Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005



# Time-Reversal Symmetry, Poincaré Recurrence, Irreversibility, and the Entropic Arrow of Time: From Mechanics to System Thermodynamics

Wassim M. Haddad, VijaySekhar Chellaboina, and Sergey G. Nersesov

*Abstract*— **In this paper, we use a large-scale dynamical systems perspective to provide a system-theoretic foundation for thermodynamics. Specifically, using a state space formulation, we develop a nonlinear compartmental dynamical system model characterized by energy conservation laws that is consistent with basic thermodynamic principles. In addition, we establish the existence of a unique, continuously differentiable global entropy function for our large-scale dynamical system, and using Lyapunov stability theory we show that the proposed thermodynamic model has convergent trajectories to Lyapunov stable equilibria determined by the system initial energies. Finally, using the system entropy, we establish the absence of Poincaré recurrence for our thermodynamic model and develop a clear connection between irreversibility, the second law of thermodynamics, and the entropic arrow of time.**

## I. Introduction

As discussed in the recent monograph [1], there have been many different presentations of classical thermodynamics with varying hypotheses and conclusions. To exacerbate matters, the careless and considerable differences in the definitions of two of the key notions of thermodynamics—namely, the notions of reversibility and irreversibility—have contributed to the widespread confusion and lack of clarity of the exposition of classical thermodynamics over the past one and a half centuries. For example, the concept of reversible processes as defined by Clausius, Kelvin, Planck, and Carathéodory have very different meanings. In particular, Clausius defines a reversible (*umkehrbar*) process as a slowly varying process wherein successive states of this process differ by infinitesimals from the equilibrium system states. Such system transformations are commonly referred to as *quasistatic* transformations in the thermodynamic literature. Alternatively, Kelvin's notions of reversibility involve the ability of a system to completely recover its initial state from the final system state. Planck introduced several notions of reversibility. His main notion of reversibility is one of *complete* reversibility and involves recoverability of the original state of the dynamical system while at the same time restoring the environment to its original condition. Unlike Clausius' notion of reversibility, Kelvin's and Planck's notions of reversibility do not require the system to exactly retrace its original trajectory in reverse order. Carathéodory's notion of reversibility involves recoverability of the system state in an adiabatic process resulting in yet another definition of thermodynamic reversibility. These subtle distinctions of (ir)reversibility are often unrecognized in the thermodynamic literature. Notable exceptions to this fact include [2], [3], with [3] providing an excellent exposition of the relation between irreversibility, the second law of thermodynamics, and the arrow of time.

In this paper, we place thermodynamics on a system-theoretic foundation so as to harmonize it with classical mechanics. In particular, we develop a novel formulation of thermodynamics that can be viewed as a moderate-sized system theory as compared to statistical thermodynamics. This middle-ground theory involves deterministic large-scale dynamical system models that bridge the gap between classical and statistical thermodynamics. Specifically, since thermodynamic models are concerned with energy flow among subsystems, we use a state space formulation to develop a nonlinear compartmental dynamical system model that is characterized by energy conservation laws capturing the exchange of energy between coupled macroscopic subsystems. Furthermore, using graph-theoretic notions, we state two thermodynamic axioms consistent with the zeroth and second laws of thermodynamics, which ensure that our large-scale dynamical system model gives rise to a thermodynamically consistent energy flow model. Specifically, using a large-scale dynamical systems theory perspective for thermodynamics, we show that our compartmental dynamical system model leads to a precise formulation of the equivalence between work energy and heat in a large-scale dynamical system.

Next, we give a deterministic definition of entropy for a large-scale dynamical system that is consistent with the classical thermodynamic definition of entropy, and we show that it satisfies a Clausius-type inequality leading to the law of entropy nonconservation. However, unlike classical thermodynamics, wherein entropy is not defined for arbitrary states out of equilibrium, our definition of entropy holds for nonequilibrium dynamical systems. Then, using Lyapunov stability theory, we show that in the absence of heat exchange with the environment our thermodynamically consistent large-scale nonlinear dynamical system model possesses a continuum of equilibria and is *semistable*, that is, it has convergent subsystem energies to Lyapunov stable energy equilibria determined by the large-scale system initial subsystem energies.

For our thermodynamically consistent dynamical system model, we further establish the existence of a *unique* continuously differentiable global entropy function for all equilibrium and nonequilibrium states. Using this global entropy function, we go on to establish a clear connection between thermodynamics and the arrow of time. Specifically, we rigorously show a *state irrecoverability* and hence a *state irreversibility* nature of thermodynamics. In particular, we show that for every nonequilibrium system state and corresponding system trajectory of our thermodynamically consistent large-scale nonlinear dynamical system, there does not exist a state such that the corresponding system trajectory completely recovers the initial system state of the dynamical system and at the same time restores the energy supplied by the environment back to its original condition. This, along with the existence of a global strictly increasing entropy function on every nontrivial system trajectory, gives a clear

This research was supported in part by the Air Force Office of Scientific Research under Grant F49620-03-1-0178.

W. M. Haddad is with the School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0150, USA (wm.haddad@aerospace.gatech.edu).

V. Chellaboina is with the Department of Mechanical, Aerospace, and Biomedical Engineering, University of Tennessee, Knoxville, TN 37996-2210, USA (Chellaboina@utk.edu).

S. G. Nersesov is with the Department of Mechanical Engineering, Villanova University, Villanova, PA 19085-1681, USA (sergey.nersesov@villanova.edu).

0-7803-9568-9/05/$20.00 ©2005 IEEE

*time-reversal asymmetry* characterization of thermodynamics, establishing an emergence of the direction of time flow.

## II. Mathematical Preliminaries

In this section we establish notation and provide a general axiomatic definition of a dynamical system. The notation used in this paper is fairly standard. Specifically, for $z \in \mathbb{R}^q$ we write $z \geq\geq 0$ (respectively, $z >> 0$) to indicate that every component of $z$ is nonnegative (respectively, positive). In this case we say that $z$ is *nonnegative* or *positive*, respectively. Let $\overline{\mathbb{R}}_+^q$ and $\mathbb{R}_+^q$ denote the nonnegative and positive orthants of $\mathbb{R}^q$, that is, if $z \in \mathbb{R}^q$, then $z \in \overline{\mathbb{R}}_+^q$ and $z \in \mathbb{R}_+^q$ are equivalent, respectively, to $z \geq\geq 0$ and $z >> 0$.

Next, we define a dynamical system as a precise mathematical object satisfying a set of axioms. For this definition, let $\mathcal{U}$ denote an input space that consists of bounded continuous $U$-valued functions on $[0,\infty)$. The set $U \subseteq \mathbb{R}^m$ contains the set of input values, that is, at any time $t \geq t_0$, $u(t) \in U$. The space $\mathcal{U}$ is assumed to be closed under the shift operator. Furthermore, we let $\mathcal{Y}$ denote an output space that consists of continuous $Y$-valued functions on $[0,\infty)$. The set $Y \subseteq \mathbb{R}^l$ contains the set of output values, that is, each value of $y(t) \in Y$, $t \geq t_0$. The space $\mathcal{Y}$ is assumed to be closed under the shift operator.

*Definition 2.1:* Let $\mathcal{D}$ be a Euclidian space with norm given by $\|\cdot\|$. A *dynamical system* on $\mathcal{D}$ is the octuple $(\mathcal{D},\mathcal{U},U,\mathcal{Y},Y,[0,\infty),s,h)$, where $s:[0,\infty)\times\mathcal{D}\times\mathcal{U}\to\mathcal{B}$ and $h:\mathcal{D}\times U\to Y$ are such that the following axioms hold:

*i)* (Continuity): $s(\cdot,\cdot,u)$ is jointly continuous for all $u \in \mathcal{U}$.

*ii)* (Consistency): $s(t_0,z_0,u)=z_0$ for all $t_0\in\mathbb{R}$, $z_0\in\mathcal{D}$, and $u\in\mathcal{U}$.

*iii)* (Determinism): $s(t,z_0,u_1)=s(t,z_0,u_2)$ for all $t\in[t_0,\infty)$, $z_0\in\mathcal{D}$, and $u_1$, $u_2\in\mathcal{U}$ satisfying $u_1(\tau)=u_2(\tau)$, $\tau\leq t$.

*iv)* (Semi-group property): $s(\tau,s(t,z_0,u),u) = s(t+\tau,z_0,u)$ for all $z_0\in\mathcal{D}$, $u\in\mathcal{U}$, and $\tau$, $t\in[t_0,\infty)$.

*v)* (Read-out map): There exists $y\in\mathcal{Y}$ such that $y(t)=h(s(t,z_0,u),u(t))$ for all $z_0\in\mathcal{D}$, $u\in\mathcal{U}$, and $t\geq t_0$.

