PHYSICAL REVIEW LETTERS

VOLUME 81 | 21 SEPTEMBER 1998 | NUMBER 12

Unified Approach to Hamiltonian Systems, Poisson Systems, Gradient Systems, and Systems with Lyapunov Functions or First Integrals

Robert I. McLachlan
Mathematics Department, Massey University, Palmerston North, New Zealand

G. R. W. Quispel
Department of Mathematics, LaTrobe University, Bundoora, Melbourne 3083, Australia

Nicolas Robidoux
Mathematics Department, Massey University, Palmerston North, New Zealand

(Received 21 April 1998)

We show that systems with a first integral (i.e., a constant of motion) or a Lyapunov function can be written as "linear-gradient systems," $\dot{x} = L(x)\nabla V(x)$, for an appropriate matrix function L, with a generalization to several integrals or Lyapunov functions. The discrete-time analog, $\Delta x / \Delta t = L\overline{\nabla} V$, where $\overline{\nabla}$ is a "discrete gradient," preserves V as an integral or Lyapunov function, respectively. [S0031-9007(98)07076-8]

PACS numbers: 03.20.+i, 02.60.Lj

I. INTRODUCTION

Integrals and Lyapunov functions—quantities that are conserved or dissipated, respectively—are fundamental in dynamics. They severely constrain the system's evolution and can be used to establish stability. There is no universal method to find such quantities, but if they are known (e.g., on physical grounds), we show that the system can be presented in a universal form which makes the conservation (dissipation) property manifest. Although elementary, this result is very general and will find many applications: Here we use it to preserve the conservation (dissipation) property under time discretization.

We start with the definition and an example of each of the classes of systems covered in this Letter.

(i) Hamiltonian systems: Hamiltonian systems are ubiquitous in physics [1]. They have the form $\dot{x} = J\nabla V(x)$; $x \in \mathbb{R}^{2n}$, where $V(x)$ denotes the Hamiltonian function; and $J := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}$, where Id denotes the identity matrix in $\mathbb{R}^n$.

Example 1. A simple Hamiltonian system is the pendulum [1] $\dot{x}_1 = x_2$, $\dot{x}_2 = -\sin(x_1)$; here $n = 1$ and $V(x_1, x_2) = \frac{1}{2}x_2^2 - \cos(x_1)$.

(ii) Poisson systems: Poisson systems also occur very frequently in physics ([1], App. 14; [2]). They have the form $\dot{x} = \Omega(x)\nabla V(x)$, $x \in \mathbb{R}^n$, where $V(x)$ again denotes the Hamiltonian function and the Poisson structure $\Omega(x)$ is an antisymmetric matrix $[\Omega^t(x) = -\Omega(x)]$, satisfying the Jacobi identity $\Omega_{jk}\partial_k\Omega_{\ell m} + \Omega_{\ell k}\partial_k\Omega_{mj} + \Omega_{mk}\partial_k\Omega_{j\ell} = 0$.

Example 2. The equations of motion of a free rigid body with moments of inertia I_1, I_2, and I_3 form a Poisson system with angular momentum $x \in \mathbb{R}^3$ and Poisson structure

$$\Omega(x) = \begin{pmatrix} 0 & x_3 & -x_2 \\ -x_3 & 0 & x_1 \\ x_2 & -x_1 & 0 \end{pmatrix} \tag{1}$$

and Hamiltonian $V(x) = \frac{1}{2}\sum_{i=1}^{3} x_i^2/I_i$ [1]. [Actually this is an example of a so-called Lie-Poisson structure, in which $\Omega(x)$ is a linear function.]

(iii) Systems with a first integral: The ordinary differential equation (ODE) $\dot{x} = f(x)$, $x \in \mathbb{R}^n$, is said to have the first integral V if $dV(x)/dt = 0$.

Example 3. A Lotka-Volterra system [3]. The ODE

$$\dot{x}_1 = e^{x_3}, \quad \dot{x}_2 = e^{x_1} + e^{x_3}, \quad \dot{x}_3 = Be^{x_1} + e^{x_2}, \tag{2}$$

where B is a parameter, possesses the integral

$$V(x_1, x_2, x_3) = e^{x_2 - x_1} + B(x_2 - x_1) - x_3 . \quad (3)$$

(iv) Gradient systems: Gradient systems arise, e.g., in dynamical systems theory [4]. They are described by $\dot{x} = -\nabla V(x)$, $x \in \mathbb{R}^n$.

