Daniel Kraft

Political Power and Socio-Economic Inequality

An Application of the Canonical Ensemble to Social Sciences

Masterarbeit

zur Erlangung des akademischen Grades
Master of Science
an der Naturwissenschaftlichen Fakultät
der Karl-Franzens-Universität Graz

Betreuer:

Univ.-Prof. Dr.phil. Wilhelm Schappacher

Institut für Mathematik und Wissenschaftliches Rechnen

Graz, am 24. Juni 2012

Abstract

I propose a model for socio-economic interaction of individuals that is based on their positions in a social space described by income, amount of contributed labour and individual power. The individuals are coupled to each other by economic principles and an abstract "energy", which corresponds roughly to the total happiness—however, it also incorporates the distribution of political power, since society does not maximise the real total happiness, but instead those making political decisions have the ability to maximise only their own. Based on this energy, the canonical ensemble of statistical mechanics is employed to produce a thermodynamic model describing the behaviour of society and its economics. This model is then analysed both theoretically and numerically with a Monte-Carlo method. The coupling of political power to social and economic status in the energy leads to the very interesting effect that this society tends to polarise itself into one privileged and very rich individual and all the others, who have to work hard in order to support the riches of the privileged one. Depending on the thermodynamic parameters, the change from an equal society with permutation symmetry between the individuals intact to this unequal state where the symmetry has been broken spontaneously takes place as a first-order phase transition. Thus my model is capable of describing and explaining the very real effect of social inequality by means of modelling the distribution of political power in addition to that of wealth.

Kurzzusammenfassung

Ich stelle ein Modell für die sozioökonomische Interaktion von Individuen einer Gesellschaft auf, das auf deren Positionen in einem sozialen Raum basiert. Dieser Raum beschreibt das Einkommen, die dafür eingebrachte Arbeitsleistung und den individuellen Anteil an politischer Macht. Durch volkswirtschaftliche Bedingungen und eine abstrakt definierte "Energie" werden die Individuen dann zueinander gekoppelt. Diese Energie misst die einzelnen Zufriedenheiten, die allerdings noch mit der jeweiligen Macht gewichtet werden — dadurch wird dem Umstand Rechnung getragen, dass üblicherweise nicht jeder gleichberechtigt an politischen Entscheidungen teilnimmt und deshalb für diese nicht gleich "wichtig" ist. Auf die Energie wird anschließend das kanonische Ensemble der statistischen Mechanik angewendet, um die Gesellschaft im thermodynamischen Gleichgewicht zu beschreiben. Dieses Modell wird dann sowohl theoretisch als auch numerisch mit einer Monte-Carlo Simulation untersucht. Die Kopplung zwischen Macht und sozialem Status über die Energie führt dazu, dass sich die Gesellschaft in einen Zustand "totaler Ungleichheit" polarisiert, in dem ein einzelnes Individuum fast alle Macht und alle Güter besitzt, während der Rest fast nichts bekommt. Abhängig von den thermodynamischen Zustandsvariablen ergibt sich in meinem Modell ein Phasenübergang erster Ordnung, wenn diese Trennung einsetzt und damit die Permutationssymmetrie zwischen den einzelnen Individuen spontan gebrochen wird. Das Modell ist somit in der Lage, den auch in der Realität beobachteten Effekt von sozialer und wirtschaftlicher Ungleichheit zu beschreiben und dadurch zu erklären, dass zusätzlich zur Verteilung von Kapital auch die von politischer Macht betrachtet wird.

Contents

1 Introduction

In the study of economic and social systems, *inequality between individuals* with respect to their economic and social situation is an omnipresent and very "real" phenomenon, as will be discussed with examples on page 12. In this work, I want to present a mathematical model that is based only on the rather simple definition in (8) and fundamental economic, physical and thermodynamic principles. Nevertheless, it is already capable of showing this effect of social unfairness. This is basically due to the fact that my fundamental principles introduce a coupling between economic and social position of an individual to its political power, so that inequality of the wealth distribution directly influences and is influenced by inequality in the distribution of power in a society—and both effects amplify each other. In the end, Theorem 8 and Theorem 9 show that the "optimal" distribution in such a society is attained when one individual is fully privileged and all others have no power and basically only work to support this single person. Furthermore, if thermodynamics is applied to the system, one can use two variables, which correspond to an abstract temperature and a measure of fairness in the political system, to characterise its states. In this picture, we will see in Subsection 4.3.3 that the system exhibits a first-order phase transition where one phase corresponds to perfect equality between every individual, and in one phase, permutation symmetry between the individuals is broken and a single one becomes privileged. While I do not know whether or not these insights can be applied in any way to help make real societies fairer (for instance, by tweaking these variables such that the "equal" phase is reached), they are certainly very interesting discoveries.

Since it is not a very far-stretched thought to apply thermodynamics to societies with lots of individuals or "actors", I was not the first to try that. See for instance the works by Jürgen Mimkes, in particular [29], [30] and also [31], or generally the topics of *econophysics* and *sociophysics*. When I initially devised the model described below, I was not aware of this fact, though—nevertheless, I believe that a short comparison of my model to those existing ones is in order. While Mimkes separates modelling of economic and social systems[1], I tried to integrate some aspects of both subjects in my model. It also seems to me that he "recycles" classical thermodynamics much more than I do, which means a more or less direct translation of its principles and terms into economics. [29, 1.2 and 1.3] apply the first and second fundamental law of thermodynamics to economics, and [29, 1.6.1] describes how production can be interpreted as Carnot process. For a homogeneous society—and in my model, individuals are always assumed to be homogeneous—, he predicts in [30, 1.3.1] the three phases *ordered*, *chaotic* and *global*, corresponding to the classical physical states *solid*, *liquid* and *gaseous*. They form much the same way as in physics when the temperature is replaced by average income and the pressure by the birth-rate according to [31][2]. In contrast, my model results in only two different phases. His main focus seems to be attraction and repulsion between individuals belonging to different groups [30, 1.2.3], or phrased differently, the wins and losses those individuals have when they are "next to" each other, meaning that they interact socially in some way [30, 1.4]. This is another aspect where he focuses on classical physical ideas—my social space is instead not related to any kind of "distance" between the individuals but represents in a more abstract way their social and economic position, including political power. As far as I have seen, there's nowhere a notion of individual power to be found among his models, while it takes a key role in mine. The interaction between individuals in my model is not described in terms of pair-interaction, but instead more abstractly by the total energy alone. Mimkes applies a lot of physical concepts and tools developed for classical thermodynamics of molecules to societies, while my model is more directly tailored to the application I have in mind. In addition to the specifically defined constraints[3] in Definition 3 and the energy in (8), it only uses the very basic and abstract framework of statistical mechanics and the canonical ensemble. This finally is the point where both approaches touch—while Mimkes does not mention the canonical ensemble and uses a slightly different mathematical starting point, he also results in the Boltzmann distribution [29, 1.4.5]. In total, I believe that I

[1] See for instance [30, 1.3.1] for one of the many places where a distinction is made between "social systems", "political systems" and "economic systems".

[2] The other and more scientific literature I considered did not give much details about his view of homogeneous societies at all, unfortunately.

[3] Mimkes also uses the term "constraints", but in his interpretation [30, 1.1.3], the constraints are not absolute, meaning that temperature allows a system to violate the constraints. In other words, his constraints are what the energy is to a physical system—it tries to minimise it, but it is not guaranteed that this is always perfectly the case. In contrast, I mean *strict* constraints which have to be fulfilled at all times, and temperature as well as maximisation of total happiness is an *additional* concept.

can well claim to have something independent of the existing models, in particular Mimkes' socio- and econophysics as discussed above, so that it indeed makes sense to study it in detail although there's already existing work based on the idea of thermodynamics used for economics and sociology.

Finally, I want to describe briefly the structure of this document: There are three main parts, in Chapter 2 the model is motivated and introduced, and then I already draw some very fundamental conclusions about it. The most important pieces in the puzzle are Definition 3, (8), Definition 6 as well as Theorem 1 and Theorem 2. The main theoretical part is Chapter 3, where I analyse the "ground state" of the model, respectively its behaviour for zero temperature, which is characterised by minimising the energy of (8) over all valid states. There one already finds that it must necessarily show social inequality, as shown by the already mentioned Theorem 8 and Theorem 9. This then leads to existence of a *unique* minimum-energy point in Theorem 10 in addition to Theorem 1, and they also allow to simplify the minimisation problem to the four-dimensional one in (22), for which further nice properties can be derived and which can also be solved numerically rather easily. Particularly interesting is also Theorem 12, which is a general result not specific to my model about continuous dependence of minimisers on the minimised function. I have not found a similar result so far in the literature, thus I proved it myself so that I can later apply it to get Theorem 13. Chapter 4 describes how to handle the case of non-zero temperature, where the Boltzmann distribution from Definition 6 has to be taken into account. Section 4.2 describes a computational method that can be used for simulation, and the results I obtained with it are presented in Section 4.3. I consider in particular Figure 5 and Figure 7 to be very interesting, especially with the phase transition described in Subsection 4.3.3 in mind. Finally, in the end I give a short conclusion and outlook over open issues which would be interesting for further research.

2 The Model

In this chapter, I will describe and justify my model, as well as comparing it a little to common literature in sociology and economy. This will be the more intuitive part of this document, while the more technical analysis of the consequences and properties of this model will be discussed in the following chapters. In statistical physics, the basic idea is that macroscopic phenomena like gas dynamics or solid state properties stem from the interaction of lots of small particles. Those can be real molecules, or more abstract things that themselves "model" a more complicated real physical entity[4]. While the fundamental laws are most of the time really simple, complex results can arise when lots of particles interact with each other. In a similar manner, society is made up of lots of individuals that interact socially. In statistical physics, one is ultimately interested in the macroscopic behaviour of a system (like the pressure or temperature of a gas), but the fundamental model is often based on microscopic states (the positions and momenta of molecules for a classical gas, or the quantum numbers of occupied states for a quantum mechanical gas). In political economy, the goal is also to describe the economic system on a high level and model society as a whole, while its possible "states" are in my model described by properties of the individuals that make it up, see Section 2.1. Only in a second step the individuals are combined to interact with each other, which in my case is done both by certain conditions the whole society is forced to obey, as well as by an energy functional, which is like a "soft" guiding force that does not ultimately forbid certain conditions, but prefers some over others. The former will lead to a set of possible conditions not for the individuals but the whole society in Section 2.2, and the latter will then describe which of those states are likely and which aren't—in other words, mathematically we'll give a probability distribution on all allowed states in Section 2.4 that describes which states may be realised by the system. This is based on the framework of the canonical ensemble in statistical mechanics [13, ch. 22, p. 183ff], with an energy functional that will be introduced in Section 2.3.

2.1 Individuals in the Social Space

Since the individuals and their positions in the space of possible states are at the core of my model, the choice of properties or *dimensions* I want to consider is very important. I my case, I'm interested mainly in the interaction between labour input and living conditions as well as social status of the people. This leads to a model that describes economic output, or GDP[5], but does not take into account the financial or labour markets of classical macroeconomics[6]. Thus the basic quantities that characterise individuals are—as stated above—*labour* the individual contributes to society, *income* it gets as reward, and a measure of *power* it has in the society. While the first two are, of course, classical components in any economic system, the last makes things ultimately very interesting and produces new effects: Namely the interaction between economy and social status (measured by income, for instance) with political decisions, which are made by the individuals together based on the individual powers. In more mathematical terms, power can simply be understood as a *weight* each individual has in society—this will be manifested ultimately in (8). Thus, my *social space* is three-dimensional:

Definition 1. Let $a_{\max}, l_{\min} > 0$ be given parameters. Then, the set of possible states for each individual is $U = [0,1] \times [0, a_{\max}] \times [l_{\min}, \infty)$. In the following, for $x \in U$ the notation $x = (m, a, l) = (m, p)$ will be adopted.

Now assume that assume (m, a, l) describes an individual as per Definition 1. Then:

m describes its *power*. The valid range $[0, 1]$ comes from the fact that m will later be interpreted as fraction of the total power, see Definition 3 and in particular (6).

a measures the *labour* the individual contributes. The bounds $[0, a_{\max}]$ are natural in that an individual may work from nothing up to some maximally possible amount, which is physically limiting.

[4] An example for this are the spins in the Ising model that may represent atoms with magnetic momenta from certain valence electrons, see [38, 7.2.2, p. 180ff]. This is clearly already a very crude abstraction.

[5] See [5, p. 22ff] or [37, gross national product, p. 202]. Since my model does not have any notion of money or a financial market, it is more exactly always *real* GDP [5, p. 24f] in the following.

[6] As is done, for instance, in [5, ch. 4, 6].

l similarly stands for the *income* of our individual. Again, the bounds $[l_{\min}, \infty)$ come from the intuitive facts that some minimal income is required to survive (which is usually called *subsistence wage* [37, subsistence, p. 421] in economics), while the income is not limited from above[7].

So if we have some values for the two parameters *maximal labour* $a_{\max}$ and *minimal income* $l_{\min}$, the set $U \subset \mathbb{R}^3$ from Definition 1 is the individuals' social space in my model. Note that the concept of social space is not new, and in fact several authors in the literature of sociology have defined such a space in more or less precise terms and with differing coordinates before. For instance, Pierre Bourdieu [16, p. 494ff] assumes that the social position of some individual is defined by what he calls its *economic*, *cultural* as well as *social capital*. Economic capital is clearly related to my l, and social capital bears some resemblance to m, but in total, Bourdieu's social space is clearly different from mine. As another example, Reinhard Kreckel defines four (primary) dimensions in his social space [22, p. 75]. Interestingly, Kreckel also introduces the notion of social status or prestige [22, p. 87ff] which has certain similarities to my m—but he considers it to be "secondary", meaning that the social space actually only consists of the four primary dimensions and prestige is a quantity derived from them rather than an independent degree of freedom on its own, as m is in my model.

Now that individuals themselves have defined states, the next step is to consider their individual *behaviour* depending on where they are in social space. Of course, each one makes his or her decisions based on the principle of *personal happiness* (see the philosophy of John Stuart Mill [24, p. 183] as one example), which in turn depends on the social status or position in social space. In lots of economic or game-theoretic models, individuals try to maximise some personal *utility function* [37, utility, p. 457]. Note that it can be very difficult in practice to find a good such function, and in fact it may even be questionable whether one can meaningfully quantify the "utility" or "happiness" an individual feels at all! See [40, ch. 4] for a treatment of utility in basic microeconomic modelling, including critical review of the concept of cardinal utility. Nevertheless, also in my model the individuals have some kind of utility function—but it will turn out that both the theoretic predictions of Chapter 3 and also the qualitative behaviour in general (see Subsection 4.3.5) do not depend strongly on the concrete form of it. Since the basic principle will later in Section 2.3 be the *minimisation* of an energy functional, my utility is however turned around and the actual function describes "unhappiness" or *strain* on the individuals. Since it may be difficult to explicitly define such a function, I leave that open to as much a degree as possible, and just require for the moment the existence of a strain function with some rather basic properties:

Definition 2. f is a proper strain function if it maps $[0, a_{\max}] \times [l_{\min}, \infty) \to \mathbb{R} \cup \{\infty\}$ and has the following properties:

- f is twice continuously differentiable[8] in the interior of its domain and continuous up to the boundary (see below for what that means with infinite values).
- For each $l \in [l_{\min}, \infty)$, $f(\cdot, l)$ is strictly increasing. Similarly, $f(a, \cdot)$ is strictly decreasing for each $a \in [0, a_{\max}]$. If $f(a, l) = \infty$, obviously non-strict behaviour is sufficient.
- $f(a, l) = \infty$ is only allowed if $x = (m, a, l) \in \partial U$[9] for all $m \in [0, 1]$. Furthermore, in this case I require that
$$\forall M > 0 : \exists \epsilon > 0 : \forall x' \in U \cap B(x, \epsilon)\,{}^{10} : f(a', l') > M$$
holds for all those x. This describes what we mean with continuity up to the boundary in the case of infinite values.

In the following, where not stated otherwise, f will always denote such a strain function.

I believe that the requirements of Definition 2 are easy to justify without losing too much generality. As already mentioned in the footnote, smoothness and continuity of f do not restrict too much. The

[7] It is limited in practice of course by the economic power or GDP of the whole economy, but that will come into play later when considering the *whole society*, see (7) which stems from (5). From the *individual's point of view*, income is unlimited.

[8] This is the maximum regularity required later, and may not be strictly necessary everywhere—but I don't think it is important to generalise every result to the absolutely minimal regularity. Especially since in practice f has to be estimated approximately anyway, and can be chosen smooth without any problems.

[9] ∂A denotes the boundary of a set A.

[10] $B(c, r)$ is the open ball around $c \in \mathbb{R}^n$ with radius $r > 0$. It is not important which norm is used for that purpose.

assumptions on strict monotonicity are also intuitive, since strain on an individual should always increase with more labour and decrease with higher income, the other factor fixed.[11] This is also one of the basic assumptions about "well-behaved preferences" in [40, p. 45]. Finally, the possibility of infinite values of f is only allowed to model that an individual feels "infinite strain" when the labour nears $a_{\max}$ or the income $l_{\min}$ as some kind of barrier, although Definition 1 of course also enforces those inequality restrictions already.

In addition to the requirements of Definition 2, we will often (but not always) assume that f is a (strictly) *convex* function. This is maybe not so intuitive as monotonicity is, but still a common assumption also in microeconomics [40, p. 46ff]. There, this property is often called *decreasing marginal utility* or *law of diminishing utility* [37, marginal utility, p. 274f] and emerges also in the principle of a *diminishing marginal rate of substitution* [40, 3.8, p. 51f]. In my context it means that, for instance, if an individual is already working very close to the physical limit, additional labour puts much more strain to it than if it works near nothing at the moment. Or in other words, people prefer situations more in the centre of social space over extreme conditions. For instance, it is preferred to work "modestly" and have "modest" income over working very hard with lots of money or working nothing but having nothing, either. Also, if two persons divide their labour and income fairly between each other, this corresponds to forming a convex combination of arguments of f. Then convexity means that in the resulting point, their strain (which is the functional value of f) is below the average strain before cooperation (on the secant)—so together they are able to improve their situation. That also seems like a plausible assumption, given that without advantages from cooperation, there would not be any society or economy to consider at all!

Lemma 1. *For all $l' > l_{min}$, f is bounded from below when restricted to a domain where the income is limited to l', which is $[0, a_{max}] \times [l_{min}, l']$.*

Proof. For (a, l) in the restricted domain, it follows from the monotonicity properties of f in Definition 2 that $f(a, l) \geq f(0, l) \geq f(0, l')$. Also, since there exist points (a, l) not on the boundary with finite strain values, $f(0, l') < \infty$. □

To conclude this section, I will now give examples for strain functions that satisfy Definition 2 and will be used later for example calculations. I think that in particular Example 1 is important, since it is strictly convex and may be[12] realistic to a certain degree. This is also the function that I'll mainly use below (mostly in Chapter 4) for numerical calculations. Example 2 and Example 3 are mostly included here because they show interesting behaviour and will be investigated later. For all those examples, I will always use $a_{\max} = 1$ and $L = l_{\min} \in (0, 1)$—see Section 2.5 below for the reason.

Example 1. Let $L \in (0, 1)$ be given and define

$$f(a, l) = e^a - \log l. \tag{1}$$

It is easy to verify that (1) defines a proper strain function as per Definition 2, f is everywhere finite on its domain (as opposed to the possibility of infinite values on the boundary) and that it is strictly convex in addition.[13] This function shows *logarithmic* increase of happiness for increasing income l, which seems plausible to me—if someone is very wealthy, the additional appreciation of more money falls rapidly (but in this case not so fast as it does in Example 3). Figure 1a shows the contour lines of this function, which are in economics usually referred to as *indifference curves* [40, 3.3, p. 36f].

Example 2. For a different example, let's set

$$f(a, l) = e^{a-l} - \log l. \tag{2}$$

Then this defines also a valid and strictly convex strain function. The indifference curves for (2) are shown in Figure 1b, and while (2) looks quite similar to (1), the qualitative behaviour of this function can be very

[11] Note that while the monotonicity must be strict, we do not require that the partial derivatives of f are bounded away from zero or something like that—thus it is well possible, and may be realistic, that the gain from more income or less labour becomes very small; it has to always stay strictly positive, though.

[12] In my opinion; of course, I have no empirical evidence for or against this hypothesis. But the basic theory as laid out here works independently of the concrete form of f, so that does not matter much.

[13] For strict convexity, note that the Hessian is $\nabla^2 f(a, l) = \begin{pmatrix} e^a & 0 \\ 0 & \frac{1}{l^2} \end{pmatrix}$, which is obviously symmetric positive definite.

different—see Example 6 below. The difference between both strain functions is that in (2), the income couples directly back to the strain that labour creates. High income makes the individuals tolerant and almost indifferent against more labour; to a certain degree this may be true also in practice, but probably not so much that l above some threshold makes a almost completely unimportant. Nevertheless, this f from (2) satisfies all formal criterions at least.

Example 3. Yet another strain function, which is also strictly convex and everywhere finite (even without restriction on the arguments) is

$$f(a, l) = e^{a^2 - l}. \tag{3}$$

This function will be used below in Example 7, and has the particular property that the merit of increasing income goes to zero even faster than before and in contrast to (1) or (2), it is bounded from below, even when the income becomes arbitrarily large. I don't think that this is realistic behaviour, as in practice, people seem to try to amass as much wealth as possible—and actual differences between rich and poor are much more striking than what this strain function gives, see Example 7 and compare l_1 for this strain function in Table 1 with the results for the others. The indifference curves of (3) are plotted in Figure 1c.

Example 4. Finally, I want to add a fourth example. In particular, I want to complement Example 1 by an alternative which was not constructed particularly to be a counter-example, but which could also be realistic:

$$f(a, l) = -\log(1 - a) - \log l \tag{4}$$

In this case, we have log-barriers to enforce $l > 0$ and $a < 1$ even when $L = 0$ and so we can disregard those two inequality constraints. The indifference curves are shown in Figure 1d. Once again, it can be easily seen via the Hessian that also f from (4) is a proper strain function and strictly convex.

2.2 Possible States of Society

The next step in building up the model is putting together the individuals of Section 2.1, so that they form a society. We will always consider the number of individuals in society fixed[14], and call this number $N \in \mathbb{N}$. Then, the society is *mainly* a collection (or tuple in mathematical terms) of N individuals (or points in social space U). But of course, in addition to that, we need some kind of "glue" between them, so that they are not merely isolated individuals, but rather cooperating with each other to form a large economy. I will only consider a *closed* economy [5, p. 48], and thus will require that *total demand equals total production* [5, p. 51]. Expressed in terms of my social space, this means that the total income of all individuals is proportional to the total labour, and the constant of proportionality (which measures labour productivity) is a model parameter. This fact is mathematically expressed in (5) below. In addition, as hinted already above in Definition 1, the power m_n of some individual measures the *fraction* of political power that individual has. That is to say, to what proportion it has influence on decisions the whole society makes affecting everyone—like law-making, decisions on taxation or the like. Thus, all powers have to be normalised in some way, which justifies (6). Reformulating the above in a more mathematical way, we arrive at the following definition for valid states of the full economy:

Definition 3. Let $\alpha > 0$ be the labour productivity. Then the set of possible states our society of $N \in \mathbb{N}$ individuals can be in is:

$$\Omega = \big\{X = (x_1, \ldots, x_N) \in U^N \mid (5) \text{ and } (6) \text{ hold}\big\}$$

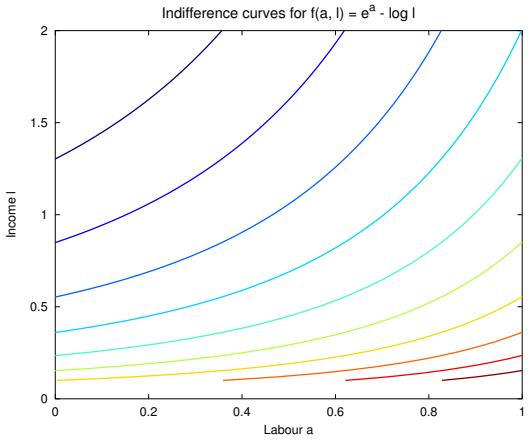
The constraints are

$$\sum_{n=1}^{N} l_n = \alpha \sum_{n=1}^{N} a_n \tag{5}$$

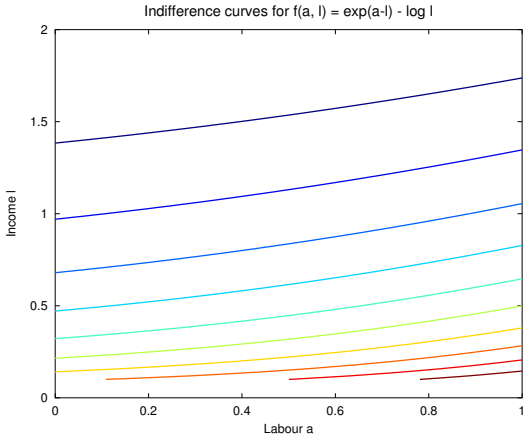
to enforce a closed economy and

$$\sum_{n=1}^{N} m_n = 1 \tag{6}$$

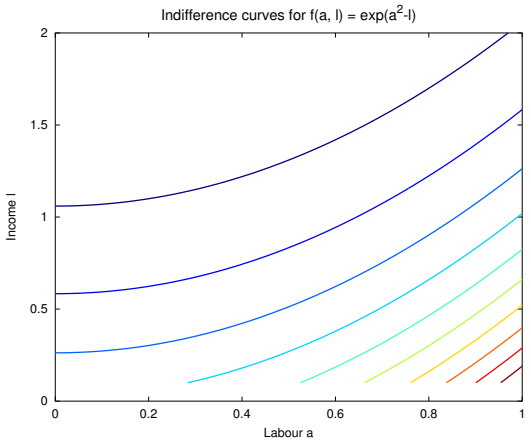
[14] See also the discussion in Section 2.4 on page 16 about the grand canonical ensemble.



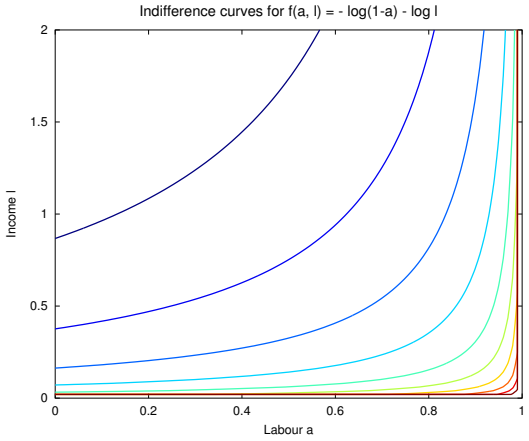
(a) Example 1, $L = 0.1$.



(b) Example 2, $L = 0.1$.



(c) Example 3, $L = 0.1$.



(d) Example 4, $L = 0.0$.

Figure 1: Indifference curves (in other words, contour lines) of the strain functions from the examples.

to normalise political powers. Where not stated otherwise, the notation $X = (x_1, \ldots, x_N)$ will be used for states and their individuals.

Lemma 2. *$\Omega \subset \mathbb{R}^{3N}$ is convex, compact and for each individual x of a valid state $X \in \Omega$, the income l is bounded:*

$$l_{min} \leq l \leq \alpha N a_{max} \tag{7}$$

Proof. Let's begin with convexity. Let $X^{(1)}, X^{(2)} \in \Omega$ and $\lambda \in (0,1)$. Define $X' = \lambda X^{(1)} + (1-\lambda)X^{(2)}$. We have to show $X' \in \Omega$. It is clear that for $n = 1, \ldots, N$, $m'_n \in [0,1]$, $a'_n \in [0, a_{\max}]$ and $l'_n \geq l_{\min}$ if these inequalities hold for $X^{(1)}$ and $X^{(2)}$ themselves.[15] It remains to verify the equality constraints (5) and (6):

$$\sum_{n=1}^{N} l'_n = \lambda \sum_{n=1}^{N} l_n^{(1)} + (1-\lambda) \sum_{n=1}^{N} l_n^{(2)} = \alpha \left(\lambda \sum_{n=1}^{N} a_n^{(1)} + (1-\lambda) \sum_{n=1}^{N} a_n^{(2)} \right) = \alpha \sum_{n=1}^{N} a'_n$$

This shows (5), similarly one can also verify (6) for X'.

