# Communications: On the linear response of mechanical systems with constraints

Carsten Hartmann,[1,a)] Christof Schütte,[1] and Giovanni Ciccotti[2]
[1]*Institut für Mathematik, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Dipartimento di Fisica, Università "La Sapienza," 00185 Roma, Italy*

(Received 19 November 2009; accepted 14 February 2010; published online 16 March 2010)

We revisit the problem of the linear response of a constrained mechanical system. In doing so, we show that the standard expressions of Green and Kubo carry over to the constrained case without any alteration. The argument is based on the appropriate definition of constrained expectations by means of which Liouville's theorem and the Green–Kubo relations naturally follow. © *2010 American Institute of Physics*. [doi:10.1063/1.3354126]

Linear response theory as originally formulated by Green[1] and Kubo[2,3] provides a tool to describe the response of a mechanical system to a small external perturbation. The typical derivation of the linear response is based on the Hamiltonian function generating the underlying dynamics.[4] In many instances, however, the equations of motion are only available in a non-Hamiltonian form, although the system is of mechanical origin. One such case is a constrained mechanical system in Cartesian (ambient-space) coordinates with explicit Lagrange multipliers, in which case a Hamiltonian formulation is not obvious.[5]

Steps toward a formulation of constrained systems in the framework of non-Hamiltonian statistical mechanics have been taken in Ref. 6. Therein, the authors demonstrate that it is still possible to derive a response result that has the familiar form, while involving additional terms that are hard to interpret and which are attributed to an apparent nonzero phase space compressibility.

In this communication, we argue that in the presence of constraints, the standard linear response result *does* in fact hold—unambigously and without any alteration from the unconstrained case. Our reasoning is based on an appropriate definition of constrained expectation values that gives rise to the standard Liouville equation for probability densities and Liouville's theorem. The linear response result of Green and Kubo then naturally follows from these ingredients.

## I. CONSTRAINED DYNAMICS

Consider a particle of unit mass assuming states $x \in \mathbf{R}^n$ with potential energy $V(x)$. The dynamics are confined to a hypersurface $\mathcal{S} \subset \mathbf{R}^n$ that is defined by

$$\mathcal{S} = \{x \in \mathbf{R}^n : \varphi(x) = 0\},$$

with a scalar function $\varphi$ and $|\nabla\varphi| \neq 0$ everywhere on $\mathcal{S}$. Newton's equations for the particle read

$$\ddot{x}_t = -\nabla V(x_t) - \lambda \nabla \varphi(x_t), \quad \varphi(x_t) = 0, \tag{1}$$

with initial values $x_0 = x$ and $\dot{x}_0 = v$ satisfying

$$\varphi(x) = 0, \quad v \cdot \nabla\varphi(x) = 0. \tag{2}$$

Here and in the following, we use the notation $x \cdot y = x^T y$ and $A:B = \mathrm{tr}(AB)$ to denote the inner products between vectors $x$ and $y$ or the double contraction of second-order tensors $A$ and $B$.

In Eq. (1), we may easily eliminate the Lagrange multiplier $\lambda$ by differentiating the constraint $\varphi(x_t)=0$ twice with respect to time. This yields

$$\ddot{x}_t \cdot \nabla\varphi(x_t) + \dot{x}_t \otimes \dot{x}_t : \nabla\nabla\varphi(x_t) = 0,$$

where we use the notation $\nabla\nabla = \nabla \otimes \nabla$ to denote the matrix of second derivatives. Hence, with Eq. (1),

$$\lambda \nabla \varphi(x) = -(\nabla V(x))^{\perp} + II(\dot{x},\dot{x}),$$

with the abbreviations

$$(\nabla V)^{\perp} = \frac{\nabla\varphi \cdot \nabla V}{|\nabla\varphi|^2} \nabla \varphi, \quad II(v,v) = \frac{v \otimes v : \nabla\nabla\varphi}{|\nabla\varphi|^2} \nabla \varphi.$$

Inserting the expression for the constraint force into the equations of motion, Eq. (1) then gives

$$\ddot{x}_t = -(P \nabla V)(x_t) - II(\dot{x}_t,\dot{x}_t), \tag{3}$$

with the notation $PX = X - X^{\perp}$. Equation (3) is called an ambient-space formulation of the differential-algebraic system Eq. (1) as it is formulated in terms of the ambient-space coordinates $(x,v)$ on $\mathbf{R}^{2n}$ rather than generalized coordinates, and it does not involve Lagrange multipliers. Introducing the new variable $v = \dot{x}$ it can be recast as

$$\dot{x}_t = v_t, \quad \dot{v}_t = -(P \nabla V)(x_t) - II(v_t,v_t). \tag{4}$$

Given initial values satisfying Eq. (2), its solution automatically stays on the set

$$\mathcal{M} = \{(x,v) \in \mathbf{R}^{2n} : \varphi(x) = 0, v \cdot \nabla\varphi(x) = 0\}.$$

