# Pareto laws for agent-based wealth distribution models

by

Milan van Roessel

to obtain the degree of Bachelor of Science
at the Delft University of Technology,
to be defended publicly on Tuesday November 21, 2023 at 09:00 AM.

| | | |
|---|---|---|
| Student number: | 5125170 | |
| Project duration: | September, 2023 – November, 2023 | |
| Thesis committee: | Prof. dr. F. H. J. Redig, | TU Delft, supervisor |
| | Dr. H. Yoldas, | TU Delft |

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

# Abstract

In this thesis, wealth distribution in a closed economic system is examined by studying the simple inclusion process (SIP). The simple inclusion process is a model coming from statistical physics that models the jumping of particles in a graph. In the model particles have attraction among each other and each site also has a characteristic attraction parameter, denoted by the variable $\alpha$. We regard the simple inclusion process as an agent based model for the economy for which sites represent agents and the particles represent wealth transferring from agent to agent. For the model, invariant measures are found that represent possible long term distributions of wealth in the system. We extend the model by looking at $\alpha$ as a parameter drawn from its own underlying probability distribution, $\psi(\alpha)$. In the view of wealth distribution, $\alpha$ can be seen as the wealth attraction an agent has. We set out to look for conditions on the distribution of $\alpha$ for which we obtain asymptotic power law behaviour of the resulting wealth distribution. It is shown that if a higher-order moment of $\psi(\alpha)$ diverges, the resulting wealth distribution will have a weak asymptotic power law lower bound. By setting stricter conditions on the distribution of $\alpha$, we obtain that the wealth distribution will be asymptotically equal to a power law.

# Contents

# 1
# Introduction

Trying to explain the distribution of wealth from simple agent based models is a challenge. By finding universalities in such a model, a possible explanation for phenomena like wealth inequality can be found. Creating effective, simple models continues to be a difficult task. Since the late 20th century, ideas of statistical physics have started to enter this area of research to better explain wealth distribution in an economy involving interacting agents. The area combining agent-based models of the economy with ideas from statistical physics is called 'econophysics', see [11] for an overview. For example, the Maxwell-Boltzmann distribution, which describes the statistical distribution of particle speeds in a gas, can be analogously applied to model the distribution of wealth among economic agents.

In this thesis, the simplified, physics-inspired model used to look at wealth distribution in an economy, is the *simple inclusion process* (SIP for short), as described in [6]. It models the jumping of particles along edges of a graph, with particles having attraction among themselves. Considering the wealth distribution perspective, particles are seen as units of wealth and the jumping of particles from site to site as the exchange of wealth between agents. The long-term distribution of the particles for the simple inclusion process, models the long-term distribution of wealth in the economic agent-based system. In this thesis, different variants of the SIP model are studied and invariant measures that give insight into the possible long term behaviour of the variants are presented. In the SIP model, the parameter $\alpha$ represents the attraction of particles for a specific site. In view of our discussion, it can be seen as the extent to which an agent attracts wealth. We start off by looking at SIP models for $\alpha$ a fixed value. Later in the discussion, the SIP model is extended with $\alpha$ seen as an independent, identically distributed random variables drawn from a distribution, which we will denote by $\psi(\alpha)$.

Different conditions on $\psi(\alpha)$ are given for which the tail of the wealth distribution will be of Pareto type. First we look at conditions for which the found wealth distribution will have an weak asymptotic Pareto (or power law) lower bound. After that we give conditions that allow us to draw stronger conclusions on the tail of the wealth distribution. These conditions on $\psi(\alpha)$ are such that the wealth distribution will be asymptotically equal to a Pareto (or power law) distribution. Since the Pareto distribution can be closely fitted to numerous examples of empirical data of the upper-part of the economy [4], it is the desired behaviour in the tail of the long term wealth distribution of our model.

## 1.1. Overview

Chapter 2 delves into the foundational principles of the probabilistic building blocks for our process, continuous-time Markov chains. Additionally some theory on invariant measures and detailed balance is given. In Chapter 3, the focus is shifted to the main model of this discussion, the SIP model. The model is explained and different variants are introduced. With the help of the theory from the previous chapter, invariant measures that model possible long term distributions of wealth in the system are found. In chapter 4, the Pareto distribution is introduced and several features of the distribution are discussed. Moreover in this chapter, the conditions on the distribution on $\alpha$ to make way for Pareto behaviour in the tail of the long term wealth distribution are presented.

# 2
# Markov chains and detailed balance

In this chapter a theoretical overview for continuous time Markov chains is presented. Some theory on invariant and reversible measures is given and the detailed balance condition is introduced. It's worth mentioning that this section draws considerably from the theory presented in [8].

## 2.1. Markov chains

Markov chains represent the probabilistic evolution of a random variable $X$ in the course of time. This random variable can take values in a set $\Omega$, which we will call the *state space*. For our discussion, we will only focus on Markov chains on a state space that is finite or at most countably infinite. A Markov chain can be constructed for both discrete and continuous time. To help understand the continuous time Markov process, let us start with the discrete time Markov process.

### 2.1.1. Discrete time

We start with the formal definition.

**Definition 2.1.1** (Discrete Markov chain)**.** Let $\Omega$ be a finite state space. A **discrete Markov chain** is a family of random variables $\{X_t : t \in \mathbb{N}\}$, taking values in $\Omega$ for which the *discrete Markov property* holds:

$$\mathbb{P}(X_n = x_n \mid X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = \mathbb{P}(X_n = x_n \mid X_{n-1} = x_{n-1}) \tag{2.1}$$

The discrete Markov property is equivalent with:

$$\mathbb{E}(f(X_n)|X_1, X_2, \dots, X_{n-1}) = \mathbb{E}(f(X_n)|X_{n-1}) \tag{2.2}$$

The term memorylessness can also be used to express the Markov property; the distribution of future states given the present and the past states only depends on the present state. Because of this property, Markov chains make for a very interesting branch of stochastics; the Markov property simplifies the process of saying something meaningful about the evolution of a stochastic variable throughout time a whole lot.

At each time step, the random variable 'jumps' from one state in the state space to the other. For a Markov process, we denote the probability to jump from $i \to j$, by $p_{ij}$. The matrix with entries $(p_{ij})$ in row $i$ and column $j$ is called the *transition matrix*, $P$. $P$ is also sometimes called the *one-step probability matrix*. For a Markov process with one-step probability matrix $P$, $P^2$ denotes the two-step probability matrix, $P^3$ the three-step probability matrix, etc.; By looking at the entry $(P^n)_{ij}$, we find the probability to go from state $i$ to state $j$ in $n$ steps. We will not present a formal proof for this property of the process. However, it can be observed in the following example: given a Markov chain $\{X_t, t \in \mathbb{N}\}$ on a state space $\Omega$ starting from $i_0$. Say we want to know the probability $\mathbb{P}(X_2 = i_2|X_0 = i_0)$, the probability to go from $i_0$ to $i_2$ in two steps:

$$\mathbb{P}(X_2 = i_2|X_0 = i_0) = \sum_{j \in \Omega} \mathbb{P}(X_2 = i_2|X_1 = j \cap X_0 = i_0)\mathbb{P}(X_1 = j|X_0 = i_0), \tag{2.3}$$

because of the Markov property, we have that

$$\mathbb{P}(X_2 = i_2|X_1 = j \cap X_0 = i_0) = \mathbb{P}(X_2 = i_2|X_1 = j).$$

With this in mind, the RHS of 2.3 evaluates to

$$\sum_{j \in \Omega} \mathbb{P}(X_2 = i_2|X_1 = j)\mathbb{P}(X_1 = j|X_0 = i_0) = (P^2)_{i_0 i_2}.$$

Considering our focus on the long term wealth distribution in a economic system, we are interested in the Markov chain's long-term behaviour. To that end, we introduce the notion of an invariant distribution.

**Definition 2.1.2** (Invariant distribution)**.** Consider a Markov process $\{X_t, t \in \mathbb{N}\}$ on a state space $\Omega$ with one-step transition matrix $P$. The vector $\pi = (\pi_i, i \in \Omega)$ is called the **invariant distribution** of the Markov chain if the following two properties hold:

a) $\pi$ is a probability distribution, i.e. $\pi_i \geq 0$ for all $i \in \Omega$ and $\sum_{i\in\Omega} \pi_i = 1$

b) $\pi = \pi P$

Because we have that for a discrete Markov chain we have that $P^n$ is the n-step probability matrix, we have the following interesting result. Consider a Markov chain that starts from $X_0$ having initial distribution $\pi$. If we are interested in the distribution of $X_n$, we can examine $\pi P^n$. If $\pi$ is an invariant distribution, $\pi P^n = \pi P P^{n-1} = P^{n-1} = \cdots = \pi$; the initial distribution of $X_0$ is not changing under the passage of time; this invariant distribution dictates a possible long-term behaviour of the Markov process.

In our discussion, we only consider *homogeneous* Markov chains. A Markov chain is called homogeneous if for all $i, j \in \Omega$, the probability $\mathbb{P}(X_{n+1} = i | X_n = j)$ does not depend on $n$, i.e.: the transition probabilities of the Markov chain do not change overtime.

### 2.1.2. Continuous

In order to make the idea of a Markov chain more relevant to real-world scenarios, it is natural to consider continuous time Markov chains. By doing so, we can model and analyse processes that unfold more naturally and continuously in various scenarios. This approach broadens the scope of Markov chains, allowing them to better capture the dynamics of real-world systems.

**Definition 2.1.3** (Continuous time Markov chain)**.** A **continuous time Markov chain** $\{X_t : t > 0\}$ on a finite state space $\Omega$, is a family of random variables for which the *continuous Markov property* holds: For all $t > 0$, $n \in \mathbb{N}$, $0 < t_1 < t_2 < \cdots < t_n < t$ and for all $f : \Omega \to \mathbb{R}$:

$$\mathbb{E}(f(X_t)|X_{t_1}, X_{t_2}, \ldots, X_{t_n}) = \mathbb{E}(f(X_t)|X_{t_n}) \tag{2.4}$$

Unfortunately, the concept of a one-step transition probability (and the probability matrix) does not carry over to the world of continuous Markov chains. Transition probabilities, which are commonly used in discrete-time Markov chains, represent the likelihood of moving from one state to another within a fixed time step. In the continuous-time context, the notion of 'time steps' becomes less meaningful because time is considered as a continuous, unbroken flow. Consequently, for continuous Markov chains, transition probabilities are replaced by transition rates, which express how quickly or frequently transitions occur between states in a continuous and uninterrupted time scale. For a continuous time Markov process $\{X_t, t \geq 0\}$ with $i, j \in \Omega$, $c(i, j) \geq 0$ denotes the transition rate to go from state $i$ to $j$. Intuitively, this transition rate resembles the probability per unit of time for the Markov process to jump from $i \to j$, with $i \neq j$. This Markov process starting from $i$ will then evolve as follows: we wait an exponential time with parameter $\lambda_i = \sum\limits_j c(i, j)$. We then jump to $j$, with probability $p(i, j) = \frac{c(i,j)}{\lambda_i}$. Note that $p(i, j)$ is indeed a probability since:

$$\sum_j p(i, j) = \sum_j \frac{c(i, j)}{\lambda_i} = \frac{1}{\lambda_i} \sum_j c(i, j) = 1$$

The fact that the exponential distribution is memoryless, i.e. for $X \sim \exp(\lambda)$, $\mathbb{P}(X > x+a|X > a) = \mathbb{P}(X > x)$ (for $x, a \geq 0$), guarantees the Markov property of the process. For a Markov process, the waiting times between jumps must be exponentially distributed. To see this, let $T_x$ denote the time at which the first jump takes place for the process starting from $x$.

$$\mathbb{P}(T_x > t + s|T_x > s) = \mathbb{P}(X_r = x, \forall r, s \leq r \leq t + s|X_u = x, \forall u, 0 \leq u \leq s),$$

by the memorylessness property,

$$\mathbb{P}(T_x > t + s|T_x > s) = \mathbb{P}(X_r = x, \forall r, s \leq r \leq t + s|X_s = x).$$

The RHS now reads the probability of $X$ to be at $x$ in the time-interval $[s, t + s]$ given the fact that $X$ is at $x$ at time $s$. This can be seen as the probability that no jump occurs in the time interval $[s, t + s]$. Since the Markov chains that we consider are homogeneous, we see that

$$\mathbb{P}(T_x > t + s|T_x > s) = \mathbb{P}(T_x > t),$$

or equivalently:

$$\frac{\mathbb{P}(T_x > t + s \cap T_x > s)}{\mathbb{P}(T_x > s)} = \mathbb{P}(T_x > t),$$

by noting that $\mathbb{P}(T_x > t + s \cap T_x > s) = \mathbb{P}(T_x > t + s)$ we get that

$$\mathbb{P}(T_x > t + s) = \mathbb{P}(T_x > s)\mathbb{P}(T_x > t) \tag{2.5}$$

If we then let $\phi(t)$ denote $\mathbb{P}(T_x > t)$, 2.5 can be rewritten as $\phi(t+s) = \phi(s)\phi(t)$. This observation, together with the fact that $\phi(t)$ is decreasing in $t$, we can conclude that $\phi(t) = e^{-\lambda t}$ for some $\lambda > 0$.