For compactness, note that $U \subset \mathbb{R}^3$ is closed by Definition 1. Thus, also $U^N \subset \mathbb{R}^{3N}$ is closed. Define

$$h : U^N \to \mathbb{R}^2, \;\; h(m_1, a_1, l_1, \ldots, m_N, a_N, l_N) = \begin{pmatrix} \sum_{n=1}^{N} m_n - 1 \\ \sum_{n=1}^{N} l_n - \alpha \sum_{n=1}^{N} a_n \end{pmatrix}.$$

Then, $\Omega = h^{-1}(\{(0,0)\})$[16] and because h is clearly continuous, Ω is relatively closed in U^N. Furthermore, as U^N is closed itself, $\Omega \subset \mathbb{R}^{3N}$ is a closed set. To show that Ω is compact, it suffices to show that it is bounded since we work in finite dimensions. Let $X \in \Omega$ be given, and pick some individual x_n of this state, $n = 1, \ldots, N$. Then obviously $|a_n| \leq a_{\max}$ and $|m_n| \leq 1$. But note also that l_n is bounded by (7) since:

$$|l_n| = l_n \leq \sum_{n=1}^{N} l_n = \alpha \sum_{n=1}^{N} a_n \leq \alpha N a_{\max}$$

So Ω is closed and bounded, and thus compact. □

Lemma 3. *f is always bounded from below when considering individuals from a valid state in Ω.*

Proof. This follows immediately from Lemma 1 in combination with (7). □

2.3 The Principle of Minimal Strain

Above in Definition 3, we derived a set of states the economy can be in based on my social space from Definition 1 in combination with constraints (5) and (6) to enforce "valid" and "realistic" outcomes. However, it is not yet clear which of those states the system may *actually* occupy. It is clear that in a real economy, not all of them are equally probable—for instance, the individuals may prefer to work more but have more income or the other way round, depending on the concrete form of the strain function f. Even more, in real societies throughout history and up to the present, there seems to be the tendency to prefer *unequal* states over equal ones—that is, some part of the population (usually the majority) has a worse social position than some other part (a governing minority), in the sense that the majority works much for little income, while the minority enjoys the production not consumed by the majority without working itself a lot. Of course, this is a very simplified view—actual political systems are much more complicated. Still, this is one of the dominating impressions when considering societies in history. But nevertheless, it is clear that *social* and *economic inequality* are omnipresent in human interaction. Since this is also one of the main motivations for my present work and in particular for (8) below, I want to give some examples from history and also literature (which is of course influenced by and portraits real conditions):

[15] In other words, U is convex.

[16] For a function $h : A \to B$ and $C \subset B$, I'll use the notation $h^{-1}(C) = \{a \in A \mid h(a) \in C\}$ to denote the *preimage* of C under the mapping h.

Roman Empire In ancient Rome, there was not even equality by law between all people, for instance with the distinction between Roman citizens and all others which had by far fewer rights, or even between free people and slaves. Furthermore, the political system was very timocratic in nature: There was a strong coupling between economic and political powers. See [41, ch. 3], for instance pages 172–174 and 258.

Middle Ages In the middle ages, a difference between *aristocracy* and *common people*, founded upon ownership of land, was present in nearly all European societies. Also during that time, feudalism with dependencies between different social classes was common. The usual form of government in European countries was monarchy. All that of course implies social inequalities. See [39, ch. 6] for a reference, especially pages 98 and 99.

19^{th} Century This time was characterised by the advent of industrial work and already the distinction between workers and capitalists, now founded upon the ownership of capital and economic wealth. Aristocrats and peasants began to lose their importance in favour of the urban middle class and industrialism, but aristocrats still had some political influence and in particular were largely over-represented in European parliaments. See [11], pages 45 and 47–48.

Europe 20^{th} Century After World War I, the aristocracy finally lost all remaining power in countries like Russia or Austria. But still even in the 20^{th} century and after both World Wars, the majority of wealth stemmed from inheritance of land. For instance, in Spain in 1962, more than half of the total land was part of just 1.8 % of all agricultural businesses. See [12], pages 55–59.

Soviet Union In the Soviet Union, private ownership of means of production was (clearly) no longer important. But nevertheless, there were distinctions between different classes (like workers in agriculture or industry). Materially by far the best position had upper leaders in politics and the party, a class referred to as "Nomenklatura". See [12], pages 658–661. Also according to [37, communism, p. 86f] the standards of income varied considerably in communist societies like Soviet Russia or China.

Present Day But not just in the past, but also the present day large social and economic inequalities remain even in western and developed countries. For instance, according to [36, p. 2] the recovery from the economic crisis in the USA is very unequal:

> In 2010, average real income per family grew by 2.3 % but the gains were very uneven. Top 1 % incomes grew by 11.6 % while bottom 99 % incomes grew only by 0.2 %. Hence, the top 1 % captured 93 % of the income gains in the first year of recovery. Such an uneven recovery can help explain the recent public demonstrations against inequality. It is likely that this uneven recovery has continued in 2011 as the stock market has continued to recover. National Accounts statistics show that corporate profits and dividends distributed have grown strongly in 2011 while wage and salary accruals have only grown only modestly. Unemployment and non-employment have remained high in 2011.

The author most probably hinted at the "Occupy Wall Street" protests [20], which are another impressive demonstration of social inequality in the USA and all over the world.

Literature One class of literature I want to mention here is that of *Utopias* (see also [37, utopia, p. 457]), like [32] or [4]. This can be interpreted as human desire for an equal society, given the class differences at the times of writing of those books. On the other side are dystopic descriptions like George Orwell's [34] or [35] ("all animals are equal but some animals are more equal than others"). The former is a projection of socialism into the future (retaining and even extending class and power differences), and the second a fable depicting the Soviet Union. In all those cases, it is evident that the authors were inspired by social inequality in their times and surroundings.

Now, let's come back to the description of my model. There, the basic principle which determines the actual states of the society is *minimisation of the total strain of the population*. For this purpose, the total strain is defined as follows:

Definition 4. Let $\gamma \in [0,1]$ be given. Then the *total strain* or *energy* of a configuration is given by:

$$\mathcal{H} : \Omega \to \mathbb{R} \cup \{\infty\}, \quad \mathcal{H}(X) = \sum_{n=1}^{N} \left(\frac{\gamma}{N} + (1-\gamma) m_n \right) f(a_n, l_n) \tag{8}$$

We adopt the convention that $0 \cdot \infty = 0$ in the cases where $m = 0$ gets multiplied with $f(a,l) = \infty$, which is possible for $\gamma = 0$.

This notion and in particular (8) are at the very heart of my model considerations. Because of that, I think it is necessary to give some remarks about this definition:

- The term *energy* as used synonymously for *total strain* in Definition 4 and elsewhere comes from the fact that below, I will interpret $\mathcal{H}$ as some kind of abstract "energy" of a physical system—in physics, there is the principle that systems (for instance, electrons in atomic orbitals [10, p. 57]) "want" to minimise the total energy. Because this resembles my postulation above that the society wants to minimise total strain, I use the term *energy* also for my model. That will be discussed in more detail later, see Section 2.4 and in particular Definition 6.

- (8) is obviously symmetric under permutation of the individuals, as is the set Ω. In other words, permuting the individuals is a symmetry of the model system at hand. However, we will see in Chapter 3 and in particular Theorem 8 and Theorem 9, that in the actual ground state with minimal energy, society is very far from "symmetry" or equality between the individuals! This means that with lowering temperature, at some point the permutation symmetry between individuals is *spontaneously broken*[17], when one individual is "selected" over all others and gets a special role. In Subsection 4.3.3, this will be discussed in more detail.

- Usually in thermodynamics, energy is an *extensive* quantity [13, p. 120]. That means that it grows proportionally with the volume of the system. In the case of (8) however, it is clear that the value of $\mathcal{H}$ does not change proportionally with changing number of individuals N, since that is normalised away by the factor $\frac{1}{N}$ and the condition (6). Thus, (8) describes an *intensive* quantity. Nevertheless, I think that mathematically this formulation in form of a convex combination is nice—and the only place where it really matters is when we want to compare results for different N, since otherwise we're just scaling $\mathcal{H}$ by a constant factor, which does not influence the "physics" or system behaviour at all. Except for Subsection 4.3.4 later, we won't do that, and thus I think (8) is fine as it is. For that exceptional case, I will explicitly introduce a scaling factor to negate this effect and get comparable results.

- The basic idea of total strain can be compared to *utilitarianism* or *welfare economics* [37, welfare economics, p. 465f] in political economy. In particular, (8) may remind the reader of the *greatest happiness principle* of Jeremy Bentham ("the greatest happiness of the greatest number", [24, p. 183]). While I believe this association is justified in some aspects, there is a very important distinction: In (8), not directly the individual utilities (or strains) are added up—in addition, they are *weighted* so that not every individual has the same "importance" or influence! It is precisely this difference that later in Section 3.2 leads to the fact that societies following my model are *not very just, fair or equal* (see page 25) which is in fact contrary to Bentham's ideals!

- The weighting factor of (8) is a convex combination of the two weights $\frac{1}{N}$ and m_n. The first term expresses the situation of complete equality between importance of individuals, and the second takes only the political powers of individuals as weighting measure into account. γ serves to control the balance between those two extremes, with $\gamma = 1$ meaning total equality (in the sense of classical utilitarianism) and $\gamma = 0$ leading to a society where importance of individuals for social decisions is completely and solely based on their power. Thus γ can be understood as parameter governing the *fairness* of the society or political system, determining how much it cares about the "common people" who do not directly possess any political power.

- I believe that in real political systems, it is well justified to assume that there's a tendency to minimise an energy of a form like (8), in particular with individuals weighted by their power. The

[17] See also, for instance, [28, p. 396] or [28, 18.4.10, p. 1026ff].

reason for this is that in general, people tend to maximise their happiness—however, not everyone is able to do so, because laws and social decisions are not made by all people but rather some form of government in a political system. And thus, depending on political power of the individuals, they are able to leverage this power in a varying degree to influence the decisions for their own welfare. Hence it makes indeed sense to take political power into account for the overall strain that society tries to minimise! Of course, not all systems are so "corrupt" that only some people have any influence on decisions—after all, modern democracy strives to give everyone the same amount of power. But even in today's developed and democratic western countries a lot of people feel that political processes are not satisfactory for the majority of them—which was impressively demonstrated by the "Occupy Wall Street" protests of 2011 in the USA and all over the world [20]. That's where γ comes into the game, as it allows to measure and control the fairness of the political system in this respect within the model.

- Finally, when we think about minimising (8), it is clear that some kind of "correlation" appears between social status (manifested as individual strains $f(a_n, l_n)$) and political power (m_n as usual). Namely, if an individual has high political power, there will be some kind of "guiding force" which leads to lower energy when that individual also has less strain. On the other hand, if someone has high social status, the system has incentive to give him or her also more weight, or power.[18] This is in contrast for instance to Kreckel's notion of prestige, which is *a-priori defined* to grow with wealth and other advantages [22, p. 90], since in my model, m is actually independent of $f(a, l)$ but *a-posteriori the system itself* leads to a correlation between them in the end. Ultimately, this can be interpreted as giving the model a timocratic tendency. This is an effect that is surely plausible and occurs also in reality, and even manifests itself in the proverb "money rules the world". See for instance the opinion expressed in [42] or by the "Occupy Wall Street" protesters [20], and the discussion about social inequality at the beginning of this section.

Now that the basic principle is defined, I'll briefly investigate some of the mathematical aspects below. Note however, that much more theoretical analysis will be done later, in particular in Chapter 3.

Definition 5. Ω is called *regular* if there is some state $X \in \Omega$ such that for all individuals x_i, $i = 1, \dots, N$, $0 < a_i < a_{\max}$ and $l_{\min} < l_i$. This means that there are "interior" points of Ω.[19]

This notion of regularity of the set of configurations Ω will always be assumed later. It is clear that any modelling makes only sense when Ω is regular, since otherwise there either aren't any valid states at all or the parameters are chosen so "tightly" that the system hasn't any real freedom to show any interesting behaviour. In this context, see also Lemma 10 and the discussion on page 19. A more practical characterisation of regularity in terms of the parameters is as follows:

Lemma 4. *Ω is regular if and only if $l_{min} < \alpha a_{max}$. If this is case, then there also exist states $X \in \Omega$ with $f(a_n, l_n) < \infty$ for all $n = 1, \dots, N$ and consequently also $\mathcal{H}(X) < \infty$.*

Proof. Assume first that Ω is regular and choose one "interior" state $X \in \Omega$ as per Definition 5. Then

$$N l_{\min} < \sum_{i=1}^{N} l_i = \alpha \sum_{i=1}^{N} a_i < \alpha N a_{\max}.$$

Assume now conversely that the parameters satisfy this inequality, choose $l^* \in (l_{\min}, \alpha a_{\max}) \neq \emptyset$, and set $a^* = \frac{l^*}{\alpha}$ as well as $m^* = \frac{1}{N}$. Then $X^* = (x^*, \dots, x^*) \in \Omega$ and satisfies all requirements of Definition 5. For the additional statement, pick again $X \in \Omega$ fulfilling the conditions of Definition 5. Then by Definition 2, $f(a_i, l_i) < \infty$ for $i = 1, \dots, N$. □

[18] Note though, that this is a just an *explanatory description*. In my model, there are no formal concepts of time, dynamics or forces—and thus it is not mathematically justified to speak of something "leading to" a different state or a "force" that tries to move the system into some direction. But to gain an intuitive understanding, I believe that this description is still valuable.

[19] Those points are however not interior in the strict formal sense, since Ω is a subset of $\mathbb{R}^{3N}$ with dimension $3N - 2$ and thus has always empty topological interior.

Lemma 5. *Let $\gamma > 0$. Then smoothness and continuity (also up to the boundary with the convention around infinite values) of Definition 2 carry over from f to $\mathcal{H}$.*[20]

Proof. Note first that $\mathcal{H}(X) = \infty$ if and only if there is at least one x_i, $i = 1, \ldots, N$, with $f(a_i, l_i) = \infty$. Because $\gamma > 0$ ensures the coefficients in front of the strains in (8) are never zero, the case $0 \cdot \infty$ never occurs. Hence the claim is clear by (8) for all $X \in \Omega$ with $\mathcal{H}(X) < \infty$. Now let $\mathcal{H}(X) = \infty$, and assume that also $f(a_k, l_k) = \infty$ for some $k = 1, \ldots, N$. Then the coefficient in front of $f(a_k, l_k)$ in $\mathcal{H}$ is bounded away from zero by $\frac{\gamma}{N}$, and thus $\mathcal{H}$ can be kept arbitrarily large around X by the continuity of $f(a_k, l_k)$ around infinity from Definition 2. □

Theorem 1. *Let $\gamma > 0$ and Ω be regular. Then $\mathcal{H}$ has a global minimum on Ω.*[21]

Proof. Since Ω is regular, we can pick $X_0 \in \Omega$ with $\mathcal{H}_0 = \mathcal{H}(X_0) < \infty$ by Lemma 4. Set $A = \mathcal{H}^{-1}(\{\infty\})$. Then by Lemma 5, for each $X \in A$ there is $r_X > 0$ such that

$$\forall X' \in \Omega \cap B(X, r_X) : \mathcal{H}(X') > \mathcal{H}_0.$$

Now define $\Omega' = \Omega \setminus \cup_{X \in A} B(X, r_X)$. This is a compact set, since we take away the union of open sets from a compact set. Since we remove all of A by this process, $\mathcal{H}$ is finite on Ω' and thus continuous in the ordinary sense. Consequently, $M = \min_{X \in \Omega'} \mathcal{H}(X)$ exists. Clearly, $M \leq \mathcal{H}_0$ because $X_0 \in \Omega'$. But since all points $X' \in \Omega \setminus \Omega'$ have values $\mathcal{H}(X') > \mathcal{H}_0$, we find that M is indeed the global minimum of $\mathcal{H}$ over the full set Ω, too. □

Lemma 6. *Let Ω be regular. Then $\mathcal{H}$ has an infimum $\mathcal{H}_0 = \inf_{X \in \Omega} \mathcal{H}(X)$ over Ω, and $\mathcal{H}_0 \in \mathbb{R}$ is finite.*

Proof. This follows immediately by the infimum principle, Lemma 3 and Lemma 4. □

Note that Lemma 6 can be proved easily and without Theorem 5. Thus it is added here as useful complement to Theorem 1 that works also for $\gamma = 0$ in case we just need a lower bound and not necessarily a real minimum that is actually attained.

2.4 Likelihood of Different Configurations

Above in Section 2.3, I discussed the basic principle that describes the "preferences" of my model system with respect to states—namely minimisation of the total energy $\mathcal{H}$ of (8). However, it is clear that in a real system, this minimum will never be perfectly achieved, because the economy described by the model will be in constant change and although the incentive is there to minimise the energy, the resulting states will be in a dynamic equilibrium fluctuating around the minimum. In physics, this manifests itself in the fact that normally for instance in a gas, not all particles will be in the absolute ground state. Instead, they show random excitations above it—this phenomenon is called *temperature* T. The higher the temperature, the more fluctuations the system shows and the higher the expectation value of the actual energy of the system will be. If the temperature is near zero on the other hand, there will be nearly no fluctuations and the energy will be near the global minimum with the system in some state close to the "perfect" one. Note however that by the third law of thermodynamics [13, p. 89f], in physical systems $T = 0$ will never be reached exactly and thus there will *always* be such a fluctuation. This corresponds to my reasoning above that it is unrealistic to assume the society will be precisely in the global minimum. Nevertheless, I believe that it makes still sense to investigate the ground state corresponding to this case in Chapter 3, as it will give interesting insights and may also be a good approximation to real systems with low temperature (see also Theorem 2 below). In this section however, I will apply the framework of the canonical ensemble [13, ch. 22, p. 183ff] to my energy $\mathcal{H}$. For this, we first introduce an additional

[20] It should be remarked that $\gamma > 0$ is indeed a necessary condition: If $\gamma = 0$ and we have the situation of $m = 0$ and $f(a, l) = \infty$, their product vanishes by convention in Definition 4 and hence $\mathcal{H}(X) < \infty$ is possible. However, as soon as m gets perturbed, $\mathcal{H}(X') = \infty$ and thus continuity can not hold in general. If the convention in Definition 4 were the other way round, the result would still be the same: Then $\mathcal{H}(X) = \infty$ in the case above, but as soon as a and l get perturbed only slightly to yield $f(a, l) < \infty$ (but still large), $\mathcal{H}(X') = 0$ and again no continuity.

[21] We'll see in Theorem 5 that also for $\gamma = 0$ there always exists a minimum. But at this moment, it is convenient to exclude that case because then we can use Lemma 5.

state variable corresponding to the *social temperature* T[22], that measures the amount of fluctuation in some society around the optimum. Furthermore, it will be convenient below to work with the *inverse temperature* $\beta = \frac{1}{kT}$ instead of T itself. I'll always use T and β interchangeably, since they are uniquely related to each other. Here, Boltzmann's constant k has been introduced—I will assume $k = 1$ in the following by choosing the unit of temperature accordingly. With all that introduced, we can state the probability distribution that describes what states the system may be found in, namely (10). Note that throughout this section, I will assume that Ω is regular.

Definition 6. Let $T > 0$ or alternatively $\beta > 0$ be given. Then

$$\mathcal{Z} = \int_\Omega e^{-\beta \mathcal{H}(X)} \, dX \tag{9}$$

is the canonical partition sum, and for a measurable set[23] $A \subset \Omega$ of states, the probability of the system being in one of the states of A is assumed to be

$$\pi_T(A) = \frac{1}{\mathcal{Z}} \int_A e^{-\beta \mathcal{H}(X)} \, dX. \tag{10}$$

Again, I believe that following the important Definition 6, some remarks are in order:

- See [13, p. 185], (22.10) and (22.8) for the physical analogues of (9) and (10).
- For the integrations, we can ignore the parts of Ω where $\mathcal{H}$ takes infinite values—first because there the exponential factors vanish, and second because they form a null set anyway.
- Ω is bounded by Lemma 2, and $e^{-\beta \mathcal{H}(X)}$ is continuous and bounded itself by Lemma 6. Thus, the integrals are well-defined. Furthermore, since by regularity of Ω we have a subset with non-zero measure where $\mathcal{H}$ is finite, $\mathcal{Z} > 0$ and thus also (10) is fully well-defined.
- (10) describes a Boltzmann distribution, and it is assumed here that the model system follows such a distribution for configurations that are measured sufficiently far apart in time (so that they can be considered to be independent).
- Instead of the canonical ensemble that is used in Definition 6, one could also use the grand canonical one [13, p. 188ff]. In this framework, the number of particles respectively individuals is variable, and maybe that can be used to model migration phenomena into or out of the considered society. I have not followed this route, though, and always assume N to be a fixed parameter.

Lemma 7. *For each $T > 0$, π_T defined by Definition 6 is actually a probability measure*[24] *on Ω.*

Proof. It is a measure since we define it as integral against a density, and $\pi_T(\Omega) = 1$ holds by definition of $\mathcal{Z}$ in (9). See also [21, p. 293]. □

To conclude this section, I want to give interesting results that relate π_T to the principle of energy minimisation in the case of low temperatures:

Lemma 8. *Assume some measurable $A \subset \Omega$ is given. Let $\mathcal{H}_0 = \inf_{X \in \Omega} \mathcal{H}(X)$ by Lemma 6 and assume that there exists $\epsilon > 0$ such that for almost all $X \in A$, $\mathcal{H}(X) > \mathcal{H}_0 + \epsilon$. That is, $\mathcal{H}$ is on A uniformly bounded away from the global infimum over Ω. Then*

$$\lim_{T \to 0} \pi_T(A) = \lim_{\beta \to \infty} \pi_T(A) = 0.$$

[22] In [30, 1.1.3], the temperature introduced there is interpreted as *tolerance* the society has against violations of social norms and the individual desires to have social contact with members of some groups but not others. Here, it can be thought of as measure for how much society tolerates higher total strain than the global minimum, or maybe how much the individuals actually care about their strain at all. I'm not completely satisfied with this interpretation, though, and thus don't advertise any interpretation at all.

[23] In this work, I will always refer to the Lebesgue integral in dimension $3N - 2$ and corresponding measurability.

[24] See for instance [21, p. 96].

Proof. I'll show the limit for $\beta \to \infty$, the one for $T \to 0$ is clearly equivalent. By definition and extending the fraction with $1 = \frac{e^{\beta \mathcal{H}_0}}{e^{\beta \mathcal{H}_0}}$, we get:

$$\pi_T(A) = \frac{\int_A e^{-\beta \mathcal{H}(X)}\, dX}{\int_\Omega e^{-\beta \mathcal{H}(X)}\, dX} = \frac{\int_A e^{-\beta(\mathcal{H}(X)-\mathcal{H}_0)}\, dX}{\int_\Omega e^{-\beta(\mathcal{H}(X)-\mathcal{H}_0)}\, dX}$$

In the numerator, $\mathcal{H}(X) - \mathcal{H}_0 > \epsilon$ for $X \in A$ by assumption, so

$$\int_A e^{-\beta(\mathcal{H}(X)-\mathcal{H}_0)}\, dX \leq \int_A e^{-\beta\epsilon}\, dX = e^{-\beta\epsilon}\, |A|\, ^{25}.$$

For the denominator, observe that $\inf_{X\in\Omega}(\mathcal{H}(X) - \mathcal{H}_0) = 0$ and $\mathcal{H}$ is continuous (we can neglect the subset where $\mathcal{H}$ is infinite as mentioned already above). Thus for all $\delta > 0$, we can find a set C with $|C| > 0$ such that $\mathcal{H}(X) - \mathcal{H}_0 < \delta$ for all $X \in C$. But then:

$$\int_\Omega e^{-\beta(\mathcal{H}(X)-\mathcal{H}_0)}\, dX \geq \int_C e^{-\beta\delta}\, dX = e^{-\beta\delta}\, |C|$$

Hence together we have:

$$\pi_T(A) \leq \frac{e^{-\beta\epsilon}\, |A|}{e^{-\beta\delta}\, |C|} = e^{-\beta(\epsilon-\delta)}\, \frac{|A|}{|C|}$$

If we pick $\delta < \epsilon$, this vanishes for $\beta \to \infty$. □

Theorem 2. *Vice versa, let $\epsilon > 0$ and $A_\epsilon = \mathcal{H}^{-1}([\mathcal{H}_0, \mathcal{H}_0 + \epsilon])$. Then*

$$\lim_{T\to 0} \pi_T(A_\epsilon) = \lim_{\beta\to\infty} \pi_T(A_\epsilon) = 1.$$

So with temperature going towards absolute zero, the energy of the system will go in probability[26] *to the global infimum of $\mathcal{H}$.*

Proof. It is clear that A_ϵ is a Borel set and hence measurable, as is $\overline{A} = \Omega \setminus A_\epsilon$. By definition of $\overline{A}$, it satisfies the requirements of Lemma 8 and the claim follows. □

2.5 Choosing the Units

So far, I have not at all considered units or dimensions in the model (except for energy versus temperature when setting $k = 1$ on page 16). In fact, the strain function f as used throughout this document is not dimensionally correct (for instance in Example 1), since we would have to introduce additional coefficients to negate the dimensions of l and a as well as produce the correct dimension for f. This is not done for reasons of clarity. Furthermore, the transformation that will be introduced in (12) makes everything dimensionless anyway and thus fixes this problem as a side-effect. Of course, for the qualitative analysis of the system, dimensions are not important and can be neglected—however, I can also use them to my advantage: By cleverly choosing the units for my analysis, I can get rid of some of the parameters in the same way I got rid of Boltzmann's constant above, and that will be done in this section. This is a common and helpful technique in theoretical physics, see for instance the usual introduction of *atomic units* [28, 14.1.3, p. 734] with $e = \hbar = m_e = \frac{1}{4\pi\epsilon_0} = 1$. But let me start with some words about dimensions in my model in general. There, naturally these dimensions occur:

Goods Measures the income l and also GDP or output of the economy. In economics, this is usually measured—after introducing prices and accounting for inflation—in (real) money.

Labour Input a into the productive system, could be hours of labour.

Productivity Dimension of α, namely goods per labour.

[25] $|A|$ denotes the (Lebesgue) measure of some measurable set A.
[26] See [21, p. 312] for convergence of random variables in general.

Power m can be chosen to be dimensionless, since it just represents the *fraction* some individual comprises (with respect to power) of the whole society. In any case, by (8) this must also be the dimension of γ, and relates individual strain to total energy dimension-wise.

Strain The strain or unhappiness of single individuals, which is the dimension of the output of f. When power is dimensionless, this is the same as energy.

Energy $\mathcal{H}$ takes the role that energy has in a physical system, and is some kind of "aggregate strain".

Temperature With $k = 1$ (see page 16), temperature and energy have the same dimension. Then one can also say that temperature is measured directly in terms of the corresponding energy kT.

Now, we consider the parameters of the model and try to cleverly choose the units used later. Excepting T and γ, which are more like dynamic state variables than parameters of the model, and N, which is a simple dimensionless number in $\mathbb{N}$, we have $a_{\max}$, $l_{\min}$, α and the function f as parameters of the model. I will call a tuple $(a_{\max}, l_{\min}, \alpha, f)$ a *set of parameters* below. Then Definition 7 defines when we consider parameter sets to be *equivalent* or "essentially the same" (meaning that they result in the same behaviour of the system), and Lemma 9 justifies this definition:

Definition 7. Consider two sets of parameters $(a_{\max}, l_{\min}, \alpha, f)$ and $(a'_{\max}, l'_{\min}, \alpha', f')$. Then they are called *equivalent* if and only if $\psi : \Omega \to \Omega'$ exists as transformation and has the following properties:

- ψ is a diffeomorphism between Ω and Ω'.[27]
- $\det D\psi$[28] as well as $\det D\psi^{-1}$ are constant over Ω.[29]
- For all $\gamma \in [0, 1]$, $\mathcal{H}' \circ \psi = \mathcal{H}$ and $\mathcal{H} \circ \psi^{-1} = \mathcal{H}'$ on the respective domains Ω and Ω'.

Lemma 9. *Let two sets of parameters be equivalent. Then for all γ, T and measurable sets $A \subset \Omega$,*

$$\pi_T(A) = \pi'_T(\psi(A)). \tag{11}$$

By symmetry the analogous result holds also for measurable sets $B \subset \Omega'$.

Furthermore, $X \in \Omega$ is a global minimum of $\mathcal{H}$ if and only if $\psi(X)$ is a global minimum of $\mathcal{H}'$. Thus there is a one-to-one correspondence between global minima in both models for zero temperature, and also between probabilities of states for $T > 0$. In other words, both systems really behave the same way.

Proof. (11) follows from the integral transformation formula [9, (4.6), p. 204] when the substitution $X = \psi^{-1}(X')$ is performed:

$$\pi_T(A) = \frac{\int_A e^{-\beta\mathcal{H}(X)}\, dX}{\int_\Omega e^{-\beta\mathcal{H}(X)}\, dX} = \frac{\int_{\psi(A)} e^{-\beta\mathcal{H}(\psi^{-1}(X'))} \det D\psi^{-1}\, dX'}{\int_{\Omega'} e^{-\beta\mathcal{H}(\psi^{-1}(X'))} \det D\psi^{-1}\, dX'} = \pi'_T(\psi(A))$$

By Definition 7 the Jacobian determinant is a constant and can thus be cancelled. Now assume that $X \in \Omega$ is a global minimum of $\mathcal{H}$ and that $X' = \psi(X)$ is no global minimum of $\mathcal{H}'$. Then there exists $Y' \in \Omega'$ with $\mathcal{H}'(Y') < \mathcal{H}'(X')$. But then also

$$\mathcal{H}(\psi^{-1}(Y')) = \mathcal{H}'(Y') < \mathcal{H}'(X') = \mathcal{H}'(\psi(X)) = \mathcal{H}(X)$$

and this contradicts the optimality of X for $\mathcal{H}$. Thus X' must be optimal for $\mathcal{H}'$ and the minimal energies are the same for both parameter sets. □

Now that the basic machinery is there, we can apply it in Theorem 3 to show that we can transform any f such that the parameters can be equivalently chosen as $a'_{\max} = 1$ and $\alpha' = 1$. This can be interpreted as choosing to measure goods in terms of labour required for their production, and setting the unit of labour to be the maximal amount a person can physically deliver. After this procedure, only $L = l'_{\min}$ and f' remain as parameters of the model. For all further analysis below this transformation is assumed.

[27] This means that ψ is bijective and both ψ and ψ^{-1} are continuously differentiable, see also [9, p. 202].

[28] $\det D\psi$ denotes in general the determinant of the Jacobian of some differentiable function ψ.

[29] This requirement sounds very restrictive, and I'm not sure whether it is really necessary. It will be fulfilled, however, for the case where I need it below, and so this does not matter.