Example 4. The system $\dot{x}_1 = -2x_1(x_1 - 1)(2x_1 - 1)$, $\dot{x}_2 = -2x_2$ is a gradient system [4] with $n = 2$ and $V(x_1, x_2) = x_1^2(x_1 - 1)^2 + x_2^2$.

(v) Systems with a Lyapunov function: The ODE $\dot{x} = f(x)$, $x \in \mathbb{R}^n$, is said to possess the Lyapunov function V if $dV(x)/dt \le 0$. These functions were introduced by Lyapunov [5] and are a crucial ingredient of his direct or second method in the study of dynamical stability [6,7]. Some sufficient conditions for the existence of a Lyapunov function are given in [8].

Example 5. [9]

$$\dot{x}_1 = -x_2 - x_1^3, \qquad \dot{x}_2 = x_1 - x_2^3 \quad (4)$$

has the Lyapunov function $V(x_1, x_2) = x_1^2 + x_2^2$.

What do the above five classes of dynamical systems have in common? A preliminary answer would be that they all possess a function $V(x)$ such that $dV(x)/dt \le 0$. That is, classes (i), (ii), and (iii) each possess a function $V(x)$ such that $dV(x)/dt \equiv 0$, and classes (iv) and (v) each possess a function $V(x)$ such that $dV(x)/dt \le 0$.

In Sect. 2 we announce the result that classes of systems (i) to (v) have even more in common: under some mild technical assumptions, they can all be written as special cases of the novel class of "linear-gradient systems." In Sect. 3 we show how these linear-gradient systems can be integrated numerically in such a way that $V(x)$ is constant or nonincreasing, as appropriate.

An extended version of this work, including proofs of the results presented here, is given in [10].

II. LINEAR-GRADIENT SYSTEMS

Our main result is the following:

Theorem 1. *Let the ODE,*

$$\dot{x} = f(x), \qquad f \in C^r, \quad (5)$$

possess a C^{r+1} Morse function $V(x)$, where

$$(a) \quad \frac{dV}{dt} = 0,$$

i.e., V is an integral; or

$$(b) \quad \frac{dV}{dt} \le 0,$$

i.e., V is a (weak) Lyapunov function; or

$$(c) \quad \frac{dV}{dt} < 0,$$

where $f(x) \neq 0$, i.e., V is a strong Lyapunov function. Then for all $\{x \mid \nabla V(x) \neq 0\}$ there exists a locally bounded C^r matrix $L(x)$ such that the ODE (5) can be rewritten in the linear-gradient form

$$\dot{x} = L(x)\nabla V(x), \quad (6)$$

where

(a) *$L(x)$ is an antisymmetric matrix, respectively (resp.),*

(b) *$L(x)$ is a negative semidefinite matrix, resp.,*

(c) *$L(x)$ is a negative definite matrix.*

Some remarks:

1. A Morse function is a function whose critical points are all nondegenerate. A negative semidefinite matrix L is a matrix such that $v^t L v \le 0$ for all vectors v. A negative definite matrix L is a matrix such that $v^t L v < 0$ for all nonzero vectors v.

2. Under a coordinate transformation $x \mapsto C(x)$ we have $L(x) \mapsto \tilde{L}(x) := dC(x)L(x)[dC(x)]^t$. This implies that the theorem is invariant under coordinate transformations because $\tilde{L}$ is antisymmetric, negative semidefinite, resp., negative definite if and only if L is.

3. The theorem has a converse: If an ODE is in linear-gradient form (6) with L antisymmetric, resp., negative semidefinite, resp., negative definite, then V is an integral, resp., weak Lyapunov function, resp., strong Lyapunov function.

4. If the sign of dV/dt (zero, nonpositive, or negative) depends on x, then L can be chosen to be antisymmetric, negative semidefinite, or negative definite, respectively, depending on x. The type of representation is not unique: At points where $dV/dt = 0$, L can be chosen to be either antisymmetric or negative semidefinite.

5. A particular $L(x)$ satisfying the requirements of the theorem is

$$L_{ij}(x) = \frac{f_i v_j - v_i f_j + \delta_{ij} \sum f_k v_k}{\sum v_k^2}, \quad (7)$$

where $v_j = \partial V/\partial x_j$. However, L in (6) yielding (5) is not unique. In particular, under further mild technical conditions, there is an L which extends smoothly through critical points of V.