We denote the dynamical system $(\mathcal{D},\mathcal{U},U,\mathcal{Y},Y,[0,\infty),s,h)$ by $\mathcal{G}$. Furthermore, we refer to the map $s(\cdot,\cdot,\cdot)$ as the *flow* or *trajectory* of $\mathcal{G}$ corresponding to $z_0\in\mathcal{D}$, and for a given $s(t,z_0,u)$, $t\geq t_0$, $u\in\mathcal{U}$, we refer to $z_0\in\mathcal{D}$ as an *initial condition* of $\mathcal{G}$. Given $t\in\mathbb{R}$ we denote the map $s(t,\cdot,\cdot):\mathcal{D}\times\mathcal{U}\to\mathcal{D}$ by $s_t(z_0,u)$. Hence, for a fixed $t\in\mathbb{R}$ the set of mappings defined by $s_t(z_0,u)=s(t,z_0,u)$ for every $z_0\in\mathcal{D}$ and $u\in\mathcal{U}$ gives the *flow* of $\mathcal{G}$. In particular, if $\mathcal{D}_0$ is a collection of initial conditions such that $\mathcal{D}_0\subset\mathcal{B}$, then the flow $s_t:\mathcal{D}_0\times\mathcal{U}\to\mathcal{B}$ is the motion of all points $z_0\in\mathcal{D}_0$ or, equivalently, the image of $\mathcal{D}_0\subset\mathcal{D}$ under the flow $s_t$, that is, $s_t(\mathcal{D}_0,\mathcal{U})\subset\mathcal{D}$, where $s_t(\mathcal{D}_0,\mathcal{U})\triangleq\{y: y=s_t(z_0,u)$ for all $z_0\in\mathcal{D}$ and $u\in\mathcal{U}\}$. Alternatively, if the initial condition $z_0\in\mathcal{D}$ is fixed and we let $[t_0,t_1]\subset\mathbb{R}$ and $u\in\mathcal{U}$, then the mapping $s(\cdot,z_0,u):[t_0,t_1]\to\mathcal{D}$ defines the *solution curve* or *trajectory* of the dynamical system $\mathcal{G}$. Given $z\in\mathcal{D}$ and $u\in\mathcal{U}$, we denote the map $s(\cdot,z,u):\mathbb{R}\to\mathcal{D}$ by $s^z(t,u)$.

The dynamical system $\mathcal{G}$ is *isolated* if $u(t)\equiv 0$. Furthermore, an *equilibrium point* of the isolated dynamical system $\mathcal{G}$ is a point $x\in\mathcal{D}$ satisfying $s(t,x,0)=x$, $t\geq t_0$. An equilibrium point $x\in\mathcal{D}_{\rm c}\subseteq\mathcal{D}$ of the isolated dynamical system $\mathcal{G}$ is *Lyapunov stable* with respect to the positively invariant set $\mathcal{D}_{\rm c}$ if, for every relatively open subset $\mathcal{N}_\varepsilon$ of $\mathcal{D}_{\rm c}$ containing $x$, there exists a relatively open subset $\mathcal{N}_\delta$ of $\mathcal{D}_{\rm c}$ containing $x$ such that $s_t(\mathcal{N}_\delta,\mathcal{U})\subset\mathcal{N}_\varepsilon$ for all $t\geq t_0$, where $\mathcal{U}=\{u:\mathbb{R}\to\mathbb{R}: u(t)\equiv 0\}$. An equilibrium point $x\in\mathcal{D}_{\rm c}$ of the isolated dynamical system $\mathcal{G}$ is called *semistable* if it is Lyapunov stable and there exists a relatively open subset $\mathcal{N}$ of $\mathcal{D}_{\rm c}$ containing $x$ such that for all initial conditions in $\mathcal{N}$, the trajectory of $\mathcal{G}$ converges to a Lyapunov stable equilibrium point, that is, $\|s(t,z,0)-y\|\to 0$ as $t\to\infty$, where $y\in\mathcal{D}_{\rm c}$ is a Lyapunov stable equilibrium point of $\mathcal{G}$ and $z\in\mathcal{D}_{\rm c}$. The isolated dynamical system $\mathcal{G}$ is said to be *semistable* if every equilibrium point of $\mathcal{G}$ is semistable.

Finally, for a given interval $[t_0,t_1]$, where $0\leq t_0<t_1<\infty$, let $\mathcal{W}_{[t_0,t_1]}$ denote the set of all possible trajectories of $\mathcal{G}$ given by

$$\mathcal{W}_{[t_0,t_1]} \triangleq \{s^z:[t_0,t_1]\times\mathcal{U}\to\mathcal{D}: s^z(\cdot,u(\cdot)) \text{ satisfies Axioms } i)-iv) \text{ of Definition 2.1, } z\in\mathcal{D}, \text{ and } u(\cdot)\in\mathcal{U}\}, \quad (1)$$

where $s^z(\cdot,u(\cdot))$ denotes the solution curve or trajectory of $\mathcal{G}$ for a given fixed initial condition $z\in\mathcal{D}$ and input $u(\cdot)\in\mathcal{U}$.

## III. Reversibility, Irreversibility, Recoverability, and Irrecoverability

The notions of reversibility, irreversibility, recoverability, and irrecoverability all play a crucial role in thermodynamic processes. In this section we define the notions of *R-state reversibility*, *state reversibility*, and *state recoverability* of a dynamical system $\mathcal{G}$. $R$-state reversibility concerns the existence of a system state with the property that a transformed system trajectory through an involution operator $R$ is an image of a given system trajectory of $\mathcal{G}$ on a specified finite time interval. State reversibility concerns the existence of a system state with the property that the resulting system trajectory is the time-reversed image of a given system trajectory of $\mathcal{G}$ on a specified finite time interval. Finally, state recoverability concerns the existence of a system state with the property that the resulting system trajectory completely recovers the initial state of the dynamical system over a finite time interval.

For the results of this section we use the definition of a dynamical system given in Definition 2.1. We start by establishing the notions of (ir)reversibility and (ir)recoverability of a dynamical system $\mathcal{G}$ defined on a Euclidian space $\mathcal{D}$.

*Definition 3.1:* Consider the dynamical system $\mathcal{G}$ defined on $\mathcal{D}$. Let $R:\mathcal{D}\to\mathcal{D}$ be an involutive operator (that is, $R^2=I_{\mathcal{D}}$, where $I_{\mathcal{D}}$ denotes the identity operator on $\mathcal{D}$) and let $s^z(\cdot,u(\cdot))\in\mathcal{W}_{[t_0,t_1]}$, where $u(\cdot)\in\mathcal{U}$. The function $s^{-z}:[t_0,t_1]\times\mathcal{U}\to\mathcal{D}$ is an *R-reversed trajectory* of $s^z(\cdot,u(\cdot))$ if there exists an input $u^-(\cdot)\in\mathcal{U}$ and a continuous, strictly increasing function $\tau:[t_0,t_1]\to[t_0,t_1]$ such that $\tau(t_0)=t_0$, $\tau(t_1)=t_1$, and

$$s^{-z}(t,u^-(t)) = Rs^z(t_0+t_1-\tau(t),u(t_0+t_1-\tau(t))), \quad t\in[t_0,t_1]. \quad (2)$$

*Definition 3.2:* Consider the dynamical system $\mathcal{G}$ defined on $\mathcal{D}$. Let $R:\mathcal{D}\to\mathcal{D}$ be an involutive operator, let $r:\mathcal{U}\times\mathcal{Y}\to\mathbb{R}$, and let $s^z(\cdot,u(\cdot))\in\mathcal{W}_{[t_0,t_1]}$, where $u(\cdot)\in\mathcal{U}$. $s^z(\cdot,u(\cdot))$ is an *R-reversible trajectory* of $\mathcal{G}$ if there exists an input $u^-(\cdot)\in\mathcal{U}$ such that $s^{-z}(\cdot,u^-(\cdot))\in\mathcal{W}_{[t_0,t_1]}$ and

$$\int_{t_0}^{t_1} r(u(t),y(t))\mathrm{d}t + \int_{t_0}^{t_1} r(u^-(t),y^-(t))\mathrm{d}t = 0, \quad (3)$$

where $y^-(\cdot)$ denotes the read-out map for the $R$-reversed trajectory of $s^z(\cdot,u(\cdot))$. Furthermore, $\mathcal{G}$ is an *R-state reversible dynamical system* if for every $z\in\mathcal{D}$, $s^z(\cdot,u(\cdot))$, where $u(\cdot)\in\mathcal{U}$, is an $R$-reversible trajectory of $\mathcal{G}$.

In classical mechanics, $R$ is a transformation which reverses the sign of all system momenta and magnetic fields, whereas in classical reversible thermodynamics $R$ can be taken to be the identity operator. Note that if $R = I_{\mathcal{D}}$, then $s^z(\cdot, u(\cdot))$, where $u(\cdot) \in \mathcal{U}$, is an $I_{\mathcal{D}}$-reversible trajectory or, simply, $s^z(\cdot, u(\cdot))$ is a *reversible trajectory*. Furthermore, we say that $\mathcal{G}$ is a *state reversible dynamical system* if and only if for every $z \in \mathcal{D}$, $s^z(\cdot, u(\cdot))$, where $u(\cdot) \in \mathcal{U}$, is a reversible trajectory of $\mathcal{G}$. Note that unlike state reversible systems, $R$-state reversible dynamical systems need not retrace every stage of the original system trajectory in reverse order, nor is it necessary for the dynamical system to recover the initial system state. The function $r(u, y)$ in Definition 3.2 is a generalized *power supply* from the environment to the dynamical system through the system's input-output ports $(u, y)$. Hence, (3) assures that the total generalized energy supplied to the dynamical system $\mathcal{G}$ by the environment is returned to the environment over a given $R$-reversible trajectory starting and ending at any given (not necessarily the same) state $z \in \mathcal{D}$. Furthermore, condition (3) assures that a reversible process completely restores the original dynamic state of a system and at the same time restores the energy supplied by the environment back to its original condition. The following result provides sufficient conditions for the existence of an $R$-reversible trajectory of a nonlinear dynamical system $\mathcal{G}$, and hence, establishes sufficient conditions for $R$-state reversibility of the dynamical system $\mathcal{G}$.

