Bucket Brigade Convergence on Markov Chains

Tom Westerdale

June 9, 2022

Abstract

The bucket brigade can be used to estimate state values in a strongly connected finite state Markov chain. Such a bucket brigade is the simplest temporal difference method, used in its simplest context. It operates with a constant step size. A bucket brigade with a smaller step size gives better value estimations. More precisely, the bucket brigade on a strongly connected finite state Markov chain converges in probability as the step size goes to zero, however rapidly or slowly it goes to zero. But in operation the step size is constant. This technical report gives the convergence proof in careful pedantic detail. I confess the proof is long and tedious, with few interesting insights. But at least we finally have a proof.

More Detailed Abstract

The bucket brigade converges in probability on a strongly connected finite state Markov chain.

We suppose a payoff number is attached to each state, the payoff received when the chain enters that state. Let $\bar{m}$ be the average payoff per time unit. Excess payoff is payoff minus $\bar{m}$. The value c_i of state i is the expectation of the total excess payoff received over all time if the chain starts in state i . (Value defined using Cesaro sum, no discount.) Each state has a certain amount of cash. We look at a particular bucket brigade. In each time unit there is a change in cash. The cash change has three components. (1) Proportion ε of the current state's cash is passed to the previous state. (2) Cash equal to ε times the excess payoff of the current state is added to the cash balance of the current state. (3) Cash equal to the amount added in component 2 is subtracted from the total cash. The apportionment of the subtraction among states is given by a constant apportionment vector. Any such vector will do.

We show that the cash balance of state i converges in probability to an equilibrium cash balance u_i as $\varepsilon \to 0$. We show that if $\bar{u}$ is the weighted average of the equilibrium cash balances, then $c_i = u_i - \bar{u}$. So a state's current cash balance minus the average cash balance is a good estimate of state value.

We then look at what happens if we eliminate component 3 or if we don't multiply by ε in component 2 or if we use payoffs instead of excess payoffs or if we make other common modifications. In most but not all cases we arrive at a similarly appropriate estimate of state value. There are no big surprises, but there are some small ones. Using payoffs instead of excess payoffs can cause problems. But any problems can be dealt with, and most things work as expected.

1 Basic Definitions

1.1 Introduction

This essay proves convergence in probability of the bucket brigade on finite state Markov chains.

The bucket brigade can be used to estimate Markov chain state values. John Holland originally proposed the bucket brigade for use in adaptive rule based systems.[1] The attractiveness of the bucket brigade lies in the way it meshes gracefully with the mechanisms of such systems. These are messier than Markov chains. One way of analysing rule based systems is to view a rule as a set of Markov chain state transitions. This gives a useful way of using what we know about Markov chains to illuminate rule based systems. An understanding of bucket brigade behavior on Markov chains gives us a solid starting point for analysis of bucket brigade behavior in rule based systems.

The bucket brigade is a particularly simple example of what we now call temporal difference methods. I think of the bucket brigade on Markov chains as a sort of temporal difference type specimen, since it captures the essence of temporal difference methods in the simplest possible way.[2] The strengths and weaknesses of

[1] for example, classifier systems

[2] The bucket brigade is a special case of what the field of Reinforcement Learning calls Linear TD(0).

the bucket brigade give us insights into the strengths and weaknesses of temporal difference methods in general.

But this essay does not discuss any of these insights, and it does not discuss the place of the bucket brigade among the various temporal diffference methods. It does not discuss rule based systems at all. Here, we discuss *only* the bucket brigade on Markov chains, and we discuss *only* its convergence.

It has always seemed obvious that the bucket brigade converges, at least in a system as simple as a Markov chain. But a convergence proof is clearly desirable. This essay provides one.

The bucket brigade iteratively adjusts a vector. We prove here that the vector converges to a limit in probability as the step size goes to zero. We prove that it so converges no matter how rapidly or slowly the step size goes to zero. This last is important. In the bucket brigade, there is no schedule of step size reduction. The step size is constant. It may be small or it may be large, but it's constant. What our convergence proof tells us, roughly, is that the vector tends to dance close to the limit most of the time, and that the smaller our step size, the closer it tends to dance. We shall make all this precise in our discussion.

It is the purpose of this essay to place beyond doubt something that seemed rather obvious in the first place. So I have given each step in pedantic detail. I admit I have made heavy weather of it, and the essay is longer than such a simple result would seem to warrant. But the various limits are interdependent, and I feel I need to make absolutely clear to the reader and to myself that the argument here is now nowhere circular. If you find any remaining errors, please let me know.

Of course I keep thinking there must be a simpler way to do this. I would love to know of any simpler proof.

The proof begins in section 2. Section 1 merely introduces our notation, gives some of our basic definitions, and states some preliminary formal facts about matrices that we will use. It also makes some further motivational comment that uses the defined quantities.

In addition to the facts mentioned in this section, our proof uses the ergodicity of our strongly connected Markov chain. The ergodicity does not follow directly from the facts mentioned in this section, but the ergodicity is a standard fact. Subsection 3.2 refers to one of the ergodicity proofs in the literature and converts the result there to the notation we need for our proof.

All vectors in this essay are row vectors. Their transposes are column vectors. Each of our vectors will be a row of N complex numbers. A vector is called *stochastic* if its entries are all non-negative real numbers and the entries sum to 1.

The vector $\mathbf{e}$ is the vector of all ones. The vector $\mathbf{e}_i$ is the vector whose i'th entry is 1 and whose other entries are all 0. Like all our vectors, these vectors are written in bold face. The italic letter e with no subscript will signify the real number whose natural logarithm is 1 . The exponential of a number t we write e^t , using the italic e .

Our matrices will be $N \times N$ matrices of complex numbers. I assume throughout that $N > 1$. A matrix is *non-negative* if all of its entries are non-negative real numbers. A matrix is row stochastic if each of its rows is a stochastic vector. So we see that a matrix B is row stochastic of and only if B is non-negative and $B\mathbf{e}^\top = \mathbf{e}^\top$. If the columns are also stochastic vectors (so $\mathbf{e}B = \mathbf{e}$) then we say the matrix is bistochastic.

We will be discussing a Markov chain that has N states. We will use the letter P for the transition probability matrix of our chain. The matrix P has entries defined as follows. If $i \to j$ is a legal chain transition, then P_{ij} is the transition probability. By that I mean that P_{ij} is the conditional probability that the next state is j , given that the current state is i . If $i \to j$ is illegal then of course $P_{ij} = 0$. The transition probability of every legal transition is positive.

We see that P is row stochastic.

Our Markov chain will be *strongly connected*, which means that for any ordered pair of states $\langle i, j\rangle$, there is a sequence of legal transitions [3] that takes the chain from state i to state j . "Strongly connected" and "irreducible" are synonyms in this essay.[4] Whether a chain is strongly connected (irreducible) or not is completely determined by which entries in its transition probability matrix are positive. The transition probability matrix is called irreducible (strongly connected) if and only if the chain is strongly connected (irreducible). So the matrix P of our chain is irreducible (strongly connected).

[3] a trajectory

[4] Another synonym is "regular".

1.2 The Useful Matrix $\hat{P}$.

We will be using a few facts about our matrix P that are trivially deducible from the Frobenius-Perron Theorem on non-negative matrices[5] and the fact that P is row stochastic and irreducible. I will call those facts the Frobenius Facts.

We already know that $\mathbf{e}^\top$ is a right eigenvector of P with eigenvalue 1. (That is, $P\mathbf{e}^\top = \mathbf{e}^\top$.) The Frobenius Facts expand on this.

> **Frobenius Facts about any row stochastic irreducible matrix P :**
> There exists a vector $\mathbf{v}$ whose every entry is a positive real and which is a left eigenvector of P with eigenvalue 1. The eigenvalue 1 has multiplicity 1. The set of left eigenvectors that have eigenvalue 1 form a one-dimensional subspace. The set of right eigenvectors that have eigenvalue 1 form a one-dimensional subspace. No eigenvalue of P has modulus greater than 1. If every entry in P is positive, then the eigenvalue 1 exceeds all the other eigenvalues in modulus.

Let us take a vector $\mathbf{v}$ like the one mentioned in the Frobenius facts. We normalize it. That is, we define the vector $\tilde{\mathbf{p}} = (\mathbf{v}\mathbf{e}^\top)^{-1}\mathbf{v}$. We see that $\tilde{\mathbf{p}}$ is a stochastic vector and that it is a left eigenvector of P with eigenvalue 1. From the Frobenius Facts, we see that every left eigenvector with eigenvalue 1 is a scalar multiple of $\tilde{\mathbf{p}}$. Thus $\tilde{\mathbf{p}}$ is the unique stochastic left eigenvector with eigenvalue 1.

We see that every entry in $\tilde{\mathbf{p}}$ is a positive real, that $\tilde{\mathbf{p}}P = \tilde{\mathbf{p}}$, and that $\tilde{\mathbf{p}}\mathbf{e}^\top = 1$.

Lemma 1 *The number 1 is not an eigenvalue of* $P - \mathbf{e}^\top\tilde{\mathbf{p}}$.

Proof: Suppose it is.[6] Then there is a non-zero vector $\mathbf{v}$ such that

$$\mathbf{v}(P - \mathbf{e}^\top\tilde{\mathbf{p}}) = \mathbf{v} \quad . \tag{1}$$

Now
$(P - \mathbf{e}^\top\tilde{\mathbf{p}})\,\mathbf{e}^\top = 0$,

so the equation
$(\mathbf{v}\,(P - \mathbf{e}^\top\tilde{\mathbf{p}}))\,\mathbf{e}^\top = \mathbf{v}\,((P - \mathbf{e}^\top\tilde{\mathbf{p}})\,\mathbf{e}^\top)$
becomes

$$\mathbf{v}\mathbf{e}^\top = 0 \quad .$$

It follows by multiplying out, that
$\mathbf{v}\,(P - \mathbf{e}^\top\tilde{\mathbf{p}}) = \mathbf{v}P - \mathbf{v}\,\mathbf{e}^\top\tilde{\mathbf{p}} = \mathbf{v}P$.
From this and equation (1), we obtain
$\mathbf{v}P = \mathbf{v}$.
So $\mathbf{v}$ is a left eigenvector of P with eigenvalue 1.
Then by the Frobenius Facts, $\mathbf{v}$ is a scalar multiple of $\tilde{\mathbf{p}}$. That is,
$\mathbf{v} = \lambda\tilde{\mathbf{p}}$ for some scalar λ.
Multiplying the last equation on the right by $\mathbf{e}^\top$ gives
$\mathbf{v}\mathbf{e}^\top = \lambda\tilde{\mathbf{p}}\mathbf{e}^\top$.
But $\mathbf{v}\mathbf{e}^\top = 0$ and $\tilde{\mathbf{p}}\mathbf{e}^\top = 1$, so
$\lambda = 0$.
Since $\mathbf{v} = \lambda\tilde{\mathbf{p}}$, we see that
$\mathbf{v} = 0$.
Contradiction.
■

We now define

$$\hat{P} = I - P + \mathbf{e}^\top\tilde{\mathbf{p}} \quad .$$

Lemma 2 $\hat{P}$ *is nonsingular.*

Proof: If $\hat{P}$ is singular then there is a non-zero vector $\mathbf{v}$ such that $\mathbf{v}\hat{P} = 0$. This means that
$\mathbf{v} = \mathbf{v}\,(P - \mathbf{e}^\top\tilde{\mathbf{p}})$,

[5] Gantmacher [1] gives a complete proof of the Frobenious-Perron Theorem.

[6] The reader may feel this proof is unnecessary. After all, we have subtracted from P the term in the spectral decomposition of P that pertains to the eigenvalue 1, so of course we have removed the eigenvalue 1. Nevertheless, I give here an elementary proof.

so $\mathbf{v}$ is an eigenvector of $P - \mathbf{e}^\top\tilde{\mathbf{p}}$ with eigenvalue 1.
But we showed that 1 is not an eigenvalue of $P - \mathbf{e}^\top\tilde{\mathbf{p}}$.
Contradiction.
■

1.3 State Probabilities

Define

$$P^{[n]} = \sum_{k=o}^{n-1}(P^k - \mathbf{e}^\top\tilde{\mathbf{p}}) \qquad .$$

$P^{[n]}$ is the n'th partial sum of a series that we want to try to sum. We note that
$P^{[n]}\mathbf{e}^\top = 0$ and
$P^{[n]}P = \sum_{k=1}^{n}(P^k - \mathbf{e}^\top\tilde{\mathbf{p}})$.
This last fact allows us to write $\sum_{k=0}^{n}(P^k - \mathbf{e}^\top\tilde{\mathbf{p}})$ in two different ways and say they are equal.
$P^{[n]} + (P^n - \mathbf{e}^\top\tilde{\mathbf{p}}) = (I - \mathbf{e}^\top\tilde{\mathbf{p}}) + P^{[n]}P$
This simplifies as follows.
$P^{[n]} + P^n = I + P^{[n]}P$
$P^{[n]} - P^{[n]}P = I - P^n$
$P^{[n]}(I - P) = I - P^n$
Since $P^{[n]}\mathbf{e}^\top$ is zero, we have
$P^{[n]}\mathbf{e}^\top\tilde{\mathbf{p}} = 0$.
Adding the last two equations gives us
$P^{[n]}\,\hat{P} = I - P^n$.
Since $\hat{P}$ is nonsingular, we can multiply on the right by $\hat{P}^{-1}$.

$$P^{[n]} = (I - P^n)\hat{P}^{-1} \tag{2}$$

Dividing by n gives
$\frac{1}{n}P^{[n]} = (\frac{1}{n}I - \frac{1}{n}P^n)\hat{P}^{-1}$.
Since the entries in P^n are all between 0 and 1 inclusive,
if we let $n \to \infty$ the right side becomes 0, and we have

$$\lim_{n\to\infty}\frac{1}{n}P^{[n]} = 0 \quad .$$

We can rewrite $P^{[n]}$ as
$P^{[n]} = \left(\sum_{k=o}^{n-1}P^k\right) - n\,\mathbf{e}^\top\tilde{\mathbf{p}}$.
Transposing gives us
$\sum_{k=0}^{n-1}P^k = P^{[n]} + n\mathbf{e}^\top\tilde{\mathbf{p}}$.
Dividing by n gives us
$\frac{1}{n}\sum_{k=0}^{n-1}P^k = \frac{1}{n}P^{[n]} + \mathbf{e}^\top\tilde{\mathbf{p}}$.
If we let $n \to \infty$, the right side goes to the limit $\mathbf{e}^\top\tilde{\mathbf{p}}$, so we obtain

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}P^k \;=\; \mathbf{e}^\top\tilde{\mathbf{p}} \quad . \tag{3}$$

From this we see that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}(\mathbf{e}_i P^k \mathbf{e}_j^\top) \;=\; \tilde{p}_j \quad . \tag{4}$$

It turns out that $\tilde{p}_j$ is the proportion of time after a visit to i that the chain spends in state j . (Note that the proportion is independent of i .) It looked to me as if the last equation said exactly that, but it doesn't. The left side is the proportion over all runs starting at i . But we are interested in the proportion in a single run. This too is $\tilde{p}_j$ in almost all runs. (It wouldn't necessarily be so if P weren't strongly connected.) This ergodicity is made precise and discussed in subsection 3.2.

So $\tilde{p}_j$ is the proportion of time the chain spends in state j, or the unconditional probability of finding the chain in state j . The probability $\tilde{p}_j$ is called the mean limiting absolute probability of state j , and the vector $\tilde{\mathbf{p}}$ is called the mean limiting absolute state probability vector.

