Professors and Grants

1. Introduction

This note is intended as a compliment and complement to B. Zhang’s very enjoyable “Coconuts and Islanders”, which motivates the Boltzmann distribution in the case where every nonnegative integer is a possible energy-level. Here, our initial focus is, instead, on Boltzmann distributions where 0 and 1 and 10 are the only possible energy-levels. Taking our cue from “Coconuts and Islanders”, we motivate by story.

Let \( N \) be a large positive integer. We analyze three systems for dispensing grant money to \( N \) professors.

The grant rules stipulate each professor receives \$0 or \$1 or \$10. Each professor is identified by a number, from 1 to \( N \).

By a dispensation, we mean a full complement of awards, with a specific amount (\$0 or \$1 or \$10) to Professor#1, a specific amount (\$0 or \$1 or \$10) to Professor#2, etc., up to and including Professor#\( N \).

The first system for awarding grants is very simple:
Congress allocates \( N \) dollars to award to the \( N \) professors.
There are many ways to dispense the \( N \) dollars.
Among all these possible award-dispensations, one is selected randomly, giving equal probability to each possible dispensation.

The main problem is to figure out:
Using this first system, for a given professor, what is the probability of being awarded \$0? \$1? \$10?

Later, second and third probabilistic award systems are described, each of which depends on three parameters \( p, q, r \) satisfying \( p, q, r > 0 \) and \( p + q + r = 1 = q + 10r \). The second system uses an iid system of random-variables, \( X_1, \ldots, X_N \) such that, \( \forall \ell, \) 
\[
\begin{align*}
\text{Pr}[X_\ell = 0] &= p, \\
\text{Pr}[X_\ell = 1] &= q, \\
\text{Pr}[X_\ell = 10] &= r.
\end{align*}
\]
For all $\ell$, the second system awards $X_\ell$ dollars to Professor$\#\ell$. The total dollar payout $X_1 + \cdots + X_N$ is then random;
if $X_1 = \cdots = X_N = 0$, it could be as small as 0 dollars,
and if $X_1 = \cdots = X_N = 10$, it could be as large as $10N$ dollars.

The **third system**
is obtained from the second, by conditioning on $X_1 + \cdots + X_N = N$, so that the total dollar payout is $N$.

**KEY POINT:** With exactly the right choice of $p,q,r$, the first and third systems are shown to be equivalent.

In §6 and §7, we show that this parameter choice is Boltzmann, meaning: ($p,q,r$) is, for some real number $\beta$,
a scalar multiple of $(e^{-0\cdot \beta}, e^{-1\cdot \beta}, e^{-10\cdot \beta})$.
That is, $\exists \beta, C \in \mathbb{R}$ s.t. $(p,q,r) = (C, Ce^{-\beta}, Ce^{-10\beta})$.

The second and third systems are accessible by basic tools of probability theory, while our main problem involves the first system.
However, once we know the first and third systems are equivalent, we can bring these probabilistic tools to bear on the main problem.

Thanks to J. Steif, for pointing out to me that the Discrete Local Limit Theorem, which is described in §9, is the right tool for the main problem, which is solved in §12.

Boltzmann distributions are often motivated by entropy, but, from our perspective,
what’s special about $(p,q,r) = (C, Ce^{-\beta}, Ce^{-10\beta})$ is:
For any $i,j,k \geq 0$, we have
$$p^i q^j r^k = C^{i+j+k} \cdot e^{-\beta(j+10k)},$$
so $p^i q^j r^k$ depends only on: $i+j+k$ and $j+10k$.

In the third system of grant awards, there exists a normalizing $S > 0$ s.t., for any dispensation in which $i$ professors receive $\$0$,
$j$ professors receive $\$1$,
$k$ professors receive $\$10$,
the probability of that dispensation is $p^i q^j r^k / S$,
which is equal to $C^{i+j+k} \cdot e^{-\beta(j+10k)} / S$. 
That probability, then, depends only on
\[ i + j + k, \text{ which is the number of professors,} \]
and \[ j + 10k, \text{ which is the total dollar payout.} \]
So, since the number of professors is constant
and the total dollar payout is constant,
we conclude: all the various award-dispensations are equally likely,
which exactly describes the first system.
Therefore, under the Boltzmann assumption,
the first and third systems are equivalent.

In §14, we expose the inequitable of a randomly selected dispensation.
In fact, assuming \( N \) is sufficiently large, we will show that:
with probability \( > 99\% \), over half of the professors receive \$0.
Thanks to V. Reiner for suggesting
applying Chebyshev’s inequality to a sum of indicator variables,
to transition from individual statistics to population statistics.

In §16 and §17 and §18, we generalize our main results to
finite sets that are different from \( \{0, 1, 10\} \).
The main theorems are Theorem 16.2 and Theorem 17.5.

In §19 and §21, we extend our earlier results to include
degenerate energy-levels, with a finite set of states.
In §22 and §24, we extend these results further to include
cases that involve a countably infinite set of states.

Thanks to C. Prouty for help with many calculations.
For some of his Python code, see §25.

2. Some notation

A box around an expression indicates that it is global,
meaning that it is fixed to the end of these notes.
Unboxed variables are freed at the end of each section, if not earlier.

Let \( \mathbb{R}^* := (-\infty) \cup \mathbb{R} \cup \{\infty\}, \quad \mathbb{Z}^* := (-\infty) \cup \mathbb{Z} \cup \{\infty\}. \)
For any \( s, t \in \mathbb{R}^* \), let
\[ (s; t) := \{x \in \mathbb{R}^* | s < x < t\}, \quad [s; t) := \{x \in \mathbb{R}^* | s \leq x < t\}, \]
\[
\begin{align*}
(s; t) &:= \{ x \in \mathbb{R}^* \mid s < x \leq t \}, \\
[s; t) &:= \{ x \in \mathbb{R}^* \mid s \leq x < t \}.
\end{align*}
\]

For any \( s, t \in \mathbb{R}^* \), let \([s..t] := (s; t) \cap \mathbb{Z}^*\), \((s..t) := [s; t) \cap \mathbb{Z}^*\), \([s..t) := (s; t) \cap \mathbb{Z}^*\), and \((s..t) := [s; t) \cap \mathbb{Z}^*\).

Let \( \mathbb{N} := [1..\infty) \) be the set of positive integers.

For any finite set \( F \), let \( \#F \) be the number of elements in \( F \).

For any infinite set \( F \), let \( \#F := \infty \). Then \( \#\mathbb{Z} = \infty = \#\mathbb{R} \).

### 3. First System of Grant Awards

Let \( N \in \mathbb{N} \). Think of \( N \) as large.

Suppose there are \( N \) professors, numbered 1 to \( N \), who apply, once per year, to the GFA (Grant Funding Agency), seeking funding for the very important work they are doing.

Each year, Congress authorizes \$N for the GFA to dispense to the \( N \) professors.

The GFA has the rule: every award is 0 or 1 or 10 dollars.

The set of grant-dispensations is represented by:

\[
\Omega := \{ \omega : [1..N] \to \{0, 1, 10\} \mid \sum_{\ell=1}^{N} [\omega(\ell)] = N \}.
\]

The GFA has set aside \( \#\Omega \) pieces of paper, and has written down all possible dispensations, one on each piece of paper.

So, for example, there is a piece of paper that says:

- Professors 1 to \( N \) each get $1.
- Another piece of paper says:
  - Professors 1 to \( N - 10 \) each get $1
  - Professors \( N - 9 \) to \( N - 1 \) each get $0
  - Professor \( N \) gets $10.

There are, of course, many, many, many other pieces of paper.

Each year, a GFA bureaucrat places all the pieces of paper in a big bin, then selects one at random and makes the awards as indicated on that piece of paper.

Under this first system of awarding grants, we have:

\[
\forall \omega \in \Omega, \quad \text{the probability that the selected grant-dispensation is } \omega \quad \text{is equal to} \quad 1/(\#\Omega).
\]

Suppose I am one of the professors. Here is our main problem:
Calculate my probability of getting $0$.
Then calculate my probability of getting $1$.
Then calculate my probability of getting $10$.
Approximate answers are acceptable.
In §5 to §12 of this note, we reformulate and then solve this problem.
Spoiler: It’s a Boltzmann distribution, approximately.

4. PARTICLES AND ENERGY

Recall that $N \in \mathbb{N}$. Think of $N$ as large.
Suppose there are $N$ particles, numbered 1 to $N$,
each of which has a certain amount of energy.
Suppose the total energy is $N$, dispensed among the $N$ particles.
Suppose physicists have somehow determined that, for any particle,
its possible energy-levels are: 0 or 1 or 10.
Recall: $\Omega = \left\{ \omega : [1..N] \rightarrow \{0, 1, 10\} \mid \sum_{\ell=1}^{N} [\omega(\ell)] = N \right\}$.
Then $\Omega$ represents the set of energy-dispensations.
Assume that physicists have somehow determined
that this system of particles has a random energy-dispensation
and that all energy-dispensations in $\Omega$ are equally probable.
That is, physicists tell us:
$\forall \omega \in \Omega, \text{ the probability that}$
the energy-dispensation is $\omega$
is equal to $1/(\#\Omega)$.
The equal probability of all energy-dispensations
is a recurring theme in microcanonical-ensemble thermodynamics,
and can often be motivated through
rules of random energy transfer between random pairs of particles.
For examples of this, either see §19 below or
search for “Coconuts and Islanders” by B. Zhang,
and, in particular, see the work leading up to
the last paragraph of §3.2 therein.
In §19 below, instead of particles exchanging energy,
there are professors exchanging dollars,
but the principle is exactly the same.
In Zhang’s exposition,
instead of particles exchanging energy, there are islanders exchanging coconuts, but the principle is exactly the same.

Returning to our $N$ particles, pick any one of them.

Problem: Calculate its probability of having energy-level 0.
Then calculate its probability of having energy-level 1.
Then calculate its probability of having energy-level 10.

Approximate answers are acceptable.

Spoiler: It’s a Boltzmann distribution, approximately.

Except for terminology, this problem is the same as the main problem (end of §3) about professors and grants.

We will go back to professors and grants.
Mathematically it makes no difference, but it’s more fun.

5. Second and third systems of grant awards

In an effort to go paperless, the GFA changes to a new system: In this second system, instead of all those pieces of paper, the GFA chooses $p, q, r > 0$ s.t. $p + q + r = 1$, and then, for each of the $N$ professors, awards $\$ 0$ with probability $p$, $\$ 1$ with probability $q$, $\$ 10$ with probability $r$.

No professor’s award depends in any way on any other professor’s; the awards are independent.

The expected payout, for any professor, is $p \cdot 0 + q \cdot 1 + r \cdot 10$ dollars. Under this second system, there is no guarantee that the total payout will be $\$N$, which is a difficulty that we will discuss later.

However, recognizing that the average award is intended to be $\$1$, the GFA chooses the numbers $p, q, r$ subject to the constraint that $p \cdot 0 + q \cdot 1 + r \cdot 10 = 1$, i.e., $q + 10r = 1$.

For each function $\omega : [1..N] \to \{0, 1, 10\}$, let

\[
\begin{align*}
i_\omega &:= \# \{ \ell \in [1..N] \mid \omega(\ell) = 0 \}, \\
j_\omega &:= \# \{ \ell \in [1..N] \mid \omega(\ell) = 1 \}, \\
k_\omega &:= \# \{ \ell \in [1..N] \mid \omega(\ell) = 10 \};
\end{align*}
\]

that is, $i_\omega$ is the number of professors awarded $\$ 0$ and $j_\omega$ is the number of professors awarded $\$ 1$ and
Then, \( \forall \omega : [1..N] \rightarrow \{0, 1, 10\} \), we have:

- the total number of awards is \( i_\omega + j_\omega + k_\omega \)
- and the total dollar payout is \( i_\omega \cdot 0 + j_\omega \cdot 1 + k_\omega \cdot 10 \), i.e., \( j_\omega + 10k_\omega \).

Then, \( \forall \omega : [1..N] \rightarrow \{0, 1, 10\} \), we have:

\[
i_\omega + j_\omega + k_\omega = N \quad \text{and} \quad j_\omega + 10k_\omega = \sum_{\ell=1}^{N} [\omega(\ell)].
\]

Recall: \( \Omega = \{ \omega : [1..N] \rightarrow \{0, 1, 10\} \mid \sum_{\ell=1}^{N} [\omega(\ell)] = N \} \).

That is, \( \Omega \) is the set of all payout functions

\[
\omega : [1..N] \rightarrow \{0, 1, 10\}
\]

s.t. the total dollar payout is \( N \).

Then: \( \forall \omega : [1..N] \rightarrow \{0, 1, 10\} \), we have:

\[
i_\omega \in \Omega \; \text{iff} \; j_\omega + 10k_\omega = N.
\]

For every \( i, j, k \in [0..N] \),

if \( i + j + k = N \) and \( j + 10k = N \),

then \( \exists \omega \in \Omega \) s.t. \( (i, j, k) = (i_\omega, j_\omega, k_\omega) \);

one such \( \omega : [1..N] \rightarrow \{0, 1, 10\} \) is described by:

\[
\omega = 0 \text{ on } [1..i], \quad \omega = 1 \text{ on } (i, i + j], \quad \omega = 10 \text{ on } (i + j..N].
\]

Let \( [A] := \{(i_\omega, j_\omega, k_\omega) \mid \omega \in \Omega\} \).

Then \( A \) is the set of all \( (i, j, k) \) s.t. \( i, j, k \in [0..N] \) and \( i + j + k = N \) and \( j + 10k = N \).

Under the second system,

- each $0 award happens with probability \( p \) and
- each $1 award happens with probability \( q \) and
- each $10 award happens with probability \( r \).

So, \( \forall \omega : [1..N] \rightarrow \{0, 1, 10\} \), under the second system,

the probability that the grant-dispensation is equal to \( \omega \)

is \( p^{i_\omega}q^{j_\omega}r^{k_\omega} \).

Let \( S := \sum_{\omega \in \Omega} p^{i_\omega}q^{j_\omega}r^{k_\omega} \).

Then \( S \) is the probability (using the second system) that \( \omega \in \Omega \), i.e., the probability that the total payout is exactly \( N \) dollars.

Assuming \( N \) is large, it turns out that \( S \) is close to zero.

So, under this second system,

the probability of paying out exactly \( N \) dollars

is very small.

Congress only allocates \( \$N \) per year for the \( N \) professors.

So, using this second system, each year,
with probability \(1 - S \approx 1\), the GFA will run a surplus or a deficit. On the other hand, since \(q + 10r = 1\), we see that, each year, the expected payout is \$1 per professor, so, each year, the expected total payout is \$N.

So these surpluses and deficits should, over time, cancel one another. Unfortunately, Congress is a paragon of fiscal responsibility, and, as soon as it becomes aware of these surpluses and deficits, it insists that the GFA never again underspend or overspend.

So the GFA changes its system one more time, as follows. Under its **third system**, each year, before announcing any of the awards publicly, the GFA writes out, in an *internal* memo, a *tentative* proposal of awards that, independently, for each of the \(N\) professors, awards \$0 with probability \(p\), \$1 with probability \(q\), \$10 with probability \(r\).

If the memo’s total award payout is NOT equal to \$N, the GFA deems the memo as unacceptable, deletes it, and starts over, making memo after memo, until an acceptable one (meaning payout exactly \$N) appears. Each memo has a probability \(S\) of being acceptable, so, each year, the GFA will likely need to repeat the memo process many times to get to a memo with total payout exactly equal to \$N.

However, as soon as that happens, the GFA uses that first acceptable memo, and makes the awards public.

Mathematically, we are conditioning on the event \(\omega \in \Omega\). So, using the third system, the probability that \(\omega \notin \Omega\) is 0.

Also, for this third system, \(\forall \omega \in \Omega\), the probability of \(\omega\) is \(p^{i^\omega}q^{j^\omega}r^{k^\omega}/S\).

The sum of these probabilities is 1:

\[
\sum_{\omega \in \Omega} \frac{p^{i^\omega}q^{j^\omega}r^{k^\omega}}{S} = \frac{1}{S} \cdot \sum_{\omega \in \Omega} p^{i^\omega}q^{j^\omega}r^{k^\omega} = \frac{1}{S} \cdot S = 1.
\]

This third system is not necessarily equivalent to the first, because in the first system, all the probabilities were \(1/(\#\Omega)\), whereas, in the third system, they are \(p^{i^\omega}q^{j^\omega}r^{k^\omega}/S\).

So a **new question** arises:
Is it possible to choose \( p, q, r > 0 \) in such a way that
\[
p + q + r = 1 \quad \text{and} \quad q + 10r = 1 \quad \text{and} \quad \forall \omega \in \Omega, \quad p^i q^j r^k \omega / S = 1 / (\#\Omega) \quad ?
\]
If yes, then, using that \((p, q, r)\),
the first and third systems are equivalent.
We will see that the answer to the new question, in fact, \textit{is} yes.
In the next two sections, assuming \( N \geq 10 \),
we will show how to compute the only \((p, q, r)\) that works.
\textbf{Spoiler:} It’s a Boltzmann distribution, \textit{exactly}.

6. \textbf{Computing \( p, q, r \) \textit{à la Boltzmann}}

As in the preceding section, \textbf{let} \( p, q, r > 0 \), \( S := \sum_{\omega \in \Omega} p^i q^j r^k \omega \).
We assume: \( p + q + r = 1 \) and \( q + 10r = 1 \).
We also assume: \( \forall \omega \in \Omega, \quad p^i q^j r^k \omega / S = 1 / (\#\Omega) \).
\textbf{We will prove} that, if \( N \geq 10 \), then
there is at most one \((p, q, r)\) that satisfies these conditions,
specifically, \((p, q, r) = (1, 9^{-1/10}, 9^{-1}) / (1 + 9^{-1/10} + 9^{-1})\).

\textbf{Define} the dot product, \( \odot \), on \( \mathbb{R}^3 \), by:
\[
\forall x, y, z, X, Y, Z \in \mathbb{R}, \quad (x, y, z) \odot (X, Y, Z) = xX + yY + zZ.
\]
For all \( u \in \mathbb{R}^3 \), \textbf{let} \( u^\perp := \{ v \in \mathbb{R}^3 \mid u \odot v = 0 \} \).
For all \( U \subseteq \mathbb{R}^3 \), \textbf{let} \( U^\perp := \{ v \in \mathbb{R}^3 \mid \forall u \in U, \ u \odot v = 0 \} \).
For all \( u, v \in \mathbb{R}^3 \), \textbf{let} \( \langle u, v \rangle_{\text{span}} \) denote the \( \mathbb{R} \)-span of \( \{u, v\} \), \textit{i.e.},
\[
\langle u, v \rangle_{\text{span}} := \{ su + tv \mid s, t \in \mathbb{R} \}.
\]
Recall (§3): \( \Omega = \{ \omega : [1..N] \rightarrow \{0, 1, 10\} \mid \sum_{\ell=1}^{N} \omega(\ell) = N \} \).
Recall (§5): \( A = \{(i, j, k, \omega) \mid \omega \in \Omega \} \).
Recall (§5): \( A \) is the set of all \((i, j, k)\) \text{ s.t. } i, j, k \in [0..N] \quad \text{and} \quad i + j + k = N \quad \text{and} \quad j + 10k = N \).
Then: \( A \) is the set of all \((i, j, k)\) \text{ s.t. } i, j, k \in [0..N] \quad \text{and} \quad (i, j, k) \odot (1, 1, 1) = N \quad \text{and} \quad (i, j, k) \odot (0, 1, 10) = N \).
For all \( a, b \in A \), we have
\[
a \odot (1, 1, 1) = N = b \odot (1, 1, 1) \quad \text{and} \quad a \odot (0, 1, 10) = N = b \odot (0, 1, 10),
\]
so we get
\[
(a - b) \odot (1, 1, 1) = 0 \quad \text{and} \quad (a - b) \odot (0, 1, 10) = 0,
\]
so \( a - b \in (1, 1, 1)^\perp \cap (0, 1, 10)^\perp \).
\textbf{Let} \( V := (1, 1, 1)^\perp \cap (0, 1, 10)^\perp \).
Then: \( \forall a, b \in A, \ a - b \in V. \)

Let \( D := \{ a - b \mid a, b \in A \} \). Then \( D \subseteq V. \)

Also, we have: \( V \subseteq (1,1,1)^\perp \) and \( V \subseteq (0,1,10)^\perp. \)

Then: \( V^\perp \supseteq (1,1,1)^\perp \) and \( V^\perp \supseteq (0,1,10)^\perp. \)

Since \((1,1,1) \in (1,1,1)^\perp \subseteq V^\perp \) and \((0,1,10) \in (0,1,10)^\perp \subseteq V^\perp, \)

we get: \( \langle (1,1,1), (0,1,10) \rangle_{\text{span}} \subseteq V^\perp. \)

Let \( W := \langle (1,1,1), (0,1,10) \rangle_{\text{span}}. \) Then: \( W \subseteq V^\perp. \)

Assume \( N \geq 10. \) Let \( a_1 := (0,N,0), \ a_2 := (9,N-10,1). \)

Then \( a_1, a_2 \in A. \) Let \( d_1 := a_2 - a_1. \) Then \( d_1 \in D. \)

Since \( W = \langle (1,1,1), (0,1,10) \rangle_{\text{span}}, \) we get: \( \dim W = 2. \)

Since \( d_1 \neq (0,0,0), \) we get: \( \dim d_1 = 2. \)

Since \( d_1 \in D \subseteq V \) and \( W \subseteq V^\perp, \) we get: \( d_1^\perp \supseteq D^\perp \supseteq V^\perp \supseteq W. \)

So, since \( \dim d_1 = 2 = \dim W, \) we get: \( d_1^\perp = D^\perp = V^\perp = W. \)

Then \( D = W. \) Recall: \( \forall \omega \in \Omega, \ p^i q^j r^k \omega/S = 1/(\#\Omega). \)

So, since \( A = \{(i_\omega, j_\omega, k_\omega) \mid \omega \in \Omega \}, \) we get:

\[ \forall (i, j, k) \in A, \ p^i q^j r^k S = 1/(\#\Omega). \]

Equivalently, \( \forall (i, j, k) \in A, \)

\[ i \cdot (\ln p) + j \cdot (\ln q) + k \cdot (\ln r) - (\ln S) = -(\ln(\#\Omega)). \]

Equivalently, \( \forall (i, j, k) \in A, \)

\[ (i,j,k) \odot (\ln p, \ln q, \ln r) = (\ln S) - (\ln(\#\Omega)). \]

Then: \( \forall a, b \in A, \)

\[ a \odot (\ln p, \ln q, \ln r) = (\ln S) - (\ln(\#\Omega)) = b \odot (\ln p, \ln q, \ln r), \]

so we get: \( (a - b) \odot (\ln p, \ln q, \ln r) = 0. \)

Then: \( \forall d \in D, \)

\[ d \odot (\ln p, \ln q, \ln r) = 0. \]

Then \( (\ln p, \ln q, \ln r) \in D^\perp. \)

Since \( (\ln p, \ln q, \ln r) \in D^\perp = W = \langle (1,1,1), (0,1,10) \rangle_{\text{span}}, \)

choose \( C > 0 \) and \( \beta \in \mathbb{R} \) s.t.

\[ (\ln p, \ln q, \ln r) = (\ln C) \cdot (1,1,1) - \beta \cdot (0,1,10). \]

Then \( (\ln p, \ln q, \ln r) = (\ln C, (\ln C) - \beta, (\ln C) - 10\beta). \)

Then \( (p,q,r) = (C, Ce^{-\beta}, Ce^{-10\beta}). \)

Then \( (p,q,r) = C \cdot (1,e^{-\beta},e^{-10\beta}). \)

So, since \( p + q + r = 1, \) we get: \( C \cdot (1 + e^{-\beta} + e^{-10\beta}) = 1. \)

Then \( C = \frac{1}{1 + e^{-\beta} + e^{-10\beta}}. \)

Then \( (p,q,r) = \frac{e^{-\beta} + 10e^{-10\beta}}{1 + e^{-\beta} + e^{-10\beta}}. \)

So, since \( q + 10r = 1, \) we get: \( \frac{1}{1 + e^{-\beta} + e^{-10\beta}} = 1. \)

Then \( e^{-\beta} + 10e^{-10\beta} = 1 + e^{-\beta} + e^{-10\beta}. \)

Then \( 0 = 1 - 9e^{-10\beta}. \)

Then \( e^{-10\beta} = 9^{-1}. \) Then \( e^{-\beta} = 9^{-1/10}. \)

Then \( (p,q,r) = \frac{(1,9^{-1/10},9^{-1})}{1 + 9^{-1/10} + 9^{-1}}. \)
So this is the only \((p,q,r)\) that can possibly work.
In the next section, we show that it does work.

7. **Showing the Boltzmann \(p,q,r\) work**

In this section, we prove the converse of the result from the preceding section.

That is, we let \((p,q,r) := \frac{(1,9^{-1/10},9^{-1})}{1 + 9^{-1/10} + 9^{-1}}\) and \(S := \sum_{\omega \in \Omega} p^i \omega q^j \omega r^k \omega\),
and we wish to show: \(p + q + r = 1\) and \(q + 10r = 1\) and 
\(\forall \omega \in \Omega\), \(p^i \omega q^j \omega r^k \omega / S = 1 / (\#\Omega)\).

Let \(\beta := (\ln 9)/10\). Then \(e^{-\beta} = 9^{-1/10}\). Then \(e^{-10\beta} = 9^{-1}\).

Then \((p,q,r) = \frac{(1,e^{-\beta},e^{-10\beta})}{1 + e^{-\beta} + e^{-10\beta}}\). Let \(C := \frac{1}{1 + e^{-\beta} + e^{-10\beta}}\).

Then \((p,q,r) = C \cdot (1,e^{-\beta},e^{-10\beta})\). Then 
\((p,q,r) = (C,Ce^{-\beta},Ce^{-10\beta})\).

Let \(K := C^N \cdot e^{-\beta \cdot N}\).

Recall (§3): \(\Omega := \{ \omega : [1..N] \rightarrow \{0,1,10\} \mid \sum_{\ell=1}^{N} [\omega(\ell)] = N \} \).

Claim: \(\forall \omega \in \Omega\), \(p^i \omega q^j \omega r^k \omega = K\).

Proof of Claim: Given \(\omega \in \Omega\), want: \(p^i \omega q^j \omega r^k \omega = K\).

Recall (§5): \(i_\omega + j_\omega + k_\omega = N\) and \(j_\omega + 10k_\omega = \sum_{\ell=1}^{N} [\omega(\ell)]\).

By definition of \(\Omega\), since \(\omega \in \Omega\), we get: \(\sum_{\ell=1}^{N} [\omega(\ell)] = N\).

Then: \(j_\omega + 10k_\omega = N\). Recall: \((p,q,r) = C \cdot (1,e^{-\beta},e^{-10\beta})\).

Then: \(p^i \omega q^j \omega r^k \omega = C_i \omega \cdot (Ce^{-\beta})^j \omega \cdot (Ce^{-10\beta})^k \omega\)
\(= C_i \omega + j_\omega + k_\omega \cdot e^{-\beta \cdot (j_\omega + 10k_\omega)} = C^N \cdot e^{-\beta \cdot N} = K\).

End of proof of Claim.

By definition of \(S\), we have: 
\(S = \sum_{\omega \in \Omega} p^i \omega q^j \omega r^k \omega\).

So, by the Claim, we get: 
\(S = (\#\Omega) \cdot K\). Then \(K/S = 1/(\#\Omega)\).

We have \(10/9 = 1 + (1/9)\). That is, \(10 \cdot 9^{-1} = 1 + 9^{-1}\).

So, since \(e^{-10\beta} = 9^{-1}\), we get: \(10e^{-10\beta} = 1 + e^{-10\beta}\).

Then: \(e^{-\beta} + 10e^{-10\beta} = 1 + e^{-\beta} + e^{-10\beta}\).

By definition of \(C\), we get: \(C \cdot (1 + e^{-\beta} + e^{-10\beta}) = 1\).

Recall: \((p,q,r) = (C,Ce^{-\beta},Ce^{-10\beta})\).

Since \(p + q + r = C \cdot (1 + e^{-\beta} + e^{-10\beta}) = 1\)
and since \(q + 10r = C \cdot (e^{-\beta} + 10e^{-10\beta}) = C \cdot (1 + e^{-\beta} + e^{-10\beta}) = 1\),
it remains only to show: \(\forall \omega \in \Omega\), 
\(p^i \omega q^j \omega r^k \omega / S = 1 / (\#\Omega)\).
Given $\omega \in \Omega$, want: $$p^\omega q^\omega r^k / S = 1 / (\# \Omega).$$
By the Claim, we get: $$p^\omega q^\omega r^k = K.$$ 
Recall: $$K / S = 1 / (\# \Omega).$$
Then: $$p^\omega q^\omega r^k / S = K / S = 1 / (\# \Omega).$$

8. COUNTABLE MEASURE THEORY

By convention, in this note, any countable set is given its discrete Borel structure.

A measure $\mu$ on a countable set $\Theta$

is completely determined by

the function $t \mapsto \mu(t) : \Theta \rightarrow [0; \infty]$, 

because: $\forall \Theta_0 \subseteq \Theta$, we have $\mu(\Theta_0) = \sum_{t \in \Theta_0} [\mu(t)]$.

The only measure on $\emptyset$ is the zero measure.

DEFINITION 8.1. Let $\Theta$ be a countable set.

Then \[ M_\Theta \] denotes the set of measures on $\Theta$,

and \[ F_M_\Theta \] := \{ $\mu \in M_\Theta$ | $\mu(\Theta) < \infty$ \},

and \[ M_\Theta^\ast \] := \{ $\mu \in M_\Theta$ | $0 < \mu(\Theta) < \infty$ \},

and \[ P_\Theta \] := \{ $\mu \in M_\Theta$ | $\mu(\Theta) = 1$ \}.

Then $M_\Theta$ is the set of measures on $\Theta$ and $F_M_\Theta$ is the set of nonzero finite measures on $\Theta$ and $P_\Theta$ is the set of probability measures on $\Theta$.

Note: $(\Theta = \emptyset) \Rightarrow (F_M_\emptyset = \emptyset = P_\emptyset)$.

DEFINITION 8.2. Let $\Theta$ be a nonempty countable set, $\mu \in F_M_\Theta$.

Let $n \in \mathbb{N}$. Then $[\mu^n]$ is defined by:

$\forall x \in \Theta^n$, $\mu^n(x) = (\mu(x_1)) \cdots (\mu(x_n))$.

The following is a basic fact, whose proof we omit:

Let $\Theta$ be a countable set, $n \in [2..\infty)$, $\mu \in F_M_\Theta$.

Let $Z \subseteq \Theta^n$, $X \subseteq \Theta^{n-1}$, $Y \subseteq \Theta$. Assume that:

under the standard bijection $\Theta^n \leftrightarrow \Theta^{n-1} \times \Theta$,

we have: $Z \leftrightarrow X \times Y$.

Then: $\mu^n(Z) = (\mu^{n-1}(X)) \cdot (\mu(Y))$.

It is common to identify $Z$ with $X \times Y$, in which case we would write:

$\mu^n(X \times Y) = ((\mu^{n-1}(X)) \cdot (\mu(Y)))$.

The countable sets that are of interest in this note all carry the discrete topology. We therefore define:
DEFINITION 8.3. Let $\Theta$ be a countable set, $\mu \in \mathcal{M}_\Theta$.
Then the \underline{support of $\mu$} is: $\text{S}_\mu := \{ t \in \Theta \mid \mu(t) \neq 0 \}$.

DEFINITION 8.4. Let $\Theta \subseteq \mathbb{R}$ be countable, $\mu \in \mathcal{M}_\Theta$.
Let $\rho \geq 1$ be real. Then $|\mu|_\rho := (\sum_{t \in \Theta} |t|^\rho \cdot (\mu(t)))^{1/\rho}$.

Note: If $\mu \in \mathcal{F} \mathcal{M}_\Theta$ and if $\# S_{\mu} < \infty$, then: $\forall \rho \geq 1, \ |\mu|_\rho < \infty$.

For any countable $\Theta \subseteq \mathbb{R}$, for any $\mu \in \mathcal{P}_\Theta$, for any real $\rho \geq 1$,
$|\mu|_\rho$ is the $L^\rho$-norm of any $\Theta$-valued random-variable with distribution $\mu$.

DEFINITION 8.5. Let $\Theta \subseteq \mathbb{R}$ be countable.

Let $\mu \in \mathcal{P}_\Theta$. Assume: $|\mu|_1 < \infty$.
Then the \underline{mean of $\mu$} is: $M_\mu := \sum_{t \in \Theta} [ t \cdot (\mu(t)) ]$.
Also, the \underline{variance of $\mu$} is: $V_\mu := \sum_{t \in \Theta} [(t - M_\mu)^2 \cdot (\mu(t))]$.

Let $\Theta \subseteq \mathbb{R}$ be countable, $\mu \in \mathcal{P}_\Theta$. Assume: $|\mu|_1 < \infty$.
Then, by subadditivity of absolute value, we get $|M_\mu| \leq |\mu|_1$.
Also, by expanding the square in the formula for $V_\mu$,
we get $V_\mu = |\mu|_2^2 - M_\mu^2$.
then $V_\mu \leq |\mu|_2^2$.

Let $\Theta \subseteq \mathbb{R}$ be countable, $\mu \in \mathcal{P}_\Theta$.

Let $X$ be a $\Theta$-valued random-variable with distribution $\mu$.
Then, for real $\rho \geq 1$, we have: $X$ is $L^\rho$ iff $|\mu|_\rho < \infty$.
If $X$ is $L^1$, then $M_\mu = \text{E}[X]$ and $V_\mu = \text{Var}[X]$.
In particular, if $X$ is $L^1$, then $M_\mu$ is the mean (aka expected value, aka average value) of $X$.

THEOREM 8.6. Let $\Theta \subseteq \mathbb{R}$ be countable, $\mu \in \mathcal{P}_\Theta$.
Assume: $|\mu|_1 < \infty$. Then: $(\# S_{\mu} \geq 2) \Leftrightarrow (V_\mu > 0)$.

The preceding result is basic. We omit its proof.

Because $\forall t \in \mathbb{Z}, \ |t| \leq t^2$, we conclude:
for any $\mathbb{Z}$-valued random-variable $X$, $\text{E}[|X|] \leq \text{E}[X^2]$.
It follows that for any $\mathbb{Z}$-valued $L^2$ random-variable $X$, we have:
$X$ is $L^1$, and so $\text{E}[X]$ is defined and finite.

Because $\forall t \in \mathbb{Z}, \ |t| \leq t^2$, we conclude:
$\forall$ countable $\Theta \subseteq \mathbb{Z}, \forall \mu \in \mathcal{M}_\Theta, \ |\mu|_1 \leq |\mu|_2^2$. 
DEFINITION 8.7. Let $\Theta$ be a countable set.

Let $\mu_1, \mu_2, \ldots \in \mathcal{P}_\Theta$. Let $\lambda \in \mathcal{P}_\Theta$.

By $[\mu_1, \mu_2, \ldots \rightarrow \lambda]$, we mean: $\forall \Theta_0 \subseteq \Theta$, $\mu_1(\Theta_0), \mu_2(\Theta_0), \ldots \rightarrow \lambda(\Theta_0)$.

Equivalently, $\forall t \in \Theta$, $\mu_1\{t\}, \mu_2\{t\}, \ldots \rightarrow \lambda\{t\}$.

For any sets $\Sigma, T$, for any function $\varepsilon : \Sigma \to T$, for any set $A$, we define $[\varepsilon A] := \{\sigma : \Sigma | \varepsilon(\sigma) \in A\}$.

For any countable set $\Sigma$, for any set $T$, for any function $\varepsilon : \Sigma \to T$, for any $\mu \in \mathcal{M}_\Sigma$, we define $[\varepsilon \mu] \in \mathcal{M}_{\underline{\varepsilon}}$ by: $\forall A \subseteq \underline{\varepsilon}$, $([\varepsilon \mu](A)) = \mu(\varepsilon A)$.

For any nonempty countable set $\Theta$, for any $\mu \in \mathcal{F}\mathcal{M}_{\underline{\Theta}}$, let $[\mathcal{N}(\mu)] := \frac{\mu}{\mu(\Theta)} \in \mathcal{P}_\Theta$; then $\forall \Theta_0 \subseteq \Theta$, $(\mathcal{N}(\mu))(\Theta_0) = \frac{\mu(\Theta_0)}{\mu(\Theta)}$, and $\mathcal{N}(\mu)$ is called the normalization of $\mu$.

