sharp$ as a special vector field. In natural coordinates we have

$$X + F^\sharp\colon (t, q^i, v^j) \mapsto (t, q^i, v^j, X^k(t, q, v) + g^{kl}F_l(t, q, v)).$$

4.3. Constraints One of the greatest conveniences of the Gibbs-Appell equations is that, interpreted properly, they are valid in the presence of constraints. In this section we define what we shall mean by constraints. We first present the general case then specialise to affine constraints. We refer the reader to Lewis and Murray [1995] for a precise statement of the equations of motion for affine constraints as well as some variational motivation. That paper also contains a detailed detailed comparison of the so-called "nonholonomic" and "vakonomic" equations of motion for constrained systems.

General Constraints

The general type of constraint we consider is defined by an m-dimensional codistribution $\boldsymbol{\beta}$ on $J^1(\mathbb{R}, Q)$. If a basis of $\boldsymbol{\beta}$ is locally given by

$$\beta^a = \beta_0^a(t, q, v)dt + \beta_i^a(t, q, v)dq^i + \hat{\beta}_i^a(t, q, v)dv^i, \qquad a = 1, \ldots, m,$$

we will impose the additional requirement that the $m \times n$ matrix with components $\hat{\beta}_i^a$ have rank m. This condition is given geometric meaning in Giachetta [1992]. A special vector field X is said to be *$\boldsymbol{\beta}$-admissible* if $\beta(X) = 0$ for every β which is a section of $\boldsymbol{\beta}$. It is straightforward to show that a special vector field X satisfies the relations

$$\beta_0^a + \beta_i^a v^i + \hat{\beta}_i^a X^i = 0, \qquad a = 1, \ldots, m \tag{4.1}$$

in natural coordinates. Denote $\mathcal{C} = \boldsymbol{\beta}^\perp \cap J^2(\mathbb{R}, Q)$ recalling that $J^2(\mathbb{R}, Q) \subset TJ^1(\mathbb{R}, Q)$. Here $\boldsymbol{\beta}^\perp$ denotes the set of vectors annihilated by the one-forms in $\boldsymbol{\beta}$. Observe that $\mathcal{C}$ is exactly the subset of $TJ^1(\mathbb{R}, Q)$ in which $\boldsymbol{\beta}$-admissible special vector fields take their value. Now recall that ρ_1^*TQ is identifiable with the set of vertical vectors on the bundle $\sigma_1 \colon J^1(\mathbb{R}, Q) \to J^0(\mathbb{R}, Q)$ by (3.1). Thus we may regard $\rho_1^*TQ \subset TJ^1(\mathbb{R}, Q)$. Therefore we may define a subbundle of ρ_1^*TQ by intersection with $\boldsymbol{\beta}^\perp$. Denote this subbundle by $\bar{\mathcal{C}}$. In natural coordinates, elements $((t, q^i, v^j), u^k)$ of ρ_1^*TQ must satisfy the relations

$$\hat{\beta}_i^a u^i = 0, \qquad a = 1, \ldots, m \tag{4.2}$$

if they are to be in $\bar{\mathcal{C}}$. Comparing (4.1) with (4.2) we see that $\mathcal{C}$ is an affine subbundle of $J^2(\mathbb{R}, Q)$ modeled on the vector subbundle $\bar{\mathcal{C}}$ of ρ_1^*TQ. We define $i_{\bar{\mathcal{C}}} \colon \bar{\mathcal{C}} \to \rho_1^*TQ$ to be the inclusion map. Note that the bundle map

$$\begin{tikzcd} \bar{\mathcal{C}} \arrow[rr, "i_{\bar{\mathcal{C}}}"] \arrow[dr] & & \rho_1^*TQ \arrow[dl] \\ & J^1(\mathbb{R}, Q) & \end{tikzcd}$$

is linear on the fibres. We shall denote by $i_{\bar{\mathcal{C}}}^T \colon \rho_1^*T^*Q \to \bar{\mathcal{C}}^*$ the fibre-wise transpose of $i_{\bar{\mathcal{C}}}$.

A curve $c \colon [a, b] \to Q$ is said to *satisfy the constraint $\boldsymbol{\beta}$* if $(j^1c)'(t) \in \boldsymbol{\beta}^\perp_{j^1c(t)}$ for $t \in [a, b]$. In natural coordinates this implies that if c is given by

$$s \mapsto (q^1(s), \ldots, q^n(s)),$$

then c satisfies the constraint $\boldsymbol{\beta}$ if and only if

$$\beta_0^a + \beta_i^a \dot{q}^i + \hat{\beta}_i^a \ddot{q}^i = 0, \qquad a = 1, \ldots, m. \tag{4.3}$$

A force $F \colon J^1(\mathbb{R}, Q) \to \rho_1^*T^*Q$ is called a *$\boldsymbol{\beta}$-constraint force* if $F(v) \in \bar{\mathcal{C}}_v^0$ for each $v \in J^1(\mathbb{R}, Q)$. We see from (4.2) that in natural coordinates a force F given by $(t, q^i, v^j) \mapsto ((t, q^i, v^j), F_k)$ is a $\boldsymbol{\beta}$-constraint force if and only if F_k is a linear combination of $\hat{\beta}_k^1, \ldots, \hat{\beta}_k^m$ for $k = 1, \ldots, n$.