Theorem 3. *For a given set of parameters $(a_{max}, l_{min}, \alpha, f)$ there exist a proper strain function f' and $L > 0$ such that the parameters are equivalent to $(1, L, 1, f')$. f' is (strictly) convex or everywhere finite if and only if f has these properties.*

Proof. We use the transformation

$$x \mapsto x' = \left(m, \frac{a}{a_{\max}}, \frac{l}{\alpha a_{\max}}\right) \quad \Leftrightarrow \quad x' \mapsto x = (m', a_{\max}a', \alpha a_{\max}l')\,. \tag{12}$$

This clearly defines a diffeomorphism. Since for $X \in \Omega$,

$$1 \cdot \sum_{n=1}^{N} a'_n = \frac{1}{a_{\max}} \sum_{n=1}^{N} a_n = \frac{1}{\alpha a_{\max}} \sum_{n=1}^{N} l_n = \sum_{n=1}^{N} l'_n,$$

$\psi(X) \in \Omega'$ and vice-versa, thus we really map Ω diffeomorphic onto Ω' when we set $L = l'_{\min} = \frac{l_{\min}}{\alpha a_{\max}}$. Since the maps are linear, the Jacobians are independent of the points. For Definition 7 it thus remains to construct f' such that $\mathcal{H}' \circ \psi = \mathcal{H}$. Define simply $f'(a', l') = f(a, l)$ with the relation of a to a' and l to l' as per (12). Then we trivially get for arbitrary $X \in \Omega$ by definition:

$$\mathcal{H}'(\psi(X)) = \sum_{n=1}^{N} \left(\frac{\gamma}{N} + (1-\gamma)m'_n\right) f'(a'_n, l'_n)) = \sum_{n=1}^{N} \left(\frac{\gamma}{N} + (1-\gamma)m_n\right) f(a_n, l_n)) = \mathcal{H}(X)$$

Since we only applied a linear scaling to the parameters of f in order to define f', it is clear that f and f' share all qualitative properties. □

Note that by the concrete form of the transformation in (12), it is clear that in the new set of parameters the quantities a' and l' for the individuals as well as the remaining parameter L are dimensionless. L is now the subsistence wage measured as fraction of the maximally possible labour a person can do, and it is also the minimal average labour necessary for the whole population to survive:

Lemma 10. *If $X \in \Omega$, then $\frac{1}{N}\sum_{n=1}^{N} a_n \geq L$.*

Proof.

$$\frac{1}{N}\sum_{n=1}^{N} a_n = \frac{1}{N}\sum_{n=1}^{N} l_n \geq \frac{1}{N} NL = L$$

□

Because of this, we have to require $L \leq 1$ or rather $L < 1$ for the model to make sense, because for $L > 1$ there aren't any admissible points at all, and for $L = 1$ there is only one point admissible where all individuals have $a = l = 1$, which is not a very interesting case. Thus in the following we always assume $L < 1$. In doing so, we automatically gain also that Ω is always regular as per Definition 5, since then we can always choose some $L < a = l < 1$ and have all individuals in that point. To conclude this section, let me again emphasise that by choosing the units of goods and labour as per (12), we have reduced the parameters of the model described above to the dimensionless subsistence wage L and the transformed strain function f'. This helps to keep things as simple as possible in the following chapters, where the model described so far will now be analysed.

3 Analysis of the Ground State

Now that my model has been introduced above in Chapter 2, I will analyse it in this chapter. As has been mentioned already on page 12, the principle behind the model is minimisation of the total energy, which is given by $\mathcal{H}$ in (8). While the "full" system behaviour is governed by the Boltzmann distribution in (10), it is also interesting to consider just the *ground state* represented by the global minimum of $\mathcal{H}$ over Ω—recall Theorem 2 for a justification of this claim. This minimisation problem has a non-linear target functional with linear equality and inequality constraints describing the admissible set Ω. Hence it is not immediately easy to solve. Note however, that $\mathcal{H}$ has additional structure—namely that it is a convex combination of the individual strains, which depend only on a and l of the single individuals; m enters only via the coefficients in this convex combination. It will turn out that this structure makes it possible to simplify the problem quite a bit, leading to (22) in Section 3.3. We will also see that in addition to Theorem 1, the global minimum is unique if f is strictly convex in Theorem 10. Finally, in Section 3.4 I will consider the effect a change in γ or N has on the minimum.

3.1 Two Extreme Cases

Before I turn to the case of general γ below in Section 3.2, I want to find the ground state for the extreme cases of full equality ($\gamma = 1$) and the opposite ($\gamma = 0$). For them, the situation is even easier than for arbitrary γ, and thus it makes sense to consider those as a "warm-up exercise" for Section 3.2.

Lemma 11. *$a \mapsto f(a,a)$ maps $[L,1]$ onto $\mathbb{R} \cup \{\infty\}$ and $a^* = \mathrm{argmin}_{a\in[L,1]} f(a,a) < \infty$ exists. If f is strictly convex, the minimum a^* is unique.*

Proof. By Definition 2, the value ∞ can (if at all) only be attained at either $a = L$ or $a = 1$. If that is the case, we can by continuity up to the boundary restrict the domain to a smaller and still compact interval such that the mapping is finite. Of course, the minimum is not affected by this procedure at all if we don't restrict too much (see also the proof of Theorem 1). Then, the map is continuous, always finite on the new domain, and by Lemma 1 bounded from below since we restrict the income argument to be in $[L,1]$. Hence it has a global minimum on the compact domain. If f is strictly convex, clearly also the mapping $a \mapsto f(a,a)$ is. Recall that strict convexity implies necessarily the existence of a unique minimum, thus also here the global minimum is unique. □

Lemma 12. *Let $n \in \mathbb{N}$, $A \subset \mathbb{R}^n$ convex and $g : A \to \mathbb{R}$ be a convex function. Assume further that $m \in \mathbb{N}$ and $a^{(1)}, \dots, a^{(m)} \in A$ as well as $\lambda_1, \dots, \lambda_m > 0$ with $\sum_{i=1}^m \lambda_i = 1$. Then*[30]

$$\sum_{i=1}^{m} \lambda_i g\left(a^{(i)}\right) \geq g\left(\sum_{i=1}^{m} \lambda_i a^{(i)}\right). \tag{13}$$

If g is strictly convex and equality holds in (13), then $a^{(1)} = a^{(2)} = \cdots = a^{(m)}$ must be the case.

Proof. As remarked already, the first part is Jensen's inequality [9, 1.3, p. 221]. For the converse result, assume that g is strictly convex and equality holds in (13), but that $a^{(i)} \neq a^{(j)}$ for some $i, j = 1, \dots, m$. Then there must exist some coordinate $k = 1, \dots, n$ with $a_k^{(i)} \neq a_k^{(j)}$, assume without loss of generality that $a_k^{(i)} < a_k^{(j)}$. We define $J = \left\{a^{(l)} \mid a_k^{(l)} \geq a_k^{(j)},\ l = 1, \dots, m\right\}$ and $I = \{1, \dots, m\} \setminus J$, then I and J are disjoint and non-empty since $a^{(i)} \in I$ and $a^{(j)} \in J$. Hence we can choose

$$p = \sum_{l\in I} \frac{\lambda_l}{N_i} a^{(l)}, \quad q = \sum_{l\in J} \frac{\lambda_l}{N_j} a^{(l)}, \tag{14}$$

where $N_i = \sum_{l\in I} \lambda_l$ and $N_j = \sum_{l\in J} \lambda_l$. p and q are then clearly convex combinations of the points in I and J, respectively, and note that $0 < N_i, N_j < 1 = N_i + N_j$. By convexity of A, $p, q \in A$. Then note

[30] This is the well-known Jensen's inequality (see [9, 1.3 and 1.4, p. 221f] and in particular the finite version which is applied to show the inequality between arithmetic and geometric mean), but since I need it and didn't find that in the common literature, I additionally present the converse result for strict convexity.

that $\sum_{l=1}^{m} \lambda_l a^{(l)} = N_i p + N_j q$ is another convex combination. Thus by convexity of g, we get

$$\sum_{l=1}^{m} \lambda_l g\left(a^{(l)}\right) \geq N_i g(p) + N_j g(q) \geq g\left(N_i p + N_j q\right). \tag{15}$$

Since equality holds in (13), also equality must hold in (15). But then strict convexity of g implies that $p = q$, which is a contradiction since by definition in (14), $p_k < a_k^{(j)} \leq q_k$. □

Theorem 4. *Let f be strictly convex and $\gamma = 1$. Then $X \in \Omega$ is a global minimum of $\mathcal{H}$ if and only if $a_n = l_n = a^*$ for all $n = 1, \ldots, N$ and the in this case unique a^* from Lemma 11.*

Proof. Note that for $\gamma = 1$, the energy of (8) simplifies to $\mathcal{H}(X) = \frac{1}{N} \sum_{n=1}^{N} f(a_n, l_n)$, which is a convex combination of strains. Hence by Lemma 12, we get for arbitrary $X \in \Omega$:

$$\mathcal{H}(X) \geq f\left(\frac{1}{N}\sum_{n=1}^{N} a_n, \frac{1}{N}\sum_{n=1}^{N} l_n\right) = f\left(\frac{1}{N}\sum_{n=1}^{N} a_n, \frac{1}{N}\sum_{n=1}^{N} a_n\right) \geq f(a^*, a^*) \tag{16}$$

In addition, trivially $\mathcal{H}(X) = f(a^*, a^*)$ holds when we choose $p_n = (a^*, a^*)$ for all $n = 1, \ldots, N$. Thus those points are actually global minima (for arbitrarily chosen m's).

Assume now reversely that some $X \in \Omega$ is a global minimum of $\mathcal{H}$, which means $\mathcal{H}(X) = f(a^*, a^*)$. Then equality must hold in (16) and by strict convexity of f, all $a_n = a_0$ and $l_n = l_0$ for some a_0, l_0 and all $n = 1, \ldots, N$. Further, by (5) we know that $a_0 = l_0$. But then $\mathcal{H}(X) = f(a_0, a_0) = f(a^*, a^*)$ and consequently $a_0 = a^*$ by the uniqueness in Lemma 11. □

Note that Theorem 4 characterises the ground state for $\gamma = 1$ completely, one only has to know the optimum labour-income balance a^* from Lemma 11 for the concrete f considered. Furthermore, this minimum is unique except for the distribution of power via the m's—but for $\gamma = 1$, they don't have any significance, so the minimum can indeed be considered to be unique for practical purposes even if it is not strictly speaking unique as point in Ω. The result is that for complete equality in the "importances" of all individuals, convexity of f leads to a completely fair society where every individual occupies the same position in social space (see also the remark on page 8). Where exactly that is depends on f and is chosen such that the optimum balance between labour and income is met (via Lemma 11). Complete fairness in this case means that for every individual, the income it gets out of the economic system exactly equals its labour contribution. This may sound like a situation where people do not cooperate with each other and instead everybody works "on his or her own". Note however that my model does not consider specialisation or different sectors of the economy, but simply aggregated labour and income, both already converted to a common unit (like money). Thus my model and in particular Theorem 4 can not imply any statements about cooperation of individuals by specialisation—it only means that the individuals do not "trade labour and income between each other" (in the sense that someone works more than she earns, so that someone else can earn more than he produces himself), which would result in some people being better off, but inevitably others standing worse.

Lemma 13. *Let $\gamma = 0$, $t \in [0, 1]$ be such that $(N-1) + t - (N-1)L \geq L$[31] and consider the map $h : [t, 1] \to \mathbb{R} \cup \{\infty\}, \tau \mapsto \mathcal{H}(X^*(\tau))$ where*

$$x_1^*(\tau) = (1, \tau, (N-1) + \tau - (N-1)L) \text{ and } x_i^*(\tau) = (0, 1, L) \text{ for } i = 2, \ldots, N.$$

Then h is well-defined and has a global minimiser $\tau^* \in [t, 1]$. If f is strictly convex, τ^* is unique.

Proof. It is easy to see that $X^*(\tau) \in \Omega$ for each $\tau \in [t, 1]$—simply verify that (5) and (6) hold—, so well-definedness is clear. Since only the first individual has non-zero weight, (8) becomes in the special case considered

$$h(\tau) = f(\tau, (N-1) + \tau - (N-1)L). \tag{17}$$

(Note again the convention $0 \cdot \infty = 0$, see page 13!) Similarly to the proof of Lemma 11, we can restrict ourselves to parts of the domain where h is finite and then get a continuous function, which is bounded from below on a compact domain, and consequently a minimiser. Strict convexity of f carries over to h, and thus in that case, the minimiser is even unique. □

[31] Such a t exists, since at least $t = 1$ always satisfies this inequality.

Theorem 5. *Let f be convex and $\gamma = 0$. Then $X^*(\tau^*)$ from Lemma 13 minimises $\mathcal{H}$ over Ω.*[32]

Proof. Let $X \in \Omega$ be arbitrary. Then $\mathcal{H}(X)$ is a convex combination of values of f, and Lemma 12 gives

$$\mathcal{H}(X) \geq f\left(\sum_{n=1}^{N} m_n a_n, \sum_{n=1}^{N} m_n l_n\right).$$

Let $\tau = \min_n a_n$, then $\sum_{n=1}^{N} m_n a_n \geq \tau$. Furthermore,

$$\sum_{n=1}^{N} m_n l_n \leq \max_n l_n \leq (N-1) + \tau - (N-1)L,$$

since this is the maximum income available by (5) to anyone assuming one person works τ and all others work maximal and have only the subsistence wage as minimal income. Consequently, monotonicity of f as per Definition 2 gives the further estimate:

$$f\left(\sum_{n=1}^{N} m_n a_n, \sum_{n=1}^{N} m_n l_n\right) \geq f(\tau, (N-1) + \tau - (N-1)L) \geq \mathcal{H}(X^*(\tau^*)) \tag{18}$$

The last inequality in (18) comes from the fact that τ^* minimises exactly this function (17) over all possible τ by Lemma 13. □

Thus, Theorem 5 shows how the minimum looks like for the case of $\gamma = 0$, which means that the power values m of the individuals are not "softened" to fairer weights in the energy by a non-zero γ. In this case in the ground state, society forms a completely unfair difference between a single privileged individual and all others, who have no power at all, consequently no influence onto the total strain, and thus the worst possible position in social space so that the privileged individual can enjoy as much wealth as possible. The optimal τ^* from Lemma 13 describes in this case the best labour-income balance for the privileged person. This result, although it is of course most extreme for $\gamma = 0$, is typical for the behaviour of my model system for any $\gamma < 1$—then, the weighting of strains by individual powers in (8) inevitably leads to social differences between a single individual and all others. This will be shown in more generality below in Section 3.2 and is also evident from the numerical results in Subsection 4.3.3, but Theorem 5 gives a first hint that it is indeed so.

Theorem 6. *Let γ be arbitrary in $[0, 1]$ and Ω be regular. Then $\mathcal{H}$ attains a global minimum on Ω.*

Proof. This follows immediately from Theorem 1 for $\gamma > 0$ and from Theorem 5 for $\gamma = 0$. □

3.2 General Minimisation

Now, I want to continue with the derivation of the global energy minimum—following up on Section 3.1, in this section I will draw some conclusions for the case of general γ. In the end, we will see that when looking for the global minimum, we can restrict ourselves to the case of

$$m_1 = 1 \text{ and } m_n = 0 \text{ for all other individuals } n = 2, \ldots, N \tag{19}$$

in Theorem 9, and Theorem 8 will give additional information about the structure of this minimum—which then leads to a large simplification of the problem at hand, as will be discussed in Section 3.3. For this goal, we'll first show that there always *exists* a minimum which has this form:

Lemma 14. *Let $\gamma < 1$ and $X \in \Omega$ be a global minimum of $\mathcal{H}$. We then get for each $n = 1, \ldots, N$: If $f(a_n, l_n) > \min_k f(a_k, l_k)$, then necessarily $m_n = 0$. That is, every individual with non-zero power must have minimal strain of all individuals.*[33]

[32] Maybe for strict convexity one could get (restricted) uniqueness similarly to before, but it is unnecessary at this place as we will see uniqueness later in Theorem 10 anyway.

[33] Of course at this point, there may be several individuals that all have non-zero power and the same minimal strain. We'll see below in Lemma 19 together with the proof of Theorem 9 that this is not actually possible, though.

Proof. Assume to the contrary that $m_n > 0$ and $f(a_n, l_n) > f(a_m, l_m)$ for some $n, m = 1, \ldots, N$. Since $m_n > 0$, it follows from (6) that $m_m < 1$. Hence, we can define a new admissible point $X' \in \Omega$ which differs from X only by setting $m'_n = 0$ and $m'_m = m_n + m_m$. Because $\gamma < 1$, it follows clearly that $\mathcal{H}(X') < \mathcal{H}(X)$, which is a contradiction. □

Theorem 7. *There exists $X \in \Omega$ which is a global minimum of $\mathcal{H}$ and fulfils (19) in addition.*

Proof. Let X' be any optimum by Theorem 6. If $\gamma = 1$, we can without any effect redefine the m's to fulfil (19). So assume $\gamma < 1$ and pick n with $m_n > 0$. Without loss of generality, I will assume that $n = 1$; if not, redefine X' by permuting the individuals accordingly. Then by Lemma 14, $f(a_1, l_1) = \min_k f(a_k, l_k)$. Define $X \in \Omega$ with $p_i = p'_i$ for $i = 1, \ldots, N$ but $m'_1 = 1$ and $m'_i = 0$ for $i = 2, \ldots, N$. Then clearly $\mathcal{H}(X) \leq \mathcal{H}(X')$ and thus X must also be optimal. □

Theorem 7 shows that we can at least always find some minimum of the form (19). However, this is also a *necessary* condition for any minimum as will be established below in Theorem 9.

Lemma 15. *Let $\gamma < 1$ and $X \in \Omega$ be such that $n \neq m$ exist with $m_n, m_m > 0$, $a_n \in (0,1) \vee l_n > L$ and $a_m \in (0,1) \vee l_m > L$. Then X can not be optimal.*

Proof. Assume X is optimal. Then by Lemma 14, we know $f(a_n, l_n) = f(a_m, l_m)$. Since for both n and m, either labour or income is not on the boundary, we can slightly change one of those in both directions. The basic idea is now to exchange labour and/or goods between both points[34]—into that direction which makes the total energy smaller. Without loss of generality, I assume below that $a_n, a_m \in (0,1)$, but the argument works the same if $l_n > L$ and/or $l_m > L$ instead.

Because f is strictly monotone and differentiable by Definition 2, we know that

$$\frac{\partial f(a_n, l_n)}{\partial a} > 0, \quad \frac{\partial f(a_m, l_m)}{\partial a} > 0. \tag{20}$$

Now assume without loss of generality that

$$\frac{\partial f(a_n, l_n)}{\partial a} \geq \frac{\partial f(a_m, l_m)}{\partial a} \tag{21}$$

and define for $\epsilon \geq 0$ a new configuration

$$X_\epsilon = (x'_1, \ldots, x'_{n-1}, (1, a_n - \epsilon, l_n), x'_{n+1}, \ldots, x'_{m-1}, (0, a_m + \epsilon, l_m), x'_{m+1}, \ldots, x'_N)$$

where $x'_i = (0, a_i, l_i)$ have only the power set to zero compared to the original points x_i. Then for ϵ small enough, this is clearly again an admissible point. Since $\mathcal{H}(X_0) = \mathcal{H}(X)$ by Lemma 14, X_0 is also optimal. Consider now this expansion based on differentiability of f:

$$\begin{aligned}
\mathcal{H}(X_0) - \mathcal{H}(X_\epsilon) &= \left(\frac{\gamma}{N} + (1-\gamma)\right) f(a_n, l_n) + \frac{\gamma}{N} f(a_m, l_m) \\
&\quad - \left(\frac{\gamma}{N} + (1-\gamma)\right) f(a_n - \epsilon, l_n) - \frac{\gamma}{N} f(a_m + \epsilon, l_m) \\
&= -\left(\frac{\gamma}{N} + (1-\gamma)\right) \frac{\partial f(a_n, l_n)}{\partial a}(-\epsilon) - \frac{\gamma}{N} \frac{\partial f(a_m, l_m)}{\partial a} \epsilon + o(\epsilon) \\
&= \epsilon \left(\frac{\gamma}{N} \left(\frac{\partial f(a_n, l_n)}{\partial a} - \frac{\partial f(a_m, l_m)}{\partial a} \right) + (1-\gamma) \frac{\partial f(a_n, l_n)}{\partial a} \right) + o(\epsilon)
\end{aligned}$$

Note that the first term with partial derivatives is ≥ 0 by (21) and the second one is > 0 by (20). Dividing by ϵ and letting $\epsilon \to 0^+$, we get:

$$-\frac{\partial \mathcal{H}(X_\epsilon)}{\partial \epsilon^+} = \frac{\gamma}{N} \left(\frac{\partial f(a_n, l_n)}{\partial a} - \frac{\partial f(a_m, l_m)}{\partial a} \right) + (1-\gamma) \frac{\partial f(a_n, l_n)}{\partial a} > 0$$

This however is a contradiction to optimality. □

[34] Either by directly exchanging labour or income between both keeping the total GDP the same (if both a or both l are not on the boundary), or by decreasing the income of one with $l > L$ and at the same time decreasing the labour of the other with $a > 0$.

Lemma 15 means that for an optimal point, one $m_n = 1$ must hold except for "corner-cases" with points on the boundary. There, lack of freedom in varying the points in social space may forbid application of the technique used in the proof. However, we will see now that those corner-cases are not possible, and thus that (19) is *always* necessary.

Lemma 16. *Let f be strictly convex, m given and $\gamma > 0$ or $m > 0$. If $X \in \Omega$ is optimal and $m_n = m_m = m$ for some $n \neq m$, then $p_n = p_m$ and thus also $x_n = x_m$.*

Proof. Let $a = \frac{a_n + a_m}{2}$ and $l = \frac{l_n + l_m}{2}$. Then define X' similar to X, with the difference of $a'_n = a'_m = a$ and $l'_n = l'_m = l$.[35] Clearly, since

$$\sum_{k=1}^{N} a'_k = \sum_{k=1}^{N} a_k = \sum_{k=1}^{N} l_k = \sum_{k=1}^{N} l'_k,$$

the new point X' also fulfils (5) and is hence admissible. Assume $p_n \neq p_m$, then necessarily either $a_n \neq a_m$ or $l_n \neq l_m$. Define $\omega = \frac{\gamma}{N} + (1-\gamma)m$. Note that $m_n = m_m$ together with (6) implies $m < 1$, and thus it follows from the assumptions that $\omega \in (0, 1)$. Strict convexity of f gives:

$$\mathcal{H}(X') = R + 2\omega f(a, l) < R + 2\omega \frac{f(a_n, l_n) + f(a_m, l_m)}{2} = \mathcal{H}(X),$$

where

$$R = \sum_{k \neq n,m} \left(\frac{\gamma}{N} + (1-\gamma)m_k \right) f(a_k, l_k)$$

is the part of the energy that is common between X and X'. This contradicts optimality of X, and thus $p_n = p_m$ must hold. □

Lemma 17. *Let $\gamma = 0$, $X \in \Omega$ be optimal and $m_n = 0$ for some $n = 1, \ldots, N$. Then $a_n = 1$ and $l_n = L$.*

Proof. If that would not be the case, we could improve the energy by adding labour $\Delta a_n = 1 - a_n$, reducing income $\Delta l_n = l_n - L$ for individual n and then giving $\Delta a_n + \Delta l_n > 0$ as additional income to some individual m with $m_m > 0$. Since individual n has no weight at all (and considering the convention on page 13), this would by strict monotonicity of f reduce the total energy and thus violate the assumption of optimality. □

Theorem 8. *Let f be strictly convex and X optimal. Then there exist two points in the labour-income subspace of Ω, p_+ and p_0, such that $p_n = p_+$ for all n with $m_n > 0$ and $p_m = p_0$ for all m with $m_m = 0$. This means that in the optimum, if we neglect the power dimension of the social space, only two different positions are possible: One for all individuals with non-zero power, and one for all the others.*

Proof. Assume first that $m_n, m_m > 0$ for $m \neq n$. I'll show that for this case $p_n = p_m$ holds, which then implies the existence of p_+ as claimed. If $\gamma = 1$, we already know that in the unique minimum given by Theorem 4 all points are equal. So assume $\gamma < 1$. Then by Lemma 14, $f(a_n, l_n) = f(a_m, l_m)$. Hence, also X' is optimal if we change only $m'_n = m'_m = \frac{m_n + m_m}{2} > 0$. But then the claim follows via Lemma 16. Now for the existence of p_0, assume $m_n = m_m = 0$ for $m \neq n$. As before, I will show that $p_n = p_m$. If $\gamma > 0$, this is immediate with Lemma 16. Assume $\gamma = 0$. But then the claim also follows by Lemma 17. □

Lemma 18. *It is not possible to have $a_n = 1$ and $l_n = L$ for all $n = 1, \ldots, N$.*

Proof. Then we would have $\sum_{n=1}^{N} a_n = N = \sum_{n=1}^{N} l_n = NL$, which implies $L = 1$. Since we assumed $L < 1$ on page 19, this is a contradiction. □

Lemma 19. *Let $\gamma < 1$ and f be strictly convex. If $X \in \Omega$ is optimal and $n \neq m$ exist with $m_n, m_m > 0$, then necessarily $a_n = a_m = 0$ and $l_n = l_m = L$.*

[35] What is done here is that the two individuals n and m exchange income and labour to make their position in social space equal. Convexity of f then implies that this only lowers their total strain, as is argued in the following. Compare also the remark about what convexity of the strain function means on page 8.

Proof. By Theorem 8, clearly $a_n = a_m = a$ and $l_n = l_m = l$ must hold. $l > L$ would violate Lemma 15, thus we know $l = L$. By the same argument, necessarily either $a = 0$ or $a = 1$. Assume $a = 1$, then by the monotonicity property of Definition 2 and Lemma 14, $p_k = (a, l) = (1, L)$ must hold for *all* $k = 1, \dots, N$ since $f(1, L)$ is as high a strain as is ever possible, but is in addition the minimal strain of all individuals. (Note that this strain must be finite since optimality in combination with Lemma 4 implies $\mathcal{H}(X) < \infty$.) This contradicts Lemma 18, leading to $a = 0$. □

Theorem 9. *Let f be strictly convex, $\gamma < 1$[36] and X be optimal. Then (19) must be fulfilled except for permutation of the individuals (see also page 13).*

Proof. I'll show that there exists some $n = 1, \dots, N$ with $m_n = 1$. Assume this is not the case, then there are $a \neq b$ with $m_a, m_b > 0$. By Theorem 8, $p_a = p_b$ follows for those. Furthermore by Lemma 14, we know that $f(a_a, l_a) = f(a_b, l_b)$. Hence, $\mathcal{H}(X) = \mathcal{H}(X')$ where we changed $m'_a = m_a + m_b$ and $m'_b = 0$, and X' is also optimal. Then Theorem 8 implies that $p_b = p'_b = p_c$ for any $c = 1, \dots, N$ with $m_c = 0$. Consequently, $p_k = p_0$ for *all* $k = 1, \dots, N$ and some suitable p_0. It follows from Lemma 19 that $p_0 = (0, L)$. This can not be however, as it violates (5) for X' and X:

$$\sum_{k=1}^{N} a_k = N a_0 = 0 \neq NL = N l_0 = \sum_{k=1}^{N} l_k$$

□

Thus, we have shown in Theorem 9 that (19) can not only be achieved for every minimum, but is also a necessary optimality condition if we neglect the special case $\gamma = 1$ and permutations. So by the special coupling of power to strain in the energy (8), we indeed enforce total inequality in the population with respect to power. Lemma 14 and Theorem 8 give even further information, namely that the privileged individual with $m_n = 1$ has unsurprisingly the best social position of everyone (see also Theorem 11 below), and all others are at the same point in social space. Consequently, we have shown that social unfairness is inherent in this model!

Lemma 20. *Let $\gamma > 0$ and f be strictly convex. If we fix values m_n for $n = 1, \dots, N$ and just vary the corresponding p's, the energy $\mathcal{H}'(p_1, \dots, p_N) = \mathcal{H}(X)$ is strictly convex and has a unique minimum $(p_1, \dots, p_N)$.*

Proof. Uniqueness of the minimum follows immediately if we establish strict convexity. Since f and thus also $\mathcal{H}'$ is twice continuously differentiable by Definition 2, it suffices to show that the Hessian $\nabla^2 \mathcal{H}'$ is positive-definite. Since f is assumed to be strictly convex, we know already that $\nabla^2 f > 0$. The special form of $\mathcal{H}'$ leads to a block-diagonal structure of its Hessian, where the blocks are the 2×2 matrices

$$\left(\frac{\gamma}{N} + (1-\gamma) m_n\right) \begin{pmatrix} \frac{\partial^2 f(a_n, l_n)}{\partial a^2} & \frac{\partial^2 f(a_n, l_n)}{\partial a \partial l} \\ \frac{\partial^2 f(a_n, l_n)}{\partial a \partial l} & \frac{\partial^2 f(a_n, l_n)}{\partial l^2} \end{pmatrix} = \left(\frac{\gamma}{N} + (1-\gamma) m_n\right) \nabla^2 f(a_n, l_n) > 0$$

and thus the claim follows. □

Note that with Lemma 20, we can provide an alternative proof of Lemma 16: Let X be optimal and consider $\mathcal{H}'$ where we fix the m's to their respective values in X. But then for $m_n = m_m$, we can exchange p_n with p_m to get another optimal point (since the energy does not change in this step). However, since the value $(p_1, \dots, p_N)$ minimising $\mathcal{H}'$ is unique, we get that $p_n = p_m$ must hold because exchanging them is not allowed to make any difference.