## II. LIOUVILLE EQUATION

Obviously, Eq. (4) is not Hamiltonian although it can be shown to be equivalent to a Hamiltonian system (e.g., by

a)Author to whom correspondence should be addressed. Electronic mail: chartman@mi.fu-berlin.de.

 

using generalized coordinates). Nonetheless, its flow shares basic properties of a Hamiltonian system such as being volume-preserving. First of all, note that

$$d\mu(z) = |\nabla \varphi(z)|^2 \delta(\varphi(z)) \delta(\dot{\varphi}(z)) dz \tag{5}$$

is the natural Liouville measure on $\mathcal{M}$ expressed in terms of the ambient-space coordinates $z=(x,v)$ (see Ref. 7 for details). Now, given an observable $f(z)$, we call

$$\langle f \rangle_\rho = \int_{\mathbf{R}^{2n}} f(z) \rho(z,t) d\mu(z)$$

the expectation of $f$ with respect to the (possibly time-dependent) probability density $\rho$. It is convenient to write Eq. (4) as the differential equation

$$\dot{z}_t = B(z_t), \tag{6}$$

with the vector field $B=(v,-P\nabla V-II)^T$ and the solution $z_t=z_t(z)$, $z_0=z$. The Liouville operator corresponding to Eq. (6) then assumes the standard form

$$L = B(z) \cdot \nabla, \tag{7}$$

whereas before and unless noted otherwise, $\nabla$ denotes the derivative with respect to the argument (here: $z$).

### A. Stationary distribution

Before we can state Liouville's theorem, we have to say what we mean by an invariant distribution. We employ what is known as the *Heisenberg picture* in quantum mechanics and call a distribution $\rho$ invariant under Eq. (6) if the expectation of an observable at time $t$

$$f(z_t) = \exp(Lt) f(z)$$

is stationary, i.e., if

$$\langle f(z_t) \rangle_\rho = \langle f(z_0) \rangle_\rho, \tag{8}$$

where the average is understood over the initial values. By construction, $f$ solves the differential equation

$$\frac{d}{dt} f(z_t) = Lf(z_t), \quad f(z_0) = f, \tag{9}$$

so stationarity of expectation values is equivalent to say that $\rho$ is a stationary solution of the Liouville equation

$$\frac{\partial}{\partial t} \rho(z,t) = L^* \rho(z,t), \quad \rho(z,0) = \rho_0(z). \tag{10}$$

Here, $\rho_0$ is understood as a probability density with respect to the constrained Liouville measure $\mu$, i.e.,

$$\int_{\mathbf{R}^{2n}} \rho_0 d\mu = 1,$$

and the adjoint $L^*$ is defined by means of the duality relation $(Lf,g)=(f,L^*g)$ with the natural scalar product $(f,g)=\langle fg \rangle_{\rho=1}$ between functions $f$ and $g$.

The Liouville Eq. (10) is in fact the usual one for $(Lf,g)=-(f,Lg)$ or $L^*=-L$, respectively. To see this, it is helpful to note that (cf. Ref. 6, p. 750)

$$\int_{\mathbf{R}^{2n}} (Lf) g |\nabla \varphi|^2 dz = -\int_{\mathbf{R}^{2n}} (Lg) f |\nabla \varphi|^2 dz,$$

which can be seen upon expanding the Liouvillian term by term and integrating by parts

$$\begin{aligned}
&\int_{\mathbf{R}^{2n}} (Lf) g |\nabla \varphi|^2 dz \\
&= \int_{\mathbf{R}^{2n}} (v \cdot \nabla_x f - (P \nabla V + II) \cdot \nabla_v f) g |\nabla \varphi|^2 dz \\
&= -\int_{\mathbf{R}^{2n}} \Big( v \cdot \nabla_x g - 2 \frac{g}{|\nabla \varphi|^2} v \otimes \nabla \varphi : \nabla\nabla\varphi - g \nabla_v \cdot II \\
&\quad - (P \nabla V + II) \cdot \nabla_v g \Big) f |\nabla \varphi|^2 dz \\
&= -\int_{\mathbf{R}^{2n}} (v \cdot \nabla_x g - (P \nabla V + II) \cdot \nabla_v g) f |\nabla \varphi|^2 dx \\
&= -\int_{\mathbf{R}^{2n}} (Lg) f |\nabla \varphi|^2 dz,
\end{aligned}$$

where in the third line we have used that $|\nabla\varphi|^2 \nabla_v \cdot II = 2v \otimes \nabla\varphi : \nabla\nabla\varphi$.