Affine Constraints

We shall define an *affine constraint* on Q to be an m-dimensional codistribution $\boldsymbol{\gamma}$ on $J^0(\mathbb{R}, Q)$. In natural coordinates a basis for $\boldsymbol{\gamma}$ may be written as

$$\gamma^a = \gamma_0^a(t,q)dt + \gamma_i^a(t,q)dq^i, \qquad a = 1, \ldots, m.$$

That the forms are linearly independent means that the $m \times n$ matrix with components γ_i^a has rank m. We may define a constraint of general type by defining $\boldsymbol{\beta}$ to be locally generated by the one-forms

$$\beta^a(t,q,v) = \boldsymbol{d}(\gamma_i^a(t,q)v^i), \qquad a = 1, \ldots, m.$$

If we expand these equations we obtain

$$\beta^a(t,q,v) = \left(\frac{\partial \gamma_0^a}{\partial t} + \frac{\partial \gamma_i^a}{\partial t}v^i\right)dt + \left(\frac{\partial \gamma_0^a}{\partial q^i} + \frac{\partial \gamma_j^a}{\partial q^i}v^j\right)dq^i + \gamma_i^a dv^i, \qquad a = 1, \ldots, m. \tag{4.4}$$

The linear independence of the one-forms γ^a is then easily seen to guarantee that the rank condition on the matrix $\hat{\beta}_i^a$ for general constraints is met.

In the affine case the solutions are restricted to the subset of $J^1(\mathbb{R}, Q)$ defined by the constraints. We denote this subset by $\mathcal{D}$. It is clear that $\mathcal{D}$ is an integral manifold of $\boldsymbol{\beta}$. This will become relevant in our discussion of the constrained Gibbs-Appell equations for affine constraints in Section 5.3. Also, notice that a $\boldsymbol{\beta}$-constraint force, when $\boldsymbol{\beta}$ is derived from an affine constraint, is a linear combination of the forms with components $\gamma_i^a, \ldots, \gamma_i^m$ in natural coordinates. The coefficients in the linear combination are classically the Lagrange multipliers.

4.4. Lagrange's Equations with External Forces and Constraints In the previous three sections we have defined the fundamental elements of Lagrangian mechanics. Now we precisely state what we shall mean by a solution of Lagrange's equations in the presence on external forces and constraints.

4.1 Definition: Let L be a regular Lagrangian on Q with Lagrangian vector field X_L. Let F be a force and let $\boldsymbol{\beta}$ be a constraint. A curve $c\colon [a,b] \to Q$ is a *solution of Lagrange's equations with force F and constraint $\boldsymbol{\beta}$* if c satisfies the constraint $\boldsymbol{\beta}$ and if there exists a $\boldsymbol{\beta}$-constraint force λ such that

$$j^2c(t) = X_L(j^1c(t)) - F^\sharp(j^1c(t)) - \lambda^\sharp(j^1c(t))$$

for each $t \in [a,b]$. □

In natural coordinates this definition gives the well-known equations

$$\ddot{q}^i = g^{ij}\left(\frac{\partial L}{\partial q^j} - \frac{\partial^2 L}{\partial v^j \partial q^k}\dot{q}^k - \frac{\partial^2 L}{\partial v^j \partial t}\right) + g^{ij}F_j(t,q,v) + g^{ij}\lambda_j(t,q,v) \tag{4.5}$$

which determine the motion when combined with the constraint equations (4.3).

5. The Gibbs-Appell Equations

In this section we present the Gibbs-Appell equations for regular Lagrangians. First we construct the Gibbs function in Section 5.1. The Gibbs-Appell equations in the absence of constraints are then easily presented in Section 5.2. The constrained case is dealt with in Section 5.3. We then present the example of Section 2 in our general framework to illustrate how the computations may be done.

5.1. The Gibbs Function In this section we present one of the main constructions of the paper. Namely, we define a generalised version of the Gibbs function which serves the purpose of the function given for the example of Section 2.

Let L be a regular Lagrangian on Q and let $a \in J^2(\mathbb{R}, Q)$. Note that we may naturally regard $X_L(\sigma_2(a)) - a$ as an element of $T_{\rho_2(a)}Q$. Therefore we can define the *Gibbs function* corresponding to L as

$$G_L(a) = \frac{1}{2} g(X_L(\sigma_2(a)) - a, X_L(\sigma_2(a)) - a). \tag{5.1}$$

In natural coordinates we have

$$G_L(t, q^i, v^j, a^k) = \frac{1}{2} g_{ij} a^i a^j + \left(\frac{\partial^2 L}{\partial v^i \partial q^j} v^j + \frac{\partial^2 L}{\partial v^i \partial t} - \frac{\partial L}{\partial q^i} \right) a^i + \frac{1}{2} g_{ij} X_L^i X_L^j.$$

Notice that the expression for X_L^i is independent of a^i. As we shall see in Sections 5.2 and 5.3, this means that the third term in the coordinate representation of the Gibbs function does not affect the equations of motion. Nevertheless, their appearance is necessary to make the function G_L well-defined on $J^2(\mathbb{R}, Q)$.