*Theorem 3.1:* Consider the dynamical system $\mathcal{G}$ defined on $\mathcal{D}$. Let $R : \mathcal{D} \to \mathcal{D}$ be an involutive operator, and let $s^z(\cdot, u(\cdot)) \in \mathcal{W}_{[t_0,t_1]}$, where $u(\cdot) \in \mathcal{U}$. Assume there exist a continuous function $V : \mathcal{D} \to \mathbb{R}$ and a function $r : \mathcal{U} \times \mathcal{Y} \to \mathbb{R}$ such that $V(z) = V(Rz)$, $z \in \mathcal{D}$, and for every $z \in \mathcal{D}$ and all $\hat{t}_0$, $\hat{t}_1$, $t_0 \le \hat{t}_0 < \hat{t}_1 \le t_1$,

$$V(s^z(\hat{t}_1, u(\hat{t}_1))) \geq V(s^z(\hat{t}_0, u(\hat{t}_0))) + \int_{\hat{t}_0}^{\hat{t}_1} r(u(t), y(t))\mathrm{d}t. \qquad (4)$$

Furthermore, assume there exists $\mathcal{M} \subset \mathcal{D}$ such that for all $\hat{t}_0$, $\hat{t}_1$, $t_0 \le \hat{t}_0 < \hat{t}_1 \le t_1$, and $s^z(t, u(t)) \notin \mathcal{M}$, $t \in [\hat{t}_0, \hat{t}_1]$, (4) holds as a strict inequality. If $s^z(\cdot, u(\cdot))$ is an $R$-reversible trajectory of $\mathcal{G}$, then $s^z(t, u(t)) \in \mathcal{M}$, $t \in [t_0, t_1]$.

It is important to note that since $V : \mathcal{D} \to \mathbb{R}$ in Theorem 3.1 is not sign definite, Theorem 3.1 also holds for the case where the inequality in (4) is reversed. The following corollary to Theorem 3.1 is immediate.

*Corollary 3.1:* Consider the dynamical system $\mathcal{G}$ defined on $\mathcal{D}$. Let $R : \mathcal{D} \to \mathcal{D}$ be an involutive operator, let $\mathcal{M} \subset \mathcal{D}$, and let $s^z(\cdot, u(\cdot)) \in \mathcal{W}_{[t_0,t_1]}$, where $u(\cdot) \in \mathcal{U}$. Assume there exists a continuous function $V : \mathcal{D} \to \mathbb{R}$ such that $V(z) = V(Rz)$, $z \in \mathcal{D}$, and for $s^z(t, u(t)) \notin \mathcal{M}$, $t \in [t_1, t_2]$, $V(s(t, z_0, u(\cdot)))$ is a strictly increasing (respectively, decreasing) function of time. If $s^z(\cdot, u(\cdot))$ is an $R$-reversible trajectory of $\mathcal{G}$, then $s^z(t, u(t)) \in \mathcal{M}$, $t \in [t_0, t_1]$.

It follows from Corollary 3.1 that if, for a given dynamical system $\mathcal{G}$, there exists an $R$-reversible trajectory of $\mathcal{G}$, then there does not exist a function of the state of the system that strictly decreases or increases in time on any trajectory of $\mathcal{G}$ lying in $\mathcal{M}$. In this case, the existence of a completely ordered time set having a topological structure involving a closed set homeomorphic to the real line cannot be established. Such systems, which include lossless Newtonian and Hamiltonian systems, are time-reversal symmetric and hence lack an inherent time direction. However, that is not the case with thermodynamic systems.

Next, we present a notion of state recoverability of a dynamical system $\mathcal{G}$.

*Definition 3.3:* Consider the dynamical system $\mathcal{G}$ defined on $\mathcal{D}$. Let $r : \mathcal{U} \times \mathcal{Y} \to \mathbb{R}$, and let $s^z(\cdot, u(\cdot)) \in \mathcal{W}_{[t_0,t_1]}$, where $u(\cdot) \in \mathcal{U}$. $s^z(\cdot, u(\cdot))$ is a *recoverable trajectory* of $\mathcal{G}$ if there exists $u^-(\cdot) \in \mathcal{U}$ and $t_2 > t_1$ such that $u^- : [t_1, t_2] \to U$,

$$s(t_2, s^z(t_1, u(t_1)), u^-(t_2)) = s^z(t_0, u(t_0)), \qquad (5)$$

and

$$\int_{t_0}^{t_1} r(u(t), y(t))\mathrm{d}t + \int_{t_1}^{t_2} r(u^-(t), y^-(t))\mathrm{d}t = 0, \qquad (6)$$

where $y^-(\cdot)$ denotes the read-out map for the trajectory $s(\cdot, s^z(t_1, u(t_1)),\ u^-(\cdot))$. Furthermore, $\mathcal{G}$ is a *state recoverable dynamical system* if for every $z \in \mathcal{D}$, $s^z(\cdot, u(\cdot))$ is a recoverable trajectory of $\mathcal{G}$.

It follows from the definition of state recoverability that the way in which the initial dynamical system state is restored may be chosen freely so long as (6) is satisfied. Hence, unlike $R$-state reversibility, it is not necessary for the dynamical system to recover the initial state of the system through an involutive transformation of the system trajectory. Furthermore, unlike state reversibility, it is not necessary for the dynamical system to retrace every stage of the original trajectory in the reverse order. However, condition (6) assures that the recoverable process completely restores the original dynamic state and at the same time restores the energy supplied by the environment back to its original condition. This notion of recoverability is closely related to Planck's notion of complete reversibility, wherein the initial system state is restored in the *totality of Nature* ("die gesamte Natur"). The following result provides a sufficient condition for the existence of a recoverable trajectory of a nonlinear dynamical system $\mathcal{G}$, and hence, establishes sufficient conditions for state recoverability of $\mathcal{G}$.

*Theorem 3.2:* Consider the dynamical system $\mathcal{G}$ defined on $\mathcal{D}$. Let $s^z(\cdot, u(\cdot)) \in \mathcal{W}_{[t_0,t_1]}$, where $u(\cdot) \in \mathcal{U}$. Assume there exist a continuous function $V : \mathcal{D} \to \mathbb{R}$ and a function $r : \mathcal{U} \times \mathcal{Y} \to \mathbb{R}$ such that for every $z \in \mathcal{D}$ and all $\hat{t}_0$, $\hat{t}_1$, $t_0 \le \hat{t}_0 < \hat{t}_1 \le t_1$, $V(s^z(t, u(t)))$ satisfies (4). Furthermore, assume there exists $\mathcal{M} \subset \mathcal{D}$ such that for all $\hat{t}_0$, $\hat{t}_1$, $t_0 \le \hat{t}_0 < \hat{t}_1 \le t_1$, and $s^z(t, u(t)) \notin \mathcal{M}$, $t \in [\hat{t}_0, \hat{t}_1]$, (4) holds as a strict inequality. If $s^z(\cdot, u(\cdot))$ is a recoverable trajectory of $\mathcal{G}$, then $s^z(t, u(t)) \in \mathcal{M}$, $t \in [t_0, t_1]$.

The following corollary to Theorem 3.2 is immediate.

*Corollary 3.2:* Consider the dynamical system $\mathcal{G}$ defined on $\mathcal{D}$. Let $\mathcal{M} \subset \mathcal{D}$, and let $s^z(\cdot, u(\cdot)) \in \mathcal{W}_{[t_0,t_1]}$, where $u(\cdot) \in \mathcal{U}$. Assume there exists a continuous function $V : \mathcal{D} \to \mathbb{R}$ such that for $s^z(t, u(t)) \notin \mathcal{M}$, $t \in [t_0, t_1]$, $V(s(t, z_0, u(\cdot))$ is a strictly increasing (respectively, decreasing) function of time. If $s^z(\cdot, u(\cdot))$ is a recoverable trajectory of $\mathcal{G}$, then $s^z(t, u(t)) \in \mathcal{M}$, $t \in [t_0, t_1]$.

As in the case of $R$-state reversibility and state reversibility, state recoverability can be used to establish a connection between a dynamical system evolving on a manifold $\mathcal{M} \subset \mathcal{D}$ and the arrow of time. However, in the case of state recoverability, the recoverable dynamical system trajectory need not involve an involutive transformation of the system trajectory, nor is it required to retrace the original system trajectory in recovering the original dynamic state. It should be noted here that state recoverability is not implied by the concepts of *reachability* and *controllability*, which play a central role in control theory. For example, one might envision, albeit with a considerable stretch of the imagination, perfectly controlled inputs that could reassemble a broken egg or even fuse water into solid cubes of ice. However, in all such cases, an external source of energy from the environment would be required to operate such an immaculate state recoverable mechanism and would violate condition (6). Clearly, state recoverability is a weaker notion than that of state reversibility since state reversibility implies state recoverability; the converse, however, is not generally true. Conversely, state irrecoverability is a logically stronger notion than state irreversibility since state irrecoverability implies state irreversibility. However, as we see in Section VI, these notions are equivalent for thermodynamic systems.