1.4 Payoffs and Values

Now $\tilde{\mathbf{p}}\hat{P} = \tilde{\mathbf{p}}$ so $\tilde{\mathbf{p}}\hat{P}^{-1} = \tilde{\mathbf{p}}$.
From the definition of $\hat{P}$, we have
$P^k\hat{P} = P^k - P^{k+1} + \mathbf{e}^\top\tilde{\mathbf{p}}$. We also have
$\mathbf{e}^\top\tilde{\mathbf{p}}\hat{P} = \mathbf{e}^\top\tilde{\mathbf{p}}$.
Subtracting the last equation from the previous one gives
$(P^k - \mathbf{e}^\top\tilde{\mathbf{p}})\hat{P} = P^k - P^{k+1}$.
Summing both sides as k goes from 0 to $n-1$ gives
$P^{[n]}\hat{P} = I - P^n$. Multiplying by $\hat{P}^{-1}$ on the right gives
$P^{[n]} = \hat{P}^{-1} - P^n\hat{P}^{-1}$.
Taking the average, we obtain
$\frac{1}{\ell}\sum_{n=0}^{\ell-1} P^{[n]} = \hat{P}^{-1} - (\frac{1}{\ell}\sum_{n=0}^{\ell-1} P^n)\hat{P}^{-1}$.
If we let $\ell \to \infty$, then equation(3) tells us that the right side goes to $\hat{P}^{-1} - (\mathbf{e}^\top\tilde{\mathbf{p}})\hat{P}^{-1}$,
which is $\hat{P}^{-1} - \mathbf{e}^\top\tilde{\mathbf{p}}$. Therefore,
$\lim_{\ell\to\infty} \frac{1}{\ell}\sum_{n=0}^{\ell-1} P^{[n]} = \hat{P}^{-1} - \mathbf{e}^\top\tilde{\mathbf{p}}$.
By the definition of Cesaro sum, the left hand side is the Cesaro sum $\sum\!\!\!\!\text{c}_{n=0}^{\infty}(P^n - \mathbf{e}^\top\tilde{\mathbf{p}})$, so the Cesaro sum exists and we have

$$\sum_{n=0}^{\infty}\!\!\!\!\!\text{c}\ (P^n - \mathbf{e}^\top\tilde{\mathbf{p}}) = \hat{P}^{-1} - \mathbf{e}^\top\tilde{\mathbf{p}} \quad . \tag{5}$$

A non-negative $N \times N$ matrix B is said to be *primitive* if there is a positive integer k such that B^k has every entry positive. Obviously, primitive matrices are irreducible. We know from the Frobenius Facts that if P is row stochastic and irreducible then 1 is an eigenvalue and all eigenvalues have modulus less than or equal to 1.

Lemma 3 *If P is a matrix that is row stochastic and primitive then every eigenvalue other than the number 1 itself has a modulus that is less than 1.*

Proof: Suppose P is row stochastic and that P^n has every entry positive.
Take any nonzero eigenvalue λ of P such that $\lambda \neq 1$. We will be done of we can show $|\lambda| < 1$.

There is an Eigenvector $\mathbf{v}$ such that $\mathbf{v}P = \lambda\mathbf{v}$. Multiplying on the right by $\mathbf{e}^\top$ gives $\mathbf{v}\mathbf{e}^\top = \lambda\mathbf{v}\mathbf{e}^\top$.
Since $\lambda \neq 1$, we have
$\mathbf{v}\mathbf{e}^\top = 0$.

We are now going to show $\lambda^n \neq 1$.
We assume $\lambda^n = 1$ and derive a contradiction.
We have both $\tilde{\mathbf{p}}P^n = \tilde{\mathbf{p}}$ and $\mathbf{v}P^n = \lambda^n\mathbf{v}$, so if $\lambda^n = 1$ then both $\tilde{\mathbf{p}}$ and $\mathbf{v}$ are eigenvectors of P^n with eigenvalue 1. Since P^n is row stochastic and irreducible, the Frobenius Facts applied to P^n say that the eigenvectors of P^n with eigenvalue 1 form a one dimensional subspace, so there is a scalar ρ such that $\mathbf{v} = \rho\tilde{\mathbf{p}}$. Then multiplying by $\mathbf{e}^\top$ on the right gives $0 = \rho$, and so $\mathbf{v} = \rho\tilde{\mathbf{p}} = 0$. But $\mathbf{v}$ is an eigenvector so $\mathbf{v} \neq 0$. Contradiction.

So we have shown that $\lambda^n \neq 1$. But λ^n is an eigenvalue of P^n , and P^n has every entry positive, so the Frobenius Facts tell us that $|\lambda^n| < 1$.

Therefore, $|\lambda| < 1$.
■

Lemma 4 *If P is row stochastic, irreducible, and primitive, then equation (3) simplifies to* $\lim_{n\to\infty} P^n = \mathbf{e}^\top\tilde{\mathbf{p}}$ *and the Cesaro sum in equation (5) can be replaced by the ordinary infinite sum.*

Proof: Suppose P is row stochastic, irreducible, and primitive. Note that $(P - \mathbf{e}^\top\tilde{\mathbf{p}})\mathbf{e}^\top = 0$.

Suppose λ is a nonzero eigenvalue of $P - \mathbf{e}^\top\tilde{\mathbf{p}}$. Then there is an eigenvector $\mathbf{v}$ such that
$\mathbf{v}(P - \mathbf{e}^\top\tilde{\mathbf{p}}) = \lambda\mathbf{v}$.
Multiplying on the right by $\mathbf{e}^\top$ gives $0 = \lambda\mathbf{v}\mathbf{e}^\top$. Since λ is nonzero, $\mathbf{v}\mathbf{e}^\top = 0$.
Therefore, multiplying out gives
$\mathbf{v}(P - \mathbf{e}^\top\tilde{\mathbf{p}}) = \mathbf{v}P$.
So we see that $\mathbf{v}P = \lambda\mathbf{v}$, and λ is an eigenvalue of P .

So every nonzero eigenvalue of $P - \mathbf{e}^\top\tilde{\mathbf{p}}$ is an eigenvalue of P .
Lemma 1 says that 1 is not an eigenvalue of $P - \mathbf{e}^\top\tilde{\mathbf{p}}$.
It follows from this and lemma 3 that since P is primitive, every eigenvalue of $P - \mathbf{e}^\top\tilde{\mathbf{p}}$ has modulus less than 1.

Then since every eigenvalue of $P - \mathbf{e}^\top \tilde{\mathbf{p}}$ has modulus less than 1, we have [7] $lim_{n\to\infty}(P - \mathbf{e}^\top \tilde{\mathbf{p}})^n = 0$.

If $n \geq 1$ then $(P - \mathbf{e}^\top \tilde{\mathbf{p}})^n = P^n - \mathbf{e}^\top \tilde{\mathbf{p}}$. Letting $n \to \infty$, we obtain $\lim_{n\to\infty} P^n = \mathbf{e}^\top \tilde{\mathbf{p}}$.

Equation (2) said that $P^{[n]} = (I - P^n)\hat{P}^{-1}$.
If we let $n \to \infty$, the right side approaches $(I - \mathbf{e}^\top \tilde{\mathbf{p}})\hat{P}^{-1}$, which is $\hat{P}^{-1} - \mathbf{e}^\top \tilde{\mathbf{p}}$. So $\lim_{n\to\infty} P^{[n]} = \hat{P}^{-1} - \mathbf{e}^\top \tilde{\mathbf{p}}$.
The left side of that last equation is $\sum_{k=0}^{\infty}(P^k - \mathbf{e}^\top \tilde{\mathbf{p}})$,
so we have equation (5), but with an ordinary infinite sum.
■

Attached to each state is a fixed payoff. The payoff of state i is m_i .
The vector $\mathbf{m}$ of payoffs is fixed and unchanging.

We define $\bar{m} = \tilde{\mathbf{p}}\,\mathbf{m}^\top$, so $\bar{m}$ is the average payoff per time unit.

We define $a_i = m_i - \bar{m}$. The quantity a_i is called the *excess payoff* of state i. The vector $\mathbf{a}$ is the vector of state excess payoffs. We note that $\tilde{\mathbf{p}}\,\mathbf{a}^\top = 0$.

Multiplying equation (5) on the right by $\mathbf{a}^\top$ gives

$$\sum_{k=0}^{\infty}{}_{\!\mathrm{C}}(P^k \mathbf{a}^\top) = \hat{P}^{-1}\mathbf{a}^\top \quad .$$

We define the vector $\mathbf{c}$ by

$$\mathbf{c}^\top = \hat{P}^{-1}\mathbf{a}^\top \quad . \tag{6}$$

So we see that

$$\mathbf{c}^\top = \sum_{k=0}^{\infty}{}_{\!\mathrm{C}}(P^k \mathbf{a}^\top) \quad . \tag{7}$$

The average total excess payoff we receive in the n time units that start with a visit to state i is $\sum_{k=0}^{n-1}(\mathbf{e}_i P^k \mathbf{a}^\top)$.
If we let $n \to \infty$ and use Cesaro sum, this becomes c_i , so in a sense, c_i is the expectation of the total excess payoff we receive on and following a visit to state i. But even if it were an ordinary sum, the statement relies on the ergodicity we mentioned when discussing (4).

We here call c_i the value of state i .

We note that
$\tilde{\mathbf{p}}\mathbf{c}^\top = 0$.

Now by the definition of $\hat{P}$, we have
$I - P + \mathbf{e}^\top \tilde{\mathbf{p}} = \hat{P}$.
Multiplying on the right by $\hat{P}^{-1}$ gives
$(I - P)\hat{P}^{-1} + \mathbf{e}^\top \tilde{\mathbf{p}} = I$.
Finally, multiplying on the right by $\mathbf{a}^\top$ gives us

$$(I - P)\mathbf{c}^\top = \mathbf{a}^\top \quad . \tag{8}$$

Multiplying on the left by $\mathbf{e}_i$ and transposing gives us

$$c_i - a_i = \sum_j P_{ij} c_j \quad . \tag{9}$$

I now want to make a motivational comment. The average payoff $\bar{m}$ is a function of P . The partial derivative of that function in the direction of increasing P_{ij} is something like $c_j - c_i + a_i$. That partial derivative is a sort of marginal utility of the transition $i \to j$. If we know the state values we can obtain the marginal utilities. Adaptation consists in changing P in an attempt to increase $\bar{m}$. The marginal utilities are important for analyzing adaptive plans, and many useful plans use them, either explicitly or implicitly. What this amounts to is rewarding actions that produce payoff. The bucket brigade is a way of estimating state values. It is employed by many adaptive plans that use marginal utilities explicitly.

All this can be made precise, but we will not do so in this essay. This essay does not discuss adaptation, only the process of estimating the state value vector $\mathbf{c}$ by use of the bucket brigade. In this essay there will be no adaptation, so the matrix P will not change. It is constant.

In this convergence discussion we turn off adaptation, so the matrix P is constant over time. Our results then will not hold exactly when the system is adapting. The hope is that they will hold approximately provided the rate of adaptation is small enough, but at present we have no proof of this.

[7] If B is a square matrix and all its eigenvalues have modulus less than 1, then $\lim_{n\to\infty} B^n = 0$.

1.5 Additional stuff

Lemma 5 *If B is an $N \times N$ row stochastic irreducible matrix and if every diagonal entry is positive, then B is primitive.*

Proof: Suppose B is row stochastic and irreducible and has every diagonal entry positive. Look at the Markov chain of which B is the state transition probability matrix. For any ordered pair of states, there is a path of *exactly* N transitions that goes from the first state to the second such that every transition in the path has nonzero probability. Therefore, every entry in B^N is positive.
■

We define the diagonal matrix D as follows.

$$\text{For each } i\ , \quad \text{we have} \quad D_{ii} \ = \ \tilde{p}_i\ .$$

We define what I call the *frequency matrix*

$$F \ = \ DP \quad .$$

We note that
$\sum_j F_{kj} \ = \ \mathbf{e}_k F \mathbf{e}^\top \ = \ \mathbf{e}_k DP\mathbf{e}^\top \ = \ \mathbf{e}_k D\mathbf{e}^\top \ = \ \mathbf{e}_k \tilde{\mathbf{p}}^\top \ = \ \tilde{p}_k \ .$ and
$\sum_i F_{ik} \ = \ \mathbf{e}F\mathbf{e}_k^\top \ = \ \mathbf{e}DP\mathbf{e}_k^\top \ = \ \tilde{\mathbf{p}}P\mathbf{e}_k^\top \ = \ \tilde{\mathbf{p}}\mathbf{e}_k^\top \ = \ \tilde{p}_k \ .$
So we have this useful equation.

$$\sum_i F_{ik} \ = \ \tilde{p}_k \ = \ \sum_j F_{kj} \tag{10}$$

I am going to use the infinity vector norm and the infinity matrix norm. If B is a complex matrix, $\|B\| = \sup_i \sum_j |B_{ij}|$.
We know that if B and C are complex matrices and if the numbers of rows and columns match up, then

$$\|BC\| \leq \ \|B\|\,\|C\| \quad . \tag{11}$$

And if the numbers of rows and columns match up then $\|B + C\| \leq \|B\| + \|C\|$.
Also $\|tB\| = |t|\,\|B\|$ for any real number t .

For norm purposes, I will treat a column vector as a matrix with just one column, so
$\|\mathbf{v}^\top\| = \sup_i |v_i|$,
This means that provided the numbers of rows and columns match up, we have
$\|B\mathbf{v}^\top\| \ \leq \ \|B\|\,\|\mathbf{v}^\top\|$,
$\|\mathbf{v}^\top + \mathbf{u}^\top\| \ \leq \ \|\mathbf{v}^\top\| + \|\mathbf{u}^\top\|$, and
$\|t\,\mathbf{v}^\top\| \ = \ |t|\,\|\mathbf{v}^\top\|$.

I am going to do something non-standard for norms of row vectors. If I treated a row vector as a one row matrix then its norm would be the sum of the moduli of its entries. I'm not going to do that. It's norm will be the *supremum* of the moduli of its entries. That is, if $\mathbf{v}$ is a row vector, then
$\|\mathbf{v}\| \ = \ \|\mathbf{v}^\top\|$.
This means that the row vector norm is incompatible with the matrix norm. It is not generally the case that $\|\mathbf{v}B\| \ \leq \ \|\mathbf{v}\|\,\|B\|$, and we must be careful not to use this inequality as if it were generally true.

Given any complex $N \times N$ matrix B , we use the exponential power series to define the matrix e^B , the exponential of matrix B .

$$e^B \ = \ \sum_{n=0}^{\infty} \frac{1}{n!} B^n$$

We shall use the following lemma, which is proved in the appendix.

Lemma 34
Let x be any positive real number.
Let B be any complex square matrix for which $\|B\| \ \leq \ 1$.
Let n be an integer variable.
As $n \to \infty$, the quantity $\|e^{tB} - (I + \frac{t}{n}B)^n\|$ goes to zero uniformly for all t in the interval $[0, x]$.

We define the matrix

$$A = D(I - P) \quad . \tag{12}$$

We will need the following fact, which is proved in the appendix.

If ρ is a real variable, then

$$\lim_{\rho\to\infty} \left\| e^{-\rho A} - \tfrac{1}{N}\,\mathbf{e}^\top\mathbf{e} \right\| = 0 \qquad . \tag{13}$$

2 Deterministic Bucket Brigade Convergence

In the bucket brigade there is a pile of cash on each Markov chain state. Cash amounts are real numbers.

The cash movement in the bucket brigade is based on two small positive numbers ε and δ , both less than 1.

When the chain takes transition $i \to j$,
(1) a proportion ε of the cash on state j is passed to state i .
(2) In addition, the payoff of state j determines an amount of cash that is added to the cash on state j .

The amount of cash added in (2) is different in different bucket brigade implementations. In some implementations, the amount of cash added is the payoff m_j . We shall prove convergence for a bucket brigade in which the cash added is the excess payoff a_j . It is easier to prove convergence for this case. Once we have proved convergence we shall look[8] at bucket brigades where the cash added is the payoff m_j , and we shall see that doing this is somewhat problematic.

Actually, in the bucket brigade we discuss, the cash added will be δa_j . I call δ the payoff scaling factor. In many implementations, $\delta = 1$. We shall be particularly interested in implementations where $\delta = \varepsilon$, because formulae are simpler in that case.

So the cash change when transition $i \to j$ occurs is the two changes (1) and (2) above, and we want to prove convergence for such a bucket brigade. But obviously we can't. Look at the total amount of cash in the system. Change (2) causes it to do an unbiased random walk. There is no convergence.