## 2.2. The Markov semigroup and generator

Consider a Markov process $\{X_t, t \geq 0\}$ and a function $f : \Omega \to \mathbb{R}$. We define

$$S_t f(x) = \mathbb{E}(f(X_t)|X_0 = x) = \mathbb{E}_x(f(X_t)). \tag{2.6}$$

Here $\mathbb{E}_x$ denotes the expectation of the process starting at $x \in \Omega$. The operation $S_t$ defined in 2.6 can be seen as the 'push-forward' of a function in the course of time. $S_t$ defines a family $\{S_t, t \geq 0\}$ of operators called the Markov semigroup. We define the Markov semigroup for $\Omega$ both infinite and finite.

**Definition 2.2.1** (Markov Semigroup)**.** Let $\Omega$ be a state space and let $\mathscr{C}(\Omega)$ denote the space of all continuous functions[1] defined on $\Omega$. If we have a Markov process $\{X_t, t \geq 0\}$ on $\Omega$ and a family of operators $\{S_t, t \geq 0\}$, with $S_t : \mathscr{C}(\Omega) \to \mathscr{C}(\Omega)$, then $\{S_t, t \geq 0\}$ is called a **Markov semigroup**, if it satisfies the properties stated in theorem 2.2.1. Moreover, the Markov semigroup connects to the Markov process on $\Omega$ via 2.6.

Let us now present the properties of the Markov semigroup.

**Theorem 2.2.1** (Semigroup properties)**.** Given a Markov process on a state space $\Omega$, $f : \Omega \to \mathbb{R}$ and $S_t$ as in 2.6, the semigroup $\{S_t, t \geq 0\}$ satisfies the following properties:

a) Identity at time zero: $S_0 = I$, i.e. for all $f, S_0 f = f$.

b) Right continuity: the map $t \to S_t f$ is right continuous.

c) Semigroup property: for all $t, s > 0$, $S_{t+s} f = S_t(S_s f) = S_s(S_t f)$.

d) Positivity: $f \geq 0$ implies $S_t f \geq 0$.

e) Normalization: $S_t 1 = 1$.

f) Contraction: $\max_x |(S_t f)(x)| \leq \max_x |f(x)|$.

Since only the proofs for property *c* and *f* are non-trivial, they are given.

*Proof.* Semigroup property:

$$S_{t+s} f(x) = \mathbb{E}_x\left(f(X_{t+s})\right)$$

We let $\mathscr{F}_s$ denote the $\sigma$-algebra generated by the random variables $\{X_r, r \leq s\}$. This is nothing other than the entire past until $s$. By using the fact that the expectation is the expectation of the conditional expectation:

$$\mathbb{E}_x\left(f(X_{t+s})\right) = \mathbb{E}_x\left(\mathbb{E}_x\left(f(X_{t+s})|\mathscr{F}_s\right)\right).$$

Because of the Markov property, conditioning on the past until time $s$ is the same as starting a new Markov chain from $X_s$:

$$\begin{aligned}\mathbb{E}_x\left(\mathbb{E}_x\left(f(X_{t+s})|\mathscr{F}_s\right)\right) &= \mathbb{E}_x\left(\mathbb{E}_{X_s}\left(f(X_t)\right)\right)\\ &= \mathbb{E}_x\left((S_t f)(X_s)\right)\\ &= S_s(S_t f)(x)\end{aligned}$$

[1] When $\Omega$ is finite, all functions defined on $\Omega$ are continuous.

By a similar calculation, we can derive $S_{t+s}f(x) = S_t(S_s f)(x)$.
Contraction property:

$$\begin{aligned}|S_t f|(x) &= \left|\mathbb{E}_x f(X_t)\right| \\ &\leq \mathbb{E}_x\left(\left|f\right|\,|(X_t)\right) \\ &\leq \mathbb{E}_x\left(\sup_y f(y)\right) \\ &= \sup_y f(y).\end{aligned}$$

Here we have used the fact that we can exchange the expectation and the absolute value at the cost of an inequality. This completes the proof. □

If $\Omega$ is finite, a function $f : \Omega \to \mathbb{R}$ can be seen as a vector mapping each value in $\Omega$ to a value in $\mathbb{R}$. $S_t$ can be seen as a matrix with elements $(S_t)_{x,y} = \mathbb{P}(X_t = y|X_0 = x)$. In this case, because of the semigroup property, $S_t = e^{tL}$ for some matrix $L$. The exponential of a matrix is defined by its Taylor series:

$$e^{tL} = \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n$$

Because of this, $e^{tL} = I + tL + \mathcal{O}(t^2)$. $L$ can be computed via

$$Lf = \lim_{t\to 0} \frac{S_t f - f}{t} \tag{2.7}$$

and is called the Markov generator. The Markov generator is an operator. An operator is a mathematical operation that maps a function in its domain to another function. Applying the Markov generator to a function can be seen as what happens in a Markov process to that function for $t$ infinitesimal. Intuitively, if we know what happens for an infinitely small time step in a process, we can chain these steps together to know what happens in the process in a longer time interval. This brings us to the following important note: given a Markov process $\{X_t, t \geq 0\}$, if we a have Markov semigroup $\{S_t, t \geq 0\}$ connecting to this Markov process via 2.6, then this Markov semigroup also uniquely determines this process. Likewise, the Markov generator connected to this Markov semigroup (via 2.7, which in turn is connected to the Markov process via 2.6) also uniquely determines this Markov process.

To conclude this section, we present the generator for the Markov process central in our discussion. The proof is excluded from this discussion.

**Theorem 2.2.2.** For a continuous time Markov chain on a finite state space $\Omega$, with rates $c(\eta,\xi)$, $\eta,\xi \in \Omega$, the Markov generator is given by:

$$Lf(\eta) = \sum_{\xi\in\Omega} c(\eta,\xi)(f(\xi) - f(\eta)) \tag{2.8}$$

## 2.3. Invariant measures

We introduce the notion of a probability measure $\mu$ on a finite state space.

**Definition 2.3.1** (Probability measure)**.** Given a finite state space $\Omega$, the mapping $\mu : \Omega \to \mathbb{R}$ is called a **probability measure**, if the following hold:

(a) $\mu(\omega) \geq 0$ for $\omega \in \Omega$

(b) $\sum_{\omega\in\Omega} \mu(\omega) = 1$

For a probability measure on $\Omega$ the integral becomes a finite sum.

$$\int_\Omega f \mathrm{d}\mu = \sum_{\omega\in\Omega} f(\omega)\mu(\omega).$$

In the rest of our discussion, if we talk about a probability measure, we are always talking about a probability measure on a finite state space. The term probability measure, probability distribution, measure and distribution are used interchangeably throughout this discussion.

### 2.3.1. Invariance

As stated before, in this discussion we work with finite state spaces. However, the notation for the results we show in this section, also translate over to Markov chains on infinite state spaces. Given a measure $\mu$, we can look at how it evolves in time. Given a Markov process on a state space and a Markov Semigroup $\{S_t, t \geq 0\}$, we denote by $\mu S_t$ the evolution of $\mu$ after time $t$. This is the distribution of the process at time $t$. We can see $\mu S_t$ as $\pi P^n$ from the discussion on discrete Markov chains. If $\{X_t, t \geq 0\}$ is a Markov chain on a state space $\Omega$, with initial distribution $\mu$, $\mu S_t$ is the unique measure such that $\int f \mathrm{d}\mu S_t = \int S_t f \mathrm{d}\mu$.

For discrete time Markov chains, we already saw the invariant distribution, $\pi$. For a continuous time Markov chain, we introduce the notion of an *invariant measure*:

**Definition 2.3.2** (Invariant measure)**.** A probability measure $\mu$ is called **invariant** if for all $t \geq 0$ and $f \in \mathscr{C}(\Omega)$:

$$\int S_t f \mathrm{d}\mu = \int f \mathrm{d}\mu \tag{2.9}$$

A process can converge to a steady state, in which the probabilities of being in a certain state does not change. A probability measure that does not change as $t$ runs, is called invariant. Important to note is that we are talking about there not being a change in distribution. If we were to run a real simulation, the Markov chain would still jump from state to state.

The following theorem connects the Markov generator and invariant measures:

**Theorem 2.3.1.** Let $L$ be the Markov generator for a continuous Markov process on finite state space $\Omega$. For all $f : \Omega \to \mathbb{R}$ in the domain of $L$, and $\mu$ a probability measure on $\Omega$, we have that $\mu$ is an invariant measure if and only if

$$\int L f \mathrm{d}\mu = 0 \tag{2.10}$$

*Proof.* Suppose $\mu$ is invariant, that is $\int S_t f \mathrm{d}\mu = \int f \mathrm{d}\mu$, then

$$\int L f \mathrm{d}\mu = \int \lim_{t \to 0} \frac{S_t f - f}{t} d\mu = \lim_{t \to 0} \frac{1}{t} \int S_t f - f \mathrm{d}\mu = 0.$$

Conversely, suppose that $\int S_t f \mathrm{d}\mu = \int f \mathrm{d}\mu$.

$$\frac{\mathrm{d}}{\mathrm{d}t} \int S_t f \mathrm{d}\mu = \int \frac{\mathrm{d}}{\mathrm{d}t} S_t f \mathrm{d}\mu = \int L S_t f \mathrm{d}\mu = 0,$$

Remember that $\Omega$ is finite: $S_t = e^{tL}$. Because of 2.10 and $S_t f$ also being in the domain of $L$, the derivative with respect to $t$ of the integral becomes zero. The integral is then constant, i.e. does not depend on $t$, so we get

$$\int S_t f \mathrm{d}\mu = \int S_0 f \mathrm{d}\mu = \int f \mathrm{d}\mu$$

by the properties of the semigroup. This completes the proof. □

### 2.3.2. Reversibility

For our further discussion, we introduce the notion of reversibility.