Let $\hat{\Theta}$ be a countable set. Let $\mu \in \mathcal{M}_{\hat{\Theta}}$. Let $\Theta \subseteq \hat{\Theta}$. Then the restriction of $\mu$ to $\Theta$, denoted $[\mu|\Theta] \in \mathcal{M}_\Theta$, is defined by: $\forall \Theta_0 \subseteq \Theta$, $([\mu|\Theta](\Theta_0)) = \mu(\Theta_0)$.

NOTE: We have $(\mu|\Theta)(\Theta) = \mu(\Theta)$. So, if $0 < \mu(\Theta) < \infty$, then:

$$\mu|\Theta \in \mathcal{F}\mathcal{M}_{\hat{\Theta}}$$

and $\mathcal{N}(\mu|\Theta) = \frac{\mu|\Theta}{\mu(\Theta)}$.

and $\forall \Theta_0 \subseteq \Theta$, $(\mathcal{N}(\mu|\Theta))(\Theta_0) = \frac{\mu(\Theta_0)}{\mu(\Theta)}$.

DEFINITION 8.8. Let $F$ be a nonempty finite set.

Then we define $[\nu_F] \in \mathcal{P}_F$ by: $\forall f \in F$, $\nu_F\{f\} = 1/(\#F)$.

Also, we define $[\nu_{\emptyset}] : \{\emptyset\} \to \{-1\}$ by: $\nu_{\emptyset}(\emptyset) = -1$.

THEOREM 8.9. Let $F$ be a nonempty finite set, $\theta \in \mathcal{P}_F$.

Assume: $\forall f, g \in F$, $\theta\{f\} = \theta\{g\}$.

Then: $\theta = \nu_F$.

Proof. Since $F$ is nonempty, choose $g_0 \in F$. Let $b := \theta\{g_0\}$.

Then: $\forall f \in F$, $\theta\{f\} = b$. Then $\sum_{f \in F} \theta\{f\} = (\#F) \cdot b$.

Since $\theta \in \mathcal{P}_F$, we get: $\theta(F) = 1$.

Since $(\#F) \cdot b = \sum_{f \in F} \theta\{f\} = \theta(F) = 1$, we get: $b = 1/(\#F)$.

Since $\forall f \in F$, $\theta\{f\} = b = 1/(\#F) = \nu_F\{f\}$, we get: $\theta = \nu_F$. □
9. The Discrete Local Limit Theorem

**DEFINITION 9.1.** Let $E \subseteq \mathbb{Z}$.

By $E$ is \underline{residue-constrained}, we mean:

\[ \exists m \in [2, \infty), \exists n \in \mathbb{Z} \quad \text{s.t.} \quad E \subseteq m\mathbb{Z} + n. \]

By $E$ is \underline{residue-unconstrained}, we mean:

$E$ is not residue-constrained.

Since $\emptyset \subseteq 2 \cdot \mathbb{Z} + 1$, we get: $\emptyset$ is residue-constrained.

For all $b \in \mathbb{Z}$, since $\{b\} \subseteq 2 \cdot \mathbb{Z} + b$, we get: $\{b\}$ is residue-constrained.

Then: $\forall$ residue-unconstrained $E \subseteq \mathbb{Z}$, $\#E \geq 2$.

We have: $\{0, 2, 10\} \subseteq 2\mathbb{Z} + 0$ and $\{1, 3, 11\} \subseteq 2\mathbb{Z} + 1$,

so $\{0, 2, 10\}$ and $\{1, 3, 11\}$ are both residue-constrained.

Here is a test for residue-unconstrainedness:

Let $E \subseteq \mathbb{Z}$. Assume $\#E \geq 2$. Let $\varepsilon_0 \in E$.

Then: ($E$ is residue-unconstrained) iff ($\gcd(E - \varepsilon_0) = 1$).

By this test, we see that:

$\{0, 1, 10\}$ and $\{2, 4, 8, 9\}$ and $\{3, 9, 13, 18\}$ are all residue-unconstrained.

**DEFINITION 9.2.** For all $\alpha \in \mathbb{R}$, for all $v \geq 0$,

define $\Phi^v_\alpha : \mathbb{R} \to (0; \infty)$ by: $\forall t \in \mathbb{R}$, $\Phi^v_\alpha(t) = \frac{\exp\left(-\frac{(t - \alpha)^2}{2v}\right)}{\sqrt{2\pi v}}$.

Note: $\Phi^v_\alpha$ is a PDF of a normal variable with mean $\alpha$ and variance $v$.

The next result is the Discrete Local Limit Theorem,

this one is stated in probability-theoretic terms:

**THEOREM 9.3.** Let $E \subseteq \mathbb{Z}$ be residue-unconstrained.

Let $X_1, X_2, \ldots$ be a $\mathbb{Z}$-valued iid sequence of $L^2$ random-variables.

Assume: $\forall n \in \mathbb{N}$, $\{t \in \mathbb{Z} | \Pr[X_n = t] > 0\} = E$.

Let $\alpha \in \mathbb{R}$, $v \geq 0$. Assume: $\forall n \in \mathbb{N}$, $E[X_n] = \alpha$ and $\text{Var}[X_n] = v$.

Then: $0 < v < \infty$, and, $\forall t_1, t_2, \ldots \in \mathbb{Z}$, as $n \to \infty$,

\[ \sqrt{n} \cdot \left[ \Pr[X_1 + \cdots + X_n = t_n] \right] - \left( \Phi^v_{\alpha_n}(t_n) \right) \to 0. \]

For a good exposition of this theorem and its proof,

search on “Terence Tao Local Limit Theorem”.

Visit the website, and then expand “read the rest of this entry”,

and then scroll down to “– 2. Local limit theorems –”.

Next is another version of the Discrete Local Limit Theorem;

this one is stated in measure-theoretic terms:
THEOREM 9.4. Let $E \subseteq \mathbb{Z}$ be residue-unconstrained.

Let $\mu \in \mathcal{P}_E$. Assume: $S_\mu = E$. Assume: $|\mu|_2 < \infty$.

Let $\alpha := M_\mu$, $v := V_\mu$. Then: $0 < v < \infty$, and, $\forall t_1, t_2, \ldots \in \mathbb{Z}$, as $n \to \infty$, $\sqrt{n} \cdot [(\mu^n \{ f \in E^n \mid f_1 + \cdots + f_n = t_n \}) - (\Phi_{n\alpha}^v(t_n))] \to 0$.

In Theorem 9.4, since $E \subseteq \mathbb{Z}$ and since $|\mu|_2 < \infty$,
we get: $|\mu|_1 < \infty$, and so $M_\mu$ and $V_\mu$ are both defined.
Moreover, $|M_\mu| \leq |\mu|_1 < \infty$, so $M_\mu$ is finite.
In Theorem 9.4, the proof that $v > 0$ is relatively simple:
Since $E$ is residue-unconstrained, we get: $\#E \geq 2$.
Since $\#S_\mu = \#E \geq 2$, by Theorem 8.6, we get: $v > 0$.
In Theorem 9.4, the proof that $v < \infty$ is relatively simple:
$v = V_\mu = |\mu|_2^2 - M_\mu^2 \leq |\mu|_2^2 < \infty$.

Here is an application of Theorem 9.3:

THEOREM 9.5. Let $E \subseteq \mathbb{Z}$ be residue-unconstrained.

Let $X_1, X_2, \ldots$ be a $\mathbb{Z}$-valued iid sequence of $L^2$ random-variables.
Assume: $\forall n \in \mathbb{N}$, $\{ t \in \mathbb{Z} \mid \Pr[X_n = t] > 0 \} = E$.

Let $\alpha \in \mathbb{R}$, $v \geq 0$. Assume: $\forall n \in \mathbb{N}$, $\Pr[X_n] = \alpha$ and $\Var[X_n] = v$.

Then: $0 < v < \infty$. Also, $\forall t_1, t_2, \ldots \in \mathbb{Z}$,
if $\{ t_n - n\alpha \mid n \in \mathbb{N} \}$ is bounded,
then, as $n \to \infty$, $\sqrt{n} \cdot (\Pr[X_1 + \cdots + X_n = t_n]) \to 1/\sqrt{2\pi v}$.

Equivalently, the conclusion says: if $\{ t_n - n\alpha \mid n \in \mathbb{N} \}$ is bounded, then
as $n \to \infty$, $\Pr[X_1 + \cdots + X_n = t_n]$ is asymptotic to $1/\sqrt{2\pi nv}$.

Proof. By Theorem 9.3, we get $0 < v < \infty$ and
we want: as $n \to \infty$, $\sqrt{n} \cdot (\Phi_{n\alpha}^v(t_n)) \to 1/\sqrt{2\pi v}$.
We have: $\forall n \in \mathbb{N}$, $\sqrt{n} \cdot (\Phi_{n\alpha}^v(t_n)) = \frac{\exp(- (t_n - n\alpha)^2 / (2nv))}{\sqrt{2\pi v}}$.

Since $\{ t_n - n\alpha \mid n \in \mathbb{N} \}$ is bounded and since $0 < v < \infty$, we get:
as $n \to \infty$, $-(t_n - n\alpha)^2 / (2nv) \to 0$.
Then: as $n \to \infty$, $\exp(-(t_n - n\alpha)^2 / (2nv)) \to 1$.
Then: as $n \to \infty$, $\sqrt{n} \cdot (\Phi_{n\alpha}^v(t_n)) \to 1/\sqrt{2\pi v}$.

We record a measure-theoretic version of Theorem 9.5:

THEOREM 9.6. Let $E \subseteq \mathbb{Z}$ be residue-unconstrained.

Let $\mu \in \mathcal{P}_E$. Assume: $S_\mu = E$.
Assume: $|\mu|_2 < \infty$. Let $\alpha := M_\mu$, $v := V_\mu$. 
Then: $0 < v < \infty$. Also, $\forall t_1, t_2, \ldots \in \mathbb{Z}$, if $\{t_n - n\alpha | n \in \mathbb{N}\}$ is bounded, then, as $n \to \infty$, $\sqrt{n} \cdot (\mu^n \{ f \in E^n | f_1 + \cdots + f_n = t_n \}) \to 1/\sqrt{2\pi v}$.

We also record the $t_n = t_0 + n\alpha$ special case of the past two theorems:

**THEOREM 9.7.** Let $E \subseteq \mathbb{Z}$ be residue-unconstrained. Let $X_1, X_2, \ldots$ be a $\mathbb{Z}$-valued iid sequence of $L^2$ random-variables. Assume: $\forall n \in \mathbb{N}, \{ t \in \mathbb{Z} | \Pr[X_n = t] > 0 \} = E$. Let $t_0, \alpha \in \mathbb{Z}, v \geq 0$. Assume: $\forall n \in \mathbb{N}, E[X_n] = \alpha$ and $\text{Var}[X_n] = v$. Then: $0 < v < \infty$, and, as $n \to \infty$, $\sqrt{n} \cdot (\Pr[X_1 + \cdots + X_n = t_0 + n\alpha]) \to 1/\sqrt{2\pi v}$.

**THEOREM 9.8.** Let $E \subseteq \mathbb{Z}$ be residue-unconstrained. Let $\mu \in \mathcal{P}_E$. Assume: $S_{\mu} = E$. Assume: $|\mu|_2 < \infty$. Let $\alpha := M_{\mu}, v := V_{\mu}$. Assume: $\alpha \in \mathbb{Z}$. Let $t_0 \in \mathbb{Z}$. Then: $0 < v < \infty$, and, as $n \to \infty$, $\sqrt{n} \cdot (\mu^n \{ f \in E^n | f_1 + \cdots + f_n = t_0 + n\alpha \}) \to 1/\sqrt{2\pi v}$.

We also record the $t_0 = 0$ special case of the past two theorems:

**THEOREM 9.9.** Let $E \subseteq \mathbb{Z}$ be residue-unconstrained. Let $X_1, X_2, \ldots$ be a $\mathbb{Z}$-valued iid sequence of $L^2$ random-variables. Assume: $\forall n \in \mathbb{N}, \{ t \in \mathbb{Z} | \Pr[X_n = t] > 0 \} = E$. Let $\alpha \in \mathbb{Z}, v \geq 0$. Assume: $\forall n \in \mathbb{N}, E[X_n] = \alpha$ and $\text{Var}[X_n] = v$. Then $0 < v < \infty$, and, as $n \to \infty$, $\sqrt{n} \cdot (\Pr[X_1 + \cdots + X_n = n\alpha]) \to 1/\sqrt{2\pi v}$.

**THEOREM 9.10.** Let $E \subseteq \mathbb{Z}$ be residue-unconstrained. Let $\mu \in \mathcal{P}_E$. Assume: $S_{\mu} = E$. Assume: $|\mu|_2 < \infty$. Let $\alpha := M_{\mu}, v := V_{\mu}$. Assume: $\alpha \in \mathbb{Z}$. Then $0 < v < \infty$, and, as $n \to \infty$, $\sqrt{n} \cdot (\mu^n \{ f \in E^n | f_1 + \cdots + f_n = n\alpha \}) \to 1/\sqrt{2\pi v}$.

10. **Average events have low information, particular case**

Suppose, in secret, I flip a coin 1000 times, then reveal to you that the total number of heads was 1000, and then ask you to guess the last flip.

The answer is that, since all the coin flips were heads, the last flip must have been a head.
Similarly, if I had told you that
the total number of heads was 0,
then you would have known that the last flip was a tail.

By contrast, if I had told you that
the total number of heads was 500,
it seems intuitively clear that
you’d have had very little information about the last flip.

We wish to generalize and formalize that intuition,
and then provide rigorous proof of the resulting formal statement.

Our main theorem is Theorem 11.5, in the next section.
In this section, we go carefully through a special case:

Let \( X_1, X_2, X_3, \ldots \) be iid random-variables such that, \( \forall n \in \mathbb{N} \),
\[
    \begin{align*}
    \Pr[X_n = -1] &= 1/2, \\
    \Pr[X_n = 0] &= 1/3, \\
    \Pr[X_n = 3] &= 1/6.
    \end{align*}
\]
Then:
\( \forall n \in \mathbb{N}, \ E[X_n] = 0 \) and \( \text{Var}[X_n] = 2. \)

Also, \( \forall n \in \mathbb{N}, \ -1 \leq X_n \leq 3 \) a.s.

For all \( n \in \mathbb{N} \), let \( T_n := X_1 + \cdots + X_n. \)
Then:
\( \forall n \in \mathbb{N}, \ -n \leq T_n \leq 3n \) a.s.

Then:
\( -1000 \leq T_{1000} \leq 3000 \) a.s.

Also, \( [T_{1000} = -1000] \Rightarrow [X_1 = \cdots = X_{1000} = -1], \)
and so \( \Pr[X_{1000} = -1 \mid T_{1000} = -1000] = 1. \)

Similarly, \( \Pr[X_{1000} = 3 \mid T_{1000} = 3000] = 1. \)

By contrast, the event \( T_{1000} = 0 \)
would seem to give very little information about \( X_{1000}. \)

It therefore seems reasonable to expect that
\[
    \begin{align*}
    \Pr[X_{1000} = -1 \mid T_{1000} = 0] &\approx 1/2, \\
    \Pr[X_{1000} = 0 \mid T_{1000} = 0] &\approx 1/3, \\
    \Pr[X_{1000} = 3 \mid T_{1000} = 0] &\approx 1/6.
    \end{align*}
\]
To make this precise, we will work “in the thermodynamic limit”,
which means: we replace 1000 by a variable \( n \in \mathbb{N}, \) and let \( n \to \infty. \)

That is, more precisely, we expect that, as \( n \to \infty, \)
\[
    \begin{align*}
    \Pr[X_n = -1 \mid T_n = 0] &\to 1/2, \\
    \Pr[X_n = 0 \mid T_n = 0] &\to 1/3, \\
    \Pr[X_n = 3 \mid T_n = 0] &\to 1/6.
    \end{align*}
\]
We will focus on proving the third of these limits;
proofs of the other two are similar.
By definition of conditional probability,

we wish to prove: \( \text{As } n \to \infty, \quad \frac{\Pr[(X_n = 3) \& (T_n = 0)]}{\Pr[T_n = 0]} \to 1/6. \)

Claim: Let \( n \in [2, \infty) \).

Then: \( \Pr[(X_n = 3) \& (T_n = 0)] = (1/6) \cdot (\Pr[T_{n-1} = -3]). \)

Proof of Claim: We have: \( T_n = X_1 + \cdots + X_{n-1} + X_n. \)

Since \( \Pr[(X_n = 3) \& (T_n = 0)] \)

\[ = \Pr[(X_n = 3) \& (X_1 + \cdots + X_{n-1} = X_n = 0)] \]

\[ = \Pr[(X_n = 3) \& (X_1 + \cdots + X_{n-1} = 3 = 0)] \]

\[ = \Pr[(X_n = 3) \& (X_1 + \cdots + X_{n-1} = -3)], \]

it follows, from independence of \( X_1, \ldots, X_n \), that

\[ \Pr[(X_n = 3) \& (T_n = 0)] \]

\[ = ( \Pr[X_n = 3] ) \cdot ( \Pr[X_1 + \cdots + X_{n-1} = -3] ). \]

So, since \( \Pr[X_n = 3] = 1/6 \) and \( X_1 + \cdots + X_{n-1} = T_{n-1} \),

we get: \( \Pr[(X_n = 3) \& (T_n = 0)] = (1/6) \cdot (\Pr[T_{n-1} = -3]). \)

End of proof of Claim.

By the claim, we wish to prove:

As \( n \to \infty, \quad \frac{(1/6) \cdot (\Pr[T_{n-1} = -3])}{\Pr[T_n = 0]} \to 1/6. \)

We wish to prove: \( \text{As } n \to \infty, \quad \frac{\Pr[T_{n-1} = -3]}{\Pr[T_n = 0]} \to 1. \)

That is, we wish to prove:

As \( n \to \infty, \quad \Pr[T_{n-1} = -3] \) is asymptotic to \( \Pr[T_n = 0]. \)

So the question becomes:

How do we get a handle on the asymptotics, as \( n \to \infty \), of

both \( \Pr[T_{n-1} = -3] \) and \( \Pr[T_n = 0] \)?

The Discrete Local Limit Theorem turns out to be just what we need.

Recall: \( \forall n \in \mathbb{N}, \ \text{E}[X_n] = 0 \) and \( T_n = X_1 + \cdots + X_n. \)

Let \( \alpha := 0 \) and \( v := 2. \) Then: both \( \forall n \in \mathbb{N}, \ n\alpha = 0 \) and \( 2\pi v = 4\pi. \)

Also, \( \forall n \in \mathbb{N}, \ \text{E}[X_n] = \alpha \) and \( \text{Var}[X_n] = v. \)

By Theorem 9.9, as \( n \to \infty, \)

\( \sqrt{n} \cdot (\Pr[X_1 + \cdots + X_n = n\alpha]) \to 1/\sqrt{2\pi v}, \)

Then: as \( n \to \infty, \quad \sqrt{n} \cdot (\Pr[T_n = 0]) \to 1/\sqrt{4\pi}, \)

so, as \( n \to \infty, \quad \Pr[T_n = 0] \) is asymptotic to \( 1/\sqrt{4\pi n}. \)

Let \( t_0 := -3. \) Then, \( \forall n \in \mathbb{N}, \ t_0 + n\alpha = -3. \)
By Theorem 9.5, as \( n \to \infty \),
\[
\sqrt{n} \cdot \left( \Pr[X_1 + \cdots + X_n = t_0 + n\alpha] \right) \to \frac{1}{\sqrt{2\pi}v}.
\]
Then: as \( n \to \infty \),
\[
\sqrt{n} \cdot \left( \Pr[T_n = -3] \right) \to \frac{1}{\sqrt{4\pi}}.
\]
Then, as \( n \to \infty \),
\[
\sqrt{n - 1} \cdot \left( \Pr[T_{n-1} = -3] \right) \to \frac{1}{\sqrt{4\pi}}.
\]
Then, as \( n \to \infty \), \( \Pr[T_{n-1} = -3] \) is asymptotic to \( \frac{1}{\sqrt{4\pi(n-1)}}, \)
which is asymptotic to \( \frac{1}{\sqrt{4\pi n}} \),
which is asymptotic to \( \Pr[T_n = 0] \).

11. AVERAGE EVENTS HAVE LOW INFORMATION, GENERAL RESULT

We generalize the result of the last section:

**THEOREM 11.1.** Let \( E \subseteq \mathbb{Z} \) be residue-unconstrained.
Let \( X_1, X_2, \ldots \) be a \( \mathbb{Z} \)-valued iid sequence of \( L^2 \) random-variables.
Assume: \( \forall n \in \mathbb{N}, \{ t \in \mathbb{Z} \mid \Pr[X_n = t] > 0 \} = E. \) Let \( \alpha, P \in \mathbb{R} \).
Assume: \( \forall n \in \mathbb{N}, \ E[X_n] = \alpha \) and \( \Pr[X_n = \varepsilon_0] = P. \) Let \( \varepsilon_0 \in E. \)
Let \( t_1, t_2, \ldots \in \mathbb{Z}. \) Assume: \( \{ t_n - n\alpha \mid n \in \mathbb{N} \} \) is bounded.
Then: as \( n \to \infty \), \( \Pr[X_n = \varepsilon_0 \mid X_1 + \cdots + X_n = t_n] \to P. \)

Part of the content of Theorem 11.1 is:
\[
\forall \text{sufficiently large } n \in \mathbb{N}, \ \Pr[X_1 + \cdots + X_n = t_n] > 0.
\]

In Theorem 11.1, I don’t know whether “\( L^2 \)” can be replaced by “\( L^1 \)”.

**Proof.** Since \( X_1, X_2, \ldots \) is identically distributed,
choose \( v > 0 \) s.t., \( \forall n \in \mathbb{N}, \ \text{Var}[X_n] = v. \)
For all \( n \in \mathbb{N} \), let \( T_n := X_1 + \cdots + X_n. \)
Want: as \( n \to \infty \), \( \Pr[X_n = \varepsilon_0 \mid T_n = t_n] \to P. \)

By Theorem 9.5, we have: \( 0 < v < \infty \) and
as \( n \to \infty \),
\[
\sqrt{n} \cdot \left( \Pr[T_n = t_n] \right) \to \frac{1}{\sqrt{2\pi}v}.
\]
Let \( D_1 := \{ t_n - n\alpha \mid n \in \mathbb{N} \}. \) By hypothesis, \( D_1 \) is bounded.
Let \( D_2 := \{ t_n - n\alpha \mid n \in [2..\infty) \} \). Then \( D_2 \subseteq D_1 \).
Let \( D_3 := \{ t_{n+1} - (n+1) \cdot \alpha \mid n \in \mathbb{N} \}. \) Then \( D_3 = D_2 \).

For all \( n \in \mathbb{N}, \)
\[
\text{let } \tilde{t}_n := t_{n+1} - \varepsilon_0; \] then \( \tilde{t}_n - n\alpha + \varepsilon_0 - \alpha = t_{n+1} - (n+1) \cdot \alpha. \)
Let \( D_4 := \{ \tilde{t}_n - n\alpha \mid n \in \mathbb{N} \}. \) Then \( D_4 + \varepsilon_0 - \alpha = D_3. \)
Since \( D_4 + \varepsilon_0 - \alpha = D_3 = D_2 \subseteq D_1, \) we get: \( D_4 \subseteq D_1 - \varepsilon_0 + \alpha. \)
So, since \( D_1 \) is bounded, we get: \( D_4 \) is bounded.
Then, by Theorem 9.5, we have:
\[
\text{as } n \to \infty, \sqrt{n} \cdot \left( \Pr[T_n = \tilde{t}_n] \right) \to \frac{1}{\sqrt{2\pi}v}.
\]
Then, as \( n \to \infty \), \( \sqrt{n - 1} \cdot (\Pr[T_{n-1} = t_{n-1}]) \to 1/\sqrt{2\pi v} \).

So, as \( n \to \infty \), \( \sqrt{n - 1} \cdot (\Pr[T_{n-1} = t_n - \varepsilon_0]) \to 1/\sqrt{2\pi v} \).

Then, as \( n \to \infty \), \( \sqrt{n} \cdot (\Pr[T_n = t_n]) \to 1/\sqrt{2\pi v} \).

Then, as \( n \to \infty \), \( \sqrt{n} \cdot (\Pr[T_{n-1} = t_n - \varepsilon_0]) \to 1 \).

Also, as \( n \to \infty \), \( \sqrt{n} \to 1 \).

Multiplying the last two limits together, we get:

as \( n \to \infty \), \( \frac{\Pr[T_{n-1} = t_n - \varepsilon_0]}{\Pr[T_n = t_n]} \to 1 \).

So, since, \( \forall n \in [2, \infty) \),

\[
\Pr[X_n = \varepsilon_0 | T_n = t_n] = \frac{\Pr[X_n = \varepsilon_0 \& (T_n = t_n)]}{\Pr[T_n = t_n]}
= \frac{\Pr[X_n = \varepsilon_0 \& (T_{n-1} + X_n = t_n)]}{\Pr[T_n = t_n]}
= \frac{\Pr[X_n = \varepsilon_0 \& (T_{n-1} + \varepsilon_0 = t_n)]}{\Pr[T_n = t_n]}
= \frac{\Pr[X_n = \varepsilon_0 \& (T_{n-1} = t_n - \varepsilon_0)]}{\Pr[T_n = t_n]}
= \frac{(\Pr[X_n = \varepsilon_0]) \cdot (\Pr[T_{n-1} = t_n - \varepsilon_0])}{\Pr[T_n = t_n]},
\]

we get: as \( n \to \infty \), \( \Pr[X_n = \varepsilon_0 | T_n = t_n] \to P \). \( \square \)

Recall (§8): \( \forall \) nonempty countable set \( \Theta \),

\( \mathcal{M}_\Theta \) is the set of measures on \( \Theta \)

and \( \mathcal{F}\mathcal{M}_\Theta^\times \) is the set of nonzero finite measures on \( \Theta \)

and \( \mathcal{P}_\Theta \) is the set of probability measures on \( \Theta \).

Recall (§8): \( \forall \) nonempty countable set \( \Theta \), \( \forall \mu \in \mathcal{F}\mathcal{M}_\Theta^\times \),

\( \mathcal{N}(\mu) \) is the normalization of \( \mu \).

Here is a measure-theoretic version of the preceding theorem:

**THEOREM 11.2.** Let \( E \subseteq \mathbb{Z} \) be residue-unconstrained.

Let \( \mu \in \mathcal{P}_E \). Assume: \( S_\mu = E \). Assume: \( |\mu|_2 < \infty \).

Let \( \alpha := M_\mu \). Assume: \( \alpha \in \mathbb{Z} \). Let \( \varepsilon_0 \in E \), \( P := \mu(\varepsilon_0) \).

Let \( t_1, t_2, \ldots \in \mathbb{Z} \). Assume: \( \{t_n - n\alpha | n \in \mathbb{N} \} \) is bounded.
For all \( n \in \mathbb{N} \), let \( \Omega_n := \{ f \in E^n \mid f_1 + \cdots + f_n = t_n \} \). Then: as \( n \to \infty \), \( (\mathcal{N}(\mu^n|\Omega_n)) \{ f \in \Omega_n \mid f_n = \varepsilon_0 \} \to P \).

Part of the content of Theorem 11.2 is:
- \( \forall \mu \), sufficiently large \( n \in \mathbb{N} \), \( \mu^n(\Omega_n) > 0 \), since, otherwise, \( \mu^n|\Omega_n \) would be the zero measure on \( \Omega_n \), and so \( \mathcal{N}(\mu^n|\Omega_n) \) would not be defined.

We record the \( t_n = t_0 + n\alpha \) special case of the preceding two theorems:

**THEOREM 11.3.** Let \( E \subseteq \mathbb{Z} \) be residue-unconstrained. Let \( t_0 \in \mathbb{Z} \).
Let \( X_1, X_2, \ldots \) be a \( \mathbb{Z} \)-valued iid sequence of \( L^2 \) random variables.
Assume: \( \forall n \in \mathbb{N}, \{ t \in \mathbb{Z} \mid \Pr[X_n = t] > 0 \} = E \).
Let \( \alpha \in \mathbb{Z}, P \in \mathbb{R} \).
Let \( \varepsilon_0 \in E \). Assume: \( \forall n \in \mathbb{N}, E[X_n] = \alpha \) and \( \Pr[X_n = \varepsilon_0] = P \).
Then: as \( n \to \infty \), \( \Pr \left[ X_n = \varepsilon_0 \mid X_1 + \cdots + X_n = t_0 + n\alpha \right] \to P \).

We record the \( t_0 = 0 \) special case of the preceding two theorems:

**THEOREM 11.4.** Let \( E \subseteq \mathbb{Z} \) be residue-unconstrained.
Let \( t_0 \in \mathbb{Z} \).
Let \( \mu \in \mathcal{P}_E \). Assume: \( S_\mu = E \). Assume: \( |\mu|_2 < \infty \).
Let \( \alpha := M_\mu \). Assume: \( \alpha \in \mathbb{Z} \).
Let \( \varepsilon_0 \in E \), \( P := \mu\{\varepsilon_0\} \).
For all \( n \in \mathbb{N} \), let \( \Omega_n := \{ f \in E^n \mid f_1 + \cdots + f_n = t_0 + n\alpha \} \).
Then: as \( n \to \infty \), \( (\mathcal{N}(\mu^n|\Omega_n)) \{ f \in \Omega_n \mid f_n = \varepsilon_0 \} \to P \).

We record the \( t_0 = 0 \) special case of the preceding two theorems:

**THEOREM 11.5.** Let \( E \subseteq \mathbb{Z} \) be residue-unconstrained.
Let \( X_1, X_2, \ldots \) be a \( \mathbb{Z} \)-valued iid sequence of \( L^2 \) random variables.
Assume: \( \forall n \in \mathbb{N}, \{ t \in \mathbb{Z} \mid \Pr[X_n = t] > 0 \} = E \).
Let \( \alpha \in \mathbb{Z}, P \in \mathbb{R} \).
Let \( \varepsilon_0 \in E \). Assume: \( \forall n \in \mathbb{N}, E[X_n] = \alpha \) and \( \Pr[X_n = \varepsilon_0] = P \).
Then: as \( n \to \infty \), \( \Pr \left[ X_n = \varepsilon_0 \mid X_1 + \cdots + X_n = n\alpha \right] \to P \).

**THEOREM 11.6.** Let \( E \subseteq \mathbb{Z} \) be residue-unconstrained.
Let \( \mu \in \mathcal{P}_E \). Assume: \( S_\mu = E \). Assume: \( |\mu|_2 < \infty \).
Let \( \alpha := M_\mu \). Assume: \( \alpha \in \mathbb{Z} \).
Let \( \varepsilon_0 \in E \), \( P := \mu\{\varepsilon_0\} \).
For all \( n \in \mathbb{N} \), let \( \Omega_n := \{ f \in E^n \mid f_1 + \cdots + f_n = n\alpha \} \).
Then: as \( n \to \infty \), \( (\mathcal{N}(\mu^n|\Omega_n)) \{ f \in \Omega_n \mid f_n = \varepsilon_0 \} \to P \).

**12. Solving the main problem**

We finally have all we need to solve the main problem (end of §3).

Let \( (p, q, r) := \frac{1, 9^{-1/10}, 9^{-1}}{1 + 9^{-1/10} + 9^{-1}} \).
We compute \((p, q, r) \approx (0.5225, 0.4194, 0.0581)\), accurate to four decimal places.

**NOTE:** \(p > 0\) and \(q > 0\) and \(r > 0\).

Again, let’s say I am one of the professors applying to the GFA.

**We will show:** Under the GFA’s first system (§3),

- my probability of getting \$0 is \(p\), approximately
- my probability of getting \$1 is \(q\), approximately
- my probability of getting \$10 is \(r\), approximately.

By the work in §7, \(p + q + r = 1 = q + 10r\) and
\[
\forall \omega \in \Omega, \quad p^{i\omega}q^{j\omega}r^{k\omega} / S = 1 / (\#\Omega).
\]

**Let** \(X_1, X_2, \ldots\) be \(\mathbb{Z}\)-valued iid random-variables s.t., \(\forall n \in \mathbb{N},\)
\[
\begin{align*}
\Pr[X_n = 0] &= p, \\
\Pr[X_n = 1] &= q, \\
\Pr[X_n = 10] &= r.
\end{align*}
\]
Then \(X_1, X_2, \ldots\) is a sequence of \(L^2\) random-variables.

Also, \(\forall n \in \mathbb{N}, \quad \mathbb{E}[X_n] = q + 10r\).

By the work in §7, \(q + 10r = 1\).

Then: \(\forall n \in \mathbb{N}, \quad \mathbb{E}[X_n] = 1\).

We model the GFA’s second system (§5) by: \(\forall \ell \in [1..N],\)
Professor\#\(\ell\) receives \(X_\ell\) dollars.

For all \(n \in \mathbb{N}, \quad \text{let} \quad T_n := X_1 + \cdots + X_n.\)

We model the GFA’s third system (§5) by: \(\forall \ell \in [1..N],\)
Professor\#\(\ell\) receives \(X_\ell\) dollars, conditioned on \(T_N = N\).

**Let** \(S := \sum_{\omega \in \Omega} p^{i\omega}q^{j\omega}r^{k\omega}.\)

By the work in §7, \(\forall \omega \in \Omega, \quad p^{i\omega}q^{j\omega}r^{k\omega} / S = 1 / (\#\Omega).\)

As we observed earlier (end of §5), it follows from this that the third system is equivalent to the first.

For definiteness, let’s assume that I am Professor\#\(N\).

Then, assuming \(N\) is large, **we wish to show:**
\[
\begin{align*}
\Pr[X_N = 0 | T_N = N] &\approx p \quad \text{and} \\
\Pr[X_N = 1 | T_N = N] &\approx q \quad \text{and} \\
\Pr[X_N = 10 | T_N = N] &\approx r.
\end{align*}
\]

To be more precise, **we wish to show:** as \(n \to \infty,\)
\[
\begin{align*}
\Pr[X_n = 0 | T_n = n] &\to p \quad \text{and} \\
\Pr[X_n = 1 | T_n = n] &\to q \quad \text{and}
\end{align*}
\]
\[ \Pr[ X_n = 10 \mid T_n = n ] \to r. \]

Let \( E := \{0, 1, 10\} \). Given \( \varepsilon_0 \in E \), let \( P := \begin{cases} p, & \text{if } \varepsilon_0 = 0 \\ q, & \text{if } \varepsilon_0 = 1 \\ r, & \text{if } \varepsilon_0 = 10, \end{cases} \)

want: as \( n \to \infty \), \( \Pr[X_n = \varepsilon_0 \mid T_n = n] \to P. \)

By definition of \( X_1, X_2, \ldots \), we get: \( \forall n \in \mathbb{N}, \Pr[X_n = \varepsilon_0] = P. \)

Let \( \alpha := 1. \) Then: \( \alpha \in \mathbb{Z} \) and \( \forall n \in \mathbb{N}, E[X_n] = \alpha. \)

We have: \( E \) is residue-unconstrained.

Also, \( \forall n \in \mathbb{N}, \{ t \in \mathbb{Z} \mid \Pr[X_n = t] > 0 \} \equiv E. \)

Then: as \( n \to \infty \), \( \Pr[X_n = \varepsilon_0 \mid T_n = n] \to P. \)

13. Probability of two professors getting zero

Under the GFA’s first system, since \( N \) is large, one would expect:

- the award amounts of two different professors are almost independent.

So, for example,

- the probability that two professors both receive zero dollars should be very close to the square of the probability that one professor receives zero dollars.

We will formalize this statement and prove it, below.

For definiteness, we will assume that

- the two professors are Professor \( \#(N - 1) \) and Professor \( \#N. \)

Let \( (p, q, r) := \frac{(1, 9^{-1/10}, 9^{-1})}{1 + 9^{-1/10} + 9^{-1}}. \)

Let \( X_1, X_2, \ldots \) be \( \mathbb{Z} \)-valued iid random-variables s.t., \( \forall n \in \mathbb{N}, \)

\[ \begin{align*} \Pr[X_n = 0] &= p, \\ \Pr[X_n = 1] &= q, \\ \Pr[X_n = 10] &= r. \end{align*} \]

Then \( X_1, X_2, \ldots \) is a sequence of \( L^2 \) random-variables.

For all \( n \in \mathbb{N}, \) let \( T_n := X_1 + \cdots + X_n. \)

Assuming \( N \) is large, our goal is therefore to prove:

\[ \Pr[ X_{N-1} = 0 = X_N \mid T_N = N ] \approx p^2. \]

To be more precise, we will prove:

- as \( n \to \infty \), \( \Pr[ X_{n-1} = 0 = X_n \mid T_n = n ] \to p^2. \)
For all $n \in \mathbb{N}$, define $\psi_n : \mathbb{Z} \to \mathbb{R}$ by:
$$\forall t \in \mathbb{Z}, \quad \psi_n(t) = \Pr[T_n = t].$$

For all $n \in \mathbb{N}$, let $a_n := \psi_n(n + 2)$, $z_n := \psi_n(n)$.

