Stochastic model of atoms in a magneto-optical trap

Jasper de Koning

Bachelor thesis, spring 2005
supervisor: dr. F.M. Spieksma

Contents

1 Introduction

In today's world, nanotechnology is one of the most promising fields, both for research and businesses. A few methods have been developed to control small amounts of atoms. One of those methods is the so-called Magneto Optical Trap. In this machine the atoms can be cooled to just a fraction above zero, thus greatly reducing the motions of the atoms.

In this paper we will study the model that was introduced in the paper: "Stochastic model for the number of atoms in a magneto-optical trap" written by A. Rukhin and I. Bebu to study the number of atoms in a trap. We'll study this model, and carefully reconsider the assumptions they made, and obtain an adjusted version of the model.

In the model that was introduced in the original paper, an extra element, a door, was added to the model to obtain a higher measure of control on the atoms in the trap. This trap-door checks each T seconds, whether there still are atoms present in the trap. If no atoms are present in the trap, the door will open the next T seconds to allow new atoms to enter the trap. If, on the other hand, atoms are present in the trap, the door will stay closed for another T seconds. We'll say that the window size equals T.

2 Original model

2.1 Mathematical description

Since all observations occur each T seconds, we can model this as a Markov chain in discrete time. For the states we pick the number of atoms that are in the MOTP. The transition probabilities are given by the modelling assumptions. In the paper they're given by the following three stochastic variables

- X_t the number of atoms in the MOTP at time t
- Y_t the number of atoms that leave the trap between time t and $t+1$
- R_t the number of atoms that arrive at the trap between time t and $t+1$. Because atoms can only enter the trap if it's empty at the beginning of the time-interval, R_t only plays a role if the trap is empty.

The original paper proceeds to tie those formulas together as follows:

$$X_{t+1} = \begin{cases} R_t & \text{if } X_t \leq Y_t \\ X_t - Y_t & \text{if } X_t > Y_t \end{cases} \tag{1}$$

Here R_t is Poisson distributed with parameter λ, and the stochastic variable Y_t is also Poisson distributed, but with parameter $\mu X_t T$.

2.2 First thoughts

When one reviews somebody elses work, one should always start by giving some positive feedback. So we'll start by saying that the arrival process R_t is a natural assumption. For we can take an arrival process Poisson distributed if the collection it departs from is large, and each item has the same departure probability. Those demands seem very reasonable in our model, because the source of the particle beam will definitely contain a large number of atoms. And it seems reasonable to give each atom of a large group of atoms, who are likely to be fired, the same chance to depart.

Unfortunately, that is the only part of the equations on which I have no negative feedback. So we'll start by looking at some glaring problems. Firstly: if $X_t \leq Y_t$ then $X_{t+1} = R_t$. This means that if the trap becomes empty during a time interval, it will be filled with R_t atoms at the beginning of the next interval. So the trap-door must have been open at the beginning of this time-interval. And that is strange, because we don't know whether the trap is empty, until we're at the end of the time interval.

Secondly: Y_t is Poisson distributed, this means that Y_t could be very big. It could even be a lot bigger than X_t. That would mean there are more atoms departing the trap in one time interval, then there were at the beginning of that interval. They have solved this by truncating the Poisson process, but this does not look like a good solution at all.

Thirdly: for a continuous process to be a Markov process, it should be memoryless. But if we assume this for the process of the departure of atoms, we'll see that we get a very strange distribution which depends on the window size T.

2.3 The original distribution of Y_t

In the last section, we noticed that the distribution of the departures has some strange attributes. This seems to indicate that we should use another

distribution. In this section we will study this distribution for two different window sizes T, and conclude that this distribution does not suffice.

We look at two separate cases. The first has window size T and the second has window size $2T$. Denote these two cases by superscript T and superscript $2T$ respectively. We will study the expected number of departures for two time-units in the first model, compared to one time-unit in the second model, with λ, μ identical in both cases and both starting with X_i atoms in the trap. Note that this means that in both cases the same amount of time passes.

Further we have to note that $Y_i^{2T} = Y_i^T + \tilde{Y}_i^T$, with $\tilde{Y}_i^T, Y_i^T$ i.i.d. Poisson distributed variables with parameter $\mu T X_i$. We continue by calculating the difference $E(Y_i^{2T}) - (E(Y_i^T) + E(Y_{i+1}^T))$.

$$\begin{aligned} E(Y_i^{2T}) - E(Y_i^T + Y_{i+1}^T) &= E(Y_i^T + \tilde{Y}_i^T) - E(Y_i^T + Y_{i+1}^T) \\ &= E(Y_i^T + \tilde{Y}_i^T - Y_i^T + Y_{i+1}^T) \\ &= E(\tilde{Y}_i^T - Y_{i+1}^T) \\ &> 0. \end{aligned}$$

One could pose that this analysis isn't complete, since we don't take into account that there may depart more than X_i atoms from the trap during the first or second T seconds. The conclusion however still holds, let's rewrite the formulas, by replacing $E(Y_i)$ with $E(\min(Y_i, X_i))$, this makes sure that we only count the atoms that leave the trap, and none more. This yields:

$$\begin{aligned} &E(\min(Y_i^{2T}, X_i)) - E(\min(Y_i^T + Y_{i+1}^T, X_i)) \\ &\quad= E(\min(Y_i^T + \tilde{Y}_i^T, X_i)) - E(\min(Y_i^T + Y_{i+1}^T, X_i)) \\ &\quad= E(\min(\tilde{Y}_i^T, X_i - Y_i^T)) - E(\min(Y_{i+1}^T, X_i - Y_i^T)) \\ &\quad= \sum_{j=0}^{\infty} P(Y_t^T = j)(E(\min(\tilde{Y}_i^T, X_i - Y_i^T) - \min(Y_{i+1}^T, X_i - Y_i^T) | Y_i^T = j)) \\ &\quad= \sum_{j=0}^{X_i} P(Y_t^T = j)\Big[\sum_{k=0}^{X_i - j} (P(\tilde{Y}_i^T = k) - P(Y_{i+1}^T = k))k \\ &\qquad\qquad + P(\tilde{Y}_i^T > X_i - j)(X_i - j) - P(Y_i^T > X_i - j)(X_i - j)\Big] \\ &\quad= \sum_{j=0}^{X_i} P(Y_t^T = j) \sum_{k=0}^{X_i - j} (P(\tilde{Y}_i^T > k) - P(Y_{i+1}^T > k)) \\ &\quad> 0. \end{aligned}$$

So there will depart more atoms in the model with window size $2T$. So how longer the time between the opening and closing of the door, the more atoms depart from the model. This seems highly unlikely, because one would expect it to be the other way around. It also gives a non-trivial correspondence between the number of atoms that depart from the trap and the time the door is open. So we can safely conclude that this may be a very poor choice for this distribution indeed.

