Thermodynamics of inequalities: 
from precariousness to economic stratification

Matteo Smerlak

Perimeter Institute for Theoretical Physics
31 Caroline Street North, N2L 2Y5 Waterloo ON, Canada

July 10, 2014

Abstract

Increasingly large social and economic inequalities are observed throughout the world. Theorists after Pareto have studied this phenomenon in terms of the tail structure of the wealth distribution at a given time. Unfortunately, this approach leaves unaddressed the dynamics of inequalities in non-equilibrium situations, e.g. under redistribution policies. Here we introduce a thermodynamical theory of inequalities based on the analogy between economic stratification and statistical entropy. Within this framework we identify the combination of upward mobility with precariousness as a fundamental driver of inequality. We formalize this statement by a “second-law” inequality displaying upward mobility and precariousness as thermodynamic conjugate variables. We estimate the time scale for the “relaxation” of the wealth distribution after a sudden change of the after-tax return on capital. Our method can be generalized to gain insight into the dynamics of inequalities in any Markovian model of socioeconomic interactions.

1 Introduction

All known human societies\(^1\) have displayed some level of economic inequality \(^2\). Yet this global imbalance is reaching alarming levels in the contemporary world: as of 2013, the 400 richest Americans have more wealth than the bottom half of all Americans combined. Indeed recent comprehensive research \(^3\) has showed that, while they have not reached the highs of the pre-1929 period, wealth inequalities in developed countries have steadily increased in the past decades. Understanding the origins and implications of these inequalities is an outstanding problem for economics, but also for society as a whole.

On the theory side, a well-established approach to this problem—pursued independently by economists \(^4\), mathematicians \(^5\), sociologists \(^7\) and physicists \(^8\) —consists in

\(^1\) Dating back to paleolithic hunter-gatherers \(^1\).
studying the equilibrium wealth distribution in stochastic models of individual (or household) income. Under general assumptions, one shows that additive income lead to exponential distributions, while multiplicative capital returns yield Pareto-like power law distributions [10–13]. These results are consistent with empirical data, both contemporary [14] and historical [15], which reveal a two-class structure with an exponential range at low wealth (where investment is negligible) and a power-law tail at high capital (where income is dominated by investment returns). Econophysicists [9] have pointed the striking similarity between this pattern and the Boltzmann-Gibbs distribution of statistical mechanics. Indeed both have the same “entropic” structure: there are many more ways to distribute a conserved quantity (be it wealth or energy) unequally than equally.

One much discussed consequence of such marked economic inequalities is the emergence of a super-elite class, the so-called “top 1%” [16], with disproportionate social, economical and political influence. But they also have more global effects, one of which is increased stratification [17]—the growth of the number of economically distinct “classes” in society. Indeed, as we will see below, “maximum entropy” wealth distributions are precisely those with the greatest stratification under global constraints on the mean wealth. This intriguing analogy between entropy and stratification points to a connection between the dynamics of inequalities and dissipation in thermal systems, extending beyond the limits of equilibrium statistical mechanics (to which it has been restricted so far).

In this paper we introduce a general framework, inspired from stochastic thermodynamics [18], to account for the dynamical origin of social inequalities. At its foundation is a general property of Markov processes known as the fluctuation theorem [368x410] (Appendix A). As we shall see, the great strength of this theorem lies in its explanatory power: given an entropy-increasing stochastic process, the fluctuation theorem elucidates the mechanism driving entropy production. In the context of social inequalities, where entropy quantifies inequality, we find that, over and above the multiplicative effect of capital return, precarious social mobility acts as a universal inequality-generating mechanism.

2 Results

2.1 Stratification index

We begin by formalizing our notion of stratification. Consider a population with (possibly nonstationary) wealth distribution $p_t(w)$, where $w \in [w_{\text{min}}, w_{\text{max}}]$ is a quantitative measure of social resource, call it wealth. Given $\delta w$ a reference wealth unit (e.g. the minimum yearly wage), we call economic stratum a segment of population with wealth in the range $[w_i, w_{i+1}]$ where $w_i = w_{\text{min}} + b \delta w$ for some number $b > 0$. We then define the stratification index $S_t$.

Originally discovered in the context of non-equilibrium statistical mechanics, this result has been successfully applied to models of evolutionary dynamics [19] and of biopoiesis [20]. More non-physics applications will likely come in the near future.
Figure 1: Comparison of the Gini and stratification indices for three familiar distributions: Pareto distributions with threshold $w_{\min}$ and tail index $\alpha$, normal distributions with mean $m$ and standard deviation $s$, and lognormal distributions with local parameter $\mu$ and scale parameter $\sigma$. Here we fix $w_{\min} = m = \mu = 1$ and $\delta w = 1$ and vary $\alpha$, $s$ and $\sigma$ respectively.

of the population at time $t$ by the expression

$$S_t = - \int p_t(w) \log p_t(w) \, dw - \log \delta w,$$

where log is the base $b$ logarithm. Mathematically, $S_t$ is the “differential entropy” of the wealth distribution $p_t(w)$. The stratification index is maximized by the uniform distribution on the interval $[w_{\min}, w_{\max}]$, in which case it simply measures the number of strata in the population. On the other hand, in a society where resources are uniformly distributed in a single stratum, $w_{\max} = w_{\min} + \delta w$, the stratification index is 0. The stratification index is an increasing function of the (more familiar) Gini index, see Fig. 1.