Since, $\forall n \in \mathbb{N}$, we have
both $z_n = \psi_n(n) = \Pr[T_n = n]$ and $\Pr[T_n = n] = \Pr[X_1 + \cdots + X_n = n] \geq \Pr[X_1 = \cdots = X_n = 1] = q^n > 0$,
we conclude: $\forall n \in \mathbb{N}, \ z_n = \Pr[T_n = n] > 0$.

**Claim:** Let $n \in [3..\infty)$. Then $\Pr[X_{n-1} = 0 = X_n | T_n = n] = p^2 \cdot \frac{a_{n-2}}{z_n}$.

**Proof of Claim:** We have $T_n = X_1 + \cdots + X_{n-2} + X_{n-1} + X_n$.

Since $\Pr[(X_{n-1} = 0 = X_n) \& (T_n = n)]$
$$= \Pr[(X_{n-1} = 0 = X_n) \& (X_1 + \cdots + X_{n-2} + X_{n-1} + X_n = n)]$$
$$= \Pr[(X_{n-1} = 0 = X_n) \& (X_1 + \cdots + X_{n-2} + 0 + 0 = n)]$$
$$= \Pr[(X_{n-1} = 0 = X_n) \& (X_1 + \cdots + X_{n-2} = n)],$$
it follows, from independence of $X_1, \ldots, X_n$, that
$$\Pr[(X_{n-1} = 0 = X_n) \& (T_n = n)]$$
$$= (\Pr[X_{n-1} = 0]) \cdot (\Pr[X_n = 0]) \cdot (\Pr[X_1 + \cdots + X_{n-2} = n]).$$

So, since $\Pr[X_{n-1} = 0] = p = \Pr[X_n = 0]$ and since $X_1 + \cdots + X_{n-2} = T_{n-2}$,
we get: $\Pr[(X_{n-1} = 0 = X_n) \& (T_n = n)] = p^2 \cdot (\Pr[T_{n-2} = n]).$

Then $\Pr[X_{n-1} = 0 = X_n | T_n = n] = \frac{\Pr[(X_{n-1} = 0 = X_n) \& (T_n = n)]}{\Pr[T_n = n]}$
$$= \frac{p^2 \cdot (\Pr[T_{n-2} = n])}{\Pr[T_n = n]} = p^2 \cdot \frac{\psi_{n-2}(n)}{\psi_n(n)} = p^2 \cdot \frac{a_{n-2}}{z_n}.$$

*End of proof of Claim.*

Because of the Claim, we want to show: as $n \to \infty$, $p^2 \cdot \frac{a_{n-2}}{z_n} \to p^2$.

**Want:** as $n \to \infty$, $\frac{a_{n-2}}{z_n} \to 1$.

We compute: $\forall n \in \mathbb{N}, \ E[X_n] = q + 10r$.

Recall (§7): $q + 10r = 1$. Then: $\forall n \in \mathbb{N}, \ E[X_n] = 1$.

We compute: $\forall n \in \mathbb{N}, \ Var[X_n] = q + 100r - 1$.

**Let** $v := q + 100r - 1$. Then: $\forall n \in \mathbb{N}, \ Var[X_n] = v$.

Since $v = (q + 10r - 1) + 90r = 0 + 90r = 90r > 0$, we get: $v > 0$.

**Let** $\tau := 1/\sqrt{2\pi v}$. Then: $\tau > 0$.

**Let** $\alpha := 1$. Then, $\alpha \in \mathbb{Z}$ and $\forall n \in \mathbb{N}, \ E[X_n] = \alpha$.

**Let** $E := \{0, 1, 10\}$. Then, $\forall n \in \mathbb{N}, \ \{t \in \mathbb{Z} | \Pr[X_n = t] > 0\} = E$. 

Also, $E$ is residue-unconstrained.

**Let** $t_0 := 2$. Then $t_0 \in \mathbb{Z}$ and $\forall n \in \mathbb{N}, \ t_0 + n\alpha = n + 2$.

By Theorem 9.7, as $n \to \infty$, $\sqrt{n} \cdot (\Pr[T_n = t_0 + n\alpha]) \to 1/\sqrt{2\pi v}$.

Then: as $n \to \infty$, $\sqrt{n} \cdot (\Pr[T_n = n + 2]) \to 1/\sqrt{2\pi v}$.

Then: as $n \to \infty$, $\sqrt{n} \cdot (\psi_n(n + 2)) \to \tau$.

Then: as $n \to \infty$, $\sqrt{n} \cdot a_n \to \tau$.

Then: as $n \to \infty$, $\sqrt{n - 2} \cdot a_{n-2} \to \tau$.

By Theorem 9.9, as $n \to \infty$, $\sqrt{n} \cdot (\Pr[T_n = n\alpha]) \to 1/\sqrt{2\pi v}$.

Then: as $n \to \infty$, $\sqrt{n} \cdot (\Pr[T_n = n]) \to 1/\sqrt{2\pi v}$.

Then: as $n \to \infty$, $\sqrt{n} \cdot (\psi_n(n)) \to \tau$.

Then: as $n \to \infty$, $\sqrt{n} \cdot z_n \to \tau$.

Then: as $n \to \infty$, $\frac{\sqrt{n - 2} \cdot a_{n-2}}{\sqrt{n} \cdot z_n} \to \frac{\tau}{\tau}$.

Then: as $n \to \infty$, $\frac{\sqrt{n - 2} \cdot a_{n-2}}{\sqrt{n} \cdot z_n} \to 1$.

Also, as $n \to \infty$, $\frac{a_{n-2}}{z_n} \to 1$.

Multiplying these last two limits, we get: as $n \to \infty$, $\frac{a_{n-2}}{z_n} \to 1$.

14. **FRACTION OF PROFESSORS GETTING A ZERO AWARD**

**Let** $(p, q, r) := \frac{(1, 9^{-1/10}, 9^{-1})}{1 + 9^{-1/10} + 9^{-1}}$.

We compute $(p, q, r) \approx (0.5225, 0.4194, 0.0581)$, accurate to four decimal places.

**Let** $X_1, X_2, X_3, \ldots$ be iid random-variables s.t., $\forall n \in \mathbb{N}$,

\begin{align*}
\Pr[X_n = 0] &= p, \\
\Pr[X_n = 1] &= q, \\
\Pr[X_n = 10] &= r.
\end{align*}

For all $n \in \mathbb{N}$, **let** $T_n := X_1 + \cdots + X_n$.

For all $\ell \in \mathbb{N}$, **let** $I_\ell$ be the indicator variable of the event: $X_\ell = 0$.

For all $n \in \mathbb{N}$, **let** $J_n := (I_1 + \cdots + I_n)/n$.

Using the GFA’s first awards system, the random-variable

\[ J_N \] conditioned on $T_N = N$

represents the fraction of professors receiving a $0 award.

In this section, **we will prove:** $\forall \varepsilon > 0$,

\[ \text{as } n \to \infty, \quad \Pr \left[ p - \varepsilon < J_n < p + \varepsilon \mid T_n = n \right] \to 1. \]

Assume, for a moment, that this is true.
Then: as $n \to \infty$, 
\[ \Pr \left[ p - 0.02 < J_n < p + 0.02 \mid T_n = n \right] \to 1. \]
From this, it follows that, if $N$ is sufficiently large, then 
\[ \Pr \left[ p - 0.02 < J_N < p + 0.02 \mid T_N = N \right] > 0.99, \]
so 
\[ \Pr \left[ p - 0.02 < J_N \mid T_N = N \right] > 0.99, \]
so 
\[ \Pr \left[ J_N > p - 0.02 \mid T_N = N \right] > 0.99. \]
Since $p \approx 0.5225$, accurate to four decimal places, we get 
\[ p - 0.02 > 0.5, \]
so 
\[ J_N > p - 0.02 \Rightarrow J_n > 0.5, \]
so 
\[ \Pr \left[ J_N > p - 0.02 \mid T_N = N \right] \leq \Pr \left[ J_N > 0.5 \mid T_N = N \right]. \]
Therefore, if $N$ is sufficiently large, then 
\[ \Pr \left[ J_N > 0.5 \mid T_N = N \right] > 0.99, \]
and so, under the GFA’s first system, with probability $> 99\%$, over half of the professors get $0.

**Given** $\varepsilon > 0$, **want**: as $n \to \infty$, 
\[ \Pr \left[ p - \varepsilon < J_n < p + \varepsilon \mid T_n = n \right] \to 1. \]
**Let** $E := \{0, 1, 10\}$. **Let** $\alpha := 1$.
Then: $\alpha \in \mathbb{Z}$ and $\forall n \in \mathbb{N}$, 
\[ \mathbb{E}[X_n] = \alpha. \]
For all $n \in \mathbb{N}$, **let** $\kappa_n := \mathbb{E} \left[ I_n \mid T_n = n \right]$. 
Then: $\forall n \in \mathbb{N}$, 
\[ \kappa_n = \Pr \left[ X_n = 0 \mid T_n = n \right]. \]
By Theorem 11.5, we get: 
\[ \text{as } n \to \infty, \quad \Pr[X_n = 0 \mid X_1 + \cdots + X_n = n\alpha] \to p. \]
That is, as $n \to \infty$, \[ \Pr[X_n = 0 \mid T_n = n] \to p. \]
That is, as $n \to \infty$, 
\[ \kappa_n \to p. \]
So, $\exists n_0 \in \mathbb{N}$ s.t., $\forall n \in [n_0, \infty)$, 
we have 
\[ p - (\varepsilon/2) < \kappa_n < p + (\varepsilon/2), \]
and so both $p - \varepsilon < \kappa_n - (\varepsilon/2)$ and $\kappa_n + (\varepsilon/2) < p + \varepsilon$, and so 
\[ [\kappa_n - (\varepsilon/2) < J_n < \kappa_n + (\varepsilon/2)] \Rightarrow [p - \varepsilon < J_n < p + \varepsilon], \]
and so 
\[ \Pr[\kappa_n - (\varepsilon/2) < J_n < \kappa_n + (\varepsilon/2) \mid T_n = n] \leq \Pr[\kappa_n - (\varepsilon/2) < J_n < \kappa_n + (\varepsilon/2) \mid T_n = n]. \]
**It therefore suffices to show:** 
\[ \text{as } n \to \infty, \quad \Pr[\kappa_n - (\varepsilon/2) < J_n < \kappa_n + (\varepsilon/2) \mid T_n = n] \to 1. \]
We have: $\forall n \in \mathbb{N}$, $T_n$ is invariant under permutation of $X_1, \ldots, X_n$, as is the joint-distribution of $X_1, \ldots, X_n$. 
Then: $\forall n \in \mathbb{N}, \forall i \in [1..n], \quad  E \left[ I_i \mid T_n = n \right] = E \left[ I_n \mid T_n = n \right].$
Then: $\forall n \in \mathbb{N}, \forall i \in [1..n], \quad  E \left[ I_i \mid T_n = n \right] = \kappa_n.$
Since, $\forall n \in \mathbb{N}$, $J_n = (I_1 + \cdots + I_n)/n$, we get: 
\[ \forall n \in \mathbb{N}, \quad E \left[ J_n \mid T_n = n \right] = (\sum_{i=1}^{n} E \left[ I_i \mid T_n = n \right]) / n. \]
Then: \( \forall n \in \mathbb{N}, \ E [ J_n | T_n = n ] = n\kappa_n/n. \)

Then: \( \forall n \in \mathbb{N}, \ E [ J_n | T_n = n ] = \kappa_n. \)

For all \( n \in \mathbb{N}, \) let \( v_n := \text{Var} [ J_n | T_n = n ]. \)

Then, by Chebyshev’s inequality, we have: \( \forall n \in \mathbb{N}, \)

\[
\Pr \left[ \kappa_n - (\varepsilon/2) < J_n < \kappa_n + (\varepsilon/2) | T_n = n \right] \geq 1 - (v_n/(\varepsilon/2)^2).
\]

**It therefore suffices to show:** as \( n \to \infty, \) \( v_n \to 0. \)

We have: \( \forall n \in \mathbb{N}, \)

\[
\text{Var} [ J_n | T_n = n ] = \left( E [ J_n^2 | T_n = n ] \right) - \left( E [ J_n | T_n = n ] \right)^2.
\]

So, since, as \( n \to \infty, \)

\[\kappa_n^2 \to p^2,\]

we want: as \( n \to \infty, \)

\[E [ J_n^2 | T_n = n ] \to p^2.
\]

For all \( n \in [2, \infty), \) let \( \lambda_n := E [ I_{n-1} \cdot I_n | T_n = n ]. \)

Then: \( \forall n \in [2, \infty), \)

\[
\lambda_n = \Pr \left[ X_{n-1} = 0 = X_n | T_n = n \right].
\]

So, by the result of §13, we get: as \( n \to \infty, \)

\[\lambda_n \to p^2.
\]

For all \( n \in \mathbb{N}, \) since \( I_n \) is an indicator variable, we get: \( I_n \in \{0,1\} \) a.s.

Then: \( \forall n \in \mathbb{N}, \)

\[
I_n = I_n^2 \text{ a.s.}
\]

Then: \( \forall n \in \mathbb{N}, \)

\[\kappa_n = E [ I_n^2 | T_n = n ].
\]

We have: \( \forall n \in \mathbb{N}, \) \( T_n \) is invariant under permutation of \( X_1, \ldots, X_n, \)

as is the joint-distribution of \( X_1, \ldots, X_n. \)

Then \( \forall n \in \mathbb{N}, \forall i \in [1..n], \)

\[
E [ I_i^2 | T_n = n ] = E [ I_{n-1} \cdot I_n | T_n = n ],
\]

so, \( \forall n \in \mathbb{N}, \forall i \in [1..n], \)

\[E [ I_i^2 | T_n = n ] = \kappa_n.
\]

Similarly, \( \forall n \in [2, \infty), \forall i, j \in [1..n], \)

if \( i \neq j, \)

\[
E [ I_i \cdot I_j | T_n = n ] = E [ I_{n-1} \cdot I_n | T_n = n ],
\]

so, \( \forall n \in [2, \infty), \forall i, j \in [1..n], \)

if \( i \neq j, \)

\[E [ I_i \cdot I_j | T_n = n ] = \lambda_n.
\]

For all \( n \in \mathbb{N}, \) for all \( i, j \in [1..n], \) let \( c_{ijn} := E [ I_i \cdot I_j | T_n = n ] \)

Then: \( \forall n \in \mathbb{N}, \forall i, j \in [1..n], \)

\[c_{ijn} = \left\{ \begin{array}{ll} \kappa_n, & \text{if } i = j \\ \lambda_n, & \text{if } i \neq j. \end{array} \right.
\]

Then: \( \forall n \in \mathbb{N}, \)

\[
\sum_{i=1}^n \sum_{j=1}^n c_{ijn} = n \cdot \kappa_n + (n^2 - n) \cdot \lambda_n.
\]

Recall: as \( n \to \infty, \)

\[\kappa_n \to p \text{ and } \lambda_n \to p^2.
\]

Since \( \forall n \in \mathbb{N}, \) \( J_n = (I_1 + \cdots + I_n)/n, \)

we get: \( \forall n \in \mathbb{N}, \)

\[
J_n^2 = \left( \sum_{i=1}^n \sum_{j=1}^n \kappa_{ijn} \right) / n^2.
\]

Then: \( \forall n \in \mathbb{N}, \)

\[E [ J_n^2 | T_n = n ] = (\sum_{i=1}^n \sum_{j=1}^n c_{ijn}) / n^2.
\]

Then: \( \forall n \in \mathbb{N}, \)

\[E [ J_n^2 | T_n = n ] = (1/n) \cdot \kappa_n + (1 - (1/n)) \cdot \lambda_n.
\]

Then: as \( n \to \infty, \)

\[E [ J_n^2 | T_n = n ] \to 0 \cdot p + 1 \cdot p^2.
\]

Then: as \( n \to \infty, \)

\[E [ J_n^2 | T_n = n ] \to p^2.
\]
15. BOLTZMANN DISTRIBUTIONS ON FINITE SETS

Recall (§8): ∀ nonempty countable set Θ,
   \( \mathcal{M}_\Theta \) is the set of measures on Θ
   and \( \mathcal{F}\mathcal{M}_\Theta^\times \) is the set of nonzero finite measures on Θ
   and \( \mathcal{P}_\Theta \) is the set of probability measures on Θ.

Recall (§11): ∀ nonempty countable set Θ, \( \forall \mu \in \mathcal{F}\mathcal{M}_\Theta^\times \),
   \( \mathcal{N}(\mu) \) is the normalization of \( \mu \).

**DEFINITION 15.1.** Let \( E \subseteq \mathbb{R} \) be nonempty and finite, \( \beta \in \mathbb{R} \).

The **unnormalized-\( \beta \)-Boltzmann distribution on \( E \)** is

the measure \( \hat{B}_E^\beta \in \mathcal{F}\mathcal{M}_E^\times \) defined by:

\[ \forall \varepsilon \in E, \quad \hat{B}_E^\beta(\varepsilon) = e^{-\beta \varepsilon}. \]

Also, the **\( \beta \)-Boltzmann distribution on \( E \)** is

\[ B_E^\beta := \mathcal{N}(\hat{B}_E^\beta) \in \mathcal{P}_E. \]

Then: \( \forall \varepsilon \in E \), we have: \( B_E^\beta(\varepsilon) = (\hat{B}_E^\beta(\varepsilon)) / (\hat{B}_E^\beta(E)). \)

**Example:** Let \( E := \{2, 4, 8, 9\} \) and let \( \beta \in \mathbb{R} \).

Then: \( \hat{B}_E^\beta(2) = e^{-2\beta}, \quad \hat{B}_E^\beta(4) = e^{-4\beta}, \quad \hat{B}_E^\beta(8) = e^{-8\beta}, \quad \hat{B}_E^\beta(9) = e^{-9\beta}. \)

Let \( C := 1/(e^{-2\beta} + e^{-4\beta} + e^{-8\beta} + e^{-9\beta}) \).

Then: \( B_E^\beta(2) = Ce^{-2\beta}, \quad B_E^\beta(4) = Ce^{-4\beta}, \quad B_E^\beta(8) = Ce^{-8\beta}, \quad B_E^\beta(9) = Ce^{-9\beta}. \)

**Example:** Let \( E := \{0, 1, 10\} \) and let \( \beta \in \mathbb{R} \).

Then: \( \hat{B}_E^\beta(0) = 1, \quad \hat{B}_E^\beta(1) = e^{-\beta}, \quad \hat{B}_E^\beta(10) = e^{-10\beta}. \)

Let \( C := 1/(1 + e^{-\beta} + e^{-10\beta}) \).

Then: \( B_E^\beta(0) = C, \quad B_E^\beta(1) = Ce^{-\beta}, \quad B_E^\beta(10) = Ce^{-10\beta}. \)

Recall (§8): For any countable set Θ, for any \( \mu \in \mathcal{M}_\Theta \),
   \( S_\mu \) is the support of \( \mu \).

Note: ∀ nonempty finite \( E \subseteq \mathbb{R}, \forall \beta \in \mathbb{R} \), we have: \( S_{\hat{B}_E^\beta} = E = S_{B_E^\beta} \).

**THEOREM 15.2.** Let \( E \subseteq \mathbb{R} \) be nonempty and finite.

Let \( \varepsilon_0 \in E, \beta, \xi \in \mathbb{R} \). Then: \( B_E^{\beta-\xi}(\varepsilon_0 - \xi) = B_{E}^{\beta}(\varepsilon_0). \)
Proof. We have: $B^E_{\beta} \{ \varepsilon_0 - \xi \} = \frac{e^{-\beta(\varepsilon_0 - \xi)}}{\sum_{\varepsilon \in E} e^{-\beta(\varepsilon - \xi)}} = \frac{\sum_{\varepsilon \in E} e^{-\beta(\varepsilon - \xi)}}{e^{-\beta(\varepsilon_0 - \xi)}} \cdot e^{-\beta \xi} = \sum_{\varepsilon \in E} e^{-\beta \varepsilon} = B^E_{\beta} \{ \varepsilon_0 \} \). \square

Recall (§8): Let $\Theta \subseteq \mathbb{R}$ be countable, $\mu \in \mathcal{P}_\Theta$. Assume $\# S_\mu < \infty$. Then $|\mu|_1 < \infty$ and $M_\mu$ is the mean of $\mu$ and $V_\mu$ is the variance of $\mu$.

Let $E \subseteq \mathbb{R}$ be nonempty and finite. Let $\beta \in \mathbb{R}$. We define:

\[
\begin{align*}
\Gamma^E_\beta & := \sum_{\varepsilon \in E} [\varepsilon \cdot e^{\beta \varepsilon}], \\
\Delta^E_\beta & := \sum_{\varepsilon \in E} [e^{\beta \varepsilon}], \\
A^E_\beta & := \Gamma^E_\beta / \Delta^E_\beta.
\end{align*}
\]

Then: $\Gamma^E_\beta = \sum_{\varepsilon \in E} [\varepsilon \cdot (B^E_{\beta} \{ \varepsilon \})]$. Also, $\Delta^E_\beta = \sum_{\varepsilon \in E} [B^E_{\beta} \{ \varepsilon \}]$, and so $\Delta^E_\beta = \hat{B}^E_{\beta} (E)$. Since $\frac{\Gamma^E_\beta}{\Delta^E_\beta} = \frac{\sum_{\varepsilon \in E} [\varepsilon \cdot (B^E_{\beta} \{ \varepsilon \})]}{\sum_{\varepsilon \in E} [B^E_{\beta} \{ \varepsilon \}]} = \sum_{\varepsilon \in E} [\varepsilon \cdot (B^E_{\beta} \{ \varepsilon \})]$, we conclude: $A^E_\beta = M_{B^E_{\beta}}$.

Then: $A^E_\beta$ is the average value of any random-variable with distribution $B^E_{\beta}$.

(“A” is for “Average.”)

**THEOREM 15.3.** Let $E \subseteq \mathbb{R}$ be nonempty and finite. Let $\beta, \xi \in \mathbb{R}$. Then: $A^E_{\beta - \xi} = A^E_\beta - \xi$.

Proof. **Want:** $M_{B^E_{\beta - \xi}} = M_{B^E_{\beta}} - \xi$.

Let $\lambda := B^E_{\beta - \xi}$, $\mu := B^E_{\beta}$. **Want:** $M_\lambda = M_\mu - \xi$.

We have: $\lambda \in \mathcal{P}_{E - \xi}$ and $\mu \in \mathcal{P}_E$.

By Theorem 15.2, we have: $\forall \varepsilon \in E$, $B^E_{\beta - \xi} \{ \varepsilon - \xi \} = B^E_{\beta} \{ \varepsilon \}$.

Then: $\forall \varepsilon \in E$, $\lambda \{ \varepsilon - \xi \} = \mu \{ \varepsilon \}$.

Since $\mu = B^E_{\beta} \in \mathcal{P}_E$, we get: $\mu (E) = 1$.

Then: $M_\lambda = \sum_{\varepsilon \in E} [\varepsilon - \xi, \lambda \{ \varepsilon - \xi \}]]$

\[
= \sum_{\varepsilon \in E} [\varepsilon - \xi, \mu \{ \varepsilon \}]]
\]

\[
= (\sum_{\varepsilon \in E} [\varepsilon, \mu \{ \varepsilon \}]) - (\sum_{\varepsilon \in E} [\xi, \mu \{ \varepsilon \}])
\]

\[
= (\sum_{\varepsilon \in E} [\varepsilon, \mu \{ \varepsilon \}]) - \xi \cdot (\sum_{\varepsilon \in E} [\mu \{ \varepsilon \}])
\]

\[
= M_\mu - \xi \cdot (\mu (E)) = M_\mu - \xi \cdot 1 = M_\mu - \xi. \square
\]
THEOREM 15.4. Let $E \subseteq \mathbb{R}$ be nonempty and finite. Then:

$$
\begin{align*}
&\text{as } \beta \to -\infty, \quad A^E_\beta \to \max E \\
&\text{and as } \beta \to \infty, \quad A^E_\beta \to \min E.
\end{align*}
$$

The proof is a matter of bookkeeping, best explained by example:

Let $E := \{2, 4, 8, 9\}$. Then $\min E = 2$ and $\max E = 9$.

Also, 
\[ A^E_\beta = \frac{2e^{-2\beta} + 4e^{-4\beta} + 8e^{-8\beta} + 9e^{-9\beta}}{e^{-2\beta} + e^{-4\beta} + e^{-8\beta} + e^{-9\beta}}, \]

and so 
\[ A^E_\beta \to 2/1 \]

and as 
\[ A^E_\beta \to 9/1, \]

and so 
\[ A^E_\beta \to \min E \]

and as 
\[ A^E_\beta \to \max E. \]

For all nonempty, finite $E \subseteq \mathbb{R}$, define $A^E_\bullet : \mathbb{R} \to \mathbb{R}$ by:

\[ \forall \beta \in \mathbb{R}, \quad A^E_\bullet (\beta) = A^E_\beta. \]

THEOREM 15.5. Let $E \subseteq \mathbb{R}$. Assume: $2 \leq \# E < \infty$.

Then: $A^E_\bullet$ is a strictly-decreasing $C^\omega$-diffeomorphism from $\mathbb{R}$ onto $(\min E; \max E)$.

Proof. Let $\kappa := \# E$. Choose $\varepsilon_1, \ldots, \varepsilon_\kappa \in \mathbb{R}$ s.t. $E = \{\varepsilon_1, \ldots, \varepsilon_\kappa\}$.

Then: $2 \leq \kappa < \infty$ and $\varepsilon_1, \ldots, \varepsilon_\kappa$ are distinct.

Then: $\forall \beta \in \mathbb{R}, A^E_\bullet (\beta) = \sum_{i=1}^\kappa \left[\varepsilon_i \cdot e^{-\beta \cdot \varepsilon_i}\right]$. Then $A^E_\bullet : \mathbb{R} \to \mathbb{R}$ is $C^\omega$.

So, by Theorem 15.4 and the $C^\omega$-Inverse Function Theorem and the Mean Value Theorem, it suffices to show: $(A^E_\bullet)' < 0$ on $\mathbb{R}$.

Given $\beta \in \mathbb{R}$, want: $(A^E_\bullet)'(\beta) < 0$.

Let $P := \sum_{i=1}^\kappa \left[\varepsilon_i \cdot e^{-\beta \cdot \varepsilon_i}\right]$, $P' := \sum_{i=1}^\kappa \left[(-\varepsilon_i^2) \cdot e^{-\beta \cdot \varepsilon_i}\right]$.

Let $Q := \sum_{j=1}^\kappa \left[e^{-\beta \cdot \varepsilon_j}\right]$, $Q' := \sum_{j=1}^\kappa \left[(-\varepsilon_j) \cdot e^{-\beta \cdot \varepsilon_j}\right]$.

Then $Q > 0$. Also, by the Quotient Rule, $(A^E_\bullet)'(\beta) = [Q P' - PQ']/Q^2$.

Want: $Q P' - PQ' < 0$.

We have: $Q P' = \sum_{i=1}^\kappa \sum_{j=1}^\kappa \left[(-\varepsilon_i^2) \cdot e^{-\beta \cdot (\varepsilon_i + \varepsilon_j)}\right]$.

We have: $P Q' = \sum_{i=1}^\kappa \sum_{j=1}^\kappa \left[(-\varepsilon_i \varepsilon_j) \cdot e^{-\beta \cdot (\varepsilon_i + \varepsilon_j)}\right]$.

Then: $Q P' - PQ' = \sum_{i=1}^\kappa \sum_{j=1}^\kappa \left[(-\varepsilon_i^2 + \varepsilon_i \varepsilon_j) \cdot e^{-\beta \cdot (\varepsilon_i + \varepsilon_j)}\right]$.

Interchanging $i$ and $j$, we get:

$Q P' - PQ' = \sum_{j=1}^\kappa \sum_{i=1}^\kappa \left[(-\varepsilon_j^2 + \varepsilon_j \varepsilon_i) \cdot e^{-\beta \cdot (\varepsilon_i + \varepsilon_j)}\right]$.

By commutativity of addition and multiplication, adding the last two equations gives:

$2 \cdot (Q P' - PQ') = \sum_{i=1}^\kappa \sum_{j=1}^\kappa \left[(-\varepsilon_i^2 - \varepsilon_j^2 + 2\varepsilon_i \varepsilon_j) \cdot e^{-\beta \cdot (\varepsilon_i + \varepsilon_j)}\right]$. 

Then: \(2 \cdot (QP' - PQ') = \sum_{i=1}^{k} \sum_{j=1}^{n} [-(\varepsilon_i - \varepsilon_j)^2 \cdot e^{-\beta(\varepsilon_i + \varepsilon_j)}].\)

Then: \(2 \cdot (QP' - PQ') < 0.\)

Then: \(QP' - PQ' < 0.\) \(\Box\)

**DEFINITION 15.6.** Let \(E \subseteq \mathbb{R}\).

Assume: \(2 \leq \#E < \infty.\) Let \(\alpha \in (\min E, \max E).\)

The \(\alpha\)-Boltzmann parameter on \(E\) is: \(BP^E_\alpha := (A^E_\alpha)^{-1}(\alpha)\).

So the \(\alpha\)-Boltzmann parameter on \(E\) is the unique \(\beta \in \mathbb{R}\) s.t. \(A^E_\beta = \alpha\).

**Example:** Computations at the end of §6 show:

\[\forall \beta \in \mathbb{R}, \text{ if } \frac{e^{-\beta} + 10e^{-10\beta}}{1 + e^{-\beta} + 10e^{-10\beta}} = 1, \text{ then } e^{-\beta} = 9^{-1/10}.\]

Then, \(\forall \beta \in \mathbb{R}, \text{ if } A^{(0,1,10)}_\beta = 1, \text{ then } \beta = (\ln 9)/10.\)

Then: \((A^{(0,1,10)}_\beta)^{-1}(1) = (\ln 9)/10.\)

Then: \(BP^{(0,1,10)}_1 = (\ln 9)/10.\)

**Example:** Let \(E := \{2, 4, 8, 9\}, \alpha := 3, \beta := BP^E_\alpha.\)

To compute \(\beta,\) we need to solve \(A^E_\beta = 3\) for \(\beta.\)

Since \(A^E_\beta\) is strictly-decreasing, there are iterative methods of solution, and we get: \(\beta \approx ?????,\) accurate to four decimals.

(Thanks to C. Prouty for these calculations. See §25.)

**THEOREM 15.7.** Let \(E \subseteq \mathbb{R}.\) Assume \(2 \leq \#E < \infty.\) Let \(\alpha \in (\min E, \max E).\) Let \(\xi \in \mathbb{R}.\) Then: \(BP^{E-\xi}_{\alpha-\xi} = BP^E_\alpha.\)

**Proof.** Let \(\beta := BP^E_\alpha.\) Want: \(BP^{E-\xi}_{\alpha-\xi} = \beta.\)

By Theorem 15.3, we get: \(A^{E-\xi}_\beta = A^E_\beta - \xi.\)

Since \(\beta = BP^E_\alpha = (A^E_\alpha)^{-1}(\alpha),\) we get: \((A^E_\alpha)(\beta) = \alpha.\)

Since \((A^{E-\xi}_\beta)(\beta) = A^{E-\xi}_\beta = A^E_\beta - \xi = ((A^E_\beta)(\beta)) - \xi = \alpha - \xi,\)

we get: \(\beta = (A^{E-\xi}_\beta)^{-1}(\alpha - \xi).\)

Then: \(BP^{E-\xi}_{\alpha-\xi} = (A^{E-\xi}_\beta)^{-1}(\alpha - \xi) = \beta.\) \(\Box\)

16. Residue-unconstrained finite sets

In this section, we generalize the result of §12 from \(\{0, 1, 10\}\) to arbitrary finite residue-unconstrained sets. Our main theorem is Theorem 16.2 below.

Recall (§8): \(\forall \text{ nonempty finite set } F, \forall f \in F, \nu_F\{f\} = 1/(\#F).\)
THEOREM 16.1. Let $E \subseteq \mathbb{Z}$ be finite and residue-unconstrained. Let $\alpha \in \mathbb{Z}$. Assume $\alpha \in (\min E; \max E)$. Let $\beta := BP_E^\alpha$. Let $t_1, t_2, \ldots \in \mathbb{Z}$. Assume: $\{t_n - n\alpha \mid n \in \mathbb{N}\}$ is bounded.

For all $n \in \mathbb{N}$, let $\Omega_n := \{f \in E^n \mid f_1 + \cdots + f_n = t_n\}$.

Let $\varepsilon_0 \in E$. Then: as $n \to \infty$, $\nu_{\Omega_n}\{f \in \Omega_n \mid f_n = \varepsilon_0\} \to B_{\beta}^E(\varepsilon_0)$.

Recall: $\nu_\emptyset(\emptyset) = -1$.

So, since $B_{\beta}^E(\varepsilon_0) > 0$, part of the content of this theorem is:

$\forall$ sufficiently large $n \in \mathbb{N}$, $\Omega_n \neq \emptyset$.

See Claim 2 in the proof below.

Proof. Let $\mu := B_{\beta}^E$. Then: $\mu \in \mathcal{P}_E$ and $S_\mu = E$.

By assumption, $E$ is finite. Then $S_\mu$ is finite.

So, since $\mu \in \mathcal{P}_E \subseteq \mathcal{F}_E$, we get: $|\mu|_1 < \infty$ and $|\mu|_2 < \infty$.

Since $\beta = BP_\alpha^E = (A_\alpha^E)^{-1}(\alpha)$, we get: $(A_\alpha^E)(\beta) = \alpha$.

So, since $M_\mu = M_{B_\beta^E} = A_\beta^E = (A_\alpha^E)(\beta)$, we get: $M_\mu = \alpha$.

For all $n \in \mathbb{N}$, define $\psi_n : \mathbb{Z} \to \mathbb{R}$ by:

$\forall t \in \mathbb{Z}$, $\psi_n(t) = \mu^\alpha \{f \in E^n \mid f_1 + \cdots + f_n = t\}$.

Then: $\forall n \in \mathbb{N}$, $\psi_n(t_n) = \mu^\alpha(\Omega_n)$.

Since $E$ is finite and residue-unconstrained, we get: $2 \leq \#E < \infty$.

Since $\#S_\mu = \#E \geq 2$, by Theorem 8.6, we get: $V_\mu > 0$.

So, since $V_\mu = |\mu|_2 - M_\mu^2 \leq |\mu|_2 < \infty$, we conclude:

$0 < V_\mu < \infty$.

Let $v := V_\mu$.

Then $0 < v < \infty$. Then $1/\sqrt{2\pi v} > 0$.

Let $\tau := 1/\sqrt{2\pi v}$. Then $\tau > 0$.

Claim 1: As $n \to \infty$, $\sqrt{n} \cdot (\psi_n(t_n)) \to \tau$.

Proof of Claim 1: By Theorem 9.10, we get:

as $n \to \infty$, $\sqrt{n} \cdot (\mu^\alpha \{f \in E^n \mid f_1 + \cdots + f_n = t_n\}) \to 1/\sqrt{2\pi v}$.

Then: as $n \to \infty$, $\sqrt{n} \cdot (\psi_n(t_n)) \to \tau$.

End of proof of Claim 1.

Since $\tau > 0$, by Claim 1, choose $n_0 \in \mathbb{N}$ s.t.

$\forall n \in [n_0, \infty)$, $\sqrt{n} \cdot (\psi_n(t_n)) > 0$.

Claim 2: Let $n \in [n_0, \infty)$. Then: $\mu^n(\Omega_n) > 0$.

Proof of Claim 2: Recall: $\psi_n(t_n) = \mu^n(\Omega_n)$. Want: $\psi_n(t_n) > 0$.

By the choice of $n_0$, we get: $\sqrt{n} \cdot (\psi_n(t_n)) > 0$. Then: $\psi_n(t_n) > 0$.

