# Classical thermodynamics and economic general equilibrium theory

Eric Smith
*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501*

Duncan K. Foley
*Department of Economics, Graduate Faculty, New School University, 65 Fifth Avenue, New York, NY 10003*
(Dated: December 15, 2005)

A long history of analogy making between neoclassical economics and physical thermodynamics has unfortunately served to obscure two important relations between the two fields: their definitions of equilibria stem from essentially the same three axioms for the mathematical representations of systems, while the classes of transformation each has chosen to emphasize, and their responses to the problem of path dependence, have led them to very different interpretations of duality in those representations. Despite these conventional differences, we show that economies in which all agents have preferences quasi-linear in some good have a trading-constraint structure isomorphic to the structure of physical systems with classical thermodynamic equations of state. Exact equivalents of thermodynamic potentials, including entropy, can be constructed, and function as the economic counterparts to *free energies*. Quasi-linear economies are the most general in which the Walrasian idea of price formation as an analog of force balance can be realized. More general economic models raise the same methodological problems as more complex physical models that exhibit path-dependence. We show how the degree of aggregatability of an economic model corresponds to which properties of equilibria retain path-independence, and to the extent to which a social-welfare function exists. A new *contour money-metric utility* defines the maximal generalization of social-welfare functions in arbitrary economies, but depends on the endowments and composition of the economy in non-quasi-linear cases, and is limited to one-dimensional contours of equilibria in non-aggregatable cases. The differences between economic and thermodynamic methodology lies in the economic focus on the irreversible movement from initial disequilibrium endowments to equilibrium through voluntary trade, in contrast to the thermodynamic recognition that only reversible transformations lead to measurement of system structure. The consequences of respecting reversibility for economic method are sketched, and alternative interpretations of the Walrasian notion of wealth preservation are presented.
JEL categories: B0, D5, D6
Keywords: Thermodynamics and economics, utility, economic potentials, economic entropy

## I. INTRODUCTION

The relation between economic and physical (particularly thermodynamic) concepts of equilibrium has been a topic of recurrent interest throughout the development of neoclassical economic theory. As systems for defining equilibria, proving their existence, and computing their properties, neoclassical economics [1, 2] and classical thermodynamics [3] undeniably have numerous formal and methodological similarities. Both fields seek to describe system phenomena in terms of solutions to constrained optimization problems. Both rely on *dual representations* of interacting subsystems: the state of each subsystem is represented by pairs of variables, one variable from each pair characterizing the subsystem's content, and the other characterizing the way it interacts with other subsystems. In physics the content variables are quantities like a subsystem's total energy or the volume in space it occupies; in economics they are amounts of various commodities held by agents. In physics the interaction variables are quantities like temperature and pressure that can be measured on the system boundaries; in economics they are prices that can be measured by an agent's willingness to trade one commodity for another.

The significance attached to these similarities has changed considerably, however, in the time from the first mathematical formulation of utility [4] to the full axiomatization of general equilibrium theory [5]. Léon Walras appears [7] to have conceptualized economic equilibrium as a balance of the gradients of utilities, more for the sake of similarity to the concept of force balance in mechanics, than to account for any observations about the outcomes of trade. Irving Fisher (a student of J. Willard Gibbs) attempted to update Walrasian metaphors from mechanics to thermodynamics [8], but retained Walras's program of seeking an explicit parallelism between physics and economics.

As mathematical economics has become more sophisticated [5] the naïve parallelism of Walras and Fisher has progressively been abandoned, and with it the sense that it matters whether neoclassical economics resembles any branch of physics. The cardinalization of utility that Walras thought of as a counterpart to energy has been discarded, apparently removing the possibility of comparing utility with any empirically measurable quantity. A long history of logically inconsistent (or simply unproductive) analogy making (see section /refsec:litrev below) has further caused the topic of parallels to fall out of favor. Paul Samuelson summarizes well [12] the current view among many economists, at the end of a one of the few methodologically sound analyses of the parallel roles of dual representation in economics and physics:

> The formal mathematical analogy between classical thermodynamics and mathematic economic systems has now been explored. This does not warrant the commonly met attempt to find more exact analogies of physical magnitudes – such as entropy or energy – in the economic realm. Why should there be laws like the first or second laws of thermodynamics holding in the economic realm? Why should "utility" be literally identified with entropy, energy, or anything else? Why should a failure to make such a successful identification lead anyone to overlook or deny the mathematical isomorphism that does exist between minimum systems that arise in different disciplines?

The view that neoclassical economics is now mathematically mature, and that it is mere coincidence and no longer relevant whether it overlaps with any body of physical theory, is reflected in the complete omission the topic of parallels from contemporary graduate texts [1].

We argue here that, despite its long history of discussion, there are important insights still to be gleaned from considering the relation of neoclassical economics to classical thermodynamics. The new results concerning this relation we present here have significant implications, both for the interpretation of economic theory and for econometrics. The most important point of this paper (more important than the establishment of formal parallels between the thermodynamics and utility economics) is that economics, because it does not recognize an equation of state or define prices intrinsically in terms of equilibrium, lacks the close relation between measurement and theory physical thermodynamics enjoys.

### A. Structure of the paper

It has been our experience that the conceptual mismatch between thermodynamics and neoclassical economics makes it very difficult for readers familiar with one field to grasp the organizational framework of the other, a situation only rendered more confusing by their superficial formal similarities. Therefore we begin in Sections II and III with brief reviews of the theories of economic exchange and classical thermodynamics. These sections fix terminology and notation, and also establish a conceptual orientation needed to understand the correspondence between the two theories. Here we also introduce canonical examples that are developed in progressively more detail throughout the rest of the paper.

By the end of Sec. III we will have shown that quasi-linear economies include an entropy function and a form of path independence identical to those of thermodynamic systems, but will not yet have demonstrated that these are the only economies that do, or explained why. We choose to begin with these constructions, which are the most important of the paper, so the reader can see where we are going, and then to descend in Sec. IV to the deeper assumptions underlying the correspondence. Here we show that the three basic "laws" defining classical thermodynamics have exact counterparts for quasi-linear economies, and in which respects they are weakened for path-dependent systems, both physical and economic. We then finish the correspondence of the full dual structure between classical thermodynamic systems and quasi-linear economies.

Sec. V makes explicit some crucial methodological implications of the demonstrated equivalences for economic theory. In particular the treatment of market disequilibrium in much of economics is flawed by the assumption that well-defined market prices exist when the economy as a whole is out of equilibrium, and the Walrasian correspondence between initial endowments and final equilibrium allocations on which much economic reasoning is based is flawed by its failure to respect the distinction between reversible and irreversible transformations.

By this point in the paper the reader will have seen the largest domain within neoclassical theory for which deterministic predictions can be made without an explicit theory of dynamics – the domain for which the correspondence to classical thermodynamics goes through entirely. In Sec. VI we consider the problem of generalizing these insights to non-quasi-linear economies, where path dependence sets in. (Since this section raises more technical mathematical issues than the rest of the paper, the reader may choose to omit it on a first reading.) We show that the concept of Gorman aggregatability is intermediate between the strong aggregatability of quasi-linear economies, and the non-aggregatability of the general case. Closed Gorman economies preserve path-independence of equilibrium prices, and admit entropies and social welfare functions that provide useful descriptions of effeciency and the value of trade. However, all of these constructions depend on the composition and endowments of each particular economy, so Gorman aggregatability is not strong enough to admit the equivalent of a thermodynamic equation of state. For economies less aggregatable than Gorman forms, we show that the generalization of the Negishi construction of social welfare is the best one can do, but that it says nothing about the indeterminacy of equilibria.

We draw together our findings in Sec. VII and relate them briefly to the literature on this problem.

## II. ECONOMIC FUNDAMENTALS

In this section we present the basic theory of an exchange economy.

### A. Commodities, utility and offer prices

In an economy with $n+1$ commodities, a commodity bundle can be written $x=(x_0,x_1,\ldots,x_n)$, and a system of prices $p=(p_0,p_1,\ldots,p_n)$. An economic agent, $j$, has a preference ordering over non-negative commodity bundles that can be represented by a quasi-concave, differentiable *ordinal utility function* $u^j[.]:R^{n+1}\to R$. At any commodity bundle $x^j$, the gradient of the utility function, $u^{j\prime}[x]$, is a vector of marginal utilities, and the agent's *marginal rate of substitution* between goods $i$ and $k$ is $u_i^{j\prime}[x]\,/u_k^{j\prime}[x]$. The marginal rate of substitution, which has the dimensions units of commodity $k$ per unit of commodity $i$, is the ratio of the maximum amount of commodity $k$ the agent will voluntarily exchange for a small amount of commodity $i$, and can also be thought of as the agent's *offer price* for commodity $i$ in terms of commodity $k$. The utility function is ordinal in the sense that any strictly monotonic transformation of $u^j[.]$ will represent the same preference ordering, and define the same system of marginal rates of substitution. In what follows we will suppress the superscript $j$ indicating the individual agent in contexts where no confusion is likely.

The mathematical dual to an agent's utility function is the *expenditure function*:

$$e\,[p,\mathcal{U}]\equiv\min_x\left[p\cdot x\mid u\,[x]\geq\mathcal{U}\right], \tag{1}$$

where $p\cdot x\equiv\sum_i p_i x_i$. Its general variation is

$$\delta e=\delta p\cdot x+p\cdot\left.\frac{\partial x}{\partial\mathcal{U}}\right|_p\delta\mathcal{U}, \tag{2}$$

as $p\cdot\delta x|_{\mathcal{U}}\equiv 0$ about $x^h\equiv\operatorname{argmin}_x\left[p\cdot x\mid u[x]\geq\mathcal{U}\right]$.

The expenditure function is derived from the *Lagrangian*

$$\mathcal{L}[x,\lambda]\equiv p\cdot x-\lambda\left(u\,[x]-\mathcal{U}\right) \tag{3}$$

by solving the *first-order conditions*

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial x}&=p-\lambda u'\,[x]=0\\ \frac{\partial\mathcal{L}}{\partial\lambda}&=\mathcal{U}-u\,[x]=0.\end{aligned} \tag{4}$$

Economists refer to the inverse $1/\lambda$ of the Lagrange multiplier as the *marginal utility of wealth.* In general $\lambda$ depends on the utility level, $\mathcal{U}$, and these $n+2$ equations must be solved simultaneously for $x,\lambda$. The commodity bundle at which expenditure is minimized, $x^h[p,\mathcal{U}]$, is the *Hicksian demand function.*[1]

[1] In writing the first-order conditions as equalities we are implicitly assuming that the minimum is interior to the nonnegative orthant of the commodity space.

#### 1. Example 1: Cobb-Douglas utility

The *Cobb-Douglas* (CD) utility function for two-commodities is:

$$u_{CD}[x_1,x_2]=\alpha_1\log[x_1]+\alpha_2\log[x_2]\,, \tag{5}$$

where we will use the ordinal property of the utility function to set $\alpha_1+\alpha_2=1$.

The offer prices for the Cobb-Douglas utility with respect to commodity 1 are $u'_{CD}[x]=(1,(\alpha_2/\alpha_1)\,(x_1/x_2))$. The expenditure function is $e_{CD}[p_1,p_2,\mathcal{U}]=\lambda=e^{\mathcal{U}}(p_1/\alpha_1)^{\alpha_1}(p_2/\alpha_2)^{\alpha_2}$, with Hicksian demand functions $\left(x_1^h=\alpha_1 e_{CD}/p_1,x_2^h=\alpha_2 e_{CD}/p_2\right)$. Note that the ratio of consumption of the two goods is completely determined by the price ratio, since $x_2/x_1=(\alpha_2/\alpha_1)\,(p_1/p_2)$.

### B. Quasi-linearity

Of great interest in understanding the relationship between thermodynamics and economic utility theory is the class of *quasi-linear* preferences where the ordinal utility function can be transformed by a monotonic transformation to the form:

$$u^j\left[x^j\right]=x_0^j+\bar{u}^j\left[\bar{x}^j\right], \tag{6}$$

where $\bar{x}\equiv(x_1,\ldots,x_n)$. The gradient of the quasi-linear utility function $u'[x]=(1,\bar{u}'[\bar{x}])$ has the same dimensions as a price system referenced to $p_0$. In the quasi-linear case the linear commodity $x_0$ is often referred to as "money". We will refer to the other commodities as *non-linear commodities.*

The first-order conditions in the Lagrangian (3) become:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial x_0}&=p_0-\lambda=0\\ \frac{\partial\mathcal{L}}{\partial\bar{x}}&=\bar{p}-\lambda\bar{u}'\,[\bar{x}]=0\\ \frac{\partial\mathcal{L}}{\partial\lambda}&=\mathcal{U}-u\,[x]=0.\end{aligned} \tag{7}$$

Thus with quasi-linear ordinal utility, the inverse of the marginal utility of wealth, $\lambda$, is equal to $p_0$ regardless of the utility level $\mathcal{U}$. Thus the Hicksian demands for the non-linear commodities are determined uniquely (assuming smooth differentiability and strict concavity of the nonlinear component of the utility function) by the first-order conditions $\bar{p}/p_0=\bar{u}'[\bar{x}]$ and we have $\bar{x}^h[\bar{p}]$, independent of the utility level $\mathcal{U}$. The Hicksian demand for the linear commodity is just $x_0^h[p,\mathcal{U}]=\mathcal{U}-\bar{u}\left[\bar{x}^h[\bar{p}]\right]$. The expenditure function for quasi-linear preferences is:

$$e_{\mathrm{QL}}[p,\mathcal{U}]=p_0\left(\mathcal{U}-\bar{u}\left[\bar{x}^h[\bar{p}]\right]\right)+\bar{p}\cdot\bar{x}^h\left[\bar{p}\right]. \tag{8}$$

When utilities are quasi-linear, the individual agents' indifference surfaces are preserved under translations of the form $(1/p_0, \bar{0})$, so that not only the first-order conditions, but all constraints on future paths of trading (such as could result from later aggregation with new agents or change in total endowment) carry no memory of the path to the trade set. In this case $\mathcal{U}^j$ itself is not directly relevant to the constraints on the demands for nonlinear commodities, which depend only on $\bar{u}^j\left[\bar{x}^j\right]$.

### 1. Example 2: Quasi-Linear CD utility

The *Quasi-Linear CD* (QLCD) utility function for three-commodities is:

$$u_{\mathrm{QLCD}}[x_0, x_1, x_2] = x_0 + \alpha_1 \log[x_1] + \alpha_2 \log[x_2]\,. \quad (9)$$

The offer prices for the QLCD utility with respect to commodity 0 are $u'_{\mathrm{QLCD}}[x] = (1, \alpha_1/x_1, \alpha_2/x_2)$. The expenditure function is $e_{\mathrm{QLCD}}[p_0, p_1, p_2, \mathcal{U}] = p_0[\mathcal{U} - \alpha_1 \log[\alpha_1 p_0/p_1] - \alpha_2 \log[\alpha_2 p_0/p_2] + \alpha_1 + \alpha_2]$, with Hicksian demand functions $(x_0^h = \mathcal{U} - \alpha_1 \log\left[x_1^h\right] - \alpha_2 \log\left[x_2^h\right], x_1^h = \alpha_1 p_0/p_1, x_2^h = \alpha_2 p_0/p_2)$. Note that the Hicksian demands for commodities 1 and 2 are independent of the utility level. The Hicksian demand for the linear commodity may be negative, indicating that when confronted by some price systems the agent will supply an amount of the linear commodity to the market, that is, incur a debt in the linear (money) commodity.

## C. Economic exchange

The economic theory of exchange economies considers a system of $m < \infty$ agents, each with its own utility function, exchanging a fixed total bundle of commodities, $w = \{w_0, \ldots, w_n\}$. The state of such a system can be described by an *allocation* $\mathbf{x} = \left\{x^1, \ldots, x^m\right\}$, where $\Sigma_j x^j = w$. An allocation is interpreted as describing the current ownership or holdings of each agent.

In general at an arbitrary allocation, the offer prices (marginal rates of substitution) of agents will differ. Under these circumstances two or more agents given the opportunity to exchange goods can find (in general many) mutually advantageous exchanges, which conserve the total quantity of commodities, but reallocate them from one agent to another. This process can be visualized for a pair of agents as in Figure 1. The dimensions of the box represent the total available commodites, and any point in the box an allocation. An arbitrary allocation at which offer prices are unequal appears in the lower-lefthand corner of the figure, with the corresponding indifference curves for the two agents drawn through it. Any exchange that leads to an allocation in the lens between these indifference curves puts both agents on a higher indifference curve, at a higher utility level. (Such exchanges are called *Pareto-improvements*.)

The allocations that cannot be overturned by voluntary exchange of commodities among the agents are the

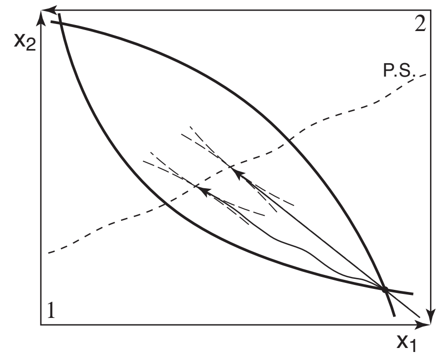


FIG. 1: The "Edgeworth-Bowley" box for a pure exchange economy with two types of agents and two goods. Axes are quantities of goods, $x_1$ and $x_2$. Origin for agents of type 1 is in the lower right, and for agent of type 2 in the upper left. The Pareto set (short-dash) is labeled P.S. Indifference curves are the heavy black lines, and the trade set is the subset of P.S. in the interior of the lens formed by these. Straight ray is the Walrasian path from initial endowments to the Walrasian allocation. Wavy ray is an arbitrary, utility-improving trading path the agents could actually take. Long-dashed curves are segments of indifference curves through the Pareto allocations attained by either path.

*Pareto allocations (PA)*, assignments of private ownership rights to the economy's resources at which the offer prices (marginal rates of substitution) of all the economic agents are proportional, and there are no mutually advantageous trading opportunities. The Pareto allocations are the dynamic equilibria of the voluntary exchange process. At a Pareto allocation the offer prices of the agents are proportional and their ratios constitute an *equilibrium price system* for the economy as a whole. The *Pareto set* is the the set of all Pareto allocations.[2] We will refer to the subset of Pareto allocations at which every agent prefers its commodity bundle to its endowment as the *trade set*.[3]

In general exchange economies the set of Pareto allocations is a continuum on which the equilibrium price system, the allocations of all commodities among agents, and agents' realized utilities all vary. This signals that the process of voluntary exchange by itself is *indeterminate* in allocating resources and forming equilibrium prices.

The problem was recognized by Hahn and Negishi [20],

---

[2] Economists often call Pareto allocations "Pareto-optima", to emphasize the idea that Pareto allocations are *efficient* in realizing all of the possible gains from voluntary exchange. In fact this property of efficiency holds only under other restrictive assumptions, in particular, the absence of external effects such as pollution which affect agent welfare outside the exchange of commodities.

[3] The trade set appears as the "contract set" in settings of bilateral negotiation between two agents.

and is illustrated in Fig. 1. The movement from an initial endowment to a Pareto allocation in a closed economy is not completely determined by the constraint of voluntary exchange. There are infinitely many distinct, continuous histories of transacting, terminating in a Pareto allocation in the trade set.

This problem of indeterminacy led Léon Walras to focus attention on those Pareto allocations at which the value of each agent's bundle of commodities at the equilibrium price system is equal to the value of the agent's endowment. We will refer to these allocations as *Walrasian allocations (WA)*.[4] Walras apparently hoped that real world voluntary exchange would lead to a Walrasian allocation, thus establishing the determinacy of the voluntary exchange process. We will argue below that the Walrasian allocations have no special properties not shared by other Pareto allocations[5]. Our consideration of parallel problems in physics will show that the Walras's general project of establishing determinacy of voluntary exchange outcomes by specifying the path of trade effectively conflated two very different types of problems.