## IV. Reversible Systems, Volume-Preserving Flows, and Poincaré Recurrence

The notion of $R$-state reversibility introduced in Section III is one of the fundamental symmetries that arises in natural science. This notion can also be characterized by the flow of a dynamical system. In particular, consider the dynamical system given by

$$\dot{z}(t) \;=\; w(z(t)), \quad z(t_0)=z_0, \quad t\in\mathcal{I}_{z_0}, \qquad (7)$$

where $z(t)\in\mathcal{D}\subseteq\mathbb{R}^q$, $t\in\mathcal{I}_{z_0}$, is the system state vector, $\mathcal{D}$ is an open subset of $\mathbb{R}^q$, $w:\mathcal{D}\to\mathbb{R}^q$ is locally Lipschitz continuous on $\mathcal{D}$, and $\mathcal{I}_{z_0}=[t_0,\tau_{z_0})$, $t_0<\tau_{z_0}\le\infty$, is the maximal interval of existence for the solution $z(\cdot)$ of (7). A function $z:\mathcal{I}_{z_0}\to\mathcal{D}$ is said to be the *solution* to (7) on the interval $\mathcal{I}_{z_0}\subseteq\mathbb{R}$ with initial condition $z(t_0)=z_0$, if $z(t)$ satisfies (7) for all $t\in\mathcal{I}_{z_0}$. Note that since $w(\cdot)$ is locally Lipschitz continuous on $\mathcal{D}$, it follows from Theorem 3.1 of [4, p. 18] that the solution to (7) is unique for every initial condition in $\mathcal{D}$ and jointly continuous in $t$ and $z_0$. In this case, the semi-group property $s(t+\tau,z_0)=s(t,s(\tau,z_0))$, $t,\tau\in\mathcal{I}_{z_0}$, and the continuity of $s(t,\cdot)$ on $\mathcal{D}$, $t\in\mathcal{I}_{z_0}$, hold. Given $t\in\mathbb{R}$, we denote the flow $s(t,\cdot):\mathcal{D}\to\mathcal{D}$ of (7) by $s_t(z_0)$ for $z_0\in\mathcal{D}$, and given $z\in\mathcal{D}$, we denote the trajectory $s(\cdot,z):\mathbb{R}\to\mathcal{D}$ of (7) by $s^z(t)$. Now, in terms of the flow $s_t:\mathcal{D}\to\mathcal{D}$ of (7), the consistency and semi-group properties of (7) can be equivalently written as $s_0(z_0)=z_0$ and $(s_\tau\circ s_t)(z_0)=s_\tau(s_t(z_0))=s_{t+\tau}(z_0)$, where "$\circ$" denotes the composition operator. Next, it follows from continuity of solutions and the semi-group property that the map $s_t:\mathcal{D}\to\mathcal{D}$ is a continuous function with a continuous inverse $s_{-t}$. Thus, $s_t$, $t\in\mathcal{I}_{z_0}$, generates a one-parameter family of homeomorphisms on $\mathcal{D}$ forming a commutative group under composition.

To show that $R$-state reversibility can be characterized by the flow of (7), let $\mathcal{R}:\mathcal{D}\to\mathcal{D}$ be a continuous map of (7) such that

$$\dot{\mathcal{R}}(z(t))=-w(\mathcal{R}(z(t))), \quad t\in\mathcal{I}_{\mathcal{R}(z_0)}. \qquad (8)$$

Now, it follows from (8) that $\mathcal{R}\circ s_t=s_{-t}\circ\mathcal{R}$, $t\in\mathcal{I}_{z_0}$. This condition, with $\mathcal{R}(\cdot)$ satisfying (8), defines an $R$-reversed trajectory of (7) in the sense of Definition 3.1 with $\tau(t)=t$.

In the context of classical mechanics involving the *configuration* manifold (space of generalized positions) $\mathcal{Q}=\mathbb{R}^n$, with governing equations given by

$$\dot{q}(t) \;=\; \left(\frac{\partial\mathcal{H}(q(t),p(t))}{\partial p(t)}\right)^{\mathrm{T}}, \quad q(t_0)=q_0, \quad t\ge t_0, \qquad (9)$$

$$\dot{p}(t) \;=\; -\left(\frac{\partial\mathcal{H}(q(t),p(t))}{\partial q(t)}\right)^{\mathrm{T}}, \quad p(t_0)=p_0, \qquad (10)$$

where $q\in\mathbb{R}^n$ denotes generalized system positions, $p\in\mathbb{R}^n$ denotes generalized system momenta, $\mathcal{H}:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ is the system Hamiltonian given by $\mathcal{H}(q,p)\triangleq\dot{q}^{\mathrm{T}}p-\mathcal{L}(q,\dot{q})$, $\mathcal{L}(q,\dot{q})$ is the system Lagrangian, and $p(q,\dot{q})\triangleq\left(\frac{\partial\mathcal{L}(q,\dot{q})}{\partial\dot{q}}\right)^{\mathrm{T}}$, the reversing symmetry $\mathcal{R}:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n\times\mathbb{R}^n$ is such that $\mathcal{R}(q,p)=(q,-p)$ and satisfies (8) trivially. In this case, $\mathcal{R}$ is an involution. This implies that if $(q(t),p(t))$, $t\ge t_0$, is a solution to (9) and (10), then $(q(-t),-p(-t))$, $t\ge t_0$, is also a solution to (9) and (10) with initial condition $(q_0,-p_0)$. In the configuration space this clearly shows the time reversal nature of lossless mechanical systems.

Reversible dynamical systems tend to exhibit a phenomenon known as *Poincaré recurrence* [5]. Poincaré recurrence states that if a dynamical system has a fixed total energy that restricts its dynamics to bounded subsets of its state space, then the dynamical system will eventually return arbitrarily close to its initial system state infinitely often. More precisely, Poincaré [5] established the fact that if the flow of a dynamical system preserves volume and has only bounded orbits, then for each open set there exist orbits that intersect the set infinitely often. In order to state the Poincaré recurrence theorem, the following definitions are needed.

*Definition 4.1:* Let $\mathcal{V}\subset\mathbb{R}^q$ be a bounded set. The *volume* $\mathcal{V}_{\mathrm{vol}}$ of $\mathcal{V}$ is defined as

$$\mathcal{V}_{\mathrm{vol}}\triangleq\int_{\mathcal{V}}\mathrm{d}\mathcal{V}. \qquad (11)$$

*Definition 4.2:* Let $\mathcal{V}\subset\mathbb{R}^q$ be a bounded set. A map $g:\mathcal{V}\to\mathcal{Q}$, where $\mathcal{Q}\subset\mathbb{R}^q$, is *volume-preserving* if for any $\mathcal{V}_0\subset\mathcal{V}$, the volume of $g(\mathcal{V}_0)$ is equal to the volume of $\mathcal{V}_0$.

The following theorem, known as Liouville's theorem [1], establishes sufficient conditions for volume-preserving flows. For the statement of this theorem, consider the nonlinear dynamical system (7) and define the divergence of $w=[w_1,...,w_q]^{\mathrm{T}}:\mathcal{D}\to\mathbb{R}^q$ by

$$\nabla\cdot w(z)\triangleq\sum_{i=1}^{q}\frac{\partial w_i(z)}{\partial z_i}, \qquad (12)$$

where $\nabla$ denotes the nabla operator, "$\cdot$" denotes the dot product in $\mathbb{R}^q$, and $z_i$ denotes the $i$th element of $z$.

*Theorem 4.1 ([1]):* Consider the nonlinear dynamical system (7). If $\nabla\cdot w(z)\equiv 0$, then the flow $s_t:\mathcal{D}\to\mathcal{D}$ of (7) is volume-preserving.

Volume preservation is the key conservation law underlying statistical mechanics. The flows of volume-preserving dynamical systems belong to one of the Lie pseudogroups of diffeomorphisms. These systems arise in incompressible fluid dynamics, classical mechanics, and acoustics. Next, we state the well known Poincaré recurrence theorem. For this result, let $g^{(n)}(z)$, $n\in\overline{\mathbb{Z}}_+$, denote the $n$-time composition operator of $g(z)$ with itself and define $g^{(0)}(z)\triangleq z$.

*Theorem 4.2:* Let $\mathcal{D}\subset\mathbb{R}^q$ be an open bounded set, and let $g:\mathcal{D}\to\mathcal{D}$ be a continuous, volume-preserving bijective (one-to-one and onto) map. Then for every open neighborhood $\mathcal{N}\subset\mathcal{D}$, there exists a dense subset $\mathcal{V}\subset\mathcal{N}$ such that for every point $z\in\mathcal{V}$, $\lim_{i\to\infty}g^{(n_i)}(z)=z$ for some sequence $\{n_i\}_{i=1}^{\infty}$, with $n_i\to\infty$ as $i\to\infty$.

It follows from Theorem 4.2 that almost every point in $\mathcal{D}\subset\mathbb{R}^q$ will return infinitely many times to any open neighborhood of itself under a continuous, volume-preserving bijective mapping which maps a bounded region $\mathcal{D}$ of a Euclidean space onto itself. For the remainder of this section we consider the nonlinear dynamical system (7) and assume that the solutions to (7) are defined for all $t\in\mathbb{R}$. Recall that if all solutions to (7) are bounded, then it follows from the Peano-Cauchy theorem [4, pp. 16, 17] that $\mathcal{I}_{z_0}=\mathbb{R}$. The following theorem shows that if a dynamical system preserves volume, then almost all trajectories return arbitrarily close to their initial position infinitely often.