It is remarkable that both the Boltzmann (exponential) and Pareto (power-law) distributions\(^3\) which describe to a good accuracy the empirical wealth distributions in the lower and higher quantiles respectively, arise as maximum stratification distributions. Indeed, the

\(^3\)The Boltzmann distribution at “inverse temperature” $\beta$ is $p_B(w) = \beta e^{-\beta w}$, with stratification index $1 - \log(\beta \delta w)$. The Pareto distribution within minimum wealth $w_{\min}$ and Pareto index $\alpha$ is $p_P(w) = \alpha w_{\min}^\alpha / w^{\alpha + 1}$, with stratification index $1 + 1/\alpha + \log(w_{\min}/\alpha \delta w)$. 

3
former corresponds to the maximum of $S$ under the constraint $\langle w \rangle = 1/\beta$, while the latter corresponds to the maximum of $S$ under the constraint $\langle \log(w/w_{\text{min}}) \rangle = 1/\alpha$. Here and below $\langle \cdot \rangle$ denotes expectation value. (One can check that $S$ is a monotonically decreasing function of $\beta$ and $\alpha$ respectively.) In other words, the lower (resp. higher) segments of society appear to be maximally stratified given a fixed mean additive (resp. multiplicative) wealth—the “poor” and “rich” segments of society are close to being at statistical equilibrium given their constraints. This observation begs a question: how is this social equilibrium reached?

2.2 A toy model

To begin investigating this question, consider the following toy model of society. Assume a finite set of “casts” $k = 1, 2, \ldots, K$, and suppose that at each time step $t$ there is a probability $\pi_+$ (resp. $\pi_-$) for each individual to move up (resp. down) one cast. Suppose that society starts off in an initial cast distribution $p_0(k)$. The probability to stay in the lowest class $k = 1$ is $1 - \pi_+$, the probability to stay in the highest class is $1 - \pi_-$. The log-ratio $O = \log(\pi_+ / \pi_-)$, which we may call opportunity, measures the tendency of individuals to climb up the social ladder; reciprocally its opposite, precariouslyness $R = \log(\pi_- / \pi_+) = -O$, measures their tendency to go down the social ladder.

In this discrete setting we take for the stratification index the Shannon entropy $S_t = -\sum_k p_t(k) \log p_t(k)$, i.e. the expected value of the surprisal $s_t(k) = -\log p_t(k)$. The fluctuation theorem for Markov chains (Appendix A.2) then states that, after any given number of time steps,

1. The probability distribution of $\Delta s - P \Delta k$ satisfies

$$\langle b^{-\Delta s + P \Delta k} \rangle = 1.$$  \hspace{1cm} (2)

This identity implies that $\Delta s - P \Delta k$ is exponentially unlikely to be negative, in the sense that for any positive number $r$,

$$\text{Prob}[\Delta s - P \Delta k \leq -r] \leq b^{-r}. \hspace{1cm} (3)$$

2. As a consequence of this identity, the variation of stratification $\Delta S = \langle \Delta s \rangle$ during the process is constrained by the second law inequality

$$\Delta S \geq P \langle \Delta k \rangle. \hspace{1cm} (4)$$

In other words, in this toy society,

$$\text{(stratification increase)} \geq \text{(precariousness)} \times \text{(upward mobility)}. \hspace{1cm} (5)$$

\footnote{This notion of equilibrium, which involves no other variable than wealth, should not be confused with other notions of economic equilibrium, such as Nash equilibrium or Pareto optimality.}

\footnote{The surprisal literally measures the “surprise” of an agent who, with no prior information, finds herself within the class $k$ at time $n$.}
Figure 2: Numerical verification of the fluctuation theorem in a toy society with \( K = 10 \) casts, for two different initial distributions \( p_0(k) \). In the top row, almost all individuals start in the lowest cast (\( p_0(k) = 91\% \) if \( k = 1 \) and \( p_0(k) = 1\% \) else). In the bottom row, the initial population is uniformly distributed over the \( K \) casts (\( p_0(k) = 1/K \)). Here \( \pi_+ = 30\% \), \( \pi_- = 50\% \), \( b = 2 \), and the statistics were computed over \( 10^4 \) realizations of the process.