**Definition 2.3.3** (Reversible measure)**.** Consider a Markov process on $\Omega$ with Markov semigroup $\{S_t, t \geq 0\}$. A measure $\mu$ is called **reversible** if for all $f, g \in \mathscr{C}(\Omega)$

$$\int (S_t f) g \mathrm{d}\mu = \int f (S_t g) \mathrm{d}\mu \tag{2.11}$$

When a measure is reversible, the Markov process $\{\eta_t, 0 \leq t \leq T\}$ starting from $\eta_0 = \mu$ is exactly distributed as its reversal, $\{\eta_{T-t}, 0 \leq t \leq T\}$, i.e. the forward Markov process is exactly the same as the backward process. In the field of thermodynamics, dynamical reversibility is the concept that for any forward motion in the dynamical system, there would be a possible reverse motion possible [9]. This is something that can be seen when the behaviour of gases in a closed system is examined. Moreover, when a dynamical system is in equilibrium, the rate of every forward transition is equal to the rate of the backward transition. In view of our discussion of wealth distribution, the notion of reversibility can be translated to the fact that for any transfer of wealth between two agents, the reverse transfer is possible and that when the system is in equilibrium, the average forward rates of wealth transfer are equal to the average backward rates.

**Lemma 2.3.1.** A measure that is reversible is invariant.

*Proof.* Take $g = 1$ in 2.11, and use $S_t 1 = 1$. □

A concept that is synonym to reversibility is *detailed balance*. It is of such importance to our discussion that it deserves its own section.

### 2.3.3. Detailed balance

**Theorem 2.3.2.** For a Markov jump process $\{\eta_t : t \geq 0\}$ on a finite state space with rates $c(\eta, \xi)$ and Markov generator

$$Lf(\eta) = \sum_{\xi} c(\eta, \xi)(f(\xi) - f(\eta)) \tag{2.12}$$

reversibility holds if and only if the detailed balance condition

$$\mu(\eta)c(\eta, \xi) = c(\xi, \eta)\mu(\xi)$$

holds, for all $\eta, \xi \in \Omega$.

*Proof.* Because we are in the finite state space, the integrals in equation 2.11 become finite sums:

$$\sum_{\eta}(S_t f(\eta) g(\eta) \mu(\eta)) = \sum_{\eta} f(\eta) S_t g(\eta) \mu(\eta).$$

This is equivalent to saying

$$\sum_{\eta}(Lf(\eta) g(\eta) \mu(\eta)) = \sum_{\eta} f(\eta) Lg(\eta) \mu(\eta),$$

because in the finite case: $S_t = e^{tL}$. If we fill in 2.12, we get:

$$\sum_{\eta}\sum_{\xi} c(\eta, \xi)\left(f(\xi) - f(\eta)\right) g(\eta)\mu(\eta) = \sum_{\eta}\sum_{\xi} c(\eta, \xi) f(\eta)\left(g(\xi) - g(\eta)\right)\mu(\eta).$$

$-c(\eta, \xi) f(\eta) g(\eta) \mu(\eta)$ appears in both sums, so this is equivalent to:

$$\sum_{\eta}\sum_{\xi} c(\eta, \xi) f(\xi) g(\eta) \mu(\eta) = \sum_{\eta}\sum_{\xi} c(\eta, \xi) f(\eta) g(\xi) \mu(\eta)$$

By taking $f(\eta) = \mathbb{1}_{\{\eta = \eta'\}}$ and $g(\xi) = \mathbb{1}_{\{\xi = \xi'\}}$ we obtain:

$$c(\xi', \eta')\mu(\xi') = c(\eta', \xi')\mu(\eta'),$$

which is precisely the detailed balance equation. □

In our discussion, we assume that detailed balance holds in the system. This is a big assumption for the model, but it makes finding invariant measures significantly easier. If we assume the detailed balance equation holds for a given measure, we also have that the measure is reversible, by theorem 2.3.2. Lemma 2.3.1 then tells us that this measure is invariant. This process of finding invariant measures for different jump processes is the main focus of the next chapter.

# 3
# The SIP-model

In this chapter, we introduce and further explore the model through which the distribution of wealth in an economic system will be examined. The simple inclusion process, (SIP for short) which was introduced in [6]. The SIP model uses a parameter, $\alpha$ which can be seen as the extent to which different agents attract wealth. Different variants of the SIP model are presented and invariant measures for the variants are shown. At the end of the chapter, the connection is made between the SIP model and a continuous redistribution model for wealth.

## 3.1. Description of the model

The SIP model is a model to model the jumping from particles from vertex to vertex in a graph, $G = (V, E)$. In the SIP model, vertices are called sites. When two sites share an edge, it is possible for a particle to jump from site to site. Particles can be at the same site and have attraction among themselves. In addition, every site also has its own characteristic attraction, $\alpha$. From figure 3.1 the significance of this model in the view of the wealth distribution in an economy becomes even more clear. Every site represents an agent and a particle represents a unit of wealth. The jumping of particles from site to site can be a model for our economy in which wealth is transferred between individuals and companies. In the real world, the transferring of wealth happens in exchange for services or goods. This transfer of services or goods is not captured in the model. In the rest of our discussion, we are going to be looking at the simple inclusion process on a graph with two sites. This can be extended to a process on a graph with multiple sites and edges by copying the model we have for two sites, to every edge in our graph. An important fact to note is that the total amount of wealth in the system stays constant. This can be justified in the light of an economy when the time scale is chosen to be large enough. If we let $\eta, \xi \in \Omega$ be two configurations of particles on the two sites, we then have a Markov process on the configuration of particles on the two sites, $\{\eta_t : t \geq 0\}$. The process can then be characterised, by the jump rates. For each pair of $\eta$ and $\xi$, there is a rate $c(\eta, \xi)$ for the jump from configuration $\eta$ to $\xi$.

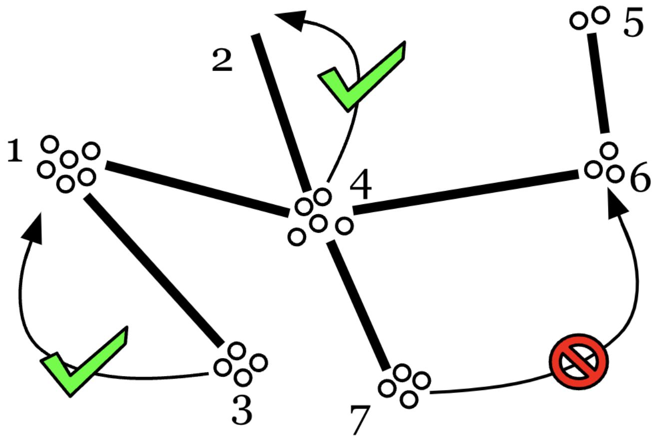


Figure 3.1: Graphical representation of the SIP model. A particle wants to jump from site 3 to site 1, this is possible. Again, the jump of a particle from site 4 to 2 is again possible, but the jump from site 7 to 6 is not, since there still is no edge connecting the two sites.

As said before, we restrict our view in this chapter to the process on a graph with only two edges. This simplification can be justified in the following way: For two sites sharing an edge, $i, j \in V$, $(i, j) \in E$, let $L_e$ be the generator working on the configuration of particles at site $i$ and $j$, $(\eta_i, \eta_j)$. Now the total generator for the configuration of particles in the entire system, $\eta$, is the sum of the generators for every edge, i.e.

$$Lf(\eta) = \sum_{(i,j)=e\in E} L_e f(\eta_i, \eta_j).$$

For the different variations of the model, we are interested in finding steady state wealth distributions. To this end, we set out to find an invariant product measure of the particles on the two sites. Recall from the previous chapter that an invariant measure is a probability measure which is invariant under the passage of time in the process; this is a steady state of the system. On top of that, we want to find a measure that is product. A product measure in this case is a measure on both sites, that is the product of a probability measure on the amount of particles of site 1, $n$, and the amount of particles of site 2, $m$; $\mu(n,m) = \mu_1(n)\mu_2(m)$. We want to find measures that are product, because we are interested in a distribution of wealth on a system not only having 2 agents, but $N$ agents for example. The measure on N agents, $\mu(x_1,\cdots,x_N)$ is not easy to calculate from the distribution in the two-agent case, except for when we have product measures, then it will be simply a product of the measures: $\mu(x_1,\ldots,x_N) = \prod_{i=1}^N \mu_i(x)$.

Let us start with the first variation of the model. The model in which at most one particle can jump from site to site at a time.

## 3.2. One particle jumps

Let $n$ denote the amount of particles on site 1 and $m$ the amount of particles on site 2. Consider the following single edge process. A particle jumps from site 1 to site 2 with rate $n(\alpha_2 + m)$ and from site 2 to site 1 with rate $m(\alpha_1 + n)$. If we denote by $(n,m)$ the amount of particles on site 1 and site 2 respectively the rates to jump from configuration to configuration become:

$$(n,m) \to (n-1, m+1) \text{ with rate } n(\alpha_2 + m) \tag{3.1}$$

$$(n,m) \to (n+1, m-1) \text{ with rate } m(\alpha_1 + n) \tag{3.2}$$

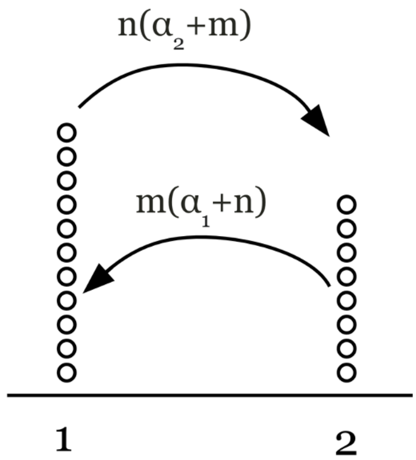


Figure 3.2: Graphical representation of the SIP model for two sites. At site 1 there are $n$ particles and at site 2 there are $m$ particles. The wealth attraction of site 1 is $\alpha_1$ and that of site 2 is $\alpha_2$. Particles jump from one site to the other with rates given by 3.1 and 3.2.

The parameters $\alpha_1$ and $\alpha_2$ are characterising parameters for the sites. For the interacting particle system, they are the characteristic parameter which identify the amount of attraction a certain site has. In the view of our model, these $\alpha$'s can be seen as the extent to which an agent attracts wealth. When $\alpha_i$ is of high value, it can be seen from 3.1 and 3.2 that the rate to jump to site $i$ will become higher.

The change of configuration of particles on the system is a Markov process. We are interested in how the configuration of particles $(n,m)$ changes in time and want to look at the steady state wealth distribution between the two agents. We therefore state the Markov generator of this process on the configuration $(n,m)$.

$$Lf(n,m) = n(\alpha_2 + m)\big(f(n-1,m+1) - f(n,m)\big) + m(\alpha_1 + n)\big(f(n+1,m-1) - f(n,m)\big) \tag{3.3}$$

We are now ready to show an invariant product measure for this variant of the simple inclusion process.