#### 1. *Example 1 continued: Exchange in a Cobb-Douglas economy*

To see the implications of these general points, consider a two-commodity economy with an endowment $(w_1, w_2)$ consisting of a large number $m$ of agents divided into two equal groups, $A$ and $B$, each of which has a Cobb-Douglas utility function with different coefficients $\left(\alpha_1^A, \alpha_2^A\right)$ and $\left(\alpha_1^B, \alpha_2^B\right)$ respectively, each summing to 1. At a Pareto allocation all agents will have the same offer prices. As we can see from section II A 1, this implies that in equilibrium every agent of the same type will hold the commodities in the same proportions. The amount of the commodities held can vary among the agents of each type subject only to the constraint that the utility exceed the utility of the agent's endowment. Thus we can characterize a Pareto allocation in a Cobb-Douglas economy by the total amount of commodities held by the agents of each type, $\left\{\left(x_1^A, x_2^A\right), \left(x_1^B, x_2^B\right)\right\}$, where $\left((m/2)\, x_1^B, (m/2)\, x_2^B\right) = \left(w_1 - (m/2)\, x_1^A, w_2 - (m/2)\, x_2^A\right)$ to respect the endowment constraint. At a Pareto allocation the weighted sum of the utilities of each type is maximized:

$$\max_{\left(x_1^A, x_1^B\right)} \lambda \left(\alpha_1^A \log\left[(m/2)\, x_1^A\right] + \alpha_2^A \log\left[(m/2)\, x_2^A\right]\right) + (1-\lambda)\left(\alpha_1^B \log\left[w_1 - (m/2)\, x_1^A\right] + \alpha_2^B \log\left[w_2 - (m/2)\, x_2^A\right]\right) \quad (10)$$

The equilibrium relative price is

$$\frac{p_2}{p_1} = \frac{w_1}{w_2} \frac{\lambda \alpha_2^A + (1-\lambda)\,\alpha_2^B}{\lambda \alpha_1^A + (1-\lambda)\,\alpha_1^B}$$

As the weight $\lambda$ varies from 0 to 1, $\left(x_1^A, x_2^A\right)$ varies from $(0,0)$ to $(w_1, w_2)$, and the equilibrium price varies from $\left(\alpha_2^B/\alpha_1^B\right)(w_1/w_2)$ to $\left(\alpha_2^A/\alpha_1^A\right)(w_1/w_2)$. Whenever $\alpha_1^A \neq \alpha_1^B$, the equilibrium price system will be indeterminate over an interval.

#### 2. *Exchange in quasi-linear economies*

We define a *quasi-linear economy* as an exchange economy in which all the agents have quasi-linear ordinal utilities with the same linear commodity, which we will refer to as "money". We will continue to call the other commodities "nonlinear commodities". The Pareto allocations of a quasi-linear economy all share the same allocation of the nonlinear commodities, and, most significantly, the same equilibrium price system. They differ only in the distribution of the linear commodity (money) among the agents. In a quasi-linear economy equilibrium prices *are* determinate, not because there is a unique Pareto allocation, but because the Pareto set is degener-

[4] In the economic literature, Walrasian allocations are often called "competitive equilibria". It would be most natural in an economic context to reserve the term "equilibrium" for Pareto allocations, since the Pareto allocations are the rest points of the dynamic process of voluntary exchange, and this usage would conform to the use of the term "equilibrium" to describe the rest points of dynamical systems in physics. Because economists associate the term "equilibrium" so closely with Walrasian allocations, we hope to avoid confusion by using a terminology that distinguishes the whole set of Pareto allocations from the subset of Walrasian allocations.

[5] Walras even appreciated, in a way, that whether one equilibrium is privileged over others depends on the particular algorithm assumed for market clearing. The Walrasian allocation may be privileged if markets clear through mechanisms akin to Walras's "auctioneer", whose process of tatônnement was introduced to give economic plausibility to this particular allocation.

ate in this sense. Figure 2 illustrates the Pareto set for a quasi-linear economy.

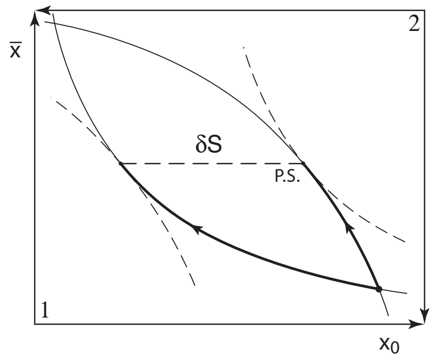


FIG. 2: All Pareto allocations for a quasi-linear economy have the same allocation of nonlinear goods, and the same equilibrium price system, and differ only in the allocation of the linear good (money) among the agents. The trade set illustrates the process by which an external speculator can extract wealth from a quasi-linear economy of two types of agents having no internal trade. Trades along indifference surfaces (heavy) are the worst agents will accept, and take place at changing prices. At the termination of trade, agents in the economy have identical (equilibrium) prices, and have lost wealth equal to the sum $\delta S$ of their economic entropies at the original allocation to the speculator.

For a given endowment, all Pareto allocations of a quasi-linear economy have the same allocations of the non-linear goods, $\left(\bar{x}^{1*}, \ldots, \bar{x}^{m*}\right)$ among the agents. This property follows from the facts that the equilibrium commodity bundle of the non-linear goods for each agent is a function only of the price vector for the nonlinear goods, $\bar{p}$, and that at equilibrium the sum of these demands is equal to the social endowment of nonlinear goods, $\bar{w}$. Thus the Pareto allocations for a quasi-linear economy differ only in the distribution of the linear good (money) among the agents, $x_0^1, \ldots, x_0^m$. The Pareto set can be visualized for a quasi-linear economy with two types of agents as in Figure 2. The Pareto set of a quasi-linear economy is a hyperplane (in the two-agent type case a line) in the subspace of the allocation space restricted to allocations of the linear good. (The trade set is the subset of this line bounded by the indifference surfaces that pass through the endowment.) The prices of the non-linear commodities $\bar{p}^* = p_0 \bar{u}^{j\prime}\left[\bar{x}^{j*}\right]$ are invariant on the Pareto set. The Hahn-Negishi indeterminacy of the path from the endowment to an equilibrium is reflected only in the final distribution of the linear good among the agents. This distribution is important and relevant economically (it determines the subjective welfare of the agents) but in the quasi-linear case has no impact on the equilibrium price system, or each agent's holding of the non-linear goods (or indeed on any possible future trades).

In a quasi-linear economy the difference between the non-linear component of an agent's utility at any allocation and on the Pareto set,

$$S^j_{\mathrm{QL}} = \bar{u}^j\left[\bar{x}^j\right] - \bar{u}^j\left[\bar{x}^{j*}\right], \tag{11}$$

is an unambiguous measure of the agent's potential gains from exchange. We define this quantity as the agent's *economic entropy.*[6]

Immediately

$$\delta S^j_{\mathrm{QL}} = \delta \bar{x}^j \cdot \left.\frac{\bar{p}}{p_0}\right|_{p=p[x^j]}. \tag{12}$$

The gradient of the economic entropy gives the agent's relative offer price $\bar{p}/p_0$ at the commodity bundle $x^j$.

The sum of the agents' economic entropies in a quasi-linear economy is maximized on the Pareto set. The economic entropy of the whole economy is strictly speaking well-defined only at Pareto allocations and is equal to this maximized value. At a Pareto allocation, by the well-known variational principle economists call the "envelope theorem", the derivatives of the maximized sum of agent entropies with respect to the quantities of commodities available to the whole economy are the equilibrium prices. Thus the economy obeys a relation equivalent to Eq. (12) obeyed by individual agents.

The sum of economic entropies is also well-defined at non-Pareto allocations, and is non-decreasing along any path from the initial endowment to the Pareto set on which the utilities of individual agents are non-decreasing (that is on paths of voluntarily acceptable exchange among the agents), because on those paths $x_0^j + \bar{u}^j\left[\bar{x}^j\right]$ is increasing for every $j$ and $\Sigma_j x_0^j = w_0$ is constant. (It is possible, however, for individual agents' economic entropies to decrease along such a path, since an agent may gain utility from increasing its holdings of the linear good.) The sum of agent entropies is a quantitative measure of how close the economy is to the Pareto set, since the difference between its value at any allocation and the maximum value it attains on the Pareto set is an unambiguous (distribution-free) measure of the economic surplus still realizable through voluntary transactions.

To measure the value of an allocation $\left\{x^j\right\}$ to the agents in a quasi-linear economy, we suppose for a moment that they cannot trade internally, and introduce the device of an external *speculator* who can mediate trade in $\bar{x}$ through $x_0$, but who holds no stock of the non-linear goods herself.

The speculator's goal is to extract the largest wealth from the economy possible by voluntary trade, which she accomplishes by exchanging $x_0$ for $\bar{x}$ with the agents along their indifference surfaces (their least-favorable accepted trades), as shown in Fig. 2. What the speculator buys of $\bar{x}^j$ from one agent she sells to the others at higher prices, maintaining zero inventory of $\bar{x}$.

---

[6] For the general derivation, see Sec. VI D below.

Because equilibrium prices for each agent in a quasi-linear economy do not depend on $x_0$, the speculator can move the agents to prices for $\bar{x}$ equivalent to their prices at any equilibrium, at the minimal acceptable $x_0^j$ for each agent $j$. She can then decouple from the economy with surplus $x_0$ leaving the economy with no further advantageous trades. The wealth obtained by the speculator is the sum of the agents' entropies at the initial allocation, because agent utilities do not increase during the extraction.

We now observe, however, that by construction the behavior of the sum of the agent economic entropies is the same when they trade with such an external speculator and when they trade with each other. In other words, if the agents can trade $\bar{x}$ internally to *any* Pareto allocation, the sum of the gains in their utilities equals the money-value of $x_0$ that they have thereby prevented an external speculator from extracting. Moreover, for any combination of speculative extraction and internal trade, the sum of $x_0$ extracted and welfare gained is constant. The sum of the economic entropies is thus an intrinsic money-metric welfare measure of the allocation in an economy, equal to its potential to deliver wealth in the form of the linear good to an external speculator or the equivalent utility value to the agents themselves through voluntary trade.

#### 3. *Example 2 continued: Quasi-linear exchange*

To see the structure of exchange in a quasi-linear economy, consider a three-commodity economy with an endowment $(w_0, w_1, w_2)$ consisting of a large number, $m$, of agents divided into two equal groups, $A$ and $B$, each of which has a QLCD utility function with $x_0$ as the linear commodity, but with different coefficients $\left(\alpha_1^A, \alpha_2^A\right)$ and $\left(\alpha_1^B, \alpha_2^B\right)$ respectively. At a Pareto allocation all agents will have the same offer prices, $p = (1, p_1, p_2)$. As we can see from section II B 1, this implies that in equilibrium every agent of type $A$ will hold the same bundle of nonlinear commodities, $x^{\bar{h}A}[p] = \left(p_0\alpha_1^A/p_1, p_0\alpha_2^A/p_2\right)$, and similarly for agents of type $B$. Since the total amount of the nonlinear commodities availabe is $(w_1, w_2)$, the equilibrium prices must satisfy:

$$\frac{(m/2)\, p_0^* \left(\alpha_1^A + \alpha_1^B\right)}{w_1} = p_1^*$$

$$\frac{(m/2)\, p_0^* \left(\alpha_2^A + \alpha_2^B\right)}{w_2} = p_2^*$$

These equilibrium prices also uniquely determine the agents' holdings of the nonlinear commodities at any Pareto allocation:

$$x_1^{A^*} = \frac{\alpha_1^A}{p_1^*} = \frac{\alpha_1^A w_1}{(m/2)\left(\alpha_1^A + \alpha_1^B\right)},$$

$$x_2^{A^*} = \frac{\alpha_2^A}{p_2^*} = \frac{\alpha_2^A w_2}{(m/2)\left(\alpha_2^A + \alpha_2^B\right)},$$

$$x_1^{B^*} = \frac{\alpha_1^B}{p_1^*} = \frac{\alpha_1^B w_1}{(m/2)\left(\alpha_1^A + \alpha_1^B\right)},$$

$$x_2^{B^*} = \frac{\alpha_2^B}{p_2^*} = \frac{\alpha_2^B w_2}{(m/2)\left(\alpha_2^A + \alpha_2^B\right)}.$$

The allocation of the linear commodity can vary on the Pareto set across agents of different types and among agents of the same type arbitrarily subject to the constraint:

$$\Sigma_{j=1}^m x_0^j = w_0.$$

The economic entropy of an agent at any bundle of nonlinear commodities is equal to the difference between the nonlinear component of utility on the Pareto set and the nonlinear part of the utility at the bundle, and for the QLCD system takes the form:

$$S_{\rm QL}^j\left[\left(x_1^j, x_2^j\right)\right] = \alpha_1^j \log\left[\frac{x_1^j}{x_1^{j*}}\right] + \alpha_2^j \log\left[\frac{x_2^j}{x_2^{j*}}\right].$$

## III. THERMODYNAMIC FUNDAMENTALS

### A. State relations and path independence in classical thermodynamics

Classical thermodynamics is based on the concept of a *state function*.

The macroscopic physical systems that are the domain of classical thermodynamics, such as confined fluids, are generally characterized instantaneously by a large number of configuration variables, and typically a small number of dynamically conserved quantities. The conserved quantities that scale together, as measures of the "size" of the system, are variables like energy $E$ or volume $V$, and are referred to as *extensive variables.*

The remainder of the configuration variables, whose values change rapidly as a result of the system's internal dynamics, are described statistically in terms of a distribution. General distributions have well-defined values of the Shannon/Boltzmann entropy $\mathcal{S}$. This description belongs to the domain of statistical mechanics, and presumes nothing about stationarity, reversibility, or equilibrium.

We will use the example of a "perfect" gas confined in a container of given volume in this exposition to illustrate the methods of classical thermodynamics. The configuration variables microscopically specifying the state in this case are the positions and momenta of the molecules that constitute the gas. The total energy of the gas is the sum of the kinetic energies of these molecules, and is higher the more rapidly on average the molecules are moving. The elementary analysis of the behavior of this simple system requires us to consider only the total energy $E$ and the volume $V$ as conserved quantities.

Statistical mechanics characterizes the equilibrium of a thermodynamic system (such as a perfect gas in a container) by a distribution over the microscopic configuration variables (in the case of the contained gas, the positions and momenta of the molecules) that maximizes informational entropy $\mathcal{S}$ among those distributions possible at given the conserved quantities $E$ and $V$. These conserved values are the *sole constraints* on distributions so defined [18], in which role they are interpreted as *extensive state variables.*[7] The maximizing value of informational entropy $S[E, V]$ is a *state function* of the other extensive state variables, and is itself extensive if the system in question interacts with other thermodynamic systems only through the boundary conditions defined by $(E,\ V)$, so that such interacting systems do not otherwise influence the distributions over the remaining internal degrees of freedom.

The hypersurface of distributions $\mathcal{S} = S[E, V]$ as $(E, V)$ is varied is called the *surface of state*. Systems such as contained perfect gases converge (or "relax" in physics jargon) to this statistical mechanical equilibrium very rapidly. As a result, macroscopic observation of thermodynamic systems normally enounters them at or very close to their equilibrium state, and therefore on the surface of state. Our interest in this paper is the interaction of macroscopic physical systems exchanging energy and volume, each of which remains at or close to equilibrium at all times. We take no position on whether there are economic parallels to the statistical mechanical processes that keep physical systems such as confined gases close to equillibrium.[8]

Once the state function is known for some system, its equilibria are completely characterized within classical thermodynamics, without further reference to statistical mechanics. Two properties from the statistical theory will remain of interest to us, however, which have classical consequences. First, the definition of $S[E, V]$ as a maximum subject to constraints implies the classical Second Law of Thermodynamics, that entropy increases as energy and volume constraints are loosened or systems are aggregated. Second, the existence of multiple paths through the space of general distributions $(E, V, \mathcal{S})$ allows many transformations to interpolate between *the same* initial and final equilibrium states, a fact recognized but not easily understood purely within the classical thermodynamic theory.

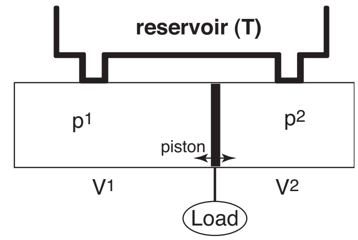


FIG. 3: Work extraction by a movable piston separating two chambers (e.g., of an ideal gas) of fixed overall volume $V^1 + V^2 \equiv V$. The piston is connected to a load, allowing it to maintain differential pressures $p^1$ and $p^2$ in the two chambers while trading volume between them, equivalent to differential prices under exchange of commodities. A reservoir at temperature $T$ permits reversible transformations while the other variable $V$ is conserved.

Equilibrium thermodynamic systems are characterized by duality between the extensive state variables and a parallel set of *intensive state variables*, whose values do not scale with overall system size. Duals are defined through the tangent plane to the surface of state at any point. Thus for example, temperature $T \equiv \partial E/\partial S|_V$ is the energetic dual to entropy, pressure $p \equiv -\partial E/\partial V|_S$ is the energetic dual to volume. For purposes of comparison to the economic theory, it is more useful to recognize $1/T \equiv \partial S/\partial E|_V$ as the entropic dual to energy, and $p/T \equiv \partial S/\partial V|_E$ as the entropic dual to volume.

### B. Example 3: Interaction of confined gases

When two thermodynamic systems with different temperature or pressure are brought into contact the exchange of energy or volume between them will continue until their temperatures and pressures are equalized. Figure 3 show a stylized experimental apparatus to illustrate this equilibration. Two perfect gases are confined in the two chambers of the apparatus, separated by a moveable piston. Initially the chambers are insulated from each other and the heat reservoir so that no energy can enter or escape, and the piston is immobilized. The gas in each chamber will relax to equilibrium at its own temperature and pressure. If the two subsystems are brought into thermal contact they can exchange energy in the form of heat, and if the piston is allowed to move freely, they can exchange volume. The problem of classical thermodynamics is to predict what configuration the interacting systems will reach.

We will discuss the similarity between this thermodynamic exchange and economic exchange in detail below.

[7] The state variable corresponding to energy $E$ is generally denoted $U$, which helps to emphasize its different role from the average energy of a general statistical distribution. We will not employ the notation shift here, because it has no counterpart for volume or other comparable constraints – notation has evolved haphazardly in physics, as in economics – and because it risks confusion with utility.

[8] It is possible to imagine economic models in which trading agents could be in internal disequilibrium (for example, where the agent is a household consisting of several individuals, and the household's collective willingness to exchange depends on its internal allocation of goods), but we do not pursue this line of thinking here.

Notice that each chamber of gas is independent of the other once the position of the piston and temperature of the reservoir (one state variable from each dual pair) are specified. We will therefore begin the comparison by considering the effect on a single chamber of changing its boundary conditions (say $V^1$), and suppress the superscript in the following subsection.

### C. Reversible and irreversible transformations

The pitfalls of pursuing purely formal analogies between utility theory and thermodynamics, without being careful about how the formalism is constructed, are illustrated by the following types of thought-experiments: Consider a perfect gas confined to a cylinder by a movable piston. Given its energy, $E$, and the volume determined by the position of the piston, $V$, it comes into equilibrium with an entropy $S[E, V]$ determined by the perfect gas law, and a well-defined temperature $T$ and pressure, $p$ given by the dual relations defined above.[9]

In the first experiment, the cylinder is maintained at a constant temperature by being brought into contact with a large thermal reservoir with which it can exchange energy in the form of heat. If the piston is slowly pulled outward, so that the system stays close to equilibrium, energy flows out of the system as work linked to the increase in volume, and flows into the system as heat from the reservoir, conserving the energy of the system.[10] At constant $E$ and higher $V$, the system has a higher entropy. This experiment is visualized in Fig. 4 as the motion along the constant-$E$ contour in the surface of state. This experiment looks familiar to an economist. The thermodynamic system appears to be trading volume for entropy with the rest of the world along an indifference surface, resulting in a fall in its marginal rate of substitution between entropy and volume, $p/T$.