Suppose for instance that all individuals started off in the lowest cast \( k = 1 \), so that \( S_0 = 0 \), and that \( \pi_- < \pi_+ \), so that downward social evolution is more likely than upward social evolution. Then the second law indicates that, as the mean cast level \( \langle k \rangle \) grows, so does the stratification index, at a rate greater than \( P \) per cast level.

### 2.3 Stochastic wealth model: “second law” inequality

Armed with this basic intuition, let us now proceed to a more realistic model of stochastic wealth growth. Let \( w_t \) be the household detrended\(^6\) wealth at time \( t \), and assume a stochastic dynamics of the form

\[
dw_t = ld dt + w_t \cdot (\rho dt + \sigma dB_t).
\]

---

\(^6\)The detrended wealth is the absolute wealth times e\(^{-gt}\) where \( g \) is the economic growth rate.
The first term describe “additive” income (labor), while the second term represents “multiplicative” income (capital returns). In the notations of [21], the mean return rate \( \rho \) is given by \( \rho = \bar{r} - g - c \), where \( \bar{r} \), \( g \) and \( c \) represent the after-tax return, growth and consumption rates respectively. We assume that the return shocks \( dB_t \) form a standard Brownian motion (Wiener process) and we use the Ito convention for stochastic differentials.\(^7\) A dictionary between the terms used in this section and more standard physics terminology is provided in Appendix D.

Stochastic wealth models such as (6) has been considered by many authors, see [22] and references therein. In particular, stochastic equations of the form (6) arise in the “random-agent” approximation of certain agent-based models [23]. In this setting one shows [8] that the stationary distribution—an inverse Gamma distribution—has a Pareto tail with exponent \( \alpha = 1 - 2\rho/\sigma^2 \), which decreases when \( \bar{r} - g \) increases [21]. (When \( \rho \geq \sigma^2/2 \), the model does not have an equilibrium distribution.)

Here we are interested in the non-equilibrium dynamics of stratification. As in the discrete case, this problem can be investigated using the fluctuation theorem for diffusion processes [24–26], see Appendix A.3. This gives the exact analogue of relations (2) and (3), with the entropy “source” term \( P \Delta k \) replaced by the Stratonovich stochastic integral

\[
\int_0^T P(w_t) \circ dw_t = \int_0^T P(w_t) \cdot dw_t - \frac{1}{2} \int_0^T P'(w_t) dt,
\]

where we defined the precariousness function

\[
P(w) = \frac{2}{\ln b} \left( \frac{\sigma^2 - \rho}{\sigma^2 w} - \frac{l}{\sigma^2 w^2} \right).\]

In particular, the second law inequality now reads\(^8\)

\[
\frac{dS_t}{dt} \geq \int P(w) j_t(w) dw \tag{9}
\]

where \( j_t(w) \) is the “social mobility flux”, i.e. the expected enrichment rate of a household with wealth \( w \).\(^9\) The inequality (9) becomes an identity in the limit of small flux \( j_t(w) \), i.e. close to statistical equilibrium.

\(^7\)Note that the special case \( l = 0 \) reduces to the geometric Brownian motion widely used in quantitative finance. In this context, our results can be interpreted as putting a lower bound on the uncertainty on an asset price in terms of its drift and volatility.

\(^8\)In this case the stratification rate can be computed exactly, as

\[
\frac{dS_t}{dt} = \int P(w) j_t(w) dw + \frac{2}{\ln b} \int \frac{j_t(w)^2}{\sigma^2 w^2 p_t(w)} dw.
\]

In thermodynamic language, the first term is the (reversible) “entropy flux” and the second term the (irreversible) “entropy production”.

\(^9\)The wealth flux is given by \( j_t(w) = (l + \rho w)p_t(w) - \sigma^2 \partial_w(w^2 p_t(w))/2 \), and satisfies the continuity equation \( \partial_k p_t(w) + \partial_w j_t(w) = 0 \).
This result sheds an interesting new light on the relationship between economic conditions and social inequalities. First, when \( \rho \leq \sigma^2 \), the precariousness function \( P(w) \) changes sign at the threshold wealth

\[
  w_* = \frac{l}{\sigma^2 - \rho}. \tag{10}
\]

This threshold has a simple interpretation: enrichment \( j_t(w) > 0 \) at \( w < w_* \) decreases the stratification index, while enrichment \( j_t(w) > 0 \) at \( w > w_* \) increases it. This result formalizes the intuitive notion that enriching the poor reduces inequalities, while enriching the rich increases inequalities.

Second, all other things being equal, this critical wealth \( w_* \) is an increasing function of \( \bar{r} - g \). This means that a larger capital return rate allows growth to have an inequality-alleviating effect on more quantiles of the population. (This is of course assuming that these quantiles actually participate in multiplicative investments—which of course is not true of the poorer strata of society). Note also that \( w_* \), like the Pareto exponent \( \alpha \), is a decreasing function of the volatility \( \sigma \). Thus, like in our toy model, the growth of inequalities can be interpreted as a consequence of risk [13].