3 Adjusted model

3.1 Differences compared to the original model

As we've seen in the previous section, the original model had some quite serious flaws. So it's time to make the necessary changes. Firstly, we have to solve the problem with the "prophetical" door. This can be done quite easily by tying the equations together a bit differently:

$$X_{t+1} = \begin{cases} R_t & \text{if } X_t = 0 \\ X_t - Y_t & \text{if } X_t \neq 0 \end{cases}$$

Secondly we'll have to choose another distribution for the departures, since the Poisson distribution clearly does not suffice. Luckily, we don't have to think long which distribution to choose, because we have little choice in the matter. We have to satisfy the following two demands: 1) the time an atom spends in the trap should not depend on the window size T; 2) the distribution has to be memoryless. This implies that the time spent by an atom in the trap should have an exponential distribution with parameter μ.

Now we've solved the most glaring problems of the original model, it's time to see what consequences our adjustments have. First we'll look what the distribution of Y_t is now, by looking at the process between the moment the atoms arrive, until the last of these atoms has departed.

We take S_i the random variable of the time that atom i spends in the trap, and X_0 the number of atoms that were present in the trap at time 0. Combining this we get:

$$P(Y_t = j) \quad = \quad P(\sum_{k=1}^{X_0} 1_{t<S_k<t+1} = j)$$

$$
\begin{aligned}
&= P(\sum_{k=1}^{X_t} 1_{S_k<t+1|S_j>t} = j) \\
&= P(\sum_{k=1}^{X_t} 1_{S_k<1} = j),
\end{aligned}
$$

which is exactly the binomial distribution with parameters X_t and $P(S_1 < T) = 1 - e^{-\mu T}$.

To sum it al up, the adjusted model becomes:

$$
X_{t+1} = \begin{cases} R_t & \text{if } X_t = 0 \\ X_t - Y_t & \text{if } X_t \neq 0 \end{cases} \quad \text{with:} \quad \begin{cases} R_t & \sim \text{Poisson}(\lambda T) \\ Y_t & \sim \text{binomial}(X_t, 1 - e^{-\mu T}) \end{cases} \tag{2}
$$

Since the distribution of R_t depends on λT and Y_t depends on μT, we can take $T = 1$ for the rest of this paper, and still get all possible processes.

The transition matrix P for $T = 1$ becomes:

$$
P_{ij} = \begin{cases} \frac{\lambda^j}{j!} e^{-\lambda} & \text{if } i = 0 \\ 0 & \text{if } i \neq 0 \text{ and } i < j \\ \binom{i}{j} (1 - e^{-\mu})^{i-j} e^{-\mu j} & \text{if } i \neq 0 \text{ and } i \geq j \end{cases} \tag{3}
$$

3.2 Properties of this Markov chain

The first thing we notice is that there is no bound on the number of atoms that arrive in the trap. As a consequence there are countably many states. This makes examining the chain more complicated. It's obvious that this chain is irreducible. So the first thing we have to do now, is to check whether this Markov chain is positive recurrent. We'll do this by proving that Foster's Criterion applies to this chain, see [2].

We apply Foster's criterion, with $\nu_i = i$ and $M = \{0\}$.

$$
\begin{aligned}
\sum_{j=1}^{\infty} P_{0j} \nu_j &= E(R_t) \\
\varepsilon + \sum_{j=1}^{\infty} P_{ij} \nu_j &= \varepsilon + \sum_{j=1}^{\infty} P_{ij} j \\
&= \varepsilon + E(X_t - Y_t | X_t = i)
\end{aligned}
$$

$$
\begin{aligned}
&= \varepsilon + i - E(Y_t | X_t = i) \\
&= \varepsilon + i - i(1 - e^{-\mu i}) \\
&= \varepsilon + i e^{-\mu} \\
&= \varepsilon + e^{-\mu} \nu_i .
\end{aligned}
$$

If we take $\varepsilon = 1 - e^{-\mu}$ then Foster's criterion holds, hence the process has a stationary distribution. But there is even exponentially fast convergence to the stationary distribution, because the following equations also hold, see [2].

$$
\begin{aligned}
\sum_{j=0}^{\infty} P_{ij} \nu_j &\leq \mu_i (1 - \epsilon), \quad \text{for } i = 1..\infty \\
\sum_{j=0}^{\infty} P_{ij} \nu_j &= E(R_t) < \infty, \quad \text{for } i = 0.
\end{aligned}
$$

This was easy. The hard part however, is to find the convergence rate. I haven't been able to compute this, but that's not from a lack of trying, for this is a notoriously hard problem.

3.3 Approximate with NxN matrix

Now we've defined a model, and checked some fundamental properties of it. We'd like to see how the value π_1 behaves for different values of λ and μ. To do this, we need some way to approximate π_1. Because it's unlikely there will arrive many more atoms than λ, it's a logical idea, to look only at the first $N - 1$ states of the Markov chain, with N suitable large.

But if we chop al but the first N states from the chain, we won't have a Markov chain anymore since $P_{0j} > 0, \quad \forall j \geq N$. So we have to modify the transition probabilities a bit. We put the error ε_N on top of the P_{00}, and hope it doesn't mess up the results too much.