6. The fact that all systems with an integral can be written in the skew-gradient form $\dot{x} = L(x)\nabla V(x)$ was, as far as we know, first published in [11]. The general case is new, although the special case of the converse with $L(x)$ symmetric negative definite is well known and forms the subject of "generalized gradient systems" in dynamical systems [4]. Special cases corresponding to a matrix L, which is the sum of a skew, Poisson part and a symmetric, dissipative part, are given in [12], and references therein.

The constructive proof of Theorem 1 is given in [10]. We now give some illustrative examples of the above theorem.

Example 6. Particle in 1D with friction [4]. Consider the ODE

$$\dot{x} = x_2, \quad \dot{x}_2 = -\frac{\partial f(x_1)}{\partial x_1} - \alpha x_2, \tag{8}$$

where $\alpha \geq 0$ is a coefficient of friction and f is a potential function. Equation (8) has the energy $V(x_1, x_2) = \frac{1}{2}x_2^2 + f(x_1)$ as a Lyapunov function and can be written in the linear-gradient form (6) as

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -\alpha \end{pmatrix} \nabla V(x_1, x_2). \tag{9}$$

For $\alpha = 0$, the system is conservative and the matrix L is antisymmetric [case (a) above]. For $\alpha > 0$, the system is dissipative, V is a (weak) Lyapunov function, and $L(x)$ is negative semidefinite [case (b) above; cf Sect. 9.4 of [13]].

Example 7. An averaged system in wind-induced oscillation [6]. Consider the system

$$\begin{aligned} \dot{x}_1 &= -\zeta x_1 - \lambda x_2 + x_1 x_2, \\ \dot{x}_2 &= \lambda x_1 - \zeta x_2 + \tfrac{1}{2}(x_1^2 - x_2^2). \end{aligned} \tag{10}$$

Here $\zeta \geq 0$ is a damping factor and λ is a detuning parameter. Guckenheimer and Holmes [6] remark that Eq. (10) is a Hamiltonian for $\zeta = 0$ and a gradient system for $\lambda = 0$. We now show that, for all allowed values of ζ and λ, Eq. (10) can be written in linear-gradient form. To this end, denote $\zeta = \rho\cos(\theta)$ and $\lambda = \rho\sin(\theta)$. Then Eq. (10) can be written in linear-gradient form $\dot{x} = L\nabla V$ with

$$L = \begin{pmatrix} -\cos(\theta) & -\sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}, \tag{11}$$

$$\begin{aligned} V(x_1, x_2) = \frac{1}{2}\rho(x_1^2 + x_2^2) - \frac{1}{2}\sin(\theta)\left(x_1 x_2^2 - \frac{x_1^3}{3}\right) \\ + \frac{1}{2}\cos(\theta)\left(\frac{x_2^3}{3} - x_1^2 x_2\right). \end{aligned} \tag{12}$$

Note that for the matrix L in this example we have $v^t L v = -\cos(\theta)|v^2|$. Therefore, in the physical regime [where $\zeta \geq 0$ and hence $\cos(\theta) \geq 0$], either the matrix L is antisymmetric and V is an integral [for $\cos(\theta) = 0$] or L is negative definite and V is a strong Lyapunov function [for $\cos(\theta) > 0$]. [Note that, for $\lambda = \zeta = 0$, we have $\rho = 0$ and we are free to choose θ. In this limit, the system possesses an integral $V_1 = x_1 x_2^2 - x_1^3/3$, as well as a Lyapunov function $V_2 = x_2^3/3 - x_1^2 x_2$, and V given by Eq. (12) represents an arbitrary linear combination of these two functions.]

Example 8. Reproduced from [11], here is the linear-gradient form for the ODE (2) in Example 3:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -e^{x_3} \\ 0 & 0 & -e^{x_1} - e^{x_3} \\ e^{x_3} & e^{x_1} + e^{x_3} & 0 \end{pmatrix} \nabla V, \tag{13}$$

where V is given by Eq. (3).

Example 9. Here is the linear-gradient form for the ODE (4) in Example 5:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \nabla V(x_1, x_2), \tag{14}$$

where $a = -(x_1^4 + x_2^4)/(x_1^2 + x_2^2)$, $b = -(x_1^2 + x_2^2 + x_2 x_1^3 - x_1 x_2^3)/(x_1^2 + x_2^2)$, and $V = x_1^2 + x_2^2$. Note that the matrix L in (14) is negative definite.