Third, the precariousness function \( P(w) \), which we saw controls the dynamics of stratification out of wealth equilibrium, turns out to be directly related to the equilibrium wealth distribution \( p_*(w) \). Indeed, the latter is simply given by

\[
p_{eq}(w) \propto b^{-V(w)} \tag{11}
\]

where

\[
V(w) = \int_{0}^{w} P(w')dw'. \tag{12}
\]
From this perspective, the relevant “conserved quantity” in the economy is V(w), and not wealth w itself (as proposed in [9] but criticized in [27]). Identifying such a conserved quantity (in physics parlance, a “potential” function) provides useful intuition for the stochastic dynamics (6): roughly speaking, each household tries to reach the minimum of V(w) (the zero-precariousness threshold value $w = w_*$), but is constantly driven away from that value by stratification-maximizing “fluctuations”. In this sense, the potential V(w) can be thought of as the mathematical expression of a Smithian “invisible hand” driving macroeconomic evolution [28].

2.4 Stochastic wealth model: relaxation time

An important question which is readily addressed in this nonequilibrium framework is that of the relaxation time of the economy [11]. Suppose that, starting from the equilibrium wealth distribution for the parameters $(l, \rho, \sigma)$ and Pareto tail exponent $\alpha = 1 - 2\rho/\sigma^2$, the detrended effective return rate $\rho$ suddenly changes to the value $\rho' = \rho + \delta\rho$ (e.g. because $\bar{r} - g$ changes according to new fiscal policies): how long will it take for the economy to reach the new statistical equilibrium with Pareto tail exponent $\alpha' = 1 - 2\rho'/\sigma^2$? A straightforward computation using the formalism above allows us to estimate the relaxation time (to first

---

**Figure 4:** Left: five sample paths $w_t$ illustrating the main qualitative feature of the stochastic wealth model (6): most paths converge to $w \simeq w_*$ and remain there forever, but some paths make wild excursions at large wealth (the “top 1%” tail of the Pareto distribution). Right: histogram of net gains $\Delta w$ after $T = 100$ and verification of the integral fluctuation relation over $10^4$ paths ($\Sigma$ denotes the stochastic integral (7)). In both plots the initial wealth is normally distributed about $w = 1$ (with standard deviation $\sigma$).  

10. Indeed, in the model (6) the equilibrium expected wealth is infinite whenever $\rho > 0$, i.e. when $\bar{r} - g > c$.  

11. I thank Thomas Piketty for suggesting this problem to me.
Figure 5: Left: the relaxation time $\tau$ as a function of $\alpha = 1 - 2\rho/\sigma^2$. Notice the strong dependence on the volatility $\sigma$ (inset: zoom on the $1.5 \leq \alpha \leq 2$ region). Right: Stratification index as a function of time, starting from the equilibrium distribution for six different values of $\delta \rho$ (the vertical line represents $t = \tau$). Here $\rho = -2.5\%$, $\sigma^2 = 10\%$ and $l = 1$ per period.

order in $\delta \rho$ as (Appendix C)

$$\tau \simeq \frac{2\psi_1(\alpha)(1 + \alpha) - 2}{\sigma^2}, \quad (13)$$

where $\psi_1$ is the trigamma function. Note that, at this order, the relaxation time $\tau$ is independent of $\delta \rho$: larger changes $\delta \rho$ lead to larger changes of the stratification index $\delta S$, but these changes always occur on the same time scale $\tau$. The latter is also independent on $l$, implying that increasing labor income does not affect the dynamics of inequalities. Finally, $\tau$ increases with $\bar{r} - g$ and is very sensitive to the volatility $\sigma$, see Fig. 5.

3 Conclusion

Using the fluctuation theorem for Markov processes as a guide, we have identified precarious mobility as a key driver of inequalities away from equilibrium. We have illustrated this idea in simple stochastic wealth models, where we obtained lower bounds on the growth of stratification over time as well as estimates of the corresponding time scales. But the scope of our approach is broader, and can be generalized to other stochastic models of socioeconomic dynamics: in Appendix B we apply the fluctuation theorem to a model of (biased) trade and find an upper bound on the so-called Theil inequality index.

Our results are complementary to earlier findings which showed that high interest rates (for given growth rate) generate fat-tailed, Pareto-like equilibrium distributions. In particular, they throw light on the question: how will the economic system “respond” to a perturbation? We saw for instance that there exists a threshold wealth $w^*$ beyond which enrichment generates more inequalities. Such knowledge provides a clear guideline regarding the quantiles which should be targeted by redistribution policies ($w < w^*$), and the ones
which should not \((w > w_s)\). In the current crisis times, we believe that developing further a “response theory” of the economic system is a pressing challenge for political economy.