Different bucket brigade implementations deal differently with changes in the total amount of cash. What we shall do is hold the total amount of cash constant so we can prove convergence in the usual sense. Since (2) has added δa_j to the total amount of cash, we now need to subtract that much cash from the system, but subtract it "equally" from the various states so we don't alter the relative cash balances of the states. But what does that mean? Should we subtract $\frac{1}{N}\delta a_j$ from each state k , or should we subtract $\tilde{p}_k \delta a_j$ from each state k ? Both schemes reduce total cash by δa_j . Which is appropriate? Or is there a third way?

We shall look at both schemes. Our cash subtraction will be governed by what I shall call the *correction vector* $\mathbf{w}$. The entries in $\mathbf{w}$ will sum to 1. From the cash balance of each state k we subtract $w_k \delta a_j$. We shall investigate various choices of $\mathbf{w}$. In each case we shall have $0 \le w_k \le 1$ for all k , and of course $\mathbf{w}\mathbf{e}^\top = 1$. The two schemes mentioned above are $\mathbf{w} = \frac{1}{N}\mathbf{e}$ and $\mathbf{w} = \tilde{\mathbf{p}}$.

Cash Change Description:
In the bucket brigade, there is a pile of cash on each state. When the chain takes transition $i \to j$, there is a change in the cash balances. The change has three components.
(1) a proportion ε of the cash of state j is passed to state i . $0 < \varepsilon < 1$.
(2) the excess payoff amount δa_j is added to the cash balance of state j .
(3) The amount $\delta a_j w_k$ is subtracted from the cash balance of every state k .

(Component (3) is omitted in many implementations, but we shall use it in our convergence proof and discuss it further when we have finished the proof.)

We write $\mathbf{v}^{(n)}$ for the vector of cash balances at time n . Its i'th entry is $v_i^{(n)}$.

We see that the three components of the cash movement add these component vectors to the current cash balance vector.

(1) $\varepsilon v_j^{(n)}(\mathbf{e}_i - \mathbf{e}_j)$
(2) $\delta a_j \mathbf{e}_j$
(3) $-\delta a_j \mathbf{w}$ (optional)

I've marked (3) optional because many implementations do not use it, but we will use it.

[8] in subsection 4.2

So the new cash balance vector is $\mathbf{v}^{(n+1)}$. The change is

$$\mathbf{v}^{(n+1)} - \mathbf{v}^{(n)} = \varepsilon v_j^{(n)}(\mathbf{e}_i - \mathbf{e}_j) + \delta a_j \mathbf{e}_j - \delta a_j \mathbf{w} \quad . \tag{14}$$

If (3) is not used then the last term on the right is not present.

As the bucket brigade runs, the cash balance vector does a random walk in vector space. To get an idea where it goes, we define what I call the *deterministic bucket brigade*, which is a simplified version of the real bucket brigade. I call the real bucket brigade the *probabilistic bucket brigade.*

Consider the change in the cash balance vector in one time unit. In the probabilistic bucket brigade, the change is the right hand side of (14). The change depends on the state transition $i \to j$.

In the deterministic bucket brigade, the change is the average of the possible changes that the probabilistic bucket brigade might make in a time unit, averaged over all state transitions. That is, the change of the cash balance vector in the deterministic bucket brigade is the right hand side of (14) averaged over all transitions $i \to j$.

$$\sum_{ij} F_{ij} \,[\, \varepsilon\, v_j^{(n)}(\mathbf{e}_i - \mathbf{e}_j) + \delta a_j \mathbf{e}_j - \delta a_j \mathbf{w}\,] \tag{15}$$

This can be written

$$\sum_{ij} F_{ij}\, \varepsilon\, v_j^{(n)}(\mathbf{e}_i - \mathbf{e}_j) + \sum_{ij} F_{ij}\, \delta a_j \mathbf{e}_j - \delta \sum_{ij} F_{ij}\, a_j \mathbf{w} \quad . \tag{16}$$

By (10) we have $\sum_i F_{ij} = \tilde{p}_j$, so we have $\sum_{ij} F_{ij} a_j = \sum_j \tilde{p}_j a_j = \tilde{\mathbf{p}}\, \mathbf{a}^\top = 0$, and we see that the third term in the expression for change is zero. That term is there because of the third (optional) component in the change. As we said, many implementations don't use the third component. In that case the third term wouldn't be there. But the third term is zero anyway, so whether or not the implementation uses the third component, the change per time step of the cash balance vector in the deterministic bucket brigade is simply this vector.

$\sum_{ij} F_{ij}\, \varepsilon\, v_j^{(n)}(\mathbf{e}_i - \mathbf{e}_j) + \sum_{ij} F_{ij}\, \delta a_j \mathbf{e}_j$

Writing the vector as a column vector, this becomes

$\sum_{ij} F_{ij}\, \varepsilon\, v_j^{(n)} \mathbf{e}_i^\top - \sum_{ij} F_{ij}\, \varepsilon\, v_j^{(n)} \mathbf{e}_j^\top + \sum_{ij} F_{ij}\, \delta a_j \mathbf{e}_j^\top$.

$$\varepsilon \sum_{ij} F_{ij}\, v_j^{(n)} \mathbf{e}_i^\top + \sum_{ij} F_{ij}\, (-\varepsilon\, v_j^{(n)} + \delta a_j)\, \mathbf{e}_j^\top \tag{17}$$

We can obtain the k'th entry in this column vector by multiplying on the left by $\mathbf{e}_k$.

The first of the two terms becomes

$\mathbf{e}_k(\varepsilon \sum_{ij} F_{ij}\, v_j^{(n)} \mathbf{e}_i^\top) = \varepsilon \sum_j F_{kj} v_j^{(n)} = \sum_j F_{kj} \varepsilon v_j^{(n)}$

Using (10), the second term becomes

$\mathbf{e}_k \sum_{ij} F_{ij}\, (-\varepsilon\, v_j^{(n)} + \delta a_j)\, \mathbf{e}_j^\top = \sum_i F_{ik}\, (-\varepsilon\, v_k^{(n)} + \delta a_k) = \sum_j F_{kj} (-\varepsilon\, v_k^{(n)} + \delta a_k)$

So the k'th entry is the following.

This is the change in the cash balance of state k in one step of the deterministic bucket brigade.

$$\sum_j F_{kj} \left(\varepsilon v_j^{(n)} - \varepsilon\, v_k^{(n)} + \delta a_k \right) \tag{18}$$

We now obtain the k'th entry of the following column vector.

$\varepsilon\, F(\mathbf{v}^{(n)})^\top + D\, (-\varepsilon\, (\mathbf{v}^{(n)})^\top + \delta \mathbf{a}^\top)$.

We multiply on the left by $\mathbf{e}_k$. The first term is this.

$\mathbf{e}_k(\varepsilon\, F(\mathbf{v}^{(n)})^\top) = \varepsilon \sum_j F_{kj} v_j^{(n)} = \sum_j F_{kj} \varepsilon v_j^{(n)}$

Using (10), we obtain the second term in the same way.

$\mathbf{e}_k(D\, (-\varepsilon\, (\mathbf{v}^{(n)})^\top + \delta \mathbf{a}^\top)) = \tilde{p}_k(-\varepsilon\, v_k^{(n)} + \delta a_k) = \sum_j F_{kj} (-\varepsilon\, v_k^{(n)} + \delta a_k)$

Adding these two terms gives us (18), so we see that the vector $\varepsilon\, F(\mathbf{v}^{(n)})^\top + D\, (-\varepsilon\, (\mathbf{v}^{(n)})^\top + \delta \mathbf{a}^\top)$ has the same k'th entry as the vector (17). The k is arbitrary, so the two vectors are identical. So the change for one time step of the cash balance vector can be written as formula (17) or as

$\varepsilon\, F(\mathbf{v}^{(n)})^\top + D\, (-\varepsilon\, (\mathbf{v}^{(n)})^\top + \delta \mathbf{a}^\top)$.

Since $F = DP$, this can be written

$(\mathbf{v}^{(n+1)})^\top - (\mathbf{v}^{(n)})^\top = D(\varepsilon P (\mathbf{v}^{(n)})^\top - \varepsilon (\mathbf{v}^{(n)})^\top + \delta \mathbf{a}^\top)$.

If we multiply by $\mathbf{e}$ on the left, the right side becomes 0, so we see that there is no change in the total

amount of cash from time n to time $n+1$. From the last equation we obtain
$\left(\mathbf{v}^{(n+1)}\right)^{\top} = \delta D\mathbf{a}^{\top} + (\varepsilon DP - \varepsilon D + I)\left(\mathbf{v}^{(n)}\right)^{\top}$.
We define

$$Q = \varepsilon DP - \varepsilon D + I \quad ,$$

so we have

$$\left(\mathbf{v}^{(n+1)}\right)^{\top} = \delta D\mathbf{a}^{\top} + Q\left(\mathbf{v}^{(n)}\right)^{\top} \quad . \tag{19}$$

Writing $\mathbf{v}^{(0)}$ as $\mathbf{v}$, we can now prove the following by induction on n .

$$\left(\mathbf{v}^{(n)}\right)^{\top} = \delta\left(\sum_{k=0}^{n-1}(Q^k D\mathbf{a}^{\top})\right) + Q^n\mathbf{v}^{\top} \tag{20}$$

We now examine the matrix Q . The matrix $-\varepsilon D + I$ is a non-negative matrix with positive diagonal entries. Therefore, Q is a non-negative matrix with positive diagonal entries. We have $Q\mathbf{e}^{\top} = \mathbf{e}^{\top}$, so Q is row stochastic. The matix Q has a positive entry everywhere P does, so Q is irreducible. Since Q is row stochastic and strongly connected (irreducible), the Frobenius Facts given in subsection 1.2 tell us that Q has a unique stochastic left eigenvector of eigenvalue 1. Let's call it $\tilde{\mathbf{q}}$. And if we apply lemma 5 to Q , we see that Q is primitive.

Now $\mathbf{e}Q = \mathbf{e}$. Q is bistochastic. Therefore, $\mathbf{e}$ is a left eigenvector with eigenvalue 1, and so by the Frobenius Facts, $\mathbf{e}$ is a scalar multiple of $\tilde{\mathbf{q}}$. Clearly the scalar is N, so
$\tilde{\mathbf{q}} = \frac{1}{N}\mathbf{e}$.

Now by analogy with $\hat{P}$, we define
$\hat{Q} = I - Q + \mathbf{e}^{\top}\tilde{\mathbf{q}}$.
By analogy with $\hat{P}$, equation (5) holds with Q in place of P , with $\tilde{\mathbf{q}}$ in place of $\tilde{\mathbf{p}}$, and with $\hat{Q}$ in place of $\hat{P}$. (Sum and inverse both existing) Furthermore, by lemma 4, since Q is primitive the Cesaro sum can be replaced by an ordinary infinite sum. So we have
$\sum_{n=0}^{\infty}(Q^n - \mathbf{e}^{\top}\tilde{\mathbf{q}}) = \hat{Q}^{-1} - \mathbf{e}^{\top}\tilde{\mathbf{q}}$.
Multiplying on the right by $D\mathbf{a}^{\top}$ and noting $\tilde{\mathbf{q}}D\mathbf{a}^{\top} = \frac{1}{N}\mathbf{e}D\mathbf{a}^{\top} = \frac{1}{N}\tilde{\mathbf{p}}\,\mathbf{a}^{\top} = 0$, we obtain
$\sum_{n=0}^{\infty}(Q^n D\mathbf{a}^{\top}) = \hat{Q}^{-1}D\mathbf{a}^{\top}$.
Lemma 4 with Q in place of P and with $\tilde{\mathbf{q}}$ in place of $\tilde{\mathbf{p}}$ also tells us that
$\lim_{n\to\infty} Q^n = \mathbf{e}^{\top}\tilde{\mathbf{q}}$.
Letting $n \to \infty$ in equation (20) and using the previous two equations gives us
$\lim_{n\to\infty}\left(\mathbf{v}^{(n)}\right)^{\top} = \delta\,\hat{Q}^{-1}D\mathbf{a}^{\top} + \mathbf{e}^{\top}\tilde{\mathbf{q}}\mathbf{v}^{\top}$.

We define
$\mathbf{u} = \lim_{n\to\infty}\mathbf{v}^{(n)}$,
so the i'th entry u_i is the limiting cash balance of state i . Using this and $\tilde{\mathbf{q}} = \frac{1}{N}\mathbf{e}$, we have

$$\mathbf{u}^{\top} = \delta\,\hat{Q}^{-1}D\mathbf{a}^{\top} + \tfrac{1}{N}\mathbf{e}^{\top}\mathbf{e}\mathbf{v}^{\top} \quad . \tag{21}$$

The limiting cash balances we call *equilibrium cash balances* because at equilibrium the cash coming into a state equals the cash going out on average. We define
$\bar{u} = \tilde{\mathbf{p}}\mathbf{u}^{\top}$.
We call $\bar{u}$ the average equilibrium cash balance.

We now examine the matrix $\hat{Q}$.
We first note that since $\mathbf{e}\hat{Q} = \mathbf{e}$, we have $\mathbf{e}\hat{Q}^{-1} = \mathbf{e}$, and consequently,
$\mathbf{e}\hat{Q}^{-1}D\mathbf{a}^{\top} = 0$. We shall use this in what follows.
We first multiply (21) on the left by $\mathbf{e}$ and obtain

$$\mathbf{e}\,\mathbf{u}^{\top} = \mathbf{e}\,\mathbf{v}^{\top} \quad . \tag{22}$$

From the definition of Q , we have
$Q = \varepsilon DP - \varepsilon D + I = -\varepsilon D(I-P) + I$. We substitute this into
$\hat{Q} = I - Q + \frac{1}{N}\mathbf{e}^{\top}\mathbf{e}$, obtaining
$\hat{Q} = \varepsilon D(I-P) + \frac{1}{N}\mathbf{e}^{\top}\mathbf{e}$. Multiplying on the right by $\hat{Q}^{-1}D\mathbf{a}^{\top}$ gives
$D\mathbf{a}^{\top} = \varepsilon D(I-P)\hat{Q}^{-1}D\mathbf{a}^{\top}$. Then multiplying on the left by $\frac{1}{\varepsilon}D^{-1}$ gives
$\frac{1}{\varepsilon}\mathbf{a}^{\top} = (I-P)\hat{Q}^{-1}D\mathbf{a}^{\top}$.

We now multiply equation (21) on the left by $(I-P)$ and use the last equation and $(I-P)\mathbf{e}^\top = 0$.
$(I-P)\mathbf{u}^\top = \frac{\delta}{\varepsilon}\mathbf{a}^\top$

We define
$\mathbf{w} = \mathbf{u} - \frac{\delta}{\varepsilon}\mathbf{c}$.
Then using the previous equation and equation (8), we have
$(I-P)\mathbf{w}^\top = 0$.
So $P\mathbf{w}^\top = \mathbf{w}^\top$, and $\mathbf{w}^\top$ is a right eigenvector of P with eigenvalue 1. So is $\mathbf{e}^\top$. Since the set of right eigenvectors with eigenvalue 1 forms a one dimensional subspace, $\mathbf{w}^\top$ is a scalar multiple of $\mathbf{e}^\top$.
So there is a scalar χ such that $\mathbf{w}^\top = \chi\mathbf{e}^\top$. In other words
$\mathbf{u}^\top - \frac{\delta}{\varepsilon}\mathbf{c}^\top = \chi\mathbf{e}^\top$.
Multiplying on the left by $\tilde{\mathbf{p}}$ (and using $\tilde{\mathbf{p}}\mathbf{c}^\top = 0$) gives $\bar{u} = \chi$.
Then solving the previous equation for $\delta\mathbf{c}^\top$ gives

$$\delta\,\mathbf{c}^\top = \varepsilon(\mathbf{u}^\top - \bar{u}\mathbf{e}^\top) \quad . \tag{23}$$

In other words, for each i we have

$$\delta\, c_i = \varepsilon(u_i - \bar{u}) \quad . \tag{24}$$

Remember we said that in most implementations, $\delta = 1$.