Hence, it follows that

$$\int_{\mathbf{R}^{2n}} (Lf) g d\mu = -\int_{\mathbf{R}^{2n}} (Lg) f d\mu + \int_{\mathbf{R}^{2n}} (L\{\delta(\varphi)\delta(\dot{\varphi})\}) fg |\nabla\varphi|^2 dz$$

with

$$\begin{aligned}
&\int_{\mathbf{R}^{2n}} (L\{\delta(\varphi)\delta(\dot{\varphi})\}) fg |\nabla\varphi|^2 dz \\
&= \int_{\mathbf{R}^{2n}} (v \cdot \nabla_x \{\delta(\varphi)\delta(\dot{\varphi})\} - (P \nabla V + II) \\
&\quad \cdot \nabla_v \{\delta(\varphi)\delta(\dot{\varphi})\}) fg |\nabla\varphi|^2 dz \\
&= \int_{\mathbf{R}^{2n}} (v \cdot \nabla\varphi \delta'(\varphi)\delta(\dot{\varphi}) + v \otimes v : \nabla\nabla\varphi \delta(\varphi)\delta'(\dot{\varphi}) \\
&\quad - (P \nabla V + II) \cdot \nabla\varphi \delta(\varphi)\delta'(\dot{\varphi})) fg |\nabla\varphi|^2 dz \\
&= \int_{\mathbf{R}^{2n}} (v \cdot \nabla\varphi \delta'(\varphi)\delta(\dot{\varphi}) \\
&\quad - (P \nabla V) \cdot \nabla\varphi \delta(\varphi)\delta'(\dot{\varphi})) fg |\nabla\varphi|^2 dz.
\end{aligned}$$

Here, the second equality follows from $II \cdot \nabla\varphi = v \otimes v : \nabla\nabla\varphi$. Finally $\dot{\varphi} = v \cdot \nabla\varphi$ with $\dot{\varphi}=0$ and $(P\nabla V)\cdot\nabla\varphi=0$, so the last integral vanishes which proves that

$$\int_{\mathbf{R}^{2n}} (Lf) g d\mu = -\int_{\mathbf{R}^{2n}} (Lg) f d\mu, \tag{11}$$

or, in other words, $L^*=-L$.

Now, Liouville's theorem readily follows: Let $A \neq \emptyset$ be any compact subset of $\mathcal{M}$. Then

$$\mu(A) = \int_{\mathbf{R}^{2n}} \chi_A d\mu \quad \text{with} \quad \chi_A(z) = \begin{cases} 1 & \text{if } z \in A \\ 0 & \text{if } z \notin A \end{cases}$$

is the phase space volume of $A$, and we have to show that $d\mu(A)/dt=0$ under the flow of Eq. (6).

To this end, we choose $f=\chi_A$ as initial condition in Eq. (9) and take the time derivative of Eq. (8) with $\rho=1$. Since $f$ solves Eq. (9), it remains to show that

$$\int_{\mathbf{R}^{2n}} Lf d\mu = 0.$$

But the integral equals $(Lf,\mathbf{1})=-(f,L\mathbf{1})$ where $L\mathbf{1}=0$. Hence the integral vanishes and, since $A$ is arbitrary, we have proved conservation of volume.

Before we come to the formulation of the linear response, a final remark is in order: it is a common fallacy that the system Eq. (6) was not volume-preserving because its ambient-space divergence is not zero. Indeed, a vector field $B$ is volume-preserving if and only if it is divergence-free.[8] But here, the solutions generated by the vector field $B$ live on the constrained phase space $\mathcal{M}\subset\mathbf{R}^{2n}$, so the appropriate notion of divergence is the divergence on $\mathcal{M}$, whereas the (unconstrained) divergence in the ambient-space variables does not tell us much about volume-preservation.