For physicists a central property of the equation of state is that it can also be used to analyze a second experiment, in which the cylinder is not connected to a thermal reservoir, and therefore cannot exchange heat. In this experiment the piston is pulled out so rapidly that it loses contact with the gas, which consequently does no work on the piston, so that internal energy is conserved. The gas is temporarily released into a disequilibrium internal configuration with respect to the larger available volume, but relaxes eventually into equilibrium with its original

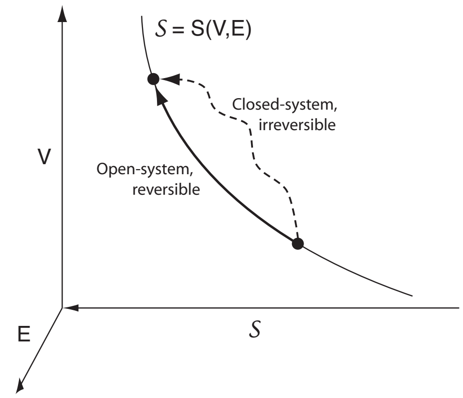


FIG. 4: Different transformations can interpolate between the same thermodynamic initial and final states. One transformation, indicated by the thick solid path along the surface of state, takes place at fixed temperature in contact with a thermal reservoir. A second transformation, indicated by the dashed line, interpolating between the same initial and final states takes place in thermal and mechanical isolation, but conserves the internal energy of the gas.

energy at the new volume. According to the equation of state it must relax to the same entropy as in the first experiment, and hence to the same state as described by the extensive and intensive variables. This visualization of the second experiment is troubling to the economist because it appears to have moved the system to a higher indifference surface (by providing it with more volume), yet the physicist locates the system on the same contour and indeed at the same point in the state space as after the first experiment. Thus it is clear that, despite their similar generation of duals, the economic indifference surface is not quite the same as a surface of state, and the statistical-mechanics coordinate system of Fig. 4 cannot correspond to an economic demand space.

The latter was to be expected, on the grounds that the non-equilibrium values of $\mathcal{S}$ in Fig. 4 have no interpretation in classical thermodynamics, and no role in assigning its duality relations, whereas every point in an economic demand space lies within some indifference surface, and through it is assigned some dual price system. The resolution of this paradox begins by replacing the coordinate $\mathcal{S}$ in a space of the same dimension, with another quantity that has a classical thermodynamic interpretation, but behaves differently depending on whether the system is undergoing a reversible or irreversible transformation. As all state variables of the system ($E$, $V$) are already drawn, the new coordinate must be a property of the environment in which the system makes its transformation.

We choose to use the reservoir, which may be coupled to the cylinder-piston system to varying degrees, but whose internal dynamics we assume remains near equilibrium no matter how the system transforms. Then we may embed the coordinates ($E$, $V$) in a space whose $(n+1)^{\mathrm{st}}$ coordinate, $\Sigma$, cumulates the loss of entropy from the

[9] Economists may be intrigued by the fact that the entropy state relation for the perfect gas law has the form: $S[E, V] = A \log[E] + B \log[V]$, that is, a Cobb-Douglas function.

[10] A property of the perfect gas law, which can be derived from footnote 9 and the relation $1/T \equiv \partial S/\partial E|_V$ of Sec. III A, is that the internal energy is a function only of its temperature. Thus the maintenance of a constant temperature through contact with the thermal reservoir implies the conservation of the internal energy of the gas in the experiment.

reservoir. One could think of this in economic terms as the system's cumulated net "entropy debt". From $\Sigma$ and $S\,[E,V]$ we could construct a "utility function" for the perfect gas, which has a quasi-linear form:

$$\mathcal{U} = \Sigma + S\,[E,V]\,. \tag{13}$$

The entropy loss from a reservoir in equilibrium at temperature $T$ equals $-dQ/T$, where $dQ$ is an increment of heat sent out through the reservoir boundary (in through the system boundary). Thus we set

$$\Sigma = -\int \frac{dQ}{T}, \tag{14}$$

and recognize that for the reversible transformation, $-\delta\Sigma$ is the entropy gain $\delta S$ by the system, formalizing the reversibility condition as $\delta\mathcal{U} = 0$ and justifying the consideration of $\mathcal{U}$ as a "utility". The system maintains its temperature and energy balance by incurring an entropy debt to the thermal reservoir.[11]

In the second experiment, the system is isolated from the reservoir, so it cannot increase or decrease its entropy debt, and $\Sigma$ remains constant. In this case the experimenter moves the thermodynamic system to a higher indifference surface by providing it with an additional endowment of volume, $V$. This raises the $S[E,V]$ component of the utility, without a compensating fall in $\Sigma$.

The source of entropy change in a classical thermal system does not matter (there is only one "kind" of entropy, arising ultimately from the statistical measure of uncertainty $\mathcal{S}$). Thus $\Sigma$ cannot be a state variable of the system itself. To create a utility diagram, for any pair $(E,V)$, we construct an "indifference surface" through $(E,V,\Sigma)$ from the state relation that results if the system reaches equilibrium at $(E,V)$ as a result of a transformation exchanging $\Sigma$, and is then probed with reversible transformations through that point. In the quasi-linear form (13) on the extended state space only the nonlinear component $S[E,V]$, like $\bar{u}[\bar{x}]$, determines the dual variables, $(1/T, p/T)$ (like $\bar{u}'[\bar{x}]$) independently of the utility and the amount of the linear component, $\Sigma$ (like $x_0$). A classical thermodynamic system exhibits no "memory" of its past interactions with the rest of the world, just as an agent with a quasi-linear utility exhibits no behavioral reaction in its demands for $\bar{x}$ to its holdings of the linear good, $x_0$.

[11] If a physical system out of equilibrium interacts with other systems out of equilibrium, temperature, and thus this entropy debt, may not be well-defined physically. But even if the subsystem itself is not in equilibrium, as long as it is interacting with other systems that are in equilibrium, they will have well-defined temperatures at which the heat flows between the systems can be cumulated to a well-defined entropy debt. Thus the parallel between classical thermodynamic systems and quasi-linear economies strictly speaking holds only for physical systems which interact with other physical systems that are in equilibrium.

### D. Thermodynamic potentials

One of the powerful insights of thermodynamics is that it is possible to generate a whole family of dual potentials to the entropy for a system. Each dual potential is a state function for the system subject to particular boundary conditions (or as economists say, "closures"). A thermodynamic system might, for example, be constrained by constant pressure rather than constant volume, and in this case it is natural to express the equation of state in terms of energy and pressure rather than energy and volume. Analogously, a small open economy trading one nonlinear commodity for money with a large world market faces a constraint on the price of that nonlinear commodity, not on its quantity. For a state function $S[E,V]$, the Legendre transform is the difference between the function and the product of one or more of the extensive variables with its corresponding intensive variable, for example $S[E,V] - \partial S/\partial E|_V E$. Differentiation shows that this transform has arguments $V$ and $\partial S/\partial E|_V = 1/T$, thus producing a new potential relating $V$ and $1/T$.

To show how Legendre transformation generates new potentials from the entropy in a thermodynamic system, we work through one particular transformation. The entropy function of extensive variables $S[E,V]$, as noted in Sec. III A, has the variations $\partial S/\partial E|_V = 1/T$, $\partial S/\partial V|_E = p/T$, so that the variation of its Legendre transform in both $E$ and $V$ is

$$\delta\left(\frac{1}{T}E + \frac{p}{T}V - S\right) = E\,\delta\left(\frac{1}{T}\right) + V\,\delta\left(\frac{p}{T}\right). \tag{15}$$

The Legendre Dual may also be written $(1/T)\,F$, where $F = E + pV - TS$ is the *Gibbs Free Energy*, to which we will return below. $F$ emerges in the context of transforming systems using $p$ as the control at constant $T$, in which

$$\left.\frac{\partial F}{\partial p}\right|_T = V. \tag{16}$$

The gradients of potentials at equilibrium define the duality relations of a thermodynamic system. These potentials are evaluable at points along (a restricted set of) transformation paths, and extremized[12] on the set of equilibria (the contract set). The potentials are defined for each subsystem, and the individual potentials add to

[12] The sign conventions in physics for Legendre transforms, like that in Eq. (15), cause *maximization* of entropy to correspond to *minimization* of the various free energies. The intuition behind this choice is that free energies are the thermodynamic generalization of mechanical energies, which give the appearance of being minimized at equilibria. It was this mechanical analogy, in an era before thermodynamics had been understood, that misled Walras into thinking his utility functions followed the model of energies.

a system-wide potential under aggregation and equilibration of subsystems. We will further consider, in the rest of the paper, the conditions under which economic potentials with similar properties can be defined. A summary of the correspondence for the quasi-linear case is provided in Table I.

## IV. PARALLELS BETWEEN THERMODYNAMICS AND ECONOMICS IN THE REPRESENTATION OF SYSTEMS

In physics it is conventional not to regard models of particular thermodynamic systems as fundamental, but to abstract the theory into a small collection of *laws* that restrict how models can be built so as to be consistent with a thermodynamic mode of description. From the preceding abbreviated reviews of exchange economies and simple $(E, V)$-type thermal systems, we are in a position to observe that the three most fundamental assumptions about system representation in economics and thermodynamics are in fact the same. The assumptions called "laws" in thermodynamics correspond to fundamental "axioms" of neoclassical economics.

### A. The three laws of thermodynamics and utility theory

> $0^{\text{th}}$ **law** (*encapsulation*):
>
> Thermodynamic systems in equilibrium have well-defined values of some set of *state variables*. These state variables are of two types: to each *extensive variable* (such as energy and volume), there corresponds a conjugate *intensive variable* (such as temperature and pressure). If two systems are brought into contact, they undergo no macroscopic changes if and only if the intensive state variables conjugate to all exchangeable quantities have the same values. In this case the extensive variables characterizing the joint system take on values equal to the sums of the extensive variables of the interacting subsystems, and the intensive variables characterizing the joint system are equal to the common values for the intensive variables of the interacting subsystems.
>
> Economic agents are characterized by their bundle of commodities. An economic agent has well-defined offer prices for all commodities, its *marginal rates of substitution* at its current bundle. If two agents are given the opportunity to trade, their holdings will remain unchanged if and only if their marginal rates of substitution for all exchangeable commodities are equal. In this case the economy consisting of the two agents has a total commodity bundle equal to the sum of the individual agents' bundles, and the marginal rates of substitution of the individual agents remain unchanged.

This law defines the category of "classical" thermodynamic systems, as those which exercise or receive constraints from other systems *only* through the interfaces defined by the state variables. Systems capable of carrying "memory" of past transformations that is not reflected in the state variables [15–17], but can affect future transformations, do not satisfy this criterion.

Thermodynamics explicitly considers the possibility that a subsystem of a larger system might not be in equilibrium. Systems not in equilibrium do not in general have well-defined values for the conjugate intensive variables, mandating the qualification in the $0^{\text{th}}$ law "thermodynamic systems *in equilibrium*". *Statistical mechanics* studies the microscopic configurations of thermodynamic systems, and provides an understanding of the process by which a thermodynamic system out of equilibrium converges (or "relaxes") to an equilibrium state with well-defined values of the conjugate intensive variables.

Economic utility theory does not generally consider the possibility of individual economic agents being out of equilibrium in this sense, so that in utility theory the offer prices (marginal rates of substitution) of an economic agent are always well-defined. Thus at the level of representation, economic agents can only be considered analogous to thermodynamic subsystems *in thermodynamic equilibrium*. On the other hand, thermodynamic subsystems in contact do not retain an equilibrium representation unless their intensive state variables take the same values in all the subsystems. There is no thermodynamic equivalent of an economic "trading path", along which agents are individually given equilibrium representations even though their intensive state variables (offer prices) differ.

*At* an equilibrium economics and thermodynamics concur in both the interpretation of dual structure and the conditions of equality on intensive state variables. They differ in the treatment of *transformations*, which of course arise in most of the interesting questions in both fields. By emphasizing reversible transformations, classical thermodynamics consistently requires equality of intensive state variables as a condition for equilibrium representation of subsystems. As a consequence energy or some other extensive quantity must be allowed to enter or leave the system, as we saw in the examples with heat flow. Utility theory has chosen to keep the equilibrium representation of agents under all conditions, but by emphasizing conservation of endowments for trades taking place away from equilibrium, has removed the condition that agents share a price system, or that their offer prices should coincide with the actual rates of exchange.

> $1^{\text{st}}$ **law** (*constraint*):

> Energy is conserved under arbitrary transformations of a closed thermodynamic system.
>
> Commodities are neither created nor destroyed by the process of exchange.

$2^{\text{nd}}$ **law** (*preference*):

> There is a partial ordering of configurations of state variables of a thermodynamic system (which can be indexed by its entropy at each configuration), and transformations that decrease the entropy of a closed system do not occur.
>
> There is a partial ordering of commodity bundles for an economic agent (which can be indexed by an ordinal utility function), and agents never voluntarily accept trades that reduce their utility.

The three laws introduce the dual structure we have seen in the representation of agents, define special roles for conservation of extensive quantities, and separate these two functions from rules specifying which way aggregates of subsystems or agents may spontaneously change. The most direct way to understand why utility *cannot* be a counterpart to energy is to recognize that energy, along with volume or other conserved quantities, defines the surface of possible configurations of aggregated systems. The function of utility is more similar to that of entropy: it identifies in which direction exchanges are permitted to occur.

We saw in sec. III C that utility is *not the same* as entropy, though, because entropy is a function of state that cannot remember everything that a "utility" like Eq. (13) for the same system could be constructed to remember. This difference has deep conseqences for both thermodynamics and economics. An aggregate of thermodynamic subsystems can undergo transformations that the systems in isolation could not, because global entropy increase is less constraining when systems can exchange heat, allowing individual-system entropy to decrease. Similarly, agents in an economy can achieve allocations that the agents in isolation cannot, because utility increase can correspond to increase or decrease of the economic entropy (11), when "money" can also be exchanged.

### B. State and conservation-based descriptions

The different uses of duality in economics and classical thermodynamics may be related back to the methodological differences in the two fields by distinguishing "state-based" from "conservation-based" diagrams, as indicated in Fig. 5. State-based diagrams (the (a) panels) attempt to identify unique equilibrium functions of boundary conditions. In classical thermodynamics, intermediate configurations off the surface of state have no assigned duality relations. Utility diagrams may be projected into a state-based form by a utility-offsetting map of all equilibria onto some reference equilibrium bundle. The duality relations associated with different (utility-distinguished) equilibria may or may not respect such a projection, either in the first-order conditions or in higher moments. For (quasi-linear) systems that do respect the projection, the duality relations may be considered properties of the *equivalence class* under the projection, rather than of the individual equilibrium allocations.

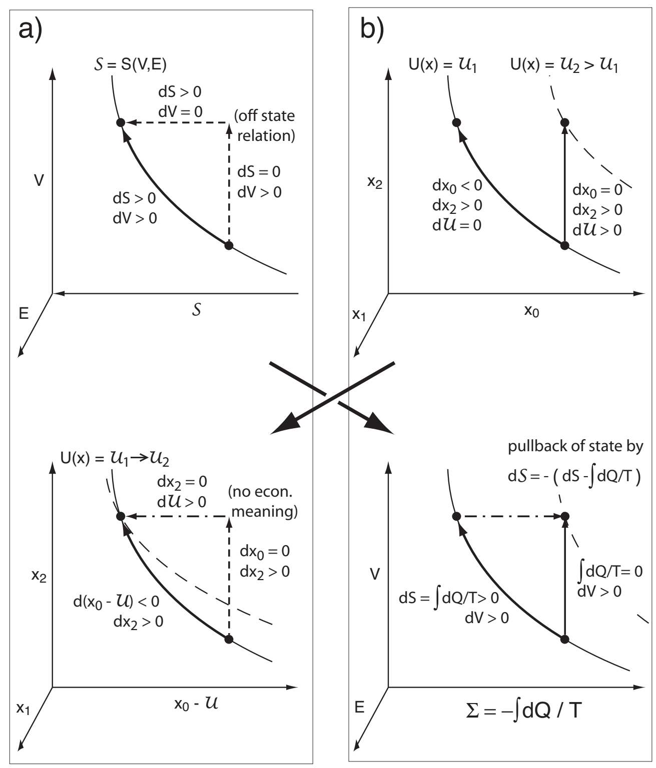


FIG. 5: State-based (a) and conservation-based (b) diagramming, the default representations of thermodynamics and neoclassical theory, respectively. One converts the utilitarian diagram to a state-based diagram by replacing one allocation coordinate (here $x_0$), with $x_0 - \mathcal{U}$ for suitably cardinalized utility $\mathcal{U}$. Conversely, one converts the thermal diagram to a conservation-based diagram by replacing non-equilibrium entropy $\mathcal{S}$ with entropy flow from a reservoir $\Sigma \equiv -\int dQ/T$, a coordinate capable of recording change in the state of the external world as well as the system. From the combination of these two correspondences one obtains the interpretations of utility and heat flow in the text. As in Fig. 4, dashed lines are off-state transformations, and here dot-dash denotes modifications, either by projection (economic) or pull-back (thermal).

In contrast, in keeping with emphasis on closed-economy conservation of endowments, economics emphasizes "conservation-based" diagrams (Fig. 5 (b) panels). For a general utility problem, there is no expectation of degeneracy in the duality relations at equilibria resulting from different trading paths. To extend the thermodynamic state to a function defined on all relaxation histories, a pullback of the duality relations by some "entropy debt" variable external to the classical system (such as $\Sigma$) is required to supplement the system description to include information about the environment.

### C. Dual potentials in economic systems

We can calculate the full suite of dual potentials to economic entropy in the quasi-linear economy. One such dual places the dependence on all nonlinear components of prices through the transformation:

$$\mathcal{F}^j_{\text{QL}} = \bar{p} \cdot \bar{x}^j - p_0 S^j_{\text{QL}}. \tag{17}$$

The variation of this object, which economists will recognize as the non-linear component of the expenditure function plus the economic surplus the agent can gain by moving to the trade set, is:

$$\delta \mathcal{F}^j_{\text{QL}} = (\delta \bar{p}) \cdot \bar{x}^j + \left( \bar{p} - p_0 \bar{u}^{j\prime}[\bar{x}^j] \right) \cdot \delta \bar{x}^j. \tag{18}$$

Since $\bar{p} = p_0 \bar{u}^{j\prime}[\bar{x}^j]$ in equilibrium we see that the derivatives of $\mathcal{F}^j_{\text{QL}}$ with respect to the prices of non-linear goods are indeed the Hicksian demands for the non-linear goods.

Various other mixed duals have natural economic interpretations in terms of a quasi-linear economy trading some subset of goods with a world market, thereby fixing the relative prices of those goods.

### D. The missing degree of freedom in the duality relation

If we define the intensive state variables in thermodynamics as the gradients of the entropy with respect to the extensive variables, as in the last line of Sec. III A, all of the state variable pairs represent measurable quantities. This property of thermodynamics differs from the most general view of utility theory, which uses both measurable and unmeasurable quantities. Marginal rates of substitution, which correspond to relative prices as long as agents trade along their indifference surfaces, are measurable properties of contracts. The *utility* and the *numéraire* for prices are not directly measurable in this way, and the latter is assumed not even to be meaningful in neoclassical theory. The equivalence relation that maps classes of indifference surfaces in quasi-linear economies onto thermodynamic equations of states reconciles this difference by keeping only the combination $\mathcal{U}^j - x^j_0 = \bar{u}^j \left[ \bar{x}^j \right]$, the economic entropy function reconstructable from the allocations $\bar{x}^j$ and their associated *relative* offer prices $\bar{p}^j / p_0$.