We see by (23) that $\mathbf{u}$ determines $\mathbf{c}$. It's also true that if we know the total amount of cash, then $\mathbf{c}$ determines $\mathbf{u}$. During bucket brigade operation, the total amount of cash remains constant. Let's write τ for the total amount of cash. Then $\mathbf{e}\mathbf{u}^\top = \tau$. For convenience, let's define the matrix
$M = \frac{1}{N}\mathbf{e}^\top\mathbf{e}$.
Let's see how $\mathbf{c}$ determines $\mathbf{u}$.
Equation (23) can be written

$$\varepsilon\bar{u}\mathbf{e}^\top \;=\; \varepsilon\mathbf{u}^\top \;-\; \delta\mathbf{c}^\top \quad . \tag{25}$$

Now $M\mathbf{e}^\top = \mathbf{e}^\top$ and $NM\mathbf{u}^\top = \mathbf{e}^\top\mathbf{e}\,\mathbf{u}^\top = \tau\mathbf{e}^\top$, so if we multiply (25) on the left by NM we obtain
$\varepsilon N\bar{u}\mathbf{e}^\top = \varepsilon\tau\mathbf{e}^\top - \delta NM\mathbf{c}^\top$.
Multiplying (25) by N gives us
$\varepsilon N\bar{u}\mathbf{e}^\top = \varepsilon N\mathbf{u}^\top - \delta N\mathbf{c}^\top$.
We subtract this from the previous equation.
$0 = \varepsilon\tau\mathbf{e}^\top - \varepsilon N\mathbf{u}^\top + \delta N(I-M)\mathbf{c}^\top$

$$\mathbf{u}^\top \;=\; \tfrac{\delta}{\varepsilon}\,(I-M)\,\mathbf{c}^\top + \;\tfrac{\tau}{N}\,\mathbf{e}^\top \tag{26}$$

Let's take a quick look at the rate of convergence to the equilibrium cash balances.
Remember that the definition of $\hat{Q}$ is
$\hat{Q} = I - Q + \frac{1}{N}\mathbf{e}^\top\mathbf{e}$. Transposing gives us
$I = \hat{Q} + Q - \frac{1}{N}\mathbf{e}^\top\mathbf{e}$.
Multiplying on the right by $\delta\,\hat{Q}^{-1}D\mathbf{a}^\top$ and remembering that $\mathbf{e}\,\hat{Q}^{-1}D\mathbf{a}^\top = 0$ yields
$\delta\,\hat{Q}^{-1}D\mathbf{a}^\top = \delta D\mathbf{a}^\top + Q(\delta\,\hat{Q}^{-1}D\mathbf{a}^\top)$.
Since Q is bistochastic, we have
$\frac{1}{N}\mathbf{e}^\top\mathbf{e}\mathbf{v}^\top = Q(\frac{1}{N}\mathbf{e}^\top\mathbf{e}\mathbf{v}^\top)$.
Adding the last two equations and using equation (21) gives us
$\mathbf{u}^\top = \delta D\mathbf{a}^\top + Q\mathbf{u}^\top$.
If we subtract this equation from equation (19), we obtain
$(\mathbf{v}^{(n+1)})^\top - \mathbf{u}^\top = Q((\mathbf{v}^{(n)})^\top - \mathbf{u}^\top)$.
Thus by induction, we have

$$(\mathbf{v}^{(n)})^\top - \mathbf{u}^\top \;=\; Q^n(\mathbf{v}^\top - \mathbf{u}^\top) \quad . \tag{27}$$

This last equation gives us a sort of convergence rate.

Suppose we start with two different cash balance vectors, $\mathbf{v}$ and $\hat{\mathbf{v}}$, and watch them change with time. After n steps, the vectors are $\mathbf{v}^{(n)}$ and $\hat{\mathbf{v}}^{(n)}$. On the next step, the change of $\mathbf{v}^{(n)}$ is given by equation (19). The corresponding change of $\hat{\mathbf{v}}^{(n)}$ is of course,
$(\hat{\mathbf{v}}^{(n+1)})^\top = \delta D\mathbf{a}^\top + Q(\hat{\mathbf{v}}^{(n)})^\top$.
If we subtract this equation from (19), we obtain

$$(\mathbf{v}^{(n+1)})^\top - (\hat{\mathbf{v}}^{(n+1)})^\top = Q\,\left((\mathbf{v}^{(n)})^\top - (\hat{\mathbf{v}}^{(n)})^\top\right) \quad . \tag{28}$$

In this way, Q governs the difference of the two vectors.

We now look at the difference $\|(\mathbf{v}^{(n)})^\top - (\hat{\mathbf{v}}^{(n)})^\top\|$ as n increases. We know it goes to zero as $n \to \infty$, since both $\mathbf{v}^{(n)}$ and $\hat{\mathbf{v}}^{(n)}$ go to $\mathbf{u}$. That is, the two vectors ultimately converge.

But we know more than that. At no stage do they diverge, even temporarily. We can see this if we apply inequality (11) to equation (28). We obtain
$\left\|(\mathbf{v}^{(n+1)})^\top - (\hat{\mathbf{v}}^{(n+1)})^\top\right\| \;\leq\; \|Q\| \; \left\|(\mathbf{v}^{(n)})^\top - (\hat{\mathbf{v}}^{(n)})^\top\right\|$
Since Q is bistochastic, we have $\|Q\| = 1$, so the result is

$$\left\|(\mathbf{v}^{(n+1)})^\top - (\hat{\mathbf{v}}^{(n+1)})^\top\right\| \;\leq\; \left\|(\mathbf{v}^{(n)})^\top - (\hat{\mathbf{v}}^{(n)})^\top\right\| \quad . \tag{29}$$

On no step do the two vectors diverge. So by induction we see the following.

$$\left\|(\mathbf{v}^{(n)})^\top - (\hat{\mathbf{v}}^{(n)})^\top\right\| \;\leq\; \left\|\mathbf{v}^\top - \hat{\mathbf{v}}^\top\right\| \quad . \tag{30}$$

If we write the vectors as row vectors, that becomes

$$\left\|\mathbf{v}^{(n)} - \hat{\mathbf{v}}^{(n)}\right\| \;\leq\; \|\mathbf{v} - \hat{\mathbf{v}}\| \quad . \tag{31}$$

Of course these comments about convergence and divergence are true only in the deterministic bucket brigade. In the probabilistic bucket brigade the situation is very different.

What have we done?

The proof in this section 2 is satisfying. The proof proved convergence of the deterministic bucket brigade to the equilibrium cash balance vector $\mathbf{u}$. And it proved that the values vector $\mathbf{c}$ is a simple function of $\mathbf{u}$. (equation (23)) The proof is short, and although it is not profound, it does have some cute tricks. And the convergence is ordinary convergence.

The bad news is of course that the deterministic bucket brigade is not implementable. The proof in this section 2 is encouraging and suggestive, but it is the probabilistic bucket brigade with which we must deal. We need to prove that the probabilistic bucket brigade converges to the same $\mathbf{u}$ to which the deterministic bucket brigade converges. That is what we now do.

Before we begin the proof, allow me a brief lament. The proof we will give has none of the attractiveness of the proof in this section 2. It is simply a brute force proof, with few cute tricks if any. Even so, it is long and complicated. I hope a simpler proof exists, but if so I have not found it.

But at least the proof we will give is a formal proof of bucket brigade convergence on Markov chains.

We give that proof in section 3.

3 Probabilistic Bucket Brigade Convergence

3.1 Selecting $\hat{\rho}$, $\hat{\ell}$, $\hat{\varepsilon}$, $\hat{n}$, and $\dot{p}$

We now investigate the performance of the probabilistic bucket brigade.

Equation (14) gives the change in the cash balance vector in one step when the proportion of cash passed back is ε .

I call ε the *scaling factor* or the *step size.* We are going to experiment with various values of ε . We will be looking at a sequence of ℓ steps, which we will call an *era.* I call ℓ the *number of steps.* We will experiment with various values of ℓ . We will define $\rho = \varepsilon\ell$. I call ρ the *distance traveled.* Of course it's not really the distance traveled, but the term is suggestive and I shall use it.

In our convergence proof in this section, the payoff scaling factor will equal the step size. This will not be relevant until we get to subsection 3.6.

In this subsection we experiment with various values of ρ , ℓ , and ε . So in this subsection, ρ , ℓ , and ε are variables. In later subsections we will determine what values we want them to have. Here, we determine only some bounds on the values we want. We call the bounds $\hat{\rho}$, $\hat{\ell}$, and $\hat{\varepsilon}$.

In the deterministic bucket brigade we used the letter $\mathbf{v}$ for the cash balance vector. In the probabilistic bucket brigade we shall use the letter $\mathbf{x}$. So we write $\mathbf{x}$ for the current cash balance vector and $\mathbf{u}$ for the equilibrium cash balance vector. Our goal is to show that $\mathbf{x}$ is close to $\mathbf{u}$, that the cash balance vector will usually be near $\mathbf{u}$, and it will be nearer if the step size is smaller. We want to show that we can get $\mathbf{x}$ as close to $\mathbf{u}$ as we like by reducing the step size. Let me make this a bit more specific.

We begin with a positive number r and a positive probability $\hat{q}$ strictly between 0 and 1. The numbers r and $\hat{q}$ can be as small as we like. Our goal is to show that if we make the step size small enough, then $\wp[\,\|\mathbf{x}-\mathbf{u}\| > r\,] \leq \hat{q}$.

(You can see that I'm writing $\wp[\textit{ statement }]$ to mean the probability that the *statement* is true.)

We define:

$$\theta = \frac{\|\mathbf{u}\| + \|\mathbf{a}\|}{2}$$

$$d = \tfrac{\theta}{r} + 1$$

$$\Delta = \min\left(\tfrac{1}{4}\,e^{-d},\ \ \tfrac{r}{4(r+\theta)}\right)$$

Note that Δ is a function of r and that the smaller r is, the smaller Δ is.
$\lim_{r\to 0}\Delta = 0$.
Since r is an arbitrarily small positive real number, so is Δ .

We define

$$M = \tfrac{1}{N}\,\mathbf{e}^\top\mathbf{e} \quad .$$

Remember that $A = D(I-P)$. (definition (12))

The ρ Condition:
We choose a positive real number $\hat{\rho}$ so that for any $\rho \geq \hat{\rho}$ we have
$\|e^{-\rho A} - M\| < \Delta$.

Equation (13) says we can do this. We also ensure that

$$\hat{\rho} \geq 1 \quad . \tag{32}$$

So we have chosen $\hat{\rho}$. We now are ready to choose $\hat{\ell}$.
We use lemma 34 for the case when x is $4\hat{\rho}$.
Then the lemma reads as follows.

Lemma 6 *Let* B *be any complex matrix for which* $\|B\| \leq 1$.
Let n *be an integer variable. As* $n\to\infty$,
the quantity $\|e^{tB} - (I+\frac{t}{n}B)^n\|$ *goes to zero uniformly for all* t *in the interval* $[\,0,\ 4\hat{\rho}\,]$.

I'm going to re-write that lemma but instead of using the variable t , I'm going to use the variable ρ , which is half t . That is, $t = 2\rho$. We replace t with 2ρ in the lemma. Of course
$t \in [\,0,\ 4\hat{\rho}\,] \implies 2\rho \in [\,0,\ 4\hat{\rho}\,] \implies \rho \in [\,0,\ 2\hat{\rho}\,]$,
so the quantity will go to zero uniformly for all ρ in the interval $[\,0,\ 2\hat{\rho}\,]$.
The lemma now looks like this.

Lemma 7 *Let* B *be any complex matrix for which* $\|B\| \leq 1$. *As* $n\to\infty$,
the quantity $\|e^{2\rho B} - (I+\frac{2\rho}{n}B)^n\|$ *goes to zero uniformly for all* ρ *in the interval* $[\,0,\ 2\hat{\rho}\,]$.

Now $\|D\| \leq 1$ and $\|-DP\| = \|DP\| \leq 1$. Therefore, by our definition (12) of A we have
$\|A\| = \|D(I-P)\| = \|D-DP\| \leq \|D\| + \|-DP\| \leq 2$.
$\|-\frac{1}{2}A\| = \frac{1}{2}\|A\| \leq 1$.
So we can use our lemma with $B = -\frac{1}{2}A$.
It now reads as follows.

Lemma 8 *As* $n\to\infty$,
the quantity $\|e^{-\rho A} - (I-\frac{\rho}{n}A)^n\|$ *goes to zero uniformly for all* ρ *in the interval* $[\,0,\ 2\hat{\rho}\,]$.

Since we have this uniform convergence, we can choose $\hat{\ell}$ as follows.

The ℓ Conditions:
We choose a positive integer $\hat{\ell}$ greater than $\hat{\rho}$ that is large enough that we have:
First Condition

$$\hat{\ell} \geq \left(2\hat{\rho}\,e^{\hat{\rho}}\right)^2\Delta^{-1} \quad ,$$

Second Condition:
For all integers n such that $n \geq \hat{\ell}$ and for all ρ in the interval $[0,\,2\hat{\rho}]$, we have

$$\left\|\left(I - \tfrac{1}{n}\,\rho A\right)^n - e^{-\rho A}\right\| < \Delta \quad .$$

So we have now chosen $\hat{\rho}$ and $\hat{\ell}$.

We define
$\hat{\varepsilon} \ = \ \hat{\rho}\,\hat{\ell}^{-1}$.
Since $\hat{\ell} \ > \ \hat{\rho}$, we have

$$\hat{\varepsilon} \ < \ 1 \quad . \tag{33}$$

We also have
$\hat{\varepsilon}\,\hat{\ell} \ = \ \hat{\rho}$.

We define $\hat{n}$ to be the smallest integer greater than

$$2\hat{\rho} + 1 + 2\tfrac{\theta}{r} \quad . \tag{34}$$

As we said at the start of this section, we begin with a positive number r and a positive probability $\hat{q}$ such that
$0 \ < \ \hat{q} \ < 1$.
We now define $\dot{p}$ to be the positive number that is the $\hat{n}$'th root of $(1+\hat{q})^{-1}$. So we have

$$\dot{p}^{\hat{n}} \ = \ (1+\hat{q})^{-1} \quad . \tag{35}$$

We see that
$\frac{1}{2} \ < \ \dot{p}^{\hat{n}} \ < \ 1$, so we have both
$0 < \dot{p} < 1$ and

$$\tfrac{1}{2} \ < \ \dot{p}^{\hat{n}} \quad . \tag{36}$$

So $\dot{p}$ is a probability strictly between 0 and 1.

3.2 Markov Chain Ergodicity

We will need to use the ergodicity of the Markov chain. That is, we need to know that as the chain runs, the proportion of times in which it visits state i is roughly $\tilde{p}_i$, and that the longer the chain runs the closer to $\tilde{p}_i$ the proportion is liable to be.

Anyone with a decent grounding in probability theory knows how to make these fairly standard ideas precise and prove what we need. Unfortunately I am not such a person and convergence of probabilities makes me uneasy. I tried to prove what we need from scratch, using the concepts, results, and notation of the previous sections. This was horribly long and fiddly, and I'm not at all sure it was correct. I think it can be done, but it was silly to waste so much time on a well known result elegantly documented in the literature.

Our Markov chain is strongly connected, so everybody knows it's ergodic. Norris [3] elegantly proves the ergodicity we need. (His proof does what I did, but he does it elegantly and my proof was a mess. He picks a state i and looks at strings of tours from i to i .)

Norris's proof works for what he calls an irreducible Markov chain. That's simply what we call strongly connected.

I now need to state Norris's result carefully so we can use it.

Norris considers the sequence of random variables
$X_0, X_1, X_2, X_3, ...$
where X_n is the state the chain is in at time n .
The chain has an initial probability distribution $\boldsymbol{\lambda}$ and the transition probability matrix P .
We can think of $\boldsymbol{\lambda}$ as a row vector. For each state j ,
$\wp[\,X_0 = j\,] \ = \ \lambda_j$.
Of course we have
$\wp[\,X_n = j\,] \ = \ \boldsymbol{\lambda} P^n \mathbf{e}_j^\top$.

For each state i , Norris defines the sequence of random variables $Y_0^{(i)}, Y_1^{(i)}, Y_2^{(i)}, Y_3^{(i)}, ...$, where $Y_n^{(i)}$ is the proportion of times up to time n that state i is visited.[9] In other words, we look at the sequence of n states $X_0, X_1, X_2, X_3, ..., X_{n-1}$ and count the number of i's. That number is $nY_n^{(i)}$, and $Y_n^{(i)}$ is the proportion of the n states that are i .