5.2. The Gibbs-Appell Equations in the Absence of Constraints It is an easy matter to write the Gibbs-Appell equations when no constraints are present and show that these equations are the same as Lagrange's equations.

5.1 THEOREM: *Let L be a regular Lagrangian and let F be a force. Let G_L be the Gibbs function defined by (5.1). A curve $c\colon [a, b] \to Q$ is a solution of Lagrange's equations if and only if c is a solution of the **Gibbs-Appell equations**,*

$$\boldsymbol{d_a} G_L(j^2 c(t)) = F(j^1 c(t)),$$

for each $t \in [a, b]$. Here we are regarding both $\boldsymbol{d_a} G_L$ and F as taking values in T^*Q.

Proof: The proof is easy in natural coordinates. If we compute $\boldsymbol{d_a} G_L$ we obtain

$$\boldsymbol{d_a} G_L = \left(\frac{\partial^2 L}{\partial v^i \partial v^j} a^j + \frac{\partial^2 L}{\partial v^i \partial q^j} v^j + \frac{\partial^2 L}{\partial v^i \partial t} - \frac{\partial L}{\partial q^i} \right) dq^i.$$

The Gibbs-Appell equations in natural coordinates are then

$$g_{ij} \ddot{q}^j + \frac{\partial^2 L}{\partial v^i \partial q^j} \dot{q}^j + \frac{\partial^2 L}{\partial v^i \partial t} - \frac{\partial L}{\partial q^i} = F_i, \qquad i = 1, \dots, n,$$

which are clearly equivalent to Lagrange's equations by inspecting (4.5). ■

5.2 REMARK: In Lewis [1995] Lagrange's equations are realised as the components of a semi-basic differential one-form, F_L, on the bundle $\rho_2\colon J^2(\mathbb{R},Q)\to Q$. In natural coordinates

$$F_L=\left(\frac{\partial^2 L}{\partial v^i\partial v^j}a^j+\frac{\partial^2 L}{\partial v^i\partial q^j}v^j+\frac{\partial^2 L}{\partial v^i\partial t}-\frac{\partial L}{\partial q^i}\right)dq^i.$$

The above theorem verifies that $F_L=\boldsymbol{d_a}G_L$. □

5.3. The Gibbs-Appell Equations in the Presence of Constraints Now we show that, properly interpreted, the Gibbs-Appell equations remain valid when constraints are imposed on the system. Let $\boldsymbol{\beta}$ be a constraint. Recall from Section 4.3 that $\boldsymbol{\beta}$ defines an affine subbundle $\mathcal{C}$ of $J^2(\mathbb{R},Q)$ modeled on a vector subbundle $\bar{\mathcal{C}}$ of ρ_1^*TQ. We can now use the construction given in Section 3.6 for the acceleration derivative restricted to affine subbundles. This defines $\boldsymbol{d}^{\mathcal{C}}_{\boldsymbol{a}}(G_L\mid\mathcal{C})\colon\mathcal{C}\to\bar{\mathcal{C}}^*$.

5.3 THEOREM: *Let L be a regular Lagrangian on Q, let F be a force, and let $\boldsymbol{\beta}$ be a constraint. Let G_L be the Gibbs function defined by (5.1). A curve $c\colon[a,b]\to Q$ is a solution of Lagrange's equations with force F and constraint $\boldsymbol{\beta}$ if and only if c satisfies the constraint $\boldsymbol{\beta}$ and if c is a solution of the* ***constrained Gibbs-Appell equations****,*

$$\boldsymbol{d}^{\mathcal{C}}_{\boldsymbol{a}}(G_L\mid\mathcal{C})(j^2c(t))=i^T_{\bar{\mathcal{C}}}(F(j^1c(t))),\tag{5.2}$$

for each $t\in[a,b]$. Here we are regarding F as taking values in $\rho_1^*T^*Q$.

Proof: We first state a simple linear algebra formula. Let E be a vector space with $F\subset E$ a subspace. Let $i\colon F\to E$ denote the inclusion and let $b\in E^*$. Then $i^*b(e)=(b\mid E)(e)$ for all $e\in F$. In particular, this implies that

$$\boldsymbol{d}^{\mathcal{C}}_{\boldsymbol{a}}(f\mid\mathcal{C})(a)=i^T_{\bar{\mathcal{C}}}\boldsymbol{d_a}f(a)\tag{5.3}$$

if $a\in\mathcal{C}$ and if f is a function on $J^2(\mathbb{R},Q)$.