Let us define this error $\varepsilon_N = \sum_{i=N}^{\infty} P_{0i}$. And since P_{0j} is very small for j large, we have just reason to assume that this will give a sufficient approximation. Let's denote each variable of this chopped Markov chain with a superscript N in front of that variable. So we'll get for the transition matrix:

$$ {}^{N}P_{ij} = \begin{cases} P_{00} + \varepsilon_N & \text{if } i = j = 0 \\ P_{0j} & \text{if } i = 0 \text{ and }) < j < N \\ 0 & \text{if } 0 < i < j < N \\ P_{ij} & \text{if } i \neq 0 \text{ and } 0 \leq j \leq i < N. \end{cases} \tag{4} $$

This obviously is a Markov chain. So now we can look how big the error is that we've introduced, by chopping so many states from the original chain. We define, given that $X_0 = i$

$$ T_{ij} = \min(t | X_t = 0), $$

and given that $X_0 = 0$,

$$ A_i = \sum_{t=0}^{\infty} 1_{\{X_t = i, T_{00} > t\}}. $$

We know:

$$ \pi_i = \frac{E(A_i)}{E(T_{00})}. $$

So if we find suitable bounds for both $\left| E(A_i) - E({}^{N}A_i) \right|$ and $\left| E(T_{00}) - E({}^{N}T_{00}) \right|$ we'll have found a bound for $\left| \pi_i - {}^{N}\pi_i \right|$. First we'll have a look at $\left| E(T_{00}) - E({}^{N}T_{00}) \right|$.

$$ \begin{aligned} E(T_{00}) &= 1 + \sum_{j=1}^{\infty} P_{0j} E(T_{j0}) \\ &= 1 + \sum_{j=1}^{N-1} P_{0j} E(T_{j0}) + \sum_{j=N}^{\infty} P_{0j} E(T_{j0}) \\ &= E({}^{N}T_{00}) + \sum_{j=N}^{\infty} P_{0j} (E(T_{j0}) - 1) \end{aligned} $$

This implies:

$$\begin{aligned}\left|E(T_{00}) - E({}^{N}T_{00})\right| &= \sum_{j=N}^{\infty} P_{0j}(E(T_{j0}) - 1) \\ &< \sum_{j=N}^{\infty} P_{0j} j/\mu \\ &= \frac{\lambda}{\mu} \sum_{j=N-1}^{\infty} P_{0j} \\ &= \frac{\lambda}{\mu}(P_{0N-1} + \varepsilon_N).\end{aligned}$$

Since evidently $E(T_{j0}) < \frac{j}{\lambda} + 1$. We'll denote the error we obtained here $\frac{\lambda}{\mu}(P_{0N-1}+\varepsilon_N)$ by ε'_N. We'll proceed with finding a bound for $\left|E(A_i) - E({}^{N}A_i)\right|$.

$$\begin{aligned}E(A_i) &= \sum_{j=0}^{\infty} P_{0j} E(A_i|X_0 = j) \\ &= \sum_{j=0}^{N-1} P_{0j} E(A_i|X_0 = j) + \sum_{j=N}^{\infty} P_{0j} E(A_i|X_0 = j) \\ &= E({}^{N}A_i) + \sum_{j=N}^{\infty} P_{0j} E(A_i|X_0 = j).\end{aligned}$$

This implies:

$$\begin{aligned}\left|E(A_i) - E({}^{N}A_i)\right| &= \sum_{j=N}^{\infty} P_{0j} E(A_i|X_0 = j) \\ &< \sum_{j=N}^{\infty} P_{0j} E(T_{j0}) \\ &< \sum_{j=N}^{\infty} P_{0j}(\frac{j}{\mu} + 1) \\ &= \varepsilon'_N + \varepsilon_N.\end{aligned}$$

Now we have those two results, we can easily combine them to obtain a bound for $\left|\pi_i - {}^{N}\pi_i\right|$. We do this by applying the following lemma.

Lemma 1. *Suppose* $\left|E(A_i) - E({}^N A_i)\right| < \delta_1$ *and* $\left|E(T_{00}) - E({}^N T_{00})\right| < \delta_2$. *Then* $\left|\pi_i - {}^N\pi_i\right| < \frac{\delta_1+\delta_2}{E(T_{00})}$.

Proof.

$$\begin{aligned}
\left|\pi_i - {}^N\pi_i\right| &= \left|\frac{E(A_i)}{E(T_{00})} - \frac{E({}^N A_i)}{E({}^N T_{00})}\right| \\
&\leq \frac{\left|E({}^N T_{00})(E(A_i) - E({}^N A_i))\right| + \left|E({}^N A_i)(E({}^N T_{00} - E(T_{00}))\right|}{E(T_{00})E({}^N T_{00})} \\
&< \frac{E({}^N T_{00})\delta_1 + E({}^N A_i)\delta_2}{E(T_{00})E({}^N T_{00})} \\
&< \frac{\delta_1 + \delta_2}{E(T_{00})},
\end{aligned}$$

since $E({}^N A_i) \leq E({}^N T_{00})$. □

If we apply lemma 1 to the obtained bounds, we get $\left|\pi_i - {}^N\pi_i\right| < \frac{2\varepsilon'_N+\varepsilon_N}{E(T_{00})}$, so we can choose N such that the error $\left|\pi_i - {}^N\pi_i\right|$ is arbitrarily small. As a last note we conclude that the stationary distribution is very easily calculated for this N x N matrix, because the matrix is of the form:

$$\begin{bmatrix} P_{00} & \dots & P_{0,N-1} \\ P_{10} & \dots & P_{N-1,0} \\ P_{i0} & \ddots & 0 \\ P_{N-1,0} & \ddots & 0 \end{bmatrix}$$

This means we start sweeping from the $(N-1)^{\text{st}}$ column, and go upwards to find the stationary distribution.