Norris discusses Markov chains with countable state sets, so his matrix P has a countable (possibly infinite) number of rows and columns. It can have a left eigenvector $\boldsymbol{\pi}$ such that $\boldsymbol{\pi} P = \boldsymbol{\pi}$, but if the

[9] Norris uses slightly different notation. He writes $V_i(n)$ instead of $nY_n^{(i)}$.

state set is infinite then $\boldsymbol{\pi}$ has an infinite number of entries. If the entries in $\boldsymbol{\pi}$ are all non-negative and they sum to 1, then we call $\boldsymbol{\pi}$ an invariant probability distribution over the states.

Of course we call the chain irreducible, or strongly connected, if for every ordered pair of states there is a sequence of legal transitions from the first state to the second and in that sequence every transition has positive probability.

The *expected return time* of state i is the expectation of the number of transitions it takes of we start in state i to return to state i . Of course this could be infinite.

Now suppose the chain is irreducible (strongly connected) and suppose there is an invariant distribution $\boldsymbol{\pi}$. Then Norris's theorem 1,7,7 says among other things that for each state i , its expected return time is π_i^{-1} .

Norris's ergodic theorem 1.10.2 says that if the chain is irreducible then with probability 1, the proportion $Y_n^{(i)}$ converges to the reciprocal of the expected return time of i as $n \to \infty$. Of course by theorem 1.7.7 that reciprocal is π_i if invariant distribution $\boldsymbol{\pi}$ exists.

So if the chain is irreducible and an invariant distribution $\boldsymbol{\pi}$ exists, then for every state i we have the convergence.

$$\wp[\; Y_n^{(i)} \longrightarrow \pi_i \quad \text{as} \quad n \to \infty \;] \;=\; 1 \;. \tag{37}$$

This holds as long as the state set is countable, finite or infinite. In section C we will use this result for infinite state sets. Here we use it for finite state sets. Our matrix P is irreducible (strongly connected) and $\tilde{\mathbf{p}}$ is an invariant distribution. So here we have the following.

$$\wp[\; Y_n^{(i)} \longrightarrow \tilde{p}_i \quad \text{as} \quad n \to \infty \;] \;=\; 1 \;. \tag{38}$$

In other words, the probability that $Y_n^{(i)}$ converges to $\tilde{p}_i$ is 1. And this is true for every state i . Furthermore, the statement is true whatever the initial distribution $\boldsymbol{\lambda}$.

What Norris proves is stronger than what we need here. As you see, Norris proves convergence with probability 1. (convergence almost everywhere) We need only convergence in probability. But convergence with probability 1 implies convergence in probability, so we have more than what we need.

I confess my knowledge of the foundations of probability theory is shaky, particularly when measurability is crucial, so I am uneasy about convergence of random variables. I found the development in Loeve [2] comforting, at least the parts I read. His explanation of measure makes sense to me. I was relieved to read on page 84 that a σ-additive set function is continuous in the sense defined there. The proof was roughly what I had guessed. This fact allows one to see that the different definitions of convergence with probability 1 are equivalent and to see how they relate to convergence in probability. The Comparison of Convergences Theorem on page 116 and its proof then make sense. This theorem includes the reason that convergence with probability 1 (almost everywhere) implies convergence in probability. Of course all this is obvious to anyone with a proper grounding in probability theory, but it wasn't obvious to me.

So since we have convergence almost everywhere, we have convergence in probability. Loeve writes these two types of convergence using similar notation at the top of page 151. Using that notation, we can write the convergence in probability of our random variables as follows.

For any positive real δ , we have
$\wp[\,|Y_n^{(i)} - \tilde{p}_i| \,\geq\, \delta\,] \;\longrightarrow\; 0 \qquad \text{as} \qquad n \to \infty$.

This is true for any state i .
And remember that it is true whatever the initial distribution $\boldsymbol{\lambda}$.

I confess that this approach to our problem looks like using a sledgehammer to crack a nut. All we need is convergence in probability for a strongly connected finite state Markov chain. I feel there should be a simple observation that makes this whole discussion unnecessary, but I don't see it. If you do, please let me know.

We said that $nY_n^{(i)}$ is the number of times the chain visits state i before time n . But actually, we are interested not in visits to states, but in visits to state transitions. We define random variable $Z_n^{(ij)}$ analogous to $Y_n^{(i)}$. The number $nZ_n^{(ij)}$ is the number of traversals of transition $i \to j$ before time n . We want to show that if $i \to j$ is a legal transition, then $Z_n^{(ij)}$ converges to F_{ij} as $n \to \infty$.

No problem. We need only construct a new Markov chain whose states are the legal transitions of the original chain. The transition probabilities of the new chain are defined in the obvious way. Since the original chain is strongly connected, so is the new chain. An initial distribution $\boldsymbol{\Lambda}$ for the new chain is a distribution over the legal transitions of the original chain. The probability Λ_{ij} is the probability that the initial transition is $i \to j$.

Since the new chain is strongly connected, the sequence
$Z_0^{(ij)}, Z_1^{(ij)}, Z_2^{(ij)}, Z_3^{(ij)}, ...$
converges in probability to F_{ij} .
That is:

> For any positive real δ we have
> $\wp[\,|Z_n^{(ij)} - F_{ij}| \ \geq\ \delta\,] \ \longrightarrow\ 0 \qquad \text{as} \qquad n \to \infty$.

This is true for any legal transition $i \to j$.
And it is true whatever the initial distribution $\mathbf{\Lambda}$.

Of course different initial distributions give us different probabilities. The initial distribution we want is this. We select an arbitrary state α of the original chain, and for each state j of the original chain we set $\Lambda_{\alpha j} = P_{\alpha j}$. If $i \neq \alpha$ we set $\Lambda_{ij} = 0$. Now the probabilities $\wp[\,|Z_n^{(ij)} - F_{ij}| \ \geq\ \delta\,]$ are those that obtain in a Markov chain with transition probability matrix P and initial state α . Those probabilities converge to zero as $n \to \infty$.

For any finite legal string of transitions σ , I write $\mathcal{F}_{ij}^{(\sigma)}$ to mean the proportion of those transitions that are $i \to j$. So if σ has length n (consists of n transitions) then $n\mathcal{F}_{ij}^{(\sigma)}$ is the number of transitions in σ that are $i \to j$.

So another way of writing the probability $\wp[\,|Z_n^{(ij)} - F_{ij}| \ \geq\ \delta\,]$ is this. We begin the original chain in state α and let it run to produce a string σ of n transitions. Our probability is
$\wp_\sigma^{\alpha n}[\,|\mathcal{F}_{ij}^{(\sigma)} - F_{ij}| \ \geq\ \delta\,]$,
where the subscript and superscripts on the $\wp_\sigma^{\alpha n}$ indicate that the probability is over all length n legal strings σ that begin with state α . The probability is determined by the transition probability matrix P . The probability of the string σ is simply the product of the probabilities of its successive transitions. The probability of a finite set of length n strings is the sum of the probabilities of the strings in it.

And we have the convergence

$$\wp_\sigma^{\alpha n}[\,|\mathcal{F}_{ij}^{(\sigma)} - F_{ij}| \ \geq\ \delta\,] \ \longrightarrow\ 0 \qquad \text{as} \qquad n \to \infty\ .$$

As n increases, the probability is over sets of increasingly longer strings σ .
We have this convergence for any positive real δ and any initial state α and any legal transition $i \to j$.

Of course what the convergence means is that for any positive probability q there is a positive integer $\mathcal{N}$ such that for any integer n greater than $\mathcal{N}$ we have
$\wp_\sigma^{\alpha n}[\,|\mathcal{F}_{ij}^{(\sigma)} - F_{ij}| \ \geq\ \delta\,] \ <\ q$.
Given δ , q , and α , the $\mathcal{N}$ we select works for our particular transition $i \to j$. There is a different $\mathcal{N}$ for each transition. But of course there are a finite number of possible transitions, so the largest of these $\mathcal{N}$'s works for all transitions. So we have the following.

Lemma 9
Given our δ , q , *and* α , *there is a positive integer* $\mathcal{N}$ *such that for any* n *greater than* $\mathcal{N}$ *we have*
$\wp_\sigma^{\alpha n}[\,|\mathcal{F}_{ij}^{(\sigma)} - F_{ij}| \ \geq\ \delta\,] \ <\ q$
for any legal transition $i \to j$.

In fact, the inequality obviously holds if there is no legal transition from i to j , so the inequality in the lemma holds for any state pair $\langle i, j \rangle$.

For any finite string σ of transitions, we define

$$wobble(\sigma) \ =\ \max_{ij} \left(\left| \mathcal{F}_{ij}^{(\sigma)} - F_{ij} \right| \right) \quad .$$

The maximum is over all legal transitions $i \to j$. Actually, it can be over all state pairs $\langle i, j \rangle$ and we get the same number.

We now prove the following lemma.

Lemma 10
Given any positive real δ , *any positive probability* q , *and any initial state* α ,
there exists a positive integer $\mathcal{N}$ *such that for any integer* n *greater than* $\mathcal{N}$ *we have*

$$\wp_\sigma^{\alpha n}[\,wobble(\sigma) \ \geq\ \delta\,] \ <\ q \quad .$$

We need to be clear what that probability means. The probability is over all legal length n strings σ that start in state α . Each string's probability is the product of the probabilities of its successive transitions. But we are looking only at strings that start at α . Then the inequality holds.

Proof:
We are given δ and q and a state α . Let
$\ddot{q} \; = \; N^{-2}q$.
Select $\mathcal{N}$ so that for any n greater than $\mathcal{N}$ we have
$\wp^{\alpha n}_{\sigma}[\,|\mathcal{F}^{(\sigma)}_{ij} - F_{ij}| \;\geq\; \delta\,] \;<\; \ddot{q}$
for every state pair $\langle i,j\rangle$.
Lemma 9 says we can do this.

Now select any n greater than $\mathcal{N}$. For every state pair $\langle i,j\rangle$,
we define $\mathcal{E}_{ij}$ to be the set of length n strings σ that begin at α and for which $|\mathcal{F}^{(\sigma)}_{ij} - F_{ij}| \;\geq\; \delta$ holds.
(If $i \to j$ is an illegal transition, then we see that $\mathcal{E}_{ij} = \emptyset$.)
Thus our previous inequality can be written
$\wp[\mathcal{E}_{ij}] \;<\; \ddot{q}$.
We define $\bar{\mathcal{E}} \;=\; \bigcup_{ij} \mathcal{E}_{ij}$.
We note that $\sigma \in \bar{\mathcal{E}}$ means that there is an ij such that $|\mathcal{F}^{(\sigma)}_{ij} - F_{ij}| \;\geq\; \delta$.
In other words, for length n strings σ beginning at α we have the following.
$\sigma \in \bar{\mathcal{E}} \quad \Longleftrightarrow \quad \max_{ij}(|\mathcal{F}^{(\sigma)}_{ij} - F_{ij}|) \;\geq\; \delta$
$\sigma \in \bar{\mathcal{E}} \quad \Longleftrightarrow \quad \mathit{wobble}(\sigma) \;\geq\; \delta$
$\wp^{\alpha n}_{\sigma}[\,\mathit{wobble}(\sigma) \;\geq\; \delta\,] \;=\; \wp[\,\bar{\mathcal{E}}\,] \;\leq\; \sum_{ij}\wp[\mathcal{E}_{ij}] \;<\; \sum_{ij}\ddot{q} \;=\; N^2\ddot{q} \;=\; q$
■

3.3 Probabilities of Excellent Epochs and Eras – Selecting $\mathcal{K}$

In section 3.1 we introduced the notion of an Era. The so called distance traveled in the Era was ρ . This was the step size times the number of steps. That is, $\rho \;=\; \varepsilon\ell$. We can reduce the step size and increase the number of steps and still have the same distance traveled. For example, suppose k is a positive integer. If the step size is $\frac{\varepsilon}{k}$ and the number of steps is $k\ell$, then the distance traveled is the same. $\rho \;=\; \frac{\varepsilon}{k}\,(k\ell)$

We can run the probabilistic bucket brigade with any step size, and we can run the deterministic bucket brigade with any step size. We are going to compare the deterministic bucket brigade and the probabilistic bucket brigade. We are going to run them both for one Era, and the two bucket brigades will both travel the same distance ρ . But the deterministic bucket brigade will do it in ℓ steps with step size ε (big steps), and the probabilistic bucket brigade will do it in $k\ell$ steps with step size $\frac{\varepsilon}{k}$ (small steps).

Of course we haven't yet decided on the size of ρ, ε, ℓ, or k, but that's what we are going to do once we decide.

Oh and by the way, remember that the quantity ρ isn't really the distance traveled. It's something like the distance traveled, but the phrase "distance traveled" is only suggestive. I will continue to use that suggestive phrase.

As the probabilistic bucket brigade runs, we get a sequence of steps. In each step there is a chain state transition and a change in the cash balance vector. So we get a sequence of chain state transitions and a sequence of cash balance vectors.

We conceptually segment the step sequence into Eras, each Era consisting of $k\ell$ steps. Each Era we segment into ℓ Epochs, each Epoch consisting of k steps. We haven't yet decided on the size of the integers ℓ and k ; we will decide that in the next section. In this section we will decide on an integer $\mathcal{K}$ that will turn out to be a sort of lower bound on the size of k .

So in this section, ℓ and k can be regarded as integer variables. Many of the quantities we discuss will be functions of ℓ and k .

In this section we ignore the sequence of cash balance vectors and concentrate on the sequence of chain state transitions. In each Epoch there is a sequence σ of state transitions. The length of σ is k . That is, σ consists of k transitions. We say that the *wobble of the Epoch* is the wobble of σ .

Definitions:
An *Excellent Epoch* is an epoch whose wobble is less than $\frac{1}{2}\,\hat{\varepsilon}\,N^{-2}$.
An *Excellent Era* is an Era every one of whose Epochs is Excellent.

How likely is it to obtain a sequence of n excellent epochs? We define
$g(i,k,n)$

to be the conditional probability that the next n epochs will all be excellent, given that the first of those epochs begins in state i .

Of course $g(i,k,1)$ is the conditional probability that the next epoch will be excellent. And the conditional probability that the next epoch will be non-excellent is
$1-g(i,k,1)$.

Of course these conditional probabilities are all functions of k . That is, they are functions of how long we decide to make the Epochs. The number k is the length of the strings σ whose wobbles we are looking at.

Lemma 10 tells us the following.

> If we are given any probability q $(0<q<1)$ and any starting state i ,
> then there exists a positive integer $\mathcal{K}$ such that
> for any integer k greater than $\mathcal{K}$ we have the following.
> If we write σ for the sequence of the next k transitions then
> $\wp[\,\text{wobble}(\sigma) \;\geq\; \frac{1}{2}\,\hat{\varepsilon}\,N^{-2}\,] \;<\; q$.

(The δ in the lemma is here $\frac{1}{2}\,\hat{\varepsilon}\,N^{-2}$,
and the $\mathcal{N}$, n , and α in the lemma are here $\mathcal{K}$, k , and i .)

Of course $\text{wobble}(\sigma) \;\geq\; \frac{1}{2}\,\hat{\varepsilon}\,N^{-2}$ simply means the epoch is non-excellent, and so
$\wp[\,\text{wobble}(\sigma) \;\geq\; \frac{1}{2}\,\hat{\varepsilon}\,N^{-2}\,] \;=\; 1-g(i,k,1)$.
So what we have just said is really this.

Lemma 11
If we are given any probability q $(0<q<1)$ *and any starting state* i ,
then there exists a positive integer $\mathcal{K}$ *such that for any integer* k *greater than* $\mathcal{K}$ *we have*
$1-g(i,k,1) \;<\; q$.

The lemma gives us a different $\mathcal{K}$ for each starting state i , but we can obviously select the largest of these $\mathcal{K}$'s and that will work for all i .

I'm going to define $p \;=\; 1-q$, so I can write that lemma as follows.

Lemma 12
If we are given any probability p $(0<p<1)$ *then there exists a positive integer* $\mathcal{K}$ *such that for any integer* k *greater than* $\mathcal{K}$ *and any starting state* i *we have*
$g(i,k,1) \;\geq\; p$.