**Acknowledgements**

I was introduced to exchange models of wealth distribution by Tom J. Carter during the Santa Fe Institute CSSS13 summer school. Helpful conversations with Ahmed Youssef are gratefully acknowledged.

**A The fluctuation theorem**

**A.1 General idea**

The (integral) fluctuation theorem is a general property of Markov stochastic processes. It states that there exists a function \(\Sigma\) of stochastic paths such that

\[
\langle b^{-\Delta s-\Sigma} \rangle = 1, \quad (14)
\]

where \(\Delta s\) is difference between the initial and final surprisal. (Recall that if \(X_t\) is the state at time \(t\) and \(p_t(X)\) is the corresponding probability density, the surprisal is defined by \(s_t = -\log p_t(X_t); \) hence \(\Delta s = \log[p_0(X_0)/p_T(X_T)]\)). The identity (14) implies that paths such that \(\Delta s - \Sigma < 0\) are exponentially unlikely, in the sense that if \(r\) is a positive number,

\[
\text{Prob}(\Delta s - \Sigma \leq -r) \leq b^{-r}. \quad (15)
\]

Furthermore, (14) implies (by convexity of the exponential function) that

\[
\langle \Delta s \rangle \geq \langle \Sigma \rangle. \quad (16)
\]

The left-hand side is nothing but the entropy production \(\Delta S = S_T - S_0\), where \(S_t\) is the entropy of the distribution \(p_t(X_t)\). Thus, the path-dependent function \(\Sigma\) can be interpreted as a stochastic entropy source, and (14) as a refinement of the second law inequality (16).

The proof of (14) is based on the idea of time-reversal. For any stochastic path \(X = (X_t)_t\) with initial distribution \(p_0(X_0)\), consider the time-reversed path \(X^\dagger = (X_{T-t})_t\). Denote \(\mathbb{P}[X]\) the probability of a path \(X\) with initial distribution \(p_0(X_0)\), and \(\mathbb{P}^\dagger[X]\) the probability of a path with initial distribution \(p_0^\dagger(X_0) = p_T(X_0)\), where \(p_T\) is the time-evolution of \(p_0\). Next define the path-dependent function \(R[X]\) by

\[
R[X] = \frac{\mathbb{P}[X]}{\mathbb{P}^\dagger[X^\dagger]} \quad (17)
\]

Then formally

\[
\sum_X \mathbb{P}[X] b^{-R[X]} = \sum_X \mathbb{P}^\dagger[X^\dagger] = 1. \quad (18)
\]
Defining $\Sigma$ by
\[ R = \Delta s - \Sigma \] (19)
immediately gives (14). To gain useful information about entropy production in a given Markov process, it therefore suffices to compute explicitly the log-ratio $R$.

Equation (15) an immediate consequence of (14) \[29\]:
\[
\text{Prob}(\Delta s - \Sigma \leq -r) = \int_{-\infty}^{-r} \text{Prob}(\Delta s - \Sigma = q)\ dq
\] (20)
\[
\leq \int_{-\infty}^{-r} \text{Prob}(\Delta s - \Sigma = q)\ b^{-q-r}\ dq
\] (21)
\[
\leq 2^{-r} \int_{-\infty}^{+\infty} \text{Prob}(\Delta s - \Sigma = q)\ dq
\] (22)
\[
\leq b^{-r}.
\] (23)

We outline below the computation of $R = \Delta s - \Sigma$ for discrete states (sec. A.2) and diffusions (A.3) in the stationary case; the nonstationary case can be treated along the exact same lines.

A.2 Markov chains

Let us begin by considering a discrete-space, discrete-time Markov chain. Let $\gamma_{ij}$ be the transition probability between states $i$ and $j$, $p_0(i)$ the initial probability distribution, and $p_T(i_T)$ the final probability distribution after $T$ time steps. Then the probability of a path $(i_0, i_1, \cdots, i_N)$ is given by
\[
\mathbb{P}(i_0, i_1, \cdots, i_N) = p_0(i_0) \prod_{k=0}^{N-1} \gamma_{i_k i_{k+1}}
\] (24)
and the probability of the reverse path $(i_N, i_{N-1}, \cdots, i_0)$ with initial distribution $p_0^\dagger(i_N) = p_T(i_N)$ is
\[
\mathbb{P}^\dagger(i_N, i_{N-1}, \cdots, i_0) = p_T(i_T) \prod_{k=0}^{N-1} \gamma_{i_k+1 i_k}.
\] (25)

Hence
\[
R(i_0, i_1, \cdots, i_N) = \log \frac{p_0(i_0)}{p_T(i_T)} + \sum_{k=0}^{N-1} \log \frac{\gamma_{i_k+1 i_k}}{\gamma_{i_k i_{k+1}}}.
\] (26)