## III. LINEAR RESPONSE

In Eq. (6), we add a small perturbation in the way that

$$\dot{z}_t^\epsilon = B(z_t^\epsilon) + \epsilon G(z_t^\epsilon)u_t, \quad \epsilon \ll 1, \tag{12}$$

where $u$ is a time-dependent scalar forcing (not necessarily smooth) and $G$ is a Hamiltonian vector field compatible with the constraint (in other words, the form of $G$ resembles that of $B$). Now set

$$L_{\text{eq}} = B(z)\cdot\nabla, \quad L_t = u_t G(z)\cdot\nabla.$$

The Liouville equation associated with Eq. (12) then reads

$$\frac{\partial}{\partial t}\rho^\epsilon(z,t) = (L_{\text{eq}}^* + \epsilon L_t^*)\rho^\epsilon(z,t), \quad \rho^\epsilon(z,0) = \rho_0(z),$$

with $L_{\text{eq}}^*=-L_{\text{eq}}$ and $L_t^*=-L_t$; the latter follows *mutatis mutandis* from Eq. (11). Applying variation of constants or Dyson's formula yields the formal solution

$$\rho^\epsilon(z,t) = \exp(tL_{\text{eq}}^*)\rho_0(z) + \epsilon\int_0^t \exp((t-s)L_{\text{eq}}^*)L_s^*\rho^\epsilon(z,s)ds,$$

for which we seek a perturbative expansion of the form

$$\rho^\epsilon = \rho_0 + \epsilon\rho_1 + \epsilon^2\rho_2 + \dots. \tag{13}$$

Let us suppose that the initial distribution $\rho(z,0)=\rho_0(z)$ is independent of $\epsilon$ and is invariant under the unperturbed dynamics, i.e., $L_{\text{eq}}^*\rho_0=0$. By plugging the ansatz Eq. (13) into the Liouville equation and equating equal powers of $\epsilon$, we recover the $\mathcal{O}(\epsilon)$-approximation

$$\rho^\epsilon(z,t) \approx \rho_0(z) + \epsilon\int_0^t \exp((t-s)L_{\text{eq}}^*)L_s^*\rho_0(z)ds.$$

Now, let

$$H:\mathbf{R}^{2n}\to\mathbf{R}, \quad H = \frac{1}{2}|v|^2 + V$$

denote the Hamiltonian of the unperturbed system. Clearly, $H$ is preserved for $\epsilon=0$. The perturbed system Eq. (12) for $\epsilon>0$ obeys the energy balance

$$\frac{d}{dt}H(z_t^\epsilon) = -\epsilon u_t J(z_t^\epsilon), \tag{14}$$

with $z_t^\epsilon=z_t^\epsilon(z)$ denoting the solution of Eq. (12) and $J(z)=G(z)\cdot\nabla H(z)$ being the dissipative flux.

### A. Green-Kubo relations

We come to our main result that generalizes the classical result of Green[1] and Kubo[2,3] for Hamiltonian systems to non-Hamiltonian systems of the form (1).

In Eq. (12), suppose that $u_t=\delta(t)$, i.e., the perturbation is impulsive at $t=0$, and the initial values $z_0^\epsilon=z$ are drawn from the canonical distribution

$$\rho_0(z) \propto \exp(-\beta H(z)), \quad \int_{\mathbf{R}^{2n}} \rho_0(z)d\mu(z) = 1.$$

Moreover, let

$$\langle f\rangle_{\rho^\epsilon} = \int_{\mathbf{R}^{2n}} f(z)\rho^\epsilon(z,t)d\mu(z)$$

denote the expectation with respect to the probability distribution $\rho^\epsilon$. By replacing $\rho^\epsilon$ in the last equation by its $\mathcal{O}(\epsilon)$-approximation we find

$$\langle f\rangle_{\rho^\epsilon} \approx \langle f\rangle_{\rho_0} + \epsilon\int_{\mathbf{R}^{2n}}\int_0^t f\exp((t-s)L_{\text{eq}}^*)L_s^*\rho_0 ds d\mu.$$