3.4 Several observations

Now we've found a way to approximate the values of π, we can create a contour plot of π_1. Since the error is smaller than $\frac{2\varepsilon'_N+\varepsilon_N}{E(T_{00})}$ we don't have to take excessive large value for N, because both ε'_N and ε'_N are small if we take $N = 2\lambda + 10$.

In all graphs we use in this paper, we'll take μ horizontal and λ vertical. In the contour plots the areas with an higher value have a lighter hue.

We can see some interesting features in this contour plot, first of all there is a "long island" at $\lambda \simeq 0,7$. Secondly we see that there is a large range of values near $\lambda = \mu = 0$. And last but not least, there is something interesting going on near λ at 7 in figure 1. We see that those contour-lines are curved, which means that the value of π_1 is not monotone! A closer look at the values of π_i for λ bigger in figure 2, shows that π_i is far from monotonous.

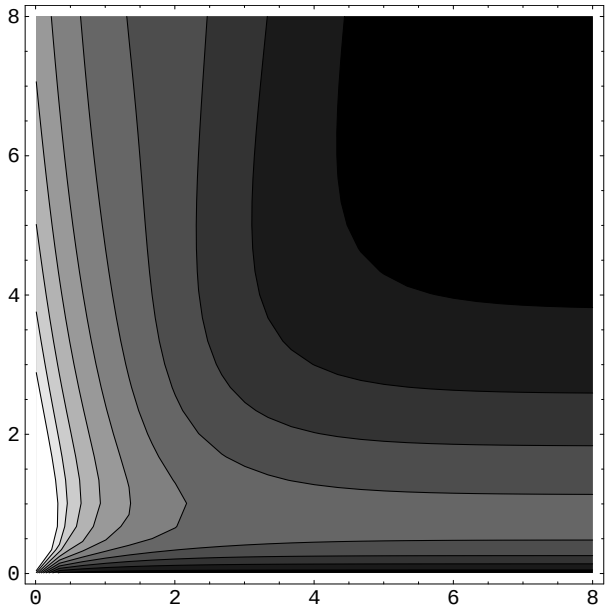

Figure 1: A contour plot of π_1 for $\mu \in [0, 01..8]$, π_1 for $\lambda \in [0, 01..8]$

4 Continuous case

4.1 Explanation

It's always hard to get hard numbers for a Markov chain, since we'll always have to work with a large and cumbersome transition matrix. And our transition matrix isn't a friendly one, it doesn't have any nice "local" properties. So it's time to look at this problem from a different angle. We've been looking at the Markov chain, but we should not forget that this Markov chain

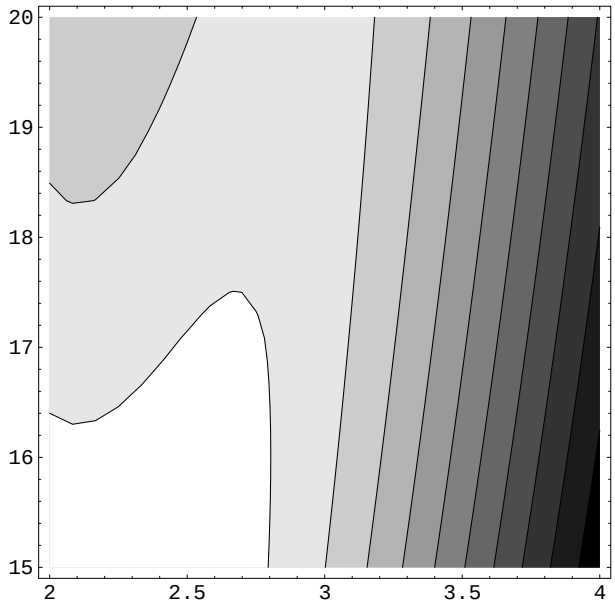

Figure 2: A contour plot of π_1 for $\mu \in [2..4]$, $\lambda \in [15..20]$

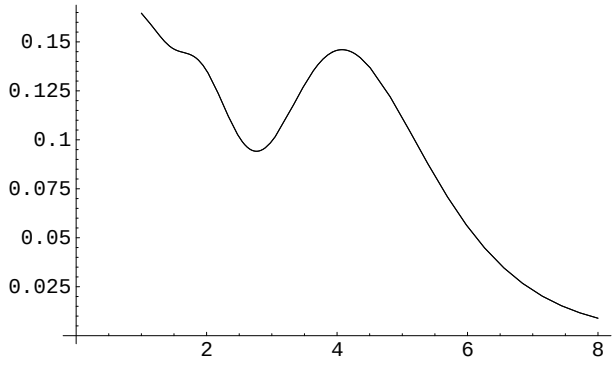

Figure 3: A line plot of π_1 for $\mu \in [1..8]$, $\lambda = e^4$

is based on a continuous departure of atoms. We can get some interesting results by looking at this continuous model.

We take $T_{i0}^C = \max_k S_k$, with S_i the exponentially distributed time that atom i remains in the trap. This means that T_{i0}^C is the random time, the last atom departing, remains in the trap.

We know that $\max_i S_i$ is distributed as $\sum_{i=1}^{X_t} F_i$, F_i exponential distributed with parameter $i\mu$. Combining this with $1 + \ln(j) > \sum_{i=1}^{j} 1/i > \ln(j+1)$, we get:

$$\frac{1}{\mu}(1 + \ln(j)) > \sum_{i=1}^{j} \frac{1}{i\mu} = E(T_{j0}^C) > \frac{\ln(j+1)}{\mu}.$$

Obviously there also is a relation between T_{i0}^C and T_{i0}. If we look at a realization of T_{i0}^C, we note that the discrete Markov chain reaches 0 at the end of the time-interval in which the last atom departed. Since the door is open for 1 time-interval, we get the following formula:

$$T_{i0}^C < T_{i0} < 1 + T_{i0}^C. \tag{5}$$

Now we have found these nice bounds for the time it takes to get from i to 0, we'd like to say something about T_{00}^C. This turns out to be considerably harder. I've tried to approximate it by using standard methods, but although this yielded a reasonable result, the formula was long and not easy to interpret. It was clear we needed another approach. In the next two sections we'll find an upper and lower bound.