By Theorem 5.1 we know that c is a solution of Lagrange's equations with force F and constraint $\boldsymbol{\beta}$ if and only if c satisfies the constraint $\boldsymbol{\beta}$ and if there exists a $\boldsymbol{\beta}$-constraint force λ so that

$$\boldsymbol{d_a}G_L(j^2c(t))=F(j^1c(t))+\lambda(j^1c(t)).$$

If we apply the mapping $i^T_{\bar{\mathcal{C}}}$ to the above equation we get the constrained Gibbs-Appell equations. Here we use (5.3) and the fact that $\lambda(j^1c(t))\in\bar{\mathcal{C}}^0_{j^1c(t)}$. Thus, if c is a solution of Lagrange's equations with force F and constraint $\boldsymbol{\beta}$, then it is a solution of the constrained Gibbs-Appell equations.

Now suppose that c is a solution of the constrained Gibbs-Appell equations. Then, using (5.3), we may write

$$i^T_{\bar{\mathcal{C}}}\boldsymbol{d_a}G_L(j^2c(t))=i^T_{\bar{\mathcal{C}}}(F(j^1c(t))).$$

This implies that $\boldsymbol{d_a}G_L(j^2c(t))-F(j^1c(t))\in\ker(i^T_{\bar{\mathcal{C}}})$. Thus

$$\boldsymbol{d_a}G_L(j^2c(t))-F(j^1c(t))=\lambda(j^1c(t))$$

for some force λ which is a section of $\boldsymbol{\beta}$. Thus c is a solution of Lagrange's equations with force F and constraint $\boldsymbol{\beta}$. ■

5.4 REMARKS:

1. Notice that in the above construction the "projection" by $i_{\bar{\mathcal{C}}}^T$ and the addition of the constraint force λ amount to the same thing. The constraint force may be viewed as a Lagrange multiplier and so we see here some suggestions of constrained variational methods. This is discussed further in Section 7.
2. If $M \subset J^1(\mathbb{R}, Q)$ is an integral manifold of $\boldsymbol{\beta}$, it is possible to restrict the construction of the constrained Gibbs-Appell equations to M. Thus we first consider the restricted bundle $J^2(\mathbb{R}, Q) \mid M$ which is an affine bundle over M modeled on the vector bundle $(\rho_1 \mid M)^*TQ$ over M. Then $\mathcal{C} \mid M$ is an affine subbundle of $J^2(\mathbb{R}, Q) \mid M$ modeled on $\bar{\mathcal{C}} \mid M$. Since the acceleration derivative only depends on the value of the function on the fibres of $J^2(\mathbb{R}, Q)$, it may still be defined in this restricted case. Similarly, the map $i_{\bar{\mathcal{C}}}^T$ is only defined on fibres and so also reduces to the case where we restrict to M. □

It is worth deriving the coordinate formulas for the constrained Gibbs-Appell equations. To do so let $(q^1, \ldots, q^n)$ be a coordinate chart for Q with $(t, q^1, \ldots, q^n, v^1, \ldots, v^n)$ the corresponding natural coordinates for $J^1(\mathbb{R}, Q)$. We coordinatise the fibres of $\mathcal{C} \subset J^2(\mathbb{R}, Q)$ in this chart with coordinates $\alpha^1, \ldots, \alpha^{n-m}$. As coordinates for the fibres of $\bar{\mathcal{C}} \subset \rho_1^*TQ$ we use $\nu^1, \ldots, \nu^{n-m}$. Therefore the inclusion of $\bar{\mathcal{C}}$ in ρ_1^*TQ may be written in our coordinates as

$$i_{\bar{\mathcal{C}}} \colon ((t, q^i, v^j), \nu^a) \mapsto ((t, q^i, v^j), u^k(\nu^a))$$

where the relations $u^k(\nu^a)$ are linear in ν^a. Thus we may write $u^k = A_a^k \nu^a$ where the matrix A_a^k is in general a function on $J^1(\mathbb{R}, Q)$. The constrained Gibbs-Appell equations in these coordinates then take the fairly familiar form

$$\frac{\partial (G_L \mid \mathcal{C})}{\partial \alpha^a} = A_a^i F_i. \tag{5.4}$$

By using the fact that the variables α^a are linear combinations of the a^is, we write (5.4) as $n - m$ second order differential equations on Q. These equations may be then combined with the m constraint equations (4.3) to obtain n second-order differential equations on Q which determine the motion.