Theorem 10. *Let f be strictly convex. Then there exists a unique minimum of $\mathcal{H}$ over Ω.*[37]

Proof. Existence was already established in Theorem 6. For $\gamma = 1$, the claim is immediate with Theorem 4. Otherwise, we know that (19) holds for the minimum by Theorem 9, which determines the power values uniquely. For $\gamma \in (0, 1)$, we can then apply Lemma 20 to get full uniqueness. It remains to consider the case $\gamma = 0$. Then Lemma 17 determines p_n uniquely for all $n = 2, \dots, N$ (where we assume that $m_1 = 1$ and permute otherwise). Uniqueness of p_1 follows by Lemma 13. □

[36] For the case $\gamma = 1$ the values of m are totally irrelevant and thus no such conclusion can be drawn.

[37] "Unique" means here of course except for the irrelevant values of m in the case $\gamma = 1$ and except for permutation of the individuals.

3.3 Four-Dimensional Minimisation

In Section 3.2 above, I have shown that the ground state energy minimum has a nice structure and also rather interesting qualitative behaviour (leading to unfairness in the society). However, it may be interesting to also calculate the minimum explicitly for a given concrete set of parameters. For that, one can still utilise the insights we got; in particular Theorem 8 and Theorem 9, as remarked also on page 25. This allows us to assume that the minimum $X^* \in \Omega$ has the form $X^* = ((1, p_1^*), (0, p_0^*), \dots, (0, p_0^*))$ where p_0^*, p_1^* minimise

$$\min_{a_0,l_0,a_1,l_1} \tilde{\mathcal{H}}(a_0, l_0, a_1, l_1) = \gamma \frac{N-1}{N} f(a_0, l_0) + \left(\frac{\gamma}{N} + (1-\gamma)\right) f(a_1, l_1) \tag{22}$$

subject to $(N-1)a_0 + a_1 = (N-1)l_0 + l_1$ as well as $a_0, a_1 \in [0,1]$ and $l_0, l_1 \geq L$. In fact, this form is completely equivalent to the original minimisation of $\mathcal{H}$ over Ω as far as the minimal energy and the minimising points are concerned: Any point admissible for (22) defines a point in Ω with the same energy, and the other way round we know from the above-mentioned theorems that the minimum of $\mathcal{H}$ over Ω can be expressed as suitable four-dimensional point for (22). However, the new problem defined by (22) is also much simpler and is—independent of the problem at hand, in particular the value of N—always only four-dimensional. If f is strictly convex, it is easy to see that also the new target function $\tilde{\mathcal{H}}$ is strictly convex (with block-diagonal Hessian), at least for $\gamma > 0$. Still, the problem does not get trivial—also in the new form, we have a *constrained* optimisation problem, with one linear equality constraint as well as box-constraints. Note that for γ near 0 or 1 and large N, the two terms in (22) have very different weight, leading to large differences in the eigenvalues of $\nabla^2 \tilde{\mathcal{H}}$ and possibly difficulties for numerical solutions.

I have not in detail considered the numerical properties of (22) because I believe that the qualitative analysis in Section 3.2 as well as the Monte-Carlo simulation for finite temperature in Chapter 4 are more important, but I still want to describe two methods for solving it. They have worked well for my example calculations and were also (in agreement with each other) used for the results presented in Example 5 below.[38]

1. Since the inequality constraints have the rather simple box form, it is easy to implement a gradient descent method where we project the resulting points after each step onto the admissible set to enforce the constraints. See [25, 6.13, p. 210ff] for a related method.

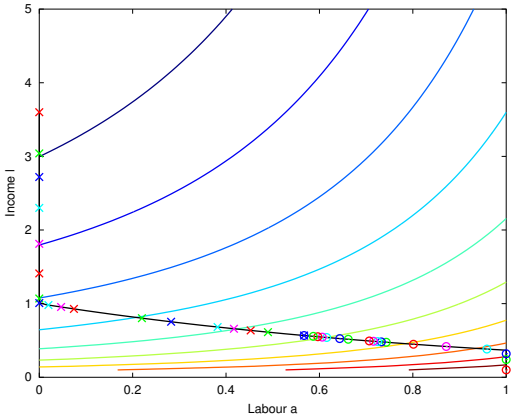
2. There are only six inequality constraints and in addition, the constraints $a_i \geq 0$ and $a_i \leq 1$, $i = 1, 2$, can not be active at the same time. This implies that there are only 36 possibilities which constraints are active in the minimum and which are not. One can try each of those, and in every case solve the corresponding problem *with only equality constraints*. I implemented this with Newton's method [19, p. 57] on the Lagrange function [19, (2.11), p. 144], which worked well for my trials.

Example 5. Consider again the strain function from Example 1 in (1), $f(a, l) = e^a - \log l$. It was already established there that this is a proper and strictly convex strain function. I solved (22) with this f numerically, and chose $L = 0.1$ as well as $N = 5$. The fairness was varied, namely

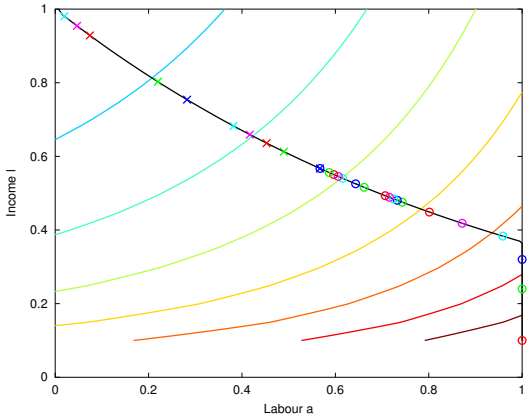
$$\gamma \in \{0.0, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.82, 0.83, 0.84, 0.85, 0.9, 0.92, 0.95, 0.96, 0.97, 0.98, 1.0\}.^{39} \tag{23}$$

Then the largest difference (in $\|\cdot\|_\infty$ norm) observed between any pair of results from the two methods above was $1.93 \cdot 10^{-6}$, and even smaller for most γ out of the range in (23). That seems to indicate that the results are accurate, since they were found by two independent methods. The optimal points are shown in Figure 2a and Figure 2b, where each γ corresponds to two points (a_0, l_0) and (a_1, l_1) of matching colour. With increasing γ, the (a_0, l_0) move inward from the lower right, and the (a_1, l_1) from the upper left. At $\gamma = 1$, they meet in the blue point near the centre of the right plots, where total fairness ensures equality between *all* individuals. One can see that the points form a continuous curve, which is drawn in solid black, when we consider the dependence on γ. This is a general property as we will see later in Theorem 13. Figure 2c and Figure 2d show the same plot for the f from Example 4 in (4) with $L = 0.0$. One can see that there's not too much qualitative difference in the behaviour of the

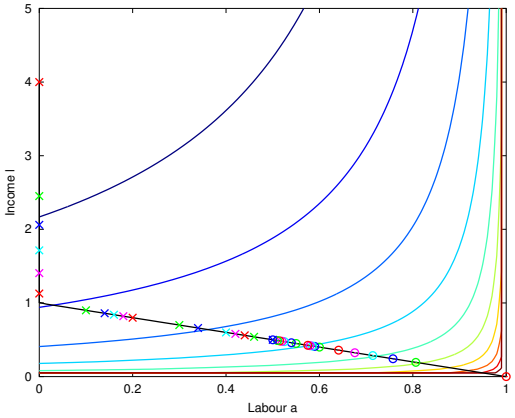
[38] I implemented these methods with GNU Octave [2] and executed the calculations on the `laptop` system of Table 2.
[39] Those values were chosen to give visibly good results but have no further meaning.



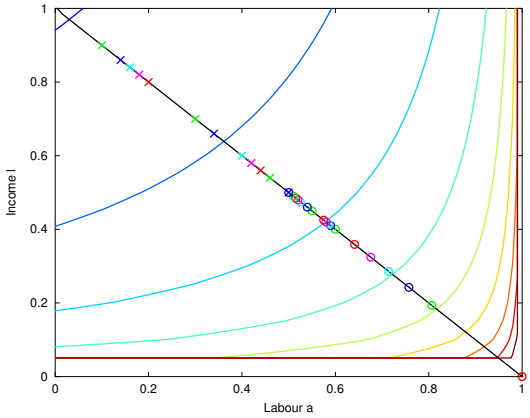
(a) (1), Full view.



(b) (1), Zoomed to central range.



(c) (4), Full view.



(d) (4), Zoomed to central range.

Figure 2: Minimum points of $\tilde{\mathcal{H}}$ from (22) for the situation in Example 5. The left and right pictures in both rows show the same situation, on the left such that all points are visible and on the right with zoom onto the centre, such that the interesting parts are better visible.

minimum between those two strain functions, although they are themselves quite different (one of them becomes infinite towards the boundary and one does not). The only really striking difference is that for (4) with log-barriers against $a = 1$ and $l = 0$, of course $a_0 < 1$ and $l_0 > 0$ for each $\gamma > 0$ (even still small ones). This is a first hint that the principal properties of the model do not depend strongly on the concrete form of f—we'll also see this below in Subsection 4.3.5.

For $\gamma = 1$ and f from (1), let's confirm this result analytically: By Theorem 4, we have to minimise $h(a) = f(a, a) = e^a - \log a$ over $a \in [L, 1]$. It is easy to verify that h has an interior minimum and is sufficiently regular, which means that we can find the minimum via the necessary optimality condition

$$h'(a) = e^a - \frac{1}{a} = 0 \quad \Leftrightarrow \quad ae^a = 1.$$

The solution of this equation can be expressed with the Lambert-W function [33, 4.13], and is then $a = W(1) \approx 0.57$. This matches the numerical result perfectly, with a difference of only $6.83 \cdot 10^{-10}$. Similarly for $\gamma = 0$, we can employ Theorem 5. In our case, all $\tau \in [0, 1]$ are allowed for Lemma 13. It is again easy to see that the optimum of (17) is this time on the boundary, namely at $\tau^* = a_1 = 0$ and hence $a_0 = 1$, $l_0 = L$ and $l_1 = (N - 1) - (N - 1)L = 4.0 - 0.4 = 3.6$. This again fits perfectly to the numerical result, the maximum difference here is $1.93 \cdot 10^{-6}$.

Theorem 11. *Consider the minimum of $\tilde{\mathcal{H}}$ for strictly convex f. Then $f(a_1, l_1) \leq f(a_0, l_0)$.*

Proof. For the proof, we make use of the equivalence of (22) to the full minimisation of $\mathcal{H}$ from (8) over Ω as discussed on page 26. Then the claim follows immediately from Lemma 14 if $\gamma < 1$ and Theorem 4 (with equality) if $\gamma = 1$. □

Theorem 11 formalises the claim that the privileged person has always a *better* position than the unprivileged mass of the population, which is intuitively clear and was already made on page 25. Note that it may also seem plausible that $a_1 \leq a_0$ or $l_0 \leq l_1$ always hold, but I have not been able to prove this. It may well be possible that more income allows more work to be done by the unprivileged people and hence it is optimal to give them both high a_0 and also high l_0. If on the other hand the claim is true, convexity of f must be crucial to show it, since with a non-convex f one can give counter-examples.[40] In any case, this is not an important fact—what really counts for social fairness is the strain (thus personal happiness) and hence Theorem 11.

Lemma 21. *Let $\gamma < 1$, f be strictly convex and (a_0, l_0, a_1, l_1) minimise (22). Then $a_0 \neq l_0$.*[41]

Proof. Assume $a_0 = l_0$. Then by the equality constraint, also $a_1 = l_1$. Furthermore, then even $a_0 = a_1 = a^*$ must hold with a^* being the unique optimum from Lemma 11: If that wouldn't be the case, then (a^*, a^*, a^*, a^*) would give a lower energy and violate the assumption of optimality. When we transform this back to the original form of the problem (see page 26), this means that $X = ((m_1, a^*, a^*), \ldots, (m_N, a^*, a^*))$ minimises $\mathcal{H}$ over Ω for arbitrary (m_n), since the choice of (m_n) does not influence the energy at all in this case. This is however a contradiction to the uniqueness of the minimum as per Theorem 10 or alternatively to Theorem 9. □

3.4 Dependence on Parameters

Let's take again a look at the minimisation problem we are facing (either the original one with objective function $\mathcal{H}$ from (8) or the reduced $\tilde{\mathcal{H}}$ in (22)): It depends on a number of parameters—but one can consider f and L to be given from the "outside" in the sense that these are related to physical properties of the individuals. On the other hand, γ and N are only related to the society as a compound system and can be changed more easily. In this section, we will investigate how the problem and in particular its solution change if γ or N are varied.

3.4.1 Changing the Fairness

With respect to γ, I'm interested in how the minimiser depends on it. Example 5 and in particular Figure 2 suggest that varying γ may lead to a continuous change in the minimum, such that one gets so to say a "curve" of optimal points for different γ's. In fact, this will really be the result in Theorem 13. But to achieve this, we need some theory about how disturbances in the objective function (which changing γ is) affect the minimising point. As far as I know, there's no theorem about this in the set of "standard tools", and I'm not aware of anything existing yet in this direction. In fact, it is clear that in general, small changes in the objective function need not necessarily lead to only small changes in the minimiser—consider for instance the case of a function with two local minima with changing relative depths; then at one point, a small change in the function may result in a jump of the global minimum from one of them to the other. However, only a small additional assumption—namely, that the function considered has a *unique* global minimum—is enough to make it work. I will give this general theoretical statement in Theorem 12, and then apply it to my model to get the promised practical result.

Lemma 22. *Let $K \subset \mathbb{R}^n$ be compact and $F \in C(K)$*[42] *be continuous. Assume that $\underline{F} = \min_{x \in K} F(x)$ is the minimal value of F over K and that the corresponding minimiser is unique. Then for all $\epsilon > 0$*

[40] For instance, $f(a, l) = \frac{a^2}{l^4}$ with $L = 0.1$, $N = 100$ and $\gamma = 0.5$. Then, $a_1 = 0$, $l_1 = 0.1$, $a_0 = 1$ and $l_0 \approx 1$, thus $l_0 > l_1$ in the optimum. But nevertheless $f(a_1, l_1) < f(a_0, l_0)$ holds!

[41] This is only a preparatory result to show $a_0 > l_0$ later in Theorem 15. This final result has to be deferred, though, since my proof needs Theorem 13.

[42] The set $C(K)$ is the set of all continuous functions $K \to \mathbb{R}$, which is a Banach space together with the (clearly well-defined) supremum-norm $\|h\|_\infty = \sup_{x \in K} |h(x)|$ for any $h \in C(K)$.

there exists a $\delta > 0$ such that for all $y \in K$:

$$|F(y) - \underline{F}| \leq \delta \quad \Rightarrow \quad \|x - y\| < \epsilon$$

This means that if we have a unique minimiser, then there's a whole region around it where F takes on values smaller than everywhere else on the domain.

Proof. Let $\underline{\mathrm{F}} = F(x)$ with $x \in K$ being the unique minimiser. Assume the opposite holds, and pick $\epsilon > 0$ such that for all $\delta > 0$ we can find $y_\delta \in K$ with $|F(y_\delta) - F(x)| \leq \delta$ but $\|x - y_\delta\| \geq \epsilon$. Then:

$$\frac{|F(y_\delta) - F(x)|}{\|x - y_\delta\|} \leq \frac{\delta}{\epsilon} \tag{24}$$

Choose a sequence $\delta_n \to 0$ with corresponding y_n. Since K is compact, we can always choose a convergent subsequence; thus assume $y_n \to y \in K$ without loss of generality. Because the denominator is bounded away from zero and we have continuity, we can pass the limit in (24) to get

$$\frac{|F(y) - F(x)|}{\|x - y\|} \leq \frac{0}{\epsilon} = 0.$$

Hence $F(y) = F(x)$ but $x \neq y$. This is a contradiction to the assumption of a unique minimiser. □

Lemma 23. *Let again $K \subset \mathbb{R}^n$ be compact and $F \in C(K)$. Let as before denote $\underline{F} = \min_{x \in K} F(x)$ the minimum of F over K.*[43] *Then for each $\epsilon > 0$ there exists a $\delta > 0$ such that for all $G \in C(K)$:*

$$\|F - G\|_\infty \leq \delta \quad \Rightarrow \quad |\underline{F} - \underline{G}| \leq \epsilon,$$

where $\underline{G} = \min_{x \in K} G(x)$. The minimal value thus depends continuously on the function being minimised.

Proof. Let $\epsilon > 0$ be given and choose $\delta = \epsilon$. Take $G \in C(K)$ as in the claim. Then $G(x) \geq F(x) - \epsilon$ for all $x \in K$ and hence $\underline{\mathrm{G}} \geq \underline{\mathrm{F}} - \epsilon$. On the other hand, if $\underline{\mathrm{F}} = F(x)$ for some suitable $x \in K$, then $G(x) \leq F(x) + \epsilon = \underline{\mathrm{F}} + \epsilon$. Hence also the minimum must satisfy $\underline{\mathrm{G}} \leq G(x) \leq \underline{\mathrm{F}} + \epsilon$. Taking both together gives the assertion. □

Theorem 12. *Let $K \subset \mathbb{R}^n$ be compact and $F \in C(K)$. Assume that F has the unique minimiser $x_0 \in K$ over K, with $\underline{F} = F(x_0)$ being the minimal value. Then for all $\epsilon > 0$ there is a $\delta > 0$ such that for all $G \in C(K)$ with $\|G - F\|_\infty < \delta$ and all minimisers $y \in K$ with $G(y) = \underline{G} = \min_{x \in K} G(x)$, $\|x_0 - y\| < \epsilon$.*

Thus, if we require that some continuous function F has a unique minimiser over a compact set, this minimiser itself depends continuously on F.[44]

Proof. Apply Lemma 22 to get δ_1 such that for all $y \in K$ with $|F(y) - \underline{\mathrm{F}}| \leq \delta_1$ we have $\|x_0 - y\| < \epsilon$. Further, use Lemma 23 to get δ_2 such that $|\underline{\mathrm{F}} - \underline{\mathrm{G}}| \leq \frac{\delta 1}{2}$ for all continuous G with $\|F - G\|_\infty \leq \delta_2$. Set $\delta = \min\left(\delta_2, \frac{\delta_1}{2}\right)$. Now pick any $G \in C(K)$ with $\|F - G\|_\infty \leq \delta$ and assume $G(y) = \underline{\mathrm{G}}$. By the definition of δ before,

$$|\underline{\mathrm{F}} - F(y)| \leq |\underline{\mathrm{F}} - \underline{\mathrm{G}}| + |G(y) - F(y)| \leq \frac{\delta_1}{2} + \frac{\delta_1}{2} = \delta_1.$$

Hence $\|x_0 - y\| < \epsilon$. □

Lemma 24. *Assume that f is everywhere finite. Consider the map $[0,1] \to C(\Omega)$, $\gamma \mapsto \mathcal{H}_\gamma$. This map is then Lipschitzian and thus in particular uniformly continuous. The same holds true for $\tilde{\mathcal{H}}$ from (22).*

[43] This time, uniqueness of the minimiser is not required.

[44] We do not need that the changed function G has a unique minimiser anymore (which is not the case for all changes, of course) to get that (each) minimiser of G is close to the unique one of F.

Proof. Since Ω is compact by Lemma 2 and f is finite and continuous up to the boundary, we can define

$$F = \max_{X \in \Omega} \sum_{n=1}^{N} |f(a_n, l_n)| .$$

But then clearly for arbitrary $\gamma, \gamma' \in [0,1]$, since $0 \leq m_n \leq 1$:

$$\begin{aligned} \|\mathcal{H}_\gamma - \mathcal{H}_{\gamma'}\|_\infty &= \left\| \frac{\gamma - \gamma'}{N} \sum_{n=1}^{N} f(a_n, l_n) + (\gamma' - \gamma) \sum_{n=1}^{N} m_n f(a_n, l_n) \right\|_\infty \\ &\leq |\gamma - \gamma'| \left\| \frac{1}{N} \sum_{n=1}^{N} |f(a_n, l_n)| + \sum_{n=1}^{N} m_n |f(a_n, l_n)| \right\|_\infty \leq |\gamma - \gamma'| \cdot F \cdot \left(1 + \frac{1}{N}\right) \end{aligned}$$

The argument for $\tilde{\mathcal{H}}$ is basically the same. □

Theorem 13. *Assume that f is everywhere finite and strictly convex. Consider the problem in (22). Then the minimiser depends continuously on the parameter γ.*

Proof. By Lemma 24 the objective function depends continuously on γ. We have a unique minimiser because of Theorem 10, and note that in the form of (22), the minimiser is *really* unique (without the need to introduce exceptions for permutations or the like). Hence we can apply Theorem 12 to get the desired result. □

In addition to continuous dependence, the solution also shows a kind of *monotonicity* in γ, namely that with increasing fairness, the strains of the unprivileged people decrease and that of the privileged individual increases. Hence it is—from a point of fairness and utilitarianism—really desirable to get high values of γ, as that increases the happiness "of the masses" or "the 99 %" [20]. As before on page 28, I do not have any results for the coordinates a or l directly—but also this time I believe that this is not important, and the really interesting result is Theorem 14:

Theorem 14. *Consider the problem in (22) and let $\gamma < \gamma'$. Then $f(a_0, l_0) \geq f(a_0', l_0')$ and conversely $f(a_1, l_1) \leq f(a_1', l_1')$, where (a_0, l_0, a_1, l_1) and (a_0', l_0', a_1', l_1') are the minimisers for γ and γ', respectively.*

Proof. Note first that the claim is obvious for $\gamma = 0$ since then $(a_0, l_0) = (1, L)$ by Lemma 17 and $f(a_0, l_0)$ is maximal, while $f(a_1, l_1)$ is minimal because it is the only component of $\tilde{\mathcal{H}}$ being minimised in (22). So let $\gamma > 0$, then both terms of $\tilde{\mathcal{H}}$ in (22) actually contribute. Let $f_0 = f(a_0, l_0)$, $f_1 = f(a_1, l_1)$ and f_0', f_1' be similarly defined. Assume that $f_0 < f_0'$ and $f_1 \leq f_1'$ (i. e., both groups are worse off in the minimum for γ' than for γ). This can not be, since then we could take the admissible point (a_0, l_0, a_1, l_1) instead of (a_0', l_0', a_1', l_1') also for γ', which would yield a lower objective value. Similarly we can rule out that $f_0 \geq f_0'$ and $f_1 > f_1'$. Assume that $f_0 < f_0'$ and $f_1 > f_1'$. Define $w = \gamma \frac{N-1}{N}$, w' in analogy, and note that $w < w'$. Optimality then gives:

$$\begin{aligned} w f_0 + (1-w) f_1 &\leq w f_0' + (1-w) f_1' \\ w' f_0' + (1-w') f_1' &\leq w' f_0 + (1-w') f_1 \end{aligned}$$

Adding both inequalities and simplifying yields $f_1 + f_0' \leq f_1' + f_0$, which contradicts our assumption that would give $f_1 + f_0' > f_1' + f_0$. Hence the only remaining possibility is the claimed assertion. □

With Theorem 13, we can now complete what was promised before on page 28 in Theorem 15. This is a quite important result for the qualitative behaviour of the model, as it finally puts on solid grounds what was implicitly assumed above already: Namely that it is really the case that the unprivileged people work more than they earn, so that the privileged individual can in turn ripe off their production for itself. Nevertheless, I think I should remark at this point that it is possible that this effect vanishes for $N \to \infty$ (which means $0 < a_0 - l_0 \to 0$), as we will see later with Example 7.

Theorem 15. *Let $\gamma < 1$, f be strictly convex and everywhere finite, and assume (a_0, l_0, a_1, l_1) minimises (22). Then $a_0 > l_0$, and consequently by the equality constraint $a_1 < l_1$.*

Proof. Define the auxiliary function $g(\gamma') = a_0^*(\gamma') - l_0^*(\gamma')$, where a_0^* and l_0^* describe the position of the minimiser depending on γ'. By Theorem 10, these are well-defined and by Theorem 13 even continuous for all $\gamma' \in [0,1]$, as is g. Because of Theorem 5, we know $g(0) = 1 - L > 0$ (see page 19). $g(\gamma) = 0$ is ruled out by Lemma 21, assume $g(\gamma) < 0$. Then by the intermediate value theorem, there is $\gamma_0 \in (0, \gamma)$ with $g(\gamma_0) = 0$. This contradicts Lemma 21, and thus $g(\gamma) > 0$ which is equivalent to the claim. □

3.4.2 Size of the Population

Now we're interested in the behaviour of the model for $N \to \infty$. I believe this is a (somewhat) realistic limit, since in real societies, the number of individuals is large. On the other hand, it is clear that if we let the size of the population grow unlimited, then the single privileged individual that theory predicts for the solution will intuitively become less and less important and it is surely not realistic to believe that this structure of the solution will be "physical" beyond some threshold for the size of the society. Instead it may be that in this case, we have to think of the individuals in our model as being groups of real people—so that the privileged individual becomes a ruling *class*, which consists of more than a single person but is still a minority compared to the class of unprivileged people. N for a real society would then be some fixed but finite value corresponding to the ratio between the size of the unprivileged and ruling classes. This is however beyond the scope of my considerations here, and I still want to present at least one interesting result about what happens for $N \to \infty$:

Theorem 16. *Consider $N \to \infty$, let $\mathcal{H}_N$ be the corresponding energies, and let X_N be the minimisers of $\mathcal{H}_N$ over Ω_N. Set $F_N = \mathcal{H}_N(X_N)$ to be the minimal values. Then the sequence (F_N) is either unbounded from below or converges to $\underline{f} = \inf_{N\in\mathbb{N}} F_N$ for $N \to \infty$. If it is unbounded, it diverges to $-\infty$.*

Proof. For $\gamma = 0$, the result follows from the known form of the optimum in Theorem 5 and monotonicity of f in Definition 2, since then with increasing size of the population, there is more surplus the privileged person can get from the growing number of unprivileged ones. So assume $\gamma > 0$ from now on. Assume the sequence is bounded from below, then $\underline{f} \in \mathbb{R}$ is well-defined. Let $\epsilon > 0$ be given and choose N_0 such that $F_{N_0} < \underline{f} + \frac{\epsilon}{3}$. Pick $a \in [L, 1]$ with $f(a,a) < \infty$ (which is possible by Lemma 11), then we can find M such that $\frac{|f(a,a)|\gamma}{M} < \frac{\epsilon}{3}$. Let $X = X_{N_0}$ denote the optimum for N_0. Without loss of generality, we can choose M such that also

$$\frac{1}{MN_0}\left|\sum_{k=1}^{N_0} f(a_k, l_k)\right| < \frac{\epsilon}{3}. \tag{25}$$

For this, note that in the energy with $\gamma > 0$ each individual has non-zero weight and thus in the optimum all $f(a_k, l_k)$ must be finite for $k = 1, \dots, N_0$. Let $N \geq MN_0$ be given and assume that $m'N_0 \leq N < (m'+1)N_0$ for some suitable $m' \in \mathbb{N}$ (integer division). We define a point $X^* \in \Omega_N$ by "duplicating" the points in X to appear each m' times with m_k changed to $\frac{m_k}{m'}$ for $k = 1, \dots, N_0$ and then adding $N - m'N_0$ points $(0, a, a)$ to get a total of N individuals. X^* is then clearly admissible. Note that

$$\left|m' \sum_{k=1}^{N_0} \frac{f(a_k,l_k)}{N} - \sum_{k=1}^{N_0} \frac{f(a_k,l_k)}{N_0}\right| = \left|\frac{m'N_0 - N}{NN_0}\right| \left|\sum_{k=1}^{N_0} f(a_k,l_k)\right| \leq \frac{N_0}{NN_0}\left|\sum_{k=1}^{N_0} f(a_k,l_k)\right| < \frac{\epsilon}{3}$$

using (25). Thus we further get:

$$m' \sum_{k=1}^{N_0} \frac{f(a_k,l_k)}{N} < \sum_{k=1}^{N_0} \frac{f(a_k,l_k)}{N_0} + \frac{\epsilon}{3} \tag{26}$$

Hence also:

$$\begin{aligned}
F_N &\leq \mathcal{H}_N(X^*) = m' \sum_{k=1}^{N_0} \left(\frac{\gamma}{N} + (1-\gamma)\frac{m_k}{m'}\right) f(a_k,l_k) + (N - m'N_0)\frac{\gamma}{N} f(a,a) \\
&< \sum_{k=1}^{N_0} \left(\frac{\gamma}{N_0} + (1-\gamma)m_k\right) f(a_k,l_k) + \gamma\frac{\epsilon}{3} + \frac{N_0}{MN_0}\gamma\,|f(a,a)| \\
&\leq F_{N_0} + \frac{\epsilon}{3} + \frac{|f(a,a)|\,\gamma}{M} < \underline{f} + \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \underline{f} + \epsilon
\end{aligned}$$

This shows convergence to f, by establishing that the sequence is kind of "quasi-decreasing".[45] Note also that we really need (26) in order to estimate the first term in the sum over k to get the second line, since we do not know the sign of the $f(a_k, l_k)$ terms and don't have an absolute value! With the same argument one can also show that in the case of unboundedness, (F_N) actually diverges to $-\infty$. □

Note that the case of unboundedness in Theorem 16 is not unrealistic and may well occur because the privileged individual gets more and more income as more and more unprivileged people are there to work (recall also Theorem 15). Consider for simplicity now the problem in the form of (22). The equality constraint there implies that for the unprivileged people, l_0 must be bounded from above. Consequently, also $f(a_0, l_0)$ is bounded from below by Lemma 1, and thus we find that the total $\tilde{\mathcal{H}}$ is unbounded from below if and only if $f(a_1, l_1)$ is unbounded.