The first term is the surprisal difference $\Delta s$, and the second term defines the entropy source as
\[
\Sigma = \sum_{k=0}^{N-1} \log \frac{\gamma_{i_k+1 i_k}}{\gamma_{i_k i_{k+1}}}.
\] (27)

\[\text{The two are related by the matrix equation } p_T = \Gamma^T p_0, \text{ where } \Gamma \text{ is the matrix with entries } \gamma_{ji}.\]
The structure of Σ provides a clear-cut explanation for the origin of entropy production: on average, entropy grows when the system makes transitions \( i_k \rightarrow i_{k+1} \) which are disfavored with respect to the reverse transitions \( i_{k+1} \rightarrow i_k \).

This result immediately generalizes to continuous-time Markov chains. In that case, \( \gamma_{ij} \) are transition rates rather than probabilities and the probability of a path \((i_0, i_1, \cdots, i_N)\) with transition times \((t_1, \cdots, t_N)\) is given by

\[
P(i_0, i_1, \cdots, i_N; t_1, \cdots, t_N) = p_0(i_0) \prod_{k=0}^{N-1} e^{-\lambda_k(t_{k+1}-t_k)} \gamma_{i_k i_{k+1}}. \tag{28}
\]

where \( \lambda_k = \sum_j \gamma_{i_k j} \). The log-ratio \( R \) is unchanged, and \( \Sigma \) is still given by (27).

### A.3 Diffusion processes

Consider now an Ito diffusion process

\[
dX_t = c(X_t) \, dt + d(X_t) \cdot dB_t \tag{29}
\]

with initial probability density \( p_0(X_0) \) and path measure \( d\mathbb{P}[X] \). Here \( dB_t \) is a standard Wiener process. Introduce the “time-reversed” process \( X_t^\dagger = X_{T-t} \), with equation

\[
dX_t^\dagger = -c(X_t^\dagger) \, dt + d(X_t^\dagger) \cdot dB_t. \tag{30}
\]

and initial probability \( p_0^\dagger(X_0^\dagger) = p_T(X_0^\dagger) \), where \( p_T \) is the time-evolution of \( p_0 \).\(^{13}\) Denote \( d\mathbb{P}^\dagger[X^\dagger] \) its path measure. Then, by the Girsanov theorem, the Radon-Nikodym derivative of \( d\mathbb{P} \) with respect to \( d\mathbb{P}^\dagger \) satisfies \(^{21}\)

\[
\ln \frac{d\mathbb{P}[X]}{d\mathbb{P}^\dagger[X^\dagger]} = \Delta s + 2 \int_0^T \left( \frac{c(X_t) - d(X_t)d'(X_t)}{d^2(X_t)} \right) \cdot dX_t \tag{31}
\]

where the Stratonovich integral is defined by

\[
\int_0^T f(X_t) \cdot dX_t = \int_0^T f(X_t) \cdot dX_t - \frac{1}{2} \int_0^T f'(X_t) \, dt. \tag{32}
\]

The entropy source function is thus given in this case by

\[
\Sigma = -\frac{2}{\ln 2} \int_0^T \left( \frac{c(X_t) - d(X_t)d'(X_t)}{d^2(X_t)} \right) \cdot dX_t. \tag{33}
\]

\(^{13}\)The final distribution \( p_T \) is the solution of the forward Kolmogorov equation with initial condition \( p_0 \).
B Inequalities from biased trade

B.1 Theil index

The entropic definition of the stratification index is reminiscent of the Theil inequality index. The two metrics, however, are conceptually different: unlike stratification, the Theil index involves $N$ different agents $n$ with wealth $w_n^t$ at time $t$; it is defined by

$$T_t = \sum_{n=1}^{N} \phi_t^n \log \phi_t^n + \log N$$  \hspace{1cm} (34)

where $\phi_t^n = w_t^n / \sum_{n=1}^{N} w_t^n$ is the fraction of the total wealth held by agent $n$. Formally, the Theil index is the difference between the maximal and observed Shannon entropy of the distribution of wealth fractions. Remarkably, the fluctuation theorem provides insight also in the dynamics of the Theil index, albeit in a dual way with respect to stratification: as we now show, the fluctuation theorem provides an upper bound on the growth of $T_t$.