End of proof of Claim 2.
Recall: \( \mu \in \mathcal{P}_E. \)
Then: \( \forall n \in \mathbb{N}, \mu^n \in \mathcal{P}_E^n, \) so \( \mu^n(\Omega_n) \leq 1. \)
So, by Claim 2, \( \forall n \in [n_0, \infty), \) \( 0 < \mu^n(\Omega_n) \leq 1. \)
Also, we have: \( \forall n \in \mathbb{N}, \) \( (\mu^n|\Omega_n)(\Omega_n) = \mu^n(\Omega_n). \)
Then: \( \forall n \in [n_0, \infty), \) \( 0 < (\mu^n|\Omega_n)(\Omega_n) \leq 1. \)
Then: \( \forall n \in [n_0, \infty), \) \( \mu^n|\Omega_n \in \mathcal{F}_\Omega^\infty. \)
Then: \( \forall n \in [n_0, \infty), \) \( \mathcal{N}(\mu^n|\Omega_n) \in \mathcal{P}_{\Omega_n}. \)

Claim 3: Let \( n \in [n_0, \infty). \) Then: \( \mathcal{N}(\mu^n|\Omega_n) = \nu_{\Omega_n}. \)

Proof of Claim 3: Let \( \theta := \mathcal{N}(\mu^n|\Omega_n), \) \( F := \Omega_n. \) Then \( \theta \in \mathcal{P}_F. \)

Want: \( \theta = \nu_F. \) By Theorem 8.9, given \( f, g \in F, \) want: \( \theta\{f\} = \theta\{g\}. \)
By Claim 2, we have: \( \mu^n(\Omega_n) > 0. \)
Since \( (\mu^n|\Omega_n)(\Omega_n) = \mu^n(\Omega_n) \) and \( \theta = \mathcal{N}(\mu^n|\Omega), \) we get: \( \theta = \frac{\mu^n|\Omega_n}{\mu^n(\Omega_n)}. \)

Want: \( \frac{(\mu^n|\Omega_n\{f\})}{\mu^n(\Omega_n)} = \frac{(\mu^n|\Omega_n\{g\})}{\mu^n(\Omega_n)}. \)
Want: \( (\mu^n|\Omega_n\{f\}) = (\mu^n|\Omega_n\{g\}). \)
Since \( f, g \in F = \Omega_n, \) we get:
\( (\mu^n|\Omega_n\{f\}) = \mu^n\{f\} \) and \( (\mu^n|\Omega_n\{g\}) = \mu^n\{g\}. \)
Want: \( \mu^n\{f\} = \mu^n\{g\}. \)
Since \#E \geq 2, we get: \( E \neq \emptyset. \) Then \( \hat{B}_E^F(E) > 0. \)

Let \( C := 1/(\hat{B}_E^F(E)). \) Then \( \mathcal{N}(\hat{B}_E^F) = C \cdot \hat{B}_E^F. \)
By definition of \( \hat{B}_E^F, \) we have: \( \forall \varepsilon \in E, \) \( \hat{B}_E^F(\varepsilon) = e^{-\beta \varepsilon}. \)
So, since \( \mu = B^E_\beta = \mathcal{N}(\hat{B}_E^F) = C \cdot \hat{B}_E^F, \) we get:
\( \forall \varepsilon \in E, \) \( \mu\{\varepsilon\} = C e^{-\beta \varepsilon}. \)
Since \( f \in F = \Omega_n, \) by definition of \( \Omega_n, \) we get: \( f_1 + \cdots + f_n = t_n. \)
Since \( g \in F = \Omega_n, \) by definition of \( \Omega_n, \) we get: \( g_1 + \cdots + g_n = t_n. \)
Since \( f_1 + \cdots + f_n = t_n = g_1 + \cdots + g_n, \)
we get: \( C^n e^{-\beta (f_1 + \cdots + f_n)} = C^n e^{-\beta (g_1 + \cdots + g_n)}. \)
Then: \( (Ce^{-\beta f_1}) \cdots (Ce^{-\beta f_n}) = (Ce^{-\beta g_1}) \cdots (Ce^{-\beta g_n}). \)
Then: \( (\mu\{f_1\}) \cdots (\mu\{f_n\}) = (\mu\{g_1\}) \cdots (\mu\{g_n\}). \)
Then:
\( \mu^n\{f\} = \mu^n\{g\}. \)

End of proof of Claim 3.

By hypothesis, \( E \) is residue-unconstrained and \( \alpha \in \mathbb{Z} \) and \( \varepsilon_0 \in E. \)
Recall: \( \mu = B^E_\beta \in \mathcal{P}_E \) and \( S_\mu = E \) and \( |\mu|_2 < \infty \) and \( M_\mu = \alpha. \)
Let \( P := \mu\{\varepsilon_0\}. \) Then \( P = B^E_\beta\{\varepsilon_0\}. \)
We want: as $n \to \infty$, 
$$\nu_{\Omega_n} \{ f \in \Omega_n \mid f_n = \varepsilon_0 \} \to P.$$ 

By Theorem 11.6, as $n \to \infty$, 
$$(\mathcal{N}(\mu^n|\Omega_n)) \{ f \in \Omega_n \mid f_n = \varepsilon_0 \} \to P.$$ 

So, by Claim 3, as $n \to \infty$, 
$$\nu_{\Omega_n} \{ f \in \Omega_n \mid f_n = \varepsilon_0 \} \to P. \quad \Box$$

We record the $t_n = n\alpha$ version of the preceding theorem:

**THEOREM 16.2.** Let $E \subseteq \mathbb{Z}$ be finite and residue-unconstrained.

Let $\alpha \in \mathbb{Z}$. Assume $\alpha \in (\min E; \max E)$. Let $\beta := \text{BP}_{\alpha}^E$.

For all $n \in \mathbb{N}$, let $\Omega_n := \{ f \in E^n \mid f_1 + \cdots + f_n = n\alpha \}$.

Let $\varepsilon_0 \in E$. Then: as $n \to \infty$, 
$$\nu_{\Omega_n} \{ f \in \Omega_n \mid f_n = \varepsilon_0 \} \to B_{\beta}^E(\varepsilon_0).$$

**Example:** Suppose $E = \{0, 1, 10\}$ and $\alpha = 1$.

Then $\Omega_N = \{ f \in E^N \mid f_1 + \cdots + f_N = N \}$,

so $\Omega_N$ represents the set of all GFA dispensations, as described in §3.

The measure $\nu_{\Omega_N}$ gives equal probability to each dispensation,

so $\nu_{\Omega_N}$ represents the GFA’s first system for awarding grants,

also described in §3.

Since $\beta = \text{BP}_{\alpha}^E = \text{BP}_{\alpha}^{E \{0,1,10\}}$, we calculate: $\beta = (\ln 9)/10$.

More calculation gives: $(B_{\beta}^E \{0\}, B_{\beta}^E \{1\}, B_{\beta}^E \{10\}) = \frac{(1,9^{-1/10},9^{-1})}{1 + 9^{-1/10} + 9^{-1}}$.

Since $N$ is large, by Theorem 16.2, we get:

$$\nu_{\Omega_N} \{ f \in \Omega_N \mid f_N = \varepsilon_0 \} \approx B_{\beta}^E(\varepsilon_0).$$

So, if I am the $N$th professor, then, under the first system,

my probability of receiving $\varepsilon_0$ dollars

is approximately equal to $B_{\beta}^E(\varepsilon_0)$.

Thus Theorem 16.2 reproduces the result of §12.

**Proof.** DELETE THIS PROOF.

Let $\mu := B_{\beta}^E$. Then: $\mu \in \mathcal{P}_E$ and $S_{\mu} = E$.

By assumption, $E$ is finite. Then $S_{\mu}$ is finite.

So, since $\mu \in \mathcal{P}_E \subseteq \mathcal{F}M_E$, we get: $|\mu|_1 < \infty$ and $|\mu|_2 < \infty$.

Since $\beta = \text{BP}_{\alpha}^E = (A_{\alpha}^E)^{-1}(\alpha)$, we get: $(A_{\alpha}^E)(\beta) = \alpha$.

So, since $M_{\mu} = M_{B_{\beta}^E} = A_{\mu}^E = (A_{\alpha}^E)(\beta)$, we get: $M_{\mu} = \alpha$.

For all $n \in \mathbb{N}$, define $\psi_n : \mathbb{Z} \to \mathbb{R}$ by:

$$\forall t \in \mathbb{Z}, \quad \psi_n(t) = \mu^n \{ f \in E^n \mid f_1 + \cdots + f_n = t \}.$$ 

Then: $\forall n \in \mathbb{N}, \quad \psi_n(n\alpha) = \mu^n(\Omega_n)$.

Since $E$ is finite and residue-unconstrained, we get: $2 \leq \#E < \infty$.

Since $\#S_{\mu} = \#E \geq 2$, by Theorem 8.6, we get: $V_{\mu} > 0$.

So, since $V_{\mu} = |\mu|_2 - M_{\mu}^2 \leq |\mu|_2 < \infty$, we conclude:

$$0 < V_{\mu} < \infty.$$
Let $v := V_\mu$. Then $0 < v < \infty$. Let $\tau := 1/\sqrt{2\pi v}$. Then $\tau > 0$.

Claim 1: As $n \to \infty$, $\sqrt{n} \cdot (\psi_n(n\alpha)) \to \tau$.

Proof of Claim 1: By Theorem 9.10, we get:

$$\sqrt{n} \cdot (\mu^n \{ f \in E^n \mid f_1 + \cdots + f_n = n\alpha \}) \to 1/\sqrt{2\pi v}.$$  

Then: as $n \to \infty$, $\sqrt{n} \cdot (\psi_n(n\alpha)) \to \tau$.

End of proof of Claim 1.

Since $\tau > 0$, by Claim 1, choose $n_0 \in \mathbb{N}$ s.t.

$$\forall n \in [n_0, \infty), \quad \sqrt{n} \cdot (\psi_n(n\alpha)) > 0.$$  

Claim 2: Let $n \in [n_0, \infty)$. Then: $\mu^n(\Omega_n) > 0$.

Proof of Claim 2: Recall: $\psi_n(n\alpha) = \mu^n(\Omega_n)$. Want: $\psi_n(n\alpha) > 0$.

By the choice of $n_0$, we get: $\sqrt{n} \cdot (\psi_n(n\alpha)) > 0$. Then: $\psi_n(n\alpha) > 0$.

End of proof of Claim 2.

Recall: $\mu \in \mathcal{P}_E$.

Then: $\forall n \in \mathbb{N}, \mu^n \in \mathcal{P}_{E^n}$, so $\mu^n(\Omega_n) \leq 1$.

So, by Claim 2,

$$\forall n \in [n_0, \infty), \quad 0 < \mu^n(\Omega_n) \leq 1.$$  

Also, we have:

$$\forall n \in \mathbb{N}, \quad (\mu^n|\Omega_n)(\Omega_n) = \mu^n(\Omega_n).$$  

Then:

$$\forall n \in [n_0, \infty), \quad 0 < (\mu^n|\Omega_n)(\Omega_n) \leq 1.$$  

Then:

$$\forall n \in [n_0, \infty), \quad \mu^n|\Omega_n \in \mathcal{F}_{\Omega_n}.$$  

Then:

$$\forall n \in [n_0, \infty), \quad \mathcal{N}(\mu^n|\Omega_n) \in \mathcal{P}_{\Omega_n}.$$  

Claim 3: Let $n \in [n_0, \infty)$. Then: $\mathcal{N}(\mu^n|\Omega_n) = \nu_{\Omega_n}$.

Proof of Claim 3: Let $\theta := \mathcal{N}(\mu^n|\Omega_n), \ F := \Omega_n$. Then $\theta \in \mathcal{P}_F$.

Want: $\theta = \nu_F$. By Theorem 8.9, given $f, g \in F$, want: $\theta\{f\} = \theta\{g\}$.

By Claim 2, we have: $\mu^n(\Omega_n) > 0$.

Since $(\mu^n|\Omega_n)(\Omega_n) = \mu^n(\Omega_n)$ and $\theta = \mathcal{N}(\mu^n|\Omega_n)$, we get: $\theta = \frac{\mu^n|\Omega_n}{\mu^n(\Omega_n)}$.

Want: $\frac{(\mu^n|\Omega_n)\{f\}}{\mu^n(\Omega_n)} = \frac{(\mu^n|\Omega_n)\{g\}}{\mu^n(\Omega_n)}$.

Want: $(\mu^n|\Omega_n)\{f\} = (\mu^n|\Omega_n)\{g\}$.

Since $f, g \in F = \Omega_n$, we get:

$(\mu^n|\Omega_n)\{f\} = \mu^n\{f\}$ and $(\mu^n|\Omega_n)\{g\} = \mu^n\{g\}$.

Want: $\mu^n\{f\} = \mu^n\{g\}$.

Since $\#E \geq 2$, we get: $E \neq \emptyset$. Then $\hat{B}_\beta^E(E) > 0$. 

Let $C := 1/(\hat{B}_E^\beta(E))$. Then $\mathcal{N}(\hat{B}_E^\beta) = C \cdot \hat{B}_E^\beta$

By definition of $\hat{B}_E^\beta$, we have: $\forall \varepsilon \in E$, $\hat{B}_E^\beta\{\varepsilon\} = e^{-\beta \varepsilon}$.

So, since $\mu = B_\beta^\varepsilon = \mathcal{N}(\hat{B}_E^\beta) = C \cdot \hat{B}_E^\beta$, we get: $\forall \varepsilon \in E$, $\mu\{\varepsilon\} = Ce^{-\beta \varepsilon}$.

Since $f \in F = \Omega_n$, by definition of $\Omega_n$, we get: $f_1 + \cdots + f_n = n\alpha$.

Since $g \in F = \Omega_n$, by definition of $\Omega_n$, we get: $g_1 + \cdots + g_n = \alpha$.

Since $f_1 + \cdots + f_n = n\alpha = g_1 + \cdots + g_n$,
we get: $C^n e^{-\beta (f_1 + \cdots + f_n)} = C^n e^{-\beta (g_1 + \cdots + g_n)}$.

Then: $\mu\{f\} \cdots (\mu\{f_n\}) = (\mu\{g_1\}) \cdots (\mu\{g_n\})$.

Then: $\mu^n\{f\} = \mu^n\{g\}$.

End of proof of Claim 3.

By hypothesis, $E$ is residue-unconstrained and $\alpha \in \mathbb{Z}$ and $\varepsilon_0 \in E$.
Recall: $\mu = B_\beta^\varepsilon \in \mathcal{P}_E$ and $S_\mu = E$ and $|\mu|_2 < \infty$ and $M_\mu = \alpha$.

Let $P := \mu\{\varepsilon_0\}$, Then $P = B_\beta^\varepsilon\{\varepsilon_0\}$,

We want: as $n \to \infty$, $\nu_{\Omega_n}\{f \in \Omega_n | f_n = \varepsilon_0\} \to P$.

By Theorem 11.6, as $n \to \infty$, $(\mathcal{N}(\mu^n|\Omega_n))\{f \in \Omega_n | f_n = \varepsilon_0\} \to P$.

So, by Claim 3, as $n \to \infty$, $\nu_{\Omega_n}\{f \in \Omega_n | f_n = \varepsilon_0\} \to P$. \hfill $\square$

**THEOREM 16.3.** Let $E \subseteq \mathbb{Z}$ be finite and residue-constrained.

Let $\alpha \in \mathbb{Q}$. Assume $\alpha \in (\min E; \max E)$. Let $\beta := B_\alpha^E$.

Let $\delta := \min\{d \in \mathbb{N} | d\alpha \in \mathbb{Z}\}$.

Let $u_1, u_2, \ldots \in \delta \mathbb{Z}$. Assume: $\{u_m - m\delta \alpha | m \in \mathbb{N}\}$ is bounded.

For all $n \in \mathbb{N}$, let $\Omega_n := \{f \in E^n | f_1 + \cdots + f_n = t_n\}$.

Let $\varepsilon_0 \in E$. Then: as $m \to \infty$, $\nu_{\Omega_{\delta m}}\{f \in \Omega_{\delta m} | f_{\delta m} = \varepsilon_0\} \to B_\beta^E\{\varepsilon_0\}$.

**Proof.** For all $n \in \mathbb{N}$, let $t_n := u_{[n]/\delta}$. Apply Theorem 16.1. MORE LATER \hfill $\square$

17. Finite sets with 0

Our goal, Theorem 17.5 below, is a generalization of Theorem 16.2 that applies to any finite subset $E$ of $\mathbb{Z}$ s.t. $0 \in E$.

We begin with some examples that use Theorem 17.5.

Up to now, we have only described the machinations of the GFA on our Earth, Earth-1218.

Recall: On Earth-1218, the GFA gives, to the N professors, grants of $0, \$1, \$10, with an average grant of $1.$
Example: On Earth-googol-plex, there are $N$ professors, and grants are $0, 4.40, 44$, with average grant $3.30$.

Assume that $N$ is large and divisible by $4$.

Under the GFA’s first system of dispensation, for any professor, what is the approximate probability of receiving $0? 4.40? 44$?

To simplify this problem, we can change monetary units to dimes (worth $0.10$ each).

Grants, in dimes, are $0, 44, 440$, with average grant $33$.

Now the grants and average grant are all integers.

Let $E := \{0, 44, 440\}$. Let $\alpha := 33$.

Since $\gcd(E \cup \{\alpha\}) = 11$, we can simplify the problem, by changing monetary units again, from dimes to “FF-dimes”, each worth $44$ dimes.

Grants, in hendeca-dimes, are $0, 1, 10$, with average grant $3/4$.

Let $E' := \{0, 1, 10\}$. Let $\alpha' := 3/4$.

Recall that $N$ is large and divisible by $4$.

Let $M := N/4$. Then $M$ is a large positive integer.

Let $\beta := \text{BP}_{\alpha'}^{E'}$ and let $g := \gcd(E')$. We calculate: $g = 4$.

For all $n \in \mathbb{N}$, let $\Omega_n := \{f' \in (E')^n \mid f'_1 + \cdots + f'_n = n\alpha'\}$.

Then $\Omega_N$ represents the set of dispensations on Earth-googol-plex.

According to Theorem 16.3, we get:

$$\forall \varepsilon_0 \in E', \nu_{\Omega_M} \{f \in \Omega_M \mid f = \varepsilon_0\} \approx B_{\beta}^{E'}\{\varepsilon_0\}.$$  

So, since $gM = 4 \cdot (N/4) = N$, we get:

$$\forall \varepsilon_0 \in E', \nu_{\Omega_N} \{f \in \Omega_N \mid f = \varepsilon_0\} \approx B_{\beta}^{E'}\{\varepsilon_0\}.$$  

Then the approximate answers are:

$$B_{\beta}^{E'}\{0\}, B_{\beta}^{E'}\{1\}, B_{\beta}^{E'}\{10\}.$$  

Computation yields: $\beta \approx ?, ?, ?$ and

$$(B_{\beta}^{E'}\{0\}, B_{\beta}^{E'}\{1\}, B_{\beta}^{E'}\{10\}) \approx (0.5506, 0.4160, 0.0334?),$$

all accurate to four decimals.

(Once again, thanks to C. Prouty for these calculations. See §25.)

Generally, supposing that $E_0$ is a set of grant awards and $\alpha_0$ is an average award, and supposing that $E_0 \subseteq \mathbb{Q}$ and $\alpha_0 \in \mathbb{Q}$, by a change of monetary units, we arrive at a new set $E$ of grant awards and a new average award $\alpha$ such that: $E \subseteq \mathbb{Z}$ and $\alpha \in \mathbb{Z}$. Let $\delta := \gcd(E)$.  

Multiplying our monetary unit by $\delta$, we arrive at: a new unit of money and a new set $E' = E/\delta$ of grant awards and a new average award $\alpha' = \alpha/\delta$

s.t. $E' \subseteq \mathbb{Z}$ is residue-unconstrained and $\alpha' \in \mathbb{Q}$.

Let $g' := \gcd(E')$.

If the number $N$ of professors is large and divisible by $g'$, then Theorem 16.3 gives the approximate probability of each grant award.

What happens if $N$ is NOT divisible by $g'$? Here is an example:

**Example:** Suppose Earth-googol-plex mints one new professor. The new number, $N + 1$, of professors, is NOT divisible by 4.

Grants, in hendeca-dimes, are still 0, 4, 40, with average grant 3.

Now the money simply cannot be dispensed, because:

- in hendeca-dimes, the total of any dispensation is divisible by 4,
  - while $3 \cdot (N + 1)$ is NOT divisible by 4.

The bureaucracy seizes up, there is pandemonium in the streets, and the military steps in to impose order.

Generally, whenever $N$ is not divisible by $g'$,

it is impossible to distribute the money, and chaos ensues.

This completes our analysis of award probabilities, assuming:

- every grant award is rational and the average grant is rational.

If every grant award is rational, but the average award is NOT rational, then the money cannot be distributed, leading to a military dictatorship.

Finally, we briefly discuss the case where NOT every grant award is in $\mathbb{Q}$.

This is somewhat complicated, and only show: It is possible that the award probabilities may not follow a Boltzmann distribution:

**Example:** On Earth-aleph-1, there are $N$ professors, and the GFA
gives,

grants of 0, $\sqrt{2}$, $\sqrt{3}$, $10 - \sqrt{2} - \sqrt{3}$ dollars,

with an average grant of 1 dollar.

Assume that $N$ is divisible by 10.

Let $M := N/10$. Then $M \in \mathbb{N}$ and there are $10M$ professors.

Moreover, there are $10M$ dollars to dispense among them.

In this situation, by linear independence of $\sqrt{2}$ and $\sqrt{3}$ over $\mathbb{Q}$,

every dispensation of awards has

- $7M$ grants of 0 dollars,
- $M$ grants of $\sqrt{2}$ dollars,
- $M$ grants of $\sqrt{3}$ dollars,
- $M$ grants of $10 - \sqrt{2} - \sqrt{3}$ dollars.

If I am one of the $10M$ professors, then I would hope to be among the lucky $M$ receiving $10 - \sqrt{2} - \sqrt{3}$ dollars.

My probability of being so lucky is: $M/(10M)$, i.e., 10%.

Extending this reasoning, we obtain:

- 70% for 0 dollars,
- 10% for $\sqrt{2}$ dollars,
- 10% for $\sqrt{3}$ dollars,
- 10% for $10 - \sqrt{2} - \sqrt{3}$ dollars.

In a Boltzmann distribution, depending on whether $\beta = 0$ or $\beta \neq 0$,
either the probabilities are all equal

or the probabilities are all distinct from one another.

Thus, the 70-10-10-10 distribution above is NOT Boltzmann.

We will need the following two basic results.

**THEOREM 17.1.** Let $U$ be a set and let $u_1, u_2, \ldots \in U$.

Let $g \in \mathbb{N}$. Then: $\exists t_1, t_2, \ldots \in U$ s.t., $\forall m \in \mathbb{N}, \ t_{gm} = u_m$.

A proof might begin: $\forall n \in \mathbb{N}, \ \text{let } t_n := \begin{cases} u_{n/g}, & \text{if } n \in g\mathbb{Z} \\ u_1, & \text{if } n \notin g\mathbb{Z}. \end{cases}$

**THEOREM 17.2.** Let $U$ be a set and let $u_1, u_2, \ldots \in U$.

Let $g \in \mathbb{N}$. Then: $\exists t_1, t_2, \ldots \in U$ s.t., $\forall m \in [2, \infty), \ t_{gm} = u_m$.

A proof might begin: $\forall n \in \mathbb{N}, \ \text{let } t_n := \begin{cases} u_{(n+1)/g}, & \text{if } n + 1 \in g\mathbb{Z} \\ u_1, & \text{if } n + 1 \notin g\mathbb{Z}. \end{cases}$

Recall (§8) the definitions of: $|\mu|_\rho$, $\mu^n$, $S_\mu$, $M_\mu$, $V_\mu$. 

We will also need a variant of the Discrete Local Limit Theorem:

**THEOREM 17.3.** Let $E \subseteq \mathbb{Z}$.

Assume: $0 \in E$ and $\#E \geq 2$. Let $g := \gcd(E)$.

Let $X_1, X_2, \ldots$ be a $\mathbb{Z}$-valued iid sequence of $L^2$ random-variables.

Assume: $\forall n \in \mathbb{N}, \{ t \in \mathbb{Z} \mid \Pr[X_n = t] > 0 \} = E$.

Let $\alpha, \nu \in \mathbb{R}$. Assume: $\forall n \in \mathbb{N}, \ E[X_n] = \alpha$ and $\text{Var}[X_n] = \nu$.

Let $t_1, t_2, \ldots \in g\mathbb{Z}$. Then, as $n \to \infty$,

$$\sqrt{n} \cdot \left[ \left( \Pr[X_1 + \cdots + X_n = t_n] \right) - g \cdot (\Phi_{\nu \alpha}^E(t_n)) \right] \to 0.$$  

*Proof.* Since $\#E \geq 2$, we get $E \neq \{0\}$, and so $\gcd(E) \in \mathbb{N}$.

So, since $g = \gcd(E)$, we get: $g \in \mathbb{N}$.

Also, since $g = \gcd(E)$, we get: $E \subseteq g\mathbb{Z}$.

Let $E' := E/g$, $\alpha' := \alpha/g$, $\nu' = \nu/g^2$.

Then: $E' \subseteq \mathbb{Z}$ and $E'$ is residue-unconstrained.

For all $n \in \mathbb{N}$, let $X'_n := X_n/g$, $t'_n := t_n/g$.

Then: $\forall n \in \mathbb{N}, \{ t' \in \mathbb{Z} \mid \Pr[X'_n = t'] > 0 \} = E'$.

Also, $\forall n \in \mathbb{N}, \ E[X'_n] = \alpha'$ and $\text{Var}[X'_n] = \nu'$ and $t'_n \in \mathbb{Z}$.

By Theorem 9.3, we get: as $n \to \infty$,

$$\sqrt{n} \cdot \left[ \left( \Pr[X'_1 + \cdots + X'_n = t'_n] \right) - (\Phi_{\nu' \alpha'}^E(t'_n)) \right] \to 0.$$  

For all $n \in \mathbb{N}$, we have: $(X'_1 + \cdots + X'_n = t'_n) \iff (X_1 + \cdots + X_n = t_n)$.

For all $n \in \mathbb{N}$, we have: $\Phi_{\nu \alpha}^E(t'_n) = g \cdot (\Phi_{\nu' \alpha'}^E(t'_n))$.

Then: as $n \to \infty$,

$$\sqrt{n} \cdot \left[ \left( \Pr[X_1 + \cdots + X_n = t_n] \right) - g \cdot (\Phi_{\nu \alpha}^E(t_n)) \right] \to 0. \quad \square$$

We record a measure-theoretic version of Theorem 17.3:

**THEOREM 17.4.** Let $E \subseteq \mathbb{Z}$, $\mu \in \mathcal{P}_E$. Assume: $S_\mu = E$.

Assume: $0 \in E$ and $\#E \geq 2$. Let $g := \gcd(E)$.

Assume: $|\mu| \leq 2 < \infty$. Let $\alpha := \max E$, $\nu := V_\mu$.

Let $t_1, t_2, \ldots \in g\mathbb{Z}$. Then, as $n \to \infty$,

$$\sqrt{n} \cdot \left[ (\mu^n\{ f \in E^n \mid f_1 + \cdots + f_n = t_n \}) - g \cdot (\Phi_{\nu \alpha}^E(t_n)) \right] \to 0.$$  

**THEOREM 17.5.** Let $E \subseteq \mathbb{Z}$ be finite. Assume $0 \in E$.

Let $\alpha \in \mathbb{Z}$. Assume: $\alpha \in (\min E; \max E)$. Let $\beta := \text{BP}_\alpha^E$.

For all $n \in \mathbb{N}$, let $\Omega_n := \{ f \in E^n \mid f_1 + \cdots + f_n = n\alpha \}$.

Let $g := \gcd(E)$. Let $\varepsilon_0 \in E$.

Then: as $m \to \infty$, $\nu_{\alpha \varepsilon_0} \{ f \in \Omega_{gm} \mid f_{gm} = \varepsilon_0 \} \to B_{\beta}^E(\varepsilon_0)$.

Recall: $\nu_0(\emptyset) = -1$.

So, since $B_{\beta}^E(\varepsilon_0) > 0$, part of the content of Theorem 17.5 is:
∀ sufficiently large \( m \in \mathbb{N} \), \( \Omega_{gm} \neq \emptyset \).

See Claim 2 in the proof below.

**Proof.** Let \( \mu := B^E_\beta \). Then: \( \mu \in \mathcal{P}_E \) and \( S_\mu = E \).

Since \( E \) is finite and since \( \mu \in \mathcal{P}_E \subseteq \mathcal{F}_E \), we get: \( |\mu|_1 < \infty \) and \( |\mu|_2 < \infty \).

Since \( \beta = B^E_\alpha = (A^E_\alpha)^{-1}(\alpha) \), we get: \( (A^E_\alpha)(\beta) = \alpha \).

So, since \( M_\mu = M_{B^E_\beta} = A^E_\beta = (A^E_\alpha)(\beta) \), we get: \( M_\mu = \alpha \).

Since \( \alpha \in (\min E; \max E) \), we get \( (\min E; \max E) \neq \emptyset \), so \( \min E \neq \max E \), so \#\( E \geq 2 \). Then \( E \neq \{0\} \).

Then \( \gcd(E) \in \mathbb{N} \), so, since \( g = \gcd(E) \), we get: \( g \in \mathbb{N} \).

For all \( n \in \mathbb{N} \), define \( \psi_n : \mathbb{Z} \to \mathbb{R} \) by:

\[
\forall t \in \mathbb{Z}, \quad \psi_n(t) = \mu^n \{ f \in E^n | f_1 + \cdots + f_n = t \}.
\]

Then: \( \forall n \in \mathbb{N} \), \( \psi_n(n\alpha) = \mu^n(\Omega_n) \).

Then: \( \forall m \in \mathbb{N} \), \( \psi_{gm}(g\alpha\mu) = \mu^{gm}(\Omega_{gm}) \).

Since \#\( S_\mu = \# E \geq 2 \), by Theorem 8.6, we get: \( V_\mu > 0 \).

Let \( v := V_\mu \). Then: \( v > 0 \).

So, since \( g \in \mathbb{N} \), we get: \( g/\sqrt{2\pi v} > 0 \).

Let \( \tau := g/\sqrt{2\pi v} \). Then: \( \tau > 0 \).

For all \( n \in \mathbb{N} \), let \( \phi_n := g \cdot \Phi_{n\alpha}^{nv} \).

Then: \( \forall n \in \mathbb{N}, \forall t \in \mathbb{R}, \quad \phi_n(t) = g \cdot \frac{\exp(-(t-n\alpha)^2/(2nv))}{\sqrt{2\pi nv}} \).

Then: \( \forall n \in \mathbb{N} \), \( \phi_n(n\alpha) = g/\sqrt{2\pi n} \).

Then: \( \forall n \in \mathbb{N} \), \( \sqrt{n} \cdot (\phi_n(n\alpha)) = \tau \).

Then: \( \forall m \in \mathbb{N} \), \( \sqrt{gm} \cdot (\phi_{gm}(g\alpha\mu)) = \tau \).

**Claim 1:** As \( m \to \infty \), \( \sqrt{gm} \cdot (\phi_{gm}(g\alpha\mu)) \to \tau \).

**Proof of Claim 1:** By Theorem 17.1, \( \forall t_1, t_2, \ldots \in g\mathbb{Z} \),

as \( n \to \infty \), \( \sqrt{n} \cdot [ (\psi_n(t_n) - (\phi_n(t_n))] \to 0 \).

Then: \( \forall t_1, t_2, \ldots \in g\mathbb{Z} \),

as \( m \to \infty \), \( \sqrt{gm} \cdot [ (\psi_{gm}(t_{gm}) - (\phi_{gm}(t_{gm}))] \to 0 \).

So, by Theorem 17.1, \( \forall u_1, u_2, \ldots \in g\mathbb{Z} \),

as \( m \to \infty \), \( \sqrt{gm} \cdot [ (\psi_{gm}(u_m) - (\phi_{gm}(u_m))] \to 0 \).

Then, as \( m \to \infty \), \( \sqrt{gm} \cdot [ (\psi_{gm}(g\alpha\mu)) - (\phi_{gm}(g\alpha\mu))] \to 0 \).

So, since \( \forall m \in \mathbb{N} \), \( \sqrt{gm} \cdot (\phi_{gm}(g\alpha\mu)) = \tau \),

we get: as \( m \to \infty \), \( \sqrt{gm} \cdot (\psi_{gm}(g\alpha\mu)) \to \tau \).

**End of proof of Claim 1.**
Since $\tau > 0$, by Claim 1, choose $m_0 \in \mathbb{N}$ s.t.
\[
\forall m \in [m_0, \infty), \quad \sqrt{gm} \cdot (\psi_{gm}(g\alpha)) > 0.
\]

**Claim 2:** Let $m \in [m_0, \infty)$. Then: $\mu^m(\Omega_{gm}) > 0$.

**Proof of Claim 2:** Since $\psi_{gm}(g\alpha) = \mu^m(\Omega_{gm})$, want: $\psi_{gm}(g\alpha) > 0$.

By the choice of $m_0$, $\sqrt{gm} \cdot (\psi_{gm}(g\alpha)) > 0$. Then: $\psi_{gm}(g\alpha) > 0$.

End of proof of Claim 2.

Since $\mu \in \mathcal{P}_E$, we get: $\forall n \in \mathbb{N}, \mu^n \in \mathcal{P}_{E^n}$, so $\mu^n(\Omega_n) \leq 1$.

Then: $\forall m \in \mathbb{N}$, $\mu^m(\Omega_{gm}) \leq 1$.

So, by Claim 2, $\forall m \in [m_0, \infty)$, $0 < \mu^m(\Omega_{gm}) \leq 1$.

Also: $\forall m \in \mathbb{N}$, $\mu^m(\Omega_{gm}) = \mu^m(\Omega_{gm})$.

Then: $\forall m \in [m_0, \infty)$, $0 < \mu^m(\Omega_{gm}) \leq 1$.

Then: $\forall m \in [m_0, \infty)$, $\mu^m(\Omega_{gm}) \in \mathcal{F} \mathcal{M}^\mathcal{g}_m$.

Then: $\forall m \in [m_0, \infty)$, $\mathcal{N}(\mu^m|\Omega_{gm}) \in \mathcal{P}_{\Omega_{gm}}$.

**Claim 3:** Let $m \in [m_0, \infty)$. Then: $\mathcal{N}(\mu^m|\Omega_{gm}) = \nu_{\Omega_{gm}}$.

**Proof of Claim 3:** Let $n := gm$. Want: $\mathcal{N}(\mu^n|\Omega_n) = \nu_{\Omega_n}$.

Let $\theta := \mathcal{N}(\mu^n|\Omega_n)$. Then $\theta \in \mathcal{P}_{\Omega_n}$. Let $F := \Omega_n$. Then $\theta \in \mathcal{P}_F$.

Want: $\theta = \nu_F$. By Theorem 8.9, given $f, h \in F$, want: $\theta\{f\} = \theta\{h\}$.

By Claim 2, we have: $\mu^m(\Omega_{gm}) > 0$. Then $\mu^n(\Omega_n) > 0$.

Since $(\mu^n|\Omega_n)(\Omega_n) = \mu^m(\Omega_{gm})$ and $\theta = \mathcal{N}(\mu^n|\Omega)$, we get: $\theta = \mu^n(\Omega_n) / \mu^m(\Omega_{gm})$.

Want: $\frac{(\mu^n|\Omega_n)f}{\mu^n(\Omega_n)} = \frac{(\mu^n|\Omega_n)h}{\mu^n(\Omega_n)}$.

Want: $\mu^n\{f\} = \mu^n\{h\}$.

Since $f, h \in F = \Omega_n$, we get:

$$(\mu^n|\Omega_n)f = \mu^n\{f\} \quad \text{and} \quad (\mu^n|\Omega_n)h = \mu^n\{h\}.$$ 

Want: $\mu^n\{f\} = \mu^n\{h\}$.

Since $\#E \geq 2$, we get: $E \neq \emptyset$. Then $\hat{B}^E_\beta(E) > 0$.

By definition of $\hat{B}^E_\beta$, we have: $\forall \varepsilon \in E, \hat{B}^E_\beta(\varepsilon) = e^{-\beta\varepsilon}$.

Let $C := 1/(\hat{B}^E_\beta(E))$. Then $\mathcal{N}(\hat{B}^E_\beta) = C \cdot \hat{B}^E_\beta$

So, since $\mu = B^E_\beta = \mathcal{N}(\hat{B}^E_\beta) = C \cdot \hat{B}^E_\beta$,

we get: $\forall \varepsilon \in E, \mu\{\varepsilon\} = C e^{-\beta\varepsilon}$.

Since $f \in F = \Omega_n$, by definition of $\Omega_n$, we get: $f_1 + \cdots + f_n = n\alpha$.

Since $h \in F = \Omega_n$, by definition of $\Omega_n$, we get: $h_1 + \cdots + h_n = n\alpha$.