4.2 Lower bound for T_{00}^C

We know

$$E(T_{00}^C) = 1 + \sum_{i=1}^{\infty} P(R_t = i) E(T_{i0}^C) = \sum_{i=1}^{\infty} P(R_t = i) \sum_{k=1}^{i} \frac{1}{k\mu}.$$

To obtain an upper-bound, we'll define the following optimisation problem:

Take p_i , $i \in [0..\infty]$, a probability distribution, with the property that it is non-decreasing on $[0..\lfloor\lambda\rfloor]$. This property clearly applies to the Poisson distribution. Denote f such a distribution, for which $\sum_{i=1}^{\infty} f_i \sum_{k=1}^{i} \frac{1}{k\mu}$ is minimal.

It is clear that we want to have as much probability-mass near 0 as possible, since $\ln(i)$ is an increasing function. But we also have to comply with the non-decreasing part. So the distribution f clearly equals:

$$f_i = \begin{cases} \frac{1}{\lfloor \lambda+1 \rfloor} & \text{if } 0 \leq i \leq \lfloor \lambda \rfloor \\ 0 & \text{otherwise} \end{cases}$$

Now we only have to calculate the value of $\sum_{i=1}^{\infty} f_i \ln i$

$$\begin{aligned} E(T_{00}^C) &= \sum_{i=1}^{\infty} P(R_t = i) \sum_{k=1}^{i} \frac{1}{k\mu} \\ &> \sum_{i=1}^{\infty} f_i \sum_{k=1}^{i} \frac{1}{k\mu} \\ &= \sum_{i=1}^{\lfloor \lambda \rfloor} \frac{1}{\lfloor \lambda+1 \rfloor} \sum_{k=1}^{i} \frac{1}{k\mu} \\ &= \sum_{k=1}^{\lfloor \lambda \rfloor} \frac{1}{\lfloor \lambda+1 \rfloor} \sum_{i=k}^{\lfloor \lambda \rfloor} \frac{1}{k\mu} \\ &= \frac{1}{\lfloor \lambda+1 \rfloor} \sum_{k=1}^{\lfloor \lambda \rfloor} \frac{\lfloor \lambda+1 \rfloor - k}{k\mu} \\ &= \sum_{k=1}^{\lfloor \lambda \rfloor} (\frac{1}{k\mu} - \frac{1}{\lfloor \lambda+1 \rfloor \mu}) \\ &= \sum_{k=1}^{\lfloor \lambda \rfloor} \frac{1}{k2\mu} + \sum_{k=1}^{\lfloor \lambda \rfloor} \frac{1}{k2\mu} - \frac{\lfloor \lambda \rfloor}{\lfloor \lambda+1 \rfloor \mu} \\ &> \frac{1}{2\mu} \ln(\lambda). \end{aligned}$$

It is easily verified that the difference of the last two terms in the right hand-side of the last equality is non-negative. This gives us the following very nice result:

$$\begin{aligned} E(T_{00}) &> 1 + E(T_{00}^C) \\ &> 1 + \tfrac{1}{2\mu} \ln(\lambda). \end{aligned} \tag{6}$$

4.3 Upper bound for T_{00}^C

Now we've found a satisfying result for the lower-bound it's time to sink our teeth in the upper-bound. We'll do this in similar fashion, but obviously with another optimisation problem.

Take f_i , $i \in [0..\infty]$, a distribution, with mean λ and maximum value of $g(f) = \sum_{i=1}^{\infty} f_i(1 + \ln(i))$. We'll restrict ourselves to $\lambda \notin \mathbf{N}$. We now assert that the following distribution f is an optimal solution for this problem, with

$$f_i = \begin{cases} \alpha & \text{if } i = \lfloor \lambda \rfloor \\ \beta & \text{if } i = \lfloor \lambda + 1 \rfloor \\ 0 & \text{otherwise} \end{cases}$$

and $\alpha\lfloor\lambda\rfloor + \beta\lfloor\lambda + 1\rfloor = \lambda$, $\alpha + \beta = 1$ and $\alpha, \beta > 0$.

Let s be a different distribution with $E(s) = \lambda$. Then there are $j, k > 0$ such that $s_{\lfloor\lambda\rfloor - j} > 0$ and $s_{\lfloor\lambda+1\rfloor+k} > 0$. We can construct a different distribution b, that yields a higher result under g, in the following manner.

$$b_i = \begin{cases} s_i - \gamma & \text{if } i = \lfloor \lambda - j \rfloor \\ s_i + \gamma & \text{if } i = \lfloor \lambda \rfloor \\ s_i + \delta & \text{if } i = \lfloor \lambda + 1 \rfloor \\ s_i - \delta & \text{if } i = \lfloor \lambda + 1 + k \rfloor \\ s_i & \text{otherwise,} \end{cases}$$

with $\gamma \leq s_{\lfloor\lambda - j\rfloor}$, $\delta \leq s_{\lfloor\lambda+1+k\rfloor}$ $\gamma, \delta > 0$ and $\gamma j - \delta k = 0$. This clearly makes b a distribution again with mean λ. We'll calculate the difference $g(b) - g(s)$. But before we prove that this indeed yields a higher value, we'll introduce a lemma to facilitate the calculation.