*Theorem 4.3:* Consider the nonlinear dynamical system (7). Assume that the flow $s_t:\mathcal{D}\to\mathcal{D}$ of (7) is volume-preserving and maps an open bounded set $\mathcal{D}_{\mathrm{c}}\subset\mathbb{R}^q$ onto itself. Then the nonlinear dynamical system (7) exhibits Poincaré recurrence, that is, almost every point $z\in\mathcal{D}_{\mathrm{c}}$

returns to every open neighborhood $\mathcal{N} \subset \mathcal{D}_\mathrm{c}$ of $z$ infinitely many times.

All Hamiltonian dynamical systems of the form (9) and (10) exhibit Poincaré recurrence since they possess volume-preserving flows and are conservative in the sense that the Hamiltonian function $\mathcal{H}(q, p)$ remains constant along system trajectories. To see this, note that with $z \triangleq [q^\mathrm{T}, p^\mathrm{T}]^\mathrm{T}$, (9) and (10) can be rewritten as

$$\dot{z}(t) = \mathcal{J}\left(\frac{\partial \mathcal{H}}{\partial z}(z(t))\right)^\mathrm{T}, \quad z(t_0) = z_0, \quad t \geq t_0, \quad (13)$$

where $z_0 \triangleq [q_0^\mathrm{T}, p_0^\mathrm{T}]^\mathrm{T} \in \mathbb{R}^{2n}$ and $\mathcal{J} \triangleq \begin{bmatrix} 0_n & I_n \\ -I_n & 0_n \end{bmatrix}$. Now, since

$$\dot{\mathcal{H}}(z) = \left(\frac{\partial \mathcal{H}}{\partial z}(z)\right) \mathcal{J} \left(\frac{\partial \mathcal{H}}{\partial z}(z)\right)^\mathrm{T} = 0, \quad z \in \mathbb{R}^{2n}, \quad (14)$$

the Hamiltonian function $\mathcal{H}(\cdot)$ is conserved along the flow of (13). If $\mathcal{H}(\cdot)$ is bounded from below and is radially unbounded, then every trajectory of the Hamiltonian system (13) is bounded. Hence, by choosing the bounded region $\mathcal{D} \triangleq \{z \in \mathbb{R}^{2n} : \mathcal{H}(z) \leq \eta\}$, where $\eta \in \mathbb{R}$ and $\eta > 0$, it follows that the flow $s_t(\cdot)$ of (13) maps the bounded region $\mathcal{D}$ onto itself. Since $\eta > 0$ is arbitrary, the region $\mathcal{D}$ can be chosen arbitrarily large. Furthermore, since (13) possesses unique solutions over $\mathbb{R}$, it follows that the mapping $s_t(\cdot)$ is one-to-one and onto. Moreover,

$$\nabla \cdot \mathcal{J}\left(\frac{\partial \mathcal{H}}{\partial z}(z)\right)^\mathrm{T} = 0, \quad z \in \mathbb{R}^{2n}, \quad (15)$$

which, by Theorem 4.1, shows that the flow $s_t(\cdot)$ of (13) is volume-preserving. Finally, since the flow $s_t(\cdot)$ of (13) is volume-preserving, continuous, and bijective, and $s_t(\cdot)$ maps a bounded region of a Euclidean space onto itself, it follows from Theorem 4.3 that the Hamiltonian dynamical system (13) exhibits Poincaré recurrence. That is, in any open neighborhood $\mathcal{N}$ of any point $z_0 \in \mathbb{R}^{2n}$ there exists a point $y \in \mathcal{N}$ such that the trajectory $s(t, y)$, $t \geq t_0$, of (13) will return to $\mathcal{N}$ infinitely many times.

Poincaré recurrence has been the main source for the long and fierce debate between the microscopic and macroscopic points of view of thermodynamics [1]. In thermodynamic models predicated on statistical mechanics, an isolated dynamical system will return arbitrarily close to its initial state of molecular positions and velocities infinitely often. If the system entropy is determined by the state variables, then it must also return arbitrarily close to its original value, and hence, undergo cyclical changes. This apparent contradiction between the behavior of a mechanical system of particles and the second law of thermodynamics remains one of the hardest and most controversial problems in statistical physics. The resolution of this paradox lies in the controversial statement that as system dimensionality increases, the recurrence time increases at an extremely fast rate. Nevertheless, the shortcoming of the mechanistic world view of thermodynamics is the absence of the emergence of damping in lossless mechanical systems. The emergence of damping is, however, ubiquitous in isolated thermodynamic systems. Hence, the development of a viable dynamical system model for thermodynamics must guarantee the absence of Poincaré recurrence. The next set of results presents sufficient conditions for the absence of Poincaré recurrence for the nonlinear dynamical system (7). For these results define the set of equilibria for the nonlinear dynamical system (7) in $\mathcal{D}$ by $\mathcal{M}_\mathrm{e} \triangleq \{z \in \mathcal{D} : w(z) = 0\}$.

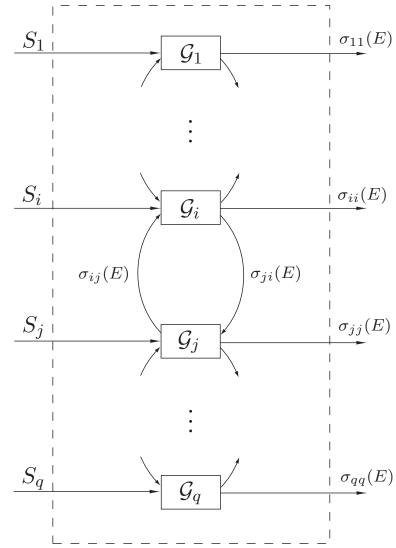


Fig. 1. Large-scale dynamical system $\mathcal{G}$.

*Theorem 4.4:* Consider the nonlinear dynamical system (7) and assume that $\mathcal{D} \setminus \mathcal{M}_\mathrm{e} \neq \emptyset$. Assume that there exists a continuous function $V : \mathcal{D} \to \mathbb{R}$ such that for every $z_0 \in \mathcal{D} \setminus \mathcal{M}_\mathrm{e}$, $V(s(t, z_0))$, $t \geq t_0$, is a strictly increasing (respectively, decreasing) function of time. Then the nonlinear dynamical system (7) does not exhibit Poincaré recurrence on $\mathcal{D} \setminus \mathcal{M}_\mathrm{e}$. That is, for some $z \in \mathcal{D} \setminus \mathcal{M}_\mathrm{e}$, there exists a neighborhood $\mathcal{N} \subset \mathcal{D} \setminus \mathcal{M}_\mathrm{e}$ such that for every $y \in \mathcal{N}$, $y \notin \omega(y)$.

The next result gives an alternative sufficient condition for the absence of Poincaré recurrence in a dynamical system. For this result, let $\mathcal{D}_\mathrm{c} \subseteq \mathcal{D}$ be a closed invariant set with respect to the nonlinear dynamical system (7).

*Theorem 4.5:* Consider the nonlinear dynamical system (7). Assume that $\mathcal{D}_\mathrm{c} \setminus \mathcal{M}_\mathrm{e} \neq \emptyset$ and assume (7) is convergent and semistable in $\mathcal{D}_\mathrm{c}$. Then the nonlinear dynamical system (7) does not exhibit Poincaré recurrence in $\mathcal{D}_\mathrm{c} \setminus \mathcal{M}_\mathrm{e}$. That is, for some $z \in \mathcal{D}_\mathrm{c} \setminus \mathcal{M}_\mathrm{e}$, there exists an open neighborhood $\mathcal{N} \subset \mathcal{D}_\mathrm{c} \setminus \mathcal{M}_\mathrm{e}$ such that for any $y \in \mathcal{N}$ the trajectory $s(t, y)$, $t \geq t_0$, does not return to $\mathcal{N}$ infinitely many times.

## V. System Thermodynamics

To formulate our state space thermodynamic model, consider the large-scale dynamical system $\mathcal{G}$ shown in Figure 1 involving energy exchange between $q$ interconnected subsystems. Let $E_i : [0, \infty) \to \overline{\mathbb{R}}_+$ denote the energy (and hence a nonnegative quantity) of the $i$th subsystem, let $S_i : [0, \infty) \to \mathbb{R}$ denote the external power (heat flux) supplied to (or extracted from) the $i$th subsystem, let $\sigma_{ij} : \overline{\mathbb{R}}_+^q \to \overline{\mathbb{R}}_+$, $i \neq j$, $i, j = 1, ..., q$, denote the instantaneous rate of energy (heat) flow from the $j$th subsystem to the $i$th subsystem, and let $\sigma_{ii} : \overline{\mathbb{R}}_+^q \to \overline{\mathbb{R}}_+$, $i = 1, ..., q$, denote the instantaneous rate of energy (heat) dissipation from the $i$th subsystem to the environment. Here we assume that $\sigma_{ij} : \overline{\mathbb{R}}_+^q \to \overline{\mathbb{R}}_+$, $i, j = 1, ..., q$, are locally Lipschitz continuous on $\overline{\mathbb{R}}_+^q$ and $S_i : [0, \infty) \to \mathbb{R}$, $i = 1, ..., q$, are bounded piecewise continuous functions of time.