Using $u_t=\delta(t)$ in the expression for $L_t^*=-L_t$, the double integral can be recast as

$$\begin{aligned}
&\int_{\mathbf{R}^{2n}}\int_0^t f\exp((t-s)L_{\text{eq}}^*)L_s^*\rho_0 ds d\mu \\
&= -\int_{\mathbf{R}^{2n}}\int_0^t f\exp((t-s)L_{\text{eq}}^*)L_s\rho_0 ds d\mu \\
&= -\int_{\mathbf{R}^{2n}}\int_0^t (\exp((t-s)L_{\text{eq}})f)(G\cdot\nabla\rho_0)\delta(s) ds d\mu \\
&= -\int_{\mathbf{R}^{2n}} (\exp(tL_{\text{eq}})f)(G\cdot\nabla\rho_0) d\mu \\
&= \beta\int_{\mathbf{R}^{2n}} f(z_t)(G\cdot\nabla H)\rho_0 d\mu,
\end{aligned}$$

where the last equality is due to $\rho_0\propto\exp(-\beta H)$. Employing the definition of the dissipative flux $J(z)=-G(z)\cdot\nabla H(z)$, it follows that

$$\int_{\mathbf{R}^{2n}} f(G\cdot\nabla H)\rho_0 d\mu = -\langle Jf(z_t)\rangle_{\rho_0}$$

with $z_t = z_t(z)$ denoting the solution to the unperturbed problem. As a consequence we recover the classical linear response result of Green and Kubo, viz.,

$$\langle f\rangle_{\rho^\epsilon} \approx \langle f\rangle_{\rho_0} - \epsilon\beta\langle Jf(z_t)\rangle_{\rho_0}. \tag{15}$$

## IV. CONCLUSION

In this communication, we have demonstrated that the common practice, namely, employing the standard linear response result by Green and Kubo in case of a system with holonomic constraints is indeed justified, although some articles that can be found in the literature suggest otherwise (e.g., Ref. 9). Although there has been no doubt that, in principle, Hamiltonian systems that are subject to holonomic constraints behave like any other natural mechanical system from the viewpoint of statistical mechanics (which is certainly not true for nonholonomic systems), the result closes a loophole in the statistical mechanics of constrained systems involving Lagrange multipliers. Consequently, the reader using linear response theory should not worry as to whether his system is subject to constraints or not.

This clearly raises the question whether the result carries over to systems that are either inherently non-Hamiltonian or which involve non-Hamiltonian perturbations. A typical representative of the first class is Nosé–Hoover dynamics while the SLLOD equations of motion belong to the second category. We stress that the ingredients we have employed to derive the Green–Kubo relations are relatively simple and are not at all tied to a Hamiltonian framework: a perturbation consistent with the equations of motion (e.g., divergence-free), a preserved measure that induces an inner product via an expectation and a formal expansion of the evolution equation for the corresponding probability densities. Addressing general dynamical systems with arbitrary non-Hamiltonian perturbations, however, is beyond the scope of this communication.

In the course of the derivation, we have also revisited the misleading statement (see, e.g., Refs. 10 and 11) that flows of constrained systems with nonvanishing ambient-space compressibility do not conserve volume. This, in fact wrong, statement is based on the misconception of taking the ambient-space divergence of a constrained vector field as indicative of being volume-preserving. The reader should also not worry about this issue.

Last but not least, all the results in this paper easily generalize to the case of multiple constraints, nontrivial mass matrices or more complicated types of perturbations (clearly, being compatible with the constraints). For the sake of readability, we refrain from presenting our results in such generality and leave it to the interested reader to fill this gap.

This work is supported by the DFG Research Center MATHEON "Mathematics for Key Technologies" in Berlin.

[1] M. S. Green, J. Chem. Phys. **22**, 398 (1954).
[2] R. Kubo, M. Yokota, and S. Nakajima, J. Phys. Soc. Jpn. **12**, 570 (1957).
[3] R. Kubo, M. Yokota, and S. Nakajima, J. Phys. Soc. Jpn. **12**, 1203 (1957).
[4] K. Hayashi and S.-I. Sasa, Physica A **370**, 407 (2006).
[5] M. E. Tuckerman, Y. Liu, G. Ciccotti, and G. J. Martyna, J. Chem. Phys. **115**, 1678 (2001).
[6] G. Ciccotti, R. Kapral, and A. Sergi, in *Handbook of Materials Modeling*, edited by S. Yip (Springer, Berlin, 2005), pp. 745–761.
[7] P. Ryckaert and G. Ciccotti, J. Chem. Phys. **78**, 7368 (1983).
[8] L. Simon, *Lectures on Geometric Measure Theory* (Proceedings of the Centre for Mathematical Analysis, Australian National University, Canberra, 1984).
[9] A. Sergi, Phys. Rev. E **67**, 021101 (2003).
[10] A. Sergi, Phys. Rev. E **69**, 021109 (2004).
[11] L. Hernández de la Pena, R. van Zon, J. Schofield, and S. B. Oppsi, J. Chem. Phys. **126**, 074105 (2007).