Since $f_1 + \cdots + f_n = n\alpha = h_1 + \cdots + h_n$,

we get: $C^n e^{-\beta(f_1 + \cdots + f_n)} = C^n e^{-\beta(h_1 + \cdots + h_n)}$. 
Then: $(Ce^{-β}f_1)\cdots(Ce^{-β}f_n) = (Ce^{-β}h_1)\cdots(Ce^{-β}h_n)$.

Then: $(\mu\{f_1\})\cdots(\mu\{f_n\}) = (\mu\{h_1\})\cdots(\mu\{h_n\})$.

Then: $\mu^\prime\{f\} = \mu^\prime\{h\}$.

End of proof of Claim 3.

We have: $\forall n \in \mathbb{N}, \forall S \subseteq \Omega_n,$

$$(\mu^\prime\{\Omega_n\})(S) = \mu^\prime(S).$$

Then: $\forall m \in \mathbb{N}, \forall S \subseteq \Omega_{gm},$

$$(\mu^\prime|m\{\Omega_{gm}\})(S) = \mu^\prime|gm(S).$$

Then, by definition of normalization, we get: $\forall m \in [m_0, \infty), \forall S \subseteq \Omega_{gm},$

$$\left(N(\mu^\prime|gm\{\Omega_{gm}\})(S) = \frac{\mu^\prime|gm(S)}{\mu^\prime|gm(\Omega_{gm})}, \right.$$ 

and so, by Claim 3,

$$\nu_{\Omega_{gm}}(S) = \frac{\mu^\prime|gm(S)}{\mu^\prime|gm(\Omega_{gm})}.$$ 

We therefore want: As $m \to \infty, \frac{\mu^\prime|gm\{f \in \Omega_{gm} | f_{gm} = \varepsilon_0\}}{\mu^\prime|gm(\Omega_{gm})} \to B^E(\varepsilon_0).$

Recall: both $\forall m \in \mathbb{N}, \psi_{gm}(gm\alpha) = \mu^\prime|gm(\Omega_{gm})$ and $\mu = B^E.$

We wish to show: As $m \to \infty, \frac{\mu^\prime|gm\{f \in \Omega_{gm} | f_{gm} = \varepsilon_0\}}{\psi_{gm}(gm\alpha)} \to \mu(\varepsilon_0).$

By hypothesis, $\varepsilon_0 \in E$ and $\forall n \in \mathbb{N}, \Omega_n = \{f \in E^n | f_1 + \cdots + f_n = na\}.$

Since $g = \gcd(E)$, we get: $E \subseteq g\mathbb{Z}.$ Recall: $\forall m \in \mathbb{N}, gm\alpha \in g\mathbb{Z}.$

So, since $\varepsilon_0 \in E \subseteq g\mathbb{Z},$ we get: $\forall m \in \mathbb{N}, \quad gm\alpha - \varepsilon_0 \in g\mathbb{Z}.$

For all $n \in \mathbb{N},$ let $s_n := n\alpha - \varepsilon_0.$ Then: $\forall m \in \mathbb{N}, \quad s_{gm} \in g\mathbb{Z}.$

Claim 4: Let $n \in [2, \infty).$

Then: $\mu^\prime\{f \in \Omega_n | f_n = \varepsilon_0\} = (\psi_{n-1}(s_n)) \cdot (\mu(\varepsilon_0))$.

Proof of Claim 4:

Because

$$\{f \in \Omega_n | f_n = \varepsilon_0\} = \{f \in E^n | (f_1 + \cdots + f_{n-1} + f_n = na) \& (f_n = \varepsilon_0)\} = \{f \in E^n | (f_1 + \cdots + f_{n-1} + \varepsilon_0 = na) \& (f_n = \varepsilon_0)\} = \{f \in E^n | (f_1 + \cdots + f_{n-1} = n\alpha - \varepsilon_0) \& (f_n = \varepsilon_0)\} = \{f \in E^n | (f_1 + \cdots + f_{n-1} = s_n) \& (f_n = \varepsilon_0)\},$$

we conclude: under the standard bijection $E^n \leftrightarrow E^{n-1} \times E,$

the subset $\{f \in \Omega_n | f_n = \varepsilon_0\} \subseteq E^n$

corresponds to $\{f \in E^{n-1} | f_1 + \cdots + f_{n-1} = s_n\} \times \{\varepsilon_0\}.$

Then $\mu^\prime\{f \in \Omega_n | f_n = \varepsilon_0\}$

$$= (\mu^\prime\{f \in E^{n-1} | f_1 + \cdots + f_{n-1} = s_n\}) \cdot (\mu(\varepsilon_0)) = \left(\psi_{n-1}(s_n)\right)(\mu(\varepsilon_0)).$$

End of proof of Claim 4.
Since $g \in \mathbb{N}$, we have: $\forall m \in [2, \infty), \; g m \in [2, \infty)$.
Then, by Claim 4, we have: $\forall m \in [2, \infty),
\mu^{gm} \{ f \in \Omega_{gm} | f_{gm} = \varepsilon_0 \} = \frac{\left(\psi_{gm-1}(s_{gm})\right) \cdot \left(\mu\{\varepsilon_0\}\right)}{(\psi_{gm-1}(s_{gm})) \cdot \left(\mu\{\varepsilon_0\}\right)}$.

**Want:** As $m \to \infty$,\[ \frac{\psi_{gm}(g\alpha)}{\psi_{gm}(gm\alpha)} \to \mu\{\varepsilon_0\}. \]

**Want:** As $m \to \infty$,\[ \frac{\psi_{gm-1}(s_{gm})}{\psi_{gm}(gm\alpha)} \to 1. \]

For all $n \in \mathbb{N}$, by definition of $\phi_n$, we get:
\[ \phi_n(s_{n+1}) = g \cdot \exp(- \frac{(s_{n+1} - n\alpha)^2}{2nv}) \cdot \frac{1}{\sqrt{2\pi v}}. \]

Then: $\forall n \in \mathbb{N}$, $\sqrt{n} \cdot [\phi_n(s_{n+1})] = g \cdot \exp(- \frac{(s_{n+1} - n\alpha)^2}{2nv}) \cdot \frac{1}{\sqrt{2\pi v}}.$

For all $n \in \mathbb{N}$, by definition of $s_{n+1}$, we have:
\[ s_{n+1} = n\alpha + \alpha - \varepsilon_0; \]
then:
\[ s_{n+1} - n\alpha = \alpha - \varepsilon_0. \]

We have: as $n \to \infty,
\begin{align*}
-\frac{(\alpha - \varepsilon_0)^2}{2nv} & \to 0. \\
\exp(- \frac{(s_{n+1} - n\alpha)^2}{2nv}) & \to 1.
\end{align*}

Recall: $\tau = g/\sqrt{2\pi v}$.

By Theorem 17.4, $\forall t_1, t_2, \ldots \in \mathbb{Z},$
\[ \lim_{n \to \infty} \sqrt{n} \cdot [\psi_n(t_n) - \phi_n(t_n)] = 0. \]

Then: $\forall t_1, t_2, \ldots \in \mathbb{Z},$
\[ \lim_{m \to \infty} \sqrt{gm - 1} \cdot [\psi_{gm-1}(t_{gm-1}) - \phi_{gm-1}(t_{gm-1})] = 0. \]

So, by Theorem 17.2, $\forall u_1, u_2, \ldots \in \mathbb{Z},$
\[ \lim_{m \to \infty} \sqrt{gm - 1} \cdot [\psi_{gm-1}(u_m) - \phi_{gm-1}(u_m)] = 0. \]

Then, as $m \to \infty$, \[ \sqrt{gm - 1} \cdot [\psi_{gm-1}(s_{gm}) - \phi_{gm-1}(s_{gm})] \to \tau. \]

So, since $m \to \infty$, \[ \sqrt{gm - 1} \cdot (\phi_{gm-1}(s_{gm})) \to \tau, \]
we get: as $m \to \infty$, \[ \sqrt{gm - 1} \cdot (\psi_{gm-1}(s_{gm})) \to \tau. \]

By Claim 1, as $m \to \infty$, \[ \sqrt{gm} \cdot (\psi_{gm}(gm\alpha)) \to \tau. \]

Dividing the last two limits, we get:
\[ \lim_{m \to \infty} \frac{\sqrt{gm - 1} \cdot (\psi_{gm-1}(s_{gm}))}{\sqrt{gm} \cdot (\psi_{gm}(gm\alpha))} = 1. \]
Also, as \( m \to \infty \),
\[
\frac{\sqrt{gm}}{\sqrt{gm-1}} \to 1.
\]

Multiplying the last two limits, we get:
\[
as m \to \infty, \quad \frac{\psi_{gm-1}(s_{gm})}{\psi_{gm}(g\alpha)} \to 1. \quad \Box
\]

18. **General finite sets of rationals**

**Let** \( E \subseteq \mathbb{Q} \). **Assume:** \( \#E < \infty \).

**Let** \( \alpha \in \mathbb{Q} \). **Assume:** \( \alpha \in (\min E; \max E) \).  

For all \( n \in \mathbb{N} \), **let** \( \Omega_n := \{ f \in E^n | f_1 + \cdots + f_n = n\alpha \} \).

**Let** \( \varepsilon_0 \in E \). **In this section,** we compute the asymptotics, as \( n \to \infty \), of \( \nu_{\Omega_n} \{ f \in \Omega_n | f_n = \varepsilon_0 \} \).

**For all** \( n \in \mathbb{N} \), **let** \( \Psi_n := \{ f \in \Omega_n | f_n = \varepsilon_0 \} \).

Then we seek the asymptotics, as \( n \to \infty \), of \( \nu_{\Omega_n}(\Psi_n) \).

Since \( \alpha \in (\min E; \max E) \), we get \( \min E \neq \max E \), so \( \#E \geq 2 \). Then: \( 2 \leq \#E < \infty \).

**Let** \( E^+ := E \cup \{ \alpha \} \). Then \( E^+ \) is finite and \( E^+ \subseteq \mathbb{Q} \).

Then: \( \exists d \in \mathbb{N} \) s.t. \( dE^+ \subseteq \mathbb{Z} \).

**Let** \( \delta := \min \{ d \in \mathbb{N} | dE^+ \subseteq \mathbb{Z} \} \).  

Then \( \delta E^+ \subseteq \mathbb{Z} \). Then \( \delta E \subseteq \mathbb{Z} \) and \( \delta \alpha \in \mathbb{Z} \).  

Also, by minimality of \( \delta \), we get: \( \delta E^+ \) is residue-unconstrained.

So, since \( \delta \varepsilon_0 \in \delta E^+ \), we get: \( \gcd(\delta E^+ - \delta \varepsilon_0) = 1 \).

Since \( \varepsilon_0 \in E \) and \( \delta E \subseteq \mathbb{Z} \), we get: \( \delta \varepsilon_0 \in \mathbb{Z} \).

**Let** \( E' := \delta E - \delta \varepsilon_0 \). Then: \( E' \subseteq \mathbb{Z} \).

**Let** \( \alpha' := \delta \alpha - \delta \varepsilon_0 \). Then: \( \alpha' \in \mathbb{Z} \).

Also, we have: \( \#E' = \#E \). Then: \( 2 \leq \#E' < \infty \).

Since \( \delta \varepsilon_0 - \delta \varepsilon_0 \in \delta E - \delta \varepsilon_0 = E' \), we get: \( 0 \in E' \).

Since \( \#E' \geq 2 \), we get: \( E' \neq \{0\} \).

Then \( \gcd(E') \in \mathbb{N} \). **Let** \( g := \gcd(E') \).

Then: \( g \in \mathbb{N} \). **Let** \( \beta := \text{BP}_{E'} \).

For all \( n \in \mathbb{N} \), **let** \( H_n := \left\{ B_{E'}^n \{ \varepsilon_0 \}, \text{ if } n \in g \mathbb{N} \right\} \)
(\( -1 \), if \( n \notin g \mathbb{N} \)).

In this section, **we prove:** as \( n \to \infty \), \( \nu_{\Omega_n}(\Psi_n) \) is asymptotic to \( H_n \).

Recall: \( 0 \in E' \). **Let** \( \varepsilon'_0 := 0 \). Then: \( \varepsilon'_0 \in E' \).

For all \( n \in \mathbb{N} \), **let** \( \Omega'_n := \{ f' \in (E')^n | f'_1 + \cdots + f'_n = n\alpha' \} \).
For all $n \in \mathbb{N}$, let $\Psi'_n := \{ f' \in \Omega'_n \mid f'_n = \varepsilon'_0 \}$.

The bijection $\varepsilon \leftrightarrow \varepsilon - \varepsilon_0 : E \leftrightarrow E'$ induces, for all $n \in \mathbb{N}$, a bijection $E^n \leftrightarrow (E')^n$; under this bijection, $\Omega_n$ corresponds to $\Omega'_n$, and $\Psi_n$ corresponds to $\Psi'_n$.

Then: $\forall n \in \mathbb{N}$, $\nu_{\Omega_n}(\Psi_n) = \nu_{\Omega'_n}(\Psi'_n)$.

**Want:** as $n \to \infty$, $\nu_{\Omega_n}(\Psi_n)$ is asymptotic to $H_n$.

By Theorem 15.2, we have: $B_{E'}^{E'}(\varepsilon'_0) = B_{E'}^{E'}(\varepsilon_0)$.

Then: $\forall n \in \mathbb{N}$, $H_n = \begin{cases} B_{E'}^{E'}(\varepsilon'_0), & \text{if } n \in g\mathbb{N} \\ -1, & \text{if } n \notin g\mathbb{N}. \end{cases}$

Recall: $\beta = \text{BP}^{E'}_{\alpha'}$.

Then, by Theorem 17.5, we have:

as $m \to \infty$, $\nu_{\Omega_{\text{gm}}}(f' \in \Omega'_{\text{gm}} | f'_m = \varepsilon'_0) \to B_{E'}^{E'}(\varepsilon'_0)$.

Then, as $m \to \infty$, $\nu_{\Omega_{\text{gm}}}(\Psi'_{\text{gm}}) \to B_{E'}^{E'}(\varepsilon'_0)$.

**It therefore suffices to show:** $\forall n \in \mathbb{N}\setminus(g\mathbb{N})$, $\nu_{\Omega_n}(\Psi'_n) = -1$.

Given $n \in \mathbb{N}\setminus(g\mathbb{N})$, want: $\nu_{\Omega_n}(\Psi'_n) = -1$.

Since $\nu_{\varnothing}(\varnothing) = -1$ and since $\Psi'_n \subseteq \Omega'_n$, it suffices to show: $\Omega'_n = \varnothing$.

Recall: $\Omega'_n := \{ f' \in (E')^n \mid f'_1 + \cdots + f'_n = n\alpha' \}$.

**Given** $f' \in (E')^n$, want: $f'_1 + \cdots + f'_n \neq n\alpha'$.

Since $g = \gcd(E')$, we get: $E' \subseteq g\mathbb{Z}$.

Since $f'_1, \ldots, f'_n \in E' \subseteq g\mathbb{Z}$, we get: $f'_1 + \cdots + f'_n \in g\mathbb{Z}$.

**It therefore suffices to show:** $n\alpha' \notin g\mathbb{Z}$.

Since $\gcd(E' \cup \{\alpha'\}) = 1$, we get: $1 = \gcd\{\alpha', g\}$.

Since $g \in \mathbb{N}$, we get: $\mathbb{N}\setminus(g\mathbb{N}) \subseteq \mathbb{Z}\setminus(g\mathbb{Z})$.

Since $n \in \mathbb{N}\setminus(g\mathbb{N}) \subseteq \mathbb{Z}\setminus(g\mathbb{Z})$, we get: $n \notin g\mathbb{Z}$.

So, since $\gcd\{\alpha', g\} = 1$, we get: $n\alpha' \notin g\mathbb{Z}$.

19. Earth-minimum-Mahlo-cardinal and the BUA

We next wish to handle thermodynamic systems in which many states may have a single energy-level.

One says that such an energy-level is “degenerate”.

In this section, we develop a whimsical example.

In the next, we will develop a general theory.

Recall that $N \in \mathbb{N}$ is large.

In a parallel universe, on Earth-minimum-Mahlo-cardinal, the BUA (Best University Anywhere) employs $N$ professors.
Each professor has a number, from 1 to $N$.

Each professor wanders the campus, carrying two bags: one red, one blue.

Each bag is closed from view, but has money in it or is empty.

By BUA rules, the amount of money in any bag is always: $0$ or $1$ or $2$ or $3$ or $4$.

The “state” of a professor is the pair $\sigma = (\sigma_1, \sigma_2)$ such that

$\sigma_1$ is the number of dollars in the professor’s red bag,

$\sigma_2$ is the number of dollars in the professor’s blue bag;

the professor’s “wealth” is $\sigma_1 + \sigma_2$ dollars.

So, if I am one of the professors, and my state is $(3, 2)$,

then I have: $\$3$ in my red bag and $\$2$ in my blue bag,

and my wealth is $\$5$.

By BUA rules, each professor’s wealth is always $\leq \$7$.

Therefore, $(4, 4)$ is not an allowable state.

Let $\Sigma := ([0..4] \times [0..4]) \setminus \{(4, 4)\}$.

Then $\Sigma$ represents the set of all allowable professor-states.

Since $\#([0..4] \times [0..4]) = 5 \cdot 5 = 25$, we get: $\#\Sigma = 24$.

Define $\varepsilon : \Sigma \to [0..7]$ by: $\forall \sigma \in \Sigma$, $\varepsilon(\sigma) = \sigma_1 + \sigma_2$.

For convenience of notation, $\forall \sigma \in \Sigma$, let $\varepsilon_{\sigma} := \varepsilon(\sigma)$.

If I am one of the professors, and if my state is $\sigma = (\sigma_1, \sigma_2) \in \Sigma$,

then I have: $\$\sigma_1$ in my red bag and $\$\sigma_2$ in my blue bag,

and my wealth is $\$\varepsilon_{\sigma}$.

Since $\varepsilon_{(3,2)} = 5 = \varepsilon_{(1,4)}$, we see that $\varepsilon$ is not one-to-one,

and we have a so-called “degeneracy” at 5.

This $\varepsilon$ has many other degeneracies.

Recall: The professors are numbered, from 1 to $N$.

At random moments, random pairs of wandering professors cross paths, and interact.

Each interaction involves three steps:

a game and then

a verbal offer and then

a rejection or a money transfer.

The first step, the game, is played as follows:

one of the two professors flips a fair coin and
if heads, then the lower-numbered professor wins and
        if tails, then the higher-numbered professor wins.

Without touching any money,
        the losing professor verbally offers $1 to the winning professor.

The losing professor then flips a fair coin, and
        if heads, then the loser’s red bag is opened and
        if tails, then the loser’s blue bag is opened.

If the loser’s open bag is empty, then
        then the winner gallantly rejects the $1 offer and
        the opened bags are closed, the interaction is over, and
        the professors continue their wanderings.

On the other hand, if the loser’s open bag is NOT empty, then,
        both of the winner’s bags are opened.

Recall that, by BUA rules, every professor’s wealth must be ≤ $7.

If the winner’s wealth is $7,
        then the winner rejects the $1 offer and
        the opened bags are closed, the interaction is over, and
        the professors continue their wanderings.

On the other hand, if the winner’s wealth is < $7,
        then the winner flips a fair coin, and
        if heads, then the winner’s red bag is closed and
        if tails, then the winner’s blue bag is closed.

At this point, the winner has one open bag, as does the loser.
Also, the loser’s open bag is NOT empty.
Recall that no bag may have more than $4.

If the winner’s open bag has $4,
        then the winner rejects the $1 offer and
        the opened bags are closed, the interaction is over, and
        the professors continue their wanderings.

On the other hand, if the winner’s open bag has ≤ $3,
        then $1 is transferred
                from the losing professor’s open bag
                to the winning professor’s open bag;
        then the opened bags are closed, the interaction is over, and
        the professors continue their wanderings.

Because of these interactions,
        the wealth of an individual professor may change over time,
but the total wealth of all of them is constant; there is “conservation of (total) wealth”.

At the BUA, this total wealth is always $N$ dollars.

A “state-dispensation” is a function $[1..N] \rightarrow \Sigma$, representing the states of all $N$ professors.

So, if, at some point in time, the state-dispensation is $\omega : [1..N] \rightarrow \Sigma$, then, for every $\ell \in [1..N]$, the state of Professor #\(\ell\) is $\omega(\ell)$, and the wealth of Professor #\(\ell\) is $\varepsilon_{\omega(\ell)}$;

therefore, the total wealth of all the professors is $\sum_{\ell=1}^{N} \varepsilon_{\omega(\ell)}$.

As we mentioned, at the BUA, that total wealth is $N$.

Let $[\Omega^*] := \{ \omega : [1..N] \rightarrow \Sigma \mid \sum_{\ell=1}^{N} \varepsilon_{\omega(\ell)} = N \}$.

Then $\Omega^*$ represents the set of all state-dispensations at the BUA.

The random interactions, described above, induce a discrete Markov-chain on $\Omega^*$.

This, in turn, induces a map $\Pi : \mathcal{P}_{\Omega^*} \rightarrow \mathcal{P}_{\Omega^*}$.

Let $T := \#\Omega^*$. Fix an ordering of $\Omega^*$, i.e., a bijection $[1..T] \leftrightarrow \Omega^*$.

The Markov-chain then has a transition-matrix $\Phi \in [0;1]^{T \times T}$, whose column-sums are all 1.

Recall that the probability-distribution $\nu_{\Omega^*} \in \mathcal{P}_{\Omega^*}$ assigns equal probability to each state-dispensation in $\Omega^*$.

That is, $\forall \omega \in \Omega^*, \nu_{\Omega^*}\{\omega\} = 1/\#\Omega^*$.

For every $\phi, \psi \in \Omega^*$, the probability of transitioning from $\phi$ to $\psi$ is equal to the probability of transitioning from $\psi$ to $\phi$.

That is, the transition-matrix $\Phi$ is symmetric.

So, since the column-sums of $\Phi$ are all 1, we get: the row-sums of $\Phi$ are all 1. Then: $\Pi(\nu_{\Omega^*}) = \nu_{\Omega^*}$.

That is, $\nu_{\Omega^*}$ is an equilibrium-distribution for the Markov-chain.

We will say that two state-dispensations $\phi, \psi \in \Omega^*$ are “adjacent”, if there is an interaction that carries $\phi$ to $\psi$.

We have: $\forall \phi, \psi \in \Omega^*, \exists m \in \mathbb{N}, \exists \omega_0, \ldots, \omega_m \in \Omega^*$

\hspace{1cm} s.t. $\phi = \omega_0$ and $\omega_m = \psi$

\hspace{1cm} and s.t. $\forall i \in [1..m]$, $\omega_{i-1}$ is adjacent to $\omega_i$. 

That is, any two state-dispensations are connected by an adjacency-path.
That is, the Markov-chain is irreducible. Recalling that some interactions result in a rejection, such interactions do not change the state-dispensation. So, a state-dispensation is sometimes adjacent to itself. That is, there are adjacency-cycles of length 1. It follows that the Markov-chain is aperiodic. So, since the Markov-chain is irreducible and since \( \Pi(\nu_{\Omega^*}) = \nu_{\Omega^*} \), by Perron-Frobenius, we get:

\[
\forall \mu \in \mathcal{P}_{\Omega^*}, \quad \mu, \Pi(\mu), \Pi(\Pi(\mu)), \Pi(\Pi(\Pi(\mu))), \ldots \rightarrow \nu_{\Omega^*}.
\]

That is, for any starting probability-distribution on \( \Omega^* \), after enough random interactions, the resulting probability-distribution on \( \Omega^* \) will be approximately equal to \( \nu_{\Omega^*} \), to any desired level of accuracy.

**Problem:** Suppose I am Professor \#N at the BUA. Suppose that the probability-distribution of state-dispensations is approximately equal to \( \nu_{\Omega^*} \).
For each \( \sigma \in \Sigma \), compute my probability of being in state \( \sigma \).
That is, \( \forall \sigma \in \Sigma \), compute \( \mu(\omega \in \Omega^* | \omega(N) = \sigma) \).
Since \( \#\Sigma = 24 \), there will be 24 answers. Approximate answers are acceptable.

To make a precise mathematical problem, we, in fact, assume the probability-distribution of state-dispensations is *exactly* equal to \( \nu_{\Omega^*} \), and we seek the exact “thermodynamic limit”, meaning:
we replace \( N \) with a variable \( n \in \mathbb{N} \), and let \( n \to \infty \).

In the next two sections, we will develop a theory to solve problems like this one.
We need only adapt our earlier methods to allow for degeneracies. Our main theorem is Theorem 21.2, and the solution to the above problem appears right before its proof.
20. Boltzmann distributions on finite sets with degeneracy

We begin by adapting our work on Boltzmann distributions to allow for degeneracies.

**DEFINITION 20.1.** Let $\Sigma$ be a nonempty finite set.

Let $\varepsilon : \Sigma \to \mathbb{R}$. Let $\beta \in \mathbb{R}$.

Then $\hat{B}_\beta^\varepsilon \in \mathcal{F}\mathcal{M}_\Sigma$ is defined by: $\forall \sigma \in \Sigma, \ \hat{B}_\beta^\varepsilon(\sigma) = e^{-\beta(\varepsilon(\sigma))}$.

Also, we define: $\hat{B}_\beta^\varepsilon := \mathcal{N}(\hat{B}_\beta^\varepsilon) \in \mathcal{P}_\Sigma$.

Then: $\forall \varepsilon \in \mathbb{E}$, we have: $B_\beta^E(\varepsilon) = (\hat{B}_\beta^E(\varepsilon)) / (\hat{B}_\beta^E(E))$.

We have: $\forall$ nonempty finite set $\Sigma$, $\forall \varepsilon : \Sigma \to \mathbb{R}$, $\forall \beta \in \mathbb{R}$, $S_{\hat{B}_\beta^\varepsilon} = \Sigma = S_{B_\beta^E}$.

**Example:** Let $\Sigma := \{0, 1, 10\}$ and let $\beta \in \mathbb{R}$.

Define $\varepsilon : \Sigma \to \mathbb{R}$ by: $\forall \sigma \in \Sigma, \ \varepsilon(\sigma) = \sigma$.

Then: $\hat{B}_\beta^\varepsilon(0) = 1$, $\hat{B}_\beta^\varepsilon(1) = e^{-\beta}$, $\hat{B}_\beta^\varepsilon(10) = e^{-10\beta}$.

Let $C := 1/(1 + e^{-\beta} + e^{-10\beta})$.

Then: $B_\beta^\varepsilon(0) = C$, $B_\beta^\varepsilon(1) = Ce^{-\beta}$, $B_\beta^\varepsilon(10) = Ce^{-10\beta}$.

**Example:** Let $\Sigma := \{2, 4, 8, 9\}$ and let $\beta \in \mathbb{R}$.

Define $\varepsilon : \Sigma \to \mathbb{R}$ by: $\forall \sigma \in \Sigma, \ \varepsilon(\sigma) = \sigma$.

Then: $\hat{B}_\beta^\varepsilon(2) = e^{-2\beta}$, $\hat{B}_\beta^\varepsilon(4) = e^{-4\beta}$, $\hat{B}_\beta^\varepsilon(8) = e^{-8\beta}$, $\hat{B}_\beta^\varepsilon(9) = e^{-9\beta}$.

Let $C := 1/(e^{-2\beta} + e^{-4\beta} + e^{-8\beta} + e^{-9\beta})$.

Then: $B_\beta^\varepsilon(2) = Ce^{-2\beta}$, $B_\beta^\varepsilon(4) = Ce^{-4\beta}$, $B_\beta^\varepsilon(8) = Ce^{-8\beta}$, $B_\beta^\varepsilon(9) = Ce^{-9\beta}$.

**Example:** Let $\Sigma := \{1, 2, 3, 4\}$ and let $\beta \in \mathbb{R}$.

Define $\varepsilon : \Sigma \to \mathbb{R}$ by: $\varepsilon(1) = 2$, $\varepsilon(2) = 4$, $\varepsilon(3) = 8$, $\varepsilon(4) = 9$.

Then: $\hat{B}_\beta^\varepsilon(1) = e^{-2\beta}$, $\hat{B}_\beta^\varepsilon(2) = e^{-4\beta}$, $\hat{B}_\beta^\varepsilon(3) = e^{-8\beta}$, $\hat{B}_\beta^\varepsilon(4) = e^{-9\beta}$.

Let $C := 1/(e^{-2\beta} + e^{-4\beta} + e^{-8\beta} + e^{-9\beta})$.

Then: $B_\beta^\varepsilon(1) = Ce^{-2\beta}$, $B_\beta^\varepsilon(2) = Ce^{-4\beta}$, $B_\beta^\varepsilon(3) = Ce^{-8\beta}$, $B_\beta^\varepsilon(4) = Ce^{-9\beta}$. 


In the preceding three examples, $\varepsilon$ is one-to-one.
That is, $\varepsilon$ has no degeneracies.
In the next, $\varepsilon$ has one degeneracy, at energy-level 9.

Example: Let $\Sigma := \{1, 2, 3, 4\}$ and define $\varepsilon : \Sigma \to \mathbb{R}$ by:

$\varepsilon(1) = 2, \quad \varepsilon(2) = 4, \quad \varepsilon(3) = 9, \quad \varepsilon(4) = 9.$

Then:

$\hat{B}_\beta^\varepsilon \{1\} = e^{-2\beta}, \quad \hat{B}_\beta^\varepsilon \{2\} = e^{-4\beta}, \quad \hat{B}_\beta^\varepsilon \{3\} = e^{-9\beta}, \quad \hat{B}_\beta^\varepsilon \{4\} = e^{-9\beta}.$

Let $C := 1/(e^{-2\beta} + e^{-4\beta} + 2 \cdot e^{-9\beta}).$
Then:

$\hat{B}_\beta^\varepsilon \{1\} = Ce^{-2\beta}, \quad \hat{B}_\beta^\varepsilon \{2\} = Ce^{-4\beta}, \quad \hat{B}_\beta^\varepsilon \{3\} = Ce^{-9\beta}, \quad \hat{B}_\beta^\varepsilon \{4\} = Ce^{-9\beta}.$

In the next example, $\varepsilon$ has many degeneracies.

Example: Let $\Sigma := \left( [0..4] \times [0..4] \right) \setminus \{(4, 4)\}$.

Let $\beta \in \mathbb{R}$ and define $\varepsilon : \Sigma \to \mathbb{R}$ by:

$\forall \sigma \in \Sigma, \quad \varepsilon(\sigma) = \sigma_1 + \sigma_2.$

Then:

$\hat{B}_\beta^\varepsilon \{(3, 2)\} = e^{-5\beta}, \quad \hat{B}_\beta^\varepsilon \{(1, 4)\} = e^{-5\beta}, \quad \hat{B}_\beta^\varepsilon \{(0, 0)\} = 1.$

Generally,

$\forall \sigma \in \Sigma, \quad \hat{B}_\beta^\varepsilon \{\sigma\} = e^{-(\sigma_1 + \sigma_2)\beta}.$

Let $C := 1/(\sum_{\sigma \in \Sigma} [e^{-(\sigma_1 + \sigma_2)\beta}]).$
Then:

$\hat{B}_\beta^\varepsilon \{(3, 2)\} = Ce^{-5\beta}, \quad \hat{B}_\beta^\varepsilon \{(1, 4)\} = Ce^{-5\beta}, \quad \hat{B}_\beta^\varepsilon \{(0, 0)\} = C.$

Generally,

$\forall \sigma \in \Sigma, \quad \hat{B}_\beta^\varepsilon \{\sigma\} = Ce^{-(\sigma_1 + \sigma_2)\beta}.$

THEOREM 20.2. Let $\Sigma$ be a nonempty finite set.

Let $\varepsilon : \Sigma \to \mathbb{R}, \quad \xi, \beta \in \mathbb{R}.$

Then:

$\hat{B}_\beta^\varepsilon = B_\beta^{\varepsilon - \xi}.$

Proof. For all $\sigma \in \Sigma$, let $\varepsilon_\sigma := \varepsilon(\sigma)$.

Since, $\forall \sigma \in \Sigma, \quad \hat{B}_\beta^\varepsilon \{\sigma\} = e^{-\beta \varepsilon_\sigma} \cdot e^{-\beta \varepsilon_{\xi - \xi}} = e^{-\beta \xi} \cdot (\hat{B}_\beta^{\varepsilon - \xi} \{\sigma\}),$

we get:

$\hat{B}_\beta^\varepsilon = e^{-\beta \xi} \cdot \hat{B}_\beta^{\varepsilon - \xi}.$

Then:

$\hat{B}_\beta^\varepsilon = \mathcal{N}(\hat{B}_\beta^\varepsilon) = \mathcal{N}(e^{-\beta \xi} \cdot \hat{B}_\beta^{\varepsilon - \xi}) = \mathcal{N}(\hat{B}_\beta^{\varepsilon - \xi}) = B_\beta^{\varepsilon - \xi}. \quad \Box$

DEFINITION 20.3. Let $\Sigma$ be a nonempty finite set, $\varepsilon : \Sigma \to \mathbb{R}$.

For all $\sigma \in \Sigma$, let $\varepsilon_\sigma := \varepsilon(\sigma)$.

For all $\beta \in \mathbb{R}$, let

$\Gamma_\beta^\varepsilon := \sum_{\sigma \in \Sigma} [\varepsilon_\sigma \cdot e^{-\beta \varepsilon_\sigma}],$

$\Delta_\beta^\varepsilon := \sum_{\sigma \in \Sigma} [e^{-\beta \varepsilon_\sigma}],$

$A_\beta^\varepsilon := \Gamma_\beta^\varepsilon / \Delta_\beta^\varepsilon.$

Then:

$\Gamma_\beta^\varepsilon = \sum_{\sigma \in \Sigma} [\varepsilon_\sigma \cdot (\hat{B}_\beta^\varepsilon \{\sigma\})].$

Then:

$\Gamma_\beta^\varepsilon$ is the integral of $\varepsilon$ wrt $\hat{B}_\beta^\varepsilon$. 
Since \( \Delta_{\beta}^\varepsilon = \sum_{\sigma \in \Sigma} [\hat{B}_{\beta}^\varepsilon(\sigma)] \),
we get: \( \Delta_{\beta}^\varepsilon = \hat{B}_{\beta}^\varepsilon(\Sigma) \).

Since \( \Gamma_{\beta}^\varepsilon = \frac{\sum_{\sigma \in \Sigma} [\varepsilon_{\sigma} \cdot (\hat{B}_{\beta}^\varepsilon(\sigma))]}{B_{\beta}^\varepsilon(\Sigma)} \),
we get: \( A_{\beta}^\varepsilon = \sum_{\sigma \in \Sigma} [\varepsilon_{\sigma} \cdot (B_{\beta}^\varepsilon(\sigma))] \).

Then: \( A_{\beta}^\varepsilon \) is the average value of \( \varepsilon \) wrt \( B_{\beta}^\varepsilon \).

(“\( A \)” is for “Average”.)

**THEOREM 20.4.** Let \( \Sigma \) be a nonempty finite set.

Let \( \varepsilon : \Sigma \to \mathbb{R}, \beta \in \mathbb{R} \). Then: \( M_{\varepsilon \cdot B_{\beta}^\varepsilon} = A_{\beta}^\varepsilon \).

**Proof.** For all \( \sigma \in \Sigma \), let \( \varepsilon_{\sigma} := \varepsilon(\sigma) \).

Because \( \Sigma \) is the disjoint union, over \( t \in \mathbb{I}_\varepsilon \), of \( \varepsilon^s\{t\} \),
we get: \( \sum_{t \in \mathbb{I}_\varepsilon} \sum_{\sigma \in \varepsilon^s\{t\}} [\varepsilon_{\sigma} \cdot (B_{\beta}^\varepsilon(\sigma))] = \sum_{\sigma \in \Sigma} [\varepsilon_{\sigma} \cdot (B_{\beta}^\varepsilon(\sigma))] \).

Also, \( A_{\beta}^\varepsilon = \sum_{\sigma \in \Sigma} [\varepsilon_{\sigma} \cdot (B_{\beta}^\varepsilon(\sigma))] \).

Then: \( \sum_{t \in \mathbb{I}_\varepsilon} \sum_{\sigma \in \varepsilon^s\{t\}} [\varepsilon_{\sigma} \cdot (B_{\beta}^\varepsilon(\sigma))] = M_{\varepsilon \cdot B_{\beta}^\varepsilon} \).

So, since \( \sum_{t \in \mathbb{I}_\varepsilon} [t \cdot ((\varepsilon_{\sigma} B_{\beta}^\varepsilon(\sigma))] = M_{\varepsilon \cdot B_{\beta}^\varepsilon} \),
we want: \( \sum_{t \in \mathbb{I}_\varepsilon} [t \cdot ((\varepsilon_{\sigma} B_{\beta}^\varepsilon(\sigma))] = \sum_{\sigma \in \Sigma} [\varepsilon_{\sigma} \cdot (B_{\beta}^\varepsilon(\sigma))] \).