When the constraints are affine we can give a bit more structure to the constrained Gibbs-Appell equations. In this case the solutions are restricted to $\mathcal{D} \subset J^1(\mathbb{R}, Q)$. We denote by $i_{\mathcal{D}} \colon \mathcal{D} \to J^1(\mathbb{R}, Q)$ the inclusion mapping. If we use coordinates (t, q^i, ν^a) for $\mathcal{D}$, we may write the inclusion as

$$i_{\mathcal{D}} \colon (t, q^i, \nu^a) \mapsto (t, q^i, v^j(\nu^a))$$

where the relations $v^j(\nu^a)$ are affine in ν^a so we may write $v^j = B_a^j \nu^a + r^j$. The matrix B_a^j and the vector r^j will in general be functions on $J^0(\mathbb{R}, Q)$. In practice, to determine B_a^j and r^j we observe that the variables ν^a are affine in v^j. Thus we may write $\nu^a = C_j^a v^j + s^a$. The affine system

$$\begin{aligned} C_j^a v^j &= \nu^a - s^a, \qquad & a &= 1, \ldots, n-m \\ \gamma_j^b v^j &= \gamma_0^b, \qquad & b &= 1, \ldots, m \end{aligned}$$

may then be solved to yield $v^j = B^j_a \nu^a + r^j$. The set $\mathcal{D}^{lift} \triangleq \mathrm{range}(Ti_{\mathcal{D}}) \cap J^2(\mathbb{R}, Q) \subset \mathcal{C}$ is invariant and so we may restrict G_L to $\mathcal{D}^{lift}$ and compute the constrained Gibbs-Appell equations on this smaller manifold as in Remark 5.4(2). In coordinates we have the inclusion of $\mathcal{D}^{lift}$ in $J^2(\mathbb{R}, Q)$ as

$$i_{\mathcal{D}^{lift}} \colon (t, q^i, \nu^k, \alpha^k) \mapsto \left(t, q^i, B^j_a \nu^a + r^j, B^k_b \alpha^b + \frac{\partial B^k_b}{\partial q^l} B^l_c \nu^c \nu^b + \frac{\partial B^k_b}{\partial t} \nu^b + \frac{\partial r^k}{\partial q^l} B^l_b \nu^b + \frac{\partial r^k}{\partial t} \right).$$

The forces which appear in the Gibbs-Appell equations may be computed in the same manner as in the general case and this gives

$$F_a = B^i_a F_i.$$

The Gibbs-Appell equations may then be computed just as in (5.4). However, in the affine case, the Gibbs function may be restricted to the smaller set $\mathcal{D}^{lift}$. We therefore expect to be able to write fewer then $2n$ differential equations to describe the system in local coordinates. Indeed, the relations $v^j = B^j_a \nu^a + r^j$ combine with (5.4) to give $2n - m$ first order differential equations the independent variables (q^i, ν^a) which determine the motion of the system on $\mathcal{D}$.

The equations (5.4) are indeed one of the most, if not the most, compact ways of representing the equations of motion of a system with constraints. This does not mean, however, that they are necessarily the most convenient or revealing. For a survey of other methods of writing the equations in the presence of constraints we refer the reader to Bloch, Krishnaprasad, Marsden, and Murray [1994].

5.4. The Example of Section 2 Revisited With the coordinate formulae of Section 5.3 we can now make better sense of the computations we performed in Section 2. Recall that the example was a particle of mass m in $Q = \mathbb{R}^3$ subject to the constraint

$$\dot{z} - y\dot{x} = 0.$$

We also considered the presence of an arbitrary force F on the system.

In our development the Lagrangian is given by

$$L = \frac{1}{2} m \left(v_x^2 + v_y^2 + v_z^2 \right)$$

and the affine constraint is defined by the one-form

$$\gamma^1 = dz - y dx.$$

An arbitrary force is given by

$$F = F_x dx + F_y dy + F_z dz$$

where the components F_x, F_y, F_z are functions on $J^1(\mathbb{R}, Q)$. Here we are thinking of the force as a map from $J^1(\mathbb{R}, Q)$ to T^*Q. An easy computation gives the Gibbs function as

$$G_L = \frac{1}{2} m \left(a_x^2 + a_y^2 + a_z^2 \right).$$

Now we turn to parameterising the necessary objects. As coordinates for $\mathcal{D} \subset J^1(\mathbb{R}, Q)$ we choose $(t, x, y, z, \nu^1 \triangleq v_y, \nu^2 \triangleq y v_x + v_z)$. The inclusion of $\mathcal{D}$ in $J^1(\mathbb{R}, Q)$ then looks like

$$i_{\mathcal{D}} \colon (t, x, y, z, \nu^1, \nu^2) \mapsto \left(t, x, y, z, \frac{1}{2y}\nu^2, \nu^1, \frac{1}{2}\nu^2\right).$$

With this we are able to compute the inclusion of $\mathcal{D}^{lift} = \mathrm{range}(Ti_{\mathcal{D}}) \cap J^2(\mathbb{R}, Q)$ in $J^2(\mathbb{R}, Q)$ to be

$$i_{\mathcal{C}} \colon (t, x, y, z, \nu^1, \nu^2, \alpha^1, \alpha^2) \mapsto \left(t, x, y, z, \frac{1}{2y}\nu^2, \nu^1, \frac{1}{2}\nu^2, \frac{1}{2y}\alpha^2 - \frac{1}{2y^2}\nu^1\nu^2, \alpha^1, \frac{1}{2}\alpha^2\right).$$