Lemma 2. *If $\alpha j - \beta k = 0$ for $\alpha, \beta, j, k > 0$ then $\alpha \ln(\lambda - j) + \beta \ln(\lambda + k) < (\alpha + \beta)\ln(\lambda)$.*

Proof.

$$\begin{aligned} &(\alpha + \beta)\ln(\lambda) - (\alpha \ln(\lambda - j) + \beta \ln(\lambda + k)) \\ &\quad > \quad (\alpha + \beta)\ln(\lambda) - (\alpha \ln(\lambda)(j + 1) + \beta \ln(\lambda)(k + 1)) \\ &\quad = \quad (\alpha + \beta)\ln(\lambda) - \ln(\lambda)(\alpha j + \beta k + \alpha + \beta) \\ &\quad = \quad (\alpha + \beta)\ln(\lambda) - \ln(\lambda)(\alpha + \beta) \\ &\quad = \quad 0. \end{aligned}$$

□

Hence, to prove that this is indeed a better distribution, we'll apply this lemma to $g(b) - g(s)$.

$$\begin{aligned}
\sum_{i=1}^{\infty}(b_i - s_i)\ln(i) &= \gamma(\ln\lfloor\lambda\rfloor - \ln\lfloor\lambda - j\rfloor) + \delta(\ln\lfloor\lambda+1\rfloor - \ln\lfloor\lambda+k+1\rfloor) \\
&> \gamma(\ln\lfloor\lambda+1\rfloor - \ln\lfloor\lambda - j + 1\rfloor) \\
&\qquad +\delta(\ln\lfloor\lambda+1\rfloor - \ln\lfloor\lambda+k+1\rfloor) \\
&= \ln\lfloor\lambda+1\rfloor(\gamma+\delta) - \gamma\ln\lfloor\lambda+1-j\rfloor - \delta\ln\lfloor\lambda+1+k\rfloor \\
&> 0.
\end{aligned}$$

So $g(b) > g(s)$. And this means that each distribution, apart from f can be adjusted to give a higher value under g. So f is clearly the optimal solution. Now we'll calculate the value of this optimal solution f:

$$\begin{aligned}
\textstyle\sum_{i=1}^{\infty} f_i \ln(i) &= \alpha\ln\lfloor\lambda\rfloor + \beta\ln\lfloor\lambda+1\rfloor \\
&= 1 + \ln(\lambda) - \alpha(\ln(\lambda) - \ln\lfloor\lambda\rfloor) + \beta(\ln(\lambda) - \ln\lfloor\lambda+1\rfloor) \\
&> 1 + \ln(\lambda).
\end{aligned}$$

Combining all the results, we get:

$$\begin{aligned}
T_{00} &< 2 + \textstyle\sum_{i=1}^{\infty} E(T_{i0}^C) \\
&< 2 + \frac{1}{\mu}\textstyle\sum_{i=0}^{\infty} P(X_t = i)(1 + \ln(i)) \\
&< 2 + \frac{1}{\mu}(1 + \textstyle\sum_{i=0}^{\infty} P(X_t = i)\ln(i)) \\
&< 2 + \frac{1}{\mu}(1 + \textstyle\sum_{i=0}^{\infty} f_i \ln(i)) \\
&< 2 + \frac{1}{\mu}(1 + \ln(\lambda)).
\end{aligned} \tag{7}$$

4.4 Studying the bounds

Now we've found those bounds, we want to know their quality. We'll examine a plot comparing T_{00} to the two bounds of T_{00}^C.

The derivation of the lower-bound was obtained considerably faster than the upper-bound. But this has come at a tremendous price of accuracy. We've lost a lot of information in the optimisation step, because the distribution does not have a mean λ, the standard deviation is way higher than λ, and the distribution is not increasing between 0 and λ. So there is a lot of room for better lower-bounds.

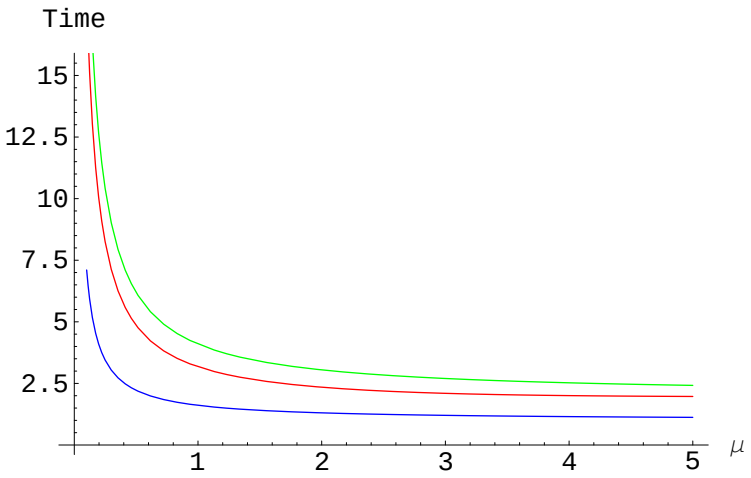

Figure 4: T_{00} with a lower- and upper-bound, for $\lambda = 3$, $\mu \in [0, 01..8]$

5 Behaviour of π_1 at $\lambda = \mu = 0$

5.1 Calculation

It's hard to say something about the behaviour of this function in general, but we're able to produce some results for specific values of λ and μ. First we'll look at π_1 for λ and μ tending to zero. To get some feeling how the function behaves, we've included a plot for these values.

It's clear that π_1 does not exist for $\lambda = \mu = 0$. So we'll have a look at some paths in parameter space leading to 0. To do this, we need a way to calculate π_1. We can do this by approximating the transition matrix P by the matrix ${}^N P$ introduced in 3.3 , with $N = 2$. Now we've got the matrix

$$ {}_2P = \begin{bmatrix} 1 - \lambda e^{-\lambda} & 1 - e^{-\mu} \\ \lambda e^{-\lambda} & e^{-\mu} \end{bmatrix}, $$

we can calculate the solutions ${}_2\pi_0$ and ${}_2\pi_1$. We obtain

$$ \begin{aligned} {}_2\pi_0 &= \frac{1 - e^{-\mu}}{1 - e^{-\mu} + \lambda e^{-\lambda}} \\ {}_2\pi_1 &= \frac{\lambda e^{-\lambda}}{1 - e^{-\mu} + \lambda e^{-\lambda}}. \end{aligned} \tag{8} $$

For the error between this formula and the π_1, we can calculate, we have:

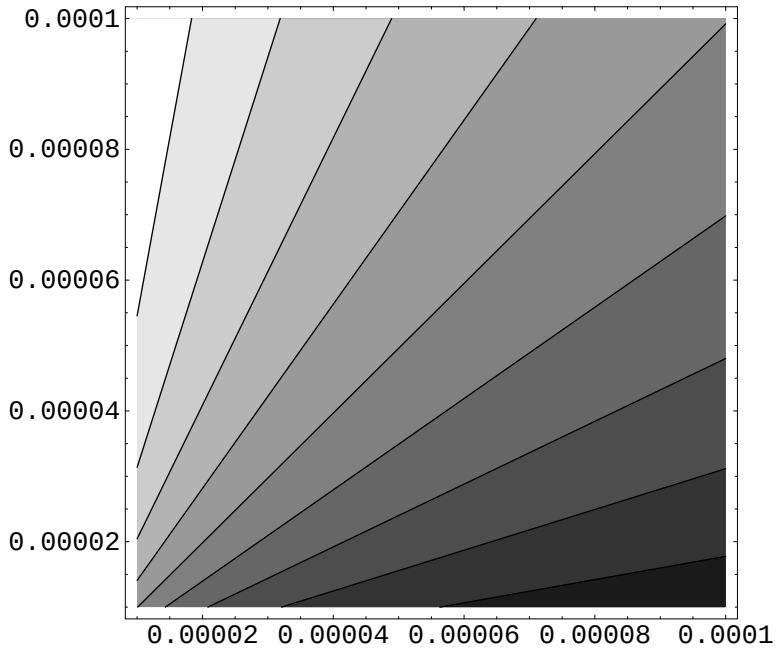

Figure 5: A contour plot of π_1 for $\mu \in [0, 01..8]$, $\lambda \in [0, 01..8]$

$\left|\pi_i - {}^N\pi_i\right| < \frac{2\delta_N}{E(T_{00})} < 2\delta_N$, with $\delta_N = \frac{\lambda}{\mu}\sum_{j=N-1}^{\infty} P_{0j} = \frac{\lambda}{\mu}(1-e^{-\lambda})$. Combining this yields $|\pi_i - {}_2\,\pi_i| < \frac{2\lambda}{\mu}(1-e^{-\lambda})$.

5.2 Linear paths

Since we see some nice straight lines in this picture, we'll first try a path to zero, with $\lambda = \alpha\mu$.

$$\begin{aligned}
{}_2\pi_1 &= \lim_{\mu\to 0} \frac{\lambda e^{-\lambda}}{1 - e^{-\mu} + \lambda e^{-\lambda}} \\
&= \lim_{\mu\to 0} \frac{\alpha\mu e^{-\alpha\mu}}{1 - e^{-\mu} + \alpha\mu e^{-\alpha\mu}} \qquad \text{by l'Hopital} \\
&= \lim_{\mu\to 0} \frac{-\alpha^2\mu e^{-\alpha\mu} + \alpha e^{-\alpha\mu}}{e^{-\mu} - \alpha^2\mu e^{-\alpha\mu} + \alpha e^{-\alpha\mu}} \\
&= \lim_{\mu\to 0} \frac{\alpha e^{-\alpha\mu}}{e^{-\mu} + \alpha e^{-\alpha\mu}} \\
&= \frac{\alpha}{1+\alpha}.
\end{aligned}$$

Further, we have

$$\begin{aligned}
|\pi_1 - {}_2\,\pi_1| &< \frac{2\lambda}{\mu}(1-e^{-\lambda}) \\
&= \lim_{\mu\to 0} \frac{2\alpha\mu}{\mu}(1-e^{-\alpha\mu}) \\
&= \lim_{\mu\to 0} 2\alpha(1-e^{-\alpha\mu}) \\
&= 0.
\end{aligned}$$

So this is the exact value of π_1 when following a path to 0, with $\lambda = \alpha\mu$:

$$\lim_{\mu\to 0,\ \lambda=\alpha/\mu} \pi_1(\lambda,\mu) = \frac{\alpha}{1+\alpha}. \tag{9}$$

5.3 More extreme paths

If we have a closer look at (9) we see that we can get π_1 arbitrarily close to 1, but never quite there. But if we take a path for which μ decreases more

rapidly than linear as compared to λ, we might get $\pi_1 = 1$. And if we'd do this the other way around, we just might get $\pi_1 = 0$.

So let's try it with the formula $\lambda^{3/2} = \mu$, (taking a square would make the error to big)

$$\begin{aligned}
{}_2\pi_1 &= \lim_{\lambda\to 0} \frac{\lambda e^{-\lambda}}{1 - e^{-\mu} + \lambda e^{-\lambda}} \\
&= \lim_{\lambda\to 0} \frac{\lambda e^{-\lambda}}{1 - e^{-\lambda^{3/2}} + \lambda e^{-\lambda}} && \text{by l'Hopital} \\
&= \lim_{\lambda\to 0} \frac{-\lambda e^{-\lambda} + e^{-\lambda}}{3/2\lambda^{1/2} e^{-\lambda^{3/2}} - \lambda e^{-\lambda} + e^{-\lambda}} \\
&= \lim_{\lambda\to 0} \frac{e^{-\lambda}}{e^{-\lambda}} \\
&= 1.
\end{aligned}$$

Calculating the error, we get:

$$\begin{aligned}
|\pi_1 -_2 \pi_1| &< \frac{2\lambda}{\mu}(1 - e^{-\lambda}) \\
&= \lim_{\lambda\to 0} \frac{2\lambda}{\lambda^{3/2}}(1 - e^{-\lambda}) \\
&= \lim_{\lambda\to 0} 2\frac{1 - e^{-\lambda}}{\lambda^{1/2}} \\
&= \lim_{\lambda\to 0} 2\frac{e^{-\lambda}}{1/2\lambda^{-1/2}} \\
&= \lim_{\lambda\to 0} 4e^{-\lambda}\lambda^{-1/2} \\
&= 0.
\end{aligned}$$

So this π_1 equals 1 if we walk to $\lambda = \mu = 0$, along the path $\lambda^{3/2} = \mu$. We will see that this is the only place where π_1 reaches 1. If we take $\lambda^{-3/2} = \mu$ we get, by executing a virtually identical calculation, that $\pi_1 = 0$.