Want: \( \forall t \in \mathbb{I}_\varepsilon, t \cdot ((\varepsilon_{\sigma} B_{\beta}^\varepsilon(\sigma])] = \sum_{\sigma \in \varepsilon^s\{t\}} [\varepsilon_{\sigma} \cdot (B_{\beta}^\varepsilon(\sigma))] \).

Given \( t \in \mathbb{I}_\varepsilon \), want: \( t \cdot ((\varepsilon_{\sigma} B_{\beta}^\varepsilon(\sigma])] = \sum_{\sigma \in \varepsilon^s\{t\}} [\varepsilon_{\sigma} \cdot (B_{\beta}^\varepsilon(\sigma))] \).

For all \( \sigma \in \varepsilon^s\{t\} \), since \( \varepsilon_{\sigma} = \varepsilon(\sigma) \in \{t\} \), we get: \( t = \varepsilon_{\sigma} \).

Want: \( t \cdot ((\varepsilon_{\sigma} B_{\beta}^\varepsilon(\sigma])] = \sum_{\sigma \in \varepsilon^s\{t\}} [t \cdot (B_{\beta}^\varepsilon(\sigma))] \).

Because \( \varepsilon^s\{t\} \) is the disjoint union, over \( \sigma \in \varepsilon^s\{t\}, \) of \( \{\sigma\} \),
we get: \( B_{\beta}^\varepsilon(\varepsilon^s\{t\}) = \sum_{\sigma \in \varepsilon^s\{t\}} [B_{\beta}^\varepsilon(\sigma)] \).

Also, \( (\varepsilon_{\sigma} B_{\beta}^\varepsilon(\sigma)] = B_{\beta}^\varepsilon(\varepsilon_{\sigma}) \).

Then: \( t \cdot ((\varepsilon_{\sigma} B_{\beta}^\varepsilon(\sigma])] = t \cdot (B_{\beta}^\varepsilon(\varepsilon_{\sigma})) = \sum_{\sigma \in \varepsilon^s\{t\}} [t \cdot (B_{\beta}^\varepsilon(\sigma))] \). \( \square \)

**THEOREM 20.5.** Let \( \Sigma \) be a nonempty finite set.

Let \( \varepsilon : \Sigma \to \mathbb{R}, \beta, \xi \in \mathbb{R} \). Then: \( A_{\beta}^{\varepsilon - \xi} = A_{\beta}^\varepsilon - \xi \).

**Proof.** We have: \( B_{\beta}^\varepsilon(\Sigma) = \sum_{\sigma \in \Sigma} [B_{\beta}^\varepsilon(\sigma)] \).

Since \( B_{\beta}^\varepsilon \in \mathcal{P}_\Sigma \), we get: \( B_{\beta}^\varepsilon(\Sigma) = 1 \).

By Theorem 20.2, we have: \( B_{\beta}^{\varepsilon - \xi} = B_{\beta}^\varepsilon - \xi \).

For all \( \sigma \in \Sigma \), let \( \varepsilon_{\sigma} := \varepsilon(\sigma) \).

Then: \( A_{\beta}^{\varepsilon - \xi} = \sum_{\sigma \in \Sigma} [\varepsilon_{\sigma} \cdot (B_{\beta}^{\varepsilon - \xi}(\sigma))] \)
\[ = \sum_{\sigma \in \Sigma} [(\varepsilon_{\sigma} - \xi) \cdot (B_{\beta}^\varepsilon(\sigma))] \]
\[ = (\sum_{\sigma \in \Sigma} [\varepsilon_{\sigma} \cdot (B_{\beta}^\varepsilon(\sigma))] - (\sum_{\sigma \in \Sigma} [\xi \cdot (B_{\beta}^\varepsilon(\sigma))]) \]
\[ = (\sum_{\sigma \in \Sigma} [\varepsilon_{\sigma} \cdot (B_{\beta}^\varepsilon(\sigma))] - \xi \cdot (\sum_{\sigma \in \Sigma} [B_{\beta}^\varepsilon(\sigma)]) \]
\[ = A_{\beta}^\varepsilon - \xi \cdot (B_{\beta}^\varepsilon(\Sigma)) = A_{\beta}^\varepsilon - \xi \cdot 1 = A_{\beta}^\varepsilon - \xi. \] \( \square \)
For any sets $\Sigma, T$, for any function $\varepsilon : \Sigma \to T$,
the image of $\varepsilon$ is: $\mathbb{I}_\varepsilon := \{ \varepsilon(\sigma) \mid \sigma \in \Sigma \} \subseteq T$.

**THEOREM 20.6.** Let $\Sigma$ be a nonempty finite set, $\varepsilon : \Sigma \to \mathbb{R}$. Then:

- as $\beta \to \infty$, $A_\beta^\varepsilon \to \min \mathbb{I}_\varepsilon$
- and as $\beta \to -\infty$, $A_\beta^\varepsilon \to \max \mathbb{I}_\varepsilon$.

The proof is a matter of bookkeeping, best explained by example:

**Proof.** For all $\sigma \in \Sigma$, let $\varepsilon_\sigma := \varepsilon(\sigma)$.

We have: $\forall \beta \in \mathbb{R}, A_\beta^\varepsilon(\beta) = \frac{2e^{-2\beta} + 4e^{-4\beta} + 9e^{-9\beta} + 9e^{-9\beta}}{e^{-2\beta} + e^{-4\beta} + e^{-9\beta} + e^{-9\beta}}$, and so

- as $\beta \to \infty$, $A_\beta^\varepsilon \to 2/1$
- and as $\beta \to -\infty$, $A_\beta^\varepsilon \to 18/2$.

For any nonempty finite set $\Sigma$, for any $\varepsilon : \Sigma \to \mathbb{R}$, define $A_\varepsilon^\star : \mathbb{R} \to \mathbb{R}$ by: $\forall \beta \in \mathbb{R}$, $A_\varepsilon^\star(\beta) = A_\beta^\varepsilon$.

**THEOREM 20.7.** Let $\Sigma$ be a finite set.

**Let** $\varepsilon : \Sigma \to \mathbb{R}$.

**Assume:** $\# \mathbb{I}_\varepsilon \geq 2$.

Then: $A_\varepsilon^\star$ is a strictly-decreasing $C^\omega$-diffeomorphism from $\mathbb{R}$ onto $(\min \mathbb{I}_\varepsilon; \max \mathbb{I}_\varepsilon)$.

**Proof.** For all $\sigma \in \Sigma$, let $\varepsilon_\sigma := \varepsilon(\sigma)$.

We have: $\forall \beta \in \mathbb{R}, A_\varepsilon^\star(\beta) = \frac{\varepsilon_\sigma \cdot e^{-\beta \varepsilon_\sigma}}{\sum_{\varphi \in \Sigma} \varepsilon_\varphi \cdot e^{-\beta \varepsilon_\varphi}}$. Then $A_\varepsilon^\star : \mathbb{R} \to \mathbb{R}$ is $C^\omega$.

So, by Theorem 20.6 and the $C^\omega$-Inverse Function Theorem and the Mean Value Theorem, it suffices to show: $(A_\varepsilon^\star)' < 0$ on $\mathbb{R}$.

**Given** $\beta \in \mathbb{R}$, **want:** $(A_\varepsilon^\star)'(\beta) < 0$.

**Let** $P := \sum_{\sigma \in \Sigma} [\varepsilon_\sigma \cdot e^{-\beta \varepsilon_\sigma}]$, $P' := \sum_{\sigma \in \Sigma} [(-\varepsilon_\sigma^2) \cdot e^{-\beta \varepsilon_\sigma}]$.

**Let** $Q := \sum_{\tau \in \Sigma} [e^{-\beta \varepsilon_\tau}]$, $Q' := \sum_{\tau \in \Sigma} [(-\varepsilon_\tau) \cdot e^{-\beta \varepsilon_\tau}]$.

Then $Q > 0$. Also, by the Quotient Rule, $(A_\varepsilon^\star)'(\beta) = [QP' - PQ']/Q^2$.

**Want:** $QP' - PQ' < 0$.

We have: $QP' = \sum_{\sigma \in \Sigma} \sum_{\tau \in \Sigma} [(-\varepsilon_\sigma^2) \cdot e^{-\beta (\varepsilon_\sigma + \varepsilon_\tau)}]$.

We have: $PQ' = \sum_{\sigma \in \Sigma} \sum_{\tau \in \Sigma} [(-\varepsilon_\sigma \varepsilon_\tau) \cdot e^{-\beta (\varepsilon_\sigma + \varepsilon_\tau)}]$.

Then: $QP' - PQ' = \sum_{\sigma \in \Sigma} \sum_{\tau \in \Sigma} [(-\varepsilon_\sigma^2 \varepsilon_\tau - \varepsilon_\sigma \varepsilon_\tau^2) - e^{-\beta (\varepsilon_\sigma + \varepsilon_\tau)}]$.
Interchanging $\sigma$ and $\tau$, we get:

$$QP' - PQ' = \sum_{\sigma \in \Sigma} \sum_{\tau \in \Sigma} \left[ (-\varepsilon_\tau^2 + \varepsilon_\tau \varepsilon_\sigma) \cdot e^{-\beta(\varepsilon_\tau + \varepsilon_\sigma)} \right].$$

By commutativity of addition and multiplication, adding the last two equations gives:

$$2 \cdot (QP' - PQ') = \sum_{\sigma \in \Sigma} \sum_{\tau \in \Sigma} \left[ (-\varepsilon_\sigma^2 - \varepsilon_\sigma^2 + 2\varepsilon_\sigma \varepsilon_\tau) \cdot e^{-\beta(\varepsilon_\sigma + \varepsilon_\tau)} \right].$$

Then:

$$2 \cdot (QP' - PQ') < 0.$$  

Then: $QP' - PQ' < 0. \quad \square$

**Definition 20.8.** Let $\Sigma$ be a finite set. Let $\varepsilon : \Sigma \to \mathbb{R}$.

Assume: $\# \Sigma \geq 2$. Let $\alpha \in (\min \varepsilon; \max \varepsilon)$.

The $\alpha$-Boltzmann parameter on $\varepsilon$ is:

$$BP_\alpha^\varepsilon := (A_\varepsilon^\alpha)^{-1}(\alpha).$$

So the $\alpha$-Boltzmann parameter on $E$ is the unique $\beta \in \mathbb{R}$ s.t. $A_\beta^\varepsilon = \alpha$.

**Example:** Let $\Sigma := \{0, 1, 10\}$, and define $\varepsilon : \Sigma \to \mathbb{R}$ by:

$$\forall \sigma \in \Sigma, \quad \varepsilon(\sigma) = \sigma.$$  

Computation shows: $A_{(\ln 9)/10}^\varepsilon = 1$. Then: $BP_1^\varepsilon = (\ln 9)/10$.

**Example:** Let $\Sigma := \{2, 4, 8, 9\}$, and define $\varepsilon : \Sigma \to \mathbb{R}$ by:

$$\forall \sigma \in \Sigma, \quad \varepsilon(\sigma) = \sigma.$$  

To evaluate $BP_5^\varepsilon$, we must solve $A_5^\varepsilon(\beta) = 5$ for $\beta$,

and, since, by Theorem 20.7, $A_5^\varepsilon$ is strictly-decreasing, there are simple iterative methods to do this.

We compute: $BP_5^\varepsilon \approx 9.3127$, accurate to four decimal places.

(Thanks to C. Prouty for this calculation. See §25.)

Next, let $\Sigma := \{1, 2, 3, 4\}$, and define $\varepsilon : \Sigma \to \mathbb{R}$ by:

$$\varepsilon(1) = 2, \quad \varepsilon(2) = 4, \quad \varepsilon(3) = 8, \quad \varepsilon(4) = 9.$$  

Then $A_5^\varepsilon = A_4^\varepsilon$, so $BP_5^\varepsilon = BP_4^\varepsilon$.

Then $BP_5^\varepsilon \approx 9.3127$, accurate to four decimal places.

**Example:** Let $\Sigma := \{1, 2, 3, 4\}$.

Let $\Sigma := \{1, 2, 3, 4\}$ and define $\varepsilon : \Sigma \to \mathbb{R}$ by:

$$\varepsilon(1) = 2, \quad \varepsilon(2) = 4, \quad \varepsilon(3) = 9, \quad \varepsilon(4) = 9.$$  

To evaluate $BP_5^\varepsilon$, we must solve $A_5^\varepsilon(\beta) = 5$ for $\beta$,

and, since, by Theorem 20.7, $A_5^\varepsilon$ is strictly-decreasing, there are simple iterative methods to do this.

We compute: $BP_5^\varepsilon \approx 9.3127$, accurate to four decimal places.

(Thanks to C. Prouty for this calculation. See §25.)
Example: Let $\Sigma := (\{0.4\} \times \{0.4\}) \setminus \{(4,4)\}$.

Define $\varepsilon : \Sigma \to \mathbb{R}$ by: $\forall \sigma \in \Sigma, \: \varepsilon(\sigma) = \sigma_1 + \sigma_2$.

To evaluate $\text{BP}^\varepsilon_1$, we must solve $A^\varepsilon_1(\beta) = 1$ for $\beta$.

and, since, by Theorem 20.7, $A^\varepsilon_1$ is strictly-decreasing, there are simple iterative methods to do this.

We compute: $\text{BP}^\varepsilon_1 \approx 1.0670$, accurate to four decimal places.

(Thanks to C. Prouty for this calculation. See §25.)

**THEOREM 20.9.** Let $\Sigma$ be a finite set.

Let $\varepsilon : \Sigma \to \mathbb{R}$. Assume: $\#\Sigma \geq 2$.

Let $\alpha \in (\min \Sigma, \max \Sigma)$. Let $\xi \in \mathbb{R}$. Then: $\text{BP}^\varepsilon_{\alpha-\xi} = \text{BP}^\varepsilon_\alpha$.

Proof. Let $\beta := \text{BP}^\varepsilon_\alpha$. Want: $\text{BP}^\varepsilon_{\alpha-\xi} = \beta$.

Since $\beta = \text{BP}^\varepsilon_\alpha = (A^\varepsilon_\alpha)^{-1}(\alpha)$, we get: $(A^\varepsilon_\alpha)(\beta) = \alpha$.

By Theorem 20.5, $A^\varepsilon_{\beta-\xi} = A^\varepsilon_\beta - \xi$.

Since $(A^\varepsilon_{\beta-\xi})(\beta) = A^\varepsilon_{\beta-\xi} = A^\varepsilon_\beta - \xi = ((A^\varepsilon_\alpha)(\beta)) - \xi = \alpha - \xi$,

we get: $\beta = (A^\varepsilon_{\alpha-\xi})^{-1}(\alpha - \xi)$.

Then: $\text{BP}^\varepsilon_{\alpha-\xi} = (A^\varepsilon_{\alpha-\xi})^{-1}(\alpha - \xi) = \beta$. \qed

**21. Degeneracy**

**THEOREM 21.1.** Let $\Sigma$ be a finite set.

Let $\varepsilon : \Sigma \to \mathbb{Z}$. Assume $\varepsilon$ is residue-constrained.

Let $\alpha \in \mathbb{Z}$. Assume $\alpha \in (\min \varepsilon; \max \varepsilon)$.

Let $\beta := \text{BP}^\varepsilon_\alpha$.

Let $t_1, t_2, \ldots \in \mathbb{Z}$. Assume: $\{t_n - n\alpha : n \in \mathbb{N}\}$ is bounded.

For all $n \in \mathbb{N}$, let $\Omega_n := \{f \in \Sigma^n \mid (\varepsilon(f_1)) + \cdots + (\varepsilon(f_n)) = t_n\}$.

Let $\sigma_0 \in \Sigma$. Then: as $n \to \infty$, $\nu_{\Omega_n} \{f \in \Omega_n \mid f_n = \sigma_0\} \to B^\varepsilon_\beta(\sigma_0)$.

Recall: $\nu_\emptyset(\emptyset) = -1$.

So, since $B^\varepsilon_\beta(\sigma_0) > 0$, part of the content of Theorem 21.1 is:

$\forall$ sufficiently large $n \in \mathbb{N}, \: \Omega_n \neq \emptyset$.

See Claim 2 in the proof below.

Proof. Since $\beta = \text{BP}^\varepsilon_\alpha = (A^\varepsilon_\alpha)^{-1}(\alpha)$, we get: $A^\varepsilon_\beta(\beta) = \alpha$.

By Theorem 20.4, we have: $M^\varepsilon_{\beta} B^\varepsilon_\beta = A^\varepsilon_\beta$.

So, since $A^\varepsilon_\beta = A^\varepsilon_\beta(\beta) = \alpha$, we get: $M^\varepsilon_{\beta} B^\varepsilon_\beta = \alpha$.

Let $\mu := B^\varepsilon_\beta$. Then: $\mu \in \mathcal{P}_\Sigma$ and $M^\varepsilon_{\beta} \mu = \alpha$.

Let $E := \varepsilon$, $\tilde{\mu} := \varepsilon \mu$. Then: $\tilde{\mu} \in \mathcal{P}_E$ and $M_{\tilde{\mu}} = \alpha$.

By hypothesis, $E$ is residue-unconstrained.

Since $\varepsilon : \Sigma \to \mathbb{Z}$, we get: $E \subseteq \mathbb{Z}$.

Since $\Sigma$ is finite, we get: $E$ is finite.
Recall: Claim 1: As $n \to \infty$, we get: $|\tilde{\mu}|_1 < \infty$ and $|\tilde{\mu}|_2 < \infty$.

For all $\sigma \in \Sigma$, let $\varepsilon_\sigma := \varepsilon(\sigma)$.

Then: $\forall n \in \mathbb{N}, \quad \Omega_n = \{ f \in \Sigma^n \mid \varepsilon f_1 + \cdots + \varepsilon f_n = t_n \}$.

For all $n \in \mathbb{N}$, define $\varepsilon^n : \Sigma^n \to E^n$ by:

$\forall f_1, \ldots, f_n \in \Sigma, \quad \varepsilon^n(f_1, \ldots, f_n) = (\varepsilon f_1, \ldots, \varepsilon f_n)$.

Then, since $\varepsilon_* \mu = \tilde{\mu}$, we get: $\forall n \in \mathbb{N}, \quad (\varepsilon^n)_* (\mu^n) = \tilde{\mu}^n$.

For all $n \in \mathbb{N}$, let $\Omega_n := \{ \tilde{f} \in E^n \mid \tilde{f}_1 + \cdots + \tilde{f}_n = t_n \}$.

Then:

\[ (\varepsilon^n)_* (\Omega_n) = \Omega_n. \]

For all $n \in \mathbb{N}$, define $\psi_n : \mathbb{Z} \to \mathbb{R}$ by:

$\forall t \in \mathbb{Z}, \quad \psi_n(t) = \tilde{\mu}^n \{ f \in E^n \mid f_1 + \cdots + f_n = t \}$.

Then: $\forall n \in \mathbb{N}, \quad \psi_n(t_n) = \tilde{\mu}^n(\Omega_n)$.

Since $E$ is finite and residue-unconstrained, we get: $2 \leq \#E < \infty$.

Since $\mu = B_\beta^\#$ and $S_{B_\beta^\#} = \Sigma$, we get: $S_{\mu} = \Sigma$.

So, since $\varepsilon : \Sigma \to \mathbb{Z}$, we get: $S_{\varepsilon_* \mu} = \mathbb{I}_\varepsilon \subseteq \mathbb{Z}$.

Recall: $\varepsilon_* \mu = \tilde{\mu}$ and $\mathbb{I}_\varepsilon = E$. Then: $S_{\tilde{\mu}} = E$.

Since $E$ is finite and $\tilde{\mu} \in \mathcal{P}_E$ and $|\tilde{\mu}|_1 < \infty$ and $\#S_{\tilde{\mu}} = \#E \geq 2$, by Theorem 8.6, we get $V_{\tilde{\mu}} > 0$.

So, since $V_{\tilde{\mu}} = |\tilde{\mu}|_2 - M_{\tilde{\mu}}^2 \leq |\tilde{\mu}|_2 < \infty$, we conclude:

$0 < V_{\tilde{\mu}} < \infty$.

Let $v := V_{\tilde{\mu}}$. Then $0 < v < \infty$. Then $1/\sqrt{2\pi v} > 0$.

Let $\tau := 1/\sqrt{2\pi v}$. Then $\tau > 0$.

Claim 1: As $n \to \infty$, $\sqrt{n} \cdot \psi_n(t_n) \to \tau$.

Proof of Claim 1: We have:

$\forall n \in \mathbb{N}, \quad \psi_n(t_n) = \tilde{\mu}^n \{ \tilde{f} \in E^n \mid \tilde{f}_1 + \cdots + \tilde{f}_n = t_n \}$.

Want: as $n \to \infty$, $\sqrt{n} \cdot \tilde{\mu}^n \{ \tilde{f} \in E^n \mid \tilde{f}_1 + \cdots + \tilde{f}_n = t_n \} \to 1/\sqrt{2\pi v}$.

Recall: $E \subseteq \mathbb{Z}$, $E$ is residue-unconstrained, $\tilde{\mu} \in \mathcal{P}_E$, $S_{\tilde{\mu}} = E$.

Then, as $n \to \infty$, $\sqrt{n} \cdot (\tilde{\mu}^n \{ \tilde{f} \in E^n \mid \tilde{f}_1 + \cdots + \tilde{f}_n = t_n \}) \to 1/\sqrt{2\pi v}$.

End of proof of Claim 1.

Since $\tau > 0$, by Claim 1, choose $n_0 \in [2, \infty)$ s.t.

$\forall n \in [n_0, \infty), \quad \sqrt{n} \cdot \psi_n(t_n) > 0$. 

Claim 2: Let $n \in [n_0..\infty)$. Then: $\mu^n(\Omega_n) > 0$.

Proof of Claim 2: Recall: $\tilde{\mu}^n(\widetilde{\Omega}_n) = \mu^n(\Omega_n)$ and $\psi_n(t_n) = \tilde{\mu}^n(\widetilde{\Omega}_n)$.
By the choice of $n_0$, we get: $\sqrt{n} \cdot (\psi_n(t_n)) > 0$. Then: $\psi_n(t_n) > 0$.
Then: $\mu^n(\Omega_n) = \tilde{\mu}^n(\widetilde{\Omega}_n) = \psi_n(t_n) > 0$.

End of proof of Claim 2.

Since $\mu \in \mathcal{P}_\Sigma$, we get: $\forall n \in \mathbb{N}$, $\mu^n \in \mathcal{P}_{\Sigma^n}$, so $\mu^n(\Omega_n) \leq 1$.
So, by Claim 2, $\forall n \in [n_0..\infty)$, $0 < \mu^n(\Omega_n) \leq 1$.
Also, we have: $\forall n \in \mathbb{N}$, $(\mu^n|\Omega_n)(\Omega_n) = \mu^n(\Omega_n)$.
Then: $\forall n \in [n_0..\infty)$, $0 < (\mu^n|\Omega_n)(\Omega_n) \leq 1$.
Then: $\forall n \in [n_0..\infty)$, $\mu^n|\Omega_n \in \mathcal{F}\mathcal{M}_{\Omega_n}$.
Then: $\forall n \in [n_0..\infty)$, $\mathcal{N}(\mu^n|\Omega_n) \in \mathcal{P}_{\Omega_n}$.
Since $\#E \geq 2$, we get: $E \neq \emptyset$.
Since $\varepsilon : \Sigma \rightarrow \mathbb{Z}$ and since $\mathcal{I}_\varepsilon = E \neq \emptyset$, we get: $\Sigma \neq \emptyset$.
Then $\hat{B}_\beta^\varepsilon(\Sigma) > 0$. Let $C : = 1/(\hat{B}_\beta^\varepsilon(\Sigma))$. Then $\mathcal{N}(\hat{B}_\beta^\varepsilon) = C \cdot \hat{B}_\beta^\varepsilon$.
By definition of $\hat{B}_\beta^\varepsilon$, we have: $\forall \sigma \in \Sigma$, $\hat{B}_\beta^\varepsilon\{\sigma\} = e^{-\beta\varepsilon\sigma}$.
So, since $\mu = B_\beta^\varepsilon = \mathcal{N}(\hat{B}_\beta^\varepsilon) = C \cdot \hat{B}_\beta^\varepsilon$,
we get: $\forall \sigma \in \Sigma$, $\mu\{\sigma\} = Ce^{-\beta\varepsilon\sigma}$.
We have: $\forall n \in \mathbb{N}$, $\forall S \subseteq \Omega_n$, $(\mu^n|\Omega_n)(S) = \mu^n(S)$.
Then: $\forall n \in \mathbb{N}$, $(\mu^n|\Omega_n)(\Omega_n) = \mu^n(\Omega_n)$.
For all $n \in \mathbb{N}$, let $z_n := \mu^n(\Omega_n)$.
Then, by Claim 2, we have: $\forall n \in [n_0..\infty)$, $z_n > 0$.
Also, since $\forall n \in [n_0..\infty)$, $\mu^n(\Omega_n) = \tilde{\mu}^n(\widetilde{\Omega}_n)$,
we conclude: $\forall n \in [n_0..\infty)$, $z_n = \tilde{\mu}^n(\widetilde{\Omega}_n)$.
For all $n \in [n_0..\infty)$, let $\lambda_n := \mathcal{N}(\mu^n|\Omega_n)$.
Then: $\forall n \in [n_0..\infty)$, $\lambda_n = (\mu^n|\Omega_n)/(\mu^n(\Omega_n))$.
Then: $\forall n \in [n_0..\infty)$, $\lambda_n = (\mu^n|\Omega_n)/z_n$.
Then: $\forall n \in [n_0..\infty)$, $\forall S \subseteq \Omega_n$, $\lambda_n(S) = (\mu^n(S))/z_n$.

Claim 3: Let $n \in [n_0..\infty)$. Then: $\lambda_n = \nu_n$.

Proof of Claim 3: Let $F := \Omega_n$. Want: $\lambda_n = \nu_F$.
Since $\lambda_n = \mathcal{N}(\mu^n|\Omega_n) = \mathcal{N}(\mu^n|F)$, we get: $\lambda_n \in \mathcal{P}_F$.
By Theorem 8.9, given $f, g \in F$, want: $\lambda_n\{f\} = \lambda_n\{g\}$.
Want: $(\mu^n\{f\})/z_n = (\mu^n\{g\})/z_n$.
Want: $(\mu^n\{f\}) = \mu^n\{g\}$.
For all $i \in [1..n]$, let $\hat{f}_i := \varepsilon f_i$ and $\hat{g}_i := \varepsilon g_i$.
Then: $\forall i \in [1..n]$, $\mu\{f_i\} = Ce^{-\beta\hat{f}_i}$ and $\mu\{g_i\} = Ce^{-\beta\hat{g}_i}$.
Since $f \in F = \Omega_n$, we get: $\varepsilon f_1 + \cdots + \varepsilon f_n = t_n$. 
Since $g \in F = \Omega_n$, we get: $\varepsilon_{g_1} + \cdots + \varepsilon_{g_n} = t_n$.

Since $\tilde{f}_1 + \cdots + \tilde{f}_n = \varepsilon_{f_1} + \cdots + \varepsilon_{f_n} = t_n$

we get: $C^n e^{-\beta (\tilde{f}_1 + \cdots + \tilde{f}_n)} = C^n e^{-\beta (\tilde{g}_1 + \cdots + \tilde{g}_n)}$.

Then: $(Ce^{-\beta \tilde{f}_1}) \cdots (Ce^{-\beta \tilde{f}_n}) = (Ce^{-\beta \tilde{g}_1}) \cdots (Ce^{-\beta \tilde{g}_n})$.

Then: $(\mu\{f_1\}) \cdots (\mu\{f_n\}) = (\mu\{g_1\}) \cdots (\mu\{g_n\})$.

Then: $\mu^n\{f\} = \mu^n\{g\}$.

End of proof of Claim 5.

Claim 4: Let $\sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}$. Then: $\mu\{\sigma\} = \mu\{\sigma_0\}$.

Proof of Claim 4: Since $\sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}$, we get: $\varepsilon(\sigma) \in \{\varepsilon_{\sigma_0}\}$.

Since $\varepsilon_{\sigma} = \varepsilon(\sigma) \in \{\varepsilon_{\sigma_0}\}$, we get: $\varepsilon_{\sigma} = \varepsilon_{\sigma_0}$.

Then: $\mu\{\sigma\} = Ce^{-\beta \varepsilon_{\sigma}} = Ce^{-\beta \varepsilon_{\sigma_0}} = \mu\{\sigma_0\}$.

End of proof of Claim 4.

Since $\varepsilon(\sigma_0) = \varepsilon_{\sigma_0} \in \{\varepsilon_{\sigma_0}\}$, we get: $\sigma_0 \in \varepsilon^*\{\varepsilon_{\sigma_0}\}$.

Then $\varepsilon^*\{\varepsilon_{\sigma_0}\} \neq \emptyset$, so $\#(\varepsilon^*\{\varepsilon_{\sigma_0}\}) \geq 1$.

Let $k := \#(\varepsilon^*\{\varepsilon_{\sigma_0}\})$. Then: $k \geq 1$.

Claim 5: $\mu(\varepsilon^*\{\varepsilon_{\sigma_0}\}) = k \cdot (\mu\{\sigma_0\})$.

Proof of Claim 5: Since $\varepsilon^*\{\varepsilon_{\sigma_0}\}$ is equal to the disjoint union, over $\sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}$, of $\{\sigma\}$,

we get: $\mu(\varepsilon^*\{\varepsilon_{\sigma_0}\}) = \sum_{\sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}} [\mu(\sigma)]$.

So, by Claim 4, we get: $\mu(\varepsilon^*\{\varepsilon_{\sigma_0}\}) = k \cdot (\mu\{\sigma_0\})$.

End of proof of Claim 5.

Claim 6: Let $n \in [n_0, \infty)$. Let $\sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}$.

Then: $\mu^n\{f \in \Omega_n \mid f_n = \sigma\} = \mu^n\{f \in \Omega_n \mid f_n = \sigma_0\}$.

Proof of Claim 6: By choice of $n_0$, we have: $n_0 \in [2, \infty)$.

Then $[n_0, \infty) \subseteq [2, \infty)$, so, since $n \in [n_0, \infty)$, we get: $n \in [2, \infty)$.

Let $X := \{f \in \Sigma^{n-1} \mid \varepsilon_{f_1} + \cdots + \varepsilon_{f_{n-1}} = t_n - \varepsilon_{\sigma}\}$.

We have

$\{f \in \Omega_n \mid f_n = \sigma\} = \{f \in \Sigma^n \mid [\varepsilon_{f_1} + \cdots + \varepsilon_{f_{n-1}} + \varepsilon_{f_n} = t_n] \& [f_n = \sigma]\}$

$= \{f \in \Sigma^n \mid [\varepsilon_{f_1} + \cdots + \varepsilon_{f_{n-1}} + \varepsilon_{\sigma} = t_n] \& [f_n = \sigma]\}$

$= \{f \in \Sigma^n \mid [\varepsilon_{f_1} + \cdots + \varepsilon_{f_{n-1}} = t_n - \varepsilon_{\sigma}] \& [f_n = \sigma]\}$,

so, under the standard bijection $\Sigma^n \leftrightarrow \Sigma^{n-1} \times \Sigma$, we have:

$\{f \in \Omega_n \mid f_n = \sigma\} \subseteq \Sigma^n$

corresponds to $X \times \{\sigma\} \subseteq \Sigma^{n-1} \times \Sigma$, respectively.
so \(\mu^n\{f \in \Omega_n \mid f_n = \sigma\} = (\mu^{n-1}(X)) \cdot (\mu\{\sigma\})\).

**Want:** \(\mu^n\{f \in \Omega_n \mid f_n = \sigma_0\} = (\mu^{n-1}(X)) \cdot (\mu\{\sigma_0\})\).

By Claim 4, we have: 
\(\mu\{\sigma\} = \mu\{\sigma_0\}\).

**Want:** \(\mu^n\{f \in \Omega_n \mid f_n = \sigma_0\} = (\mu^{n-1}(X)) \cdot (\mu\{\sigma_0\})\).

Since \(\sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}\), we get: 
\(\varepsilon(\sigma) \in \{\varepsilon_{\sigma_0}\}\).

Since \(\varepsilon_{\sigma} = \varepsilon(\sigma) \in \{\varepsilon_{\sigma_0}\}\), we get: 
\(\varepsilon_{\sigma} = \varepsilon_{\sigma_0}\).

Then \(X = \{f \in \Sigma^{n-1} \mid \varepsilon_{f_1} + \cdots + \varepsilon_{f_{n-1}} = t_n - \varepsilon_{\sigma_0}\}\).

We have:
\[
\{f \in \Omega_n \mid f_n = \sigma_0\} = \{f \in \Sigma^n \mid [\varepsilon_{f_1} + \cdots + \varepsilon_{f_{n-1}} + \varepsilon_{f_n} = t_n] \land [f_n = \sigma_0]\}
\]
\[
= \{f \in \Sigma^n \mid [\varepsilon_{f_1} + \cdots + \varepsilon_{f_{n-1}} + \varepsilon_{\sigma_0} = t_n] \land [f_n = \sigma_0]\}
\]
\[
= \{f \in \Sigma^n \mid [\varepsilon_{f_1} + \cdots + \varepsilon_{f_{n-1}} = t_n - \varepsilon_{\sigma_0}] \land [f_n = \sigma_0]\},
\]
so, under the standard bijection \(\Sigma^n \leftrightarrow \Sigma^{n-1} \times \Sigma\), we have:
\[
\{f \in \Omega_n \mid f_n = \sigma_0\} \subseteq \Sigma^n
\]
corresponds to 
\[
X \times \{\sigma_0\} \subseteq \Sigma^{n-1} \times \Sigma,
\]
so 
\(\mu^n\{f \in \Omega_n \mid f_n = \sigma_0\} = (\mu^{n-1}(X)) \cdot (\mu\{\sigma_0\})\).

**End of proof of Claim 6.**

**Claim 7:** Let \(n \in [n_0, \infty)\).

Then: 
\(\tilde{\mu}^n\{\tilde{f} \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0}\} = k \cdot (\mu^n\{f \in \Omega_n \mid f_n = \sigma_0\})\).

**Proof of Claim 7:** Recall: 
\(\tilde{\mu}^n = (\varepsilon^n)_* (\mu^n)\).

Since 
\[(\varepsilon^n)_* \{\tilde{f} \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0}\} = \{f \in \Omega_n \mid f_n \in \varepsilon^*\{\varepsilon_{\sigma_0}\}\},
\]
we get: 
\(\mu^n((\varepsilon^n)_* \{\tilde{f} \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0}\}) = \mu^n\{f \in \Omega_n \mid f_n \in \varepsilon^*\{\varepsilon_{\sigma_0}\}\}\).

Then: 
\[(\varepsilon^n)_* (\mu^n) = \mu^n\{f \in \Omega_n \mid f_n \in \varepsilon^*\{\varepsilon_{\sigma_0}\}\} = \mu^n\{f \in \Omega_n \mid f_n \in \varepsilon^*\{\varepsilon_{\sigma_0}\}\}.
\]

**Want:** 
\(\mu^n\{f \in \Omega_n \mid f_n \in \varepsilon^*\{\varepsilon_{\sigma_0}\}\} = k \cdot (\mu^n\{f \in \Omega_n \mid f_n = \sigma_0\})\).

Since 
\[
\{f \in \Omega_n \mid f_n \in \varepsilon^*\{\varepsilon_{\sigma_0}\}\}
\]
is the disjoint union, over \(\sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}\), of
\[
\{f \in \Omega_n \mid f_n = \sigma\},
\]
we get: 
\(\mu^n\{f \in \Omega_n \mid f_n \in \varepsilon^*\{\varepsilon_{\sigma_0}\}\} = \sum_{\sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}} [\mu^n\{f \in \Omega_n \mid f_n = \sigma\}]\).

So, since \(k = \#(\varepsilon^*\{\varepsilon_{\sigma_0}\})\), by Claim 6, we get:
\(\mu^n\{f \in \Omega_n \mid f_n \in \varepsilon^*\{\varepsilon_{\sigma_0}\}\} = k \cdot (\mu^n\{f \in \Omega_n \mid f_n = \sigma_0\})\).

**End of proof of Claim 7.**

Recall: 
\(\forall n \in [n_0, \infty)\), \(\lambda_n = \mathcal{N}(\mu^n|\Omega_n)\).

For all \(n \in [n_0, \infty)\), let \(\tilde{\lambda}_n := \mathcal{N}(\tilde{\mu}^n|\tilde{\Omega}_n)\).