B.2 An exchange model

Suppose that $N$ agents trade their wealth among themselves, following the rule that $n$ transfers to $m$ an amount (proportional to his wealth $w_t^n$) at the rate $J_{nm}$. The exchange rates $J_{nm}$ need not be symmetric; we define the trade bias from $n$ to $m$ by $b_{nm} = \log(J_{nm}/J_{mn})$. The wealth fractions are governed by the rate equations

$$\frac{d\phi_t^n}{dt} = \sum_{m \neq n} (J_{mn}\phi_t^m - J_{nm}\phi_t^n).$$  \hspace{1cm} (35)

Importantly, these equations can be interpreted as the master equations of a continuous-time Markov chain $n_t$. In this interpretation, wealth can be thought of as a token that randomly changes hands among the agents; $\phi_t^n$ is then the probability to find the token with agent $n$ at time $t$. Applying the fluctuation theorem to this chain, we obtain the statistical constraint (2), now with $-B = -\sum_t b_{nm_{t+1}}$ as a source term (playing the role of $P\Delta k$ in sec. 2.2). The second law inequality takes the form

$$\Delta T \leq \langle B \rangle,$$  \hspace{1cm} (36)

i.e. the increase of the Theil index is smaller than the expected cumulated bias along the chain. In particular, if all exchanges are unbiased, we obtain the intuitive result that inequalities must decrease over time.

To illustrate this result, consider a trading society with just three agents $f_1, f_2$ and $u$ ($f$ standing for “fair” and $u$ for “unfair”). Assume that $J_{f_1f_2} = J_{f_2f_2} = J_0$, $J_{fu} = J_{fu} = J_+$ and $J_{uf_1} = J_{uf_2} = J_-$ with $b = \log(J_+/J_-) > 0$. Then the second law (36) implies that, while the presence of $u$ clearly creates inequalities, the Theil index cannot grow more than $b$ times the mean number of unreciprocated exchanges from $f_{1,2}$ to $u$. 

13
Figure 6: Equilibrium wealth distribution $p_{eq}(w) \propto b^{-V(w)}$ of the stochastic model (6) for various values of $\rho = \bar{r} - g - c$, corresponding to different Pareto tail exponent $\alpha = 1 - 2\rho/\sigma^2$. Here $l = 1$ and $\sigma^2 = 10\%$ per period.

C Computation of the relaxation time (13)

The equilibrium distribution $p_{eq}(w)$ for the stochastic model (6) is the inverse gamma distribution
\[
p_{eq}(w) = \frac{\beta^\alpha e^{-\beta/w}}{\Gamma(\alpha) w^{\alpha+1}}
\]  
where $\alpha = 1 - 2\rho/\sigma^2$ and $\beta = 2l/\sigma^2$. Its differential entropy is given by
\[
-\int_0^\infty p_{eq}(w) \log p_{eq}(w) dw = \frac{\alpha + \ln(2l/\sigma^2) + \ln \Gamma(\alpha) - (1 + \alpha)\psi(\alpha)}{\ln b},
\]
where $\Gamma$ and $\psi$ are the gamma and digamma functions. Now, suppose that at time $t = 0$, the effective return rate $\rho = \bar{r} - g - c$ changes to a different value $\rho' = \rho + \delta \rho$. We can estimate the time before the the wealth distribution settles to a new equilibrium with tail exponent $\alpha' = 1 - 2\rho'/\sigma^2$ as
\[
\tau \simeq \left( \frac{\partial S_{eq}}{\partial \rho} \delta \rho \right) / \left( \frac{dS_t}{dt} \right) \bigg|_{t=0}.
\]
From (38) we compute
\[
\frac{\partial S_{eq}}{\partial \rho} = \frac{-2}{\sigma^2} \left( \frac{1 - \psi_1(\alpha)(1 + \alpha)}{\ln b} \right).
\]
Next we estimate the denominator of (39) using
\[
\left. \frac{dS_t}{dt} \right|_{t=0} \simeq \int_0^\infty P(w) j_0(w) dw.
\]
where the initial flux $j_0(w)$ is given by

$$j_0(w) = (l + \rho' w)p_{eq}(w) - \frac{\partial}{\partial w} \left( \frac{\sigma^2 w^2 p_{eq}}{2} \right)$$

(42)

$$= (l + \rho' w)p_{eq}(w) - (l + \rho w)p_{eq}(w)$$

(43)

$$= (\delta \rho w)p_{eq}(w).$$

(44)

This gives

$$\frac{dS_t}{dt} \bigg|_{t=0} \simeq \delta \rho \int_0^\infty w P(w) p_{eq}(w) dw.$$  

(45)

The integral on the right-hand side can be evaluated explicited using (8) and (37), yielding

$$\int_0^\infty w P(w) p_{eq}(w) dw = \frac{1}{\ln b}. $$

(46)

Plugging (40) and (46) into (39) gives (13).

\section{Dictionary}

Here we provide for the reader’s convenience a dictionary relating the concepts used in this paper, notably in sec. 2.3, to their physics counterparts.

| Economics          | Physics                                    |
|--------------------|--------------------------------------------|
| Stratification index | Gibbs entropy                              |
| Precariousness      | (Driving force)/(temperature)              |
| Mobility flux       | Probability current                        |
| Potential          | Potential energy                           |
| Surprisal          | Stochastic or local entropy                |

\section{References}

[1] H. Pringle, “The ancient roots of the 1%,” Science 344 no. 6186, (May, 2014) 822–825.