III. DISCRETE GRADIENTS AND THE NUMERICAL INTEGRATION OF LINEAR-GRADIENT SYSTEMS

For differential equations whose time evolution has particular structural properties, such as preservation of Lagrangian structure [14], symplectic structure, phase space volume, symmetries, or conserved quantities, it is desirable to mimic these properties in any numerical integration [15]. This is particularly useful in long-time integrations. One can also view the discrete-time analogs as interesting physical systems in their own right [16]. Note that, in general, it is impossible to preserve symplectic structure *and* all first integrals simultaneously [17]. In this Letter we concentrate on preserving integrals (and Lyapunov functions).

A major application of the linear-gradient formulation (6) is that it has a simple and elegant discrete-time analog; moreover, this analog is also a universal representation for systems of each class.

Definition 1. [18] *Let $V(x)$ be a differentiable function. Then $\overline{\nabla}V(x, x')$ is a* ***discrete gradient*** *of V if it is continuous and*

$$\begin{aligned} \overline{\nabla}V(x, x')(x' - x) &= V(x') - V(x), \\ \overline{\nabla}V(x, x) &= \nabla V(x). \end{aligned} \tag{15}$$

Discrete gradients are not unique. Several examples of discrete gradients are given in [10,18,19].

Definition 2. *The function V is an integral of the map $x \mapsto x'$ if $V(x') = V(x), \forall x$. It is a weak Lyapunov function if $V(x') \leq V(x), \forall x$. It is a strong Lyapunov function if $V(x') < V(x)$ for all x such that $x \neq x'$.*

Theorem 2. [10] *Let the map $x \mapsto x'$ be defined implicitly by*

$$\left(\frac{\Delta x}{\Delta t} = \right)\frac{x' - x}{\tau} = \tilde{L}(x, x', \tau)\overline{\nabla}V(x, x'), \tag{16}$$

where $\overline{\nabla}V$ is any discrete gradient, $\tilde{L}$ is a matrix function, and τ represents a time step. Then $V(x)$ is an integral, resp., weak Lyapunov function, resp., strong Lyapunov function of the map if $\tilde{L}$ is antisymmetric, resp., negative semidefinite, resp., negative definite. Conversely, for any map with such a V, and any discrete gradient $\overline{\nabla}$, at points such that $\overline{\nabla}V(x, x') \neq 0$ there exists such an $\tilde{L}$ that the map takes the form (16).

It follows that (16) is a discrete approximation to the linear-gradient system (6) that preserves integrals, resp., Lyapunov functions, provided the method is consistent, i.e., $\tilde{L}(x, x, 0) = L(x)$.

Equations similar to (15) and (16) have appeared in many energy-conserving schemes for Hamiltonian systems [20–23], although the first axiomatic presentation was [18], and the first application to all systems with an integral was [24].

IV. CONCLUDING REMARKS

(i) In this Letter, for simplicity, we have restricted our discussion to the case of *one* first integral or Lyapunov function. In [10] we show that an n-dimensional ODE with $m \leq n - 1$ integrals and/or Lyapunov functions $V_1, \ldots, V_m$ can be written in the "multilinear-gradient" form

$$\dot{x} = L(x)\nabla V_1 \ldots \nabla V_m, \quad x \in \mathbb{R}^n, \tag{17}$$

where $L(x)$ is an $(m + 1)$ tensor. Structure-preserving integrators for Eq. (17) have also been constructed, generalizing Eq. (16).

(ii) Associated with (multi)linear-gradient systems of the form (17) there is also a formulation in terms of a bracket:

$$\frac{df(x)}{dt} = \{f, V_1, \ldots, V_m\}_L, \tag{18}$$

where the bracket is defined by

$$\{f_1, \ldots, f_p\}_L := \sum_{i_1,\ldots,i_p} L_{i_1,\ldots,i_p} \frac{\partial f_1}{\partial x_{i_1}} \cdots \frac{\partial f_p}{\partial x_{i_p}}, \tag{19}$$

where $p = m + 1$. This bracket satisfies the Leibniz rule in each of its variables:

$$\{f_1, \ldots, f_{j-1}, \phi(g_1, \ldots, g_k), f_{j+1}, \ldots, f_p\}_L = \sum_{i=1}^{k} \left(\frac{\partial \phi}{\partial g_i}\right) \{f_1, \ldots, f_{j-1}, g_i, f_{j+1}, \ldots, f_p\}_L, \tag{20}$$