If we wish to describe the duality between $\bar{x}^j$ and $\bar{p}^j / p_0$ directly, there is no reason to introduce $p_0$ as a variable component of prices. We may use $x_0$ as numéraire and consider the natural dual $(1/p_0)\, \mathcal{F}^j_{\text{QL}}$ to economic entropy $S^j_{\text{QL}}$.[13] Using Eq. (18) and the cancellation of the coefficient of $\delta \bar{x}^j$, the resulting gradient is

$$\delta \left( \frac{\mathcal{F}^j_{\text{QL}}}{p_0} \right) = \delta \left( \frac{\bar{p}}{p_0} \right) \cdot \bar{x}^j. \tag{19}$$

We will return to the problem of constructing a dual theory, involving only the measurable marginal rates of substitution, for more general economies in Sec. VI C.

### E. Quasi-linear economies and classical thermodynamic systems

These results demonstrate the exact formal equivalence of quasi-linear economies to classical thermodynamic systems[14], and identify the economic forms of entropy and other potentials in the quasi-linear case. Perhaps the best way to describe this equivalence is our observation above that classical thermodynamic systems, with state spaces extended to include entropy debts to reservoirs, are quasi-linear economies in which the linear good is the cumulated entropy debt of the system to the rest of the world.

Later in this paper we investigate in detail how far it is possible to generalize the analytical methods of classical thermodynamics to systems that do not have a linear good. It is important to realize that this issue is not peculiar to economic systems. Physical systems that do not exhibit path-independence, and therefore retain a memory of the path through which they reached equilibrium (for example, in their micro-state) pose the same analytical problems as more general economies. In such systems the equilibrium boundary forces can depend on the path by which the system relaxed, and there is no well-defined entropy, nor equation of state linking aggregate system contents such as volume and energy rigorously to observable boundary forces holding the system in its pseudo-equilibrium.

---

[13] We introduced $\mathcal{F}^j_{\text{QL}}$ in Equations (17,18) because it corresponds strictly to the familiar physical Gibbs free energy $F$ appearing in Eq. (16). While in physics energy connects thermodynamics to many other descriptions such as mechanics, and it is experimentally convenient to compute intensive state variables as gradients of the internal energy at equilibrium, from the informational viewpoint in which *entropy* is the privileged extensive state variable, it is more natural to consider the elementary dual $(1/T)\, F$ corresponding to $(1/p_0)\, \mathcal{F}^j_{\text{QL}}$.

[14] More precisely, the *equivalence class* of utility surfaces modulo $x_0$ holdings, which determines prices and the unambiguous allocations $\bar{x}$ at equilibria, is equivalent to the description of a classical thermodynamic system.

## V. ECONOMIC AND THERMODYNAMIC METHOD

Both the existence of an exact correspondence of quasi-linear economies to classical thermodynamic systems, and the *non-existence* of such a correspondence for more general economies (demonstrated in later sections by the failure of the quasi-linear correspondence to generalize in all respects) carry important implications for economic theory and method. In physics, the failure of classical thermodynamics to make correct predictions for systems that do not obey the laws of Sec. IV indicates the need for more explicit theories of dynamics. The same is *a fortiori* true for non-quasi-linear economies, a fact that has not been acknowledged in general equilibrium theory.

### A. Economy-wide prices exist only near equilibrium

The essential insight from the $0^{\text{th}}$ law in Sec. IV is that thermodynamic systems only *have* well-defined intensive state variables that can be predicted from their extensive state variables when they are at (or very near) equilibrium. This restriction follows from the understanding, arising ultimately from statistical mechanics [18], that intensive state variables reflect *constraints* placed on the system by its boundary conditions (which may include interactions with other thermodynamic systems). Only when a thermodynamic system has relaxed to the point that its state variables represent its *only* constraints, can the intensive and extensive values be predicted from each other using the equation of state. Thermodynamic systems like the two chambers of gas in Fig. 3 can have well-defined intensive state variables that differ, only if they are individually in equilibrium with different boundary conditions provided by a separator like the piston.

General equilibrium theory seeks to predict economy-wide prices from endowments and the structure of agent preferences, and like intensive state variables, these prices are well-defined for economic equilibria (Pareto allocations). However, general equilibrium theory continues to represent agents has having well-defined prices (their offer prices) at non-Pareto allocations, even though those prices differ among agents, and even when they are not taken to reflect the presence of explicit constraints on the individual agents' holdings. This weaker requirement for the existence of "prices", than the requirement for intensive thermodynamic state variables, has created the widespread tendency of economists to conceptualize markets out of equilibrium as having a well-defined price (albeit not the equilibrium price). At disequilibria "price" thus comes to refer to logically different and sometimes inconsistent concepts.

One root of these inconsistencies is a persistent confusion in economic thinking between two different logical experiments. In the first experiment we start with an economy close to equilibrium, so that equilibrium prices are well-defined, and a single small economic agent is perturbed away from its equilibrium configuration, say, by a shock to the agent's commodity bundle. This agent finds itself with offer prices different from the economy's equilibrium prices. In this situation it makes sense to suppose that the agent will trade back to equilibrium at the economy's equilibrium prices, moving on a budget set that passes through the agent's perturbed endowment point. The agent's final position in this experiment is well-determined by the equilibrium prices, the agent's utility function, and the perturbed endowment point. This analysis makes sense precisely because the economy as a whole is close to equilibrium and large compared to the perturbed agent, so that it can absorb the agent's adjustment process with a negligible disturbance.

In the second experiment we imagine a shock that changes the offer prices of all of the economic agents in different ways. In this experiment the whole economy has been moved far from equilibrium, and as a result well-defined economy-wide prices do not exist. Voluntary exchange among the agents can restore equilibrium by bringing agents' offer prices into equality or proportionality, but this process is not determinate, and cannot be represented by the movement of each agent along a well-defined budget set. The logical confusion at the heart of the Walrasian theory of equilibrium arises from the attempt to conflate these two experiments.

### B. Reversible and irreversible transformations

Fig. 4 shows that the understanding of prices as constraints, such as those that arise at equilibria, remains compatible with finite transformations as long as they are conducted *reversibly.* Reversible transformations in thermodynamics are those that move each subsystem within its surface of state. In economics they are the transformations that move each agent within a single indifference surface.

Thermodynamics categorically distinguishes reversible from irreversible transformations. It generates deterministic predictions by measuring system structure through reversible transformations. Outcomes of irreversible transformations can then be predicted only if they suffer no path dependence and therefore coincide with the outcomes of other transformations that could be performed reversibly. Many physical systems, such as spin glasses [15], granular matter [16], and other materials with friction [17] can posess large sets of equilibria consistent with a single macroscopic boundary condition, but reached by different paths. *For these there is no macroscopic identification of preferred outcomes*, and a specific theory of dynamics is needed. The indeterminacy of the Hahn-Negishi trading process in Fig. 1 results in the same way from the irreversibility of trade to a Pareto allocation in a closed economy. Only when the inevitable path dependence of the equilibrium is limited to a degenerate coordinate under an equivalence relation (such as

$x_0$-equivalence in the quasi-linear economy) are the remaining properties of equilibrium determinate (*i.e.*, they could be obtained through an alternative reversible transformation). If there is no such degeneracy, prices and possibly higher-order derivatives of the indifference surface can vary among Pareto allocations, and some more detailed description of the dynamics of trade is required to predict the rates of exchange along actual trading histories and at the resulting equilibrium.

We have operationalized reversible transformations in economics in Sec. II C with the device of the external speculator, who can make small transactions with the agents of the economy, gradually inducing agents to exchange one commodity for another by offering terms of trade just a little better than their offer prices. The speculator can in this way vary the total holdings of the economy without allowing agents to move to higher utility levels[15]. The manifold revealed by reversible transformations summarizes the structure of the economy in the same sense that it reveals the equation of state of a thermodynamic system, with the caveat that in path-dependent economies indifference surfaces at different utility levels may not be degenerate in any of the allocation components.

The basic scenario of Walrasian economics envisions a group of economic agents each of whom initially holds an endowment bundle of commodities at which their offer prices differ, trading through markets to an equilibrium at which their offer prices coincide. Trade from a non-Pareto endowment to a Pareto allocation is inherently irreversible, since it is impossible to induce economic agents to voluntarily accept transactions that move them to a lower indifference surface. While the existence theorems for Walrasian allocations claim to treat only final allocations and not transformations, the definition of "wealth preservation" that assigns a particular Pareto allocation (or a discrete set of Pareto allocations) to any endowment is necessarily that realized by one particular trading path, the path described by Walrasian auction.

Walrasian theory thus attempts to overcome Hahn-Negishi indeterminacy by combining a theory of prices and equilibria with particular assumptions about the dynamics of irreversible trading. The goal of economic general equilibrium theory is to derive a particular Pareto allocation directly from the endowments and preferences of individual agents. The correspondence we have established between economics and thermodynamics shows that this goal is as unattainable as the goal of predicting the exact path of irreversible transformations in thermodynamic systems. While the Pareto set itself is derivable purely from the preferences of economic agents and the aggregate economic endowment, the path from an arbitrary non-Pareto endowment allocation to the Pareto set is not determined by purely by preferences and endowments. Only at Pareto allocations do preferences directly imply a law of one price, and there is no principle implying the same constraint on arbitrary, utility-improving trades.

[15] This scenario may not be as abstract as it might at first seem. A central bank making very small changes in the interest rate on bank reserves by absorbing the excess supply or demand of reserves might be viewed as approximating such an external speculator.

### C. Interpreting economies close to the Pareto set

Adopting the analytical assumption that an economy is at or close to the Pareto set requires some rethinking of the interpretive substructure of economic theory and econometrics. Much economics implicitly or explicitly adopts the interpretation of Hicks' *Value and Capital* [21], in which economic time is periodized into "weeks". On Sunday night of each Hicksian week all the agents receive their endowment of commodities, and are thus on Monday morning far from the Pareto set. On Monday a market occurs, which reallocates the initial endowments (through what we realize now is an irreversible transformation) to final Pareto allocation commodity bundles, and the agents spend the rest of the week actually consuming those bundles. Within the framework of this parable, the actual measured transaction flows of the economy correspond to the irreversible transformations associated with the achievement of a Pareto allocation (or, with the auxiliary hypothesis of the Walrasian auction, a Walrasian allocation).

If we want to adopt the point of view of reversible transformations, it makes more sense to interpret the commodity bundles of agents as *stocks*, such as stocks of consumer durables (the food in the refrigerator, for example). The availability of well-organized markets permits agents to keep close to their desired stocks at equilibrium prices at all times. Since agents are human beings who get hungry, wear out clothes, and in general deplete stocks, it is necessary for them to make transactions more or less continuously to keep close to their desired stocks (selling their labor-power, paying their rent, buying food, and so forth). These transactions, which generate national income, are not in this way of thinking the result of irreversible movements from far-from-Pareto endowments to a Pareto allocation, but the result of agents' constant effort to maintain their desired stocks given equilibrium prices. Something like Hicks' Sunday night, in which the economy and its agents are suddenly moved to a point far from the Pareto set, occurs only rarely as the result of external shocks to the system.

If we regard actual data on economic transactions as arising in this way, conventional specifications of demand functions in which flow transactions are a function of market prices and incomes are inappropriate. The prices at which transactions in a close-to-Pareto allocation economy take place are in fact equilibrium prices, which we can thus observe directly. The quantities transacted, however, depend on the dynamics of consumption and depreciation of stocks, which require specific modeling.

The assumption that agents generally remain close to desired stocks, and that the economy can as a result be analyzed with the concept of reversible transformations, is a strong abstraction. For example, an agent who loses her job typically feels that she has been forcibly (irreversibly) moved to a lower utility level. Real economies experience shocks (wars, revolutions, depressions, and technological innovations, for example) that intuitively seem to be best understood as irreversible transformations. The gradual processes of economic growth and development move agents to higher indifference surfaces, but on a time scale much longer than that of the establishment of market prices. We would like to emphasize the notion that the method of reversible transformations is best adapted to analyzing ongoing economies operating more or less normally.

## VI. THERMODYNAMIC POTENTIALS AND PATH DEPENDENCE IN GENERAL ECONOMIES

In sections II and III we demonstrated by construction the equivalence of quasi-linear economies, in which indifference surfaces are redundant under translation, to classical thermodynamic systems, and observed that the thermodynamic surface of state should be understood as the *equivalence class* of quasi-linear indifference surfaces modulo the quantity of the linear good $x_0$. Non-classical thermal systems [15–17] correspond to more general economies than the quasi-linear class in that they appear externally to be governed by the same set of extensive state variables, but under transformations are found to have significant path-dependence, exhibiting endpoint memory, hysteresis, and other phenomena that macroscopically would be associated with a *multi-valued* state relation on $(E, V)$. Similarly, economies that are not quasi-linear in any commodity have at least one degree of freedom per agent that distinguishes among equilibria, and on which prices or more complex constraints on subsequent trade can depend.

### A. The strong aggregatability of quasi-linear economies

Quasi-linear economies are equivalent to thermodynamic systems because they are aggregatable in a very strong sense, as economists have appreciated for decades. Any agent has the equivalent of an entropy, her $\bar{u}^j\left[\bar{x}^j\right]$, which is independent of the economy in which she is embedded and in particular independent of the preferences and endowments of other agents. These economic entropy analogues aggregate in the same way as thermodynamic entropies, because they are non-arbitrary money measures of (part of) the utility. Their sum is related to a so-called "social welfare" function (to which we return in Sec. VI D), which for any economy takes a unique maximum value on all of the economy's Pareto allocations. The equivalence class of the Pareto allocations corresponds to the uniquely defined thermodynamic equilibrium.

Thus the problem of "micro-foundations" for a quasi-linear economy is completely and convincingly resolved. The economic entropy and related potentials derived from it by Legendre transformation give a transparent understanding of the dynamics of trade in such an economy and the relation between the behavior of individual agents and aggregate economic objects such as equilibrium prices. On the other hand, this aggregation is not simply a process of "adding up" the individual agents in the economy. They do interact to form equilibrium prices, and there is an emergent property of the economy which cannot be reduced to or deduced from the properties of the individual agents: the distribution of wealth at the Pareto allocation finally reached.

### B. Transformations in general economies

With these clarifications in hand, we examine how a combination of structure determination based on reversible economic transformations, and the projection of economic equilibria onto appropriate utility-equivalence classes, can be used to describe the common features of general voluntary trading paths from non-Pareto endowments to a Pareto allocation. Along the way we will identify the maximal generalization of the Negishi construction of social welfare functions and its relation to Gorman aggregatability, the extent to which the Walrasian idea of a potential whose gradients are prices can be realized, and the proper relation of that potential to thermodynamic concepts including energy and entropy. We follow the logical sequence from simplest to most complex systems, starting with the determination of structure and redundancy for a single agent, relating that structure to the movement of a collection of agents to a Pareto allocation in a closed economy, and then extending the treatment to economies in contact with large external markets that serve as reservoirs for commodities.

We will find that the useful features of the thermodynamic correspondence are not lost all at once as one relaxes the assumption of quasi-linearity. Rather, a natural social welfare function, and path-independent prices and Walrasian potentials on the Pareto set, are retained for the class of Gorman-aggregatable economies, though higher derivatives of the indifference surfaces at equilibria generally become path dependent. If we abandon Gorman aggregatability, equilibrium prices can also become path dependent, and the concept of Walrasian potentials and social welfare functions are constructable only along one-parameter contours within the Pareto set.

### C. Price boundary conditions and measurement of system structure

The thermodynamic equivalent to the economic concept of "revealing preference" is the determination of the state relation by manipulating a system to achieve reversible transformations. The simplest economic system in which to measure such structure is an individual agent facing well-defined transaction prices who is allowed to adjust her stocks of goods to their desired level at those prices.

Two problems, which we were able to avoid for quasi-linear economies, introduce path dependence in general economies and cause revealed preferences to be less predictive than thermodynamic equations of state. The numéraire for prices is arbitrary, so that the number of economically meaningful price components is one fewer than the number of components of the allocation bundle $x^j$. Duality alone therefore cannot determine $x^j$ unambiguously. The extra coordinate that specifies $x^j$ in neoclassical theory is the level of utility (or equivalently the marginal utility of wealth), but unlike a thermodynamic state variable, utility is *not* assumed to be measurable in general. We therefore ask what is the best we can do to remove the unmeasurable numéraire and utility level from a theory of dual structure for general economies.

The variation (2) of the expenditure function (1) gives $n$ out of $n+1$ components of the Hicksian demand $x^h_i(p,\mathcal{U})$ corresponding to any $n$ relative price components of $p$, which we are free to choose. For quasi-linear economies we escaped the need to identify the last component from a dual relation because $\mathcal{U}$ and a particular component $x^h_0$ were *degenerate* quantities under an equivalence relation on indifference surfaces. As a result it was also possible to use transformations within indifference surfaces to infer properties of utility-increasing exchanges. The degeneracy in $x_0$ identified the price component $p_0$ as the one it is unnecessary to vary, so that $x_0$ effectively became numéraire.

As suggested by the lower panel of Fig. 5(a), no equivalence relation will map the full set of indifference surfaces to a well-defined state relation with respect to all allocations, for a general utility function. We can, however, regard a one-parameter family of allocations as equivalent for some purposes. Suppose we do this for a contour of Pareto allocations of some exchange economy, so that we may regard the difference between equilibria along that contour as changes in a degenerate "heat flow" coordinate, analogously to the degeneracy of the Pareto set for quasi-linear economies. It does not matter which component $p_i$ we use as normalization, as long as utility is increasing in the conjugate commodity $x_i$ (so that $x_i$ can be used as an index for the utility level).

Now separate the normalization from relative price variations by writing $\delta p = (\delta p_i)\, p/p_i + p_i \delta\left(p/p_i\right)$. From Eq. (2) it then follows that

$$\delta\left(\frac{e}{p_i}\right) = \delta\left(\frac{p}{p_i}\right)\cdot x + \frac{p}{p_i}\cdot\left.\frac{\partial x}{\partial\mathcal{U}}\right|_p \delta\mathcal{U}. \tag{20}$$

$x_i$ is effectively numéraire in the relative expenditure function $(e/p_i)$ (though we need not assume that $p_i$ is held constant as a normalization condition, or even relative to other agents at distributions far from the Pareto set).

As in the quasilinear case $e$ can vary with $\mathcal{U}$, but this variation can be canceled by working instead with an expenditure function $\left(e\left[p,\mathcal{U}\right]/p_i - u_{c,i}\left[\mathcal{U}\right]\right)$ offset by a suitable cardinalization $u_{c,i}\left[\mathcal{U}\right]$ of ordinal $\mathcal{U}$, as was done to form the economic Gibbs potential in Sec. IV D. Given any contour $c$ that intersects each indifference surface only once transversally, a contour-dependent cardinal utility of money-metric form in numéraire $x_i$ is defined by

$$u_{c,i}\left[\mathcal{U}\right] \equiv \int^{u[x_c]=\mathcal{U}} dx_c\cdot\left(\frac{p}{p_i}\right). \tag{21}$$

Here $dx_c$ is the differential commodity vector along contour $c$, $p$ is the agent's offer price vector defined by duality through the indifference surface at the bundle $x_c$, and the integration contour terminates on the value of $x_c$ in the indifference surface $u[x]=\mathcal{U}$. Intuitively we measure the utility of the agent by the distance along the contour from a reference commodity bundle.[16] At arguments $(p,\mathcal{U})$ dual to $x$ on this contour it follows that $(p/p_i)\cdot\left.\partial x/\partial\mathcal{U}\right|_c \equiv du_{c,i}/d\mathcal{U}$. For smooth indifference surfaces, $p[x_c]\cdot\left(\left.\partial x/\partial\mathcal{U}\right|_c - \left.\partial x/\partial\mathcal{U}\right|_p\right) = 0$ (by transversality), and $\mathcal{U}$-dependence is removed from the variation of $\left(e\left[p,\mathcal{U}\right]/p_i - u_{c,i}\left[\mathcal{U}\right]\right)$ along $c$. It then follows from Eq. (20) that

$$\delta\left(\frac{e}{p_i} - u_{c,i}\right) = \delta\left(\frac{p}{p_i}\right)\cdot x, \tag{22}$$

the generalization of Eq. (19). We thus obtain something that looks like a utility-independent state relation by removing the dependence on $x_i$, but in general can accomplish this only for points in the arbitrarily chosen contour $c$.