**Theorem 3.2.1.** For $0 \le \lambda < 1$, the SIP process with Markov generator given by 3.3 has invariant product measure:

$$\mu(n,m) = \frac{\lambda^{n+m}}{n!m!}\frac{\Gamma(\alpha_1+n)\Gamma(\alpha_2+m)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}(1-\lambda)^{\alpha_1+\alpha_2} \tag{3.4}$$

*Proof.* We are going to find an expression for $\mu(n,m) = \mu_1(n)\mu_2(m)$ by assuming that detailed balance holds:

$$\mu(n,m)c(n,m;n-1,m+1) = \mu(n-1,m+1)c(n-1,m+1;n,m)$$

By substituting the given rates from 3.1 and 3.2 in the detailed balance equation, we get

$$\mu_1(n)\mu_2(m)n(\alpha_1+m) = \mu_1(n-1)\mu_2(m+1)(m+1)(\alpha_2+n-1)$$

or equivalently,

$$\frac{\mu_1(n)}{\mu_1(n-1)}\frac{n}{\alpha_1+n-1} = \frac{\mu_2(m+1)}{\mu_2(m)}\frac{m+1}{\alpha_2+m}. \tag{3.5}$$

Since the LHS and RHS of 3.5 are equal and dependent on a different variable, they must be equal to a constant $\lambda$. The condition $0 \le \lambda \le 1$ is needed to ensure that $\mu$ will be a probability measure. We write 3.5 as

$$\frac{\mu_1(n)}{\mu_1(n-1)}\frac{n}{\alpha_1+n-1} = \frac{\mu_2(m+1)}{\mu_2(m)}\frac{m+1}{\alpha_2+m} = \lambda.$$

We can make a telescoping product:

$$\begin{aligned}\frac{\mu_1(n)}{\mu_1(0)} &= \prod_{k=1}^{n}\frac{\mu_1(k)}{\mu_1(k-1)}\\ &= \prod_{k=1}^{n}\frac{\lambda(\alpha_1+k-1)}{k}\\ &= \frac{\lambda^n}{n!}\prod_{k=1}^{n}(\alpha_1+k-1)\\ &= \frac{\lambda^n}{n!}(\alpha_1)(\alpha_1+1)\cdots(\alpha_1+n-1)\end{aligned}$$

For the gamma function, the property $\Gamma(x+1) = x\Gamma(x)$ then tells us that

$$(\alpha_1)(\alpha_1+1)\cdots(\alpha_1+n-1) = \frac{\Gamma(\alpha_1+n)}{\Gamma(\alpha_1)}$$

We see from this calculation that $\mu_1(n) = \frac{\lambda^n}{n!}\frac{\Gamma(\alpha_1+n)}{\Gamma(\alpha_1)}$, however we need to make sure that $\mu_1(n)$ is a probability measure, i.e. $\sum_{k=0}^{\infty}\mu_1(k) = 1$. To find the normalisation term, we calculate the following infinite sum:

$$\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)} = (1-\lambda)^{-\alpha}, \tag{3.6}$$

so the normalisation term for this measure is $(1-\lambda)^{\alpha}$. To see that 3.6 holds, we are going to calculate the Taylor expansion of $(1-\lambda)^{-\alpha}$ around $\lambda = 0$.

$$(1-\lambda)^{-\alpha} = \sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\left.\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}(1-\lambda)^{-\alpha}\right|_{\lambda=0} \tag{3.7}$$

Now calculating $\left.\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}(1-\lambda)^{-\alpha_1}\right|_{\lambda=0}$ explicitly for a few different $n$:

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}\lambda}(1-\lambda)^{-\alpha_1}\right|_{\lambda=0} &= \alpha_1\\ \left.\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2}(1-\lambda)^{-\alpha_1}\right|_{\lambda=0} &= \alpha_1(\alpha_1+1)\\ \left.\frac{\mathrm{d}^3}{\mathrm{d}\lambda^3}(1-\lambda)^{-\alpha_1}\right|_{\lambda=0} &= \alpha_1(\alpha_1+1)(\alpha_1+2)\\ &\cdots\end{aligned}$$

By using once more the fact that for the gamma function, $\Gamma(x+1) = x\Gamma(x)$, we see that:

$$\frac{\mathrm{d}^n}{\mathrm{d}\lambda^n}(1-\lambda)^{-\alpha_1}\bigg|_{\lambda=0} = \frac{\Gamma(\alpha_1+n)}{\Gamma(\alpha_1)}$$

From this we can conclude that the LHS of 3.6 is exactly the Taylor expansion of $(1-\lambda)^{-\alpha}$ around $\lambda = 0$. This gives us that

$$\mu_1(n) = \frac{\lambda^n}{n!}\frac{\Gamma(\alpha_1+n)}{\Gamma(\alpha_1)}(1-\lambda)^{\alpha_1}. \tag{3.8}$$

Through almost the exact same calculation we find

$$\mu_2(m) = \frac{\lambda^m}{m!}\frac{\Gamma(\alpha_2+m)}{\Gamma(\alpha_2)}(1-\lambda)^{\alpha_2}$$

Since we assumed the measure on both sites to be product, we obtain 3.4 by simply multiplying $\mu_1$ and $\mu_2$ □

We now formulate a property of the measure found in theorem 3.2.1. To show the dependence of $\mu$ on $\alpha$, we introduce the notation $\mu_\alpha$.

**Theorem 3.2.2.** Consider the Markov process with generator given by 3.3 and the stationary product measure for this process given by 3.8. Let $X$ and $Y$ be two independent random variables, such that $X \sim \mu_\alpha$ and $Y \sim \mu_{\alpha'}$, we have that

$$X + Y \sim \mu_{\alpha+\alpha'} \tag{3.9}$$

*Proof.* Another way of stating 3.9 is that for all $n \in \mathbb{N}$,

$$\sum_{k=0}^{n} \mu_\alpha(k)\mu_{\alpha'}(n-k) = \frac{\lambda^n}{n!}\frac{\Gamma(\alpha+\alpha'+n)}{\Gamma(\alpha+\alpha')}(1-\lambda)^{\alpha+\alpha'}.$$

This is equivalent to showing

$$\sum_{n=0}^{\infty}\zeta^n\sum_{k=0}^{n} \mu_\alpha(k)\mu_{\alpha'}(n-k) = \sum_{n=0}^{\infty}\zeta^n\frac{\lambda^n}{n!}\frac{\Gamma(\alpha+\alpha'+n)}{\Gamma(\alpha+\alpha')}(1-\lambda)^{\alpha+\alpha'}. \tag{3.10}$$

We can evaluate the RHS to

$$\begin{aligned}
\sum_{n=0}^{\infty}\zeta^n\frac{\lambda^n}{n!}\frac{\Gamma(\alpha+\alpha'+n)}{\Gamma(\alpha+\alpha')}(1-\lambda)^{\alpha+\alpha'} &= (\frac{1-\lambda}{1-\lambda\zeta})^{\alpha+\alpha'} \\
&= (\frac{1-\lambda}{1-\lambda\zeta})^{\alpha}(\frac{1-\lambda}{1-\lambda\zeta})^{\alpha'} \\
&= \sum_{n_1=0}^{\infty}\zeta^{n_1}\frac{\lambda^{n_1}}{n_1!}\frac{\Gamma(\alpha+n_1)}{\Gamma(\alpha)}(1-\lambda)^{\alpha}\sum_{n_2=0}^{\infty}\zeta^{n_2}\frac{\lambda^{n_2}}{n_2!}\frac{\Gamma(\alpha'+n_2)}{\Gamma(\alpha')}(1-\lambda)^{\alpha'} \\
&= \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty}\zeta^{n_1+n_2}\frac{\lambda^{n_1}}{n_1!}\frac{\Gamma(\alpha+n_1)}{\Gamma(\alpha)}(1-\lambda)^{\alpha}\frac{\lambda^{n_2}}{n_2!}\frac{\Gamma(\alpha'+n_2)}{\Gamma(\alpha')}(1-\lambda)^{\alpha'}
\end{aligned}$$

By changing the indices of summation to $n = n_1 + n_2$ and $k = n_1 \in \{0, \ldots, n\}$, we obtain that the RHS of 3.10 equals

$$\sum_{n=0}^{\infty}\zeta^n\sum_{k=0}^{n}\frac{\lambda^k}{k!}\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}(1-\lambda)^{\alpha}\frac{\lambda^{n-k}}{(n-k)!}\frac{\Gamma(\alpha'+n-k)}{\Gamma(\alpha')}(1-\lambda)^{\alpha'},$$

which is precisely the LHS of 3.10. This completes the proof. □

## 3.3. Uniform redistribution model

We are now going to consider a variation on the SIP model described in the previous chapter where more particles can hop in one step. The process starts, once more, with $n$ particles on the first site and $m$ particles on the second site. We define the following Markov process on the configuration of particles $(n, m)$: An exponential clock with parameter $\lambda = 1$ is ticking between the two sites, when it goes off, the procedure of redistribution begins. At time of redistribution, $k$ particles go to the first place and the rest of the particles go to the second place, with $k \sim$ discrete uniform $(0, n+m)$. The rates for this process are given by:

$$(n, m) \to (k, n+m-k) \text{ with rate } \mathbb{P}(X = k) \tag{3.11}$$

Since $k$ is uniform, $\mathbb{P}(X = k) = \frac{1}{n+m+1}$. The Markov generator for this process is:

$$Lf(n, m) = \sum_{k=0}^{n+m} \frac{1}{n+m+1} (f(k, n-m-k) - f(n, m)) \tag{3.12}$$

The following theorem gives an invariant product measure for the uniform jump process.

**Theorem 3.3.1.** The Markov jump process with generator given by 3.12 has an invariant product measure

$$\mu(n, m) = (1-\lambda)^2 \lambda^{n+m}. \tag{3.13}$$

*Proof.* We prove that this is an invariant product measure by showing that detailed balance holds:

$$\mu(n, m) c(n, m; k, n+m-k) = \mu(k, n+m-k) c(k, n+m-k; n, m)$$

filling in the rates from 3.11, we get

$$\mu(n, m) \frac{1}{n+m+1} = \mu(k, n+m-k) \frac{1}{n+m+1} \tag{3.14}$$

Now filling in 3.13:

$$(1-\lambda)^2 \lambda^{n+m} \frac{1}{n+m+1} = (1-\lambda)^2 \lambda^{k+n+m-k} \frac{1}{n+m+1}$$

we see that both sides are equal, and complete the proof. □

We can obtain another invariant measure, $\mu'$ by conditioning the measure from 3.13 on the total wealth in the system to be a fixed quantity, $s$.

$$\begin{aligned} \mu'(n, m) &= \mu(n, m | n+m = s) \\ &= \frac{(1-\lambda)^2 \lambda^{n+m}}{\sum\limits_{n', m' : n'+m' = s} (1-\lambda)^2 \lambda^{n'+m'}} \\ &= \frac{1}{n+m+1} \mathbb{1}_{\{n+m=s\}} \end{aligned}$$

We show that this is also an invariant measure for the uniform jump process.

**Theorem 3.3.2.** The uniform jump process with the total quantity of wealth on the two sites being $s$, has invariant measure

$$\mu(n, m) = \frac{1}{n+m+1} \mathbb{1}_{\{n+m=s\}} \tag{3.15}$$

*Proof.* Once more, we prove that this is an invariant product measure by showing that detailed balance holds:

$$\begin{gathered} \mu(n, m) c(n, m; k, n+m-k) = \mu(k, n+m-k) c(k, n+m-k; n, m) \\ \frac{1}{n+m+1} \mathbb{1}_{\{n+m=s\}} \frac{1}{n+m+1} = \frac{1}{n+m+1} \mathbb{1}_{\{k+n+m-k=s\}} \frac{1}{n+m+1} \end{gathered}$$

We complete the proof by seeing that the LHS and RHS are equal. □

In the next section, we look at a more general case of the simple inclusion process.