An *energy balance* for the $i$th subsystem yields

$$E_i(T) = E_i(t_0) + \sum_{j=1,\, j\neq i}^{q} \int_{t_0}^{T} [\sigma_{ij}(E(t)) - \sigma_{ji}(E(t))]\mathrm{d}t$$

$$-\int_{t_0}^{T}\sigma_{ii}(E(t))\mathrm{d}t+\int_{t_0}^{T}S_i(t)\mathrm{d}t,\quad T\geq t_0,\qquad(16)$$

or, equivalently, in vector form,

$$E(T) = E(t_0)+\int_{t_0}^{T}w(E(t))\mathrm{d}t-\int_{t_0}^{T}d(E(t))\mathrm{d}t + \int_{t_0}^{T}S(t)\mathrm{d}t,\quad T\geq t_0,\qquad(17)$$

where $E(t) \triangleq [E_1(t),...,E_q(t)]^{\mathrm{T}}$, $d(E(t)) \triangleq [\sigma_{11}(E(t)),...,\sigma_{qq}(E(t))]^{\mathrm{T}}$, $S(t) \triangleq [S_1(t),..., S_q(t)]^{\mathrm{T}}$, $t\geq t_0$, and $w=[w_1,...,w_q]^{\mathrm{T}}:\overline{\mathbb{R}}_+^q\to\mathbb{R}^q$ is such that

$$w_i(E)=\sum_{j=1,\,j\neq i}^{q}[\sigma_{ij}(E)-\sigma_{ji}(E)],\quad E\in\overline{\mathbb{R}}_+^q.\qquad(18)$$

It is important to note that the exchange of energy between subsystems in (16) is assumed to be a nonlinear function of all the subsystems, that is, $\sigma_{ij}=\sigma_{ij}(E)$, $E\in\overline{\mathbb{R}}_+^q$, $i\neq j$, $i,j=1,...,q$. This assumption is made for generality and would depend on the complexity of the diffusion process. For example, thermal processes may include evaporative and radiative heat transfer as well as thermal conduction giving rise to complex heat transport mechanisms. However, for simple diffusion processes it suffices to assume that $\sigma_{ij}(E)=\sigma_{ij}(E_j)$, wherein the energy flow from the $j$th subsystem to the $i$th subsystem is only dependent (possibly nonlinearly) on the energy in the $j$th subsystem. Similar comments apply to system dissipation.

Note that (16) yields a conservation of energy equation and implies that the energy stored in the $i$th subsystem is equal to the external energy supplied to (or extracted from) the $i$th subsystem plus the energy gained by the $i$th subsystem from all other subsystems due to subsystem coupling minus the energy dissipated from the $i$th subsystem to the environment. Equivalently, (16) can be rewritten as

$$\dot{E}(t)=w(E(t))-d(E(t))+S(t),\quad E(t_0)=E_0,\quad(19)$$

where $t\geq t_0$ and $E_0\triangleq[E_{10},...,E_{q0}]^{\mathrm{T}}$, yielding a *power balance* equation that characterizes energy flow between subsystems of the large-scale dynamical system $\mathcal{G}$. Equation (19) shows that the rate of change of energy, or power, in the $i$th subsystem is equal to the power input (heat flux) to the $i$th subsystem plus the energy (heat) flow to the $i$th subsystem from all other subsystems minus the power dissipated from the $i$th subsystem to the environment. Furthermore, since $w(\cdot)-d(\cdot)$ is locally Lipschitz continuous on $\overline{\mathbb{R}}_+^q$ and $S(\cdot)$ is a bounded piecewise continuous function of time, it follows that (19) has a unique solution over the finite time interval $[t_0,\tau_{E_0})$. If, in addition, the power balance equation (19) is *input-to-state stable*, then $\tau_{E_0}=\infty$.

Equation (17) or, equivalently, (19) is a statement of the *first law of thermodynamics* as applied to *isochoric transformations* (i.e., constant subsystem volume transformations) for each of the subsystems $\mathcal{G}_i$, $i=1,...,q$, with $E_i(\cdot)$, $S_i(\cdot)$, $\sigma_{ij}(\cdot)$, $i\neq j$, and $\sigma_{ii}(\cdot)$, $i,j=1,...,q$, playing the role of the $i$th subsystem internal energy, rate of heat supplied to (or extracted from) the $i$th subsystem, heat flow between subsystems due to coupling, and the rate of energy (heat) dissipated to the environment, respectively. To further elucidate that (17) is essentially the statement of the principle of the conservation of energy, let the total energy in the large-scale dynamical system $\mathcal{G}$ be given by $U\triangleq\mathbf{e}^{\mathrm{T}}E$, where $\mathbf{e}^{\mathrm{T}}\triangleq[1,...,1]$ and $E\in\overline{\mathbb{R}}_+^q$, and let the net energy received by the large-scale dynamical system $\mathcal{G}$ over the time interval $[t_1,t_2]$ be given by

$$Q\triangleq\int_{t_1}^{t_2}\mathbf{e}^{\mathrm{T}}[S(t)-d(E(t))]\mathrm{d}t,\qquad(20)$$

where $E(t)$, $t\geq t_0$, is the solution to (19). Then, premultiplying (17) by $\mathbf{e}^{\mathrm{T}}$ and using the fact that $\mathbf{e}^{\mathrm{T}}w(E)\equiv 0$, it follows that $\Delta U=Q$, where $\Delta U\triangleq U(t_2)-U(t_1)$ denotes the variation in the total energy of the large-scale dynamical system $\mathcal{G}$ over the time interval $[t_1,t_2]$. This is a statement of the first law of thermodynamics for isochoric transformations of the large-scale dynamical system $\mathcal{G}$ and gives a precise formulation of the equivalence between the variation in system internal energy and heat.

It is important to note that the large-scale dynamical system model (19) does not consider work done by the system on the environment nor work done by the environment on the system. Hence, $Q$ can be physically interpreted as the net amount of energy that is received by the system in forms other than work. The extension of addressing work performed by and on the system can be easily addressed by including an additional state equation, coupled to the power balance equation (19), involving volume (deformation) states for each subsystem. Since this extension does not alter any of the conceptual results of this paper, it is not considered in this paper for simplicity of exposition. Work performed by the system on the environment and work done by the environment on the system is addressed in [1].

For our large-scale dynamical system model $\mathcal{G}$, we assume that $\sigma_{ij}(E)=0$, $E\in\overline{\mathbb{R}}_+^q$, whenever $E_j=0$, $i,j=1,...,q$. In this case, $w(E)-d(E)$, $E\in\overline{\mathbb{R}}_+^q$, is essentially nonnegative. The above constraint implies that if the energy of the $j$th subsystem of $\mathcal{G}$ is zero, then this subsystem cannot supply any energy to its surroundings nor dissipate energy to the environment. Moreover, we assume that $S_i(t)\geq 0$ whenever $E_i(t)=0$, $t\geq t_0$, $i=1,...,q$, which implies that when the energy of the $i$th subsystem is zero, then no energy can be extracted from this subsystem. Under these assumptions, it can be shown (see [1] for details) that the solution $E(t)$, $t\geq t_0$, to (19) is nonnegative for all nonnegative initial conditions $E_0\in\overline{\mathbb{R}}_+^q$.

## VI. Entropy and Irreversibility

The nonlinear power balance equation (19) can exhibit a full range of nonlinear behavior, including bifurcations, limit cycles, and even chaos. However, a thermodynamically consistent energy flow model should ensure that the evolution of the system energy is diffusive (parabolic) in character with convergent subsystem energies. As established in Section IV, such a system model would guarantee the absence of Poincaré recurrence. Otherwise, the thermodynamic model would violate the second law of thermodynamics, since subsystem energies (temperatures) would be allowed to return to their starting state and thereby subverting the diffusive character of the dynamical system. Hence, to ensure a thermodynamically consistent energy flow model, we require the following axioms. For the statement of these axioms, we first recall the following graph-theoretic notions.

*Definition 6.1 ([6]):* A *directed graph* $G(\mathcal{C})$ associated with the *connectivity matrix* $\mathcal{C}\in\mathbb{R}^{q\times q}$ has *vertices* $\{1,2,...,q\}$ and an *arc* from vertex $i$ to vertex $j$, $i\neq j$, if and only if $\mathcal{C}_{(j,i)}\neq 0$. A *graph* $G(\mathcal{C})$ associated with the connectivity matrix $\mathcal{C}\in\mathbb{R}^{q\times q}$ is a directed graph for which the *arc set* is symmetric, that is, $\mathcal{C}=\mathcal{C}^{\mathrm{T}}$. We say that $G(\mathcal{C})$ is *strongly connected* if for any ordered pair of vertices $(i,j)$,

$i \neq j$, there exists a *path* (i.e., a sequence of arcs) leading from $i$ to $j$.

Recall that the connectivity matrix $\mathcal{C} \in \mathbb{R}^{q\times q}$ is *irreducible*, that is, there does not exist a permutation matrix such that $\mathcal{C}$ is cogredient to a lower-block triangular matrix, if and only if $G(\mathcal{C})$ is strongly connected (see Theorem 2.7 of [6]). Let $\phi_{ij}(E) \triangleq \sigma_{ij}(E) - \sigma_{ji}(E)$, $E \in \overline{\mathbb{R}}_+^q$, denote the net energy flow from the $j$th subsystem $\mathcal{G}_j$ to the $i$th subsystem $\mathcal{G}_i$ of the large-scale dynamical system $\mathcal{G}$.