Eq. (22) has the form of a Legendre transformation of an economic entropy, though in general it depends on a contour $c$ which in turn is chosen for each agent within the Pareto set of the economy in question. With these restrictions, and up to a reference scale for prices that

[16] In the most general case $p_i$ may be defined with respect to a basis that depends on position in the contour $c$, but the contour utility construction is only particularly useful in cases where $p_i$ is fixed either by contour or by boundary condition, so we do not pursue the more general case here.

will be determined from context, the expenditure function offset by utility is the economic equivalent to a Gibbs potential for an individual agent. Before considering variations in general open systems, however, we relate the characterization of system structure gained from interaction with a single agent, to the process through which a collection of such agents trade from a non-Pareto endowment to a Pareto allocation, in a conventional closed economy. These will in part determine how the entropy can be naturally defined from its dual.

### D. Trading to the Pareto set in a closed economy

In a closed thermal system, equilibrium is defined by the condition that the sum of entropies of the component subsystems is maximized, on the space of configurations preserving the initial totals of the extensive variables (energy, volume, etc). The maximizing value must be unique, if the equilibrium is to be path-independent. The equivalent to entropy maximization in economics is social welfare maximization, and the existence of a well-defined entropy requires the same conditions as existence of a well-defined welfare function, though the two are in general not identical.

If $(e/p_i - u_{c,i})$ is the equivalent of a Gibbs potential, the economic entropy is its Legendre dual, by the reverse of the construction used in Sec. IV C. To make sense, such an entropy must be invariant on the trade set, independent of the path of trading or its terminal Pareto allocation. Clearly any single component $x_i$ satisfies $\sum_j x_i^j = w_i$ under arbitrary exchanges within a closed economy. Moreover, if the contour $c^j$ defining utility $u_{c,i}^j$ for each agent $j$ is the projection of a single contour in the Pareto set onto $j$'s allocation variables $x^j$, the sum of utilities $\sum_j u_{c,i}^j$ of Eq. (21) is invariant under change of Pareto allocation *remaining within that contour*. (This sum of contour utilities defines an intrinsic money-measure of the total contour $c$ in numéraire $x_i$.)

If we write $p = (p_i, \bar{p})$ with respect to the (as yet unspecified) coordinates of Sec. VI C, the relative-price degrees of freedom $\bar{p}/p_i$ define a basis for Legendre transform of $(e/p_i - u_{c,i})$. Specifically, the Legendre transform of $e$ is $e - \bar{p}\cdot\bar{x} = p_i x_i$ in this basis, and a candidate for an $x_i$-metric entropy for agent $j$ is defined by

$$\begin{aligned} -S_{c,i}^j &= \left. \frac{e^j\left[p,\mathcal{U}^j\right] - \bar{p}\cdot\bar{x}^j}{p_i} - u_{c,i}^j\left[\mathcal{U}^j\right] \right|_{p=p[x^j],\mathcal{U}^j=u^j[x^j]} \\ &= x_i^j - u_{c,i}^j\left[\mathcal{U}^j\left[x^j\right]\right]. \end{aligned} \tag{23}$$

From Eq. (22) it follows immediately that about any point in $c$,

$$\delta S_{c,i}^j = \delta\bar{x}^j \cdot \left.\frac{\bar{p}}{p_i}\right|_{p=p[x^j]}. \tag{24}$$

Since all agents $j$ have common dual prices at any Pareto allocation, we also immediately obtain the desired property

$$\sum_j \delta S_{c,i}^j = 0, \tag{25}$$

for movements among Pareto allocations within $c$.

For allocations not in $c$, the sum of individual agent entropies preserves the property of monotone increase with a (local) maximum on $c$, though its derivatives with $\bar{x}$ are no longer in general properly normalized prices as in Eq. (24). (Since $\sum_j x_i^j = w_i$ for all closed-economy trades, the only change in $\sum_j S_{c,i}^j$ comes from utility changes, which are individually nondecreasing.)

So far, no restriction has been made on the metric $x_i$ for $S_{c,i}^j$. For trade to equilibrium in the most general closed exchange economy, there may be no preferred choice, and the $S_{c,i}^j$ add no particular insight to what is already specified by the trade set. For such cases the $S_{c,i}^j$ are mostly interesting as the widest possible generalization of the Negishi construction of social welfare functions [19]. We may make them degenerate along a one-dimensional set of equilibria for each agent, because a monotone rescaling of ordinal utility admits an arbitrary function of one real variable. There are, however, intermediate degrees of aggregatability short of quasi-linearity, for which appropriately chosen $S_{c,i}^j$ provide natural money measures of the value of trade.

### E. Gorman aggregatability

If we choose commodity coordinates so that $\bar{p}^* \equiv \bar{0}$ at an equilibrium price system $p^*$, the Legendre conjugate demand variations $\delta\bar{x}^j$ for each agent $j$ are locally unconstrained by the preferences of other agents, and from Eq. (24) it follows that each agent maximizes $S_{c,i}^j$ over her own commodity holdings $\bar{x}^j$ independently (see Fig. 6 below). In the quasi-linear case if we take $p_i = p_0$ of the linear good, the $x_0$ independence implied by Eq. (24) on the Pareto set extends to the whole demand space, and with it independence of the utility level $\mathcal{U}^j$.

Generalizing from the quasi-linear case, Terence Gorman [22] has shown that the demand functions of a collection of different economic agents will coincide with the demand function of a single economic agent (so that the economy is *aggregatable in the Gorman sense*) if and only if the indirect utility[17] for each agent can be put in the form [2]

$$v^j\left[p, m^j\right] = \mathbf{b}[p]\, m^j + \mathbf{a}^j[p]. \tag{26}$$

[17] The indirect utility is defined as a function of prices and a wealth constraint $v(p,m)$. The value of the wealth coordinate $m$ equals the expenditure function $e[p,\mathcal{U}]$ at the maximal $\mathcal{U}$ affordable at that $(p,m)$.

The function $\mathbf{b}[p]$ is *common* to all the agents in an aggregatable economy. The function $\mathbf{a}^j[p]$ is idiosyncratic to each agent. The equivalent expression for the expenditure function is

$$e^j_{\text{Agg}}\left[p,\mathcal{U}^j\right] = \frac{1}{\mathbf{b}[p]}\left(\mathcal{U}^j - \mathbf{a}^j[p]\right). \tag{27}$$

quasi-linear utility satisfies the Gorman conditions with $\mathbf{b}[p] = 1/p_0$ and $\mathbf{a}^j[p] = \bar{u}^j\left[\bar{x}^{jh}[p]\right] - (1/p_0)\,\bar{p}\cdot\bar{x}^{jh}[p]$.

The demand function for agent $j$ in a Gorman economy is then

$$x^j\left[p, m^j\right] = \beta[p]\, m^j + \alpha^j[p], \tag{28}$$

with $\beta[p] = -\mathbf{b}'[p]\,/\mathbf{b}[p]$, and $\alpha^j[p] = -\mathbf{a}^{j\prime}[p]\,/\mathbf{b}\,[p]$. The aggregate demand of the economy has the same form, by construction:

$$x[p, M] = \beta[p]\, M + \sum_j \alpha^j[p], \tag{29}$$

with $M = \sum_j m^j$.[18]

For commodity vector $w$ held collectively in a Gorman economy, the equilibrium price system $p^*$ satisfies:

$$\beta[p^*]\, e[p^*,\mathcal{U}^*] = w - \sum_j \alpha^j[p^*]. \tag{30}$$

By an affine transformation of the commodity space, we may set $\mathbf{b}[p^*] \equiv 1$, $\mathbf{a}^j[p^*] \equiv 0$ if we wish, in which case the indirect utility (26) becomes the money-metric utility at equilibrium prices [2].

Pareto allocations in Gorman economies are related by redistribution of bundles proportional to $\beta[p^*]$ among the agents, a change in $x^{jh}$ that preserves $p = p^*$ in Eq. (28). The set of expenditures $m^j$ and utility values $\mathcal{U}^j$ changes under such a re-allocation, but not the first-order conditions, so that the equilibrium price has no memory of the path taken to the trade set.

Non-quasi-linear Gorman economies exhibit a much weaker degree of equivalence to classical thermodynamic systems than quasi-linear economies, so that welfare functions are defined for closed economies only in endowment- and preference-dependent forms, and Walrasian potentials for prices are only guaranteed to give the correct normalization within the Pareto set. The first-order conditions at equilibria are path-independent, but higher-order derivatives of the indifference surfaces will differ at different Pareto allocations.

In a Gorman economy, any contour utility in the Pareto set equals (up to an additive constant) the direct money-metric utility at equilibrium price $p^*$, so that all $\sum_j S^j_{c,i}$ have $c$-independent degenerate maxima on the Pareto set, and we may drop the $c$ index on $S$. To consistently treat this similarity of Gorman to quasi-linear systems, it is natural to make the Gorman bundle (the basis element in the simplex of Pareto allocations) the wealth transfer on which prices do not depend, by normalizing prices so that $p \cdot \beta[p^*] \equiv 1$. To study the individual entropy-maximizing problem of agents, the natural metric for entropy is $p^* \cdot x^j$, and the corresponding normalization $p^*_i$ in Eq. (23) may be replaced by a coordinate-independent form $p^* \cdot \beta[p^*] \equiv 1$ (with an overall scaling depending on the endowments, which in other circumstances may not be appropriate). The individual agent entropy is then

$$\begin{aligned} S^j_{\text{Agg}} &= e\left[p^*,\mathcal{U}^j\left[x^j\right]\right] - p^* \cdot x^j \\ &= \left.\frac{\left[\left(\mathbf{b}[p]\,p - \mathbf{b}[p^*]\,p^*\right)\cdot x^j + \mathbf{a}^j[p] - \mathbf{a}^j[p^*]\right]}{\mathbf{b}[p^*]}\right|_{p=p[x^j]}. \end{aligned} \tag{31}$$

As we have seen above, for an economy of agents with utilities quasi-linear in a specific commodity $x_0$, $\beta[p^*]$ does not depend on $w$ or the idiosyncratic $\bar{u}^j\left[\bar{x}^j\right]$. If we use $x_0$ rather than $p^* \cdot x$ as numéraire, replacing a Gorman-form entropy which is individually maximized with an entropy independent of $x_0$ and $\mathcal{U}$ at all $x$, Eq. (23) reduces to

$$S^j_{\text{QL}} = p^*_0\left(\bar{u}^j\left[\bar{x}^j\right] - \bar{u}^j\left[\bar{x}^{j*}\right]\right). \tag{32}$$

Immediately

$$\delta S^j_{\text{QL}} = p^*_0\, \delta\bar{x}^j \cdot \left.\frac{\bar{p}}{p_0}\right|_{p=p[x^j]}. \tag{33}$$

If $x_0$ is numéraire this variation gives properly normalized prices at all demands. If not, because quasi-linear utilities are aggregatable without respect to total endowments or other-agent $\bar{u}^j\left[\bar{x}^j\right]$, it is generally preferable to remove the normalization $p^*_0$ (added through $p^* \cdot \beta[p^*]$ for notational convenience in the more general Gorman case (31)), to produce $S^j_{\text{QL}}$ as in Sec. II C, with gradient (12) giving relative price $\bar{p}/p_0$ dual to $\bar{x}^j$ in any economy.

The function of entropy is independent of which of these conventions we choose. For economies where all relative prices are invariant among equilibria, it amounts to a convention for the units of measurement. In cases where we can construct entropies independent of a wealth coordinate, however, we gain the interpretation of economic equivalents to thermodynamic effective potentials in open systems, where irreversible and reversible transformations can be compared directly.

---

[18] Note that $\mathbf{b}[p]$ and $\mathbf{a}^j[p]$ cannot be quite arbitrary functions. Utility $v^j\left[p, m^j\right]$ should be homogeneous of order zero in prices, or

$$\begin{aligned} 0 &= \left.\frac{\partial}{\partial\epsilon} v^j\left[(1+\epsilon)\,p, (1+\epsilon)\,m^j\right]\right|_{\epsilon=0} \\ &= \mathbf{b}[p]\left[\left(1 - p\cdot\beta[p]\right) m^j - p\cdot\alpha^j[p]\right]. \end{aligned}$$

Since $p$ and $m^j$ vary independently, $v^j\left[p, m^j\right]$ is homogeneous at all arguments iff $p\cdot\beta[p] = 1$, $p\cdot\alpha^j[p] = 0$.

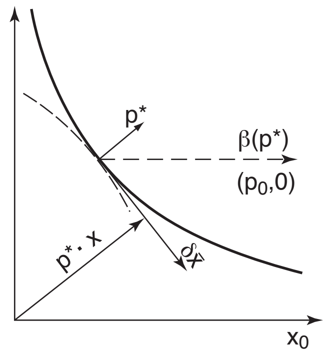


FIG. 6: The choice of reference price permits several different maximized entropies for a closed system trading from a non-Pareto endowment to a Pareto allocation. Numéraire $p^* \cdot x$ (natural for general Gorman case, well-defined but endowment-dependent for the quasi-linear special case) describes the individual's optimization of the relatively defined $\delta \bar{x}$ as a taker of externally-given prices. Numéraire $x_0$ (quasi-linear only) defines endowment- and other agent preference-independent entropy with no functional dependence on either $x_0$ or $\mathcal{U}$. Such context independence is not possible for more general Gorman economies, because the equilibrium price system $p^*$ depends on the endowment and agent preferences. Because a quasi-linear example is shown here, the linear good $(x_0, 0)$ also defines the Gorman bundle $\beta[p^*]$.

### F. Micro-foundations again

These reflections reveal that Gorman economies are aggregatable only in a limited and methodologically incomplete sense. The problem is that while it is possible to define entropy-like functions for individual agents in a particular non-quasi-linear Gorman economy, these functions depend on the economy-wide endowment of commodities and the preferences of all the agents who participate. They provide no real micro-foundation for analyzing the outcome in a Gorman economy. We cannot define individual agent entropies in terms of which we can analyze the dynamics of trade to the Pareto set, independent of the economic context in which the agents are embedded. Although Gorman hoped, and some later economists assumed, that Gorman aggregatability would provide rigorous micro-foundations for macroeconomics (in the form of a "representative agent" paradigm), these observations show that this hope is not actually fulfilled.

### G. Transformations in partly open economies

Reversible transformations preserving the economy's aggregate endowment of commodities are not possible in closed systems. A boundary condition with at least one fixed price and open demand coordinate is needed to compare reversible and irreversible transformations directly. Fixed price for some goods introduces the notion of a trading "reservoir". The idealized reservoir, in which prices do not change under any finite exchange of demands, is not only an analytical convenience; it can be a good first approximation for situations like that of a small-country economy trading some commodities but not others with a world market. The world market effectively fixes prices in the small economy for the traded commodities. We now relate trading to equilibrium in such an economy to the closed-economy entropy maximization.

For a closed economy we can arrive at any point on a contour $c$ within the Pareto set by maximizing the constrained Lagrangian

$$\delta \left[ \sum_j S^j_{c,i} \left[ x^j \right] - \eta \cdot \left( \sum_j x^j - w \right) \right] = 0. \tag{34}$$

The construction of the contour utilities (21) gives immediately that in the vector $\eta$ of Lagrange multipliers, $\eta_i = 0$ and $\bar{\eta} = \bar{p}[x_c] / p_i[x_c]$ at any equilibrium $x_c \in c$. The optimization procedure does not discriminate Pareto allocations, just as the Hahn-Negishi procedure of continuous transacting does not. $\sum_j S^j_{c,i} \left[ x^j \right]$ is the maximal generalization of the Negishi social welfare function, capable of identifying all equilibria on a one-parameter family within the Pareto set, rather than only a discrete subset as in the original Negishi construction.

Suppose, now that we consider such an economy embedded in a world market exchanging two goods $x_{i'}$ and $x_i$ at fixed price $p_{i'} / p_i|_{\mathrm{Res}}$. This time let one of the goods $(x_{i'})$ be numéraire and the other $(x_i)$ the metric for entropy. Write the resulting vector of goods $(x_i, x_{i'}, \tilde{x})$, and similarly for $w$, $p$, and $\eta$. Equilibrium of the reservoir-coupled economy is given by

$$\delta \left[ \sum_j \left( S^j_{c,i} \left[ x^j \right] - \left. \frac{p_{i'}}{p_i} \right|_{\mathrm{Res}} x^j_{i'} \right) \right. \\ \left. - \eta_i \left( \sum_j x^j_i - w_i \right) - \tilde{\eta} \cdot \left( \sum_j \tilde{x}_j - \tilde{w} \right) \right] = 0. \tag{35}$$

Variation of the contour utility sets $p_{i'}[x_c] / p_i[x_c] = p_{i'} / p_i|_{\mathrm{Res}}$, and as before $\eta_i = 0$ and $\tilde{\eta} = \tilde{p}[x_c] / p_i[x_c]$.

If the economy consists of a single agent, all demand components except $x_i$ and $x_{i'}$ are fixed, and the quantity

$$S^j_{c,i} \left[ x^j \right] - \left. \frac{p_{i'}}{p_i} \right|_{\mathrm{Res}} x^j_{i'} = u_{c,i} \left[ \mathcal{U} \right] - x_i - \left. \frac{p_{i'}}{p_i} \right|_{\mathrm{Res}} x^j_{i'} \tag{36}$$

is maximized over $x_i$ and $x_{i'}$. The construction (36) corresponds to physical $S - (1/T) \, E \equiv - (1/T) \, A$, where $A \equiv E - TS$ is called the *Helmholtz Free Energy*. $p_i / p_{i'}|_{\mathrm{Res}}$ corresponds to the temperature $T$ (the energy-price, or numéraire-price, of entropy) set by the reservoir. Physical equilibria with boundaries on extensive quantities other than energy, and a fixed-temperature (intensive) boundary condition setting $\partial S/\partial E = 1/T$ are identified by minimization of $A$. Likewise, economic equilibria on an arbitrary contour $c$ consistent with relative-price boundary conditions on two commodities are identified with minimization of

$$\sum_j \mathcal{A}^j_{c,i'} \equiv \sum_j x^j_{i'} - \left.\frac{p_i}{p_{i'}}\right|_{\text{Res}} S^j_{c,i} = \sum_j x^j_{i'} + \left.\frac{p_i}{p_{i'}}\right|_{\text{Res}} \left(x_i - u_{c,i}\left[\mathcal{U}\right]\right), \quad (37)$$

subject to whatever additional (extensive) constraints apply to non-traded goods. (The generalization of the contour appropriate to agents interacting with a reservoir, in which $x_i$ and $x_{i'}$ are not individually conserved, but the combination $p_i x_i + p_{i'} x_{i'}$ is conserved, is taken up in the appendix.)

From Eq. (24), it follows that about any equilibrium

$$\delta \mathcal{A}^j_{c,i'} = -\delta \tilde{x}^j \cdot \left.\frac{\tilde{p}}{p_{i'}}\right|_{\tilde{p}=\tilde{p}\left[x^j, p_{\text{Res}}\right]}. \quad (38)$$

$\mathcal{A}^j_{c,i'}$ is a Walrasian potential of the fixed components of the allocation, whose gradient gives the conjugate components of price, *independent of whether the variations correspond to reversible or irreversible transformations.* Note that $\mathcal{A}^j$ is neither the utility nor the entropy, but rather a dual defined in the context of a partly open system enabling reversible transformations, through which the duality relations are defined as the allocation boundary conditions change. Eq. (38) is the economic equivalent of the physical variation

$$\left.\frac{\partial A}{\partial V}\right|_T = -p \quad (39)$$

which identifies the gradient of the Helmholtz potential as (minus) the force conjugate to volume.