## 3.4. Beta binomial redistribution model

Now we are going to consider a more general case of the SIP model. We show that the simple inclusion process from section 3.2 is a particular case of this more general SIP model, both having the same stationary product measures. As with the uniform jump process from section 3.3, when the exponential clock goes off, $k$ particles redistribute to the first agent and the rest to the second agent. But in this model, $k$ is beta binomially distributed with parameters $\alpha_1, \alpha_2$ ($k \sim \text{BetaBin}(n+m; \alpha_1, \alpha_2)$). The rates of this process are given by:

$$(n, m) \to (k, n+m-k) \text{ with rate } \mathbb{P}(X=k), \tag{3.16}$$

where $\mathbb{P}(X=k)$ is given by:

$$\mathbb{P}(X=k) = \binom{n+m}{k} \frac{1}{B(\alpha_1, \alpha_2)} \int_0^1 p^{k+\alpha_1-1}(1-p)^{n+m-k+\alpha_2-1} \mathrm{d}p \tag{3.17}$$

$$= \binom{n+m}{k} \frac{B(\alpha_1+k, \alpha_2+n+m-k)}{B(\alpha_1, \alpha_2)} \tag{3.18}$$

A random variable that is beta binomially distributed with parameters $(n, m, \alpha_1, \alpha_2)$ is a random variable that is binomially distributed with parameter $p$. Contrary to the binomial distribution, $p$ is not a fixed value, $p$ is a variable that is beta distributed with parameters $\alpha_1$ and $\alpha_2$.

Note how for $\alpha_1 = \alpha_2 = 1$ the model with $k \sim \text{Betabin}(n+m; \alpha_1, \alpha_2) = \text{Betabin}(n+m; 1, 1)$ is exactly the model of section 3.3:

$$\begin{aligned}
\mathbb{P}(X=k) &= \binom{n+m}{k} \frac{B(\alpha_1+k, \alpha_2+n+m-k)}{B(\alpha_1, \alpha_2)} \\
&= \binom{n+m}{k} \frac{B(1+k, 1+n+m-k)}{B(1, 1)} \\
&= \frac{\Gamma(n+m+1)}{\Gamma(k+1)\Gamma(n+m-k+1)} \frac{\Gamma(k+1)\Gamma(n+m-k+1)}{\Gamma(n+m+2)} \\
&= \frac{1}{n+m+1}
\end{aligned}$$

**Theorem 3.4.1.** The stationary product measure for this model with jump rates given in 3.16 has a stationary product measure given by:

$$\mu(n, m) = \mu_1(n)\mu_2(m),$$

with

$$\mu_1(n) = \frac{\lambda^n}{n!} \frac{\Gamma(\alpha_1+n)}{\Gamma(\alpha_1)} (1-\lambda)^{\alpha_1} \tag{3.19}$$

$$\mu_2(m) = \frac{\lambda^m}{m!} \frac{\Gamma(\alpha_2+m)}{\Gamma(\alpha_2)} (1-\lambda)^{\alpha_2} \tag{3.20}$$

*Proof.* We are going to proof that this is an invariant product measure for this process, by showing that detailed balance holds.

$$\mu(n, m)c(n, m; k, n+m-k) = \mu(k, n+m-k)c(k, n+m-k; n, m) \tag{3.21}$$

note that the rate to go from configuration $(n, m)$ to $(k, n+m-k)$ is equal to the probability $\mathbb{P}(X=k)$ and the rate to go the opposite direction, from $(k, n+m-k)$ to $(n, m)$ is the probability $\mathbb{P}(X=n)$, filling this in in 3.21, we obtain

$$\mu(n, m)\mathbb{P}(X=k) = \mu(k, n+m-k)\mathbb{P}(X=n).$$

By applying what we know from 3.18 we get that

$$\mu_1(n)\mu_2(m)\binom{n+m}{k}\frac{B(\alpha_1+n, \alpha_2+n+m-k)}{B(\alpha_1, \alpha_2)} = \mu_1(k)\mu_2(n+m-k)\binom{n+m}{n}\frac{B(\alpha_1+n, \alpha_2+m)}{B(\alpha_1, \alpha_2)}$$

$$\mu_1(n)\mu_2(m)\binom{n+m}{k}B(\alpha_1+n, \alpha_2+n+m-k) = \mu_1(k)\mu_2(n+m-k)\binom{n+m}{n}B(\alpha_1+n, \alpha_2+m)$$

By filling in 3.19 and 3.20 we get:

$$\frac{\lambda^n}{n!}\frac{\Gamma(\alpha_1+n)}{\Gamma(\alpha_1)}(1-\lambda)^{\alpha_1}\frac{\lambda^m}{m!}\frac{\Gamma(\alpha_2+m)}{\Gamma(\alpha_2)}(1-\lambda)^{\alpha_2}\binom{n+m}{k}\lambda B(\alpha_1+k,\alpha_2+n+m-k)=$$

$$\frac{\lambda^k}{k!}\frac{\Gamma(\alpha_1+k)}{\Gamma(\alpha_1)}(1-\lambda)^{\alpha_1}\frac{\lambda^{(n+m-k)}}{(n+m-k)!}\frac{\Gamma(\alpha_2+n+m-k)}{\Gamma(\alpha_2)}(1-\lambda)^{\alpha_2}\binom{n+m}{n}B(\alpha_1+n,\alpha_2+m),$$

by using that $\binom{n+m}{k}=\frac{n!}{k!(n+m-k)!}$ a lot of terms cancel on both sides, leaving us with:

$$\Gamma(\alpha_1+n)\Gamma(\alpha_2+m)B(\alpha_1+k,\alpha_2+n+m-k)=\Gamma(\alpha_1+k)\Gamma(\alpha_2+n+m-k)B(\alpha_1+n,\alpha_2+m)$$

By using the identity: $B(x,y)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ we obtain:

$$\Gamma(\alpha_1+n)\Gamma(\alpha_2+m)\frac{\Gamma(\alpha_1+k)\Gamma(\alpha_2+n+m-k)}{\Gamma(\alpha_1+\alpha_2+n+m)}=\Gamma(\alpha_1+k)\Gamma(\alpha_2+n+m-k)\frac{\Gamma(\alpha_1+n)\Gamma(\alpha_2+m)}{\Gamma(\alpha_1+\alpha_2+n+m)}$$

We see that the LHS and RHS are equal, and thus completing the proof. □

In the next section, we look at how this discrete model relates to a continuous wealth distribution model.

## 3.5. Continuous redistribution model

We introduce the continuous variant of the wealth distribution model between two sites. Consider a continuous amount of wealth on two sites, $x$ and $y$. We define a Markov process on the amount of wealth as follows: At each time an exponential clock rings, a parameter $U$ is drawn from a distribution and the redistribution of the wealth goes as follows:

$$(x,y)\rightarrow((x+y)U,(x+y)(1-U)), \tag{3.22}$$

When U is uniformly distributed, the process described is exactly the KMP model as described in [7]. This model describes the energy transfer in a system of one-dimensional oscillators. When U has a beta distribution with parameters $s$ and $s$, the process is the same as the thermalised Brownian Energy process. This is explored in depth in [3].

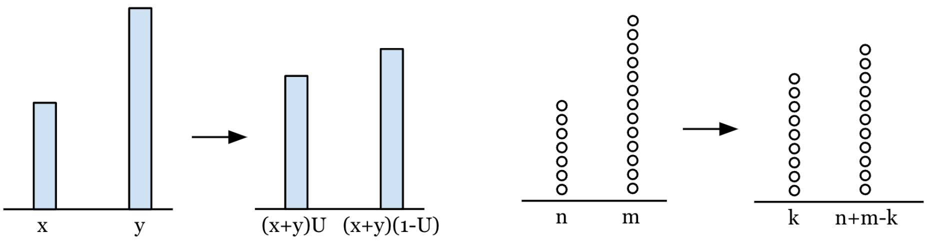


Figure 3.3: The redistribution process for the continuous wealth distribution model and the discrete SIP model.

The continuous model can be seen as a discrete model with a ton of particles, after a proper rescaling. The continuous model can be approximated by the discrete model with a lot of particles after a proper rescaling. In this discussion, we won't provide a formal proof for this assertion; instead, we offer an intuitive explanation. Consider a configuration of wealth on the two sites, $(x,y)$, $x,y\geq 0$. If we define the discrete jump process starting with $(n,m)=(\lfloor xN\rfloor,\lfloor yN\rfloor)$, then as $N\rightarrow\infty$, the discrete jump process $(\frac{n(t)}{N},\frac{m(t)}{N})$ will approach the continuous redistribution process starting from $(x,y)$. In [10] it is shown that for a jump process where $k$ particles get distributed to one site and the rest to the other site, that for $k\sim\text{BetaBin}(n+m;\alpha,\alpha)$ the generator corresponding to this discrete process will converge to the continuous process with redistribution given as in 3.22, with $U\sim\text{Beta}(\alpha,\alpha)$. It can be shown that a continuous process with redistribution as in 3.22, with $U\sim U(0,1)$, can be approximated in the way we described above by a discrete jump process with $k\sim$ discrete uniform $(0,n+m)$.

# 4
# The Pareto distribution and random $\alpha$'s

In the previous chapter we looked at invariant measures for different variants of the SIP model. In this chapter we zoom in on the variant where one particle can jump at a time. However, the properties we derive inn this chapter will also hold for the other redistribution models we have seen. In the previous chapter, $\alpha$ was said to be a fixed value. In this chapter, we look at the case where $\alpha$ has its own distribution, $\psi(\alpha)$. This results in the marginal probability distribution $\nu(n) = \int_0^\infty \mu_\alpha(n)\psi(\alpha)\mathrm{d}\alpha$. Pareto (or power law)-like behaviour is found empirically for the distribution of wealth in an economy [4] . We therefore are interested in finding Pareto (or power law)-like behaviour for $\nu(n)$. First, we will give conditions on $\psi(\alpha)$ for which $\nu(n)$ will have a weak asymptotic power law lower bound. After that, we give the condition on $\psi(\alpha)$ for which we will obtain asymptotic equality to a power law of the partial sums of the sequence $(\nu(n)n(n-1)\cdots(n-k+1))_{n\geq 0}$ for some $k \in \mathbb{N}$. Additionally, this same condition on $\psi(\alpha)$ in combination with a condition on the coefficients of this sequence, will result in asymptotic power law behaviour of $\nu(n)$.

We start the chapter of by introducing the Pareto distribution.

## 4.1. Pareto distribution

The Pareto distribution is a probability distribution discovered by the Italian mathematician Vilfredo Pareto who used this distribution to model the wealth of the richest people in Italy at the end of the 19$^{\text{th}}$ century. As we will see in section 4.1.2, the distribution has since been used numerous times to model the wealth distribution of the upper part of the economy. The Pareto distribution has many variants. In view of our discussion, we will only look at the Pareto type I distribution. The Pareto type I distribution is of the form

$$\phi(x) = x^{-\beta} \tag{4.1}$$

with $\beta > 0$ being a shape parameter. In subsequent references to the Pareto distribution, it is specifically the Pareto type I distribution being mentioned. The Pareto distribution is a distribution that follows a power law. When a distribution follows a power law, a change in one variable generates a change in another variable proportional to a power, as can be observed in 4.1. Throughout this chapter, the terms power law and Pareto are used interchangeably.

In the next section, two features of the Pareto distribution are explained that give rise to inequality in light of wealth distribution; *scale invariance* and the notion of *fat tails*.

### 4.1.1. Inequality

When examining the probability density function of a random variable that is exponentially distributed, it becomes evident that it diminishes quickly towards the tail. The further we come from the origin, the lower the probability becomes of finding a random variable with that value. The tail of a Pareto distribution decays slowly. In figure 4.1 we see that it does so way slower than that of an exponential distribution, a distribution that is not fat-tailed. The tail of a fat-tailed distribution decays according to a power law. When the tail of a distribution decays like a power law, the distribution is said to be *fat-tailed.* To further characterise this feature, we need the notion of asymptotic equality:

**Definition 4.1.1** (Asymptotic equality)**.** We say that $f$ and $g$ are **asymptotically equal**, i.e. $f(x) \asymp g(x)$ if and only if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1. \tag{4.2}$$

When a distribution is said to be fat-tailed, there's a non-negligible probability of observing values far from the mean. For $X_i$ identical independently distributed random variables distributed according to some fat-tailed distribution, we have that

$$\sum_{i=1}^{N} X_i \asymp \max_{1\leq i\leq N} X_i. \tag{4.3}$$

The proof of 4.3 is left out of the discussion, but the statement tells us that as $N$ grows very large, the size of the sum of $X_i's$ will get dictated by the largest $X_i$. For a wealth distribution that is governed by a Pareto distribution, this means that if we keep adding more and more agents to our model, the total wealth in the system will become asymptotically equal to the wealth of the richest agent. This is inequality in its essence.