Lemma 25. *Let $\gamma < 1$. Then unboundedness in Theorem 16 is the case if and only if*

$$\lim_{l\to\infty} f(0, l) = -\infty, \tag{27}$$

which by the monotonicity property of f is equivalent to the requirement that $f(a, l)$ diverges for some arbitrary $a \in [0, 1]$.

Proof. Consider the problem again for simplicity in the form (22). Since $L < 1$ is generally assumed (see page 19), we can choose $L < l_0 < a_0 < 1$. Set $a_1 = 0$ and $l_1 = (N-1)(a_0 - l_0)$. Then this is a feasible point for (22) if only N is large enough. Hence the minimal value of $\tilde{\mathcal{H}}$ will be even less than this point's energy, which is

$$\tilde{\mathcal{H}}(a_0, l_0, a_1, l_1) = \gamma \frac{N-1}{N} f(a_0, l_0) + \left(\frac{\gamma}{N} + (1-\gamma)\right) f(0, l_1).$$

In the limit $N \to \infty$, the first term converges and the second one goes to $-\infty$ by assumption, since we require $\gamma < 1$. The other direction is of course clear: If $\tilde{\mathcal{H}}$ is unbounded from below, it must be the case that f is unbounded in the limit of large income (which is the only way it can be unbounded from below and thus make $\tilde{\mathcal{H}}$ go to $-\infty$). □

It seems again plausible that not only the total energy—as we found in Theorem 16—has a limit for $N \to \infty$, but in fact also the individual components a_0, l_0, a_1, l_1 do (of course, for l_1 the limit may well be ∞). I have no result on this, however.[46] One may even think that $a_1 \to 0$ must be the case, with the following argument: If more and more unprivileged individuals work and the income l_1 grows towards infinity, then the additional income which the privileged individual can earn "by itself" via $a_1 > 0$ becomes negligible and thus it is optimal to have $a_1 = 0$. This, however, is not true, as will be shown in Example 6. If we, on the other hand, simply want to know whether there are *any* limits to the coordinates, this seems to be the case there.

Example 6. Consider the strain function from Example 2, $f(a, l) = e^{a-l} - \log l$ in (2). I chose $L = 0.1$ and $\gamma = 0.5$ and calculated the optimum numerically with Newton's method as described on page 26 for different (increasing) N. The results are shown in the first group of columns in Table 1. As can be seen from the table, in this case it is optimal also for the privileged individual to work as much as possible, independent of N. This is because with the special strain used (as mentioned already in Example 2), the penalty for it to work decreases very rapidly with its growing income. Another interesting observation in this example is that the values of l_0 do *not behave monotonically*—they increase for N from 5 to 20, and then decrease from 20 to 10,000. So while monotonic behaviour could also have seemed plausible, it is not true in general!

[45] I believe it could well be really a decreasing sequence, but don't have a proof or counter-example for this. It does not matter for the argument, though.

[46] If one wanted to show this, one idea could be to eliminate l_1 from (22) via the equality constraint. Then notice that in the new form, the only way in which the domain continues to depend on N is via the requirement $l_1 = a_1 + (N-1)(a_0 - l_0) \geq L$, which is however very likely to be inactive anyway. In particular, (27) is a sufficient condition, as can be seen from the proof of Lemma 25. So one may be tempted to apply Theorem 12 to this new problem, arguing that the coefficients in (22) converge for $N \to \infty$ and thus the energy does not change much for N large enough, and so the minimising coordinates form a Cauchy sequence. I think that in special cases, this may work—but note that, in particular if f is unbounded as per (27), one can not easily estimate the difference in the energy with l_1 eliminated to be small: Changing N also affects l_1 and if $N \to \infty$ implies $l_1 \to \infty$ and this in turn $f(a_1, l_1) \to -\infty$, the energy *will* change drastically for fixed a_0, l_0, a_1 when changing N_0 to some N' if only N' is large enough, even if N_0 is also large. Thus, probably something different from this idea has to be used.

One can also see that $a_1 = 1$ must be the case for any N with a theoretical consideration: Assume that some (a_0, l_0) is fixed already and consider $\tau = a_1$ to be the independent variable. Then the energy being minimised becomes

$$h(\tau) = \tilde{\mathcal{H}}(a_0, l_0, \tau, \tau + (N-1)(a_0 - l_0)) = \gamma \frac{N-1}{N} f(a_0, l_0) + \left(\frac{\gamma}{N} + (1-\gamma)\right) f(\tau, \tau + (N-1)(a_0 - l_0)).$$

One can easily see that

$$\frac{\partial h(\tau)}{\partial \tau} = \left(\frac{\gamma}{N} + (1-\gamma)\right) \frac{-1}{\tau + (N-1)(a_0 - l_0)} < 0,$$

where we used $a_0 - l_0 > 0$ by Theorem 15. Thus, the optimal τ must be at the upper boundary, $\tau = a_1 = 1$ for all γ and N.

Example 7. It is also interesting to consider what the strain function $f(a, l) = e^{a^2 - l}$ from Example 3 in (3) does for growing N—see the second set of columns in Table 1. In this case, since the benefit from additional income of the privileged individual vanishes rapidly, there is no strong incentive to get $l_1 \to \infty$ fast and thus one finds that $a_0 - l_0 \to 0$. The "magic" value $a = l = \frac{1}{2}$, to which three out of the four coordinates converge for $N \to \infty$ in this example, is the optimum as per Lemma 11. This can be seen easily by the necessary optimality condition for the corresponding minimisation. Furthermore, in contrast to Example 6 and Example 8, unboundedness from Lemma 25 does *not* apply here.

Example 8. Finally, the last columns of Table 1 show the behaviour of $f(a, l) = e^a - \log l$ from Example 1 and (1). In this case, neither of the "special cases" above applies—l_1 grows towards infinity with $a_0 - l_0$ bounded away from zero, and we have some arbitrary limits inside the domain. One can however see that in this case, *all* individuals benefit from growing N. This is not a general property (see falling l_0 for large N in Example 6), but fits in well with the infimum property of $\mathcal{H}$ as per Theorem 16.

	$f(a,l) = e^{a-l} - \log l$				$f(a,l) = e^{a^2-l}$				$f(a,l) = e^{a} - \log l$			
N	a_0	l_0	a_1	l_1	a_0	l_0	a_1	l_1	a_0	l_0	a_1	l_1
5	1.000	0.571	1.000	2.716	0.500	0.142	0.500	1.933	0.959	0.383	0.000	2.301
10	1.000	0.638	1.000	4.260	0.500	0.260	0.500	2.658	0.902	0.406	0.000	4.464
20	1.000	0.666	1.000	7.342	0.500	0.348	0.500	3.392	0.876	0.416	0.000	8.742
50	1.000	0.663	1.000	17.526	0.500	0.421	0.500	4.353	0.862	0.422	0.000	21.540
100	1.000	0.661	1.000	34.610	0.500	0.454	0.500	5.069	0.857	0.424	0.000	42.858
200	1.000	0.659	1.000	68.780	0.500	0.473	0.500	5.777	0.855	0.425	0.000	85.489
500	1.000	0.659	1.000	171.294	0.500	0.488	0.500	6.704	0.854	0.426	0.000	213.381
1,000	1.000	0.659	1.000	342.149	0.500	0.493	0.500	7.402	0.853	0.426	0.000	426.533
2,000	1.000	0.658	1.000	683.861	0.500	0.496	0.500	8.098	0.853	0.426	0.000	852.835
5,000	1.000	0.658	1.000	1,708.995	0.500	0.498	0.500	9.016	0.853	0.426	0.000	2,131.744
10,000	1.000	0.658	1.000	3,417.553	0.500	0.499	0.500	9.710	0.853	0.426	0.000	4,263.258

Table 1: Numerical minimisers of (22) for the different strain functions used in Example 6, Example 7 and Example 8, showing their behaviour for increasing N. I always used $L = 0.1$ and $\gamma = 0.5$, the values have been calculated with Newton's method (page 26).

4 Results for Finite Temperature

Following the detailed analysis of the ground state with minimum energy above in Chapter 3, in this chapter I want to investigate and discuss the behaviour of the model for the case of non-zero temperature $T > 0$. In that case, there no longer exists some $X \in \Omega$ that precisely defines the state of the system. Instead, it can only be described by the probability distribution π_T on Ω, which was introduced in Definition 6 and (10). The interesting quantities to consider are now *expectation values* of certain observables and how they behave when we change the parameters of the system (mainly β and γ, but also N). The observables that interest me as well as some analytical properties of them are introduced in Section 4.1. Note however, that for the most part in this chapter, I won't be able to present theoretical calculations—only numerical results are available. In Section 4.2 I describe the Monte-Carlo procedure that I developed and used for these calculations (which is based on the Metropolis algorithm [15, 4.1.3, p. 78f]), and in Section 4.3 I finally present some "beautiful pictures" and my results in general. Section 4.4 is a small addition, discussing how the parameters β and γ can be estimated if one has knowledge of a particular measured configuration $X \in \Omega$.

4.1 Observables and Expectation Values

Observables are physical quantities that can be calculated from particular configurations of the system, and then describe its current state in a more "compressed" and usable form. For instance, in ordinary thermodynamics, the configurations of a gas describe in detail every single particle; more interesting and useful are however quantities like energy, pressure or entropy (more precisely, their expectation values when the thermal distribution of configurations is taken into account). This is exactly what we're going to consider also for our model in this section. In mathematical terms, such an observable is a *random variable* on the probability space that π_T from Definition 6 defines on Ω:

Definition 8. Let $\Omega' = \mathcal{H}^{-1}(\mathbb{R})$ be the set where the energy is finite[47], and define $\Omega'' = \Omega' \times [0,1]$ to be this set extended by the possible ranges of the parameter γ. Then an observable is a continuous function $\mathcal{O} : \Omega'' \to \mathbb{R}$, which is furthermore infinitely differentiable with respect to the parameter γ.[48]

Lemma 26. *Let $\mathcal{O}$ be an observable. Then its expectation value*[49]

$$\langle \mathcal{O} \rangle = \int_\Omega \mathcal{O}(X)\, d\pi_T(X) = \frac{1}{\mathcal{Z}} \int_\Omega \mathcal{O}(X) e^{-\beta \mathcal{H}(X)}\, dX \tag{28}$$

is well-defined for arbitrary $\beta \in \mathbb{R}$ and $\gamma \in [0,1]$. Furthermore, $\langle \mathcal{O} \rangle$ depends continuously and even infinitely differentiably on β and γ.

Proof. Note first that Ω and Ω' differ only by a set of measure zero, as was remarked already above. Thus we can restrict the integration to Ω', and there $\mathcal{O}$, $\mathcal{H}$ and consequently the integrand are well-defined and continuous. Furthermore, recall the remark on page 16 that $\mathcal{Z} > 0$. Thus well-definedness is clear. For regularity, note that the observable $\mathcal{O}$ does not depend on β and is infinitely smooth in γ by Definition 8. The other dependences on these parameters are the exponential Boltzmann factor from π_T and $\mathcal{Z}$. Consider first the exponential factor: Its dependence on β is only the one explicit occurrence in (10) or (28). γ enters only as linear coefficient in (8) into $\mathcal{H}$, and it is clear that both of these are infinitely differentiable. For regularity of $\mathcal{Z}$, note that because of its definition in (9), the same argument (with $\mathcal{O} = 1$) applies. Then the result follows by the well-known result about differentiating an integral with respect to a parameter, see for instance [8, (8.11.2), p. 177]. □

The following quantities are all observables in the sense of Definition 8 (which is easy to see) and may be of particular interest. These are also the observables that will be calculated in the simulation results in Section 4.3:

[47] Note that $\pi_T(\Omega') = 1$ (or equivalently, its complement is a null set), and thus this restriction does not affect any of the following integrations or expectation values. See also page 16.

[48] One may want to generalise this definition so that observables need not be that regular in γ, but this is not necessary for my purposes. All the observables I will consider (see below for a list) won't depend on γ at all or if they do (as $\mathcal{H}$ does), are infinitely differentiable.

[49] See also [38, (7.8), p. 172].

Total Energy This is simply what was introduced already as energy of the system, namely $\mathcal{H}$ defined in (8). Since the behaviour of the system is formulated in terms of the energy in Definition 6, it is clear that this observable plays a key role.

Fair Energy In principle, energy or strain was introduced to describe personal (un)happiness. However, the total energy $\mathcal{H}$ does that only indirectly, after the individuals are weighted by their relative importances. This is of course crucial to the model considered, but as a consequence, $\mathcal{H}$ does not describe something like the total happiness of the society very well; the result is biased towards happiness of only certain individuals. Thus I believe that it makes also sense to introduce the *fair energy*:

$$\mathcal{H}_f(X) = \frac{1}{N}\sum_{n=1}^{N} f(a_n, l_n) = \mathcal{H}(X; \gamma = 1)$$

$\mathcal{H}_f$ allows us now to meaningfully quantify social welfare.

Output (GDP) From an economic point of view, the total production of the system is also very interesting. In my model, this is the "equal" value from the constraint (5):

$$\mathcal{Y}(X) = \sum_{n=1}^{N} a_n = \sum_{n=1}^{N} l_n$$

Thus $\mathcal{Y}$ is either the sum of all labour or the sum of all income. Confer also [5, p. 23f], namely that the GDP in classical economics can also be calculated either as sum of production or sum of incomes in an economy.

Gini Coefficient Since my work is mainly concerned with social and economic *fairness*, I also include a classical measure of inequality, namely the *Gini coefficient* [27, p. 771]. There are basically four quantities whose distribution differences could be used to measure inequality in my model: The labour (a_n), the income (l_n), the power (m_n) and the individual strains ($f(a_n, l_n)$). The strains would be most logical to describe fairness, but they have the problem that for them it is not really the case that some total, fixed amount of "happiness" is distributed among the individuals (which is the basic idea behind the Gini coefficient). From the others, I choose income to be the quantity upon which the Gini coefficient is based, because income and wealth are what the Gini coefficient usually measures (see [17], even its title) and because it is the only quantity where individual values are not constrained to a predefined interval ($[0, 1]$ in case of a as well as m). Then the Gini coefficient can be expressed like this:[50]

$$\mathcal{G}(X) = \frac{2\sum_{n=1}^{N} n l_{\sigma(n)}}{N\sum_{n=1}^{N} l_n} - \frac{N+1}{N}, \tag{29}$$

where $\sigma \in S(N)$[51] is a permutation that orders the individuals by their income, such that we get $l_{\sigma(n)} \leq l_{\sigma(n+1)}$ for $n = 1, \ldots, N-1$.

Uprising Finally one observable which is motivated from the physical framework, namely the generalised force [13, (8.2), p. 56] corresponding to the parameter γ:

$$\mathcal{U}(X) = \frac{\partial \mathcal{H}(X)}{\partial \gamma} = \sum_{n=1}^{N} \left(\frac{1}{N} - m_n\right) f(a_n, l_n) = \mathcal{H}_f(X) - \mathcal{H}(X; \gamma = 0) \tag{30}$$

I'll call $\mathcal{U}$ *uprising* in the following, since it loosely describes the resistance of the population with respect to changes of the political fairness parameter. The last form in (30) shows that the uprising also describes the difference in energy between the cases of total fairness ($\gamma = 1$) and $\gamma = 0$.

Lemma 27. *Let $\mathcal{O}$ be an observable. Then its dependence on the parameters is characterised by:*

$$\frac{\partial \langle \mathcal{O} \rangle}{\partial \beta} = -Cov(\mathcal{H}, \mathcal{O})\,^{52}, \quad \frac{\partial \langle \mathcal{O} \rangle}{\partial \gamma} = \left\langle \frac{\partial \mathcal{O}}{\partial \gamma} \right\rangle - \beta\, Cov(\mathcal{U}, \mathcal{O})$$

[50] See [44], but (29) is also not hard to derive from the definition of the Gini coefficient via the Lorenz curve directly.
[51] $S(N)$ denotes the symmetric group of permutations of the set $\{1, \ldots, N\} \subset \mathbb{N}$.
[52] For two random variables A and B, $\mathrm{Cov}(A, B)$ denotes their *covariance*.

Proof. This is straight-forward calculation, using again [8, (8.11.2), p. 177]. Note first that

$$\frac{\partial \mathcal{Z}}{\partial \beta} = \frac{\partial}{\partial \beta} \int_{\Omega'} e^{-\beta \mathcal{H}(X)} \, dX = - \int_{\Omega'} \mathcal{H}(X) e^{-\beta \mathcal{H}(X)} \, dX = -\mathcal{Z} \langle \mathcal{H} \rangle \tag{31}$$

and similarly for the parameter γ:

$$\frac{\partial \mathcal{Z}}{\partial \gamma} = -\beta \int_{\Omega'} \frac{\partial \mathcal{H}(X)}{\partial \gamma} e^{-\beta \mathcal{H}(X)} \, dX = -\beta \mathcal{Z} \langle \mathcal{U} \rangle \tag{32}$$

Furthermore, we know by Definition 8 that $\mathcal{O}$ itself may depend on γ but not on β explicitly. With (31), we then get for the first relation:

$$\begin{aligned} \frac{\partial \langle \mathcal{O} \rangle}{\partial \beta} &= \frac{\partial}{\partial \beta} \left(\frac{1}{\mathcal{Z}} \int_{\Omega'} \mathcal{O}(X) e^{-\beta \mathcal{H}(X)} \, dX \right) \\ &= \frac{\mathcal{Z} \langle H \rangle}{\mathcal{Z}^2} \int_{\Omega'} \mathcal{O}(X) e^{-\beta \mathcal{H}(X)} \, dX - \frac{1}{\mathcal{Z}} \int_{\Omega'} \mathcal{O}(X) \mathcal{H}(X) e^{-\beta \mathcal{H}(X)} \, dX \\ &= \langle \mathcal{H} \rangle \langle \mathcal{O} \rangle - \langle \mathcal{O}\mathcal{H} \rangle = -\mathrm{Cov}\left(\mathcal{H}, \mathcal{O}\right) \end{aligned}$$

For the second result, we similarly use (32):

$$\begin{aligned} \frac{\partial \langle \mathcal{O} \rangle}{\partial \gamma} &= \frac{\beta \mathcal{Z} \langle \mathcal{U} \rangle}{\mathcal{Z}^2} \int_{\Omega'} \mathcal{O}(X) e^{-\beta \mathcal{H}(X)} \, dX + \frac{1}{\mathcal{Z}} \int_{\Omega'} \left(\frac{\partial \mathcal{O}(X)}{\partial \gamma} - \beta \mathcal{O}(X) \frac{\partial \mathcal{H}(X)}{\partial \gamma} \right) e^{-\beta \mathcal{H}(X)} \, dX \\ &= \beta \langle \mathcal{U} \rangle \langle \mathcal{O} \rangle + \left\langle \frac{\partial \mathcal{O}}{\partial \gamma} \right\rangle - \beta \langle \mathcal{O}\mathcal{U} \rangle = \left\langle \frac{\partial \mathcal{O}}{\partial \gamma} \right\rangle - \beta \mathrm{Cov}\left(\mathcal{U}, \mathcal{O}\right) \end{aligned}$$

□

Lemma 28. *Let f be strictly convex.*[53] *$\mathrm{Var}(\mathcal{H})$*[54] *and $\mathrm{Var}(\mathcal{U})$ are both strictly positive for any $\beta > 0$ and $\gamma \in [0, 1]$.*

Proof. Note that $\mathrm{Var}\left(\mathcal{H}\right) = \left\langle \left(\mathcal{H} - \langle \mathcal{H} \rangle\right)^2 \right\rangle$ and $\mathcal{Z} > 0$, so it suffices to show

$$\mathcal{Z} \cdot \mathrm{Var}\left(\mathcal{H}\right) = \int_{\Omega'} \left(\mathcal{H}(X) - \langle \mathcal{H} \rangle\right)^2 e^{-\beta \mathcal{H}(X)} \, dX > 0.$$

Since the integrand is continuous and non-negative, it is enough to find one $X \in \Omega'$ where it is not zero. For this, it is sufficient if we have $X, Y \in \Omega'$ with $\mathcal{H}(X) \neq \mathcal{H}(Y)$, since then at least one of their energies differs from the energy expectation and because both are in Ω', the exponential factor is also non-zero. From Lemma 11 we know that there exist $a^*, a' \in [L, 1]$ with $f(a^*, a^*) < f(a', a') < \infty$: To see this, one can take a^* to be the unique minimum and a' to be any other point which is not on the boundary and thus also $f(a', a') < \infty$. Define $X = (x, \ldots, x)$ and $Y = (y, \ldots, y)$ with $x = \left(\frac{1}{N}, a^*, a^*\right)$ and $y = \left(\frac{1}{N}, a', a'\right)$. Then $X, Y \in \Omega'$ per construction, but also

$$\mathcal{H}(X) = f(a^*, a^*) \neq \mathcal{H}(Y) = f(a', a').$$

Similarly one can show that $\mathrm{Var}\left(\mathcal{U}\right) > 0$ with $X, Y \in \Omega'$ and $\mathcal{U}(X) \neq \mathcal{U}(Y)$. Here we don't even need strict convexity of f: Assume that $L < a_1 < a_0 < 1$, then $f(a_1, a_0) < f(a_0, a_1) < \infty$ by Definition 2. Choose $a' \in (L, 1)$ and note that

$$\begin{aligned} X &= ((1, a_1, a_0), (0, a_0, a_1), (0, a', a'), \ldots, (0, a', a')) \\ Y &= ((0, a_1, a_0), (1, a_0, a_1), (0, a', a'), \ldots, (0, a', a')) \end{aligned}$$

again define points in Ω'. From (30) it is easy to see that they furthermore satisfy $\mathcal{U}(X) > \mathcal{U}(Y)$, which completes the proof. □

[53] We have to rule out cases like $f(a, l) = a - l$ with $\gamma = 1$, in which case the energy is really *always* constantly zero and thus $\mathrm{Var}\left(\mathcal{H}\right) = 0$. This is not the most general assumption, but one we have already made above frequently and which is not a significant restriction.

[54] For a random variable A, $\mathrm{Var}\left(A\right)$ is its variance. Note that $\mathrm{Var}\left(A\right) = \mathrm{Cov}\left(A, A\right)$.

Theorem 17. *Let f be strictly convex. Then $\langle \mathcal{H} \rangle$ is strictly decreasing with respect to β (or strictly increasing with respect to the temperature T) and $\langle \mathcal{U} \rangle$ is strictly decreasing with respect to γ.*

Proof. This follows from Lemma 27 and Lemma 28, as then the covariances become variances and they are strictly positive. Thus the overall derivatives are always strictly negative. □

With respect to the energy, Theorem 17 is well-known from physics; think for instance of the relation $E = \frac{3}{2}kTN$ for an ideal gas, or [13, (9.19), p. 71] in general. The result for the uprising is also nice, since that means that the uprising decreases with increasing fairness in society—which is intuitive. In physical terms, this can be compared to the result that the pressure of a gas decreases with increasing volume, since pressure is the generalised force corresponding to volume [13, (8.8), p. 58].

4.2 The Monte-Carlo Simulation Method

The next question is how to actually calculate expectation values of observables according to (28). Note that analytical calculation is far too difficult, even the integral for $\mathcal{Z}$ alone in (9) is not tractable in general.[55] Thus we have to resort to numerical techniques, and in this section I want to describe the method I implemented and used with success for the results presented below in Section 4.3. It is a version of the famous Metropolis algorithm, which is successfully employed for example in theoretical physics on the topic of *Lattice QCD*. See for instance [15, sec. 4, p. 73ff], both for the method in general as used here and its application to this field. Since the method in general is well-known, I will describe it only briefly and focus instead on the particular details of my model and implementation below. A thorough discussion of the Metropolis algorithm can be found in the above-mentioned textbook or in other texts.

To approximate (28), we use *Monte-Carlo integration* [15, (4.3), p. 74]: Instead of integrating over the whole domain, which can in the Riemann-sense be understood as summing up infinitely many little "pieces", we randomly choose points and sum only those. If the points are well distributed over the domain and enough, this discrete sum will then be near the correct value of the integral. In particular for (28), if we manage to randomly generate points $X_k \in \Omega$, say $i = 1, \ldots, K$ with K large, according to the distribution induced by π_T, we can assume that

$$\langle \mathcal{O} \rangle \approx \frac{1}{K} \sum_{k=1}^{K} \mathcal{O}(X_k) \tag{33}$$

is a good approximation. The remaining difficulty is to *actually generate* points according to π_T. This is what the Metropolis algorithm is all about.

The basic idea is to generate a so-called *Markov chain* [15, 4.1.2, p. 75ff], which is in mathematical terms one trajectory of a *Markov process* in Ω. This means that we start with some $X_0 \in \Omega$, and then apply *update steps*, which stochastically change X_0 to some $X_1 \in \Omega$ and so forth. For the process to be Markovian, it must be possible to describe the update by a set of *transition probabilities*[56] $P(X \to Y)$. Let $P(X)$ denote the probabilities of points $X \in \Omega$ from the distribution we want to simulate (in our case π_T), then for the process to be in equilibrium, we need the *balance equation* [15, (4.12), p. 77]

$$\sum_Y P(Y \to X) P(Y) = \sum_Y P(X \to Y) P(X) \tag{34}$$

to hold for every $X \in \Omega$, so that X is neither a sink nor source of probability. A particular sufficient condition for (34) is the *detailed balance condition* [15, (4.15), p. 78]:

$$P(Y \to X) P(Y) = P(X \to Y) P(X) \tag{35}$$

for all $X, Y \in \Omega$. The Metropolis algorithm (Algorithm 1) describes how to perform update steps in practice, such that the resulting Markov process fulfils (35). Note that the details of how a candidate configuration is chosen in Line 1 will be discussed below; for simplicity, I assume we have the case of

[55] If $\mathcal{Z}$ were known, one can see that $\langle \mathcal{H} \rangle$ as well as $\langle \mathcal{U} \rangle$, at least, could be calculated as derivatives of $\mathcal{Z}$, without the need to solve also the integral (28) in addition. This is basically (31) and (32) above. See also [38, (7.26), p. 175] for $\langle \mathcal{H} \rangle$. This does not change the fact that $\mathcal{Z}$ itself is already too complicated, though.

[56] Since Ω is not discrete, in fact all "probabilities" in the discussion of the Markov chain are actually *probability densities*. Probability is however the term used in [15], so I stick to it.

symmetric a-priori selection probabilities [15, (4.18), p. 78].[57] The particular beauty of Algorithm 1 is that all we have to calculate in Line 2 is the *quotient* of two probabilities and not the probabilities themselves—in our case, this is

$$\frac{P(X')}{P(X)} = \frac{\pi_T(X')}{\pi_T(X)} = \frac{e^{-\beta\mathcal{H}(X')}}{e^{-\beta\mathcal{H}(X)}} = e^{\beta(\mathcal{H}(X)-\mathcal{H}(X'))} \tag{36}$$

where in particular $\mathcal{Z}$ drops out. If that wouldn't be the case and $P(X)$ directly required, we needed to know $\mathcal{Z}$ already to carry out the algorithm, which would defeat its whole purpose. Fortunately, (36) can actually be calculated easily in practice. All we need to know is the change in energy between the original and proposed configuration.

Algorithm 1 Update step in the Metropolis algorithm. See also [15, 4.1.3, p. 78f].

Require: $P(\cdot)$ is some probability density we want to simulate, $X_n \in \Omega$ is already given
Ensure: $X_{n+1} \in \Omega$ is generated such that the process satisfies (35) and thus (34)
1: randomly choose a candidate configuration X' based on X_n ▷ see Subsection 4.2.1
2: calculate $\rho \leftarrow \frac{P(X')}{P(X)}$ as in (36)
3: with the *acceptance probability* $\min(1, \rho)$ set $X_{n+1} \leftarrow X'$, otherwise reject X' and set $X_{n+1} \leftarrow X_n$

4.2.1 Generation of Candidate Configurations

One very important aspect in the implementation of the Metropolis algorithm as outlined above is the question, how candidate configurations in Line 1 of Algorithm 1 are generated. In our case, we have to somehow change the coordinates m, a and l of selected individuals randomly within their allowed intervals while keeping (5) and (6) satisfied. My first attempt at this was to select some single individual and then randomly change the coordinates; as a second step, scale the values of all (m_n), (a_n) or (l_n) by a common factor so that the equality constraints are again satisfied. This works in principle, but it turned out that it is very disadvantageous that for each single change, *the whole population* is affected by the scaling. Furthermore, it is not always easy to ensure that after scaling all individuals are within the allowed range; for instance, that not some a gets scaled beyond 1 or some l smaller than L. Thus, I think this "single individual" approach is not a suitable method.