In the general case there is no natural discrimination in the choice of $x_{i'}$ for numéraire and $x_i$ for entropy metric. Moreover, the usefulness of the Walrasian potential interpretation (24) is limited, both because of the dependence on contour $c$ and the possibility of an $x$-dependent factor scaling prices even in Gorman economies away from the Pareto set. However, in quasi-linear economies with a natural choice $x_0$ (the degenerate coordinate in the equation of state) for entropy measure and any other numéraire, the Walrasian interpretation of $\mathcal{A}^j$ as a price potential leads to another interpretation equivalent to the physical concept of "free energy" in a thermal system. Suppose that the reservoir trades a numéraire good $x_1$ for $x_0$, so that the temperature $T \leftrightarrow p_0/p_1$ becomes the money price of entropy in the small market, and $x_1 + (p_0/p_1)\, x_0$ defines its balance-of-payments budget constraint. From such a case we can identify $\sum_j \mathcal{A}^j \left[x^j\right]$ as an intrinsic money measure of the value of endowments $\left\{x^j\right\}$, equivalent to a physical measure of available potential energy.

### H. Transformations with all derivative constraints

Equilibria with all price boundary conditions, immediately, are defined by extremization of the full Legendre dual to entropy:

$$\delta \left[\sum_j \left(S^j_{c,i}\left[x^j\right] - \left.\frac{\bar{p}}{p_i}\right|_{\text{Res}} \bar{x}^j\right)\right] = 0 \quad (40)$$

As noted in Sec. VI D, when we first constructed the contour entropy

$$S^j_{c,i}\left[x^j\right] - \left.\frac{\bar{p}}{p_i}\right|_{\text{Res}} \cdot \bar{x}^j = u^j_{i,c}\left[x^j\right] - \left.\frac{p}{p_i}\right|_{\text{Res}} \cdot x^j \quad (41)$$

may be regarded as equivalent to $-\left(1/T\right) F$, where $F = E + pV - TS$ is the the Gibbs free energy introduced in Sec. III D. Its economic equivalent,

$$\mathcal{F}^j_{c,i'} \equiv \left.\frac{p}{p_{i'}}\right|_{\text{Res}} \cdot x^j - \left.\frac{p_i}{p_{i'}}\right|_{\text{Res}} u_{c,i}\left[\mathcal{U}\right], \quad (42)$$

is the potential minimized at equilibria under all intensive boundary constraints. $\mathcal{F}^j$ shares with $\mathcal{A}^j$ the interpretation of a potential for profit extraction, only in the context of trades through which price, rather than volume, is the controlling interface to the goods $\tilde{x}$.

With the Gibbs potential we have come full circle, to the original familiar problem of maximizing utility subject to a constraint on prices. The difference of this construction from conventional economic definitions of the expenditure function or the indirect utility is that the endowment wealth constraint does not appear in the maximization; only relative prices appear. This weakening of the extremization problem is what allows Eq. (41) to specify all equilibria in a contour $c$ (or in the Pareto set of a Gorman economy), rather than a single Pareto allocation associated with a particular wealth constraint.

### I. Summary of duality and transformations

The physical entropy and its various duals, together, provide a suite of *thermodynamic potentials* through which the state relations and transformation structure of classical thermodynamics are defined. Temperature is distinguished among the intensive state variables (and therefore used as the scale factor for the Gibbs and Helmholtz potentials) because it relates entropy changes to changes in energy, the fundamental (extensive) constraint on the space of accessible states and the physical equivalent of the numéraire and balance-of-payments constraint.

The thermodynamic potentials are constructed so that the gradient of each with respect to its arguments is the *energetic* dual vector (see Eq. (16) and Eq. (39)). For this reason the thermodynamic potentials also define limits of available energy under transformations of their arguments as boundary conditions. The gradients of Helmholtz-type potentials (the mixed duals to the entropy) with respect to their extensive arguments produce the conjugate forces, whose balance identifies thermodynamic equilibria. It was therefore to be expected that, in those cases where an economic entropy has a natural definition, the economic Helmholtz potentials constructed from it would be implementations of Walrasian potentials for utilitarian "force balance". The correspondence is exact for quasi-linear economies, and valid at the level of first-order conditions in Gorman economies.

However, the Walrasian potentials (free energies) do not correspond to physical energies, though the goods in which they are denominated do; remember from Sec. IV that the function of energy is to constrain the possible configurations. Nor do they correspond to utilities (remember that even the quasi-linear economic entropy is not the utility). The essential difference between utility and entropy is that entropy can be exchanged through heat (or money) flow, while utility cannot. Thus it can be stipulated that utility increases for agents whether they are isolated or embedded in economies, whereas the entropies of idividual agents embedded in economies can increase or decrease, under suitable circumstances. The inequivalence of entropy and utility captures the way markets make advantageous trades available to agents, while respecting the global constraint that all trades must be volutary.

### J. How restrictive is quasi-linearity?

Our analysis has touched on, but not thoroughly discussed, several issues concerning aggregatability that may yet result in confusion. A final example that ties them together can aid interpretation, and also address questions of applicability.

Economists have already well understood that quasi-linearity buys strong aggregatability, but have regarded the loss of generality from assuming quasi-linear utilities too great a price to pay for some additional path independence of the properties of equilibria. Quasi-linearity also distinguishes one good ($x_0$) from all others, and we may ask what characteristics of real markets create such a distinction. Concerning this question, our characterization of $x_0$ as "money" in Sec. II B may be misleading, depending on which concepts are included in the notion of "money".

Recall that $x_0$ is the one degenerate commodity under the equivalence class of quasi-linear indifference surfaces, and the one commodity that does not have the interpretation of a "state variable" of the economy, under the thermodynamic correspondence. We must extend our notion of state to include an "entropy debt" to the environment, in a somewhat artificial way, to construct a thermodynamic system in which the counterpart to $x_0$ has meaning. In the neoclassical theory of pure exchange, this is an acceptable interpretation for a medium of exchange, in keeping with arbitrariness of $x_0$ as numéraire. The remaining "state variables" of the economy are the non-money commodities whose holdings have nontrivial effects on relative prices.

Yet in many economic models, money is a commodity in its own right, often serving as the measure of wealth. In such cases there is no reason why utility should be quasi-linear in money any more than in other commodities. If we want some measure of wealth to serve as numéraire, then $x_0$ should be a different commodity, whose holding by any agent does not directly affect prices, but which can continue to serve as a medium of exchange and the metric for economic entropy, as was assumed in Sec. VI G. The following example will show that the most natural contexts for quasi-linearity often motivate such a separation.

We consider familiar dividend-discount models from finance. A risky asset returns dividends, drawn independently from a stationary distribution at a sequence of times with interval $\delta t$. There is a market for the asset and also for money, which may be borrowed in return for a promise to pay service on all outstanding debt. Now $\delta t$ may be a property of the *asset*, responsible for its stochastic returns, but suppose we do *not* wish to assume that $\delta t$ is a property of the market's institutional structure. In other words, the decisions to borrow or repay capital, or to buy or sell the risky asset, may be made at any time. Such an assumption is appropriate for many financial (and other) markets, which may exchange a range of risky assets characterized by different intrinsic intervals $\delta t$. If $M$ is an agent's (money) wealth, and the agent borrows $\delta M$, the only debt service incurred per period, consistent with the assumption of time continuity of the markets, is $\delta D = \delta M \, r \delta t$. The interest rate $r$ need not be the same for all periods, but it must be a smooth function of time as $\delta t \to 0$.

Let $N$ be the number of shares of the risky asset some agent holds in any period of length $\delta t$. We will show that a natural utility model for the problem of pricing the asset is based on a commodity bundle $(x_0, x_1, x_2) \equiv (-D, M, N)$. Capital can change by purchase of shares at price $p_N$ as well as by borrowing, so that a general exchange satisfies

$$\delta M = -p_N \delta N + \frac{1}{r\delta t} \delta D. \tag{43}$$

Prices in $M$-metric are thus $(p_0, p_1, p_2) \equiv (1/r\delta t, 1, p_N)$, and correspond under the thermodynamic association to physical $(T, 1, p)$[19]. The accounting identity (43) corresponds to the physical constraint of energy conservation,

$$\delta E = -p \, \delta V + \delta Q, \tag{44}$$

with $\delta Q = T \delta S$ if the transformation is reversible.

[19] We have denoted share price $p_N$ to distinguish it from the thermal equivalent, pressure $p$.

We now ask what class of utilities can mirror the time continuity provided by the market. We suppose that the only intrinsic value of the asset on very short times is its dividend, for which we let $d$ denote the value in the current timestep in a single realization. Any cardinal intertemporal utility capable of evaluating the holding of assets for an arbitrarily short time must be a function of $Nd - D$ for the current period, and must have regular limit as $Nd - D \to 0$. If we express the utility as a power series in $Nd - D$, and divide by its first derivative at $Nd - D = 0$, the remaining term takes a form equivalent to a wealth change in the period

$$\Delta W \equiv Nd - D + \phi[M] \, . \tag{45}$$

$\phi[M]$ may come from a deterministic component of income accruing to money $M$ in the period, but may also reflect any possible future utility of wealth that is impacted by holding $M$ through the current period. Higher-order terms in $Nd - D$ are $\mathcal{O}(\delta t^2)$ (in expectation), and vanish as $\delta t \to 0$. Terms at lower order in $\delta t$ are incompatible with a sensible *continuum limit*, meaning that a succession of short periods of total length $K\delta t$ in which no action is taken must have the same utility as if the basic period were $K\delta t$ and the dividend process were simply accumulated through the period.

We can remove the uncertainty about the dividend process in a number of ways, of which the Constant Absolute Risk Aversion (CARA) model is an algebraically convenient example. CARA transforms $\Delta W$ of Eq. (45) into a cardinal form

$$u_{\text{card}} \equiv -\exp\left\{-\Delta W/\nu\right\} , \tag{46}$$

and assumes that dividends $d$ are distributed normally, with mean $\langle d \rangle \equiv \bar{d}$ and variance $\left\langle \left(d - \bar{d}\right)^2 \right\rangle \equiv \bar{d}^2 \sigma^2$. $\nu$ is the agent's risk tolerance, measured in units of $M$. The ordinal utility resulting from averaging Eq. (46) over $d$ is

$$u_{\text{ord}} = -\exp\left\{ - \left[ N\bar{d}\left(1 - \frac{N\bar{d}}{2\nu}\sigma^2\right) - D + \phi[M] \right] /\nu \right\} , \tag{47}$$

equivalent under monotone transformation to the so-called *certainty equivalent of wealth*[20]

$$\mathcal{U} \equiv N\bar{d}\left(1 - \frac{N\bar{d}}{2\nu}\sigma^2\right) - D + \phi[M] \, . \tag{48}$$

The utility $\mathcal{U}$ differs from the QLCD utility (9) only in having a polynomial rather than logarithmic dependence on $N$ ($x_2$) and arbitrary (concave) dependence $\phi$ on $M$ ($x_1$).

[20] A better name in this case would be "certainty equivalent of income".

The expenditure function is

$$e\left[p, \mathcal{U}\right] = M + p_N N - \left[ \phi\left[M\right] + N\bar{d}\left(1 - \frac{N\bar{d}}{2\nu}\sigma^2\right) - \mathcal{U} \right] /r\delta t, \tag{49}$$

where the minimizing values of $M$ and $N$ satisfy

$$r\delta t = \frac{\partial \phi}{\partial M} \tag{50}$$

and

$$\left(r\delta t\right) p_N = \bar{d}\left(1 - \frac{N\bar{d}}{\nu}\sigma^2\right) . \tag{51}$$

These marginal rates of substitution are also the rates of exchange in reversible trades (43).

The $\mathcal{U}$- and $D$-independent entropy is then

$$S \equiv \mathcal{U} + D = N\bar{d}\left(1 - \frac{N\bar{d}}{2\nu}\sigma^2\right) + \phi\left[M\right] = \bar{u}\left[x_1, x_2\right] , \tag{52}$$

in the notation of Eq. (32). Like the ideal gas of thermodynamics, the entropy (52) is separable. From Eq. (49), the Gibbs potential is

$$\mathcal{F} = M + p_N N - \frac{1}{r\delta t} S, \tag{53}$$

and the corresponding Helmholtz potential

$$\mathcal{A} = M - \frac{1}{r\delta t} S. \tag{54}$$

Eq (50) and (51) give immediately upon differentiation that

$$p_N = -\left.\frac{\partial \mathcal{A}}{\partial N}\right|_{r\delta t} \tag{55}$$

for arbitrary transformations. Thus $\mathcal{A}$ is the Walrasian potential for price.

We see from this example, and in particular the forms (52-54), that the *debt service* $D$ is the natural quasi-linear commodity and metric for entropy, though it is not the measure of wealth, the store of value across periods, or even the numéraire, all of which roles are filled by $M$. The characteristics that generically lead to utility linear in $D$ make it equally intuitive that $D$ should be a degenerate coordinate of the indifference surfaces. $D$ maps value across time, but itself has arbitrary scale proportional to $\delta t$. Unlike a stock variable like $M$ or $N$, it is a flow used to maintain equilibrium in the other commodities.

## VII. DISCUSSION

Classical thermodynamics came to formulate its dual variables as *state variables*, related through a conceptually central function called *entropy*. The entropy, which expresses the forces as functions of the coordinates through a relation called the *equation of state*, came to be understood as serving this role only for systems *at equilibrium*. Thus the very existence of a dual representation for thermodynamic systems came to be tied intrinsically to the condition of equilibration. The use of thermodynamic equilibrium does not preclude transformations, but it is restricted to paths that lie within the surface traced out by equilibrium equation of state, the *reversible transformations*. Reversible transformations are used to measure the values of all the state variables and to deduce the equation of state, which characterizes the structure of the thermodynamic system at all possible equilibria, and through all possible reversible transformations.

While only reversible transformations can be used to measure system structure, thermodynamic prediction is not restricted to cases in which only reversible transformations have occurred. The endpoints of irreversible transformations – those for which intermediate configurations may be far from equilibrium – can be predicted *if* they are known to be path-independent, and could thus have been reached by alternative reversible transformations. *Classical* thermodynamics applies to those systems for which the equation of state is a complete description at equilibrium, meaning that there are no other hidden quantities capable of retaining a memory of paths of transformation.

Neoclassical economics, in contrast, has not come to limit the use of dual representations of agents to conditions of equilibrium, but rather assumes that such representations exist under all conditions. It has also come to emphasize transformations *to equilibrium in closed economies*, a class of transformations that are incompatible with maintaining conditions of equilibrium along the transformation paths. Because of these choices, economic theory has not developed a role for concepts equivalent to reversible transformations, or a special class of path-independent systems.

These differences matter because it is well-known in physics that the thermodynamic formalism cannot predict the outcomes of irreversible transformations in path-dependent systems (of which many kinds have been studied [15–17]). The same is true in economics: the prices and allocations that result from trading from initial non-equilibrium endowments to an equilibrium cannot be deduced purely from the dual representation of the preferences of agents. Yet this is the stated goal of general equilibrium theory [5], which purports to achieve it by adding to the dual representation of preferences the Walrasian condition for *wealth preservation* at eventual equilibrium prices during disequilibrium trading.

Wealth preservation as it is used in the fundamental theorems of general equilibrium theory is not implied by economic axioms defining utility, and can be empirically wrong even if a dual representation of preferences is justified[21]. Unfortunately, this means that the theoretical device of wealth preservation doesn't really overcome the problem of indeterminacy in path-dependent economies.

General equilibrium theory simultaneously identifies prices with the structure of agent preferences and the dynamic process of disequilibrium trading. Doing so changes the interpretation of prices from the thermodynamic concept of state variables to the economic concept of transactions ratios along paths leading from nonequilibrium endowments to an equilibrium. But whereas state variables measured in reversible transformations are deterministic consequences of the structure of thermodynamic systems, the transaction ratios of trades to equilibrium are not deterministic and need not coincide with the prices derived from general equilibrium theory, since the latter arise from a false solution to the problem of indeterminacy.

We have derived in this paper the interpretation of equilibria, transformations, and price systems that arises from a consistent methodological correspondence in the representation of agent preferences and thermodynamic systems. We find that, in economics as in thermodynamics, there is a useful role for reversible transformations, and an important class of path-independent systems (comprising subsets of observables in the the so-called quasi-linear economies) for which deterministic predictions of an important subset of equilibrium variables (the prices and allocation of the nonlinear commodities) can be made. The thermodynamic correspondence cannot overcome the problem of indeterminacy in more general path-dependent economies, but it allows us to understand the differing categories of path dependence that are possible and their relation to the economic notion of *aggregatability*, and clarifies the circumstances in which the auxiliary notion of wealth preservation from general equilibrium theory applies.

We have shown that the formal structure of utility theory already includes functional analogues to thermodynamic quantities like energy and entropy, which is interesting in its own right because it emphasizes how much of the representation of inanimate systems economics has adopted for the modeling of rational agents. For path-independent economies the correspondences are strict and constructive. Utility is not entropy, but it is related to a well-defined economic entropy that expresses prices as functions of quantities of goods held, and identifies which transformations are possible under voluntary exchange, the same functions entropy performs in thermodynamics. Though energy occupies a privileged position in much of physics, it is only one of many state variables within the thermodynamic formalism, whose roles

[21] We have shown in Sec. V that wealth preservation is intrinsically a theory of *market function* and not of preferences, and the existence of real trading systems that produce very different outcomes automatically implies that any single algorithm for wealth preservation can be valid for at most one of them.

are collectively to specify which states it is possible for combined systems to adopt. The quantities of commodities held perform the equivalent function for exchange in closed economies. Entropy ceases to exist as a well-defined function of the state of path-dependent systems in both physics and economics, but in cases where some of its properties are preserved, a generalization of the construction from quasi-linear economies continues to be possible. In these cases again we observe that it is not the differences between inanimate objects and rational agents that has distinguished thermodynamics and utility theory, but the approach they have taken to path dependence and indeterminacy.

### A. A century of analogy-making between economics and physics

Walras [4] apparently thought of utility as a measurable quantity analogous to potential energy in mechanics, whose gradient (with a minus sign) is the force exerted by a system on its boundaries. Economic equilibrium was to correspond to force balance, and vector equality of forces to equality of prices. There were two flaws in the Walrasian correspondence, one surmountable, one fatal.

The surmountable flaw was that Walras sought a correspondence with rational mechanics [7], the physics of his day. We now understand that mechanics cannot cause convergence to equilibrium. Purely mechanical systems started at rest with all forces balanced can remain in that state, but systems started away from equilibrium will oscillate about it forever. A mechanical analogue in economics would require not only forces, but dynamical equivalents to inertia.

This flaw was recognized and corrected by Fisher [8] after it was understood that thermodynamics, and not mechanics, is the correct physical theory to explain how disequilibrium systems can converge to equilibrium and remain there. The thermodynamic law of increase of entropy subject to constraints (on total energy, volume, etc.) [18] explains how systems can leave disequilibrium regions of configuration space and never return, so that eventually the equilibrium configurations are the only ones left available to them. The potential minimized at physical equilibria is not the energy, as was believed in Walras's time, but a quantity called *free energy*, which also receives contributions from the entropy.

Fisher's doctoral advisor was J. W. Gibbs, one of the principle architects of the new field of *statistical mechanics*, which was just then providing a microscopic foundation for the rules of thermodynamics, in particular the law of increase of entropy. Fisher went on from correcting Walras's error of analogy, in light of new physical understanding, to commit an even more elementary logical error, in proposing a point-by-point formal correspondence of utility theory with elements not only from thermodynamics, but also from Gibbs's statistical mechanics. Fisher proposed that the economic agent corresponds to the microscopic particle, but continued to compare the agent's commodities and utility to thermodynamic state variables like energy and entropy. The error is that microscopic particles in statistical mechanics are stochastic objects, and the thermodynamic state variables are deterministic functions obtained from averages over distributions of microscopic particle states. Fisher was thus drawing analogies of the economic agent to both stochastic and deterministic physical objects at the same time.

We emphasize here that thermodynamics is a classical, deterministic formal system, which is internally consistent whether or not it is derived from a statistical mechanics foundation [3]. To the extent that utility theory has correspondences with physics, they are with thermodynamics and not statistical mechanics. Much current work in "econophysics" is devoted to replacing the assumptions of utility theory with statistical models of agents and markets, but that is *not* what this paper is about. Statistical mechanics has been important in physics because, in addition to explaining equilibrium thermodynamics, it has provided some mechanisms to predict the behavior of nonequilibrium transformations. It is an interesting question whether there is a similar statistical foundation for neoclassical theory that can also be extended to disequilibrium trading, but we take no position on that question here.