Since, \(\forall n \in \mathbb{N}\), 
\((\varepsilon^n)_*(\mu^n) = \tilde{\mu}^n\) and 
\((\varepsilon^n)_*\tilde{\Omega}_n = \Omega_n\),
we get \(\forall n \in \mathbb{N}\), 
\((\varepsilon^n)_*(\mu^n|\Omega_n) = \tilde{\mu}^n|\tilde{\Omega}_n\).
so \( \forall n \in [n_0, \infty), \quad (\varepsilon^n)_*(N(\mu^n|\Omega_n)) = N(\pmb L^n|\pmb \tilde \Omega_n). \)

Then: \( \forall n \in [n_0, \infty), \quad (\varepsilon^n)_*(\lambda_n) = \lambda_n. \)

Recall: \( \forall n \in [n_0, \infty), \quad z_n = \tilde \mu^n(\Omega_n). \)

We have: \( \forall n \in [n_0, \infty), \forall S \subseteq \tilde \Omega_n, \quad (\pmb L^n|\tilde \Omega_n)(S) = \tilde \mu^n(S). \)

Then: \( \forall n \in [n_0, \infty), \quad (\pmb L^n|\tilde \Omega_n)(\tilde \Omega_n) = \tilde \mu^n(\tilde \Omega_n). \)

Dividing the last two equations gives:

\( \forall n \in [n_0, \infty), \forall S \subseteq \tilde \Omega_n, \quad (N(\pmb L^n|\tilde \Omega_n))(S) = (\tilde \mu^n(S))/\tilde \mu^n(\tilde \Omega_n)). \)

Then: \( \forall n \in [n_0, \infty), \forall S \subseteq \tilde \Omega_n, \quad \lambda_n(S) = (\tilde \mu^n(S))/z_n. \)

Recall: \( \forall n \in [n_0, \infty), \forall S \subseteq \Omega_n, \quad \lambda_n(S) = (\mu^n(S))/z_n. \)

**Claim 8:** Let \( n \in [n_0, \infty). \)

Then: \( \lambda_n\{f \in \tilde \Omega_n | \tilde f_n = \varepsilon_{\sigma_0}\} = k \cdot (\lambda_n\{f \in \Omega_n | f_n = \sigma_0\}). \)

**Proof of Claim 8:**

By Claim 7, \( \tilde \mu^n\{f \in \tilde \Omega_n | \tilde f_n = \varepsilon_{\sigma_0}\} = k \cdot (\mu^n\{f \in \Omega_n | f_n = \sigma_0\}). \)

Dividing this last equation by \( z_n \) yields:

\( \lambda_n\{f \in \Omega_n | f_n = \varepsilon_{\sigma_0}\} = k \cdot (\lambda_n\{f \in \Omega_n | f_n = \sigma_0\}). \)

**End of proof of Claim 8.**

**Let** \( P := \mu\{\sigma_0\} \) and \( \tilde P := \tilde \mu\{\varepsilon_{\sigma_0}\}. \)

By Claim 5, we have: \( \mu(\varepsilon^*\{\varepsilon_{\sigma_0}\}) = k \cdot (\mu\{\sigma_0\}). \)

Recall: \( \tilde \mu = \varepsilon^*\mu. \)

Since \( \tilde P = \tilde \mu\{\varepsilon_{\sigma_0}\} = (\varepsilon^*\mu)\{\varepsilon_{\sigma_0}\} = \mu(\varepsilon^*\{\varepsilon_{\sigma_0}\}) = k \cdot (\mu\{\sigma_0\}) = k \cdot P, \)

we get: \( \tilde P/k = P. \)

Recall: \( \mu = B^\varepsilon_\beta \) and \( \alpha = M\tilde \mu \) and \( \tilde \mu \in \mathcal{P}_E \) and \( S_{\tilde \mu} = E. \)

Recall: \( E \) is residue-unconstrained and \( |\tilde \mu|_2 < \infty. \)

Since \( \sigma_0 \in \Sigma \) and since \( E = I_{\varepsilon}, \) we get: \( \varepsilon(\sigma_0) \in E. \) Then \( \varepsilon_{\sigma_0} \in E. \)

Recall: \( \tilde P = \tilde \mu\{\varepsilon_{\sigma_0}\}. \)

By hypothesis, \( \alpha \in \mathbb{Z}. \)

By Theorem 11.2, as \( n \to \infty, \quad N(\tilde \mu^n|\tilde \Omega_n)\{\tilde f \in \tilde \Omega_n | \tilde f_n = \varepsilon_{\sigma_0}\} \to \tilde P. \)

Recall: \( \forall n \in [n_0, \infty), \quad \lambda_n = N(\tilde \mu^n|\tilde \Omega_n). \)

Then:

\( \lambda_n\{f \in \tilde \Omega_n | \tilde f_n = \varepsilon_{\sigma_0}\} \to \tilde P. \)

So, by Claim 8, as \( n \to \infty, \quad k \cdot (\lambda_n\{f \in \Omega_n | f_n = \sigma_0\}) \to \tilde P. \)

Then: \( \lambda_n\{f \in \Omega_n | f_n = \sigma_0\} \to \tilde P/k. \)

So, by Claim 3, as \( n \to \infty, \quad \nu_{\Omega_n}\{f \in \Omega_n | f_n = \sigma_0\} \to \tilde P/k. \)

So, since \( \tilde P/k = P = \mu\{\sigma_0\} = B^\varepsilon_\beta\{\sigma_0\}, \) we get:

\( \nu_{\Omega_n}\{f \in \Omega_n | f_n = \sigma_0\} \to B^\varepsilon_\beta\{\sigma_0\}. \) \( \Box \)

**THEOREM 21.2.** Let \( \Sigma \) be a finite set.

Let \( \varepsilon : \Sigma \to \mathbb{Z}. \) Assume \( I_{\varepsilon} \) is residue-unconstrained.
Let $\alpha \in \mathbb{Z}$. Assume $\alpha \in (\min \mathbb{I}_\varepsilon; \max \mathbb{I}_\varepsilon)$. Let $\beta := \mathcal{B}P_\varepsilon^\varepsilon$.

For all $n \in \mathbb{N}$, let $\Omega_n := \{ f \in \Sigma^n | (\varepsilon(f_1)) + \cdots + (\varepsilon(f_n)) = n\alpha \}$. Let $\sigma_0 \in \Sigma$. Then: as $n \to \infty$, $\nu_{\Omega_n} \{ f \in \Omega_n \mid f_n = \sigma_0 \} \to B_{\beta}^\varepsilon(\sigma_0)$.

Recall: $\nu_{\emptyset}(\emptyset) = -1$.

So, since $B_{\beta}^\varepsilon(\sigma_0) > 0$, part of the content of Theorem 21.2 is:

- $\forall$ sufficiently large $n \in \mathbb{N}$, $\Omega_n \neq \emptyset$.

See Claim 2 in the proof below.

**Example:** Suppose $\Sigma = \{0, 1, 10\}$ and $\alpha = 1$.

Suppose, also, $\forall \sigma \in \Sigma$, $\varepsilon(\sigma) = \sigma$.

Then $\Omega_N$ represents the set of all GFA dispensations to the $N$ professors.

Since $\nu_{\Omega_N}$ gives equal probability to each dispensation,

$\nu_{\Omega_N}$ represents the GFA’s first system for awarding grants.

Since $\beta = \mathcal{B}P_\varepsilon^\varepsilon = \mathcal{B}P_1^\varepsilon$, we calculate: $\beta = (\ln 9)/10$.

More calculation gives: $(B_{\beta}^\varepsilon(0), B_{\beta}^\varepsilon(1), B_{\beta}^\varepsilon(10)) = (1, 9^{-1/10}, 9^{-1})$.

Since $N$ is large, by Theorem 21.2, we get:

$\nu_{\Omega_N} \{ f \in \Omega_N \mid f_N = \sigma_0 \} \approx B_{\beta}^\varepsilon(\sigma_0)$.

So, if I am the $N$th professor, then, under the first system, my probability of receiving $\sigma_0$ dollars is approximately equal to $B_{\beta}^\varepsilon(\sigma_0)$.

Thus Theorem 21.2 reproduces the result of §12.

**Example:** Suppose $\Sigma = ([0..4] \times [0..4]) \setminus \{(4, 4)\}$.

Suppose, also, $\alpha = 1$ and $\forall \sigma \in \Sigma$, $\varepsilon(\sigma) = \sigma_1 + \sigma_2$.

Then $\Omega_N$ represents the set of all state-distributions at the BUA.

Since $\beta := \mathcal{B}P_\varepsilon^\varepsilon$, we get: $\beta \approx 1.0670$, accurate to four decimal places.

Let $M \in \mathbb{R}^{5 \times 5}$ be the matrix defined by: $M_{55} = 0$ and

$\forall (i, j) \in ([1..5] \times [1..5]) \setminus \{(5, 5)\}$, $M_{ij} = B_{\beta}^\varepsilon((i - 1, j - 1))$.

Then $M \approx \begin{bmatrix}
0.4345 & 0.1495 & 0.0514 & 0.0177 & 0.0061 \\
0.1495 & 0.0514 & 0.0177 & 0.0061 & 0.0021 \\
0.0514 & 0.0177 & 0.0061 & 0.0021 & 0.0007 \\
0.0177 & 0.0061 & 0.0021 & 0.0007 & 0.0002 \\
0.0061 & 0.0021 & 0.0007 & 0.0002 & 0
\end{bmatrix}$

all accurate to four decimal places.

(Thanks to C. Prouty for these calculations. See §25.)

According to Theorem 21.2, this answers
the problem formulated near the end of §19. Since \( B^0_{\beta}(0,0) = M_{11} = 0.4345 \), it is possible (cf. §14) to prove:

If \( N \) is sufficiently large, then, with probability > 99%, over 43% of the BUA professors have $0$ wealth.

\textbf{Proof.} DELETE THIS PROOF

Since \( \beta = BP^e_{\alpha} = (A_{\cdot}^e)^{-1}(\alpha) \), we get: 
\[ A_{\cdot}^e(\beta) = \alpha. \]
By Theorem 20.4, we have: 
\[ M_{\varepsilon B^0_{\beta}} = A_{\cdot}^e. \]
So, since \( A_{\cdot}^e = A_{\cdot}^e(\beta) = \alpha \), we get: 
\[ M_{\varepsilon B^0_{\beta}} = \alpha. \]
Let \( \mu := B^0_{\beta} \). Then: 
\[ \mu \in \mathcal{P}_{\Sigma} \] and \( M_{\varepsilon \mu} = \alpha. \)
Let \( E := \mathbb{I}_e, \tilde{\mu} := \varepsilon_{\ast \mu}. \) Then: 
\[ \tilde{\mu} \in \mathcal{P}_E \] and \( M_{\tilde{\mu}} = \alpha. \)
By hypothesis, \( E \) is residue-unconstrained.

Since \( \varepsilon : \Sigma \rightarrow \mathbb{Z} \), we get: 
\[ E \subseteq \mathbb{Z}. \]
Since \( \Sigma \) is finite, we get: 
\[ E \text{ is finite}. \]
So, since \( \tilde{\mu} \in \mathcal{P}_E \subseteq \mathcal{F}_M_E \), we get: 
\[ |\tilde{\mu}|_1 < \infty \text{ and } |\tilde{\mu}|_2 < \infty. \]
For all \( \sigma \in \Sigma \), let \( \varepsilon_{\sigma} := \varepsilon(\sigma). \)
Then: 
\[ \forall n \in \mathbb{N}, \quad \Omega_n = \{ f \in \Sigma^n \mid \varepsilon_{f_1} + \cdots + \varepsilon_{f_n} = n\alpha \}. \]
For all \( n \in \mathbb{N} \), define \( \varepsilon^n : \Sigma^n \rightarrow E^n \) by:
\[ \forall f_1, \ldots, f_n \in \Sigma, \quad \varepsilon^n(f_1, \ldots, f_n) = (\varepsilon_{f_1}, \ldots, \varepsilon_{f_n}). \]
Then, since \( \varepsilon_{\ast \mu} = \tilde{\mu} \), we get: 
\[ \forall n \in \mathbb{N}, \quad (\varepsilon^n)_{\ast \mu} = \tilde{\mu}^n. \]
For all \( n \in \mathbb{N} \), let 
\[ \tilde{\Omega}_n := \{ \tilde{f} \in E^n \mid \tilde{f}_1 + \cdots + \tilde{f}_n = n\alpha \}; \]
then 
\[ (\varepsilon^n)_{\ast \tilde{\Omega}_n} = \Omega_n. \]
Then: 
\[ \forall n \in \mathbb{N}, \quad (\varepsilon^n)_{\ast \mu}(\tilde{\Omega}_n) = \mu^n(\Omega_n). \]
Then: 
\[ \forall n \in \mathbb{N}, \quad ((\varepsilon^n)_{\ast \mu})_{\ast \mu}(\tilde{\Omega}_n) = \mu^n(\Omega_n). \]
Then: 
\[ \forall n \in \mathbb{N}, \quad \tilde{\mu}^n(\Omega_n) = \mu^n(\Omega_n). \]
For all \( n \in \mathbb{N} \), define \( \psi_n : \mathbb{Z} \rightarrow \mathbb{R} \) by:
\[ \forall t \in \mathbb{Z}, \quad \psi_n(t) = \tilde{\mu}^n(\tilde{f} \in E^n \mid \tilde{f}_1 + \cdots + \tilde{f}_n = t). \]
Then: 
\[ \forall n \in \mathbb{N}, \quad \psi_n(n\alpha) = \tilde{\mu}^n(\tilde{\Omega}_n). \]
Since \( E \) is finite and residue-unconstrained, we get: 
\[ 2 \leq \#E < \infty. \]
Since \( \mu = B^0_{\beta} \) and \( S_{B^0_{\beta}} = \Sigma \), we get: 
\[ S_{\mu} = \Sigma. \]
So, since \( \varepsilon : \Sigma \rightarrow \mathbb{Z} \), we get: 
\[ S_{\varepsilon \mu} = \mathbb{I}_e \subseteq \mathbb{Z}. \]
Recall: \( \varepsilon_{\ast \mu} = \tilde{\mu} \) and \( \mathbb{I}_e = E \). Then: 
\[ S_{\tilde{\mu}} = E. \]
Since \( E \) is finite and \( \tilde{\mu} \in \mathcal{P}_E \) and \( |\tilde{\mu}|_1 < \infty \) and \( \#S_{\tilde{\mu}} = \#E = 2 \),
by Theorem 8.6, we get \( V_{\tilde{\mu}} > 0 \).
So, since \( V_{\tilde{\mu}} = |\tilde{\mu}|_2 - M_{\tilde{\mu}}^2 \leq |\tilde{\mu}|_2 < \infty \), we conclude:
\[ 0 < V_{\tilde{\mu}} < \infty. \]
Let \( v := V_{\tilde{\mu}}. \) Then 
\[ 0 < v < \infty. \] Then \( 1/\sqrt{2\pi v} > 0. \)
Let \( \tau := 1/\sqrt{2\pi v}. \) Then \( \tau > 0. \)
Claim 1: As $n \to \infty$, $\sqrt{n} \cdot (\psi_n(n\alpha)) \to \tau$.

Proof of Claim 1: We have:

$$\forall n \in \mathbb{N}, \; \psi_n(n\alpha) = \tilde{\mu}^n \{ \tilde{f} \in E^n \mid \tilde{f}_1 + \cdots + \tilde{f}_n = n\alpha \}. $$

Want: as $n \to \infty$, $\sqrt{n} \cdot (\tilde{\mu}^n \{ \tilde{f} \in E^n \mid \tilde{f}_1 + \cdots + \tilde{f}_n = n\alpha \}) \to 1/\sqrt{2\pi v}$.

Recall: $E \subseteq \mathbb{Z}$, $E$ is residue-unconstrained, $\tilde{\mu} \in \mathcal{P}_E$, $S_{\tilde{\mu}} = E$, $|\tilde{\mu}| < \infty$, $\alpha = M_{\tilde{\mu}}$, $v = V_{\tilde{\mu}}$, $\alpha \in \mathbb{Z}$.

Then, by Theorem 9.10, we get:

$$\sqrt{n} \cdot (\tilde{\mu}^n \{ \tilde{f} \in E^n \mid \tilde{f}_1 + \cdots + \tilde{f}_n = n\alpha \}) \to 1/\sqrt{2\pi v}.$$ 

End of proof of Claim 1.

Since $\tau > 0$, by Claim 1, choose $n_0 \in [2, \infty)$ s.t. $\forall n \in [n_0, \infty)$, $\sqrt{n} \cdot (\psi_n(n\alpha)) > 0$.

Claim 2: Let $n \in [n_0, \infty)$. Then: $\mu^n(\Omega_n) > 0$.

Proof of Claim 2: Recall: $\tilde{\mu}^n(\tilde{\Omega}_n) = \mu^n(\Omega_n)$ and $\psi_n(n\alpha) = \tilde{\mu}^n(\tilde{\Omega}_n)$.

By the choice of $n_0$, we get: $\sqrt{n} \cdot (\psi_n(n\alpha)) > 0$. Then: $\psi_n(n\alpha) > 0$.

Then: $\mu^n(\Omega_n) = \tilde{\mu}^n(\tilde{\Omega}_n) = \psi_n(n\alpha) > 0$.

End of proof of Claim 2.

Since $\mu \in \mathcal{P}_\Sigma$, we get: $\forall n \in \mathbb{N}$, $\mu^n \in \mathcal{P}_{\Sigma^n}$, so $\mu^n(\Omega_n) \leq 1$.

So, by Claim 2, $\forall n \in [n_0, \infty)$, $0 < \mu^n(\Omega_n) \leq 1$.

Also, we have: $\forall n \in \mathbb{N}$, $(\mu^n|\Omega_n)(\Omega_n) = \mu^n(\Omega_n)$.

Then: $\forall n \in [n_0, \infty)$, $0 < (\mu^n|\Omega_n)(\Omega_n) \leq 1$.

Then: $\forall n \in [n_0, \infty)$, $\mu^n|\Omega_n \in \mathcal{F} \mathcal{M}_{\Omega_n}^\times$.

Then: $\forall n \in [n_0, \infty)$, $\mathcal{N}(\mu^n|\Omega_n) \in \mathcal{P}_{\Omega_n}$.

Since $\#E \geq 2$, we get: $E \neq \emptyset$.

Since $\varepsilon : \Sigma \to \mathbb{Z}$ and since $\mathbb{I}_\varepsilon = E \neq \emptyset$, we get: $\Sigma \neq \emptyset$.

Then $\hat{B}_\varepsilon(\Sigma) > 0$. Let $C := 1/(\hat{B}_\varepsilon(\Sigma))$. Then $\mathcal{N}(\hat{B}_\varepsilon) = C \cdot \hat{B}_\varepsilon$.

By definition of $\hat{B}_\varepsilon^e$, we have: $\forall \sigma \in \Sigma$, $\hat{B}_\varepsilon^e \{ \sigma \} = e^{-\beta \varepsilon \sigma}$.

So, since $\mu = B_\varepsilon^e = \mathcal{N}(\hat{B}_\varepsilon^e) = C \cdot \hat{B}_\varepsilon^e$, we get: $\forall \sigma \in \Sigma$, $\mu \{ \sigma \} = C e^{-\beta \varepsilon \sigma}$.

We have: $\forall n \in \mathbb{N}$, $\forall S \subseteq \Omega_n$, $(\mu^n|\Omega_n)(S) = \mu^n(S)$.

Then: $\forall n \in \mathbb{N}$, $(\mu^n|\Omega_n)(\Omega_n) = \mu^n(\Omega_n)$.

For all $n \in \mathbb{N}$, let $z_n := \mu^n(\Omega_n)$.

Then, by Claim 2, we have: $\forall n \in [n_0, \infty)$, $z_n > 0$.

Also, since $\forall n \in [n_0, \infty)$, $\mu^n(\Omega_n) = \tilde{\mu}^n(\tilde{\Omega}_n)$,
we conclude: \( \forall n \in [n_0, \infty), \quad z_n = \mu^n(\Omega_n). \)

For all \( n \in [n_0, \infty), \) let \( \lambda_n := \mathcal{N}(\mu^n|\Omega_n). \)

Then: \( \forall n \in [n_0, \infty), \quad \lambda_n = (\mu^n|\Omega_n)/(\mu^n(\Omega_n)). \)

Then: \( \forall n \in [n_0, \infty), \quad \lambda_n = (\mu^n|\Omega_n)/z_n. \)

Then: \( \forall n \in [n_0, \infty), \forall S \subseteq \Omega_n, \quad \lambda_n(S) = (\mu^n(S))/z_n. \)

Claim 3: Let \( n \in [n_0, \infty). \) Then: \( \lambda_n = \nu_{\Omega_n}. \)

Proof of Claim 3: Let \( F := \Omega_n. \) Want: \( \lambda_n = \nu_F. \)

Since \( \lambda_n = \mathcal{N}(\mu^n|\Omega_n) = \mathcal{N}(\mu^n|F), \) we get: \( \lambda_n \in \mathcal{P}_F. \)

By Theorem 8.9, given \( f, g \in F, \) want: \( \lambda_n\{f\} = \lambda_n\{g\}. \)

Want: \( (\mu^n\{f\})/z_n = (\mu^n\{g\})/z_n. \)

Want: \( \mu^n\{f\} = \mu^n\{g\}. \)

For all \( i \in [1..n], \) let \( \tilde{f}_i := \varepsilon_{f_i} \) and \( \tilde{g}_i := \varepsilon_{g_i}. \)

Then: \( \forall i \in [1..n], \mu\{\tilde{f}_i\} = Ce^{-\beta_{\tilde{f}_i}} \) and \( \mu\{g_i\} = Ce^{-\beta_{\tilde{g}_i}}. \)

Since \( f \in F = \Omega_n, \) we get: \( \varepsilon_{f_1} + \cdots + \varepsilon_{f_n} = n\alpha. \)

Since \( g \in F = \Omega_n, \) we get: \( \varepsilon_{g_1} + \cdots + \varepsilon_{g_n} = n\alpha. \)

Since \( \tilde{f}_1 + \cdots + \tilde{f}_n = \varepsilon_{f_1} + \cdots + \varepsilon_{f_n} = n\alpha \)

we get: \( Ce^{-\beta_{\tilde{f}_1}} \cdots Ce^{-\beta_{\tilde{f}_n}} = (Ce^{-\beta_{\tilde{g}_1}} \cdots (Ce^{-\beta_{\tilde{g}_n}}). \)

Then: \( (\mu\{\tilde{f}_1\}) \cdots (\mu\{f_n\}) = (\mu\{g_1\}) \cdots (\mu\{g_n\}). \)

Then: \( \mu^n\{f\} = \mu^n\{g\}. \)

End of proof of Claim 3.

Claim 4: Let \( \sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}. \) Then: \( \mu\{\sigma\} = \mu\{\sigma_0\}. \)

Proof of Claim 4: Since \( \sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}, \) we get: \( \varepsilon(\sigma) \in \varepsilon_{\sigma_0}. \)

Since \( \varepsilon_{\sigma} = \varepsilon(\sigma) \in \varepsilon_{\sigma_0}, \) we get: \( \varepsilon_{\sigma} = \varepsilon_{\sigma_0}. \)

Then: \( \mu\{\sigma\} = Ce^{-\beta_{\varepsilon_{\sigma}}} = Ce^{-\beta_{\varepsilon_{\sigma_0}}} = \mu\{\sigma_0\}. \)

End of proof of Claim 4.

Since \( \varepsilon(\sigma_0) = \varepsilon_{\sigma_0} \in \varepsilon_{\sigma_0}, \) we get: \( \sigma_0 \in \varepsilon^*\{\varepsilon_{\sigma_0}\}. \)

Then \( \varepsilon^*\{\varepsilon_{\sigma_0}\} \neq \varnothing, \) so \( \#(\varepsilon^*\{\varepsilon_{\sigma_0}\}) \geq 1. \)

Let \( k := \#(\varepsilon^*\{\varepsilon_{\sigma_0}\}). \) Then: \( k \geq 1. \)

Claim 5: \( \mu(\varepsilon^*\{\varepsilon_{\sigma_0}\}) = k \cdot (\mu\{\sigma_0\}). \)

Proof of Claim 5: Since \( \varepsilon^*\{\varepsilon_{\sigma_0}\} \) is equal to

the disjoint union, over \( \sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}, \) of \( \{\sigma\}, \)

we get: \( \mu(\varepsilon^*\{\varepsilon_{\sigma_0}\}) = \sum_{\sigma \in \varepsilon^*\{\varepsilon_{\sigma_0}\}} [\mu(\sigma)]. \)
So, by Claim 4, we get: \( \mu(\varepsilon^s(\varepsilon_{\sigma_0})) = k \cdot (\mu(\sigma_0)) \).

End of proof of Claim 5.

Claim 6: Let \( n \in [n_0, \infty) \). Let \( \sigma \in \varepsilon^s(\varepsilon_{\sigma_0}) \).

Then: \( \mu^n\{f \in \Omega^n \mid f_n = \sigma\} = \mu^n\{f \in \Omega^n \mid f_n = \sigma_0\} \).

Proof of Claim 6: By choice of \( n_0 \), we have: \( n_0 \in [2, \infty) \).

Then \( [n_0, \infty) \subseteq [2, \infty) \), so, since \( n \in [n_0, \infty) \), we get: \( n \in [2, \infty) \).

Let \( X := \{f \in \Sigma^{n-1} \mid \varepsilon_{f_1} + \cdots + \varepsilon_{f_{n-1}} = n\alpha - \varepsilon_\sigma\} \).

We have \( \{f \in \Omega^n \mid f_n = \sigma\} \) (Claim 4)

Let \( \{f \in \Omega^n \mid f_n = \sigma_0\} \) (Claim 5)

proof of Claim 6.

Claim 7: Let \( n \in [n_0, \infty) \).

Then: \( \tilde{\mu}^n\{\tilde{f} \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0}\} = k \cdot (\mu^n\{f \in \Omega^n \mid f_n = \sigma_0\}) \).

Proof of Claim 7: Recall: \( \tilde{\mu}^n = (\varepsilon^n)_s(\mu^n) \).

Since \( (\varepsilon^n)_s\{\tilde{f} \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0}\} = \{f \in \Omega^n \mid f_n \in \varepsilon^s(\varepsilon_{\sigma_0})\} \),

we get: \( \mu^n((\varepsilon^n)_s\{\tilde{f} \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0}\}) = \mu^n\{f \in \Omega^n \mid f_n \in \varepsilon^s(\varepsilon_{\sigma_0})\} \).

Then: \( ((\varepsilon^n)_s(\mu^n))\{\tilde{f} \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0}\} = \mu^n\{f \in \Omega^n \mid f_n \in \varepsilon^s(\varepsilon_{\sigma_0})\} \).
Then: \[ \tilde{\mu}^n \{ f \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0} \} = \mu^n \{ f \in \Omega_n \mid f_n = \varepsilon_{\sigma_0} \}. \]

**Want:** \[ \mu^n \{ f \in \Omega_n \mid f_n \in \varepsilon^* \{ \varepsilon_{\sigma_0} \} \} = k \cdot (\mu^n \{ f \in \Omega_n \mid f_n = \sigma_0 \}). \]

Since \( \{ f \in \Omega_n \mid f_n \in \varepsilon^* \{ \varepsilon_{\sigma_0} \} \} \) is the disjoint union, over \( \sigma \in \varepsilon^* \{ \varepsilon_{\sigma_0} \} \), of \( \{ f \in \Omega_n \mid f_n = \sigma \} \), we get: \[ \mu^n \{ f \in \Omega_n \mid f_n \in \varepsilon^* \{ \varepsilon_{\sigma_0} \} \} = \sum_{\sigma \in \varepsilon^* \{ \varepsilon_{\sigma_0} \}} (\mu^n \{ f \in \Omega_n \mid f_n = \sigma \}). \]

So, since \( k = \#(\varepsilon^* \{ \varepsilon_{\sigma_0} \}) \), by Claim 6, we get:

\[ \mu^n \{ f \in \Omega_n \mid f_n \in \varepsilon^* \{ \varepsilon_{\sigma_0} \} \} = k \cdot (\mu^n \{ f \in \Omega_n \mid f_n = \sigma_0 \}). \]

**End of proof of Claim 7.**

Recall: \( \forall n \in [n_0, \infty), \lambda_n = N(\mu^n | \Omega_n). \)

For all \( n \in [n_0, \infty) \), let \( \tilde{\lambda}_n := N(\tilde{\mu}^n | \tilde{\Omega}_n). \)

Since, \( \forall n \in N, (\varepsilon^n)_* (\mu^n) = \tilde{\mu}^n \) and \( (\varepsilon^n)^* \tilde{\Omega}_n = \Omega_n, \)

we get \( \forall n \in N, (\varepsilon^n)_* (\mu^n | \Omega_n) = \tilde{\mu}^n | \tilde{\Omega}_n, \)

so \( \forall n \in [n_0, \infty), (\varepsilon^n)_* (\mathcal{N}(\mu^n | \Omega_n)) = \mathcal{N}(\tilde{\mu}^n | \tilde{\Omega}_n). \)

Then: \( \forall n \in [n_0, \infty), (\varepsilon^n)_* (\lambda_n) = \tilde{\lambda}_n. \)

Recall: \( \forall n \in [n_0, \infty), \lambda_n = \mathcal{N}(\tilde{\mu}^n | \tilde{\Omega}_n), z_n = \tilde{\mu}^n (\tilde{\Omega}_n) \).

We have: \( \forall n \in [n_0, \infty), \forall S \subseteq \tilde{\Omega}_n, (\tilde{\mu}^n | \tilde{\Omega}_n)(S) = \tilde{\mu}^n(S). \)

Then: \( \forall n \in [n_0, \infty), \forall S \subseteq \tilde{\Omega}_n, (\tilde{\mu}^n | \tilde{\Omega}_n)(\tilde{\Omega}_n) = \tilde{\mu}^n(\tilde{\Omega}_n). \)

Dividing the last two equations gives:

\( \forall n \in [n_0, \infty), \forall S \subseteq \tilde{\Omega}_n, \mathcal{N}(\tilde{\mu}^n | \tilde{\Omega}_n)(S) = (\tilde{\mu}^n(S))/(\tilde{\mu}^n(\tilde{\Omega}_n)). \)

Then: \( \forall n \in [n_0, \infty), \forall S \subseteq \tilde{\Omega}_n, \lambda_n(S) = (\tilde{\mu}^n(S))/z_n. \)

Recall: \( \forall n \in [n_0, \infty), \forall S \subseteq \Omega_n, \lambda_n(S) = (\mu^n(S))/z_n. \)

**Claim 8:** Let \( n \in [n_0, \infty). \)

Then: \( \tilde{\lambda}_n \{ f \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0} \} = k \cdot (\lambda_n \{ f \in \Omega_n \mid f_n = \sigma_0 \}). \)

**Proof of Claim 8:**

By Claim 7, \( \tilde{\mu}^n \{ f \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0} \} = k \cdot (\mu^n \{ f \in \Omega_n \mid f_n = \sigma_0 \}). \)

Dividing this last equation by \( z_n \) yields

\( \tilde{\lambda}_n \{ f \in \tilde{\Omega}_n \mid \tilde{f}_n = \varepsilon_{\sigma_0} \} = k \cdot (\lambda_n \{ f \in \Omega_n \mid f_n = \sigma_0 \}). \)

**End of proof of Claim 8.**

**Let** \( P := \mu \{ \sigma_0 \} \) and \( \tilde{P} := \tilde{\mu} \{ \varepsilon_{\sigma_0} \}. \)

By Claim 5, we have: \( \mu(\varepsilon^* \{ \varepsilon_{\sigma_0} \}) = k \cdot (\mu \{ \sigma_0 \}). \)

Recall: \( \tilde{\mu} = \varepsilon_* \mu. \)

Since \( \tilde{P} = \tilde{\mu} \{ \varepsilon_{\sigma_0} \} = (\varepsilon_* \mu) \{ \varepsilon_{\sigma_0} \} = \mu(\varepsilon^* \{ \varepsilon_{\sigma_0} \}) = k \cdot (\mu \{ \sigma_0 \}) = k \cdot P, \)

we get: \( \tilde{P}/k = P. \)
Recall: \(\mu = B_\beta^\varepsilon\) and \(\alpha = M_\mu\) and \(\tilde{\mu} \in \mathcal{P}_E\) and \(S_\mu = E\).

Recall: \(E\) is residue-unconstrained and \(|\tilde{\mu}|_2 < \infty\).

Since \(\sigma_0 \in \Sigma\) and since \(E = \mathbb{I}_\varepsilon\), we get: \(\varepsilon(\sigma_0) \in E\). Then \(\varepsilon(\sigma_0) \in E\).

Recall: \(\tilde{P} = \tilde{\mu}\{\varepsilon_{\sigma_0}\}\). By hypothesis, \(\alpha \in \mathbb{Z}\).

By Theorem 11.6, as \(n \to \infty\), \(N(\tilde{\mu}^n|\tilde{\Omega}_n)\{f \in \tilde{\Omega}_n | \tilde{f}_n = \varepsilon_{\sigma_0}\} \to \tilde{P}\).

Recall: \(\forall n \in [n_0,\infty), \tilde{\lambda}_n = N(\tilde{\mu}_n|\tilde{\Omega}_n)\).

Then: as \(n \to \infty\), \(\tilde{\lambda}_n\{f \in \tilde{\Omega}_n | \tilde{f}_n = \varepsilon_{\sigma_0}\} \to \tilde{P}\).

So, by Claim 8, as \(n \to \infty\), \(k \cdot (\lambda_n\{f \in \Omega_n | f_n = \sigma_0\}) \to \tilde{P}/k\).

Then: as \(n \to \infty\), \(\lambda_n\{f \in \Omega_n | f_n = \sigma_0\} \to \tilde{P}/k\).

So, by Claim 3, as \(n \to \infty\), \(\nu\{f \in \Omega_n | f_n = \sigma_0\} \to \tilde{P}/k\).

So, since \(\tilde{P}/k = P = \mu\{\sigma_0\} = B_\beta^\varepsilon\{\sigma_0\}\), we get:

\[ \text{as } n \to \infty, \nu\{f \in \Omega_n | f_n = \sigma_0\} \to B_\beta^\varepsilon\{\sigma_0\}. \]

\[ \Box \]

Here is another approach to proving Theorem 21.1:

By density of the set of injective functions \(\Sigma \to \mathbb{R}\) in the topological space of all functions \(\Sigma \to \mathbb{R}\), we reduce to the case where \(\varepsilon\) is injective.

Then the proof can follow the proof of Theorem 17.5, which avoids many of the complications in the proof of Theorem 21.1 given above.

22. \(\infty\)-properness

**Definition 22.1.** Let \(\Sigma\) be a set, \(\varepsilon : \Sigma \to \mathbb{R}\).

By \(\varepsilon\) is \(\infty\)-proper, we mean: \(\forall t \in \mathbb{R}, \#\{\sigma \in \Sigma | \varepsilon(\sigma) \leq t\} < \infty\).

Note that, for any finite set \(\Sigma\), for any \(\varepsilon : \Sigma \to \mathbb{R}\), we have: \(\varepsilon\) is \(\infty\)-proper.

**Definition 22.2.** Let \(\Sigma\) be a set, \(\varepsilon : \Sigma \to \mathbb{R}\).

Then \([\text{IP}_\varepsilon]\) denotes the set of sequences \((\sigma_1, \sigma_2, \ldots)\) s.t.

\(\Sigma = \{\sigma_1, \sigma_2, \ldots\}\) and \(\sigma_1, \sigma_2, \ldots\) are distinct

and \(\varepsilon(\sigma_1) \leq \varepsilon(\sigma_2) \leq \cdots\) and as \(n \to \infty\), \(\varepsilon(\sigma_n) \to \infty\).

In the preceding definition, “\(\text{IP}\)” stands for “Infinity-proper”.

Sequences in \([\text{IP}_\varepsilon]\) will be called: “\(\infty\)-proper sequences for \(\varepsilon\)”.

Note: If \([\text{IP}_\varepsilon] \neq \emptyset\), then \(\Sigma\) is countably infinite.

The following is basic, and we omit the proof.

**Theorem 22.3.** Let \(\Sigma\) be an infinite set, \(\varepsilon : \Sigma \to \mathbb{R}\).

Then: \((\varepsilon\text{ is }\infty\text{-proper}) \iff ([\text{IP}_\varepsilon] \neq \emptyset)\).
In particular, for any set \( \Sigma \), for any \( \varepsilon : \Sigma \to \mathbb{R} \), we have: if \( \varepsilon \) is \( \infty \)-proper, then \( \Sigma \) is countable.