6 Discussion

Firstly, we have the island near $\mu = 0,77$ for λ big. It is not hard to prove that for μ big, π_1 is almost optimal for this value. Due to time-constraints I haven't been able to give it a separate section, to discuss it in detail. We

can note however, that if μ is big, the probability to arrive at $X_t = 1$ from a non-empty state gets smaller when μ increases. So we can try to look only at the probability to get to $X_t = 1$ from $X_t = 0$. Furthermore, if μ is big, we can take $E(T_{00}) = 1 + P(R_t > 0)$ as a reasonable approximation.

We proceed to calculate an estimated value of π_1, given by inserting this two approximations in the formula $\pi_1 = EA_1/E(T_{00})$. Optimising this formula for given μ, yields an optimal value for $\lambda = 1 + \text{ProductLog}(\frac{1}{2e}) \sim 0.77$.

As stated before, it was very hard to calculate the convergence speed of π_1. In section 3.2 we proved that the convergence speed is exponential, but we couldn't give a specific value. We have tried a lot of approaches to tackle this problem, but in the end we had to give in because of time constraints. The best approach seems to be to use coupling times. This still isn't easy to solve, but by dividing the state-space in two parts (states in which at least one of both states is zero, and the states where neither is zero), one gets a nice systems which probably is easy to solve. At the end of the paper, we have included some plots of the approximate convergence speed, obtained by taking the second largest eigenvalue of a big NxN matrix. This clearly doesn't give an exact value, but it gives a good idea how wildly the convergence speed fluctuates.

The last point of interest is the calculation of an optimal value for λ given μ, and vice versa. In the previous sections we have obtained upper and lower-bounds which give bounds to the area we have to search for the optimal values. It can be shown that the value of π_1 in the vicinity of a λ, μ cannot change too wildly and is subjected to some constraints. This way we can get a reasonable approximation for the optimal solutions, complete with an error margin.

The proof for this construction will not be given in this paper. However, there is a plot included at the end of the paper, which was obtained using these methods. The plot shows the optimal value for λ, given μ. The optimal value for μ, given λ turns out to be $\mu \to 0$, which seems logical.

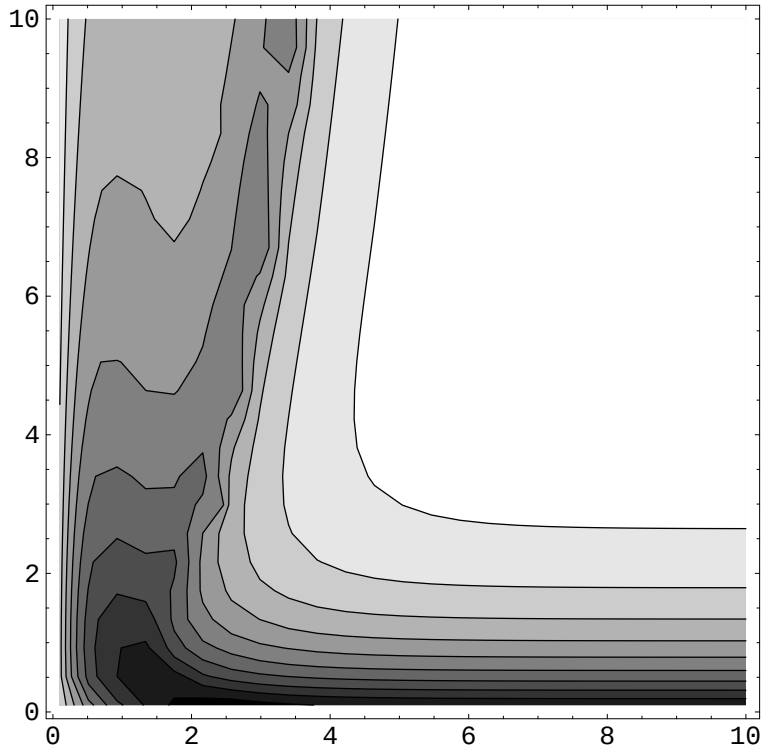

Figure 6: An approximation of the convergence-speed for π_1 with $\mu \in [0,01..10]$, $\lambda \in [0,01..10]$

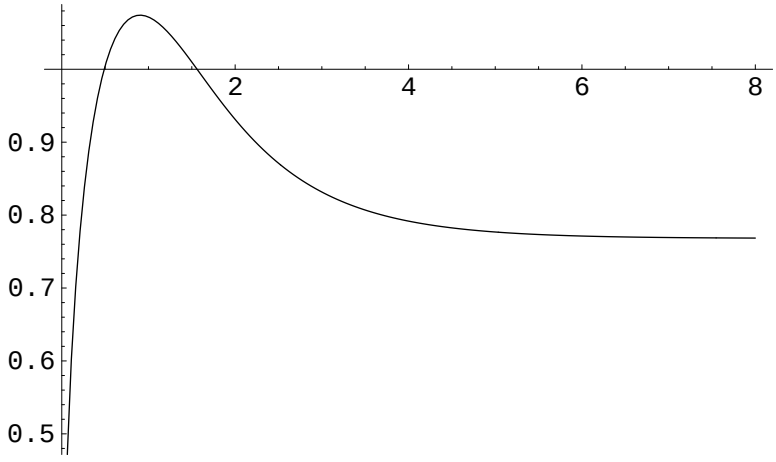

Figure 7: The value of λ for which π_1 is maximal, given μ, $\mu \in [0,01..8]$

References

[1] A. Rukhin, I. Bebu, *Stochastic model for the number of atoms in a magneto-optical trap.* To appear in Prob. Eng. Inf. Sc. 2006.

[2] A. Hordijk and F.M. Spieksma (1992), On ergodicity and recurrence properties of a Markov chain with an application to an open Jackson network. *Adv. Appl. Prob.* **24**, 343–376.