The fatal flaw in the Walrasian correspondence was the assumption that cardinal forms for utilities could be constructed *a priori*, whose gradients would equal prices for all the equilibria of arbitrarily composed economies. Such a condition is equivalent to requiring that equilibria be maxima of a global *social welfare function*, which is the sum of the utilities of all agents in an economy. This amounts to postulating a universal addition rule for cardinal utilities.

In the neoclassical theory of preferences, price normalization is arbitrary,[22] and utility gradients need only be parallel in equilibrium. This fact was exploited by Negishi [19] to represent Walrasian allocations as maxima of social welfare functions constructed by adding arbitrary cardinal utilities *after* multiplying these by appropriate scalar weights. We develop in Sec. VI C a maximal generalization of Negishi's construction by introducing a new contour money-metric utility, which allows us to define a social welfare function on an arbitrary one-parameter set of equilibria for any economy, rather than an arbitrarily chosen single equilibrium as in the Negishi construction. Though some aspects of the utility/potential correspondence can be salvaged in this way, the lack of a natural cardinalization for general utilities reflects the deeper fact that welfare functions simply do not exist for general path-dependent economies.

Lisman [9] recognized Fisher's logical inconsistency, and (correctly) associated the economic agent with the

[22] See discussion in Ref. [5], p. 28.

quasi-deterministic macroscopic thermal system, as we will do here. He labeled utility an "analogon" to entropy, but proposed no specific relation between the two. He also (correctly) identified the function of resource constraints with that of energy conservation in thermodynamics: specifying the surface of possible configurations within which utility or entropy is maximized. Lisman then fell back into the trap of formal analogy, by proposing that the expenditure function $p \cdot x$ corresponds to the term $pV$ in the physical ideal gas equation of state, because both are products of dual state variables. We show, starting in Sec. IV, what the expenditure function means under the correct correspondence, and that in general it has nothing to do with the particular equation of state of the ideal gas.

Other formal analogies have been pursued without any basis in methodological correspondence. Bryant [10] attempts to reproduce the ideal-gas equation of state by inventing a quantity "productive content" as a placeholder for temperature in that equation. "Productive content" is assigned a number but is not measurable even in principle, in contrast to temperature, which is the foremost measurable quantity in thermodynamics, usually considered the object of its "0th law" (discussed in Sec. IV A). More recently Saslow [11] has proposed a somewhat eclectic mapping between thermodynamic and economic variables, which, however, leaves the exact relation of utility and entropy and the range of models in which it holds undefined.

A few authors have found correct associations without proposing specious analogies, though all have stopped short of a methodological correspondence or complete interpretation. Bródy [13] et. al., in macroeconomics, recognize that resource conservation constrains the surface of possible configurations for coupled subsystems, and thus corresponds functionally to energy in thermodynamics. Candeal [14] et. al., demonstrate the mathematical equivalence of the utility representation problem to that of entropy representation in physics, though they propose no specific relation of the two functions.

Samuelson's [12] is the most complete (correct) relation of the theories of thermodynamic and economic equilibrium, emphasizing duality of state variables and the correspondence of expenditure functions to certain thermodynamic free energies. His concern is measurability of preferences and identification of equilibria, and thus the interpretation of economic dual variables, but not transformations or economies in which the equilibrium allocations change as a result of net flow of goods. Transformations are, however, essential, because they provide the mechanism for measuring the values of state variables and empirically inferring the equation of state.

## VIII. CONCLUSIONS

In representing the behavior of individuals (or households) as the maximization of well-behaved utility functions (which represent transitive, convex preference orderings) under constraints, marginalist and neoclassical economics effectively regards individuals as equilibrated thermodynamic systems in which a well-defined equation of state links extensive variables (commodity bundles) to intensive variables (marginal rates of substitution). This point of view has become very widely propagated through its adoption as the foundation of the economics curriculum and as the starting point for an enormous theoretical-empirical literature on economic problems. This conception of the individual leads to an analysis of real economic phenomena as the interaction of the individual agents (analogous to equilibrated subsystems in thermodynamics) and hence as a version of thermodynamic equilibrium.

History seems to show that no economist has had a clear understanding of the full methodological implications of this thermodynamic perspective for economics as an explanatory science. The stumbling blocks appear to have been that in thermodynamics there is no natural role for accumulated heat flow as a state variable, and thus no reason to expand the equation of state into a quasi-linear function akin to a utility. Conversely, economists have appropriately regarded quasi-linear economies as too constrained to represent the full range of economic phenomena. As a result economics has developed a general theory (Walrasian general equilibrium theory), but on methodologically flawed foundations which foreclose the tight and scientifically fertile connection between theory and measurement enjoyed by thermodynamics.

Economics unwittingly found itself coping with the complex phenomenon of path-dependency in its attempt to theorize general economic interactions not constrained to the quasi-linear case. The device of the Walrasian auctioneer is an unsuccessful attempt to finesse the issues raised by path-dependency without coming to grips with them (through positing a determinate transactions path for irreversible transformations from disequilibrium endowments to the Pareto set). In the process both economics and physics lost sight of the common ground that underlies their approach to complex systems. We hope that a better understanding of the fact that physics and economics face the same problems in applying thermodynamic reasoning to path-dependent systems can foster a more fertile interchange of ideas.

This paper takes no stand on how appropriate the abstract thermodynamic point of view is to understanding real economic phenomena. It is possible to argue that the formal equivalence of utility-based economics and thermodynamics reveals the inadequacy of this approach to deal with human social phenomena. It is also possible to argue that a correct application of the methodology implied by the thermodynamic approach can greatly strengthen the empirical explanatory power of economic theory. However one argues, though, those economists who commit themselves to utility theory as a basic framework of analysis will posess a clearer conceptual system when the theory of preferences and duality is disentangled from the problems associated with disequilibrium, dynamics, and inevitably institutions.

### Acknowledgment

DES is grateful to acknowledge Insight Venture Partners for support.

## APPENDIX A: ECONOMIC INTERPRETATION OF THE HELMHOLTZ "FREE ENERGY"

Here we generalize the case diagrammed in Fig. 2, of a speculator interacting with an economy of agents with utilities quasi-linear in $x_0$, to include a reservoir. The agents can freely trade $x_1$ for $x_0$ with a world market at price $p_0/p_1$, giving us a well-defined counterpart to temperature. Under this correspondence, the problem of the speculator, the economy, and the world market is the counterpart to the system of piston, gases, and thermal reservoir shown in Fig. 3. Again we suppose that the agents cannot trade internally, and that only capital ($x_1$ or $x_0$ interconvertible through the reservoir) is of interest to the speculator, who can mediate trade in $\tilde{x}$. The physical counterparts are that the gases cannot exchange energy except through the mediating piston or the thermal reservoir, and energy is the only quantity that the piston can extract from the system.

The speculator exchanges $x_1 + (p_0/p_1)\, x_0$ for $\tilde{x}$ with the agents along their indifference surfaces as shown in Fig. 7, again maintaining zero inventory of $\tilde{x}$. Thus $\tilde{x}^j$ plays for each agent $j$ the role of volume ($V$) shared between two chambers in Fig. 3[23]. Work can be extracted from the system if the initial pressures (energy-prices of volume) are different, by trading volume between the two chambers, at pressures that differ on the two sides of the piston, with the excess $\int \left(p^1 - p^2\right) dV^1$ extracted by the load.

Because equilibrium prices for each agent in the small economy do not depend on $x_0$, the speculator can move the agents to prices on $x_1, \tilde{x}$ equivalent to their prices at any closed-economy Pareto allocation, at the minimal acceptable $x_0^j$ for each agent $j$. She can then decouple from the economy with surplus $x_0$ convertible to $x_1$ by the reservoir, leaving the small economy with no further advantageous trades. The wealth obtained by the speculator is $-\sum_j \delta\mathcal{A}^j = (p_0/p_1) \sum_j \delta x_0^j$, because agent utilities did not increase during the extraction.

[23] In the closed-economy analysis of Sec. II C, the speculator comprised both reservoir and load together, and since in all cases the condition in which she left the economy when she finally decoupled had common $\bar{x}$, only net $x_0$ in the system differed between internally-obtained equilibria and maximal extraction, to serve as a measure of wealth.

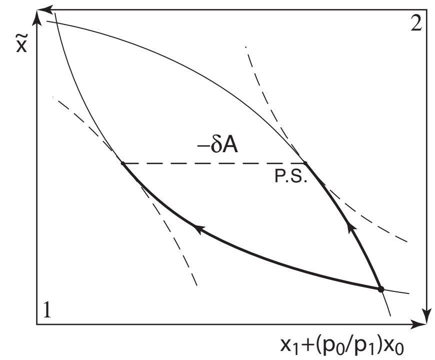


FIG. 7: An external speculator extracts wealth from a small economy of two agents having no internal trade. $\bar{x} = (x_1, \tilde{x})$. If $\tilde{x}$ is the controlled variable, the instantaneously equilibrated $x_1[\tilde{x}]$ is determined by the indifference surfaces and the reservoir price $p_0/p_1$. Trades along indifference surfaces (heavy) are the worst agents will accept, and take place at different prices. At the termination of trade, agents in the small economy have identical prices, and have lost wealth equal to $-\delta A$ to the speculator.

As for the entropy in Sec. II C, $\delta\mathcal{A}^j$ depends only on $\delta x_1$, $\delta\tilde{x}$, and does not discriminate reversible from irreversible transformations. For any combination of speculative extraction and internal trade, the sum of capital extracted and welfare gained is a function only of $\delta x_1$, $\delta\tilde{x}$ of the transformation. The economic Helmholtz potential is an intrinsic money-metric welfare measure of the allocation of an economy in contact with a reservoir, equal to its potential to deliver money wealth (now a nonlinear good) to an external market through voluntary trade (generally through a combined loss of $x_1$ and $x_0$ mediated by the reservoir). This interpretation coincides with the interpretation of the physical Helmholtz potential $A^j$ as the maximum *work* extractable from a system through change of a volume boundary, when the system also contacts a reservoir for *heat* (energy accompanied by entropy). $A^j$ differs from the mechanical energy in the system (here equivalent to the balance-of-payments constraint $\sum_j x_1^j$), and is called the *free energy* of the reservoir-coupled system.

The interpretation of the Helmholtz free energy as a maximum potential for profit extraction remains generally valid with respect to any contour $c$ of equilibria, as shown in the Appendix, though in general it comes to depend on the endowments and preferences of other agents, and on the contour chosen. In the problem of aggregating only two sub-economies, each internally equilibrated, the Pareto set is one dimensional, so well-defined entropies with respect to that aggregation problem exist for arbitrary reservoirs.

### 1. Example 1 continued: Cobb-Douglas entropies and the valuation of trade

The ability to define entropy and contour-independent free energies as individual agent functions, which are independent of the endowment and preferences of other agents in the economy in which they are embedded, is lost even for simple Gorman-aggregatable utilities such as Cobb-Douglas. However, the interpretation of the entropy component of the free energies as a money-metric welfare measure remains useful to assign an intrinsic value to the outcomes of trade, as in Sec. II C. A Cobb-Douglas example is shown in Fig. 8. The entropy is defined as in Eq. (31), with the Gorman bundle that defines the Pareto set also defining the wealth measure through the normalization convention $p \cdot \beta[p^*] \equiv 1$. The welfare gained in trade between any two distributions is the difference between the change in $p^* \cdot x^j$ and the the money-measure of the segment of Pareto set between the indifference surfaces for the distributions. In the example, "wealth-preserving" trade outcomes are shown, so that $p^* \cdot x^j$ does not change, and the utility increase is also the entropy increase.

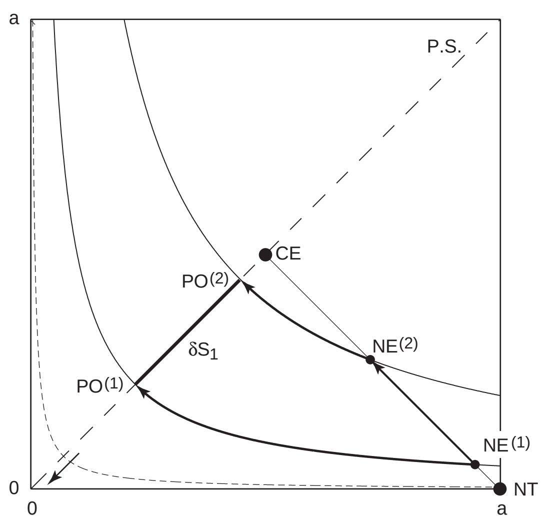


FIG. 8: Entropy welfare measure in a Cobb-Douglas economy for two goods with initial endowment $(a, 0)$ for the agent shown, and total endowment $(a, a)$. NT is the No-Trade solution. $\mathrm{NE}^{(1)}$ and $\mathrm{NE}^{(2)}$ are other outcomes, which might be reached as Nash Equilibria of non-cooperative games, and CE is the wealth-preserving Walrasian allocation. $\mathrm{PO}^{(1)}$ and $\mathrm{PO}^{(2)}$ are Pareto allocations indifferent to the respective non-cooperative equilibria, and $\delta S_1$ is the entropy measure of utility of $\mathrm{NE}^{(2)}$ over $\mathrm{NE}^{(1)}$. The value of NT is asymptotically that of zero endowment, as indicated by the dashed line.

Assigning money values to outcomes of trade generates interesting measures of the *efficiency* of trades that take agents toward the Pareto set, but may not reach it. These efficiencies are defined for *all* allocations in the economy, and provide one of the more useful applications of economic entropy in economies with even weak (Gorman) aggregatability.

An alternative form to Eq. (5) for the utility of an arbitrary number of goods to a Cobb-Douglas agent $j$ is

$$u^j_{\mathrm{CD}}\left[x^j\right] = \prod_i \left(x^j_i\right)^{\alpha_i}, \tag{A1}$$

where $\{\alpha_i\}$ are exponents common to all agents in the economy. As before, let the endowments in the closed economy be $\sum_j x^j_i = w_i$.

It is straightforward to show that the Hicksian demands for the utilities (A1) are of Gorman form (28),

$$x^{jh}_i\left[p\right] = \frac{\hat{\alpha}_i}{p_i} m^j, \tag{A2}$$

with $\hat{\alpha}_i \equiv \alpha_i / \sum_{i'} \alpha_{i'}$, so that for this economy

$$\beta_i\left[p\right] = \frac{\hat{\alpha}_i}{p_i}. \tag{A3}$$

As required $p \cdot \beta[p] = \sum_i \hat{\alpha}_i \equiv 1$. $\mathbf{a}^j[p]$ and $\alpha^j[p]$ are identically zero.

Equilibrium prices follow from equating the sum of Hicksian demands to the endowments:

$$p^*_i = \frac{\hat{\alpha}_i}{w_i}\left(p^* \cdot w\right), \tag{A4}$$

where $(p^* \cdot w)$ may be chosen for normalization of prices. The Gorman bundle at a Pareto allocation is

$$\beta[p^*] = \frac{w}{p^* \cdot w}. \tag{A5}$$

Some algebra gives the expenditure function at equilibrium prices,

$$e\left[p^*, \mathcal{U}^j\left[x^j\right]\right] = \prod_i \left(\frac{p^*_i x^j_i}{\hat{\alpha}_i}\right)^{\hat{\alpha}_i}. \tag{A6}$$

It is helpful to recognize that $\hat{\alpha}_i$ defines a normalized distribution over commodities $i$, and that another such distribution is defined by the allocation of wealth to commodities at price $p^*$ and bundle $x^j$:

$$\hat{\omega}^j_i \equiv \frac{p^*_i x^j_i}{p^* \cdot x^j}. \tag{A7}$$

In terms of these distributions we can write

$$e\left[p^*, \mathcal{U}^j\left[x^j\right]\right] = \left(p^* \cdot x^j\right) e^{-D\left(\hat{\alpha} \| \hat{\omega}^j\right)}, \tag{A8}$$

where

$$D\left(\hat{\alpha} \parallel \hat{\omega}^j\right) \equiv \sum_i \hat{\alpha}_i \log \frac{\hat{\alpha}_i}{\hat{\omega}^j_i} \tag{A9}$$

is the positive semidefinite Kullback-Leibler divergence [23]. $D\left(\hat{\alpha} \parallel \hat{\omega}^j\right)$ vanishes only on $\hat{\alpha} = \hat{\omega}^j$, the allocation for agent $j$ in the projection of the Pareto set onto $x^j$.

The Gorman entropy (31) evaluates for Cobb-Douglas economies to

$$S^j_{\rm CD} = \left( p^* \cdot x^j \right) \left( e^{-D\left(\hat{\alpha} \| \hat{\omega}^j \right)} - 1 \right). \tag{A10}$$

The sum over agents $j$,

$$\sum_j S^j_{\rm CD} = \sum_j \left( p^* \cdot x^j \right) e^{-D\left(\hat{\alpha} \| \hat{\omega}^j \right)} - p^* \cdot w, \tag{A11}$$

is a sum of the exponentiated distance of each agent's wealth allocations at equilibrium prices from the equilibrium allocation, weighted by that agent's overall wealth. Its maximum, when all agents have bundles proportional to $\beta[p^*]$, is zero for all wealth distributions in the Pareto set.

The variation of the expenditure (A8) is

$$\begin{aligned} \frac{\partial}{\partial x^j_i} e\left[ p^*, \mathcal{U}^j\left[ x^j \right] \right] &= \frac{\hat{\alpha}_i}{x^j_i} e\left[ p^*, \mathcal{U}^j\left[ x^j \right] \right] \\ &= \left( \frac{p^* \cdot w}{p \cdot w} \right) \left( \sum_i \frac{\hat{\alpha}_i w_i}{x^j_i} \right) \left( \sum_i \frac{\hat{\alpha}_i x^j_i}{w_i} \right) e^{-D\left(\hat{\alpha} \| \hat{\omega}^j \right)} p_i . \end{aligned} \tag{A12}$$

If we normalize prices according to $p \cdot \beta[p^*] = (p \cdot w) \,/\, (p^* \cdot w) \equiv 1$, Eq. (A12) and some algebra give

$$\delta S^j_{\rm CD} = \left[ \sum_i \frac{\hat{\alpha}^2_i}{\hat{\omega}^j_i} e^{-D\left(\hat{\alpha} \| \hat{\omega}^j \right)} p - p^* \right] \cdot \delta x^j \tag{A13}$$

In the Pareto set the gradient of $S^j_{\rm CD}$ is $p - p^* = 0$, and otherwise it gives weight $> 1$ to the relative prices $\bar{p} \equiv p - p^*$, and a residual component proportional to $p^*$ from incomplete cancellation in the "numéraire" component.

From the sum of entropies (A11) we may define a natural and intuitive measure of the efficiency of trade. Suppose for convenience that initial endowments are distributed to maximize a failure of the double coincidence of wants in the economy, by giving each agent at most one good. The initial entropy (A11) of the economy is then $-p^* \cdot w$. The entropy of any Pareto optimum is zero, defining an idealized gain from trade of $p^* \cdot w$. For a trade algorithm leading to any distribution $\left\{ x^j \right\}$ in the demand space, the efficiency $\varepsilon \left[ \left\{ x^j \right\} \right]$ may be defined as the ratio of $\delta \sum_j S^j_{\rm CD}$ to the ideal gain $p^* \cdot w$. Its intuitive evaluation in terms of the Kullback-Leibler divergences and agent wealths is

$$\varepsilon \left[ \left\{ x^j \right\} \right] = \sum_j \left( \frac{p^* \cdot x^j}{p^* \cdot w} \right) e^{-D\left(\hat{\alpha} \| \hat{\omega}^j \right)}. \tag{A14}$$

All outcomes have efficiencies $\in [0, 1]$, with the maxima identifying the Pareto set. (More generally we would normalize the efficiency by the potential gains to trade from the whatever pre-trade allocation the agents have, preserving its range $\varepsilon \left[ \left\{ x^j \right\} \right] \in [0, 1]$ for all Pareto-improvements from the initial allocation.)