**THEOREM 22.4.** Let \( \Sigma \) be a nonempty set, \( \varepsilon : \Sigma \to \mathbb{R} \).
Assume: \( \varepsilon \) is \( \infty \)-proper. Then: \( \exists \zeta_0 \in \mathbb{I}_{\varepsilon} \text{ s.t., } \forall \sigma \in \Sigma, \varepsilon(\sigma) \geq \zeta_0 \).

Proof. Since \( \Sigma \neq \emptyset \), we get: \( \mathbb{I}_{\varepsilon} \neq \emptyset \).
In case \( \mathbb{I}_{\varepsilon} \) is finite, let \( \zeta_0 := \min \mathbb{I}_{\varepsilon} \).
We therefore assume: \( \mathbb{I}_{\varepsilon} \) is infinite. Then \( \Sigma \) is infinite.
By Theorem 22.3, choose \( \sigma_1, \sigma_2, \ldots \in \mathbb{I}_{\varepsilon} \).
Let \( \zeta_0 := \varepsilon(\sigma_1) \). Then \( \zeta_0 = \varepsilon(\sigma_1) \in \mathbb{I}_{\varepsilon} \).
Given \( \sigma \in \Sigma \), want: \( \varepsilon(\sigma) \geq \zeta_0 \).
Since \( \sigma \in \Sigma = \{\sigma_1, \sigma_2, \ldots\} \), choose \( i \in \mathbb{N} \) s.t. \( \sigma = \sigma_i \).
Then: \( \varepsilon(\sigma) = \varepsilon(\sigma_i) \geq \varepsilon(\sigma_{i-1}) \geq \cdots \geq \varepsilon(\sigma_1) = \zeta_0 \).

**THEOREM 22.5.** Let \( \Sigma \) be a set. Let \( \varepsilon : \Sigma \to \mathbb{R} \) be \( \infty \)-proper.
Then: \( \mathbb{I}_{\varepsilon} \) is bounded below and \( \forall t \in \mathbb{I}_{\varepsilon}, \varepsilon^*\{t\} \) is finite.

The preceding is basic; we omit the proof.
When \( \varepsilon \) is \( \mathbb{Z} \)-valued, the converse is also true:

**THEOREM 22.6.** Let \( \Sigma \) be a set. Let \( \varepsilon : \Sigma \to \mathbb{Z} \).
Then: \( [ \varepsilon \text{ is } \infty \text{-proper} ] \) \iff \( [ ( \mathbb{I}_{\varepsilon} \text{ is bounded below} ) \& ( \forall t \in \mathbb{I}_{\varepsilon}, \varepsilon^*\{t\} \text{ is finite} ) ] \).

The preceding is basic; we omit the proof.
The following two results are corollaries.

**THEOREM 22.7.** Let \( \Sigma \) be a set. Let \( \varepsilon : \Sigma \to \mathbb{Z} \) be injective.
Then: \( [ \varepsilon \text{ is } \infty \text{-proper} ] \) \iff \( [ \mathbb{I}_{\varepsilon} \text{ is bounded below} ] \).

**THEOREM 22.8.** Let \( \Sigma \subseteq \mathbb{Z} \).
Define \( \varepsilon : \Sigma \to \mathbb{R} \) by: \( \forall \sigma \in \Sigma, \varepsilon(\sigma) = \sigma \).
Then: \( [ \varepsilon \text{ is } \infty \text{-proper} ] \) \iff \( [ \Sigma \text{ is bounded below} ] \).

23. **Boltzmann distributions on countable sets**

We begin by generalizing our work on Boltzmann distributions to allow for countably infinite sets of states.

**DEFINITION 23.1.** Let \( \Sigma \) be a nonempty countable set, \( \varepsilon : \Sigma \to \mathbb{R} \).
Then, \( \forall \beta \in \mathbb{R}, \hat{B}_\beta^\varepsilon \in \mathcal{M}_\Sigma \) is defined by: \( \forall \sigma \in \Sigma, \hat{B}_\beta^\varepsilon(\sigma) = e^{-\beta \cdot \varepsilon(\sigma)} \).
**DEFINITION 23.2.** Let $\Sigma$ be a set, $\varepsilon : \Sigma \to \mathbb{R}$, $\beta \in \mathbb{R}$.

Then:

$$\Delta_\beta^\varepsilon := \sum_{\sigma \in \Sigma} [e^{-\beta(\varepsilon(\sigma))}] \in [0; \infty].$$

Let $\Sigma$ be a set, $\varepsilon : \Sigma \to \mathbb{R}$.

If $\Sigma = \emptyset$, then: $\varepsilon$ is the empty function and, $\forall \beta \in \mathbb{R}$, $\Delta_\beta^\varepsilon = 0$.

If $\Sigma \neq \emptyset$, then:

- $\forall \beta \in \mathbb{R}$, $\Delta_\beta^\varepsilon > 0$.
- $\forall \beta \in \mathbb{R}$, $\Delta_\beta^\varepsilon = \infty$.

Let $\Sigma$ be a nonempty countable set, $\varepsilon : \Sigma \to \mathbb{R}$.

Then:

$$\forall \beta \in \mathbb{R}, \quad \Delta_\beta^\varepsilon = \sum_{\sigma \in \Sigma} [\hat{B}_\beta^\varepsilon(\sigma)], \quad \text{so} \quad \Delta_\beta^\varepsilon = \hat{B}_\beta(\Sigma).$$

**DEFINITION 23.3.** Let $\Sigma$ be a set, $\varepsilon : \Sigma \to \mathbb{R}$.

Then the **Delta-finite-set** of $\varepsilon$ is:

$$DF_\varepsilon := \{ \beta \in \mathbb{R} | \Delta_\beta^\varepsilon < \infty \}.$$

Let $\Sigma$ be a countable set, $\varepsilon : \Sigma \to \mathbb{R}$.

If $\Sigma = \emptyset$, $\varepsilon$ is the empty function, and there is nothing to say.

If $\Sigma$ is nonempty and finite,

- we have already ($\S$20) developed a satisfactory Boltzmann theory.

We will therefore focus on the case: $\Sigma$ is infinite.

Since $\varepsilon : \Sigma \to \mathbb{R}$, we get: $\varepsilon^* \mathbb{R} = \Sigma$.

Since $(-\infty; 0] \cup [0; \infty) = \mathbb{R}$, we get: $(\varepsilon^* (-\infty; 0]) \cup (\varepsilon^* [0; \infty)) = \varepsilon^* \mathbb{R}$.

Since $(\varepsilon^* (-\infty; 0]) \cup (\varepsilon^* [0; \infty)) = \varepsilon^* \mathbb{R} = \Sigma$,

assuming $\Sigma$ is infinite, we conclude:

either $\varepsilon^* (-\infty; 0]$ is infinite or $\varepsilon^* [0; \infty)$ is infinite.

Depending on which of these is true, the Boltzmann theory will follow one of two parallel lines of development.

Both lines are similar to one another, and replacing $\varepsilon$ by $-\varepsilon$ interchanges them.

In the sequel, we will only pursue the line $\varepsilon^* [0; \infty)$ is infinite.

**THEOREM 23.4.** Let $\Sigma$ be a set, $\varepsilon : \Sigma \to \mathbb{R}$.

Assume: $\varepsilon^* [0; \infty)$ is infinite. Then: $DF_\varepsilon \subseteq (0; \infty)$.

*Proof.* Given $\beta \in DF_\varepsilon$, **want:** $\beta \in (0; \infty)$. **Want:** $\beta > 0$.

Assume $\beta \leq 0$.

**Want:** Contradiction.

For all $\sigma \in \Sigma$, let $\varepsilon_\sigma := \varepsilon(\sigma)$.

We have: $\forall \sigma \in \varepsilon^* [0; \infty), \varepsilon_\sigma \geq 0$.

So, since $\beta \leq 0$, we get: $\forall \sigma \in \varepsilon^* [0; \infty), -\beta \varepsilon_\sigma \geq 0$.

Then: $\forall \sigma \in \varepsilon^* [0; \infty), e^{-\beta \varepsilon_\sigma} \geq 1$.

So, since $\varepsilon^* [0; \infty)$ is infinite, we get: $\sum_{\sigma \in \varepsilon^* [0; \infty)} [e^{-\beta \varepsilon_\sigma}] = \infty$. 


Since \[ \Delta_{\beta}^\varepsilon = \sum_{\sigma \in \Sigma} [e^{-\beta \varepsilon \sigma}] \geq \sum_{\sigma \in \varepsilon [0; \infty]} [e^{-\beta \varepsilon \sigma}] = \infty, \]
we get: \[ \beta \notin \text{DF}_\varepsilon. \] Contradiction. □

Let \( \Sigma \) be a nonempty set, let \( \varepsilon : \Sigma \to \mathbb{R} \).
If \( \text{DF}_\varepsilon = \emptyset \), then, for all \( \beta \in \mathbb{R} \), since \( \hat{B}_\beta^\varepsilon (\Sigma) = \Delta_{\beta}^\varepsilon = \infty \),
we see that \( \hat{B}_\beta^\varepsilon \) cannot be normalized, i.e., there is no \( B_\beta^\varepsilon \).
So, if \( \text{DF}_\varepsilon = \emptyset \), we have no Boltzmann distributions to study.
So, going forward, we will only consider situations where \( \text{DF}_\varepsilon \neq \emptyset \).
So, since we are only looking at cases in which \( \varepsilon^* [0; \infty) \) is infinite,
the next result tells us that, in the situations of interest to us,
\( \varepsilon \) will be \( \infty \)-proper,
which implies that \( \Sigma \) is countable.

**THEOREM 23.5.** Let \( \Sigma \) be a set, \( \varepsilon : \Sigma \to \mathbb{R} \).
Assume: \( \varepsilon^* [0; \infty) \) is infinite and \( \text{DF}_\varepsilon \neq \emptyset \).
Then: \( \varepsilon \) is \( \infty \)-proper.

**Proof.** Given \( t \in \mathbb{R} \), let \( \Sigma_0 := \{ \sigma \in \Sigma | \varepsilon(\sigma) \leq t \} \), want: \#\( \Sigma_0 \) < \( \infty \).
Since \( \text{DF}_\varepsilon \neq \emptyset \), choose \( \beta \in \text{DF}_\varepsilon \). Then: \( \Delta_{\beta}^\varepsilon < \infty \).
It suffices to show: \( (\#\Sigma_0) \cdot e^{-\beta t} < \infty \).
By Theorem 23.4, \( \text{DF}_\varepsilon \subseteq (0; \infty) \). So, since \( \beta \in \text{DF}_\varepsilon \), we get: \( \beta > 0 \).
For all \( \sigma \in \Sigma \), let \( \varepsilon_\sigma := \varepsilon(\sigma) \).
Then: \( \Delta_{\beta}^\varepsilon = \sum_{\sigma \in \Sigma} [e^{-\beta \varepsilon_\sigma}] \).
Since \( \beta > 0 \) and since \( \forall \sigma \in \Sigma_0 \), \( t \geq \varepsilon(\sigma) = \varepsilon_\sigma \),
we get: \( \forall \sigma \in \Sigma_0, -\beta \cdot t \leq -\beta \cdot \varepsilon_\sigma \).
Then \( (\#\Sigma_0) \cdot e^{-\beta t} = \sum_{\sigma \in \Sigma_0} [e^{-\beta t}] \leq \sum_{\sigma \in \Sigma_0} [e^{-\beta \varepsilon_\sigma}] \leq \sum_{\sigma \in \Sigma} [e^{-\beta \varepsilon_\sigma}] = \Delta_{\beta}^\varepsilon < \infty. \) □

Recall (§8) the definitions of: \( |\mu|_\rho, \mu^n, S_\mu, M_\mu, V_\mu \).

**THEOREM 23.6.** Let \( \Sigma \) be a set, \( \varepsilon : \Sigma \to \mathbb{R} \), \( \tilde{\beta} \in \text{DF}_\varepsilon \), \( \rho, \beta \in \mathbb{R} \).
Assume: \( \varepsilon \) is \( \infty \)-proper and \( \rho \geq 1 \) and \( \beta > \tilde{\beta} \). Then: \( |B_{\beta}^\varepsilon|_\rho < \infty \).

**Proof.** By Theorem 22.4, choose \( \zeta_0 \in I_\varepsilon \) s.t., \( \forall \sigma \in \Sigma, \varepsilon(\sigma) \geq \zeta_0 \).
For all \( \sigma \in \Sigma \), let \( \varepsilon_\sigma := \varepsilon(\sigma) \).
Then: \( \forall \sigma \in \Sigma, \varepsilon_\sigma \geq \zeta_0 \).
Let \( \beta := \beta - \tilde{\beta} \). Then: \( \beta > 0 \).
Since \( \beta > 0 \), we get: \( \zeta \to \infty, |\zeta|_\rho \cdot e^{-\beta \zeta} \to 0 \).
So, since \( \zeta \to |\zeta|_\rho \cdot e^{-\beta \zeta} : [\zeta_0; \infty) \to \mathbb{R} \) is continuous,
by the Extreme Value Theorem, choose \( M \in \mathbb{R} \) s.t.,
\( \forall \zeta \geq \zeta_0, |\zeta|_\rho \cdot e^{-\beta \zeta} \leq M. \)
Then: \( \forall \sigma \in \Sigma, |\varepsilon_\sigma|_\rho \cdot e^{-\beta \varepsilon_\sigma} \leq M. \)
Want: \( |B_{\beta}^\varepsilon|_\rho < \infty \). We have: \( |B_{\beta}^\varepsilon|_\rho = \sum_{\sigma \in \Sigma} [|\varepsilon_\sigma|_\rho \cdot e^{-\beta \varepsilon_\sigma}] \).
So, since $-\beta = -\tilde{\beta} - \tilde{\beta}$, we get: $|B_{\tilde{\beta}}^\varepsilon|_\rho = \sum_{\sigma \in \Sigma} \left[ (|\varepsilon_\sigma|^\rho \cdot e^{-\tilde{\beta} \varepsilon_\sigma}) \cdot (e^{-\tilde{\beta} \varepsilon_\sigma}) \right]$.

Since $\tilde{\beta} \in \text{DF}_\varepsilon$, we get: $\Delta^\varepsilon_{\tilde{\beta}} < \infty$. Then: $M \cdot \Delta^\varepsilon_{\tilde{\beta}} < \infty$.

Then: $|B_{\tilde{\beta}}^\varepsilon|_\rho = \sum_{\sigma \in \Sigma} \left[ (|\varepsilon_\sigma|^\rho \cdot e^{-\tilde{\beta} \varepsilon_\sigma}) \cdot (e^{-\tilde{\beta} \varepsilon_\sigma}) \right]$ 
\leq \sum_{\sigma \in \Sigma} \left[ M \cdot (e^{-\tilde{\beta} \varepsilon_\sigma}) \right] 
= M \cdot \left( \sum_{\sigma \in \Sigma} [e^{-\tilde{\beta} \varepsilon_\sigma}] \right) = M \cdot \Delta^\varepsilon_{\tilde{\beta}} < \infty. \quad \Box$

**THEOREM 23.7.** Let $\Sigma$ be a nonempty set, $\varepsilon : \Sigma \to \mathbb{R}$, $\beta \in \mathbb{R}$.

Assume: $\varepsilon$ is $\infty$-proper and $|\hat{B}_\beta^\varepsilon|_1 < \infty$. Then: $0 < \Delta^\varepsilon_{\beta} < \infty$.

**Proof.** Since $\Sigma \neq \emptyset$, we get: $\Delta^\varepsilon_{\beta} > 0$. Want: $\Delta^\varepsilon_{\beta} < \infty$.

In case $\Sigma$ is finite, we get: $\Delta^\varepsilon_{\beta} = \sum_{\sigma \in \Sigma} [e^{-\beta \varepsilon_\sigma}] < \infty$.

We therefore assume: $\Sigma$ is infinite.

By Theorem 22.3, $\text{IP}_\varepsilon \neq \emptyset$, so choose $(\sigma_1, \sigma_2, \ldots) \in \text{IP}_\varepsilon$.

For all $n \in \mathbb{N}$, let $\varepsilon_n := \varepsilon(\sigma_n)$.

By definition of $\text{IP}_\varepsilon$, we get as $n \to \infty$, $\varepsilon_n \to \infty$.

so **choose** $n_0 \in \mathbb{N}$ s.t., $\forall n \in [n_0, \infty)$, $\varepsilon_n \geq 1$.

Then: $\forall n \in [n_0, \infty)$, $\varepsilon_n > 0$, and so $|\varepsilon_n| = \varepsilon_n$.

We have: $|\hat{B}_\beta^\varepsilon|_1 = \sum_{n=1}^{\infty} [\varepsilon_n] \cdot e^{-\beta \varepsilon_n}$ and $\Delta^\varepsilon_{\beta} = \sum_{n=1}^{\infty} [e^{-\beta \varepsilon_n}]$.

Since $\Delta^\varepsilon_{\beta} = \sum_{n=1}^{\infty} [e^{-\beta \varepsilon_n}] = (\sum_{n=1}^{n_0-1} [e^{-\beta \varepsilon_n}] + (\sum_{n=n_0}^{\infty} [e^{-\beta \varepsilon_n}])$, and since $\sum_{n=n_0}^{\infty} [e^{-\beta \varepsilon_n}] < \infty$, we want: $\sum_{n=n_0}^{\infty} [e^{-\beta \varepsilon_n}] < \infty$.

For all $n \in [n_0, \infty)$, since $1 \leq \varepsilon_n = |\varepsilon_n|$ and since $e^{-\beta \varepsilon_n} > 0$, we get: $1 \cdot e^{-\beta \varepsilon_n} \leq |\varepsilon_n| \cdot e^{-\beta \varepsilon_n}$.

Then: $\sum_{n=n_0}^{\infty} [e^{-\beta \varepsilon_n}] \leq \sum_{n=n_0}^{\infty} [|\varepsilon_n| \cdot e^{-\beta \varepsilon_n}] 
\leq \sum_{n=1}^{\infty} [|\varepsilon_n| \cdot e^{-\beta \varepsilon_n}] = |\hat{B}_\beta^\varepsilon|_1 < \infty. \quad \Box$

**THEOREM 23.8.** Let $\Sigma$ be a set, $\varepsilon : \Sigma \to \mathbb{R}$, $\beta, \tilde{\beta} \in \mathbb{R}$.

Assume: $\varepsilon$ is $\infty$-proper and $\beta > \tilde{\beta} \in \text{DF}_\varepsilon$. Then: $\beta \in \text{DF}_\varepsilon$.

**Proof.** By Theorem 23.6, we have: $|\hat{B}_\beta^\varepsilon|_1 < \infty$.

Then, by Theorem 23.7, we have: $\Delta^\varepsilon_{\beta} < \infty$. Then $\beta \in \text{DF}_\varepsilon$. \quad \Box

**THEOREM 23.9.** Let $\Sigma$ be a set, $\varepsilon : \Sigma \to \mathbb{R}$.

Assume: $\varepsilon^*[0; \infty)$ is infinite and $\text{DF}_\varepsilon \neq \emptyset$. Let $\beta_0 := \inf \text{DF}_\varepsilon$.

Then: $(0 \leq \beta_0 < \infty)$ and $(\text{DF}_\varepsilon = [\beta_0; \infty))$ or $\text{DF}_\varepsilon = (\beta_0; \infty))$.

**Proof.** Since $\text{DF}_\varepsilon \neq \emptyset$, we get $\inf \text{DF}_\varepsilon < \infty$.

By Theorem 23.4, we get $\text{DF}_\varepsilon \subseteq (0; \infty)$, so $\inf \text{DF}_\varepsilon \geq \inf(0; \infty)$.

Then $\inf \text{DF}_\varepsilon \geq 0$. Then $0 \leq \inf \text{DF}_\varepsilon < \infty$.

So, since $\beta_0 = \inf \text{DF}_\varepsilon$, we get: $0 \leq \beta_0 < \infty$.

It remains only to show: $\text{DF}_\varepsilon = [\beta_0; \infty)$ or $\text{DF}_\varepsilon = (\beta_0; \infty)$. 

Want: \((\beta_0; \infty) \subseteq \text{DF}_\varepsilon \subseteq [\beta_0; \infty)\).

Since \(\beta_0 = \inf \text{DF}_\varepsilon\), we get: \(\text{DF}_\varepsilon \subseteq [\beta_0; \infty)\).

It remains to show: \((\beta_0; \infty) \subseteq \text{DF}_\varepsilon\).

Given \(\beta \in (\beta_0; \infty)\), want: \(\beta \in \text{DF}_\varepsilon\).

By Theorem 23.5, we have: \(\varepsilon\) is \(\infty\)-proper.

Since \(\beta > \beta_0 = \inf \text{DF}_\varepsilon\), choose \(\tilde{\beta} \in \text{DF}_\varepsilon\) s.t. \(\tilde{\beta} < \beta\).

Then, by Theorem 23.8, we get: \(\beta \in \text{DF}_\varepsilon\). \(\square\)

Note that, in both cases of the conclusion of the preceding theorem, the interior of \(\text{DF}_\varepsilon\) is \((\beta_0; \infty)\).

In the following definition, “IDF” stands for “Interior-Delta-Finite”.

**DEFINITION 23.10.** Let \(\Sigma\) be a set, \(\varepsilon : \Sigma \to \mathbb{R}\).
Assume: \(\varepsilon\) is \(\infty\)-proper and \(\text{DF}_\varepsilon \neq \emptyset\). Let \(\beta_0 := \inf \text{DF}_\varepsilon\).

Then: \([\text{IDF}_\varepsilon] := (\beta_0; \infty)\).

**THEOREM 23.11.** Let \(\Sigma\) be a set, \(\varepsilon : \Sigma \to [0; \infty)\).

For all \(k \in \mathbb{N}\), let \(n_k := \#(\varepsilon^*[k-1; k))\).

Let \(\beta \geq 0\). Then: \((\beta \in \text{DF}_\varepsilon) \iff \left(\sum_{k=1}^{\infty} n_k e^{-\beta k} < \infty\right)\).

*Proof.* Proof of \(\Rightarrow:\)

Assume: \(\beta \in \text{DF}_\varepsilon\). **Want:** \(\sum_{k=1}^{\infty} n_k e^{-\beta k} < \infty\).

MORE LATER.

End of proof of \(\Rightarrow\).

Proof of \(\Leftarrow:\)

Assume: \(\sum_{k=1}^{\infty} n_k e^{-\beta k} < \infty\). **Want:** \(\beta \in \text{DF}_\varepsilon\).

MORE LATER.

End of proof of \(\Leftarrow\). \(\square\)

**THEOREM 23.12.** Let \(\Sigma\) be a set, \(\varepsilon : \Sigma \to [0; \infty)\).

For all \(k \in \mathbb{N}\), let \(n_k := \#(\varepsilon^*[k-1; k))\).

Assume: \(\forall k \in \mathbb{N}, \ n_k > e^{k^2}\). Then: \(\text{DF}_\varepsilon = \emptyset\).

**THEOREM 23.13.** Let \(\Sigma\) be a set, \(\varepsilon : \Sigma \to [0; \infty), \beta_0 \geq 0\).

For all \(k \in \mathbb{N}\), let \(n_k := \#(\varepsilon^*[k-1; k))\).

Assume: as \(k \to \infty, \ n_k e^{-\beta_0 k} \to 1\). Then: \(\text{DF}_\varepsilon = (-\beta_0; \infty)\).

**THEOREM 23.14.** Let \(\Sigma\) be a set, \(\varepsilon : \Sigma \to [0; \infty), \beta_0 > 0\).

For all \(k \in \mathbb{N}\), let \(n_k := \#(\varepsilon^*[k-1; k))\).

Assume: as \(k \to \infty, \ k^2 n_k e^{-\beta_0 k} \to 1\). Then: \(\text{DF}_\varepsilon = [-\beta_0; \infty)\).
DEFINITION 23.15. Let \( \Sigma \) be a set, \( \varepsilon : \Sigma \rightarrow [0; \infty), \beta \in \mathbb{R}. \) If \( |\hat{B}_\beta^\varepsilon|_1 < \infty, \) then: \( \Gamma^\varepsilon_\beta := \sum_{\sigma \in \Sigma} [(\varepsilon(\sigma)) \cdot \epsilon^{-\beta(\varepsilon(\sigma))}] . \)

If \( \Sigma = \emptyset, \) then: \( \varepsilon \) is the empty function and \( \forall \beta \in \mathbb{R}, \Gamma^\varepsilon_\beta = 0. \)

If \( |\hat{B}_\beta^\varepsilon|_1 < \infty, \) then \( \Gamma^\varepsilon_\beta = \sum_{\sigma \in \Sigma} [\varepsilon(\sigma) \cdot (\hat{B}_\beta^\varepsilon(\sigma))], \)

and it follows that \( \Gamma^\varepsilon_\beta \) is the integral of \( \varepsilon \) wrt \( \hat{B}_\beta^\varepsilon. \)

We have: \( |\hat{B}_\beta^\varepsilon|_1 = \sum_{\sigma \in \Sigma} [\varepsilon(\sigma) \cdot (\hat{B}_\beta^\varepsilon(\sigma))], \)

So, by subadditivity of absolute value, if \( |\hat{B}_\beta^\varepsilon|_1 < \infty, \) then \( |\Gamma^\varepsilon_\beta| \leq |\hat{B}_\beta^\varepsilon|_1. \)

Let \( \Sigma \) be a set, \( \varepsilon : \Sigma \rightarrow \mathbb{R}, \beta \in \mathbb{R}. \) Then:

\( (0 < \Delta^\varepsilon_\beta < \infty) \Leftrightarrow (0 < \hat{B}_\beta^\varepsilon(\Sigma) < \infty) \Leftrightarrow (\hat{B}_\beta^\varepsilon \in \mathcal{F}\mathcal{M}_\Sigma^\varepsilon). \)

DEFINITION 23.16. Let \( \Sigma \) be a set, \( \varepsilon : \Sigma \rightarrow \mathbb{R}, \beta \in \mathbb{R}. \) Assume: \( 0 < \Delta^\varepsilon_\beta < \infty. \) Then: \( \hat{B}_\beta^\varepsilon := \mathcal{N}(\hat{B}_\beta^\varepsilon) \in \mathcal{P}_\Sigma. \)

THEOREM 23.17. Let \( \Sigma \) be a set, \( \varepsilon : \Sigma \rightarrow (0; \infty). \) Assume: \( \varepsilon^*[0; \infty) \) is infinite and \( \text{DF}_\varepsilon \neq \emptyset. \) Then: as \( \beta \rightarrow \infty, \Gamma^\varepsilon_\beta \rightarrow 0 \) and \( \Delta^\varepsilon_\beta \rightarrow 0. \)

Proof. LATER. \( \square \)

DEFINITION 23.18. Let \( \Sigma \) be a set, \( \varepsilon : \Sigma \rightarrow \mathbb{R}. \) Assume: \( 0 < \Delta^\varepsilon_\beta < \infty \) and \( |\hat{B}_\beta^\varepsilon|_1 < \infty. \) Then: \( \hat{A}_\beta^\varepsilon := \Gamma^\varepsilon_\beta / \Delta^\varepsilon_\beta. \)

THEOREM 23.19. Let \( \Sigma \) be a set, \( \varepsilon : \Sigma \rightarrow \mathbb{R}, \beta \in \mathbb{R}. \) Assume: \( 0 < \Delta^\varepsilon_\beta < \infty \) and \( |\hat{B}_\beta^\varepsilon|_1 < \infty. \) Then: \( \hat{A}_\beta^\varepsilon = M_{\varepsilon\cdot\hat{B}_\beta^\varepsilon}. \)

Proof. LATER. \( \square \)

Note that, by Theorem 23.7, if \( \Sigma \neq \emptyset \) and \( \varepsilon \) is \( \infty \)-proper, then \( (|\hat{B}_\beta^\varepsilon|_1 < \infty) \Rightarrow (0 < \Delta^\varepsilon_\beta < \infty). \)

Without the assumption of \( \infty \)-properness,

this implication fails, see Theorem 23.21.

Even with \( \infty \)-properness, \( (|\hat{B}_\beta^\varepsilon|_1 < \infty) \Leftrightarrow (0 < \Delta^\varepsilon_\beta < \infty), \) as follows:

THEOREM 23.20. Let \( \Sigma := [3..\infty). \)

Define \( \varepsilon : \Sigma \rightarrow \mathbb{R} \) by: \( \forall k \in \Sigma, \varepsilon(k) = (\ln k) + 2 \cdot (\ln(\ln k)). \)

Let \( \beta := 1. \) Then: \( 0 < \Delta^\varepsilon_\beta < \infty \) and \( |\hat{B}_\beta^\varepsilon|_1 = \infty. \)

Proof. Since \( \Sigma \neq \emptyset, \) we get: \( \Delta^\varepsilon_\beta > 0. \)

Want: \( \Delta^\varepsilon_\beta < \infty \) and \( |\hat{B}_\beta^\varepsilon|_1 = \infty. \) For all \( k \in \Sigma, \) let \( \varepsilon_k := \varepsilon(k). \)
Since $\Delta_\beta^\varepsilon = \sum_{k \in \Sigma} [e^{-\beta \varepsilon_k}] = \sum_{k=3}^{\infty} [e^{-\varepsilon_k}] = \sum_{k=3}^{\infty} [e^{-1 - \varepsilon_k}] < \infty$, it remains only to show: $|\hat{B}_\beta^\varepsilon|_1 = \infty$.

It suffices to show: $|\hat{B}_\beta^\varepsilon|_1 \geq \infty$.

We have: $\forall k \in [3, \infty), k > e$, so $\ln k > 1$, so $\ln(\ln k) > 0$.

For all $k \in [3, \infty)$, since $\varepsilon_k = (\ln k) + 2 \cdot (\ln(\ln k)) > 1 + 2 \cdot 0 = 1 > 0$,

we get: $|\varepsilon_k| = \varepsilon_k$.

We have: $|\hat{B}_\beta^\varepsilon|_1 = \sum_{k \in \Sigma} [\frac{\varepsilon_k}{k} \cdot e^{-\varepsilon_k}] = \sum_{k=3}^{\infty} [\frac{\varepsilon_k}{k} \cdot e^{-\varepsilon_k}]$

$= \sum_{k=3}^{\infty} \left[ \frac{\varepsilon_k}{e^{\varepsilon_k}} \right] = \sum_{k=3}^{\infty} \left[ \frac{\varepsilon_k}{\varepsilon_k} \right] = \sum_{k=3}^{\infty} \left[ \frac{(\ln k) + 2 \cdot (\ln(\ln k))}{e^{(\ln k) + 2 \cdot (\ln(\ln k))}} \right]$

$= \sum_{k=3}^{\infty} \left[ \frac{(\ln k) + 2 \cdot (\ln(\ln k))}{k \cdot (\ln k)^2} \right]$

$\geq \sum_{k=3}^{\infty} \left[ \frac{\ln k}{k \cdot (\ln k)^2} \right] = \sum_{k=3}^{\infty} \left[ \frac{1}{k \cdot (\ln k)^2} \right] = \infty$.  

\[ \square \]

**THEOREM 23.21.** Let $\Sigma := \mathbb{N}$. Let $\beta := 1$.

Define $\varepsilon : \Sigma \to \mathbb{R}$ by: $\forall k \in \Sigma, \varepsilon(k) = 1/k^2$.

Then: $|\hat{B}_\beta^\varepsilon|_1 < \infty$ and $\Delta_\beta^\varepsilon = \infty$.

**Proof.** For all $k \in \Sigma$, let $\varepsilon_k := \varepsilon(k)$. Then: $\forall k \in \mathbb{N}, \varepsilon_k = 1/k^2$.

We have: $\forall k \in \mathbb{N}$, both $|\varepsilon_k| = 1/k^2$ and $-1 \cdot \varepsilon_k = 1 - 1/k^2$.

Since $|\hat{B}_\beta^\varepsilon|_1 = \sum_{k=1}^{\infty} [\varepsilon_k \cdot e^{-\varepsilon_k}]$

$= \sum_{k=1}^{\infty} [\varepsilon_k \cdot e^{-1 - \varepsilon_k}]$

$= \sum_{k=1}^{\infty} [(1/k^2) \cdot e^{-1/k^2}]$

$\leq \sum_{k=1}^{\infty} [(1/k^2) \cdot 1] = \sum_{k=1}^{\infty} [1/k^2] < \infty$,

it remains only to show: $\Delta_\beta^\varepsilon = \infty$.

We have: as $k \to \infty$, $e^{-1/k^2} \to 1$. Then: $\sum_{k=1}^{\infty} [e^{-1/k^2}] = \infty$.

Then: $\Delta_\beta^\varepsilon = \Delta_1^\varepsilon = \sum_{k=1}^{\infty} [e^{-1 - \varepsilon_k}] = \sum_{k=1}^{\infty} [e^{-1/k^2}] = \infty$.  

\[ \square \]

**DEFINITION 23.22.** Let $\Sigma$ be a set, $\varepsilon : \Sigma \to \mathbb{R}$.

Then $A^\varepsilon : \text{IDF}_\varepsilon \to \mathbb{R}$ is defined by: $\forall \beta \in \text{IDF}_\varepsilon, A^\varepsilon_\beta(\beta) = A^\varepsilon_\beta$.

**THEOREM 23.23.** Let $\Sigma$ be a set, $\varepsilon : \Sigma \to \mathbb{R}$.

Assume: $\varepsilon^*(0; \infty)$ is infinite and $\text{DF}_\varepsilon \neq \emptyset$.

Let $\beta_0 := \inf \text{DF}_\varepsilon$. Then: as $\beta \to \beta_0^+$, $A^\varepsilon_\beta(\beta) \to \infty$.

**Proof.** LATER.
THEOREM 23.24. Let $\Sigma$ be a set, $\varepsilon : \Sigma \to \mathbb{R}$.
Assume: $\varepsilon^* [0; \infty)$ is infinite and $DF_\varepsilon \neq \emptyset$.
Then: as $\beta \to \infty$, $A_\varepsilon^*(\beta) \to \min I_\varepsilon$.

Proof. LATER. $\square$

24. Countably infinite sets of states

MORE LATER
25. Appendix: Python code

Thanks once again to C. Prouty, for writing the Python code to do the Boltzmann computations in this paper:

First code: The GFA and 0, 2, 20 dollar awards, with average 3 dollars.

```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

def F(beta):
    z = np.zeros(3)
    z[0] = 1
    z[1] = np.exp(-2 * beta)
    z[2] = np.exp(-20 * beta)
    return z

def G(beta):
    z = np.zeros(3)
    z[0] = 0
    z[1] = 2 * np.exp(-2 * beta)
    z[2] = 20 * np.exp(-20 * beta)
    return z

def f(beta):
    return np.sum(F(beta))

def g(beta):
    return np.sum(G(beta))

def bisection(minval, maxval, y, fn):
    mid = (maxval + minval) / 2
    while((fn(mid) - y) ** 2 > 0.0000001):
        if(fn(mid) < y):
            maxval = mid
        else:
            minval = mid
    mid = (maxval + minval) / 2
    return mid

fn = lambda x: g(x) / f(x)
```
target = bisection(-25, 25, 3, fn)
b = 0.07410049 # hard-coded result of bisection
r = F(b) / f(b)
df = pd.DataFrame(r)
df.to_excel("results2.xlsx", index=False)
betas = np.linspace(-25,25,100000)
z = np.zeros(len(betas))
for i in range(len(betas)):
z[i] = fn(betas[i])
plt.plot(betas,z)
plt.show()

Second code: The BUA and red bags and blue bags

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
def F(beta):
z = np.zeros(25).reshape(5,5)
for i in range(5):
    for j in range(5):
z[i,j] = np.exp(-(i+j)*beta)
z[4,4] = 0
return z
def G(beta):
z = np.zeros(25).reshape(5,5)
for i in range(5):
    for j in range(5):
z[i,j] = (i+j) * np.exp(-(i+j)*beta)
z[4,4] = 0
return z
def f(beta):
return np.sum(F(beta))
def g(beta):
return np.sum(G(beta))
def bisection(minval, maxval, y, fn):
mid = (maxval + minval) / 2
while((fn(mid) - y) ** 2 > 0.0000001):
    if(fn(mid) < y):
        maxval = mid
    else:
        minval = mid
    mid = (maxval + minval) / 2
return mid

fn = lambda x: g(x) / f(x)
target = bisection(-25, 25, 1, fn)
b = 1.06697083 # hard-coded result of bisection
r = F(b) / f(b)
df = pd.DataFrame(r)
df.to_excel("results5.xlsx", index=False)
betas = np.linspace(-25, 25, 100000)
z = np.zeros(len(betas))
for i in range(len(betas)):
    z[i] = fn(betas[i])
plt.plot(betas, z)
plt.show()