Rather, in the update procedure I finally used, the equality constraints are taken more seriously. Instead of updates of single individuals, (most of the time) they are updated *in pairs*, where some quantity of power, labour or income is *exchanged* between the two partners. In that way, the equality constraints are not changed, and not the whole population but only two individuals have to be updated at once. The only draw-back with this method is that in the way I just described it, this would mean that the output of the economy can not be changed by any of those steps—which is clear, since only exchanging labour or income between two partners does neither alter $\sum_{n=1}^{N} a_n$ nor $\sum_{n=1}^{N} l_n$. After all, this was the whole point of ensuring (5) is automatically fulfilled after an update step! To mitigate this, let's introduce an additional procedure: With a certain chance, instead of doing a pair-exchange, a single individual is updated by increasing or decreasing *both* its labour a_n and its income l_n *simultaneously* by the same random amount. This also keeps (5) satisfied, but changes $\mathcal{Y}$ by that amount. One then only has to care about the amount which is exchanged, so that it is not too much and the values stay in the allowed range. But this time, this is simply a matter of reducing the interval from which the random value is chosen accordingly, and quite simple. Since there are now different types of updates described above, multiple of these update steps for each individual are combined into a *sweep* over the whole population. The procedure used for such an update sweep is given in details in Algorithm 2. The parameters $\mu, \delta, \alpha, \lambda$ and q are used to configure this algorithm.[58] Note that in the case of $\mu = \delta = \alpha = \lambda = 1$, one can see

[57] See page 40 below for a discussion of this matter. I think that my update in Algorithm 2 is too complicated to ensure this, but believe that in approximation, I can make this assumption without affecting the outcome of the calculation. It would probably also be very hard if not impossible to calculate the exact correction factors for [15, (4.16), p. 78].

[58] For the actual simulation, I chose $\mu = \delta = \alpha = \lambda = 0.3$ and $q = \frac{1}{2}$. This was however arbitrary, and since it gave satisfying results from the beginning, I did not take the trouble to try and benchmark different values. Thus these are probably not the best ones, but they did fine for me.

that this update procedure is *ergodic*[59] and also *symmetric* in the sense that we may both come from X to Y in one of the single updates, as well as from Y back to X: For the single-individual change in a and l, this is clear since the change can happen in both directions and for $\delta = 1$ to any value which is not strictly prohibited by the constraints. For the pair-exchange, we always give something *from n to k*, but during the whole sweep, the pair can also be matched together the other way round and thus the exchange happen in the other direction. With $\mu = \alpha = \lambda = 1$, again the reverse exchange is perfectly possible, since then everything allowed by the inequalities can actually be the outcome.

Algorithm 2 Update sweep over the whole population, this is the "fundamental unit" of the generation of configurations and thus at the heart of the Monte-Carlo simulation.

Require: $X \in \Omega$ is a given initial configuration, $P(\cdot)$ the probability distribution we want to simulate ▷ which is given in Definition 6
Require: parameters $\mu, \delta, \alpha, \lambda \in (0, 1]$, $q \in (0, 1)$ are fixed
Ensure: X is updated to a new configuration following the distribution $P(\cdot)$ and fulfilling (35)

for $n = 1 \to N$ **do**
  randomly choose partner individual k different from n
  randomly select $\Delta m \in [0, \mu \cdot m_n]$
  update $m_n' \leftarrow m_n - \Delta m$, $m_k' \leftarrow m_k + \Delta m$ ▷ give power over to partner
  propose changed state as per Algorithm 1

  if random chance of q applies **then** ▷ do single-individual update
    if random chance of 50 % applies **then** ▷ increase a and l
      set $\overline{a} \leftarrow 1 - a_n$ ▷ maximum increase is up to $a_n = 1$
    else ▷ decrease a and l
      set $\overline{a} \leftarrow \max(-a_n, L - l_n)$ ▷ maximum decrease is to $a_n = 0$ or $l_n = L$, note $\overline{a} \leq 0$
    end if
    randomly select Δa between 0 and $\delta \cdot \overline{a}$
    update $a_n' \leftarrow a_n + \Delta a$, $l_n' \leftarrow l_n + \Delta a$ ▷ increase or decrease a and l simultaneously
    propose changed state as per Algorithm 1
  else ▷ exchange labour and income between partners
    randomly choose partner individual k different from n ▷ possibly different from k before
    set $\overline{a} \leftarrow \min(a_n, 1 - a_k)$ ▷ maximum labour exchange is to $a_n = 0$ or $a_k = 1$
    randomly select $\Delta a \in [0, \alpha \cdot \overline{a}]$
    update $a_n' \leftarrow a_n - \Delta a$, $a_k' \leftarrow a_k + \Delta a$ ▷ give labour over to partner
    propose changed state as per Algorithm 1

    randomly choose partner individual k different from n ▷ possibly different from k before
    set $\overline{l} \leftarrow l_n - L$ ▷ maximum income exchange is to $l_n = L$
    randomly select $\Delta l \in [0, \lambda \cdot \overline{l}]$
    update $l_n' \leftarrow l_n - \Delta l$, $l_k' \leftarrow l_k + \Delta l$ ▷ give income over to partner
    propose changed state as per Algorithm 1
  end if
end for

4.2.2 Initialisation and Equilibration

Next, we have to talk about how to *initialise* the procedure: I discussed above how to generate a new configuration from an existing one (which led to Algorithm 2), but it is not yet clear where the starting point for this update sequence comes from. For a basic discussion of this topic in general, see [15, 4.4.1 and 4.4.2, p. 91f]. As before, I will describe now the specific details about my own implementation. For initialisation, my program chooses a particular admissible point $X_0 \in \Omega$:

$$X_0 = (x_0, \ldots, x_0), \quad x_0 = \left(\frac{1}{N}, a_0, a_0\right), \quad a_0 = \frac{L+1}{2} \in (L, 1) \tag{37}$$

[59] Which means that in only one sweep, it is possible (though maybe very unlikely) to reach any $Y \in \Omega$ starting from an arbitrary $X \in \Omega$. See [15, p. 77].

In some sense, X_0 from (37) can be regarded as being "in the centre" of the admissible set Ω. It has the particular advantage that it is a somewhat distinguished point, easy to calculate and admissible for all cases. However, I do not claim that it has something to do with a "typical" or correct configuration, whatever the parameters are. For small γ and low temperatures, where the system is expected to be near the unequal optimum of Theorem 5, this is surely not the correct point. And even for the equal situation of Theorem 4, there's nothing to tell us that the optimal a^* from Lemma 11 is near a_0 from (37)! Because of this, we have to bring the system into equilibrium after initialising it with (37)—by performing sufficiently many update sweeps as per Algorithm 2, but of course it is not easy to say what "sufficiently many" are. The basic observation here is that after the start, the measured values of observables will approach the correct value, while randomly fluctuating around some "trend" which converges towards the correct expectation value. That is, usually the measured values will increase or decrease steadily shortly after the initialisation (because the values are "very wrong" and there's a strong tendency to get them right), and when the system nears equilibration, the behaviour from update to update will get more and more like a random distribution around a fixed mean value—see Example 9 and in particular Figure 3 for how this looks like in practice. So my approach is to perform as many updates as is necessary to get a "flat" regression line for the last k measurements of some observable—this is formalised in Algorithm 3. To equilibrate an initial configuration chosen according to (37) in my actual calculations, I performed Algorithm 3 with respect to the energy $\mathcal{H}$ and subsequently with respect to uprising $\mathcal{U}$, with the parameters chosen as $k = 100$ and $\epsilon = 1 \cdot 10^{-6}$. The idea behind the double application is that the expected energy depends on β and the expected uprising on γ (see Theorem 17), so if both of them are equilibrated, then the system is likely adapted to both the chosen β and γ. I think that this approach and the parameters are good enough, but of course one has to be aware that there is not really a precise way to determine good equilibration and these are only heuristics!

Algorithm 3 My routine for "automatic" equilibration.

Require: X_0 is chosen according to (37), f, β and γ are given
Require: $k \in \mathbb{N}$, $\epsilon > 0$ are parameters, and $\mathcal{O}$ is some observable
Ensure: X is updated to some configuration which is assumed to be equilibrated such that measured values of $\mathcal{O}$ follow the exact distribution of (28) for the given β and γ
1: perform k updates according to Algorithm 2 to get configurations $X_1, \dots, X_k$
2: set $t \leftarrow k$
3: **repeat**
4: perform update $X_t \to X_{t+1}$ according to Algorithm 2
5: increment $t \leftarrow t + 1$
6: define k data points: $(1, \mathcal{O}(X_{t-k+1})), \dots, (k, \mathcal{O}(X_t))$
7: calculate mean value $\overline{\mathcal{O}}$ and incline σ' of the regression line for them
8: set $\sigma = \frac{\sigma'}{\overline{\mathcal{O}}}$ to be the "relative incline"
9: **until** $|\sigma| \leq \epsilon$

Example 9. I now want to conclude this discussion about equilibration with an example of actual equilibration runs from my calculations. This will hopefully make the description above clearer, but will also demonstrate some of the characteristics and problems with the approach I decided upon. In this example, I use the strain function from Example 1, and a population size of $N = 50$. As mentioned already above, the parameters were chosen as $k = 100$ and $\epsilon = 1 \cdot 10^{-6}$, which were also the parameters used throughout the actual calculations in Section 4.3. Then, Figure 3 shows the evolution of energy and uprising over "time" (corresponding to update sweeps, not real time of course) for two sets of parameters. The plots in the left column always show the measured values directly after initialisation, and the right column shows the values of *the same run* but after a double application of Algorithm 3 for equilibration. Since on the x-axis always the number of sweeps since initialisation is given, one can see there how many equilibration steps were taken until the exit condition applied.

In Figure 3a and Figure 3b, the evolution of the energy is shown for low temperature (high β) and $\gamma = 0.0$, corresponding to parameters that will force the system to be near the point of Theorem 5. Furthermore, the high β makes updates which increase the energy unlikely in Line 2 of Algorithm 1. This can be seen in that the energy curves have a clear downward trend, even after equilibration. However, in Figure 3b one also sees that from time to time, also updates are accepted that *increase* the energy—and

subsequently the downward movement continues. Note that Figure 3b visually still looks like there's a continued decrease and the system not yet in equilibration, but note that the shown time interval is not the one which satisfied the exit condition in Algorithm 3 but was measured *afterwards*; thus it may look like this by pure chance. Furthermore, as mentioned already above, with low temperature, there is *always* seemingly a downward trend in energy simply because energy increasing updates are suppressed by the Metropolis update, until from time to time some increase gets accepted nevertheless. Finally, note the y-axis' scaling in Figure 3b—the changes in energy there are only a fraction of the changes at the beginning in Figure 3a! But still, it can also very well be that the system is *really* not yet fully equilibrated and this introduces an error. The error should be small, however, since the changes in energy have become very small at least, even if they have still a downward trend. Producing a configuration which satisfies the condition in Line 9 of Algorithm 3, even if by chance, should always ensure that at least there's no longer a *strong* trend. Figure 3c and Figure 3d show *the same run*, but now the values of uprising. There, basically the same observations can be made, except that the value now *increases* away from 0 since the configurations get more and more unequal. Apart from that, the comments above apply again, and one can even see that the sudden energy increase in Figure 3b is accompanied by a decrease in uprising in Figure 3d at the same update sweep. Similarly, also the curves in Figure 3a and Figure 3c are clearly correlated and look like mirror images of each other to some degree.

The plots in Figure 3e and Figure 3f show the energy for another run, this time with smaller β (thus higher temperature) and higher γ. The higher temperature clearly leads to fluctuations which are much larger, and in addition, one can see in Figure 3e that the system is in equilibrium already shortly after the beginning, as it takes only about 10 sweeps to bring it into a state where the large fluctuations dominate over a clear trend. The reason for why my equilibration procedure still performed more than 6,000 sweeps is that it took that long until the fluctuations fulfilled the exit condition in Line 9 of Algorithm 3 *by chance*, since the required tolerance of $\epsilon = 1 \cdot 10^{-6}$ is much smaller than the fluctuations themselves for the chosen temperature.

4.2.3 Measurements and Decorrelation

Now, we are able to construct an initialised point with Algorithm 3 and then subsequently update it with Algorithm 2. However, since following configurations are constructed not independently but instead from an earlier one by the update sweeps, it is obvious that the configurations generated are not independent but *correlated* to each other. This is of course an unwanted effect, since for best results in the Monte-Carlo integration (33) we would like to have the samples stochastically independent. Note that correlation between samples does not introduce a bias in the expectation value, but only increases the statistical error [15, (4.63), p. 95]—thus one could in principle still use the correlated data points for (33). One can even do a detailed analysis of the error introduced by correlation between the samples, see for instance [15, 4.5.2, p. 94ff]—however, for simplicity (and since quantitative accuracy is not so important for me anyway[60]) my approach is different: I perform a number t_d of update sweeps for *decorrelation* between taking actual samples for the integration, and then assume that those samples are independent. In my calculations, I chose $t_d = 100$—see Example 10 for this value in particular and the discussion below about auto-correlation in general. Note that from Figure 5, one can see that most inaccuracies from the calculation are restricted to a lower-dimensional line in the parameter space, and thus one could probably make the simulation more efficient if the decorrelation time was chosen longer for this small region and much shorter for the rest. This was not implemented because of simplicity, but would be a possible route for tuning the algorithm further.

All that now remains for calculating the expectation values of some observables is to use (33) with generated and decorrelated configurations. Note that this is a procedure that is perfectly suited to parallel computing: If p threads are available, one can simply split the K measurements necessary for (33) into equal parts and let each of the threads generate only $\frac{K}{p}$ configurations. For this, one can initialise and equilibrate one configuration at the beginning, and then let each thread start its own Markov chain from that. In the end for the averaging in (33), the configurations generated by all threads are combined. Since usually the by far most expensive task in the simulation is generation of sample configurations,

[60] Mostly, I'm interested in the qualitative behaviour of the results in Section 4.3, and also want to demonstrate that in principle my model can be calculated numerically. Moreover, it would be difficult to get meaningful quantitative data anyway, since for that it is mandatory to have a strain function f that corresponds to reality—a task which seems hard to impossible to me.

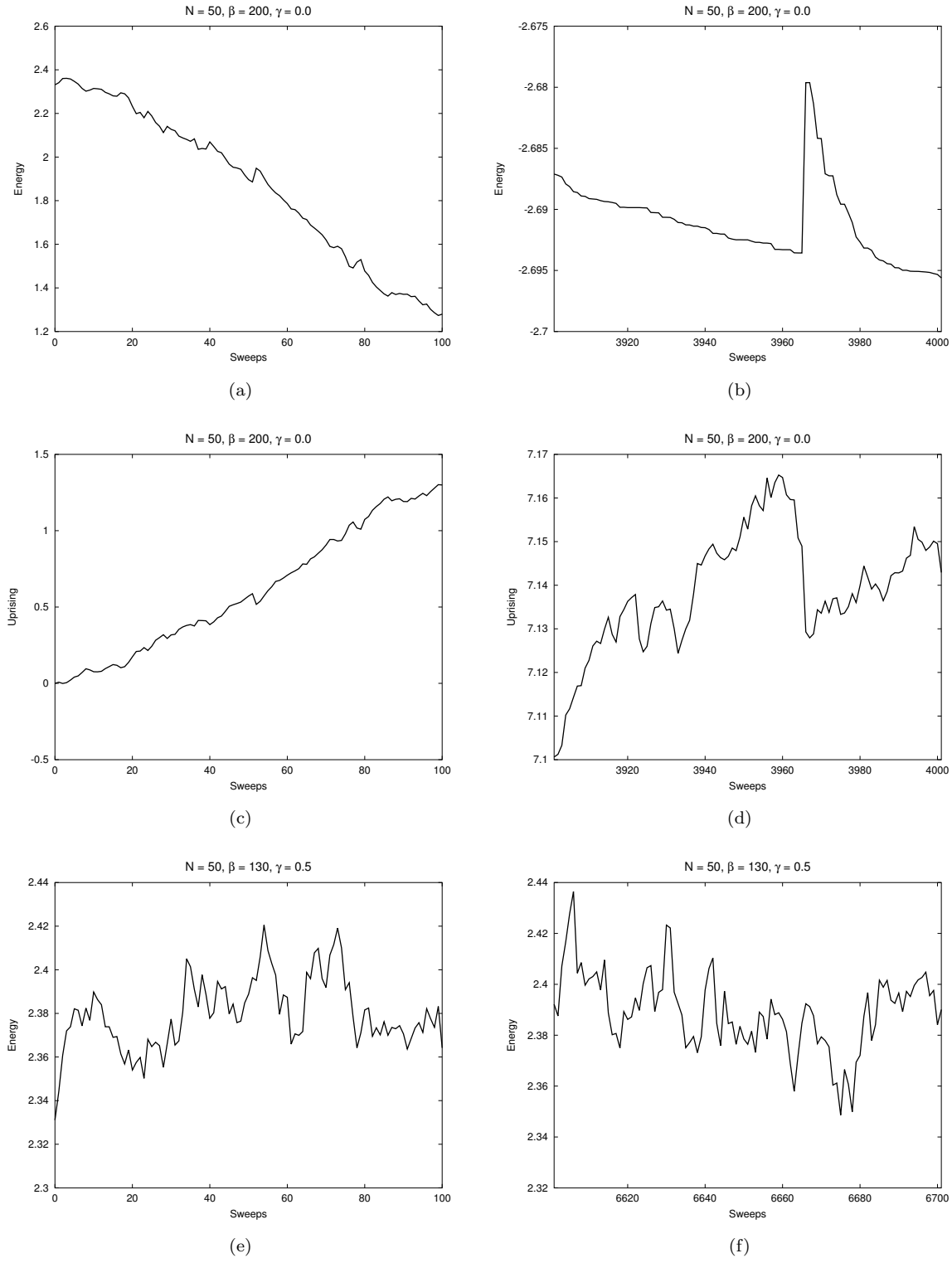


Figure 3: Evolution of observable values with update sweeps for Example 9. Left and right plots always show parts of the same run, but once before and once after equilibration was performed. Note the different scaling of the y-axes between left and right!

this should result in good scaling behaviour with the number of threads.[61] This whole procedure is now summarised in Algorithm 4; compare this also to [15, Table 4.1, p. 90].

Algorithm 4 The complete and parallelised Monte-Carlo simulation procedure to calculate observable expectation values.

Require: N, β, γ, f, t_d, K and p are given
Require: $\mathcal{O}$ is an observable we're interested in
Ensure: approximate $\langle \mathcal{O} \rangle$ for the given parameters
1: initialise $X \in \Omega$ as per (37)
2: apply Algorithm 3 to X twice, with respect to $\mathcal{H}$ and then with respect to $\mathcal{U}$
3: branch into p threads, with a local copy of X each, and execute p times in parallel:
4: **for** $t = 1 \to \frac{K}{p}$ **do**
5: perform t_d update sweeps as per Algorithm 2 on the local copy of X
6: measure $\mathcal{O}_t^{(i)} \leftarrow \mathcal{O}(X)$, where i is the thread number
7: **end for**
8: use (33) with $\mathcal{O}_t^{(i)}$, $i = 1, \dots, p$ and $t = 1, \dots, \frac{K}{p}$, to get the result for $\langle \mathcal{O} \rangle$

Before discussing my choice of $t_d = 100$, I have to introduce the *auto-correlation time* as a new concept, which describes roughly the "time" (meaning again update sweeps in the computation) needed to decorrelate configurations. To this end, we define the *auto-correlation function* $C_{\mathcal{O}}(t)$:

$$C_{\mathcal{O}}(t) = \operatorname{Cov}\left(\mathcal{O}(X), \mathcal{O}(X_{+t}\right) = \langle \mathcal{O}(X)\mathcal{O}(X_{+t}) \rangle - \langle \mathcal{O}(X) \rangle \langle \mathcal{O}(X_{+t}) \rangle = \langle \mathcal{O}(X)\mathcal{O}(X_{+t}) \rangle - \langle \mathcal{O}(X) \rangle^2 , \quad (38)$$

where $\mathcal{O}$ is some observable and X_{+t} denotes the configuration generated from X after t update sweeps in *the same Markov chain*. Note that $C_{\mathcal{O}}(0) = \operatorname{Var}(\mathcal{O})$ and in the following, we will use the *normalised* version of (38) instead, so that we can compare results for different observables between each other:

$$\Gamma_{\mathcal{O}}(t) = \frac{C_{\mathcal{O}}(t)}{C_{\mathcal{O}}(0)} \quad (39)$$

See also [15, (4.61), p. 94]. Usually $\Gamma_{\mathcal{O}}$ shows exponential decay with increasing time t, like

$$\Gamma_{\mathcal{O}}(t) = e^{-\frac{t}{\tau_{\text{exp}}}}.$$

τ_{exp} is the so-called *exponential auto-correlation time*. This is the interesting quantity, where we want $t_d \approx \tau_{\text{exp}}$ at least. To calculate $C_{\mathcal{O}}$, one can simply generate configurations and approximate the expectation values in (38) in the same way as before, namely via (33). Then, to calculate τ_{exp} from $C_{\mathcal{O}}$, note that

$$\tau_{\text{exp}} \approx \tau_{\text{int}} = \frac{1}{2} + \sum_{t=1}^{\infty} \Gamma_{\mathcal{O}}(t). \quad (40)$$

Here we introduced the *integrated auto-correlation time* τ_{int}, which can be easily calculated numerically as sum. See [15, (4.65), p. 95] for a discussion of this definition and why this approximates τ_{exp}.

Example 10. For a concrete example, choose $N = 20$ and the f from Example 1. As observables I chose the social space coordinates of some fixed individual, so that one can see how well they are "randomised" by updates. In particular,

$$\mathcal{O}_1(X) = m_1, \quad \mathcal{O}_2(X) = a_1, \quad \mathcal{O}_3(X) = l_1,$$

and then Γ_m, Γ_a and Γ_l defined accordingly per (39). The normalised auto-correlation functions as well as the integrated auto-correlation times are shown for different parameter values β and γ in Figure 4. For the approximation of expectation values in (38) in this calculation, $K = 500{,}000$ single measurements were made, and the sum in (40) was truncated after $t_{\max} = 1{,}000$ time steps, where all auto-correlations had already decayed to near zero. Figure 4b shows the same curves as Figure 4a in a logarithmic scale, where one can clearly see that the curves show nice exponential behaviour at least for the first 20 sweeps.

[61] See also Subsection 4.3.6 below, which partially confirms this claim. Furthermore, profiling for a particular example case also showed that well over 90 % of the total run-time was really spent in the update procedure.

Later on, the value is already so small that the random noise dominates. However, it can also be seen that especially for large β (which means low temperature and thus less fluctuations which can decorrelate configurations quickly) and small γ the behaviour can be not so nice, for instance in Figure 4e. Another demonstration of this issue can be found in Figure 7b below. I think my choice of $t_d = 100$ seems nevertheless justified, but one has to keep in mind that for certain parameter values this may increase the statistical error as discussed already above on page 42.

4.3 Simulation Results

Now, I finally want to present the results of the actual Monte-Carlo simulations that I performed according to the method described above in Section 4.2. It has to be admitted that my focus was on the general, qualitative behaviour of the model instead of quantitative accuracy and analysis in every detail. Thus, below I only present results that give a rough (but in my opinion very interesting) overview of the behaviour of my model. The analysis will be done in Subsection 4.3.1–Subsection 4.3.4 with $L = 0.1$ and the strain function from Example 1 since I believe this is the most realistic f from those discussed above, but I also performed basic calculations with other strain functions in Subsection 4.3.5 and the qualitative behaviour and in particular the phase transition discussed in Subsection 4.3.3 seems to be independent of the concrete form of f. Finally in Subsection 4.3.6 I give a short summary about the approximate algorithm performance and calculation times for the results presented.

4.3.1 Observable Expectation Values

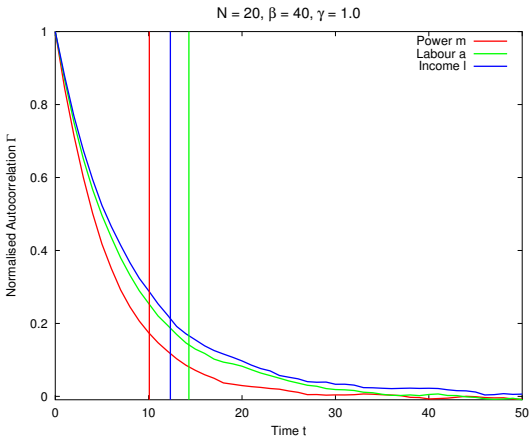
Let's begin with the expectation values of the observables listed above on page 36 for different parameters β and γ—this is plotted in Figure 5. There, one can see two things that were already derived analytically above: The monotonic behaviour of $\mathcal{H}$ and $\mathcal{U}$ from Theorem 17, and that for increasing β the energy approaches the global minimum further and further from Theorem 2. But maybe as a first surprise, one can easily see that the behaviour of the different observables is qualitatively very similar to each other, although the actual observables are not particularly related, at least not all of them. Of course, one can argue that energy and fair energy are related, or that uprising and Gini coefficient both describe the same phenomenon (social inequality). But I don't think one can easily explain why for instance output and the Gini coefficient should be related, since they describe "orthogonal" aspects. The only difference is that while energy decreases for more and more unequal situations ($\beta \to \infty$ and $\gamma \to 0$), which is clear by Theorem 17 and Theorem 14, the other four observables *increase* for this limit—this is also clear for $\mathcal{H}_f$, $\mathcal{G}$ and $\mathcal{U}$, since they actually measure inequality in some sense. But interestingly, also the GDP of the society grows with increasing unfairness—this is because in my model, the unprivileged individuals have to contribute more and more labour and thus the GDP is increased for the benefit of the privileged individual to their disadvantage; see also Theorem 15. Thus in my economic model, large GDP is actually *bad*—which is opposed to the usual notion of large GDP (per capita, but in my model N is a fixed constant anyway) corresponding to a higher standard of living [5, p. 26]. The reason for this is that because of social inequality and the effect power has in my model, the large GDP does not benefit the masses but only the powerful few.[62] Because of this similarity between the different observables, I will concentrate without loss of generality on the energy $\mathcal{H}$ as observable for all the following discussions. The second very interesting observation to make is that all of the expectation values show a *discontinuity*[63] along a certain curve in the (β, γ) parameter space—this striking phenomenon will be discussed on its own in Subsection 4.3.3.

4.3.2 Low Temperature Limit

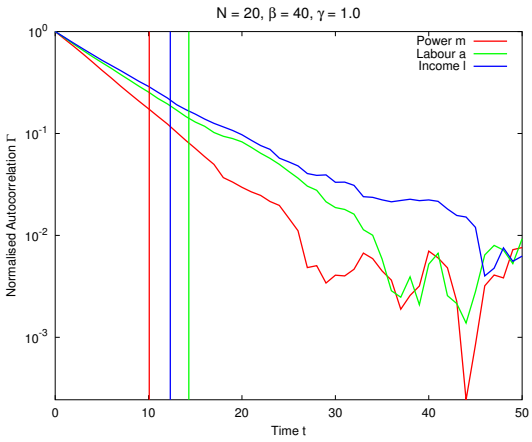
In the limit of low temperatures or $\beta \to \infty$, we know by Theorem 2 that the random configurations the system realises should approach more and more the global minimum which was excessively discussed in Chapter 3. This is also impressively demonstrated by Figure 6: In Figure 6a, different energy expectation curves (depending each on $\gamma \in [0, 1]$) are shown, where each curve is calculated with successively lower temperature, together with the curve calculated by minimisation of (22) with Newton's method as described on page 26. One can see very nicely that with $\beta \to \infty$ the curves approach the minimum curve

[62] Whether or not this argument also applies to real economies is something I want to leave open at this point.

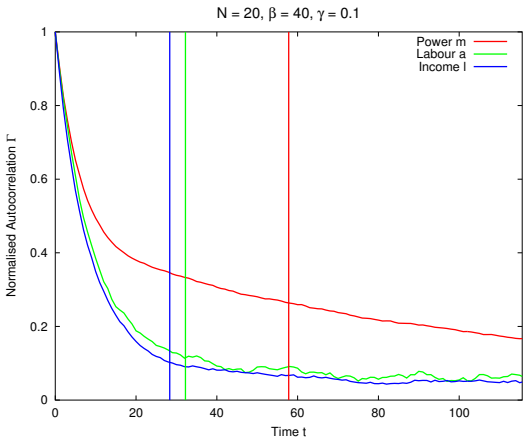
[63] In fact, this term is not exactly correct, but for reasons of clarity I'll stick to "discontinuity" below. See Subsection 4.3.4.



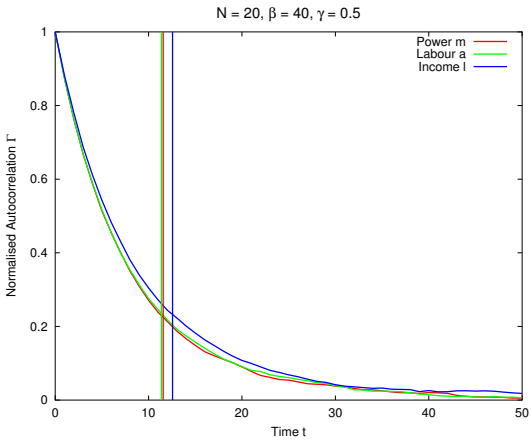
(a)



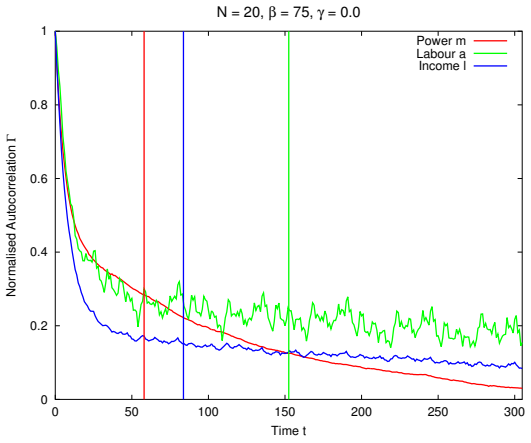
(b) Logarithmic scale shows exponential decay.