We define
$\hat{g}(k,n) \;=\; \min_i g(i,k,n)$.
Then the lemma can be written as follows.

Lemma 13
If we are given any probability p $(0<p<1)$ *then there exists a positive integer* $\mathcal{K}$ *such that for any integer* k *greater than* $\mathcal{K}$ *we have*
$\hat{g}(k,1) \;\geq\; p$.

We now ask what the probability is that the next n epochs are all excellent. That's $g(i,k,n)$.

Lemma 14
If we are given any probability p $(0<p<1)$ *and any* k *for which*
$\hat{g}(k,1) \;\geq\; p$,
then for every positive integer n *we have*
$\hat{g}(k,n) \;\geq\; p^n$.

Proof:
We select a k for which $\hat{g}(k,1) \;\geq\; p$ and we prove $\hat{g}(k,n) \;\geq\; p^n$ by induction on n .
The $n=1$ case is a tautology.
We assume $\hat{g}(k,n) \;\geq\; p^n$ and we prove $\hat{g}(k,n+1) \;\geq\; p^{n+1}$.

We take any starting state i and we look at all the sequences of $n+1$ epochs that begin in chain state i . We suppose the chain is in state i and we ask what the probability is of the next $n+1$ epochs all being excellent. That's $g(i,k,n+1)$.

But suppose we know what the first epoch is going to be. Suppose we know that in the first epoch the chain will go through the sequence τ of k transitions. Suppose that that sequence τ would leave the chain in state j . Now what is the probability that all $n+1$ epochs are excellent?
If τ is excellent, then it's simply the probability that the following n epochs are all excellent. That is, if τ is excellent then the probability is $g(j,k,n)$.
And of course if τ is non-excellent then the probability is zero.

So to recap. If the first epoch sequence is τ , taking the chain from state i to state j , then the probability that all $n+1$ epochs are excellent is:
$g(j,k,n)$ if τ is excellent,
0 if τ is non- excellent.

By the induction assumption,
$g(j,k,n) \ \geq \ \hat{g}(k,n) \ \geq \ p^n$.
Therefore, the probability that all the epochs are excellent is:
greater than or equal to p^n if τ is excellent,
0 if τ is non- excellent.

That probability is conditional on the sequence of the first epoch being τ . But it might be any legal sequence starting at i . To get the unconditional probability we need to average over all possible τ's . Some of these τ's are excellent and some are non-excellent. The proportion that are excellent is $g(i,k,1)$.
So we see that the unconditional probability that all $n+1$ epochs are excellent is greater than or equal to $g(i,k,1)\,p^n$. In other words,
$g(i,k,n+1) \ \geq \ g(i,k,1)\,p^n$.
Since $g(i,k,1) \ \geq \ \hat{g}(k,1) \ \geq \ p$, we have
$g(i,k,n+1) \ \geq \ p^{n+1}$.
This holds for any start state i .
Let's choose the i for which the left side is minimum.
$\hat{g}(k,n+1) \ \geq \ p^{n+1}$.
■

So if we have a k for which $\hat{g}(k,1) \ \geq \ p$, then it is also true that $\hat{g}(k,n) \ \geq \ p^n$ for all positive integers n . So we can re-write lemma 13 as follows.

Lemma 15
If we are given any probability p $(0 < p < 1)$ *and any integer* n ,
then there exists a positive integer $\mathcal{K}$ *such that for any integer* k *greater than* $\mathcal{K}$ *we have*
$\hat{g}(k,n) \ \geq \ p^n$.

The exponent in the last line is annoying. We can get rid of it.

Lemma 16
If we are given any probability p $(0 < p < 1)$ *and any integer* n ,
then there exists a positive integer $\mathcal{K}$ *such that for any integer* k *greater than* $\mathcal{K}$ *we have*
$\hat{g}(k,n) \ \geq \ p$.

This is true because if we are given a p strictly between 0 and 1 then its n'th root is also strictly between 0 and 1 and we can use that root in the previous lemma and get the conclusion we want.

We now are ready to select the positive integer $\mathcal{K}$ that we will use.
We have already selected integer $\hat{\ell}$, and we have defined the probability $\dot{p}$.

We now select a positive integer $\mathcal{K}$ such that for any integer k greater than $\mathcal{K}$ we have

$$\hat{g}(k,\ 2\,\hat{\ell}+1) \ \geq \ \dot{p} \qquad . \tag{39}$$

Lemma 16 says we can do this.

Suppose we have two positive integers, n_1 and n_2 . From the definition of g it is obvious that
$n_1 \leq n_2 \quad \Longrightarrow \quad g(i,k,n_1) \ \geq \ g(i,k,n_2)$,
and so from the definition of $\hat{g}$ we have

$$n_1 \ \leq \ n_2 \quad \Longrightarrow \quad \hat{g}(k,n_1) \ \geq \ \hat{g}(k,n_2) \quad . \tag{40}$$

3.4 Selecting k, ε, ℓ, and ρ

We are now going to determine k and our step size, which I shall temporarily call $\bar{\varepsilon}$.
It must be small enough.

The step size can be *any* positive real number less than or equal to
$(\mathcal{K}+1)^{-1}\hat{\varepsilon}$.

So we have

$$\bar{\varepsilon} \leq (\mathcal{K}+1)^{-1}\hat{\varepsilon} \quad . \tag{41}$$

Since $\mathcal{K}$ is a positive integer, $\mathcal{K}+1 \geq 2$ and so $\bar{\varepsilon} \leq \frac{1}{2}\hat{\varepsilon}$.
Adding $\frac{1}{2}\hat{\varepsilon} - \bar{\varepsilon}$ to both sides gives us

$$\tfrac{1}{2}\hat{\varepsilon} \leq \hat{\varepsilon} - \bar{\varepsilon} \tag{42}$$

We define k to be $\bar{\varepsilon}^{-1}\hat{\varepsilon}$ rounded down to the nearest integer.
Thus we have

$$\bar{\varepsilon}^{-1}\hat{\varepsilon} - 1 < k \leq \bar{\varepsilon}^{-1}\hat{\varepsilon} \quad . \tag{43}$$

From (41) we have
$\mathcal{K}+1 \leq \bar{\varepsilon}^{-1}\hat{\varepsilon}$,
and from (43) we have
$\bar{\varepsilon}^{-1}\hat{\varepsilon} - 1 < k$.
From these two inequalities, we have

$$k > \mathcal{K} \quad .$$

Multiplying (43) by $\bar{\varepsilon}$ gives us
$\hat{\varepsilon} - \bar{\varepsilon} < k\bar{\varepsilon} < \hat{\varepsilon}$.
From this and (42) we obtain
$\frac{1}{2}\hat{\varepsilon} < k\bar{\varepsilon} < \hat{\varepsilon}$.

We define
$\varepsilon = k\bar{\varepsilon}$.

We see from the last inequality that

$$\varepsilon \in \left[\tfrac{1}{2}\hat{\varepsilon},\ \hat{\varepsilon}\right] \tag{44}$$

We see that $\bar{\varepsilon} = \frac{\varepsilon}{k}$.
This is our step size for the probabilistic bucket brigade.
We shall refer to the step size as
$\frac{\varepsilon}{k}$.

We now select an ℓ compatible with ε .
The number ℓ will be $\hat{\rho}\,\varepsilon^{-1}$ rounded up to the nearest integer.
And we define $\rho = \varepsilon\ell$.
We now have various consequences.

We have

$$\hat{\rho}\,\varepsilon^{-1} \leq \ell \leq \hat{\rho}\,\varepsilon^{-1} + 1 \quad . \tag{45}$$

Multiplying by ε gives
$\hat{\rho} \leq \rho \leq \hat{\rho} + \varepsilon$.
From (33) and (32) we have $\hat{\varepsilon} < 1 \leq \hat{\rho}$, so $\hat{\rho} + \hat{\varepsilon} \leq 2\hat{\rho}$. Therefore,
$\hat{\rho} \leq \rho \leq 2\hat{\rho}$.

$$\rho \in [\hat{\rho},\ 2\hat{\rho}] \tag{46}$$

From (32) we have

$$\rho \geq 1 \quad . \tag{47}$$

From (45) we have $\hat{\rho}\,\varepsilon^{-1} \leq \ell$.
From the definition of $\hat{\varepsilon}$, we have $\hat{\ell} = \hat{\rho}\,\hat{\varepsilon}^{-1}$.
Using these two facts and $\varepsilon \leq \hat{\varepsilon}$, we obtain
$\hat{\ell} = \hat{\rho}\,\hat{\varepsilon}^{-1} \leq \hat{\rho}\,\varepsilon^{-1} \leq \ell$.
So we have

$$\ell \geq \hat{\ell} \quad . \tag{48}$$

From (44) we have
$\varepsilon \;\geq\; \frac{1}{2}\hat{\varepsilon}$.
$\varepsilon^{-1} \;\leq\; 2\hat{\varepsilon}^{-1}$
$\hat{\rho}\,\varepsilon^{-1} \;\leq\; 2\hat{\rho}\,\hat{\varepsilon}^{-1}$
Since $\hat{\varepsilon}\hat{\ell} \;=\; \hat{\rho}$, the right side is $2\,\hat{\ell}$.
$\hat{\rho}\,\varepsilon^{-1} \;\leq\; 2\,\hat{\ell}$
From (45) we have $\ell \;\leq\; \hat{\rho}\,\varepsilon^{-1} + 1$, so

$$\ell \;\leq\; 2\,\hat{\ell} + 1 \tag{49}$$

Since $\rho \;\leq\; 2\hat{\rho}$, we have
$\rho^2 e^{\rho} \;\leq\; (2\hat{\rho})^2\, e^{(2\hat{\rho})} \;=\; \left(2\hat{\rho}\, e^{\hat{\rho}}\right)^2$.
From the First ℓ Condition we have
$\hat{\ell}\,\Delta \;\geq\; \left(2\hat{\rho}\, e^{\hat{\rho}}\right)^2$.
From these two inequalities and $\ell\,\Delta \;\geq\; \hat{\ell}\,\Delta$, we obtain
$\ell\,\Delta \;\geq\; \rho^2 e^{\rho}$.
Dividing by ℓ gives us

$$\Delta \;\geq\; \varepsilon\rho\, e^{\rho} \quad . \tag{50}$$

This will be useful later.

We selected $\hat{\ell}$ so that it met the Second ℓ Condition, so from (46) and (48) we have

$$\left\| \left(I - \tfrac{1}{\ell}\,\rho A\right)^{\ell} \,-\, e^{-\rho A} \right\| \;<\; \Delta \quad . \tag{51}$$

Since $\rho \;=\; \varepsilon\ell$ we can write this as
$\|\,(I - \varepsilon A)^{\ell} \,-\, e^{-\rho A}\,\| \;<\; \Delta$.

We selected $\hat{\rho}$ so that it met The ρ Condition. Therefore, since $\rho \geq \hat{\rho}$ (by (46)) we have
$\|\, e^{-\rho A} \,-\, M\,\| \;<\; \Delta$.
The last two inequalities give us

$$\|(I - \varepsilon A)^{\ell} - M\| \;<\; 2\Delta \quad . \tag{52}$$

Now let's look at the deterministic bucket brigade with step size ε . We begin with initial cash balance vector $\mathbf{v}$. After n steps, the cash balance vector is $\mathbf{v}^{(n)}$.
By equation (27) and the obvious $Q \;=\; I - \varepsilon A$, we have
$(\mathbf{v}^{(n)})^{\top} - \mathbf{u}^{\top} = \; Q^n(\mathbf{v}^{\top} - \mathbf{u}^{\top}) = (I - \varepsilon A)^n(\mathbf{v}^{\top} - \mathbf{u}^{\top})$.
If $n = \ell$ this becomes
$(\mathbf{v}^{(\ell)})^{\top} - \mathbf{u}^{\top} \;=\; (I - \varepsilon A)^{\ell}(\mathbf{v}^{\top} - \mathbf{u}^{\top})$.
By equation (22) we have
$0 \;=\; -M(\mathbf{v}^{\top} - \mathbf{u}^{\top})$.
Adding the last two equations gives this.
$(\mathbf{v}^{(\ell)})^{\top} - \mathbf{u}^{\top} = ((I - \varepsilon A)^{\ell} - M)(\mathbf{v}^{\top} - \mathbf{u}^{\top})$
Thus we have
$\|(\mathbf{v}^{(\ell)})^{\top} - \mathbf{u}^{\top}\| \;\leq\; \|(I - \varepsilon A)^{\ell} - M\|\; \|\mathbf{v}^{\top} - \mathbf{u}^{\top}\|$,
and hence by inequality (52) we have

$$\|(\mathbf{v}^{(\ell)})^{\top} - \mathbf{u}^{\top}\| \;\leq\; 2\Delta\; \|\mathbf{v}^{\top} - \mathbf{u}^{\top}\| \quad . \tag{53}$$

3.5 Good Epochs and Eras

We now return to the probabilistic bucket brigade.

Since $k \;>\; \mathcal{K}$, inequality (39) holds, that is,
$\hat{g}(k,\, 2\,\hat{\ell} + 1) \;\geq\; \dot{p}$.
From (40) and (49) we have
$\hat{g}(k, \ell) \;\geq\; \hat{g}(k,\, 2\,\hat{\ell} + 1)$.
The last two inequalities give us

$$\hat{g}(k, \ell) \;\geq\; \dot{p} \quad . \tag{54}$$

Definitions:
An *Good Epoch* is an epoch whose wobble is less than $N^{-2}\varepsilon$.
An *Good Era* is an Era every one of whose Epochs is Good.

The definition of Good Epoch is just like the definition of Excellent Epoch except that the wobble limit is $N^{-2}\varepsilon$ instead of $\frac{1}{2}\hat{\varepsilon}N^{-2}$. By (44) we have
$\frac{1}{2}\hat{\varepsilon}N^{-2} \leq N^{-2}\varepsilon$,
so we see that every Excellent Epoch is Good, and consequently, every Excellent Era is Good.

We look at the successive Eras, each of ℓk steps. Suppose we are at the end of an Era.

Definition of p_i :
We write p_i for the conditional probability that the next era will be good given that the current chain state is i .

So suppose we are at the end of an Era and the current state is i . The probability that the next Era will be Good is p_i . Of course the probability that the next Era will be Excellent is $g(i,k,\ell)$. Since every Excellent Era is Good, the probability the next Era is Good is at least as big as the probability it will be Excellent. In other words
$p_i \geq g(i,k,\ell)$.
We have a p_i for each chain state i .
Using inequality (54), we have
$p_i \geq g(i,k,\ell) \geq \hat{g}(k,\ell) \geq \dot{p}$.

So for any state i ,
$$p_i \geq \dot{p} \quad . \tag{55}$$

3.6 A Norm Bound

We look at one step of the probabilistic bucket brigade. Suppose $\mathbf{x}$ is the cash balance vector at the start of the step, and suppose the state transition is $i \to j$.

The change in the cash balance vector in the next time step is the right side of (14) except that we are writing our current vector as $\mathbf{x}$, and the step size is $\frac{\varepsilon}{k}$ rather than ε . So the change in the cash balance vector is
$$\tfrac{\varepsilon}{k}\, x_j^{(n)}(\mathbf{e}_i - \mathbf{e}_j) \;+\; \delta\, a_j \mathbf{e}_j \;-\; \delta\, a_j \mathbf{w} \quad . \tag{56}$$
Multiplying on the right by $\mathbf{e}^\top$ gives 0. This is the change in total cash.

Remember that the third term comes from the optional third component of the Cash Movement Description. We said that many implementations omit this component. In that case, the change in total cash is δa_i . Then the average change in cash per time step is $\sum_i \tilde{p}_i \delta a_i = 0$. So we see that the total amount of cash does an unbiased random walk. This adds a complication to our convergence discussion. We discuss this complication in subsection 4.3.

We shall assume that our bucket brigade includes the third component of the Cash Movement Description, and consequently the total amount of cash remains constant. This simplifies matters so that we can concentrate on the difficult issues.

The quantity δ is a payoff scaling factor. It scales the payoff used in the bucket brigade.