[2] T. Piketty and E. Saez, “Inequality in the long run,” Science 344 no. 6186, (May, 2014) 838–843.

[3] T. Piketty, Capital in the Twenty-First Century. Harvard University Press, Mar., 2014.

[4] D. G. Champernowne, “A Model of Income Distribution,” Econ. J. 63 no. 250, (June, 1953) 318.
[5] B. Mandelbrot, “The Pareto-Levy Law and the Distribution of Income,”
    *International Economic Review* **1** no. 2, (May, 1960) 79.

[6] H. Kesten, “Random difference equations and Renewal theory for products of random
    matrices,” *Acta Math.* **131** no. 1, (Dec., 1973) 207–248.

[7] J. Angle, “The Surplus Theory of Social Stratification and the Size Distribution of
    Personal Wealth,” *Soc. Forces* **65** no. 2, (Dec., 1986) 293–326.

[8] J.-P. Bouchaud and M. Mézard, “Wealth condensation in a simple model of
    economy,” *Physica A* **282** no. 3-4, (July, 2000) 536–545.

[9] A. Dragulescu and V. M. Yakovenko, “Statistical mechanics of money,”
    *Eur. Phys. J. B* **17** no. 4, (Oct., 2000) 723–729.

[10] M. Nirei, “Pareto Distributions in Economic Growth Models,” Tech. Rep. 09-05, 2009.

[11] *Econophysics of wealth distributions*. Springer Milan, Milano, 2005.

[12] V. M. Yakovenko, “Colloquium: Statistical mechanics of money, wealth, and income,”
    *Rev. Mod. Phys.* **81** no. 4, (Dec., 2009) 1703–1725.

[13] “The Distribution of Wealth and Fiscal Policy in Economics With Finitely Lived
    Agents,” *Econometrica* **79** no. 1, (2011) 123–157.

[14] A. Drăgulescu and V. M. Yakovenko, “Exponential and power-law probability
    distributions of wealth and income in the United Kingdom and the United States,”
    *Physica A* **299** no. 1-2, (Oct., 2001) 213–221.

[15] A. Y. Abul-Magd, “Wealth distribution in an ancient Egyptian society,”
    *Phys. Rev. E* **66** no. 5, (Nov., 2002) 057104.

[16] A. Atkinson, T. Piketty, and E. Saez, “Top Incomes in the Long Run of History,”
    tech. rep., National Bureau of Economic Research, Cambridge, MA, Cambridge, MA,
    Oct., 2009.

[17] D. B. Grusky, *Social Stratification*. Class, Race, and Gender in Sociological
    Perspective. Westview Press, 2001.

[18] U. Seifert, “Stochastic thermodynamics, fluctuation theorems, and molecular
    machines,” *Rep. Prog. Phys.* **75** (2012) 126001.

[19] V. Mustonen and M. Lassig, “Fitness flux and ubiquity of adaptive evolution,”
    *Proc. Natl. Acad. Sci. USA* **107** no. 9, (Mar., 2010) 4248–4253.

[20] J. L. England, “Statistical physics of self-replication,” *J. Chem. Phys.* **139** no. 12,
    (2013) 121923.
[21] T. Piketty and G. Zucman, “Wealth and Inheritance in the Long Run,” in *Handbook of Income Distribution*. North Holland.

[22] X. Gabaix, “Power Laws in Economics and Finance,” *Ann. Rev. Econ.* 1 no. 1, (2009) 255–294.

[23] B. M. Boghosian, “Kinetics of wealth and the Pareto law,” *Phys. Rev. E* 89 no. 4, (Apr., 2014) 042804.

[24] J. L. Lebowitz and H. Spohn, “A Gallavotti-Cohen-Type Symmetry in the Large Deviation Functional for Stochastic Dynamics,” *J. Stat. Phys.* 95 no. 1/2, (1999) 333.

[25] C. Maes, “On the origin and the use of fluctuation relations for the entropy,” *Séminaire Poincaré* 2 (2003) 29–62.

[26] U. Seifert, “Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem,” *Phys. Rev. Lett.* 95 no. 4, (2005) 4.

[27] B. Hayes, “Computing Science: Follow the Money,” *American Scientist* (2002) 400–405.

[28] M. H. Cohen and I. I. Eliazar, “Econophysical visualization of Adam Smith’s invisible hand,” *Physica A* 392 no. 4, (Feb., 2013) 813–823.

[29] C. Jarzynski, “Nonequilibrium work relations: foundations and applications,” *Eur. Phys. J. B* 64 no. 3-4, (July, 2008) 331–340.