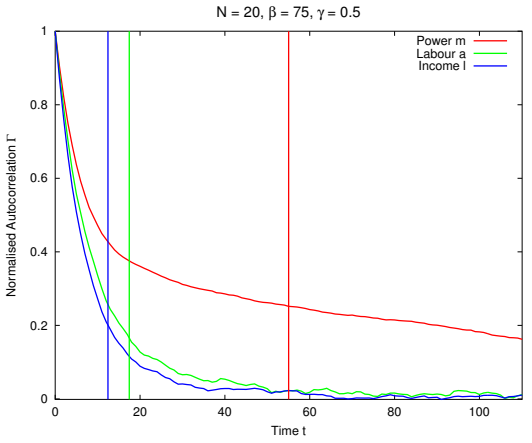
(c)



(d)

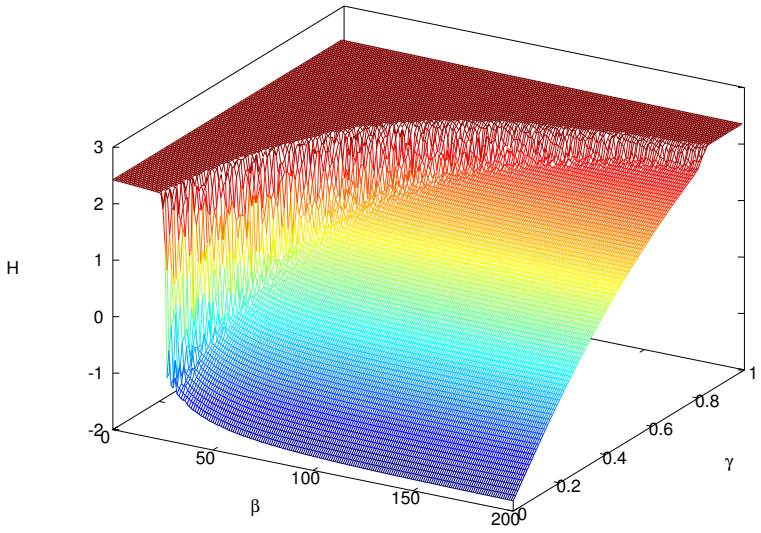


(e)

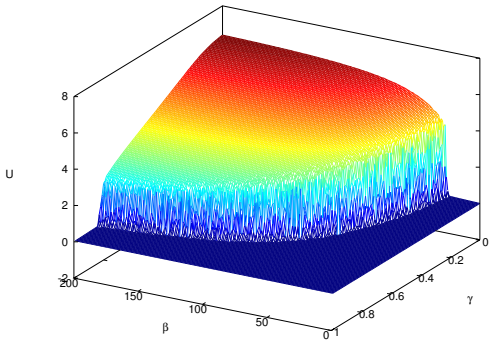


(f)

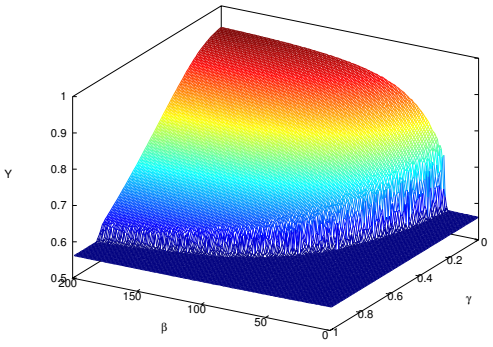
Figure 4: Normalised auto-correlation function Γ for different parameters, see Example 10. The vertical lines show the positions of the corresponding integrated auto-correlation times (40). Note the different scaling of the x-axes!



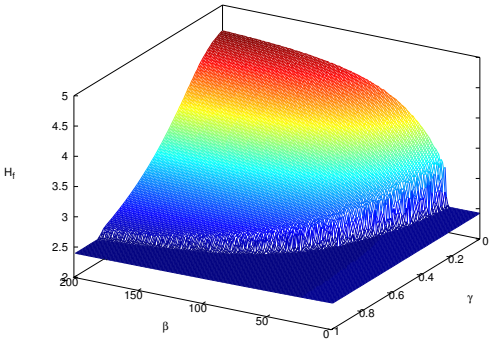
(a) Energy $\mathcal{H}$.



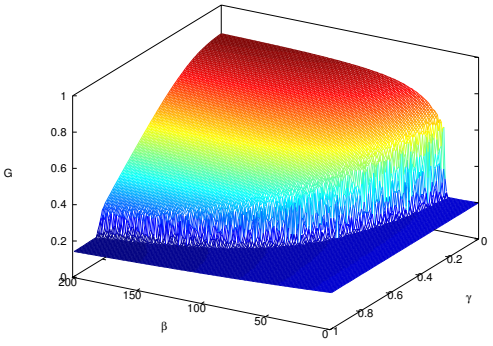
(b) Uprising $\mathcal{U}$.



(c) Output / GDP $\mathcal{Y}$.

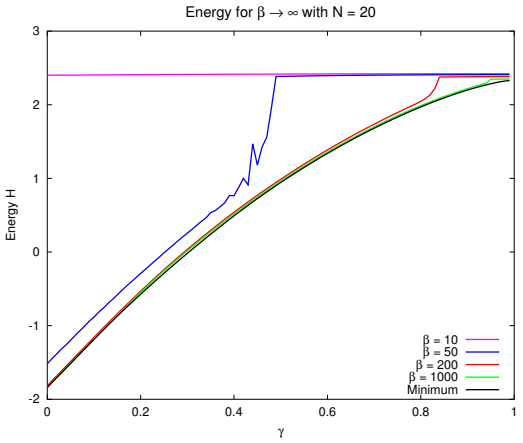


(d) Fair energy $\mathcal{H}_f$.

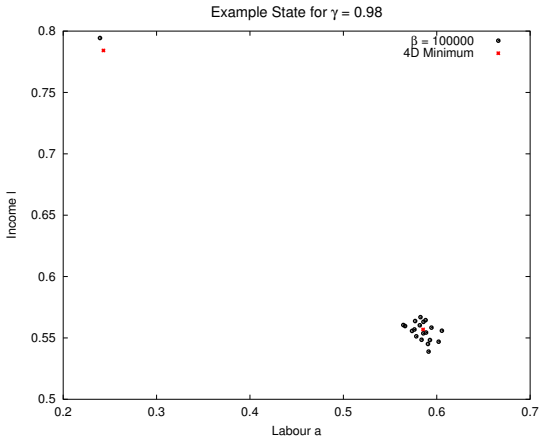


(e) Gini coefficient $\mathcal{G}$.

Figure 5: Expectation values of the observables from page 36 for $N = 20$ and different β, γ. For each expectation value, $K = 10{,}000$ was used in (33). Note that the smaller four pictures have a different viewing angle than the large plot of $\mathcal{H}$ at the top!



(a) Energy $\mathcal{H}$ with varying γ for lowering temperatures.



(b) Typical configuration in the (a, l) subspace.

Figure 6: Comparison of Monte-Carlo results for low temperature (high β) with results obtained by the "exact" minimisation as described on page 26. For the Monte-Carlo results in Figure 6a, $K = 50{,}000$ measurements were made in (33) for each γ.

from above, as they should do by Theorem 2 and Theorem 17. One can also see that the jump at the discontinuity in the curve decreases in height and moves with increasing β to the right, which matches the behaviour visible from Figure 5 and extends it for larger β than shown there. During this process, the curves approach the limit curve which is continuous as was predicted already by Theorem 13 (in combination with continuity of the energy itself). Figure 6b in addition shows a "typical" configuration (which was chosen fully randomly, even without requiring that it should match the expectation values or something like that) for $\beta = 100{,}000$ and $\gamma = 0.98$[64] together with the positions (a_1, l_1) and (a_0, l_0) expected for the privileged and unprivileged individuals from solving (22). There one can again see that Theorem 2 is fulfilled. Note that this—besides being interesting for its own sake—provides a further confirmation that my numerical methods work as expected, since the completely independent approaches of minimisation and the Monte-Carlo simulation produce indeed matching results for the cases where Theorem 2 allows us to compare them to each other.

4.3.3 Two Phases

The discontinuity visible in Figure 5 shows that the system exhibits a *phase transition*[65] along a specified curve in the (β, γ) parameter space—here I want to discuss this effect, since it is a very interesting finding. Note that for the analysis, I chose $N = 10$ as opposed to $N = 20$ which was used above very often. The reason for this is that the transition between both phases becomes more rapid with increasing N (see also Figure 6 below). With $N = 20$ and $\gamma = 0.0$, the transition happens somewhere between $\beta = 23.6761$ and $\beta = 23.6783$. It is thus much easier to demonstrate the basic properties for lower N. Additionally, $\gamma = 0.0$ was chosen to get the largest effect, and for the histograms shown below, 500,000 measurements were made each.

One finds that with the given parameters the phase transition happens at about $\beta = 16.3$. Figure 7 shows some results measured at exactly this critical point: In Figure 7a, a histogram[66] of energies for different configurations is shown—it forms two peaks, with about half of the values centring around 2.4 and the other half around 0.2. Compare these results to the histograms in Figure 8—this clearly shows that at the critical parameter values, two "types" of configurations coexist. I propose to call those two

[64] If β is further increased, the "cloud" of points becomes more and more concentrated. Similarly, if γ is decreased, the two centres themselves move further apart from each other. The concrete values of the parameters were simply chosen to produce a visually appealing result.

[65] See [38, 7.3, p. 184ff], [45] or [13, VI, p. 299ff].

[66] See [18, 15.11, p. 633ff] for a more sophisticated method related to the crude analysis I do here based on histograms.

phases *left* and *right*, respectively, corresponding to left and right sides of the jump in Figure 5a.[67] When the temperature is changed, one of those two types prevails—and the energy difference between them leads to the jump in the energy expectation value. In physical terms, this energy difference is *latent heat* [13, p. 172] that is stored or gets released at the phase transition. The existence of latent heat shows that we have a so-called *first-order* transition [45, p. 1]. For this phase transition, the uprising $\mathcal{U}$ can be used as order parameter (see [38, p. 184] or [13, p. 304f]), which changes at the phase transition from $\mathcal{U} = 0$ to $\mathcal{U} > 0$ as can be seen in Figure 5b or Figure 9b. If one generates a histogram as in Figure 7a but for $\mathcal{U}$ instead of $\mathcal{H}$, one also finds two peaks set apart from each other by a non-zero jump. Figure 7b shows the first 6,000 samples (each taken in a distance of t_d update sweeps already) used for the histogram. One can see also in this figure that there seem to be two types, and the system frequently jumps between both. Note that it also shows a tendency to remain in one of the two phases for some time, even though decorrelation updates have been performed—at the phase transition, decorrelation becomes more difficult and actually a larger t_d would have to be used; see also page 45. This effect is sometimes called *critical slowing down* [38, p. 190]. In the calculations for this subsection, I repeated the usual equilibration procedure (Line 2 of Algorithm 4) a hundred times in order to be sure that even near the critical temperature an equilibrium has been reached, which is especially important to find the exact critical parameter values. In Figure 7c and Figure 7d finally, two "typical" configurations belonging to each of the two phases is shown. They are actually *the same* configurations in both plots, only once projected to (m, l) and once to (a, l). There one can see what the distinction between the two phases is—the green dots corresponding to the left phase show permutation symmetry between the individuals, in other words, the society is equal except for random fluctuations. In the right phase depicted by red dots, this symmetry is clearly broken as a single individual is now privileged over the others. It has almost all the power and much higher income than any other individual—to compensate, the others have lower income than individuals in the left phase have. As a comparison, one can also produce histograms for other temperatures. This is shown in Figure 8: In Figure 8a, only the left phase is visible, Figure 8b and Figure 8c are already close to the transition temperature on either side but not as close as Figure 7, and in Figure 8d, only the right phase is present.

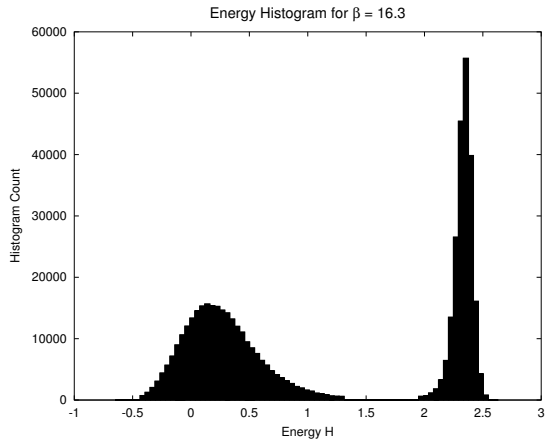
Note that depending on the phase, the individual positions in social space fluctuate either around one common centre or two centres (recall Theorem 8). Even for the right phase with one individual around the privileged position and all others around the (different) unprivileged one, the model never shows a *continuous* distribution of power or income. That is to say, societies described by it don't have a *middle class*. This is in contrast to classical theories of sociology, see for instance "Pareto's law" [17, p. 73] or [23, p. 21]. On the other hand, [23, p. 24] claims that in a global picture, there's a real trend towards polarisation into just two extremes without any kind of middle ground in-between. However, I believe that interpretation of my results in a real, sociological context is beyond the scope of the present work.

4.3.4 Infinite Volume Limit

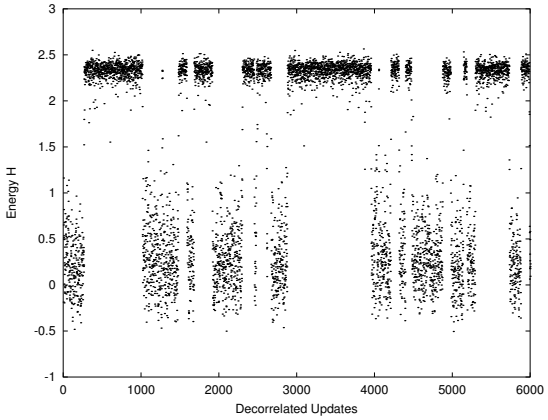
While above I always talked about "discontinuities" with respect to the phase transition described in Subsection 4.3.3, according to Lemma 26 the expectation values can't really be discontinuous. However, they behave *nearly discontinuously* (have a very steep incline but no real jump) at the critical parameter values. See also [13, p. 306] for a general discussion of this fact in thermodynamics. With growing N, this behaviour gets more and more extreme forming a real discontinuity for the limit $N \to \infty$. Here, I want to demonstrate this effect also with numerical results by comparing simulation results for different N with each other. But before I can do so, I have to remind that my energy $\mathcal{H}$ as defined in (8) is an *intensive* quantity and not an extensive one, as the energy usually is—this was already discussed above on page 13. It did not matter so far, but now for comparing different N with each other, it does. Therefore, I introduce a new "scaled" inverse temperature $\tau = \frac{\beta}{N}$. Consequently,

$$\langle \mathcal{O} \rangle_\tau (\tau) = \langle \mathcal{O} \rangle_\beta (N\tau) = \frac{1}{\mathcal{Z}} \int_\Omega \mathcal{O}(X) e^{-\tau N \mathcal{H}(X)} \, dX \tag{41}$$

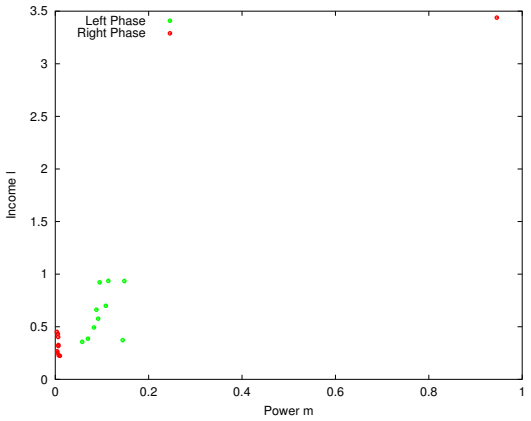
[67] In a discussion, someone suggested to call the phases "communist" and "capitalist" for obvious reasons. Since in practice however, real communist systems did not achieve the desired equality (see page 12), I don't think this is a good term. Note that the chosen names "left" and "right" also coincide with the usual associations of political ideas to either of those sides.



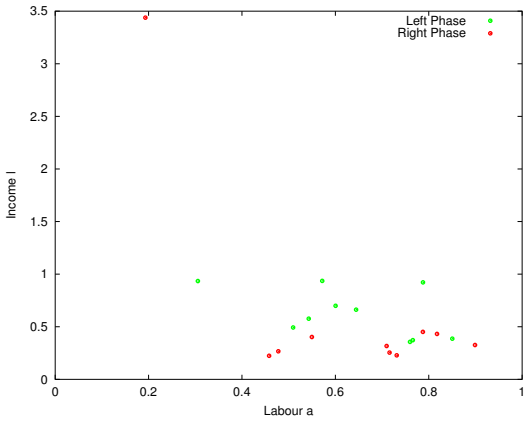
(a) Histogram of energies with two peaks of about equal mass.



(b) Time evolution of the first data values.

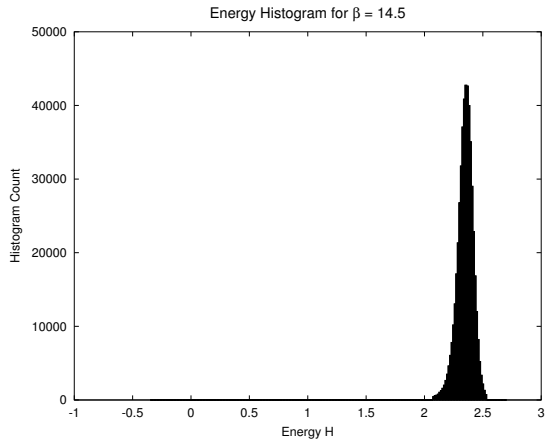


(c) Typical configurations for both phases.

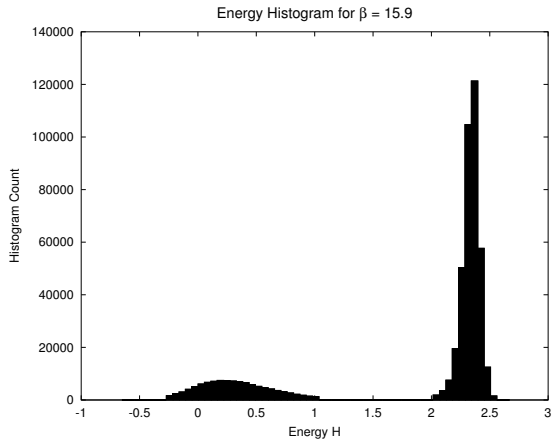


(d) Typical configurations for both phases.

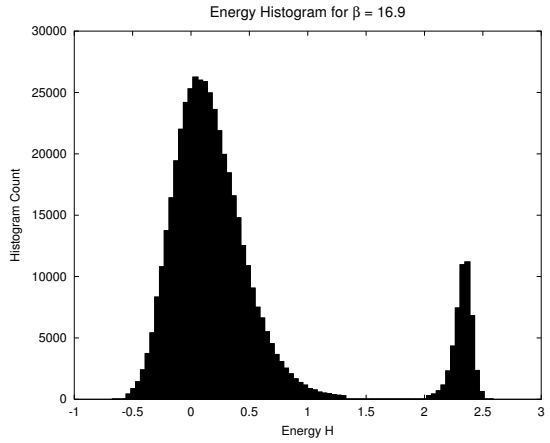
Figure 7: Behaviour of the model for $N = 10$, $\gamma = 0.0$ and $\beta = 16.3$, which is near the critical temperature. Note that single steps on the x-axis in Figure 7b correspond to t_d updates (i. e., already "decorrelated" samples)! All four plots were produced from the same run, especially the top two ones are based on the same data. Also the bottom two plots show the same two configurations, but projected onto different coordinates of the social space.



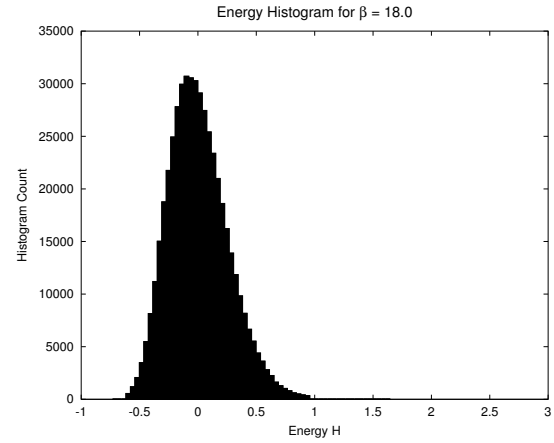
(a) Only left phase.



(b) Mostly left phase.

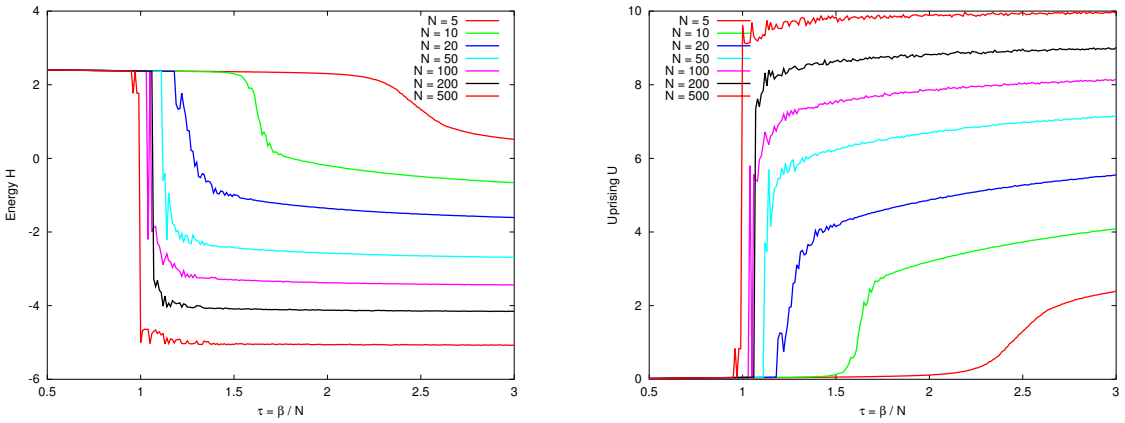


(c) Mostly right phase.



(d) Only right phase.

Figure 8: Energy histograms for $N = 10$, $\gamma = 0.0$ and temperatures not exactly at the phase transition. Note that the y-axes are scaled differently between the plots, but also the individual histogram bars have different width and thus the counts different meanings. In total, in each histogram there's the same number of data points (namely 500,000) distributed between the bins.



(a) Energy $\mathcal{H}$.

(b) Uprising $\mathcal{U}$.

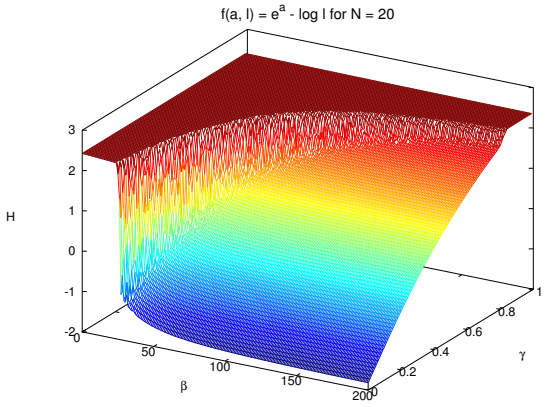
Figure 9: Jump in the observables for $N \to \infty$ with $\gamma = 0.0$ and different $\tau = \frac{\beta}{N}$. $K = 50{,}000$ was used.

is the analogue of (28) expressed in terms of τ instead of β. If we fix τ, it is clear that now the energy appears in (41) only scaled by N and thus can be considered to be "effectively extensive". After this transformation, Figure 9 shows the behaviour of $\langle \mathcal{H} \rangle_\tau$ and $\langle \mathcal{U} \rangle_\tau$ for $\gamma = 0.0$[68] around the critical temperature for different N. One can nicely see that the discontinuity gets more extreme and happens at a well-defined τ (but not β) for $N \to \infty$. Note that in Figure 9, it looks as if the jump would happen precisely at $\tau = 1$ in the limit—I don't know if this is purely by chance or has some reason, but note that for instance with the strain function from (2) it happens decidedly a little *before* $\tau = 1$ for $N = 50$ and $N = 100$, and thus this can not be some unknown general principle (although even for that case, it is at least *near* $\tau = 1$). For $N = 20$, it can also be seen from Figure 10 that the discontinuity for $\gamma = 0.0$ is at (very) different β between the different strain functions.

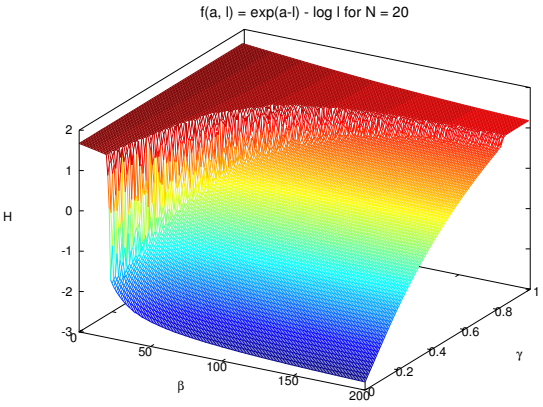
4.3.5 Other Strain Functions

So far, I have only presented calculation results for f from (1) first introduced in Example 1. This is because I believe that of all the example strain functions given in Section 2.1, this is the "most realistic" one. However, since I also don't have strong arguments in favour of this claim, I now want to demonstrate that the choice of f does in fact not too much influence the macroscopic and qualitative behaviour of the model. This is evident from Figure 10: Figure 10a is the same plot as Figure 5a, and Figure 10b–Figure 10d show the same quantity (namely the energy expectation $\langle \mathcal{H} \rangle$) for the other strain functions. One can see that they differ of course in details, but their overall behaviour is the same—and in particular, all of them show the phase transition discussed above in Subsection 4.3.3. While the four strain functions shown are of course still only particular examples, I believe that they differ quite a bit in their structure—and hence these findings allow me to claim that the actual form of f really seems to be not so important if only the basic behaviour is of interest. Note, though, that of course it *is* important if one is interested in the precise numeric values of the observables—for instance as will be the case in Section 4.4. Similarly, also the actual form of configurations *can be* very different for the different strain functions, as we've seen for instance in Example 6. In relation to this example, one finds that also random configurations generated by the Monte-Carlo method have, in accordance to the theory, $a_i \approx 1$ for *all* $i = 1, \ldots, N$ (including the privileged one) with the strain function (2). But the *qualitative* behaviour, especially with respect to the observed phase transition, is still similar for the different strain functions.

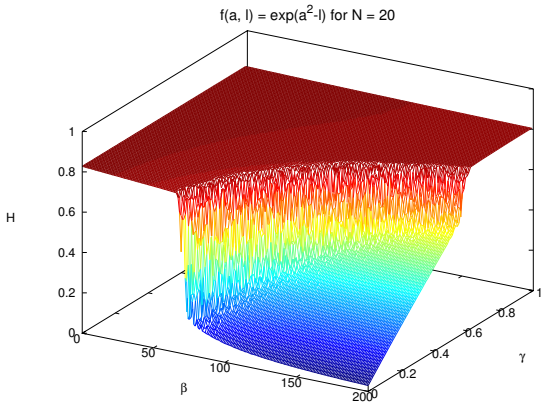
[68] For $\gamma = 0.0$ the jump is largest and thus the effect best visible, but also other values of γ show the same behaviour.



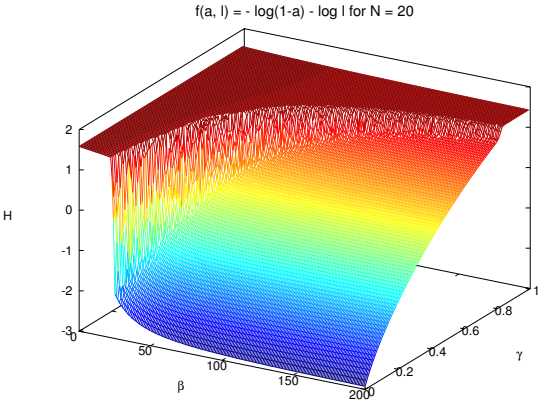
(a) (1), $L = 0.1$.



(b) (2), $L = 0.1$.



(c) (3), $L = 0.1$.



(d) (4), $L = 0.0$.

Figure 10: Comparison of the energy expectation $\langle \mathcal{H} \rangle$ for different strain functions. In all cases, $N = 20$ and $K = 10{,}000$ were chosen.

	CPU	Cores	RAM	Operating System	GCC	Octave
`laptop`	Intel Core i3	2+2[69]	4 GiB	Debian 7 "Wheezy"	4.6.3	3.2.4
`dualhex`	2x AMD Opteron	12	[70]28 GiB	Ubuntu 12.04 LTS	4.6.3	3.2.4

Table 2: Details of the computers used for numerical simulations.