**Axiom *i*)** For the connectivity matrix $\mathcal{C} \in \mathbb{R}^{q\times q}$ associated with the large-scale dynamical system $\mathcal{G}$ defined by

$$\mathcal{C}_{(i,j)} \triangleq \begin{cases} 0, & \text{if } \phi_{ij}(E) \equiv 0, \\ 1, & \text{otherwise,} \end{cases} \quad i \neq j, \ i,j = 1,...,q, \tag{21}$$

and

$$\mathcal{C}_{(i,i)} \triangleq -\sum_{k=1,\, k\neq i}^{q} \mathcal{C}_{(k,i)}, \quad i = j, \quad i = 1,...,q, \tag{22}$$

rank $\mathcal{C} = q-1$, and for $\mathcal{C}_{(i,j)} = 1$, $i \neq j$, $\phi_{ij}(E) = 0$ if and only if $E_i = E_j$.

**Axiom *ii*)** For $i,j = 1,...,q$, $(E_i - E_j)\phi_{ij}(E) \leq 0$, $E \in \overline{\mathbb{R}}_+^q$.

The fact that $\phi_{ij}(E) = 0$ if and only if $E_i = E_j$, $i \neq j$, implies that subsystems $\mathcal{G}_i$ and $\mathcal{G}_j$ of $\mathcal{G}$ are *connected*; alternatively, $\phi_{ij}(E) \equiv 0$ implies that $\mathcal{G}_i$ and $\mathcal{G}_j$ are *disconnected*. Axiom $i$) implies that if the energies in the connected subsystems $\mathcal{G}_i$ and $\mathcal{G}_j$ are equal, then energy exchange between these subsystems is not possible. This statement is consistent with the *zeroth law of thermodynamics*, which postulates that temperature equality is a necessary and sufficient condition for thermal equilibrium. Furthermore, it follows from the fact that $\mathcal{C} = \mathcal{C}^{\mathrm{T}}$ and rank $\mathcal{C} = q-1$ that the connectivity matrix $\mathcal{C}$ is irreducible, which implies that for any pair of subsystems $\mathcal{G}_i$ and $\mathcal{G}_j$, $i \neq j$, of $\mathcal{G}$ there exists a sequence of connectors (arcs) of $\mathcal{G}$ that connect $\mathcal{G}_i$ and $\mathcal{G}_j$. Axiom $ii$) implies that energy flows from more energetic subsystems to less energetic subsystems and is consistent with the *second law of thermodynamics*, which states that heat (energy) must flow in the direction of lower temperatures. Furthermore, note that $\phi_{ij}(E) = -\phi_{ji}(E)$, $E \in \overline{\mathbb{R}}_+^q$, $i \neq j$, $i,j = 1,...,q$, which implies conservation of energy between lossless subsystems. With $S(t) \equiv 0$, Axioms $i$) and $ii$) along with the fact that $\phi_{ij}(E) = -\phi_{ji}(E)$, $E \in \overline{\mathbb{R}}_+^q$, $i \neq j$, $i,j = 1,...,q$, imply that at a given instant of time, energy can only be transported, stored, or dissipated but not created, and the maximum amount of energy that can be transported and/or dissipated from a subsystem cannot exceed the energy in the subsystem.

Next, we show that the classical Clausius equality and inequality for reversible and irreversible thermodynamics over cyclic motions are satisfied for our thermodynamically consistent energy flow model. For this result $\oint$ denotes a cyclic integral evaluated along an arbitrary closed path of (19) in $\overline{\mathbb{R}}_+^q$; that is, $\oint \triangleq \int_{t_0}^{t_{\mathrm{f}}}$ with $t_{\mathrm{f}} \geq t_0$ and $S(\cdot) \in \mathcal{U}$ such that $E(t_{\mathrm{f}}) = E(t_0) = E_0 \in \overline{\mathbb{R}}_+^q$.

*Proposition 6.1:* Consider the large-scale dynamical system $\mathcal{G}$ with power balance equation (19), and assume that Axioms $i$) and $ii$) hold. Then for all $E_0 \in \overline{\mathbb{R}}_+^q$, $t_{\mathrm{f}} \geq t_0$, and $S(\cdot) \in \mathcal{U}$ such that $E(t_{\mathrm{f}}) = E(t_0) = E_0$,

$$\int_{t_0}^{t_{\mathrm{f}}} \sum_{i=1}^{q} \frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \mathrm{d}t = \oint \sum_{i=1}^{q} \frac{\mathrm{d}Q_i(t)}{c + E_i(t)} \leq 0, \tag{23}$$

where $c > 0$, $\mathrm{d}Q_i(t) \triangleq [S_i(t) - \sigma_{ii}(E(t))]\mathrm{d}t$, $i = 1,...,q$, is the amount of net energy (heat) received by the $i$th subsystem over the infinitesimal time interval $\mathrm{d}t$, and $E(t)$, $t \geq t_0$, is the solution to (19) with initial condition $E(t_0) = E_0$. Furthermore, (23) holds as an equality if and only if there exists a continuous function $\alpha : [t_0, t_{\mathrm{f}}] \to \overline{\mathbb{R}}_+$ such that $E(t) = \alpha(t)\mathbf{e}$, $t \in [t_0, t_{\mathrm{f}}]$.

Inequality (23) is a generalization of Clausius' inequality for reversible and irreversible thermodynamics as applied to large-scale dynamical systems and restricts the manner in which the system dissipates (scaled) heat over cyclic motions. It follows from Axiom $i$) and (19) that for the *adiabatically isolated* large-scale dynamical system $\mathcal{G}$ (that is, $S(t) \equiv 0$ and $d(E(t)) \equiv 0$), the energy states given by $E_{\mathrm{e}} = \alpha\mathbf{e}$, $\alpha \geq 0$, correspond to the equilibrium energy states of $\mathcal{G}$. Thus, as in classical thermodynamics, we can define an *equilibrium process* as a process in which the trajectory of the large-scale dynamical system $\mathcal{G}$ moves along the equilibrium manifold $\mathcal{M}_{\mathrm{e}} \triangleq \{E \in \overline{\mathbb{R}}_+^q : E = \alpha\mathbf{e}, \alpha \geq 0\}$ corresponding to the set of equilibria of the isolated system $\mathcal{G}$. The power input that can generate such a trajectory can be given by $S(t) = d(E(t)) + u(t)$, $t \geq t_0$, where $u(\cdot) \in \mathcal{U}$ is such that $u_i(t) \equiv u_j(t)$, $i \neq j$, $i,j = 1,...,q$. Our definition of an equilibrium transformation involves a continuous succession of intermediate states that differ by infinitesimals from equilibrium system states and thus can only connect initial and final states, which are states of equilibrium. This process need not be slowly varying, and hence, equilibrium and quasistatic processes are not synonymous in this paper. Alternatively, a *nonequilibrium process* is a process that does not lie on the equilibrium manifold $\mathcal{M}_{\mathrm{e}}$. Hence, it follows from Axiom $i$) that for an equilibrium process $\phi_{ij}(E(t)) = 0$, $t \geq t_0$, $i \neq j$, $i,j = 1,...,q$, and thus, by Proposition 6.1, inequality (23) is satisfied as an equality. Alternatively, for a nonequilibrium process it follows from Axioms $i$) and $ii$) that (23) is satisfied as a strict inequality.

Next, we give a deterministic definition of entropy for the large-scale dynamical system $\mathcal{G}$ that is consistent with the classical thermodynamic definition of entropy.

*Definition 6.2:* For the large-scale dynamical system $\mathcal{G}$ with power balance equation (19), a function $\mathcal{S} : \overline{\mathbb{R}}_+^q \to \mathbb{R}$ satisfying

$$\mathcal{S}(E(t_2)) \geq \mathcal{S}(E(t_1)) + \int_{t_1}^{t_2} \sum_{i=1}^{q} \frac{S_i(t) - \sigma_{ii}(E(t))}{c + E_i(t)} \mathrm{d}t \tag{24}$$

for any $t_2 \geq t_1 \geq t_0$ and $S(\cdot) \in \mathcal{U}$ is called the *entropy* function of $\mathcal{G}$.

Next, we establish the existence of a *unique*, *continuously differentiable* entropy function for $\mathcal{G}$ for equilibrium and nonequilibrium processes. This result answers the long-standing question of how the entropy of a nonequilibrium state of a dynamical process should be defined [7], [8], and establishes its global existence and uniqueness.

*Theorem 6.1:* Consider the large-scale dynamical system $\mathcal{G}$ with power balance equation (19), and assume that Axioms $i$) and $ii$) hold. Then the function $\mathcal{S} : \overline{\mathbb{R}}_+^q \to \overline{\mathbb{R}}_+^q$ given by

$$\mathcal{S}(E) = \mathbf{e}^{\mathrm{T}}\mathbf{log}_e(c\mathbf{e} + E) - q\log_e c, \quad E \in \overline{\mathbb{R}}_+^q, \tag{25}$$

where $\mathbf{log}_e(c\mathbf{e} + E) \triangleq [\log_e(c + E_1), ..., \log_e(c + E_q)]^{\mathrm{T}}$ and $c > 0$ is a unique (modulo a constant of integration), continuously differentiable entropy function of $\mathcal{G}$. Furthermore, for $E(t) \notin \mathcal{M}_{\rm e}$, $t \geq t_0$, where $E(t)$, $t \geq t_0$, denotes the solution to (19) and $\mathcal{M}_{\rm e} = \{E \in \overline{\mathbb{R}}_+^q : E = \alpha\mathbf{e}, \alpha \geq 0\}$, (25) satisfies (24) as a strict inequality.