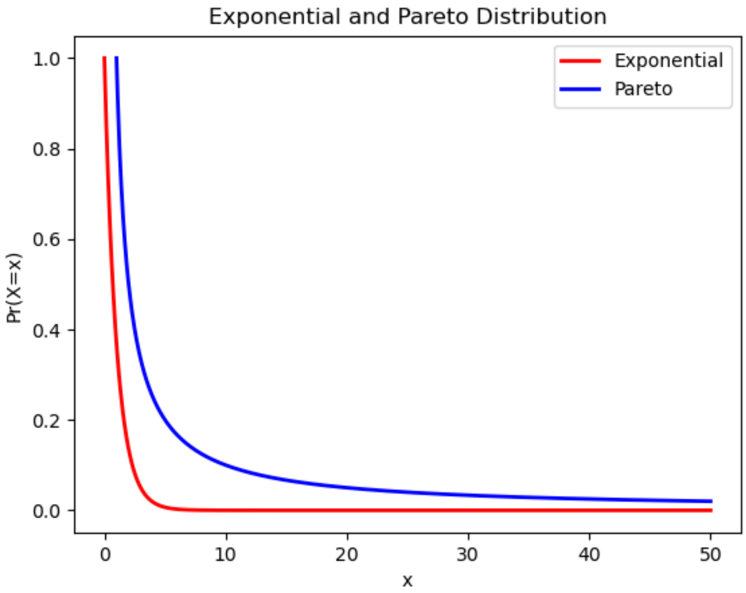


Figure 4.1: Plot of a Pareto distribution with shape parameter $\beta = 1$ and an exponential distribution with $\lambda = 1$. The exponential distribution goes to a probability of zero very quickly compared to the Pareto distribution. From this picture it can be seen that the tail of the Pareto distribution is *fat*.

Consider the wealth for an agent $X$ to be exponentially distributed; given the fact that someone has a certain amount of wealth, $x$, the probability that we find someone that has twice this amount of wealth, becomes lower as $x$ grows:

$$\mathbb{P}(X \geq 2x | X \geq x) = \frac{e^{-2x}}{e^{-x}} = e^{-x}.$$

When the wealth of agents $X$ is Pareto distributed, this probability becomes:

$$\mathbb{P}(X \geq 2x | X \geq x) = \frac{2x^{-\beta}}{x^{-\beta}} = 2^{-\beta}, \tag{4.4}$$

Note how the probability in 4.4 is not dependent on $x$. A probability distribution that has this property, is called *scale invariant*. The scale invariance of the Pareto distribution implies that if we find an agent with more than €100, the probability of finding an agent with more than €200 is just as big as the probability of finding an agent having more than €2.000.000 having found an agent with €1.000.000. The characteristic of scale invariance also sets the stage for the occurrence of outliers and, consequently, contributes to the emergence of inequality.

### 4.1.2. Empirical examples

The reason that we want to show a Pareto distribution in the distribution of wealth in the simple inclusion process becomes clear from empirical data. Data for looking at the distribution of wealth in the world is not readily available. Banks, for good reasons, do not have this data publicly available. However, the Pareto distribution can be empirically observed in other data that are closely related to wealth. In [4] an overview of articles is given that look at distribution of incomes in certain countries. They all find that the upper part[1] of the incomes follow a Pareto distribution.

The occurrence of the Pareto distribution in the distribution of wealth is not only observed in the wealth distribution of recent times. In [1], the wealth distribution of the upper class in ancient Egypt is researched. Excavations of the ancient city of Akhetaten gave valuable insights on the distribution of wealth of the $14^{\text{th}}$ century BC, the time that this city was populated. It was found that the distribution of the wealth of the upper class of the population was of Pareto type. The remarkable conclusion from this research is that the specific parameters for the Pareto fit on the data are not far of the parameters that are found for wealth distributions nowadays.

## 4.2. Quenched and annealed random $\alpha$

In the preceding discussion, we talked about $\alpha$ being the parameter that defined the wealth attraction of a specific agent. In this preceding discussion, $\alpha$ was a fixed value for every agent. For a specific value of $\alpha$ we found the invariant measure for the one jump process. From now on, to highlight the dependence on $\alpha$, let us denote this invariant measure with $\mu_\alpha$, i.e.

$$\mu_\alpha(n) = \frac{\lambda^n}{n!}\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)}(1-\lambda)^\alpha. \tag{4.5}$$

If we make the connection between our agent model and the real world, it is not entirely clear as to why the wealth attraction of some people, companies or banks is higher than that of others. It is of course dependent on many different factors, but the reason behind why the intrinsic wealth attraction for a specific agent is a certain value cannot be said explicitly. To model this uncertainty and perceived randomness, we will look at the different $\alpha$'s for agents to be independent random variables drawn from an underlying distribution $\psi(\alpha)$. We assume that this wealth attraction is non-negative, $\psi(\alpha) = 0$ for $\alpha < 0$ and since $\psi(\alpha)$ is a probability distribution, we have that $\psi(\alpha) \geq 0$ and $\int_0^\infty \psi(\alpha)\mathrm{d}\alpha = 1$.

The expression for $\mu_\alpha$ as given by 4.5, is what is called a *quenched* probability distribution. This is the distribution with $\alpha$ drawn from the distribution and fixed. If in our distribution we want to take the underlying probability distribution of $\psi(\alpha)$ into account, there is the notion of the *annealed* probability distribution. This can be seen as the wealth distribution averaged over different values of $\alpha$. We let $\nu(n)$ denote the annealed probability distribution. We can write $\nu(n)$ as the following marginal distribution:

$$\nu(n) = \int_0^\infty \mu_\alpha(n)\psi(\alpha)\mathrm{d}\alpha. \tag{4.6}$$

The rest of our discussion will be dedicated to giving conditions on $\psi(\alpha)$ for which the tail of our annealed distribution will show power law like behaviour. In the next section, we will give the first condition on $\psi(\alpha)$ that causes $\nu(n)$ to have a weak asymptotic power law lower bound.

## 4.3. Conditions for weak asymptotic power law lower bound

In this section, we set out to find a condition on $\psi(\alpha)$ for which the tail of the distribution of $\nu$ will have a weak power law lower bound, i.e. for some $\gamma \geq 0$

$$\limsup \nu(n) n^\gamma = \infty. \tag{4.7}$$

To emphasise this weak power law lower bound, we introduce for 4.7 the notation:

$$\nu(n) \gtrsim n^{-\gamma}. \tag{4.8}$$

Note that this is not a strict lower power law lower bound for $\nu(n)$. It only tells us that after a certain $n_0$, $\nu(n)$ has infinitely many points that are above the graph of $n^{-\gamma}$. It may very well be that after $n_0$, $\nu(n)$ also has many points below $n^{-\gamma}$.

To find the right condition, we are going to define the generating function, $G(z)$:

$$G(z) = \sum_{n=0}^{\infty} \nu(n) z^n \tag{4.9}$$

[1] The upper 5-10%.

In the rest of our discussion, when referring to the generating function $G(z)$, we mean the function as described above, with $\nu(n)$ as given in 4.6. The following definition we will need in our further discussion.

**Definition 4.3.1** ($\text{k}^{\text{th}}$ moment)**.** Let $X$ be a random variable and $\phi$ be its underlying probability distribution, i.e. $X \sim \phi$. The **$\mathbf{k^{th}}$ moment** of X is defined for $k \in \mathbb{N}_{\geq 1}$ and is equal to $\mathbb{E}(X^k) = \int_{-\infty}^{\infty} x^k \phi(x)\mathrm{d}x$.

In the rest of our discussion, when talking about the $\text{k}^{\text{th}}$ moment of a distribution $\phi$, we mean the $\text{k}^{\text{th}}$ moment of a random variable $X$ that is distributed according to $\psi(\alpha)$.

The following theorem will be helpful in finding the right conditions on $\psi(\alpha)$ to get the desired power law lower bound.

**Theorem 4.3.1.** The $\text{k}^{\text{th}}$ moment of $\psi(\alpha)$ diverges, if and only if $\left.\frac{\mathrm{d}^k}{\mathrm{d}z^k}G(z)\right|_{z=1}$ diverges.

*Proof.* To prove this, we first are going to calculate $\left.\frac{\mathrm{d}^k}{\mathrm{d}z^k}G(z)\right|_{z=1}$.

$$\begin{aligned} G(z) &= \sum_{n=0}^{\infty} \nu(n) z^n \\ &= \sum_{n=0}^{\infty} \int_0^{\infty} \mu_\alpha(n)\psi(\alpha)\mathrm{d}\alpha z^n \end{aligned}$$

because $\nu(n)$ is a probability measure, the integral converges uniformly so we can exchange summation and integration:

$$\begin{aligned} &= \int_0^{\infty} \sum_{n=0}^{\infty} \mu_\alpha(n)\psi(\alpha)\mathrm{d}\alpha z^n \\ &= \int_0^{\infty} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)}(1-\lambda)^\alpha \psi(\alpha)\mathrm{d}\alpha z^n \\ &= \int_0^{\infty} \sum_{n=0}^{\infty} \frac{(z\lambda)^n}{n!}\frac{\Gamma(\alpha+n)}{\Gamma(\alpha)}(1-\lambda)^\alpha \psi(\alpha)\mathrm{d}\alpha \end{aligned}$$

By applying what we know from 3.7 we obtain

$$G(z) = \int_0^{\infty} \left(\frac{1-\lambda}{1-z\lambda}\right)^\alpha \psi(\alpha)\mathrm{d}\alpha \tag{4.10}$$

Now we can take the derivative:

$$\begin{aligned} \frac{\mathrm{d}^k}{\mathrm{d}z^k}G(z) &= \frac{\mathrm{d}^k}{\mathrm{d}z^k}\int_0^{\infty} \left(\frac{1-\lambda}{1-z\lambda}\right)^\alpha \psi(\alpha)\mathrm{d}\alpha \\ &= \int_0^{\infty} \frac{\mathrm{d}^k}{\mathrm{d}z^k}\left(\frac{1-\lambda}{1-z\lambda}\right)^\alpha \psi(\alpha)\mathrm{d}\alpha \\ &= \int_0^{\infty} (1-\lambda)^\alpha \lambda^k \frac{\alpha(\alpha+1)\cdots(\alpha+k)}{(1-z\lambda)^{\alpha+k}}\psi(\alpha)\mathrm{d}\alpha \end{aligned} \tag{4.11}$$

By evaluating at $z=1$ we get:

$$\left.\frac{\mathrm{d}^k}{\mathrm{d}z^k}G(z)\right|_{z=1} = \int_0^{\infty} \frac{\alpha(\alpha+1)\cdots(\alpha+k)}{(1-\lambda)^k}\lambda^k \psi(\alpha)\mathrm{d}\alpha.$$

Remember that the $\text{k}^{\text{th}}$ moment of $\psi$ is given by $\int_0^{\infty} \alpha^k \psi(\alpha)\mathrm{d}\alpha$.