$j = 1, \ldots, p$. Conversely, the tensor L is defined by the fundamental brackets $L_{i_1,\ldots,i_p} = \{x_{i_1}, \ldots, x_{i_p}\}_L$. It follows that $V(x)$ is an integral, resp., weak Lyapunov function, resp., strong Lyapunov function of the (multi)linear-gradient system (17) if and only if $W = 0\ \forall x$, resp., $W \leq 0\ \forall x$, resp., $W < 0$ for all x such that $|\nabla V(x)| \neq 0$, where $W := \{V, V_1, \ldots, V_m\}_L$. It also follows that V is an integral (resp., Lyapunov function) of the system (17) if and only if V_j is an integral (resp., Lyapunov function) of the system

$$\dot{x} = \tilde{L}(x)\nabla V_1 \ldots \nabla V_{j-1} \nabla V \nabla V_{j+1} \ldots \nabla V_m, \tag{21}$$

where

$$\tilde{L}_{i_1,\ldots,i_{j-1},i_j,i_{j+1},\ldots,i_m} = L_{i_j,\ldots,i_{j-1},i_1,i_{j+1},\ldots,i_m}. \tag{22}$$

Special cases of the bracket (19) are the Poisson bracket and the Nambu bracket [25].

(iii) It will be interesting to investigate whether there are topological obstacles to carrying over the results of this paper to the case of non-Euclidean phase spaces.

(iv) We hope to address the numerical order of accuracy of the integrator (16) in a forthcoming publication.

ACKNOWLEDGMENTS

We are indebted to the Marsden Fund of the Royal Society of New Zealand, and Lottery Science New Zealand, for their financial support. G.R.W.Q. is also grateful for a La Trobe University Large Central Grant and an Australian Research Council Small Grant supporting this project.

[1] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, New York, 1989), 2nd ed.
[2] P. J. Olver, *Applications of Lie Groups to Differential Equations* (Springer, New York, 1993), 2nd ed.
[3] B. Grammaticos, J. Moulin-Ollagnier, A. Ramani, J.-M. Strelcyn, and S. Wojciechowski, Physica (Amsterdam) **163A**, 683 (1990).
[4] Morris Hirsch and Stephen Smale, *Differential Equations, Dynamical Systems, and Linear Algebra* (Academic Press, New York, 1974).
[5] A. M. Lyapunov, *The General Problem of the Stability of Motion* (Taylor & Francis, London, 1992).
[6] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields,* (Springer, New York, 1983).
[7] N. Rouche, P. Habets, and M. LaLoy, *Stability Theory by Liapunov's Direct Method* (Springer, New York, 1977).
[8] Taro Yoshizawa, *Stability Theory by Liapunov's Second Method* (Mathematical Society of Japan, Tokyo, 1966), Chap. 5.
[9] D. A. Sanchez, *Ordinary Differential Equations and Stability Theory* (Freeman, San Francisco, 1968).
[10] R. I. McLachlan, G. R. W. Quispel, and N. Robidoux, Philos. Trans. R. Soc. London A (to be published).
[11] G. R. W. Quispel and H. W. Capel, Phys. Lett. A **218**, 223–228 (1996).
[12] A. Bloch, P. S. Krishnaprasad, J. E. Marsden, and T. S. Ratiu, Commun. Math. Phys. **175**, 1–42 (1996).
[13] J. K. Hale and H. Koçak, *Dynamics and Bifurcations* (Springer, Berlin, 1991).
[14] J. M. Wendlandt and J. E. Marsden, Physica (Amsterdam) **106D**, 223–246 (1997).

[15] A. M. Stuart and A. R. Humphries, *Dynamical Systems and Numerical Analysis* (Cambridge University Press, New York, 1996).
[16] T. D. Lee, J. Stat. Phys. **46**, 843–860 (1987).
[17] Z. Ge and J. E. Marsden, Phys. Lett. A **133**, 135–139 (1988).
[18] Oscar Gonzalez, J. Nonlinear Sci. **6**, 449–467 (1996).
[19] Amiram Harten, Peter D. Lax, and Bram van Leer, SIAM Rev. **25**, 35–61 (1983).
[20] A. J. Chorin, T. J. R. Hughes, J. E. Marsden, and M. McCracken, Commun. Pure Appl. Math. **31**, 205–256 (1978).
[21] L. Gotusso, Appl. Math. Comput. **17**, 129–136 (1985).
[22] T. Itoh and K. Abe, J. Comput. Phys. **77**, 85 (1988).
[23] Yu. A. Kriksin, Comput. Math. Math. Phys. **33**, 181–191 (1993).
[24] G. R. W. Quispel and G. S. Turner, J. Phys. A **29**, L341–L349 (1996).
[25] L. Takhtajan, Commun. Math. Phys. **160**, 295–315 (1994).