Note that it follows from Axiom $i)$ and Axiom $ii)$ that the entropy function given by (25) satisfies (24) as an equality for an equilibrium process and as a strict inequality for a nonequilibrium process. Hence, it follows from Theorem 4.4 that the isolated (i.e., $S(t) \equiv 0$ and $d(E) \equiv 0$) large-scale dynamical system $\mathcal{G}$ does not exhibit Poincaré recurrence in $\overline{\mathbb{R}}_+^q \setminus \mathcal{M}_{\rm e}$. The entropy expression given by (25) is identical in form to the Boltzmann entropy for statistical thermodynamics. Due to the fact that the entropy given by (25) is indeterminate to the extent of an additive constant, we can place the constant of integration $q \log_e c$ to zero by taking $c = 1$. Since $\mathcal{S}(E)$ given by (25) achieves a maximum when all the subsystem energies $E_i$, $i = 1, ..., q$, are equal [1], the entropy of $\mathcal{G}$ can be thought of as a measure of the tendency of a system to lose the ability to do useful work, lose order, and settle to a more homogenous state.

Finally, using the system entropy function given by (25) we show that our large-scale dynamical system $\mathcal{G}$ with power balance equation (19) is state irreversible for every nontrivial (nonequilibrium) trajectory of $\mathcal{G}$. For this result, let $\mathcal{W}_{[t_0,t_1]}$ denote the set of all possible energy trajectories of $\mathcal{G}$ over the time interval $[t_0, t_1]$ given by

$$\mathcal{W}_{[t_0,t_1]} \triangleq \{s^E : [t_0, t_1] \times \mathcal{U} \to \overline{\mathbb{R}}_+^q : s^E(\cdot, S(\cdot)) \text{ satisfies (19)}\}, \tag{26}$$

and let $\mathcal{M}_{\rm e} \subset \overline{\mathbb{R}}_+^q$ denote the set of equilibria of the isolated system $\mathcal{G}$ given by $\mathcal{M}_{\rm e} = \{E \in \overline{\mathbb{R}}_+^q : \alpha\mathbf{e}, \alpha \geq 0\}$.

*Theorem 6.2:* Consider the large-scale dynamical system $\mathcal{G}$ with power balance equation (19), and assume Axioms $i)$ and $ii)$ hold. Furthermore, let $s^E(\cdot, S(\cdot)) \in \mathcal{W}_{[t_0,t_1]}$, where $S(\cdot) \in \mathcal{U}$. Then $s^E(\cdot, S(\cdot))$ is an $I_q$-reversible trajectory of $\mathcal{G}$ if and only if $s^E(t, S(t)) \in \mathcal{M}_{\rm e}$, $t \in [t_0, t_1]$.

Theorem 6.2 establishes an equivalence between (non)equilibrium and state (ir)reversible thermodynamic systems. Furthermore, Theorem 6.2 shows that for every $E_0 \notin \mathcal{M}_{\rm e}$, the large-scale dynamical system $\mathcal{G}$ is state irreversible. In addition, since state irrecoverability implies state irreversibility and, by Theorem 6.2, state irreversibility is equivalent to $E(t) \notin \mathcal{M}_{\rm e}$, $t \geq t_0$, it follows from Theorem 3.2 that state (ir)reversibility and state (ir)recoverability are equivalent for our thermodynamically consistent large-scale dynamical system $\mathcal{G}$. Hence, in the remainder of the paper we use the notions of (non)equilibrium, state (ir)reversible, and state (ir)recoverable dynamical processes interchangeably.

## VII. Semistability and the Arrow of Time

For the isolated large-scale dynamical system $\mathcal{G}$, (24) yields the fundamental inequality

$$\mathcal{S}(E(t_2)) \geq \mathcal{S}(E(t_1)), \quad t_2 \geq t_1. \tag{27}$$

Inequality (27) implies that, for any dynamical change in an isolated large-scale dynamical system $\mathcal{G}$, the entropy of the final state can never be less than the entropy of the initial state. Inequality (27) is often identified with the second law of thermodynamics as a statement about entropy increase. Furthermore, it follows from Theorem 6.1 that for an isolated large-scale dynamical system $\mathcal{G}$ the entropy function (25) is a strictly increasing function of time along the trajectories of (19) with initial conditions in $\overline{\mathbb{R}}_+^q \setminus \mathcal{M}_{\rm e}$. Hence, it follows from Theorem 4.4 that the isolated large-scale dynamical system $\mathcal{G}$ does not exhibit Poincaré recurrence in $\overline{\mathbb{R}}_+^q \setminus \mathcal{M}_{\rm e}$. This result can also be arrived at using the fact that our thermodynamically consistent large-scale dynamical system $\mathcal{G}$ is semistable.

*Theorem 7.1:* Consider the large-scale dynamical system $\mathcal{G}$ with power balance equation (19) with $S(t) \equiv 0$ and $d(E) \equiv 0$, and assume that Axioms $i)$ and $ii)$ hold. Then for every $\alpha \geq 0$, $\alpha\mathbf{e}$ is a semistable equilibrium state of (19). Furthermore, $E(t) \to \frac{1}{q}\mathbf{e}\mathbf{e}^{\rm T}E(t_0)$ as $t \to \infty$ and $\frac{1}{q}\mathbf{e}\mathbf{e}^{\rm T}E(t_0)$ is a semistable equilibrium state.

Theorem 7.1 shows that the isolated (i.e., $S(t) \equiv 0$ and $d(E) \equiv 0$) large-scale dynamical system $\mathcal{G}$ is semistable. Hence, it follows from Theorem 4.5 that the isolated large-scale dynamical system $\mathcal{G}$ does not exhibit Poincaré recurrence in $\overline{\mathbb{R}}_+^q \setminus \mathcal{M}_{\rm e}$. Next, using the system entropy function given by (25), we show that our large-scale isolated dynamical system $\mathcal{G}$ with power balance equation (19) is state irreversible for all nonequilibrium trajectories of $\mathcal{G}$ establishing a clear connection between our thermodynamic model and the arrow of time.

*Theorem 7.2:* Consider the large-scale dynamical system $\mathcal{G}$ with power balance equation (19) with $S(t) \equiv 0$ and $d(E) \equiv 0$, and assume Axioms $i)$ and $ii)$ hold. Furthermore, let $s^E(\cdot, 0) \in \mathcal{W}_{[t_0,t_1]}$. Then for every $E_0 \notin \mathcal{M}_{\rm e}$, there exists a continuously differentiable function $\mathcal{S} : \overline{\mathbb{R}}_+^q \to \mathbb{R}$ such that $\mathcal{S}(s^E(t, 0))$ is a strictly increasing function of time. Furthermore, $s^E(\cdot, 0)$ is an $I_q$-reversible trajectory of $\mathcal{G}$ if and only if $s^E(t, 0) \in \mathcal{M}_{\rm e}$, $t \in [t_0, t_1]$.

Theorem 7.2 shows that for every $E_0 \notin \mathcal{M}_{\rm e}$, the isolated dynamical system $\mathcal{G}$ is state irreversible. This gives a clear connection between our thermodynamic model and the arrow of time. In particular, it follows from Corollary 3.1 and Theorem 7.2 that there exists a function of the system state that strictly increases in time on any nonequilibrium trajectory of $\mathcal{G}$ if and only if there does *not* exist a nonequilibrium reversible trajectory of $\mathcal{G}$. Thus, the existence of the continuously differentiable entropy function given by (25) for $\mathcal{G}$ establishes the existence of a completely ordered time set having a topological structure involving a closed set homeomorphic to the real line. This fact follows from the inverse function theorem of mathematical analysis and the fact that a continuous strictly monotonic function is a topological mapping (i.e., a homeomorphism), and conversely every topological mapping of a strictly monotonic function's domain onto its codomain must be strictly monotonic. This topological property gives a clear time-reversal asymmetry characterization of our thermodynamic model establishing an emergence of the direction of time flow.

## References


[1] W. M. Haddad, V. Chellaboina, and S. G. Nersesov, *Thermodynamics: A Dynamical Systems Approach*. Princeton, NJ: Princeton University Press, 2005.

[2] P. Bridgman, *The Nature of Thermodynamics*. Cambridge, MA: Harvard University Press, 1941, reprinted (Gloucester, MA: Peter Smith, 1969).

[3] J. Uffink, "Bluff your way in the second law of thermodynamics," *Stud. Hist. Phil. Mod. Phys.*, vol. 32, pp. 305–394, 2001.

[4] J. K. Hale, *Ordinary Differential Equations*, second edition ed. New York: Wiley, 1980, reprinted by Krieger, Malabar, 1991.

[5] H. Poincaré, "Sur le probléme des trois corps et les équations de la dynamique," *Acta Math.*, vol. 13, pp. 1–270, 1890.

[6] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. New York, NY: Academic Press, Inc., 1979.

[7] J. Meixner, "On the foundation of thermodynamics of processes," in *A Critical Review of Thermodynamics*, E. B. Stuart, B. Gal-Or, and A. J. Brainard, Eds. Mono Book Corp., pp. 37-47, 1970.

[8] B. Lavenda, *Thermodynamics of Irreversible Processes*. London: Macmillan, 1978, reprinted (New York: Dover, 1993).