In our formal argument we will always set the payoff scaling factor equal to the step size. What we are going to do is prove convergence of the probabilistic bucket brigade when the payoff scaling factor δ equals the step size $\frac{\varepsilon}{k}$. We will do this by comparing that probabilistic bucket brigade with a deterministic bucket brigade in which the payoff scaling factor δ equals the step size ε . (Of course we have already proved deterministic bucket brigade convergence for any step size and any value of δ .)

In subsection 4.1 we discuss what happens if the scaling factor doesn't equal the step size.

Although the size of δ doesn't affect whether the deterministic bucket brigade converges, it does affect what it converges to. It converges to $\mathbf{u}$, but what is $\mathbf{u}$? It is the equilibrium cash balance vector to which the deterministic bucket brigade converges, the deterministic bucket brigade with step size ε . So if we look at equation (23), we see that if $\delta = 1$ then we have $\mathbf{c}^\top = \varepsilon(\mathbf{u}^\top - \bar{u}\mathbf{e}^\top)$, whereas of δ equals the step size ε , then we have $\mathbf{c}^\top = \mathbf{u}^\top - \bar{u}\mathbf{e}^\top$. The value of $\mathbf{u}$ is different. The deterministic bucket brigade we use in our proof will have the payoff scaling factor δ equal to the step size ε .

So in our probabilistic bucket brigade, $\delta = \frac{\varepsilon}{k}$, and formula (56), the change in the cash balance vector in one step, can be written as follows.
$\frac{\varepsilon}{k}\left(x_j^{(n)}(\mathbf{e}_i - \mathbf{e}_j) + a_j(\mathbf{e}_j - \mathbf{w})\right)$

We define the *change vector*.

$$\eta(\mathbf{x}, i, j) \;=\; x_j(\mathbf{e}_i - \mathbf{e}_j) \;+\; a_j(\mathbf{e}_j - \mathbf{w}) \quad . \tag{57}$$

So if $\mathbf{x}$ is the cash balance vector at the start of the step and if $i \to j$ is the state transition, then the new cash balance vector at the end of the step will be
$\mathbf{x} + \frac{\varepsilon}{k}\,\eta(\mathbf{x}, i, j)$.
I call the difference between the old and new vectors the *jump vector*. The jump vector is

$$\tfrac{\varepsilon}{k}\,\eta(\mathbf{x}, i, j) \quad . \tag{58}$$

In the probabilistic bucket brigade the cash balance vector doesn't necessarily approach the equilibrium vector u . Indeed it can wander farther and farther from u . We need some sort of bound on how far it can wander in an Era. So suppose the initial cash balance vector at the start of the Era is $\mathbf{v}$.

We define
$\beta = e^{\rho}(\|\mathbf{v}\| + \|\mathbf{a}\|)$.

For each non-negative integer n less than or equal to $k\ell$, we define the set
$\mathcal{S}_n$ to be the set of all cash balance vectors whose norm is less than or equal to
$(1+\frac{\varepsilon}{k})^{-n}\beta \;-\; \|\mathbf{a}\|$.

We note that

$$0 \;\le\; n \;\le\; k\ell \quad\Longrightarrow\quad \mathcal{S}_n \;\subseteq\; \mathcal{S}_0 \quad . \tag{59}$$

Using inequality (94), we have the following;
$(1+\frac{\varepsilon}{k})^{k\ell} \;=\; (1+\frac{\rho}{k\ell})^{k\ell} \;<\; e^{\rho}$
$(1+\frac{\varepsilon}{k})^{k\ell}(\|\mathbf{v}\| + \|\mathbf{a}\|) \;<\; e^{\rho}(\|\mathbf{v}\| + \|\mathbf{a}\|) \;=\; \beta$
$\|\mathbf{v}\| + \|\mathbf{a}\| \;<\; (1+\frac{\varepsilon}{k})^{-k\ell}\beta$
$\|\mathbf{v}\| \;<\; (1+\frac{\varepsilon}{k})^{-k\ell}\beta \;-\; \|\mathbf{a}\|$

$$\mathbf{v} \in \mathcal{S}_{k\ell} \tag{60}$$

So we see that $\mathbf{v} \in \mathcal{S}_{k\ell}$, and indeed if $0 \le n \le k\ell$ then $\mathbf{v} \in \mathcal{S}_n$.

Now here is the point of the $\mathcal{S}_n$ definition.

Lemma 17 *If* $0 < n \le k\ell$ *and the current cash balance vector* $\mathbf{x}$ *is in* $\mathcal{S}_n$ *, then if we run the probabilistic bucket brigade for one time step, the next cash balance vector will be a member of* $\mathcal{S}_{n-1}$ *.*

Proof:
Let the current cash balance vector be $\mathbf{x}$ and let the next vector be $\mathbf{y}$. Assume $\mathbf{x} \in \mathcal{S}_n$.
Suppose the current state transition is $i \to j$. Then the change vector at $\mathbf{x}$ is this.
$\eta(\mathbf{x}, i, j) \;=\; x_j(\mathbf{e}_i - \mathbf{e}_j) \;+\; a_j(\mathbf{e}_j - \mathbf{w})$
Since $0 \le w_k \le 1$ for all k , we have $\|\mathbf{e}_j - \mathbf{w}\| \;\le\; 1$. Therefore we have this.
$\|\eta(\mathbf{x}, i, j)\| \;\le\; |x_j| \,+\, |a_j| \;\le\; \|\mathbf{x}\| + \|\mathbf{a}\|$
Note that the right hand side is independent of the transition $i \to j$.
$\mathbf{y} \;=\; \mathbf{x} + \frac{\varepsilon}{k}\,\eta(\mathbf{x}, i, j)$
$\|\mathbf{y}\| \;\le\; \|\mathbf{x}\| \,+\, \frac{\varepsilon}{k}\|\eta(\mathbf{x}, i, j)\| \;\le\; \|\mathbf{x}\| \,+\, \frac{\varepsilon}{k}(\|\mathbf{x}\| + \|\mathbf{a}\|) \;=\; (1 \,+\, \frac{\varepsilon}{k})\|\mathbf{x}\| \,+\, \frac{\varepsilon}{k}\|\mathbf{a}\|$
Since $\mathbf{x} \in \mathcal{S}_n$ we have this.
$\|\mathbf{x}\| \;\le\; (1+\frac{\varepsilon}{k})^{-n}\beta - \|\mathbf{a}\|$
From the last two inequalities we have the following.

$$\begin{aligned}
\|\mathbf{y}\| \;&\le\; (1+\tfrac{\varepsilon}{k})\|\mathbf{x}\| + \tfrac{\varepsilon}{k}\|\mathbf{a}\| \\
&\le\; (1+\tfrac{\varepsilon}{k})[\,(1+\tfrac{\varepsilon}{k})^{-n}\beta - \|\mathbf{a}\|\,] + \tfrac{\varepsilon}{k}\|\mathbf{a}\| \\
&=\; (1+\tfrac{\varepsilon}{k})^{1-n}\beta - (1+\tfrac{\varepsilon}{k})\|\mathbf{a}\| + \tfrac{\varepsilon}{k}\|\mathbf{a}\| \\
&=\; (1+\tfrac{\varepsilon}{k})^{-(n-1)}\beta \;-\; \|\mathbf{a}\|
\end{aligned}$$

Therefore,
$\mathbf{y} \in \mathcal{S}_{n-1}$
■

Lemma 18 Norm Bound Fact:
During the entire Era, every cash balance vector has norm less than or equal to $\beta \;-\; \|\mathbf{a}\|$ *.*

Proof:
Suppose we start the Era at vector $\mathbf{v}$ and take n steps, where $0 \le n \le k\ell$. Suppose that takes us to vector $\mathbf{x}$. Now (60) tells us that $\mathbf{v} \in \mathcal{S}_{k\ell}$. Lemma 17 used n times tells us that $\mathbf{x} \in \mathcal{S}_{k\ell-n}$. From (59) we have $\mathcal{S}_{k\ell-n} \subseteq \mathcal{S}_0$, so $\mathbf{x} \in \mathcal{S}_0$. So $\|\mathbf{x}\| \le \beta - \|\mathbf{a}\|$.
■

This all works because the number of steps in the Era is $k\ell$, the step size is $\frac{\varepsilon}{k}$, their product (distance traveled) is ρ , and the quantity β embodies ρ and the initial vector $\mathbf{v}$ of this particular Era.

For clarity, it's worth noting that the proof of lemma 17 is even stronger than we need. If we take n steps in the bucket brigade, these will be determined by the sequence σ of n state transitions. That sequence σ will be a legitimate transition sequence in two senses. First, each transition will be a legal transition. Second, each transition in the sequence must be from the state that the previous transition moved the chain to. The point is that we can think of a sequence that is illegitimate in one or both senses, just a sequence of n transitions that may be legitimate or illegitimate and may link up or not. Even then, we still have
$\|\eta(\mathbf{x}, i, j)\| \le \|\mathbf{x}\| + \|\mathbf{a}\|$
for each transition, which is what we used in the proof of the lemma. So the lemma still holds, and taking n such illegitimate steps still ends up at a vector in $\mathcal{S}_0$. But all our sequences will be legal in both senses.

3.7 Sanitizing an Epoch

We are going to examine just one Epoch, consisting of k small steps. Suppose at the start of these k steps we have cash balance vector $\hat{\mathbf{x}}$, which I shall call our initial vector, although it's really initial only for these k steps. And suppose at the start of the Epoch, the chain is in state $\hat{\imath}$, which I shall call the initial state.

By the Norm Bound Fact, none of the cash balance vectors in the entire Era will have norm exceeding $\beta - \|\mathbf{a}\|$, so none of the vectors in the Epoch will have norm exceeding $\beta - \|\mathbf{a}\|$.

Now if we take the probabilistic bucket brigade from $\hat{\mathbf{x}}$, it gives us a k-step trajectory through the vector space and a k-step sequence of chain states. We want to know what the final vector is of that k-step trajectory of vectors, the vector at the end of the Epoch.

But of course there are many possible final vectors. Each k-step sequence of chain states has a probability, and that is the probability of the associated k-step trajectory through the vector space. Thus we have a probability distribution over the k-step trajectories and hence a probability distribution over the possible final vectors.

What we want to show is that if we take the probabilistic bucket brigade for k small steps, the final vector we end up at is likely to be close to where we would be if we took just one big step using the deterministic bucket brigade. This is all assuming that when we start we are at initial vector $\hat{\mathbf{x}}$ and the chain is in initial state $\hat{\imath}$. Let's try to make all this precise.

If $\mathbf{x}$ is any vector on our k step trajectory, then we have the following.
Recall that the definition of η is this.
$\eta(\mathbf{x}, i, j) = x_j(\mathbf{e}_i - \mathbf{e}_j) + a_j(\mathbf{e}_j - \mathbf{w})$
Taking infinity norms of row vectors, we have
$\|\eta(\mathbf{x}, i, j)\| \le |x_j| + |a_j| \le \|\mathbf{x}\| + \|\mathbf{a}\|$.
By the Norm Bound Fact we have $\|\mathbf{x}\| \le \beta - \|\mathbf{a}\|$. Therefore

$$\|\eta(\mathbf{x}, i, j)\| \le \beta \quad . \tag{61}$$

We now let $\mathbf{x}^{(n)}$ be the cash balance vector on the n'th small step of the total of k small steps in the Epoch. Let's look at the next small step. Our step size is $\frac{\varepsilon}{k}$, so if the state transition is $i \to j$, we have
$\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} + \frac{\varepsilon}{k}\, \eta(\mathbf{x}^{(n)}, i, j)$.

$$\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)} = \frac{\varepsilon}{k}\, \eta(\mathbf{x}^{(n)}, i, j) \tag{62}$$

$\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\| = \frac{\varepsilon}{k}\, \|\eta(\mathbf{x}^{(n)}, i, j)\| \le \frac{\varepsilon}{k}\beta$
Then by induction on n , we have
$\|\mathbf{x}^{(n)} - \hat{\mathbf{x}}\| \le n\frac{\varepsilon}{k}\beta \le \varepsilon\beta$.
The definition of η gives us the next two equations.
$\eta(\mathbf{x}^{(n)}, i, j) = x_j^{(n)}(\mathbf{e}_i - \mathbf{e}_j) + a_j(\mathbf{e}_j - \mathbf{w})$
$\eta(\hat{\mathbf{x}}, i, j) = \hat{x}_j(\mathbf{e}_i - \mathbf{e}_j) + a_j(\mathbf{e}_j - \mathbf{w})$
Taking their difference gives us

$\eta(\mathbf{x}^{(n)}, i, j) - \eta(\hat{\mathbf{x}}, i, j) \;=\; (x_j^{(n)} - \hat{x}_j)(\mathbf{e}_i - \mathbf{e}_j)$.
$\|\eta(\mathbf{x}^{(n)}, i, j) - \eta(\hat{\mathbf{x}}, i, j)\| \;\leq\; |x_j^{(n)} - \hat{x}_j| \;\leq\; \|\mathbf{x}^{(n)} - \hat{\mathbf{x}}\| \;\leq\; \varepsilon\beta$

$$\|\eta(\mathbf{x}^{(n)}, i, j) - \eta(\hat{\mathbf{x}}, i, j)\| \;\leq\; \varepsilon\beta \tag{63}$$

Now let's look at our trajectory of k moves starting at $\hat{\mathbf{x}}$ and taking us to the final vector. Each move consists of a jump vector. The jump vector for the $(n+1)$'th move is given by equation (62). The initial vector plus the sum of the k jump vectors gives us the final vector.

Note that for each move, the jump vector is calculated using the vector at the start of the move. (The right hand side of (62) uses $\mathbf{x}^{(n)}$.) I am going to *sanitize* the trajectory by *sanitizing* each move. In the sanitized move, the jump vector is calculated using not the vector $(\mathbf{x}^{(n)})$ at the start of the move, but using instead the initial vector $(\hat{\mathbf{x}})$. Thus the correct jump vector is
$\frac{\varepsilon}{k}\,\eta(\mathbf{x}^{(n)}, i, j)$,
whereas the sanitized jump vector is
$\frac{\varepsilon}{k}\,\eta(\hat{\mathbf{x}}, i, j)$.

We can get a nice bound on the change that sanitization causes in the jump vector. We see by equation (63) that it is
$\frac{\varepsilon}{k}\,\|\eta(\mathbf{x}^{(n)}, i, j) - \eta(\hat{\mathbf{x}}, i, j)\| \;\leq\; \frac{1}{k}\varepsilon^2\beta$.
Since there are k jump vectors, the change that sanitization causes in the sum of the jump vectors is less than or equal to $\varepsilon^2\beta$.

So we see that although sanitization moves the final vector, the distance the final vector is moved is less than or equal to $\varepsilon^2\beta$. In other words, if we take the final vector of the original trajectory and subtract the final vector of the sanitized trajectory, the norm of that difference vector is less than or equal to $\varepsilon^2\beta$.

Let's call the final vector of the original trajectory the *final vector* $\hat{\mathbf{x}}_\sigma$. And let's call the final vector of the sanitized trajectory the *sanitized final vector* $\hat{\mathbf{s}}_\sigma$. In both cases I've written a subscript σ because the vector depends on what σ is. Of course they both also depend on what the initial vector $\hat{\mathbf{x}}$ is, but I haven't indicated that.

What we have shown is

$$\|\hat{\mathbf{x}}_\sigma - \hat{\mathbf{s}}_\sigma\| \;\leq\; \varepsilon^2\beta \quad . \tag{64}$$

3.8 Good and Bad Epochs

Now let's compare an epoch of the probabilistic bucket brigade with a single step of the deterministic bucket brigade. They each begin with initial cash balance vector $\hat{\mathbf{x}}$. The probabilistic bucket brigade takes k steps, each with step size $\frac{\varepsilon}{k}$. The deterministic bucket brigade takes one step with step size ε . Remember that we said we are using a deterministic bucket brigade in which the payoff scaling factor equals the step size. $(\delta = \varepsilon)$

Let's see what happens in the one step of the deterministic bucket brigade. In theory, the one deterministic step should give us a sort of average change. That is, the vector it takes us to should be $\hat{\mathbf{x}}$ plus

$$\sum_{ij} F_{ij}\varepsilon\,\eta(\hat{\mathbf{x}}, i, j) \quad . \tag{65}$$

Using the definition of η , we see that this is
$\sum_{ij} F_{ij}[\varepsilon\hat{x}_j(\mathbf{e}_i - \mathbf{e}_j) + \varepsilon a_j\mathbf{e}_j - \varepsilon a_j\mathbf{w}]$.
This is formula (15) section 2 with $\delta = \varepsilon$.
(In (15) the change was from $\mathbf{v}^{(n)}$ and here it's from $\hat{\mathbf{x}}$.)