### 2. Helmholtz potentials for general preferences

The construction of Pareto-set contour utilities in Sec. VI B is straightforward for closed economies, and reservoir-stabilized trade can be included if the reservoir is represented as a large equilibrated agent within the economy. However, the purpose of idealizing a trading reservoir is to use the symmetries it imposes on prices to simplify the study of the rest of the economy. These symmetries are more transparent when the reservoir is made a boundary condition rather than an explicitly modeled agent, and the representation of utilities and equilibria in the economy changed accordingly. This simplification is implicit in the definition of the economic equivalents to thermodynamic potentials.

### 3. Contour utilities for general preferences

To perform the general construction, consider first the subspace of demand components for agent $j$, and suppress the superscript. With respect to this subspace, the formal condition that a parametrized contour $c[\lambda]$ intersect every indifference surface of a utility $u$ once, transversally, is

$$\left. \frac{\partial u}{\partial x} \right|_{x = c[\lambda]} \cdot \frac{dc}{d\lambda} \neq 0. \tag{A15}$$

Letting $x_c$ correspond to some point $c[\lambda]$, and writing $dx_c$ for the demand differential $(dc/d\lambda)\, d\lambda$, the direct contour utility on $c$, referenced to units of good $x_i$, is

$$\mu_{c,i}[x] \equiv u_{c,i}[u[x]], \tag{A16}$$

with $u_{c,i}[\mathcal{U}]$ given by Eq. (21). The relative price along $c$ in Eq. (21) is given by

$$\frac{p}{p_i} = \left. \frac{\partial u / \partial x}{\partial u / \partial x_i} \right|_{x_c}. \tag{A17}$$

A value of $\lambda$ is assigned to every point in the commodity space by $u$. Denoted $\lambda_u[x]$, it is the point on $c$ with the same utility as $x$:

$$\left. u[c[\lambda]] \right|_{\lambda = \lambda_u[x]} \equiv u[x]. \tag{A18}$$

Reducing notation, we write $c[x] \equiv c[\lambda_u[x]]$.

Then by transversality (A15), Eq. (A17) can be written

$$\left. \frac{p}{p_i} \right|_{c[\lambda]} \cdot \frac{dc}{d\lambda} = \frac{du/d\lambda}{\left. \partial u / \partial x_i \right|_{c[\lambda]}}, \tag{A19}$$

from which immediately

$$\begin{aligned} \frac{\partial \mu_{c,i}[x]}{\partial x} &= \frac{\left. \partial u / \partial x \right|_x}{\left. \partial u / \partial x_i \right|_{c[x]}} \\ &= \left( \frac{\left. \partial u / \partial x_i \right|_x}{\left. \partial u / \partial x^i \right|_{c[x]}} \right) \left. \frac{p}{p_i} \right|_x . \end{aligned} \tag{A20}$$

$\left. (p/p_i) \right|_x$ is the price relative to $p_i$ at $x$, related to the gradient of $\mu_{c,i}$ by an $x$- and $c$-dependent rescaling.

### 4. Edgeworth boxes with reservoirs

A reservoir for two goods defines a constraint

$$\delta \left(p_i x_i + p_{i'} x_{i'}\right) = 0 \qquad \text{(A21)}$$

on all trades in which it participates, with $p_{i'}/p_i|_{\text{Res}}$ independent of volume traded. As in the text, we take $x_{i'}$ to be numéraire, and $x_i$ to denominate entropy. $a \equiv x_{i'} + \left(p_i/p_{i'}\right) x_i$ is a balance-of-payments coordinate, not changed by reservoir trade, while some independent combination like $b \equiv x_{i'} - \left(p_i/p_{i'}\right) x_i$, measures the extent of trade with the reservoir. We transform from coordinates $(x_i, x_{i'}, \tilde{x})$ of the text to "diagonal" coordinates

$$x \equiv (a, b, \tilde{x}) \, . \qquad \text{(A22)}$$

The value of any bundle $x$ at prices $p$ decomposes as

$$p \cdot x \equiv p_{i'} a + \tilde{p} \cdot \tilde{x}. \qquad \text{(A23)}$$

Writing in the same basis $x_c \equiv (a_c, b_c, \tilde{x}_c)$, the contour utility at constant reservoir prices becomes

$$\mu_{c,i}\left[x\right] = \frac{p_{i'}}{p_i} \int_{\lambda_0}^{\lambda_u[x]} \left(da_c + \tilde{p}[x_c] \cdot d\tilde{x}_c\right) . \qquad \text{(A24)}$$

The most general case in which economic Helmholtz potentials with the properties of free energies can be defined on the whole Pareto set, with unrestricted utilities, is that for two agents (interacting with the reservoir), for which the Pareto set is one-dimensional. This case is readily visualized in a modification of the Edgeworth box, projected to the coordinates $(a, \tilde{x})$ conserved in interactions with the reservoir, and effective ordinal utilities on these. The projection works because, in coordinates $(a, b, \tilde{x})$, $u$ is still concave, but no longer monotonic in $b$. For any equilibrium supported at price $p_{i'}/p_i$, there is a $b_u[a, \tilde{x}]$ for which

$$\left.\frac{\partial u}{\partial b}\right|_{a, b_u[a,\tilde{x}], \tilde{x}} \equiv 0. \qquad \text{(A25)}$$

A *reservoir effective utility*

$$u_r[a, \tilde{x}] \equiv u[a, b_u\left[a, \tilde{x}\right], \tilde{x}] \qquad \text{(A26)}$$

then satisfies

$$\begin{aligned} \left.\frac{\partial u_r}{\partial a}\right|_{a,\tilde{x}} &= \left.\frac{\partial u}{\partial a}\right|_{a, b_u[a,\tilde{x}], \tilde{x}}, \\ \left.\frac{\partial u_r}{\partial \tilde{x}}\right|_{a,\tilde{x}} &= \left.\frac{\partial u}{\partial \tilde{x}}\right|_{a, b_u[a,\tilde{x}], \tilde{x}}. \end{aligned} \qquad \text{(A27)}$$

The Edgeworth box for two agent bundles $\left(a^1, \tilde{x}^1\right)$, $\left(a^2, \tilde{x}^2\right)$ in equilibrium with the reservoir, and effective utilities $u_r^1$, $u_r^2$, is shown in Fig. 9. Under arbitrary interaction of only those two agents and the reservoir, there is a fixed quantity of $a^1 + a^2 \equiv a^{\text{TOT}}$, and of $\tilde{x}^1 + \tilde{x}^2 \equiv \tilde{x}^{\text{TOT}}$.

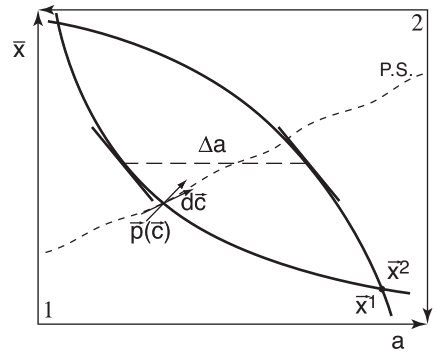


FIG. 9: The Edgeworth box for projected allocations $(a, \tilde{x})$, in the presence of reservoir lending. Indifference curves are of ${u_r}^1$, ${u_r}^2$, and the associated Pareto set (short-dashed) is labeled P.S. Maximal profit potential from initial allocations $\vec{x}^1$, $\vec{x}^2$ (shown in projection) is the long-dash segment labeled $\Delta a$.

In particular, the Pareto set for the three-component economy (the agents and reservoir) projects onto a single curve in $(a, \tilde{x})$. We may take segments of the curves inferior to indifference surfaces $u_r^1$ and $u_r^2$ in Fig. 9 to define the contours $c^1$ and $c^2$ respectively, evaluating $b^1$ and $b^2$ by Eq. (A25). The sum of contour utilities (A24) for any point shared by the two agents at common dual prices is a measure of the projection of the Pareto set onto $(a, \tilde{x})$. Since the sum $b_{u^1} + b_{u^2}$ at a Pareto allocation is not fixed in general, the pair of curves $\left(a_c, b_{u^1}\left[a_c, \tilde{x}_c\right], \tilde{x}_c\right)$, $\left(a_c, b_{u^2}\left[a_c, \tilde{x}_c\right], \tilde{x}_c\right)$ will be called a *generalized* Pareto set for this economy.

With this definition of contours $c^j$, the Helmholtz potential for each agent $j$ in Eq. (37) takes the form

$$\mathcal{A}^j\left[x^j\right] \equiv a^j - \left.\frac{p_i}{p_{i'}}\right|_{\text{Res}} \mu^j_{c,i}\left[x^j\right] . \qquad \text{(A28)}$$

We can now state and prove the following theorem:

**Theorem 1** *For ordinal $\mathcal{U}^1$ and $\mathcal{U}^2$ representing strictly convex, insatiable preferences, $\mathcal{A}^1\left[x^1\right] + \mathcal{A}^2\left[x^2\right]$ has the following properties:*

1. *$\mathcal{A}^1 + \mathcal{A}^2 = const. \; \forall \left(x^1, x^2\right)$ in the generalized Pareto set.*

2. *$\delta\left(\mathcal{A}^1 + \mathcal{A}^2\right) \leq 0$ for all voluntary trades involving only $x^1$, $x^2$, and the reservoir, with strict inequality for initial conditions not in the generalized Pareto set ($\delta$ denotes change in value).*

3. *The maximum reduction in $\mathcal{A}^1 + \mathcal{A}^2$ from voluntary trades, leaving the reservoir-untraded goods $\tilde{x}^1 + \tilde{x}^2 = \tilde{x}_{TOT}$, is the maximum profit an external trader can extract from the agent-reservoir system through voluntary trading.*

4. *This maximum extractable profit is strictly decreasing under any voluntary trades involving only $x^1$*

*and $x^2$ and the reservoir, and vanishes on the generalized trade set.*

**Proof:** By construction of the Edgeworth box, for any $\left(a^1, \tilde{x}^1\right), \left(a^2, \tilde{x}^2\right)$ in the Pareto set of $u^1_r, u^2_r$, $\mathcal{A}^1 + \mathcal{A}^2$ depends only on the endpoints of the two utility integration contours:

$$\begin{aligned} \mathcal{A}^1 + \mathcal{A}^2 &= a^{\rm TOT} - \int_{\lambda_0^1}^{\lambda_0^2} da_c + \tilde{p}[x_c] \cdot d\tilde{x}_c \\ &\equiv \mathcal{A}^{\rm TOT}_{CC}. \end{aligned} \tag{A29}$$

This proves point *1*.

For general $a^1 + a^2 \equiv a^{\rm TOT}$, $\tilde{x}^1 + \tilde{x}^2 = \tilde{x}^{\rm TOT}$, $\mathcal{A}^1 + \mathcal{A}^2$ differs only by the length of the segment of the Pareto set between the indifference surfaces $u^1_r\left[a^1, \tilde{x}^1\right]$ and $u^2_r\left[a^2, \tilde{x}^2\right]$ (a positive quantity by construction); hence

$$\mathcal{A}^1 + \mathcal{A}^2 = \mathcal{A}^{\rm TOT}_{CC} + \int_{\lambda_{u^1}\left[x^1\right]}^{\lambda_{u^2}\left[x^2\right]} da_c + \tilde{p}[x_c] \cdot d\tilde{x}_c. \tag{A30}$$

From Equations (A20) and (A30) it follows that, for any trades involving only the agents and reservoir,

$$\delta\left(\mathcal{A}^1 + \mathcal{A}^2\right) = -\frac{\delta u^1}{\partial u^1 / \partial x^1_{i'}\big|_{c^1(x^1)}} - \frac{\delta u^2}{\partial u^2 / \partial x^2_{i'}\big|_{c^2(x^2)}}. \tag{A31}$$

For voluntary trades $\delta u^1 \geq 0$, $\delta u^2 \geq 0$, with strict inequality of one utility for any transformations not in the generalized Pareto set. For convex preferences $x_{i'}$ can be chosen to represent a desirable commodity everywhere, so both $\partial u^j / \partial x^j_{i'}$ are positive, proving *2*.

For any trades involving an external agent, if the residual holdings $\tilde{x}^1 + \tilde{x}^2 \neq \tilde{x}^{\rm TOT}$, there is no mechanism to convert the bundle extracted to cash, using only the agents and reservoir. Therefore the maximum profit extractable is the

$$\Delta a = \max\left[a^{\rm TOT} - \left(a^1 + a^2\right)\right] \tag{A32}$$

that occurs on the initial indifference curves, where

$$\left.\frac{\partial \tilde{x}^1}{\partial a^1}\right|_{u^1_r} = \left.\frac{\partial \tilde{x}^2}{\partial a^2}\right|_{u^2_r} \tag{A33}$$

and $\tilde{x}^1 + \tilde{x}^2 = \tilde{x}^{\rm TOT}$. From Eq. (A27) it follows that the gradients of $u^1$ and $u^2$ are parallel, so that $\delta\left(\Delta a\right)$ is $\max\left[a^{\rm TOT} - \left(a^1 + a^2\right)\right]$ in the full configuration space, constrained by $\tilde{x}^{\rm TOT}$ and equilibrium with the exchange reservoir. Since motion along indifference curves preserves $\mu_{c,i}[x]$, the differential trades between the initial and final configurations satisfy

$$\delta\left(\mathcal{A}^1 + \mathcal{A}^2\right) = \delta\left(a^1 + a^2\right) \tag{A34}$$

giving *3*.

Finally, the expression for the change in $a^j$ in Eq. (A32), from any change in initial configuration, is

$$\delta a^j = \frac{\delta u^j_r}{\partial u^j_r / \partial a^j} - \tilde{p} \cdot \delta \tilde{x}^j, \tag{A35}$$

where $\tilde{p}$ is the same for both agents. Since voluntary trades increase either $u^1_r$ or $u^2_r$, and $\delta\tilde{x}^1 + \delta\tilde{x}^2 = 0$, it follows that $\Delta a \leq 0$ in Eq. (A32) for voluntary changes in initial conditions, with strict inequality when the initial state is not in the generalized Pareto set, giving *4*. QED.

The theorem demonstrates that functionally in systems of this type, $\mathcal{A}$ is the monetary equivalent of a free energy, while Eq. (A20) with the definition (A24) gives the relative scaling of the gradient of $\mathcal{A}$ with respect to the components of $\tilde{x}$ from properly normalized $\tilde{p}$ (no rescaling for $x$ in the generalized Pareto set, or at any $x$ for utility quasi-linear in $x_i$).

---


[1] Mas-Collel, A., M. D. Whinston, and J. R. Green , "Microeconomic Theory", Oxford Univ. Press, New York, 1995.

[2] Varian, H. R., "Microeconomic Analysis", third edition, Norton, New York, 1992, ch. 7 and ch. 10.

[3] Fermi, E., *Thermodynamics*, Dover, New York, 1956.

[4] L. Walras, Economique et Mécanique, *Bulletin de la Société Vaudoise de Sciences Naturelles* **45**, 1909, 313-325.

[5] Debreu, G., "Theory of Value", Yale Univ. Press, New Haven, CT, 1987

[6] Popper, K. R., "Conjectures and refutations; the growth of scientific knowledge" Basic Books, New York, 1962

[7] Mirowski, P., "More Heat than Light", Cambridge Univ. Press, Cambridge, England, 1989.

[8] Fisher, I., Mathematical Investigations in the Theory of Value and Prices (doctoral thesis), *Transactions of the Connecticut Academy* Vol. IX, July 1892

[9] Lisman, J. H. C., Econometrics, Statistics and Thermodynamics, The Netherlands Postal and Telecommunications Services, The Hague, Holland, 1949, Ch. IV.

[10] Bryant, J., A thermodynamic approach to economics, 36-50, Butterworth and Co., 1982.

[11] Saslow, W. M., An economic analogy to thermodynamics, Am. J. Phys. **67**, 1999, 1239-1247.

[12] Samuelson, P. A., "Structure of a Minimum Equilibrium System", (R. W. Pfouts, ed., *Essays in Economics and Econometrics: A Volume in Honor of Harold Hotelling*, University of North Carolina Press, 1960), reprinted in J. E. Stiglitz ed., *The Collected Scientific Papers of Paul*

*A. Samuelson*, M.I.T Press, Cambridge, Mass., 1966, 651-686
[13] Bródy, A., K. Martinas, and K. Sajo, An essay in macroeconomics, *Acta Oeconometrica* **35**, 1985, 337-343
[14] Candeal, Juan C., Juan. R. De Miguel, Esteban Induráin, and Ghanshyam B. Mehta, Utility and Entropy, *Economic Theory* **17**, 2001, 233-238.
[15] Fischer, K. H., and Hartz, J. A., and Hertz; *Spin glasses*, Cambridge U. Press, New York, 1991; Mezard, M., Parisi, G., and Virasoro, M. A., *Spin glass theory and beyond*, World Scientific, Singapore, 1987.
[16] Herrmann, H. J., Hovi, J.-P., and Luding, S., eds., *Physics of Dry Granular Media*, Kluwer, Dordrecht, 1997.
[17] Bhushan, B.; Israelachvili, J. N.; Landman, U., Nanotribology: friction, wear and lubrication at the atomic scale, *Nature* **374**, 1995, 607-616
[18] Jaynes, E. T., "Information theory and statistical mechanics", *Phys. Rev.* 106, 620–630, 108, 171–190, reprinted in R. D. Rosenkrantz, ed., *E. T. Jaynes: Papers on Probability, Statistics, and Statistical Physics*, Dordrecht: Reidel, 1989.
[19] Negishi, T., Welfare Economics and the Existence of an Equilibrium for a Competitive Economy. *Metroeconomica* **12** (1960), 92-97.
[20] Hahn, F., and T. Negishi, A Theorem on Nontatonnement Stability, *Econometrica* **30**, 1962, 463-469.
[21] Hicks, J. R., *Value and capital : an inquiry into some fundamental principles of economic theory*, Oxford: Clarendon, 1946.
[22] Gorman, T., Community preference fields, *Econometrica* **21**, 1953, 63-80.
[23] Cover, T. and Thomas, J., *Elements of Information Theory* New York: Wiley, 1991.

| | | | |
|---|---|---|---|
| commodities (other than $x_0$) | $\bar{x}$ | extensive state variables | $(E, V)$ |
| flow of $x_0$ in trade | $\delta x_0$ | entropy change from heat flow | $\delta\Sigma = -\frac{\delta Q}{T}$ |
| economic entropy | $S_{\mathrm{QL}}[\bar{x}] \equiv \bar{u}[\bar{x}] = \mathcal{U} - x_0$ | thermodynamic entropy | $S[E, V]$ |
| offer prices relative to $p_0$ | $\frac{\bar{p}}{p_0} \equiv \bar{u}'[\bar{x}]$ | intensive state variables | $\left(\frac{1}{T}, \frac{p}{T}\right) \equiv \left(\frac{\partial S}{\partial E}\Big\vert_V, \frac{\partial S}{\partial V}\Big\vert_E\right)$ |
| money-metric ($x_0$) value of trade | $\delta S_{\mathrm{QL}} = \delta\bar{x} \cdot \frac{\bar{p}}{p_0}$ | energy conservation see Eq. (44) | $\delta S = \frac{1}{T}\delta E + \frac{p}{T}\delta V$ |
| expenditure net of utility | $\frac{e[p,\mathcal{U}]}{p_0} - \mathcal{U} = \bar{x} \cdot \frac{\bar{p}}{p_0} - \bar{u}[\bar{x}]$ | Gibbs free energy (relative to $T$) | $\frac{1}{T}F \equiv \frac{1}{T}E + \frac{p}{T}V - S$ |

TABLE I: Correspondence of the quasi-linear economy to a classical thermodynamic system from Sections II and III. We have set $\bar{u}[x^*] = 0$ to simplify the presentation.