We then compute the restriction of G_L to $\mathcal{C}$ as

$$G_L \mid \mathcal{C} = \frac{1}{2}m\left(\frac{1}{4y^2}(\alpha^2)^2 + \frac{1}{4y^4}(\nu^1)^2(\nu^2)^2 - \frac{1}{2y^3}\alpha^2\nu^1\nu^2 + (\alpha^1)^2 + \frac{1}{4}(\alpha^2)^2\right).$$

To compute how the forces enter the equations of motion, note that the matrix B^i_a is given by

$$B = \begin{bmatrix} 0 & \frac{1}{2y} \\ 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

Thus we compute

$$F_1 = F_y, \qquad F_2 = \frac{1}{2y}F_x + \frac{1}{2}F_z.$$

The Gibbs-Appell equations are then easily computed to be

$$\begin{aligned} m\alpha^1 &= F_y \\ \left(\frac{m}{4y^2} + \frac{m}{4}\right)\alpha^2 - \frac{m}{4y^3}\nu^1\nu^2 &= \frac{1}{2y}F_x + \frac{1}{2}F_z. \end{aligned}$$

Of course we regard these as differential equations since $\alpha^i = \dot{\nu}^i$, $i = 1, 2$. On top of these equations we have the equations which define the constraints,

$$\dot{z} - y\dot{x} = 0,$$

and the equations which define the coordinates ν^1, ν^2,

$$\nu^1 = \dot{y}, \qquad \nu^2 = y\dot{x} + \dot{z}.$$

These may be inverted as in Section 2 to give $(\dot{x}, \dot{y}, \dot{z})$ in terms of (ν^1, ν^2). Of course the equations obtained here agree with the equations obtained in Section 2.

6. The Gibbs Function on Riemannian Manifolds

When the Lagrangian is given by the kinetic energy with respect to a Riemannian metric, Lagrange's equations yield the equations for geodesics. It turns out that in this case the Gibbs function has an interesting form when thought of as a function on TTQ.

Let (Q, g) be a Riemannian manifold. Recall that the Riemannian connection defines a bundle map $K\colon TTQ \to TQ$ such that the following diagram commutes.

$$\begin{tikzcd} TTQ \arrow[r, "K"] \arrow[d, "\tau_{TQ}"'] & TQ \arrow[d, "\tau_Q"] \\ TQ \arrow[r, "\tau_Q"'] & TQ \end{tikzcd}$$

The subbundle $HQ = \ker(K)$ has the property of making a complement to $VQ = \ker(T\tau_Q)$ in TTQ. In natural coordinates for TTQ the map K has the form

$$K(q^i, v^j, u^k, w^l) = (q^i, w^j + \Gamma^j_{kl} u^k v^l)$$

where Γ^j_{kl} are the *Christoffel symbols* for the Riemannian metric. On TTQ we may now define a function

$$\tilde{G}_g(w) = \frac{1}{2} g(K(w), K(w)).$$

In natural coordinates we have

$$\tilde{G}_g(q^i, v^j, u^k, w^l) = \frac{1}{2} g_{ij} w^i w^j + g_{ij} w^i \Gamma^j_{kl} u^k v^l + \frac{1}{2} g_{ij} \Gamma^i_{kl} \Gamma^j_{rs} u^k v^l u^r v^s.$$

Now recall the canonical involution of TTQ. This is a map J such that the following diagram commutes.

$$\begin{tikzcd} TTQ \arrow[rr, "J"] \arrow[dr, "T\tau_Q"'] & & TTQ \arrow[dl, "\tau_{TQ}"] \\ & TQ & \end{tikzcd}$$

In natural coordinates the map J has the form

$$J(q^i, v^j, u^k, w^l) = (q^i, u^j, v^k, w^l).$$

We shall denote by $\mathrm{Fix}(J)$ the set of points in TTQ which are fixed by J. For each $t \in \mathbb{R}$ there is a diffeomorphism from the two-jets of maps from $\mathbb{R}$ to Q whose source is t to $\mathrm{Fix}(J)$. In natural coordinates this diffeomorphism is given by

$$(t, q^i, v^j, a^k) \mapsto (q^i, v^j, v^k, a^l).$$

By restricting the function $\tilde{G}_g(w)$ on TTQ to $\mathrm{Fix}(J)$ we may define a function G_g on $J^2(\mathbb{R}, Q)$. This function has the form

$$G_g(t, q^i, v^j, a^k) = \frac{1}{2} g_{ij} a^i a^j + g_{ij} a^i \Gamma^j_{kl} v^k v^l + \frac{1}{2} g_{ij} v^i v^j + \frac{1}{2} g_{ij} \Gamma^i_{kl} \Gamma^j_{rs} v^k v^l v^r v^s.$$

in natural coordinates for $J^2(\mathbb{R}, Q)$. The following result is easily proven.