The limit comparison test tells us that when $\lim_{x\to b}\frac{f(x)}{g(x)} = C < \infty$ for $f(x), g(x) > 0$, then $\int_a^b f(x)\mathrm{d}x$ diverges if and only if $\int_a^b g(x)\mathrm{d}x$ diverges.

$$\begin{aligned}\lim_{\alpha\to\infty}\frac{\alpha^k\psi(\alpha)}{\alpha(\alpha+1)\cdots(\alpha+k)\psi(\alpha)(1-\lambda)^{-k}\lambda^k} &< \lim_{\alpha\to\infty}\frac{\alpha^k\psi(\alpha)}{\alpha^k\psi(\alpha)(1-\lambda)^{-k}\lambda^k}\\ &= \frac{(1-\lambda)^k}{\lambda^k}\\ &< \infty,\end{aligned}$$

this completes the proof. □

The following theorem and the subsequent lemma give the conditions on $\psi(\alpha)$ for $\nu(n)$ to have an asymptotic power law lower bound.

**Theorem 4.3.2.** Divergence of $\left.\frac{\mathrm{d}^k}{\mathrm{d}z^k}G(z)\right|_{z=1}$ implies that $\forall\epsilon>0: \nu(n)\gtrsim n^{-(k+1+\epsilon)}$.

*Proof.* We are going to proof this by contraposition. Suppose the implication does not hold. This means from 4.7 that $\exists\epsilon>0$: $\limsup \nu(n)n^{k+1+\epsilon}<\infty$. In other words, $\exists C>0, n_0\in\mathbb{N}: \forall n\geq n_0,\ \nu(n)\leq Cn^{-(k+1+\epsilon)}$. We will show that $\left.\frac{\mathrm{d}^k}{\mathrm{d}z^k}G(z)\right|_{z=1}$ will converge.

$$\begin{aligned}\frac{\mathrm{d}^k}{\mathrm{d}z^k}G(z) &= \frac{\mathrm{d}^k}{\mathrm{d}z^k}\sum_{n=0}^{\infty}\nu(n)z^n\\ &= \sum_{n=0}^{\infty}\frac{\mathrm{d}^k}{\mathrm{d}z^k}\nu(n)z^n\\ &= \sum_{n=0}^{\infty}\nu(n)n(n-1)\dots(n-k+1)z^{n-k}.\end{aligned} \tag{4.12}$$

Evaluating in $z=1$ gives us:

$$\left.\frac{\mathrm{d}^k}{\mathrm{d}z^k}G(z)\right|_{z=1} = \sum_{n=0}^{\infty}\nu(n)n(n-1)\dots(n-k+1).$$

We split this sum up in two terms, one that sums from 0 to $n_0-1$ and one that sums from $n_0$ to $\infty$:

$$\begin{aligned}\left.\frac{\mathrm{d}^k}{\mathrm{d}z^k}G(z)\right|_{z=1} &= \sum_{n=0}^{n_0-1}\nu(n)n(n-1)\dots(n-k+1)+\sum_{n=n_0}^{\infty}\nu(n)n(n-1)\dots(n-k+1)\\ &\leq \sum_{n=0}^{n_0-1}\nu(n)n(n-1)\dots(n-k+1)+\sum_{n=n_0}^{\infty}Cn^{-(k+1+\epsilon)}n^k\\ &= \sum_{n=0}^{n_0-1}\nu(n)n(n-1)\dots(n-k+1)+C\sum_{n=n_0}^{\infty}n^{-(1+\epsilon)}\\ &\leq \sum_{n=0}^{n_0-1}\nu(n)n(n-1)\dots(n-k+1)+C\sum_{n=0}^{\infty}n^{-(1+\epsilon)}\end{aligned} \tag{4.13}$$

Since the finite sum of a series of finite terms is finite, the left sum in 4.13 converges. The right sum of 4.13 converges, since $\sum_{n=0}^{\infty}\frac{1}{n^\delta}$ converges for all $\delta>0$. This concludes the proof. □

The following lemma is a direct result of the previous theorem and gives the condition on $\psi(\alpha)$ to get the weak asymptotic power law lower bound for $\nu(n)$.

**Lemma 4.3.1.** When the k$^{\text{th}}$ moment of $\psi(\alpha)$ diverges, $\nu(n)\gtrsim n^{-\gamma}$ for some $\gamma\geq 0$.

*Proof.* This follows from combining theorem 4.3.1 and 4.3.2, and taking $\gamma=k+2$ for example. □

Lemma 4.3.1 now tells us that for $v(n)$ to show asymptotic power-law behaviour in the tail, we need $\psi(\alpha)$ for which a higher order moment diverges.

In this chapter, we have shown a condition on $\psi(\alpha)$ for which we got a weak asymptotic power law lower bound for the wealth distribution, $v(n)$. In the next section, we will present conditions on $\psi(\alpha)$ through which we can draw stronger conclusion about the asymptotic behaviour of $v(n)$. This will be done by using two Tauberian theorems. The first theorem connects the divergent behaviour of a power series with the divergent behaviour of its coefficients and the second one relates the behaviour of a function at infinity, with the behaviour of its Laplace transform at 0.

## 4.4. Power law asymptotics

For this section, we need to slightly extend the definition of asymptotic equality. When in 4.2 the limit does not go infinity, but approaches some constant $a$ (from the left or the right), asymptotic equality is denoted by

$$f(x) \asymp g(x) \quad (x \to a-) \text{ or } f(x) \asymp g(x) \quad (x \to a+),$$

meaning

$$\lim_{x\to a-} \frac{f(x)}{g(x)} = 1 \text{ or } \lim_{x\to a+} \frac{f(x)}{g(x)} = 1.$$

To find the proper conditions on $\psi(\alpha)$, we will need two Tauberian theorems. A Tauberian theorem is a statement of the following form: when a sequence or function exhibits regular behaviour, an average of that sequence or function also displays regular behaviour. In the first Tauberian theorem we present, the way in which a certain power series diverges, gets connected to the divergence of the coefficients. In the second theorem, the connection is made between how the Laplace transform of a function diverges and the divergence of the function itself.

Before we are ready to present the relevant Tauberian theorems, we need to introduce three more definitions.

**Definition 4.4.1** (Slowly varying function)**.** A measurable function $\ell : \mathbb{R} \to \mathbb{R}$ is said to be **slowly varying** if and only if for all $a > 0$:

$$\lim_{x\to\infty} \frac{\ell(ax)}{\ell(x)} = 1.$$

A function that is slowly varying is a function whose behaviour at infinity is similar to that of a converging function.

**Definition 4.4.2** (Ultimately increasing/decreasing)**.** A sequence $(a_n)_{n\geq 0} \in \mathbb{R}$ is said to be **ultimately increasing** if $\exists \mathrm{K} \in \mathbb{N}$ such that for all $n \geq \mathrm{K}$ $a_{n+1} \geq a_n$. A sequence is said to be ultimately decreasing if the same conditions hold but $a_{n+1} \leq a_n$.

**Definition 4.4.3** (Ultimately monotone)**.** A sequence $(a_n)_{n\geq 0} \in \mathbb{R}$ is said to be **ultimately monotone** if $\exists \mathrm{K} \in \mathbb{N}$ such that the sequence $(a_n)_{\geq \mathrm{K}}$ is ultimately increasing or decreasing.

We are now ready to introduce the two Tauberian theorems for our discussion, *Karamata's Tauberian Theorem* and *Karamata's Tauberian theorem for Power series.* Both of these are presented and proven in [2].

**Theorem 4.4.1** (Karamata's Tauberian Theorem for Power Series)**.** If $a_n \geq 0$ and the power series $A(s) = \sum_{n=0}^{\infty} a_n s^n$ converges for $s \in [0, 1)$ then for $c, \rho \geq 0$ and $\ell$ slowly varying,

$$\sum_{k=0}^{n} a_k \asymp cn^{\rho} \frac{\ell(n)}{\Gamma(1+\rho)} \tag{4.14}$$

if and only if

$$A(s) \asymp c \frac{\ell\left(\frac{1}{1-s}\right)}{(1-s)^{\rho}} \quad (s \to 1-). \tag{4.15}$$

If $c\rho > 0$ and $(a_n)_{n\geq 0}$ is ultimately monotone, both are equivalent to

$$a_n \asymp cn^{\rho-1} \frac{\ell(n)}{\Gamma(\rho)}. \tag{4.16}$$

**Theorem 4.4.2** (Karamata's Tauberian Theorem)**.** Let $f$ be a non-decreasing right-continuous function on $\mathbb{R}$ with $f(x) = 0$ for all $x < 0$, if $\ell$ varies slowly and $c, \rho \geq 0$, the following statements are equivalent:

$$f(x) \asymp cx^{\rho} \frac{\ell(x)}{\Gamma(1+\rho)} \tag{4.17}$$

$$(\mathscr{L} f)(s) \asymp cs^{-\rho} \ell\left(\frac{1}{s}\right) \quad (s \to 0+). \tag{4.18}$$

Here, $(\mathscr{L} f)(s)$ denotes the Laplace transform of $f(x)$, i.e.:

$$(\mathscr{L} f)(s) = \int_0^{\infty} e^{-sx} f(x) \mathrm{d}x$$

With these two Tauberian theorems in hand, we are now ready to present the condition on $\psi(\alpha)$ asserting power law asymptotics on $\nu(n)$.

**Theorem 4.4.3.** For $0 \leq \lambda < 1$, $u > 0$ and $k \in \mathbb{N}$, such that $k > u$, if

$$\psi(\alpha) \asymp \left(\frac{1}{1+\alpha}\right)^u, \tag{4.19}$$

then that implies for $\nu(n)$ as defined in 4.6

$$\sum_{l=0}^{n} \nu(l) l(l-1) \cdots (l-k+1) \asymp C_1 n^{-(u+1)}, \tag{4.20}$$

for some constant $C_1$. Additionally, when the sequence $(\nu(n)n(n-1)\cdots(n-k+1))_{n\geq 0}$ is ultimately monotone, 4.19 implies

$$\nu(n) \asymp C_2 n^{-(u+1)}, \tag{4.21}$$

for some constant $C_2$.

*Proof.* We are going to proof this by applying theorem 4.4.2 and then 4.4.1.