4.3.6 Calculation Times

Finally, I also want to give some "meta-data" about the results presented above, namely the actual performance of my simulation routines on the machines used to find what was shown in this section before. Table 2 lists the characteristics of the two machines used for calculations, where `laptop` is my private notebook and `dualhex` [14] a computer of the Institut für Mathematik und Wissenschaftliches Rechnen, University of Graz. Table 3 lists the calculation times that were necessary to produce the numerical results included in this section, but note that since I don't consider this to be an important aspect, they should be taken only as very approximate since the machines were partially under load by other processes while measuring and the timings were produced by a single run and are not statistical averages. The Monte-Carlo method of Section 4.2 was implemented in C++11 [43] with parallelisation via OpenMP [3], compiled by GCC [1]. The raw data was then post-processed to produce the figures with GNU Octave [2]. The results indicate that the performance is dominated by the time needed for update sweeps via Algorithm 2, as was already claimed above on page 44. This follows from the fact that the time necessary for one update sweep and per N, as shown in the last column of Table 3, is roughly constant for a single machine with fixed number of threads but over different tasks, different numbers of update sweeps and also to some degree different N. Note that by Algorithm 2, the time for one update sweep should scale linearly with N, thus it makes sense to give the time also per N and not only per update sweep. This of course only gives a measure of performance on a low level. For a high-level estimate, it would be more interesting to consider the run-time necessary to get results with statistical error smaller than some threshold—that would then also measure performance of equilibration and decorrelation properties of the chosen update method. Here I didn't do that, though, as I only worked with a single update algorithm and rather simplistic equilibration and decorrelation techniques. For the results corresponding to Figure 10, it should be noted that in this case because of different f also the calculation effort was different, since f appears at the very heart of the algorithm and has to be evaluated very often. Note that the times given are *CPU user time* as reported by `time`, and thus this time already accounts for multiple threads running in parallel.

Nevertheless, the times reported in Table 3 differ for different numbers of threads, which indicates that efficiency is hurt by parallelisation, although the basic procedure in Algorithm 4 seems relatively easy to parallelise. In order to investigate this further, Table 4 shows the real times corresponding to the runs for task Subsection 4.3.3. However, be aware that these times are still not taken very carefully and the warnings above apply again! Since the number of update sweeps was not exactly but nearly the same for all of them, it seems fair to compare them directly to each other. The results for `laptop` look plausible, where two threads give a good although not perfect speed-up; when one takes into account that—as mentioned in the footnote—the task considered is especially unsuited for parallelisation because of extended equilibration, this seems like a good result. Four threads improve it further, but not nearly

[69] Two physical cores with Intel's hyper-threading [6], posing as four SMP cores to the operating system.

[70] Note that this value differs from [14] which states 32 GiB, since one 4 GiB unit broke down since the report.

[71] This is only a rough indication of what exactly was done in each calculation, since most of the time, there were more things calculated (e. g., for different parameter values) than actually shown in the end.

[72] Because of causal dependence between earlier and later data points in the Markov chains, the time-series can not be calculated in parallel (at least not by the simple parallelisation method described in Algorithm 4).

[73] Binary search to find the critical temperature for given γ, leading to the results for N = 20 mentioned on page 48.

[74] This task was chosen for comparison of different numbers of threads because it has a relatively small run-time. Note however that especially for the histogram, the equilibration time was higher than for the other tasks (as described on page 49), and the equilibration was only performed with a single thread in every case.

[75] This is slightly different to the update count in the line above, because of the randomness inherent in the Monte-Carlo simulation (since a different number of updates was performed in automatic equilibration described in Subsection 4.2.2). While my code can give reproducible results when the random-number generator is seeded correctly, this only works when the number of threads is the same, too.

Task[71]	**Machine**	**Threads**	N	**Updates** K	**Time** t	t/KN
Figure 3	`laptop`	[72]1	10	22,524		
			50	77,224		
					5.5 s	1.3 ms
Figure 4	`dualhex`	12	20	4,004,255,206	354,000 s	4.4 ms
Search Subsection 4.3.3[73]	`dualhex`	12	20	966,221,607	80,010 s	4.1 ms
Subsection 4.3.3[74]	`dualhex`	12	10	258,213,229	11,715 s	4.5 ms
		8	10	[75]258,213,230	14,115 s	5.5 ms
		6	10	258,213,228	10,630 s	4.1 ms
		4	10	258,213,228	6,975 s	2.7 ms
		2	10	258,213,230	5,010 s	1.9 ms
		1	10	258,213,230	3,840 s	1.5 ms
	`laptop`	4	10	258,213,228	5,055 s	2.0 ms
		2	10	258,213,230	3,720 s	1.4 ms
		1	10	258,213,230	3,285 s	1.3 ms
Figure 6	`dualhex`	12	20	6,011,888,472	564,240 s	4.7 ms
Parts of Figure 9	`dualhex`	12	5	1,260,491,664		
			10	1,258,451,343		
			20	1,256,855,474		
			50	1,256,131,500		
					532,140 s	5.0 ms
Parts of Figure 9	`dualhex`	12	100	1,256,132,888		
			200	1,256,123,531		
					2,873,145 s	7.6 ms
Parts of Figure 9	`dualhex`	12	500	1,256,254,802	4,788,570 s	7.6 ms
Figure 5	`dualhex`	12	20	20,370,566,491	1,845,210 s	4.5 ms
Figure 10b	`dualhex`	12	20	20,349,102,497	1,125,240 s	2.8 ms
Figure 10c	`dualhex`	12	20	20,532,987,101	993,780 s	2.4 ms
Figure 10d	`dualhex`	12	20	20,347,408,014	1,795,680 s	4.4 ms

Table 3: Timings for the calculation results presented above.

Machine	Threads	Time	Speed-Up[76]
`dualhex`	12	1,380 s	2.79
	8	1,995 s	1.93
	6	[77]1,995 s	1.93
	4	1,920 s	2.01
	2	2,595 s	1.48
	1	3,850 s	1.00
`laptop`	4	1,395 s	2.37
	2	1,935 s	1.71
	1	3,300 s	1.00

Table 4: Real times for runs of Subsection 4.3.3 in Table 3 with different numbers of threads as well as the scaling efficiency derived from them.

as much as they should for good scaling—however, keep in mind that `laptop` only has two physical cores, and it is not given that hyper-threading works very well for the problem at hand! On the other hand, the timings for `dualhex` look quite odd in both tables—especially that around 6 and 8 threads the times are not even monotone. Thus, I believe there must be something foul with those. It could be for instance that this machine, since it is a shared computer and I don't have root privileges on it, throttles the maximum allowed CPU usage at certain times or under certain conditions. But since run-time performance is only a very minor aspect in my considerations, I'll leave it like that.

4.4 Parameter Estimation

Finally, I want to describe a method for estimating the parameters β and γ of a real system under the assumptions that it follows my model, that f, L and N are fixed, and that somehow a particular sample configuration $X \in \Omega$ has been measured. I'm not sure if this is really of practical relevance, since measuring the state of a society in Ω is probably quite tricky to impossible (see also the remark on page 60). Nevertheless, I believe that these discussions will be interesting at least from a theoretical point of view.

Theorem 18. *Assume that β and γ can be estimated via the maximum-likelihood method[78]. Then the estimates are such that*

$$\langle \mathcal{H} \rangle = \mathcal{H}(X), \quad \langle \mathcal{U} \rangle = \mathcal{U}(X), \tag{42}$$

where $X \in \Omega$ is the measured configuration for which estimation is performed.

Proof. If β and γ are the maximum-likelihood estimations, the likelihood

$$\mathcal{L}(\beta, \gamma) = \frac{e^{-\beta \mathcal{H}(X)}}{\mathcal{Z}}$$

is per definition maximal for them. Thus the first-order necessary optimality condition implies that $0 = \frac{\partial \mathcal{L}}{\partial \beta} = \frac{\partial \mathcal{L}}{\partial \gamma}$ must hold. We can calculate those derivatives as done above, using again [8, (8.11.2), p. 177] and also (31) as well as (32):

$$\frac{\partial \mathcal{L}}{\partial \beta} = \frac{-\mathcal{H}(X) e^{-\beta \mathcal{H}(X)}}{\mathcal{Z}} - \frac{e^{-\beta \mathcal{H}(X)}}{\mathcal{Z}^2} \frac{\partial \mathcal{Z}}{\partial \beta} = \frac{e^{-\beta \mathcal{H}(X)}}{\mathcal{Z}} \left(\langle \mathcal{H} \rangle - \mathcal{H}(X) \right) \tag{43}$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = \frac{-\beta e^{-\beta \mathcal{H}(X)}}{\mathcal{Z}} \frac{\partial \mathcal{H}(X)}{\partial \gamma} - \frac{e^{-\beta \mathcal{H}(X)}}{\mathcal{Z}^2} \frac{\partial \mathcal{Z}}{\partial \gamma} = \frac{\beta e^{-\beta \mathcal{H}(X)}}{\mathcal{Z}} \left(\langle \mathcal{U} \rangle - \mathcal{U}(X) \right) \tag{44}$$

With this, (42) follows immediately. □

[76] This is the ratio of real time of the single-threaded run to that of the multi-threaded one, and should equal the number of threads for perfect scaling.

[77] It is *not* a mistake that (approximately) the same value appears here again!

[78] See for instance [7, 33.2, p. 498ff].

Theorem 18 formally justifies a heuristic idea one could have for estimating the parameters: Since $\langle\mathcal{H}\rangle$ is monotone in β by Theorem 17 and $\mathcal{H}(X)$ is the best guess we have for the true $\langle\mathcal{H}\rangle$, it seems plausible to adjust β in order to get $\langle\mathcal{H}\rangle = \mathcal{H}(X)$. The same goes for $\mathcal{U}$ and γ. Nevertheless, it is not guaranteed that (42) can always be solved, and not even, that a maximum-likelihood estimator exists—Lemma 29 gives some thoughts about concavity of $\mathcal{L}$ and thus existence of a maximum, but I was not able to find a definite result. Furthermore, since $\mathcal{H}$ (but not $\mathcal{U}$) depends explicitly on γ, we have to recalculate $\mathcal{H}(X)$ each time γ is adjusted! But the best we can do for now is assume that it all works, then a possible approach for actual calculation of β and γ from given $X \in \Omega$ is to simulate configurations, check $\langle\mathcal{H}\rangle$ and $\langle\mathcal{U}\rangle$ against $\mathcal{H}(X)$ and $\mathcal{U}(X)$, and adapt β and γ accordingly in order to minimise a function like

$$\Delta(\beta,\gamma) = |\langle\mathcal{H}\rangle - \mathcal{H}(X)| + |\langle\mathcal{U}\rangle - \mathcal{U}(X)| \tag{45}$$

with some method of choice, since this corresponds to solving (42). This will be discussed a little further below in Example 11. Or to give another idea, one could fix either β or γ and in turn do a binary search on the then-monotone function $\langle\mathcal{U}\rangle$ or $\langle\mathcal{H}\rangle$, respectively, and iterate this process.

Lemma 29. *If (42) holds, the Hessian of $\mathcal{L}$ with respect to β and γ is:*

$$\nabla^2\mathcal{L} = -\frac{e^{-\beta\mathcal{H}(X)}}{\mathcal{Z}} \begin{pmatrix} \mathit{Var}(\mathcal{H}) & \beta\,\mathit{Cov}(\mathcal{H},\mathcal{U}) \\ \beta\,\mathit{Cov}(\mathcal{H},\mathcal{U}) & \beta^2\,\mathit{Var}(\mathcal{U}) \end{pmatrix}$$

Thus $\mathcal{L}$ is strictly concave if this matrix is negative-definite, or equivalently by the minor criterion:

$$\mathit{Var}(\mathcal{H})\ \mathit{Var}(\mathcal{U}) > \mathit{Cov}(\mathcal{H},\mathcal{U})^2.$$

Proof. This is again straight-forward calculation similar to the proof of Theorem 18. If we use (43) and (44), differentiate once more and apply the product rule, note that by (42) only derivatives of the parenthesised term contribute since this term vanishes when left alone. Then the result follows easily with Lemma 27. □

Example 11. Here I want to illustrate the idea of minimising (45) a little further. For that, I generated a single random configuration with $N = 20$, $\beta = 150$, $\gamma = 0.5$, the usual f from (1) and $L = 0.1$. The parameters are chosen to get a point in the right phase, since it is clear from Figure 5 that it would be nearly impossible to approximate the point from $\mathcal{H}$ and $\mathcal{U}$ in the left phase, where both have flat expectation values. Figure 11 shows Δ for this configuration based on the data used also in Figure 5 for the expectation values. One can see that a valley with almost zero error forms, which is clear since $\langle\mathcal{H}\rangle$ and $\langle\mathcal{U}\rangle$ have similar shape already in Figure 5 and so parameters along a curve in the (β,γ) space all fulfil (42) to some degree. This makes it of course difficult to find the exact point, but the actual accuracy possible depends probably on the calculation time available to calculate expectation values precisely. With the data I had available already from Figure 5, the minimal Δ is found at $\beta = 141$ and $\gamma = 0.5$ and is $\Delta(141, 0.5) = 9.5 \cdot 10^{-4}$. This is not too far off the actual parameters of $\beta = 150$ and $\gamma = 0.5$, where $\Delta(150, 0.5) = 3.2 \cdot 10^{-2}$. Note that the estimation can of course never be exact, since already the random configuration chosen at the beginning was very likely not exactly at the expectation values. With larger N, however, statistical mechanics tells us that the effect of this fluctuation should become smaller [13, (22.20), p. 187].

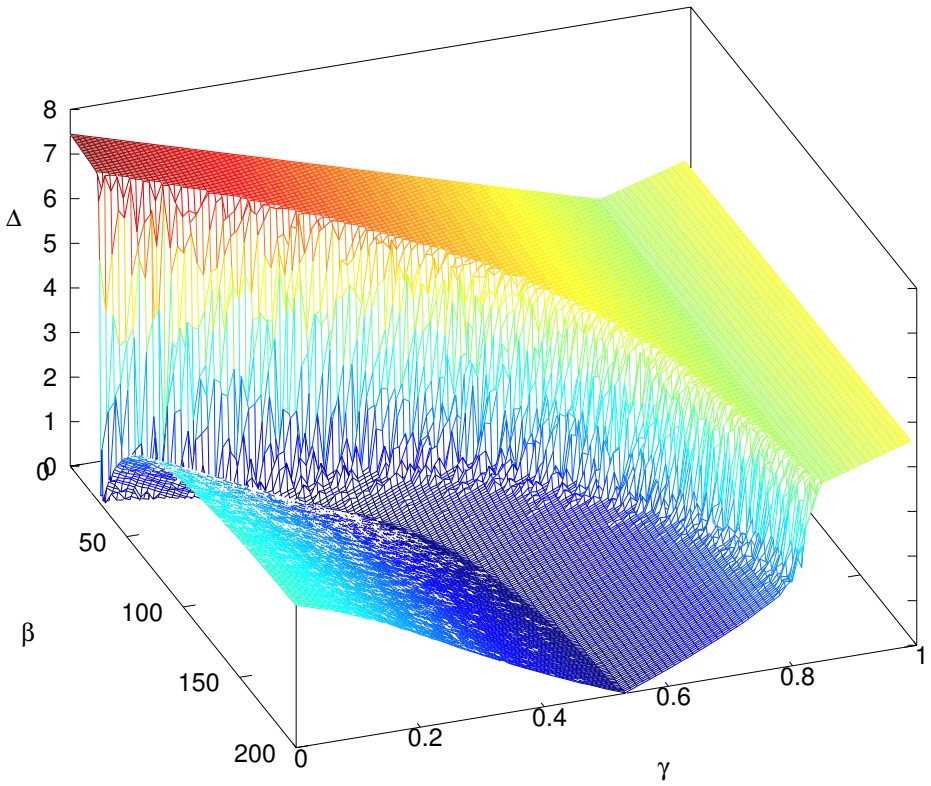


Figure 11: Error Δ for Example 11. The point we try to find is at $\beta = 150$ and $\gamma = 0.5$.

5 Conclusion & Outlook

Above, I presented a model for an economic and political society. Because a coupling of political power to social status exists, I explained that it should have a tendency to develop social inequality, which is an effect also observed in reality as discussed extensively on page 12. In addition, I claimed that—as is done in my model—one can apply the framework of statistical mechanics to problems in economy and sociology. I believe that the results shown above justify both those claims: For one, it was shown in Section 3.2 that the model really tends to break permutation symmetry corresponding to equality between all participants in favour of an energetically better state which is however totally unequal. Secondly, simulations based on techniques developed for the calculation of physical systems described by statistical mechanics done in Chapter 4 showed very interesting behaviour, confirming the inequality result. In Subsection 4.3.3 they even led to the conclusion that the transition from social equality to unequal states shows the characteristics of a first-order phase transition, so that my model can be—depending on the chosen parameters β and γ—classified to be in either of two phases: Left corresponding to intact permutation symmetry, or right in which one individual becomes privileged above all others. Concluding, I can say that the model gave very interesting insights and that all my expectations with respect to it worked out. I believe that I was able to show everything that is necessary for a comprehensive initial study of the introduced model. But of course, as always in science, there remain a lot of things open for further research. As the final discussion in this document, I want to mention some of those open aspects as I have them in mind:

- While in Chapter 3 a lot of fundamental things have been shown, I was not able to establish some other results which may be true and seem plausible. In particular, I was not able to deduce much about the behaviour of the minimiser of (22) for the limit $N \to \infty$, although I believe that there probably is a limit point. Subsection 3.4.2 only gives very basic first ideas in this direction. If the existence of this limit point can be established, a further question is how it depends itself on γ. Another interesting open point is whether or not one can strengthen Theorem 11 in order to yield relations between the individual coordinates (a_0, l_0) and (a_1, l_1) instead of "only" the strains corresponding to them. Even if some of my conjectures are not true, one could try to find counter-examples to disprove them.

- I think that the introduction of γ to control fairness in a society, as done above in (8), is not very nice—mathematically, it makes sense to introduce it via convex combinations, and it allows nice conclusions to be drawn. However, it also leads to the fact that power is *always* distributed totally unevenly as $m_1 = 1$ and $m_n = 0$ for $n = 2, \dots, N$ (which was shown in Theorem 9), rather than "softening" this distinction for increasing γ. Instead, γ merely weakens *the practical effect* m has. Also, for $\gamma = 1$, m is rendered completely irrelevant. In short, while γ as introduced works to control fairness, it also has some odd implications for the coordinate m in social space. Thus, one could think of modifying the definition of $\mathcal{H}$ so that γ does not play directly into the weights of individual strains, but instead an additional penalty term is added which, depending on γ, increases the energy for unequal power distributions. It would surely be interesting to set up and analyse such a modified model.

- For real societies, of course the thermodynamic limit $N \to \infty$ may be interesting instead of fixing a number of discrete individuals. One possibility is to consider my model for large or increasing N—see the remark above but also what was already done in Subsection 4.3.4. A different approach to reach this goal is to get away from discrete individuals and instead model the state of our society as a probability measure in social space giving a *density* of people. The constraints (5) and (6) can then probably be formulated as conditions on expectation values of random variables corresponding to the coordinates in social space with respect to this new measure (not to be confused with expectation values as used above, which were always with respect to π_T). This then leads to a problem of *infinite-dimensional* optimisation, and possibly some of the techniques used for the results in Section 3.2 are still applicable. Note though that also in ordinary thermodynamics where N is much, much larger than any society can ever be, one usually considers the particles still as discrete and rather takes $N \to \infty$ directly without such a formal transition to a "continuous" problem.

- In my model, I assume that productivity α is constant. In fact, I set $\alpha = 1$ in Section 2.5. This is valid to describe a stationary situation where the means of production (capital, machines, and so on) do not change, but of course in reality, productivity usually grows over time because of accumulation of capital and also technological progress. Thus one could try to analyse how a change in α affects the model behaviour with everything else[79] fixed.

- I have not very thoroughly investigated the phase-transition described in Subsection 4.3.3, so there's still a lot of potential for further research on its detailed properties. For instance, critical exponents and universality [38, p. 188].

- Although I have given indications that the concrete form of the strain function f is not very important (see Subsection 4.3.5), it would still be interesting to try more than the ones I did and possibly even to think about ways to find "real" strain functions other than by guessing what might be plausible. This could be done either by modelling personal preferences in some way, or maybe even by empirical experiments as are common nowadays in experimental economics [26].

- So far I have neglected one facet of mathematical modelling which is in fact very important: Whether or not the model actually describes something useful. In my case, the only indications that my model might have something to do with reality are that it leads to social inequality which is also observed in nature, and that it is partially based on assumptions which are made frequently in economics. However, more extensive tests could be done. It should not be too difficult to compare its qualitative behaviour in certain aspects to corresponding behaviour of real economies—provided that the correct aspects can be found. And if the problem of a realistic strain function can be solved, one can even try quantitative comparisons—for instance by estimating positions in my social space of some sample of a country's population[80] and comparing that to simulation results. It may even be simpler than that—as discussed in Section 4.4, for estimating β and γ of a real society it is enough to know $\mathcal{H}$ and $\mathcal{U}$. Maybe a similar technique can be used on other observables like $\mathcal{Y}$ and $\mathcal{G}$, which are readily available for real economies, although I don't have a formal justification for it.

- In addition, there exists of course a lot of literature on related topics by more classical authors of economy and sociology. I have only sometimes referred to corresponding sources, and it would be interesting to compare my approach in more depth to existing models and theories about the topics I touched above (mostly social inequality).

- I have not given any interpretation of what the temperature T (or correspondingly β) means for a society, except for a short remark on page 16. It would be interesting to consider this, and also what thermal equilibrium between two societies implies.

- My model only describes a single, closed economy. But of course, in reality one has to consider economics *globally* [23] where lots of single economies with very different local conditions interact with each other. If it becomes clear how temperature maps to reality, one could try to model this global interaction by putting together multiple societies described each by my model for different N, β, γ and possibly even f and L to account for local differences, and put those in thermodynamic contact to each other. If the grand canonical ensemble is used as mentioned on page 16, even migration between those societies can be described.

So far, so good—as you can see, there are indeed still *a lot* of open issues. But if you've really made it so far as here, I ought to thank you for your interest in my work rather than telling you what I want you to do for my model in the future! And if you've been only fascinated marginally as much as I was by the results in particular in Section 4.3, I can be glad because you must have enjoyed it very much.

[79] In particular also f, recall that f was modified in Theorem 3 where $\alpha = 1$ was chosen.

[80] Although of course actually figuring out more or less exact coordinates, in particular a and m, by a questionnaire seems difficult to me.

References

[1] GNU Compiler Collection. `http://gcc.gnu.org/`.

[2] GNU Octave. `http://www.octave.org/`.

[3] OpenMP.org. `http://openmp.org/`.

[4] Edward Bellamy. *Looking Backward, 2000 to 1887*, volume 624 of *Project Gutenberg*. Project Gutenberg, P.O. Box 2782, Champaign, IL 61825-2782, USA, 1996.

[5] Olivier Blanchard. *Macroeconomics*. Pearson Prentice Hall, fourth edition, 2006.

[6] Intel Corporation. Intel Hyper-Threading Technology. `http://www.intel.com/content/www/us/en/architecture-and-technology/hyper-threading/hyper-threading-technology.html`. Online; accessed 2012-05-29.

[7] Harald Cramér. *Mathematical Methods of Statistics*. Princeton University Press, Princeton, eleventh edition, 1966.

[8] Jean Dieudonné. *Foundations of Modern Analysis*, volume 10-I of *Pure and Applied Mathematics*. Academic Press, New York / London, second edition, 1969.

[9] Jürgen Elstrodt. *Maß- und Integrationstheorie*. Springer, Berlin / Heidelberg, sixth edition, 2009.

[10] Wolfgang Finkelnburg. *Einführung in die Atomphysik*. Springer, Berlin / Heidelberg / New York, twelfth edition, 1976.

[11] Wolfram Fischer, editor. *Europäische Wirtschafts- und Sozialgeschichte von der Mitte des 19. Jahrhunderts bis zum ersten Weltkrieg*, volume 5 of *Handbuch der europäischen Wirtschafts- und Sozialgeschichte*. Klett-Cotta, Stuttgart, 1985.

[12] Wolfram Fischer, editor. *Europäische Wirtschafts- und Sozialgeschichte vom ersten Weltkrieg bis zur Gegenwart*, volume 6 of *Handbuch der europäischen Wirtschafts- und Sozialgeschichte*. Klett-Cotta, Stuttgart, 1987.

[13] Torsten Fließbach. *Statistische Physik*, volume IV of *Lehrbuch zur Theoretischen Physik*. Spektrum Akademischer Verlag, Heidelberg / Berlin, third edition, 1999.

[14] Peter Gantner and Hannes Hattenberger. Rechnerorganisation I: Dual Server'10 AMD Hexalon. `http://www.uni-graz.at/~haasegu/Lectures/RO-I/WS09/Hexalon/Bericht_RO1.pdf`. Online; accessed 2012-05-29.

[15] Christof Gattringer and Christian B. Lang. *Quantum Chromodynamics on the Lattice: An Introductory Presentation*, volume 788 of *Lecture Notes in Physics*. Springer, Berlin / Heidelberg, 2010.

[16] Anthony Giddens, Christian Fleck, and Marianne Egger de Campo. *Soziologie*. Nausner & Nausner, Graz / Wien, third edition, 2009.

[17] Corrado Gini. On the Measure of Concentration with Special Reference to Income and Wealth. *Colorado College Publication, General Series*, 208:73–80, September 1936.

[18] Harvey Gould, Jan Tobochnik, and Wolfgang Christian. *An Introduction to Computer Simulation Methods: Applications to Physical Systems*. Pearson Addison Wesley, San Francisco, third edition, 2007.

[19] Magnus R. Hestenes. *Optimization Theory: The Finite Dimensional Case*. John Wiley & Sons, New York / London / Sydney / Toronto, 1975.

[20] Saif Khalid. Occupy Wall Street: 'We are the 99 %'. *Al Jazeera English*, December 2011. `http://www.aljazeera.com/indepth/spotlight/aljazeeratop102011/2011/12/20111226125438417351.html`. Online; accessed 2012-03-15.

[21] John F. C. Kingman and S. J. Taylor. *Introduction to Measure and Probability*. Cambridge University Press, London, second edition, 1973.

[22] Reinhard Kreckel. *Politische Soziologie der sozialen Ungleichheit*. Campus Verlag, Frankfurt / New York, third edition, 2004.

[23] Reinhard Kreckel. Soziologie der sozialen Ungleichheit im globalen Kontext. *Der Hallesche Graureiher*, 4, 2006.

[24] Heinz D. Kurz, editor. *Klassiker des ökonomischen Denkens*, volume 1. C. H. Beck, München, 2008.

[25] George Leitmann, editor. *Optimization Techniques With Applications to Aerospace Systems*. Academic Press, New York / London, third edition, 1965.

[26] Ulrike Leopold-Wildburger. Some Selected Topics in Experimental Economics. *International Transactions of Operational Research*, 4(3):165–174, 1997.

[27] M. Mackintosh, V. Brown, N. Costello, G. Dawson, G. Thompson, and A. Trigg. *Economics and changing economies*. International Thomson Business Press, London, 1996.

[28] Dieter Meschede. *Gerthsen Physik*. Springer, Berlin / Heidelberg, 23rd edition, 2006.

[29] Jürgen Mimkes. A Thermodynamic Formulation of Economics. *Econophysics & Sociophysics: Trends & Perspectives*, pages 1–34, 2006. Viley-VCH, Weinheim.

[30] Jürgen Mimkes. A Thermodynamic Formulation of Social Science. *Econophysics & Sociophysics: Trends & Perspectives*, pages 277–308, 2006. Viley-VCH, Weinheim.

[31] Jürgen Mimkes and Patrick Illinger. Physik der menschlichen Gesellschaft: Formeln der Freiheit. *Süddeutsche Zeitung*, February 2011. `http://www.sueddeutsche.de/wissen/physik-der-menschlichen-gesellschaft-formeln-der-freiheit-1.1061313`. Online; accessed 2012-05-31.

[32] Thomas More. *Utopia*, volume 2130 of *Project Gutenberg*. Project Gutenberg, P.O. Box 2782, Champaign, IL 61825-2782, USA, 2000.

[33] National Institute of Standards and Technology. Digital Library of Mathematical Functions. `http://dlmf.nist.gov/`. Online; accessed 2012-06-02.

[34] George Orwell. *1984*. Signet Classic, Penguin Books, New York, 1977.

[35] George Orwell. *Animal Farm*. Penguin Books, London, 2000.

[36] Emmanuel Saez. Striking it Richer: The Evolution of Top Incomes in the United States (Updated with 2009 and 2010 estimates). `http://elsa.berkeley.edu/~saez/saez-UStopincomes-2010.pdf`, March 2012. Online; accessed 2012-03-15.

[37] Harold S. Sloan and Arnold J. Zurcher. *Dictionary of Economics*. Barnes & Noble Books, New York / Hagerstown / San Francisco / London, fifth edition, 1970.

[38] Jos Thijssen. *Computational Physics*. Cambridge University Press, second edition, 2007.

[39] Jan A. van Houtte, editor. *Europäische Wirtschafts- und Sozialgeschichte im Mittelalter*, volume 2 of *Handbuch der europäischen Wirtschafts- und Sozialgeschichte*. Klett-Cotta, Stuttgart, 1980.

[40] Hal R. Varian. *Intermediate Microeconomics: A Modern Approach*. W. W. Norton & Company, New York / London, third edition, 1993.

[41] Friedrich Vittinghoff, editor. *Europäische Wirtschafts- und Sozialgeschichte in der römischen Kaiserzeit*, volume 1 of *Handbuch der europäischen Wirtschafts- und Sozialgeschichte*. Klett-Cotta, Stuttgart, 1990.

[42] Dalip Singh Wasan. Money Rules the World. *Articlesbase*, January 2008. http://www.articlesbase.com/law-articles/money-rules-the-world-296824.html. Online; accessed 2012-03-15.

[43] Wikipedia. C++11 — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=C%2B%2B11&oldid=494830497. Online; accessed 2012-05-29.

[44] Wikipedia. Gini coefficient — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Gini_coefficient&oldid=495106104. Online; accessed 2012-05-30.

[45] Julia M. Yeomans. *Statistical Mechanics of Phase Transitions*. Clarendon Press, Oxford, third edition, 1994.