6.1 Proposition: *Let (Q, g) be a Riemannian manifold and define a Lagrangian by $L(v) = \frac{1}{2}g(v, v)$. Then $G_g = G_L$.*

Proof: The proof consists of a simple comparison of (5.1) with the definition of G_g in natural coordinates. ■

6.2 Remarks:

1. On a Riemannian manifold one may define a Riemannian metric on its tangent bundle called the *Sasaki metric* (see Sasaki [1958], Sasaki [1962]). This metric is defined by

$$\tilde{g}(w_1, w_2) = g(T\tau_Q(w_1), T\tau_Q(w_2)) + g(K(w_1), K(w_2)).$$

Notice that the Gibbs function for the geodesic system differs from the kinetic energy of the Sasaki metric exactly by the kinetic energy of the original system regarded as a function on $J^2(\mathbb{R}, Q)$. To be precise,

$$G_g(w) = \frac{1}{2}\tilde{g}(w, w) - \frac{1}{2}g(T\tau_Q(w), T\tau_Q(w)).$$

2. One can see why the expression "acceleration energy" is sometimes used for the Gibbs function. Indeed, the Gibbs function in the Riemannian case is simply one-half of the norm squared of the "geometric acceleration" with respect to the original metric.

3. In Giachetta [1992] it is shown that a special vector field (in particular a Lagrangian vector field) defines an Ehresmann connection on the bundle $\rho_1 \colon J^1(\mathbb{R}, Q) \to Q$. In the case when the special vector field is the Lagrangian vector field corresponding to the kinetic energy on a Riemannian manifold, this Ehresmann is easily seen to be the Riemannian connection corresponding to the metric. Therefore, we may ask whether the construction we performed above generalises to arbitrary regular Lagrangians. It turns out that it does not, and it is just in the Riemannian case that the construction of the Gibbs function in Section 5.1 agrees with the function G_g as we defined it above. □

7. Gauss' Principle of Least Constraint

Gauss' principle of least constraint roughly says that, among the accelerations which satisfy the given constraints and external forces, the one which determines the equations of motion is that which minimises some positive definite function of acceleration. We have the machinery developed which makes it easy to make this a precise statement for our general development.

7.1 Theorem: (Gauss' Principle of Least Constraint) *Let L be a regular Lagrangian on Q, let F be a force, and let $\boldsymbol{\beta}$ be a constraint. Let $\mathcal{C}$ be the subset of $J^2(\mathbb{R}, Q)$ defined by the constraints $\boldsymbol{\beta}$. Fix $v \in J^1(\mathbb{R}, Q)_{t,q}$. Suppose that $c\colon [a, b] \to Q$ is a solution of*

Lagrange's equations through v with force F and constraint $\boldsymbol{\beta}$. Then $j^2c(t)$ is the unique element of $(\sigma_2 \mid \mathcal{C})^{-1}(v)$ which minimises the function

$$G(L,F)(a) = g(X_L(v) - F^\sharp(v) - a, X_L(v) - F^\sharp(v) - a)$$

restricted to $\mathcal{C} \subset J^2(\mathbb{R},Q)$. Here we are thinking of $X_L - F^\sharp$ as a special vector field taking its values in $J^2(\mathbb{R},Q)$.

Proof: Note that the function $G(L,F) \mid \mathcal{C}$ is a smooth convex function on the affine space $\mathcal{C}_v$. Therefore, its minimum will be the unique point where its derivative vanishes. That is to say, $G(L,F)$ has its unique minimum at the point $a \in \mathcal{C}_v$ where $\boldsymbol{d}^{\mathcal{C}}_{\boldsymbol{a}}(G(L,F) \mid \mathcal{C})(a) = 0$. By (5.3) we have

$$\boldsymbol{d}^{\mathcal{C}}_{\boldsymbol{a}}(G(L,F) \mid \mathcal{C})(a) = i^T_{\mathcal{C}}\boldsymbol{d}_{\boldsymbol{a}}G(L,F)(a)$$

for $a \in \mathcal{C}$. In natural coordinates we have

$$G(L,F) = g_{ij}\left(a^i a^j - 2a^i\left(X^j_L - g^{jk}F_k\right) + (X^i_L - g^{ik}F_k)(X^j_L - g^{jl}F_l)\right).$$

We may compute

$$\boldsymbol{d}_{\boldsymbol{a}}G(L,F)(a) = 2\left(\boldsymbol{d}_{\boldsymbol{a}}G_L(a) - F(v)\right).$$

This implies that

$$\boldsymbol{d}^{\mathcal{C}}_{\boldsymbol{a}}(G(L,F) \mid \mathcal{C})(a) = 0 \quad \iff \quad \boldsymbol{d}^{\mathcal{C}}_{\boldsymbol{a}}(G_L)(a) = i^T_{\mathcal{C}}(F(v))$$

which proves the theorem by Theorem 5.3. ■

7.2 Remarks:

1. Observe that when no constraints are present, then we may select $a \in \sigma_2^{-1}(v)$ so that the absolute minimum value, zero, for $G(L,F)$ is attained. In this case we immediately obtain the equations of motion $j^2c(t) = X_L(j^1c(t)) - F(j^1c(t))$.