Let us start by defining $\chi(\alpha) = \alpha(\alpha+1)\cdots(\alpha+k)\psi(\alpha)$. We have that

$$\chi(\alpha) \asymp \frac{\alpha^k}{(1+\alpha)^u}$$

Now define $\rho = k - u$ and let $C' = \Gamma(1+\rho)$, we see that

$$\chi(\alpha) \asymp C' \alpha^{\rho} \frac{1}{\Gamma(1+\rho)}. \tag{4.22}$$

Since $C', \rho > 0$, by theorem 4.4.2 and choosing $\ell = 1$, we have that 4.22 is equivalent with

$$(\mathscr{L}\chi)(s) \asymp C' s^{-\rho} \quad (s \to 0+). \tag{4.23}$$

We let $\beta = \frac{\lambda}{1-\lambda}$ and we define $f(s) = \log(1+\beta s)$. In 4.23, instead of evaluating the Laplace transform of $\chi$ in the point $s$, we can also evaluate it in the point $f(s)$. By the way we have defined $f(s)$, we can rewrite 4.23 to:

$$\int_0^{\infty} \chi(\alpha) e^{-\alpha f(s)} \mathrm{d}\alpha \asymp C' f(s)^{-\rho} \quad (s \to 0+). \tag{4.24}$$

We note that $f(s) \asymp \beta s$ for $s \to 0$. Now 4.24 is equivalent with saying

$$\beta^k \int_0^{\infty} \chi(\alpha) e^{-\alpha f(s)} \mathrm{d}\alpha \asymp C'' s^{-\rho} \quad (s \to 0+),$$

with $C'' = C' \beta^{-\rho+k}$. The $\beta^k$ term on the LHS will prove to be beneficial later on. By the definition of the natural logarithm, we get that

$$\beta^k \int_0^{\infty} \chi(\alpha) \left(\frac{1}{1+\beta s}\right)^{\alpha} \mathrm{d}\alpha \asymp C'' s^{-\rho} \quad (s \to 0+),$$

or equivalently

$$\beta^k \int_0^\infty \alpha(\alpha+1)\cdots(\alpha+k)\left(\frac{1}{1+\beta s}\right)^\alpha \psi(\alpha)\mathrm{d}\alpha \asymp C'' s^{-\rho} \quad (s \to 0+). \tag{4.25}$$

We apply the change of variables $s = 1-z$ and by filling in $\beta = \frac{\lambda}{1-\lambda}$, 4.25 becomes

$$\frac{\lambda^k}{(1-\lambda)^k} \int_0^\infty \alpha(\alpha+1)\cdots(\alpha+k)\left(\frac{1-\lambda}{1-z\lambda}\right)^\alpha \psi(\alpha)\mathrm{d}\alpha \asymp C'' s^{-\rho} \quad (z \to 1-).$$

For $z$ close to 1, this is equivalent with saying

$$\frac{\lambda^k}{(1-z\lambda)^k} \int_0^\infty \alpha(\alpha+1)\cdots(\alpha+k)\left(\frac{1-\lambda}{1-z\lambda}\right)^\alpha \psi(\alpha)\mathrm{d}\alpha \asymp C''(1-z)^{-\rho} \quad (z \to 1-),$$

or in different form:

$$\int_0^\infty \lambda^k (1-\lambda)^\alpha \frac{\alpha(\alpha+1)\cdots(\alpha+k)}{(1-z\lambda)^{\alpha+k}} \asymp C''(1-z)^{-\rho} \quad (z \to 1-).$$

In the LHS we can recognise 4.11:

$$\frac{\mathrm{d}^k}{\mathrm{d}z^k} G(z) \asymp C''(1-z)^{-\rho} \quad (z \to 1-) \tag{4.26}$$

or equivalently by what we know from 4.12

$$\sum_{n=0}^\infty \nu(n) n(n-1)\cdots(n-k+1) z^{n-k} \asymp C''(1-z)^{-\rho} \quad (z \to 1-). \tag{4.27}$$

We recognise the LHS of 4.27 as a power series with coefficients $\nu(n)n(n-1)\cdots(n-k+1)$. We can now apply theorem 4.4.1. From it we find that 4.27 holds if and only if

$$\sum_{l=0}^n \nu(l) l(l-1)\cdots(l-k+1) \asymp C'' \frac{n^\rho}{\Gamma(1+\rho)}.$$

By taking $C_1 = C''$, this proves 4.20. When we have the additional requirement that the series $(\nu(n)n(n-1)\cdots(n-k+1))_{n\geq 0}$ is ultimately monotone, with the help of theorem 4.4.1, with 4.27 we can conclude

$$\nu(n) n(n-1)\cdots(n-k+1) \asymp C'' \frac{n^{\rho-1}}{\Gamma(\rho)}, \tag{4.28}$$

by taking $n(n-1)\cdots(n-k+1)$ to the other side, we get:

$$\nu(n) \asymp C'' \frac{n^{-(u+1)}}{\Gamma(\rho)} \tag{4.29}$$

Now by taking $C_2 = \frac{C''}{\Gamma(\rho)}$, we prove 4.21, thus completing the proof. □

# 5
# Conclusion

In this thesis, the simple inclusion process was studied in the view of wealth distribution in a closed economic system. SIP models the jumping of particles from site to site, with particles having attraction among themselves. The attention was shifted to this process on two sites, and invariant product measures for different variants of the process were found. To obtain the wealth distribution on a graph $G = (V, E)$, we observe that the graph model is essentially the aggregate of the simple two-agent model for each pair of sites connected by an edge. Because the invariant measures found for the two agent model are product, the long term wealth distribution on the model with more sites is simply the product of the invariant measures found in the two agent model.

In the simple inclusion process there is a variable $\alpha$, which can be seen as the extent to which every site attracts particles. In view of our discussion of the wealth distribution, this is the extent to which every agent attracts wealth. Let $\alpha_1$ be this for site 1 and $\alpha_2$ for site 2. When $n$ and $m$ denote the amount of particles on site 1 and site 2 respectively, it was shown that for a simple inclusion process in which one particle can jump from site to site, an invariant product measure on the two sites is of the form:

$$\mu(n,m) = \frac{\lambda^{n+m}}{n!m!}\frac{\Gamma(\alpha_1+n)\Gamma(\alpha_2+m)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}(1-\lambda)^{\alpha_1+\alpha_2} \tag{5.1}$$

This simple inclusion process with one particle jumping was shown to be a special form of the more general simple inclusion process. In this process, there is an exponential clock between the two agents. When the clock rings, $k$ of the total wealth goes to the first site and the rest goes to the second site, with $k \sim \text{BetaBin}(n+m;\alpha_1,\alpha_2)$. This process has an invariant product measure that is exactly 5.1. We also looked at a uniform jump process on two sites, where at the ringing of an exponential clock, $k$ particles get distributed to the first site and the rest to the second, but now with $k \sim$ discrete uniform $(0, n+m)$. This variant of the SIP model was shown to have an invariant prooduct measure of the form

$$\mu(n,m) = \lambda^{n+m}(1-\lambda)^2.$$

By conditioning this measure on the total wealth being a fixed value, $s$, we obtained another invariant product measure

$$\mu(n,m) = \frac{1}{n+m+1}\mathbb{1}_{n+m=s}.$$

In the rest of the discussion, we looked at the long term distribution of wealth in the system, with the wealth attraction parameter for an agent $\alpha$ not fixed, but independent identically distributed random variables distributed according to $\psi(\alpha)$. This resulted in the marginal wealth distribution

$$\nu(n) = \int_0^\infty \mu_\alpha(n)\psi(\alpha)\mathrm{d}\alpha$$

We then set out to look for conditions on $\psi(\alpha)$ for which the tail of the wealth distribution $\nu(n)$ would have Pareto (or power law) behaviour. It was shown that when a higher moment of $\psi(\alpha)$ diverges, that we have a weak asymptotic power law lower bound, i.e. for some $\gamma \geq 0$:

$$\nu(n) \gtrsim n^{-\gamma},$$

meaning $\limsup \nu(n)n^\gamma = \infty$. With the help of two Tauberian theorems, we made the condition on $\psi(\alpha)$ stronger. With $u > 0$, and $k \in \mathbb{N}$, such that $k > u$ for

$$\psi(\alpha) \asymp \left(\frac{1}{\alpha+1}\right)^u \tag{5.2}$$

we proved asymptotic equality of the partial sums of the series $(v(n)n(n-1)\cdots(n-k+1))_{n\geq 0}$, i.e.

$$\sum_{l=0}^{n} v(l)l(l-1)\cdots(l-k+1) \asymp C_1 n^{-(u+1)}$$

for some constant $C_1$. With the additional condition that $(v(n)n(n-1)\cdots(n-k+1))_{n\geq 0}$ is an ultimately monotone sequence, we proved that the long term wealth distribution for an agent is asymptotically equal to a power law:

$$v(n) \asymp C_2 u^{-(u+1)}$$

for some constant $C_2$.

# 6
# Discussion and future work

In this study, the objective was to explore a statistical physical model in the view of wealth distribution in an agent-based model representing the economy. Conditions on the distribution of a specific parameter of the model were presented, for which the resulting wealth distribution would be of Pareto type. With this relatively simple model, it was shown to be possible to model power law asymptotics for the wealth distribution. We are going to present possible ways in which the SIP model central in this discussion can be extended to better represent real-world dynamics.

One way to extend the model, is by adding for every agent a predisposition to save a certain portion of their wealth during every transaction. This introduces a layer of complexity, allowing for a more nuanced understanding of economic interactions and the impact of individual savings behaviours on the overall wealth dynamics. The SIP model with added propensity to save for the agents is presented and further explored in [5]. A second way to extend the model is by introducing reservoirs in the system. These reservoirs allow for in- and outflow of particles in the system, in [3] this is discussed in detail. These wealth reservoirs can be seen as entities that inject or absorb wealth in the economic system. Since we are assuming detailed balance in our system to hold, in order to find stationary measures, there can be no net in- or outflow from the reservoirs. This brings us to a point of discussion, the notion of detailed balance.

Within the simple inclusion process, a crucial but simplifying assumption was the adherence of the model to detailed balance. Detailed balance is an assumption, influencing the equilibrium properties of the model. A direction for extending the model involves relaxing the constraint of detailed balance, allowing for a more dynamic and non-equilibrium representation of the wealth distribution model. This adjustment could provide a deeper exploration of the system's behaviour under conditions where traditional equilibrium assumptions no longer hold, potentially revealing new insights into the underlying dynamics of the modelled wealth distribution.

An interesting tangent arising from the conversation in this thesis is one that is more philosophical or sociological in nature. We saw that in the relatively simple model of our discussion, every agent has its own wealth attraction parameter. By the way the model is defined, a higher $\alpha$ means a higher attraction of wealth. For an individual or agent in the system, what does this parameter $\alpha$ entail in the real world? And how can an individual increase its magnitude? In the view of wealth inequality, we saw conditions on the distribution of $\alpha$ for which we obtained a wealth distribution that closely resembled the distribution of wealth in the world; one in which there is inequality. A question that arises is, for what kind of distribution of $\alpha$ can we get a wealth distribution that is less prone to inequality? And once obtained such a distribution for $\alpha$, how can this be realised in the real world?

# Bibliography


[1] Adel Y Abul-Magd. Wealth distribution in an ancient egyptian society. *Physical Review E*, 66(5):057104, 2002.

[2] Nicholas H Bingham, Charles M Goldie, and Jef L Teugels. *Regular variation*. Number 27. Cambridge university press, 1989.

[3] Gioia Carinci, Cristian Giardina, Claudio Giberti, and Frank Redig. Duality for stochastic models of transport. *Journal of Statistical Physics*, 152:657–697, 2013.

[4] Bikas K Chakrabarti, Anirban Chakraborti, Satya R Chakravarty, and Arnab Chatterjee. *Econophysics of income and wealth distributions*. Cambridge University Press, 2013.

[5] Pasquale Cirillo, Frank Redig, and Wioletta Ruszel. Duality and stationary distributions of wealth distribution models. *Journal of Physics A: Mathematical and Theoretical*, 47(8):085203, 2014.

[6] Cristian Giardina, Frank Redig, and Kiamars Vafayi. Correlation inequalities for interacting particle systems with duality. *Journal of Statistical Physics*, 141:242–263, 2010.

[7] C Kipnis, Carlo Marchioro, and E Presutti. Heat flow in an exactly solvable model. *Journal of Statistical Physics*, 27:65–74, 1982.

[8] Frank Redig. Basic techniques in interacting particle systems, September 2014.

[9] Richard Chace Tolman. *The principles of statistical mechanics*. Courier Corporation, 1979.

[10] B. van Ginkel, F. Redig, and F. Sau. Duality and stationary distributions of the "immediate exchange model" and its generalizations. *Journal of Statistical Physics*, 163:92–112, 2016.

[11] Victor M Yakovenko and J Barkley Rosser Jr. Colloquium: Statistical mechanics of money, wealth, and income. *Reviews of modern physics*, 81(4):1703, 2009.