2. In Lanczos [1970] Gauss' principle is related to d'Alembert's principle. Here it is pointed out that Gauss' principle makes a variational principle out of the non-variational d'Alembert's principle. Indeed, Gauss' principle of least constraint can be seen to take an infinite-dimensional problem in the calculus of variations and make it into a finite-dimensional one. Furthermore, the finite-dimensional constrained variational problem on $\sigma_2^{-1}(v)$ is a simple one since the function to be minimised is convex. See Lewis and Murray [1995] for a discussion of variational methods for systems with constraints. □

8. Conclusions

In this paper we have generalised the Gibbs-Appell equations for particles and rigid bodies to general (regular) Lagrangians. Our formulation is shown to give the same equations of motion as Lagrange's equations in the presence of general classes of forces and constraints. When the Lagrangian is kinetic energy with respect to a Riemannian metric, the Gibbs function is shown to be related to the Sasaki metric on the tangent bundle of the configuration manifold. We are able to use our formulation to give a general version of Gauss' principle of least constraint.

References

Appell, P. 1900a. Développements sur une forme nouvelle des équations de la Dynamique. *Journal de Mathématiques Pures et Appliquées. Neuvième Sér*, **6**, 5–40.

Appell, P. 1900b. Sur une forme générale des équations de la Dynamique. *Journal für die Reine und Angewandte Mathematik*, **121**, 310–319.

Bloch, Anthony M., Krishnaprasad, P. S., Marsden, Jerrold E., and Murray, Richard M. 1994. *Nonholonomic Mechanical Systems with Symmetry*. Technical report CDS 94-013. Caltech Department of Control and Dynamical Systems. Submitted to Archives for Rational Mechanics and Analysis. Available electronically from `http://avalon.caltech.edu/cds`.

Desloge, Edward A. 1988. The Gibbs-Appell equations of motion. *American Journal of Physics*, **56**(9), 841–846.

Gauss, C. F. 1829. Uber ein neues allgemeines grundgesatz der mechanik. *Journal für die Reine und Angewandte Mathematik*, **4**, 232–235.

Giachetta, Giovanni. 1992. Jet methods in nonholonomic mechanics. *Journal of Mathematical Physics*, **33**(5), 1652–1665.

Gibbs, J. Willard. 1879. On the fundamental formulae of dynamics. *American Journal of Mathematics*, **II**, 49–64.

Goldschmidt, Hubert L. 1967. Integrability criteria for systems of nonlinear partial differential equations. *Journal of Differential Geometry*, **1**, 269–307.

Goldschmidt, Hubert L., and Sternberg, Schlomo. 1973. The Hamilton-Jacobi formalism in the calculus of variations. *Université de Grenoble. Annales de l'Institut Fourier*, **23**(1), 203–267.

Golubitsky, Martin, and Guillemin, Victor. 1973. *Stable Mappings and Their Singularities*. New York-Heidelberg-Berlin: Springer-Verlag.

Hermann, Robert. 1982. The differential geometric structure of general mechanical systems from the Lagrangian point of view. *Journal of Mathematical Physics*, **23**, 2077–2089.

Kalaba, Robert E., and Udwadia, Firdaus E. 1993. Lagrangian mechanics, Gauss's principle, quadratic programming and generalized inverses: New equations of motion for nonholonomically constrained discrete mechanical systems. *Quarterly of Applied Mathematics*, **52**(2), 229–241.

Kane, Thomas R. 1983. Formulations of dynamical equations of motion. *American Journal of Physics*, **51**(11), 974–977.

Lanczos, Cornelius. 1970. *The Variational Principles of Mechanics.* Fourth edition. 30 Lesmill Road, Don Mills, Toronto, Canada: University of Toronto Press.

Lewis, Andrew D. 1995. *Aspects of Geometric Mechanics and Control of Mechanical Systems.* Ph.D. thesis, California Institute of Technology.

Lewis, Andrew D., and Murray, Richard M. 1995. Variational principles for constrained systems: Theory and experiment. *International Journal of Nonlinear Mechanics.* To appear.

Ray, John R. 1992. Geometry of constraints and the Gauss-Appell principle of least constraint. *The Journal of the University of Kuwait. Science*, **19**, 11–15.

Sasaki, Shigeo. 1958. On the differential geometry of tangent bundles of Riemannian manifolds. *The Tôhoku Mathematical Journal. Second Series*, **10**, 338–354.

Sasaki, Shigeo. 1962. On the differential geometry of tangent bundles of Riemannian manifolds II. *The Tôhoku Mathematical Journal. Second Series*, **14**, 146–155.

Sharf, I., D'eleuterio, G. M. T., and Hughes, P. C. 1992. On the dynamics of Gibbs, Appell, and Kane. *European Journal of Mechanics. A/Solids*, **11**, 145–155.

Townsend, Miles A. 1991. Equivalence of Kane's, Gibbs-Appell's and Lagrange's equations. *Journal of Guidance*, **15**(5), 1289–1292.

Udwadia, Firdaus E., and Kalaba, Robert E. 1992. A new perspective on constrained motion. *Proceedings of the Royal Society. London. Series A. Mathematical and Physical Sciences*, **439**(1906), 407–410.