So we see that expression (65) is indeed the change vector for one step of the deterministic bucket brigade with step size ε and taking the step from position $\hat{\mathbf{x}}$.

So we see that this one step of the deterministic bucket brigade with step size ε takes us from vector $\hat{\mathbf{x}}$ to what I shall call the *Target Vector* $\hat{\mathbf{t}}$.
$\hat{\mathbf{t}} \;=\; \hat{\mathbf{x}} \;+\; \sum_{ij} F_{ij}\varepsilon\,\eta(\hat{\mathbf{x}}, i, j)$

We want to compare this with an Epoch of the probabilistic bucket brigade. During the Epoch, the probabilistic bucket brigade takes k steps, each with step size $\frac{\varepsilon}{k}$. That takes us to a final vector that we called $\hat{\mathbf{x}}_\sigma$. How far is that final vector liable to be from the Target Vector?

Taking k steps of the probabilistic bucket brigade leads us along a trajectory to a final vector. If we sanitize that trajectory, the sanitized trajectory leads us to what we called the sanitized final vector $\hat{\mathbf{s}}_\sigma$. We have seen that the distance between the final vector and the corresponding sanitized final vector is no greater than $\varepsilon^2\beta$. We now ask how far the sanitized final vector is from the Target Vector.

Suppose σ is the sequence of state transitions during the Epoch. Remember from the definition of $\mathcal{F}_{ij}^{(\sigma)}$ that $k\mathcal{F}_{ij}^{(\sigma)}$ is the number of transitions in σ that are $i \to j$.
Then the sanitized final vector $\hat{\mathbf{s}}_\sigma$ is
$\hat{\mathbf{x}} + \sum_{ij}(k\mathcal{F}_{ij}^{(\sigma)})\frac{\varepsilon}{k}\,\eta(\hat{\mathbf{x}},i,j)$.
So we have
$\hat{\mathbf{s}}_\sigma = \hat{\mathbf{x}} + \sum_{ij}\mathcal{F}_{ij}^{(\sigma)}\varepsilon\,\eta(\hat{\mathbf{x}},i,j)$.
We see the similarity of this formula to the formula for the Target Vector.

Let's compare the sanitized final vector with the target vector.
$\hat{\mathbf{s}}_\sigma - \hat{\mathbf{t}} = \varepsilon\sum_{ij}(\mathcal{F}_{ij}^{(\sigma)} - F_{ij})\,\eta(\hat{\mathbf{x}},i,j)$.
$\|\hat{\mathbf{s}}_\sigma - \hat{\mathbf{t}}\| \le \varepsilon\sum_{ij}|\mathcal{F}_{ij}^{(\sigma)} - F_{ij}|\,\|\eta(\hat{\mathbf{x}},i,j)\|$.
By inequality (61) we have
$\|\eta(\hat{\mathbf{x}},i,j)\| \le \beta$ for all i and j . Therefore
$\|\hat{\mathbf{s}}_\sigma - \hat{\mathbf{t}}\| \le \varepsilon\beta\sum_{ij}|\mathcal{F}_{ij}^{(\sigma)} - F_{ij}| \le \varepsilon\beta N^2\,(\max_{ij}|\mathcal{F}_{ij}^{(\sigma)} - F_{ij}|) = \varepsilon\beta N^2\,\text{wobble}(\sigma)$.
By definition, a Good Epoch is one in which $\text{wobble}(\sigma) < N^{-2}\varepsilon$. Therefore, in a Good Epoch,

$$\|\hat{\mathbf{s}}_\sigma - \hat{\mathbf{t}}\| \le \varepsilon^2\beta \quad . \tag{66}$$

Now how far can the final vector $\hat{\mathbf{x}}_\sigma$ (unsanitized) be from the Target Vector in a Good Epoch? In other words, how big can $\|\hat{\mathbf{x}}_\sigma - \hat{\mathbf{t}}\|$ be?
We use (64) and (66).
$\|\hat{\mathbf{x}}_\sigma - \hat{\mathbf{t}}\| \le \|\hat{\mathbf{x}}_\sigma - \hat{\mathbf{s}}_\sigma\| + \|\hat{\mathbf{s}}_\sigma - \hat{\mathbf{t}}\| \le \varepsilon^2\beta + \varepsilon^2\beta = 2\varepsilon^2\beta$.
Thus we have this lemma.

Lemma 19
Suppose a Good Epoch of the probabilistic bucket brigade has initial cash balance vector $\hat{\mathbf{x}}$ and final cash balance vector $\hat{\mathbf{x}}_\sigma$. And suppose a single step of the deterministic bucket brigade takes us from $\hat{\mathbf{x}}$ to cash balance vector $\hat{\mathbf{t}}$. Then we have

$$\|\hat{\mathbf{x}}_\sigma - \hat{\mathbf{t}}\| \le 2\varepsilon^2\beta \quad . \tag{67}$$

(Of course this lemma is assuming the probabilistic bucket brigade step size is $\frac{\varepsilon}{k}$, and the deterministic bucket brigade step size is ε .)

> **Reminder regarding β :**
> In this section and the previous section we have been discussing an Epoch. The initial vector of the Epoch is $\hat{\mathbf{x}}$. But the definition of β uses $\mathbf{v}$, the initial vector of the entire Era. This is the β that appears in the formulae of the last two sections.

3.9 Good Eras

Lemma 20
Suppose the Era begins with a sequence of n Good Epochs. Suppose the vector at the start of the Era is $\mathbf{v}$, the sequence of nk state transitions is σ , and the final vector at the end of the sequence is $\mathbf{x}_\sigma$. Then if we run the deterministic bucket brigade for n steps starting at vector $\mathbf{v}$ and using step size ε , and if the vector we arrive at that way is $\hat{\mathbf{t}}$, then we have
$\|\mathbf{x}_\sigma - \hat{\mathbf{t}}\| \le 2\varepsilon^2\beta n$.

Proof:
We prove this by induction on n . (trivial if $n = 0$.)
We assume the lemma holds for n and prove it for $n+1$.
So suppose we have a sequence of $n+1$ good epochs. Suppose the sequence starts at vector $\mathbf{v}$. Suppose the sequence of state transitions is $\tau\sigma$, where τ has k transitions and σ has nk transitions. Suppose the vector at the end of the first epoch is $\mathbf{x}_\tau$ and the final vector is $\mathbf{x}_{\tau\sigma}$.
We define the following two target vectors. If we run the deterministic bucket brigade for $n+1$ steps starting at $\mathbf{v}$, then we get vector $\hat{\mathbf{t}}$. If we run the deterministic bucket brigade for n steps starting at $\mathbf{x}_\tau$, then we get vector $\hat{\mathbf{t}}_\tau$.

By the induction assumption we have

$$\|\mathbf{x}_{\tau\sigma} - \hat{\mathbf{t}}_\tau\| \;\le\; 2\varepsilon^2\beta n \tag{68}$$

Now suppose a single step of the deterministic bucket brigade carries $\mathbf{v}$ to $\mathbf{y}$. Since the first epoch is a good one, inequality (67) gives us
$\|\mathbf{x}_\tau - \mathbf{y}\| \;\le\; 2\varepsilon^2\beta$.

Now we run the deterministic bucket brigade for n steps starting at $\mathbf{x}_\tau$ and starting at $\mathbf{y}$. These two runs take us to $\hat{\mathbf{t}}_\tau$ and $\hat{\mathbf{t}}$ respectively. Then the inequality (31) for this case is
$\|\hat{\mathbf{t}}_\tau - \hat{\mathbf{t}}\| \;\le\; \|\mathbf{x}_\tau - \mathbf{y}\|$.
The last two inequalities give us
$\|\hat{\mathbf{t}}_\tau - \hat{\mathbf{t}}\| \;\le\; 2\varepsilon^2\beta$.
Using this and (68) gives this.
$\|\mathbf{x}_{\tau\sigma} - \hat{\mathbf{t}}\| \;\le\; \|\mathbf{x}_{\tau\sigma} - \hat{\mathbf{t}}_\tau\| + \|\hat{\mathbf{t}}_\tau - \hat{\mathbf{t}}\| \;\le\; 2\varepsilon^2\beta n + 2\varepsilon^2\beta \;=\; 2\varepsilon^2\beta(n+1)$
■

Now we are going to run the probabilistic bucket brigade for an Era. The initial cash balance vector is $\mathbf{v}$. We write σ for the entire transition sequence and write $\hat{\mathbf{x}}_\sigma$ for the final vector.

We compare this with running the deterministic bucket brigade for ℓ steps, taking us to the final vector $\hat{\mathbf{t}}$. The step size is $\frac{\varepsilon}{k}$ for the probabilistic bucket brigade and ε for the deterministic bucket brigade.

If in the last lemma we set $n = \ell$, we have this lemma.

Lemma 21
Suppose we have a Good Era with initial vector $\mathbf{v}$ *, transition sequence* σ *, and final vector* $\hat{\mathbf{x}}_\sigma$ *. And suppose if we run the deterministic bucket brigade on* $\mathbf{v}$ *for* ℓ *steps with step size* ε *we arrive at vector* $\hat{\mathbf{t}}$ *. Then we have*
$\|\hat{\mathbf{x}}_\sigma - \hat{\mathbf{t}}\| \;\le\; 2\varepsilon\rho\beta$.

We now look at the deterministic bucket brigade and write inequality (53) using infinity norms of row vectors. It looks like this.
$\|\mathbf{v}^{(\ell)} - \mathbf{u}\| \;\le\; 2\Delta\,\|\mathbf{v} - \mathbf{u}\|$
Of course $\mathbf{v}^{(\ell)}$ is the result of running the deterministic bucket brigade for ℓ steps starting with initial vector $\mathbf{v}$. That's what $\hat{\mathbf{t}}$ is. In other words, $\mathbf{v}^{(\ell)} \;=\; \hat{\mathbf{t}}$. So we have
$\|\hat{\mathbf{t}} - \mathbf{u}\| \;\le\; 2\Delta\,\|\mathbf{v} - \mathbf{u}\|$.
For convenience, I am going to define

$$\nu \;=\; \|\mathbf{v} - \mathbf{u}\| \quad ,$$

the distance our starting vector is from the equilibrium cash balance vector. Thus we have,

$$\|\hat{\mathbf{t}} - \mathbf{u}\| \;\le\; 2\Delta\nu \quad . \tag{69}$$

Using the definition of θ from section 3.1 we have
$\|\mathbf{v}\| + \|\mathbf{a}\| \;\le\; \|\mathbf{v} - \mathbf{u}\| + \|\mathbf{u}\| + \|\mathbf{a}\| \;=\; \nu + 2\theta$.
Multiplying by e^ρ gives

$$\beta \;\le\; e^\rho(\nu + 2\theta) \quad . \tag{70}$$

From inequality (50) we obtain
$\Delta(\nu + 2\theta) \;\ge\; \varepsilon\rho\, e^\rho(\nu + 2\theta)$.
This and (70) give us

$$\Delta\,(\nu + 2\theta) \;\ge\; \varepsilon\rho\beta \quad . \tag{71}$$

Lemma 22
In any Good Era we have $\quad \|\hat{\mathbf{x}}_\sigma - \mathbf{u}\| \;\le\; 4\Delta(\nu + \theta)$.

Proof:
Using lemma 21 and inequalities (69) and (71), we have
$\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\| \;\le\; \|\hat{\mathbf{x}}_\sigma - \hat{\mathbf{t}}\| \,+\, \|\hat{\mathbf{t}} - \mathbf{u}\| \;\le\; 2\varepsilon\rho\beta \,+\, 2\Delta\nu \;\le\; 2\Delta(\nu + 2\theta) + 2\Delta\nu \;=\; 4\Delta(\nu + \theta)$.
■

3.10 The *real position* of a Cash Balance Vector

Look at the definition of Δ .
From $\Delta \;\le\; \frac{1}{4}\,e^{-d}$ we have

$$d \;\le -\log(4\Delta) \quad . \tag{72}$$

From $\Delta \;\le\; \frac{r}{4(r+\theta)}$ we have

$$4\Delta(r+\theta) \;\le\; r \quad . \tag{73}$$

We chose $\hat{n}$ to be the smallest integer greater than or equal to $2\hat{\rho}+1+2\frac{\theta}{r}$.
(See (34) for the definition.)
By (46) we have $\rho \;\le 2\hat{\rho}$. Therefore we have

$$\rho+1+2\,\frac{\theta}{r} \;\le\; \hat{n} \quad . \tag{74}$$

We define

$$\hat{w} \;=\; \log(r) \quad .$$

Since r is is positive but small, $\hat{w}$ is negative with a large magnitude.
By the property of log we have $\log(1+2\,\frac{\theta}{r}) \;\le\; 2\,\frac{\theta}{r}$.
$\log(r+2\theta) \;=\; \log(1+2\,\frac{\theta}{r}) \,+\, \log(r) \;\le\; 2\,\frac{\theta}{r} \,+\, \hat{w}$
$\rho \,+\, 1 \,+\, \log(r+2\theta) \;\le\; \rho \,+\, 1 \,+\, 2\,\frac{\theta}{r} \,+\, \hat{w} \;\le\; \hat{n} \,+\, \hat{w}$.

$$\rho \,+\, 1 \,+\, \log(r+2\theta) \;\le\; \hat{n} \,+\, \hat{w} \tag{75}$$

For each cash balance vector $\mathbf{x}$, we are interested in the log of its distance from $\mathbf{u}$.
We call that $\vartheta(\mathbf{x})$. That is,

$$\vartheta(\mathbf{x}) \;=\; \log(\|\mathbf{x}-\mathbf{u}\|) \quad .$$

We call $\vartheta(\mathbf{x})$ the *real position* of $\mathbf{x}$.
For completeness, we say that $\vartheta(\mathbf{u}) = -\infty$.

We have been discussing the behavior of the probabilistic bucket brigade through the whole ℓk steps of an Era, taking the system from initial vector $\mathbf{v}$ to final vector $\hat{\mathbf{x}}_\sigma$. At the start of the ℓk steps of the Era, the real position is $\vartheta(\mathbf{v})$. At the end, it's $\vartheta(\hat{\mathbf{x}}_\sigma)$.
By the Norm Bound Fact, lemma 18, we have
$\|\hat{\mathbf{x}}_\sigma\| \;\le\; \beta - \|\mathbf{a}\| \;\le\; \beta$.
From the definition of θ and ν we have
$\|\mathbf{u}\| \;\le\; \|\mathbf{u}\|+\|\mathbf{a}\| \;=\; 2\theta \;\le\; \nu+2\theta$.
From the last two inequalities, (47), and (70) we have

$$\begin{aligned}
\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\| \;&\le\; \|\hat{\mathbf{x}}_\sigma\| \,+\, \|\mathbf{u}\| \\
&\le\; \beta + (\nu+2\theta) \\
&\le\; e^{\rho}(\nu+2\theta) \,+\, e^{\rho}(\nu+2\theta) \\
&=\; 2e^{\rho}(\nu+2\theta) \\
&\le\; e^{\rho+1}(\nu+2\theta) \quad .
\end{aligned}$$

So we have
$\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\| \;\le\; e^{\rho+1}(\nu+2\theta)$.
Now $\nu+2\theta$ is non-negative. If it is positive, we can take logs of both sides and obtain

$$\vartheta(\hat{\mathbf{x}}_\sigma) \;\;\le\;\; \rho+1+\log(\nu+2\theta) \quad . \tag{76}$$

(Of course $\vartheta(\hat{\mathbf{x}}_\sigma)$ might be $-\infty$.)
On the other hand, if $\nu+2\theta \;=\; 0$ then the previous inequality tells us that $\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\| \;=\; 0$, and so $\vartheta(\hat{\mathbf{x}}_\sigma) \;=\; -\infty$, and inequality (76) holds trivially ($-\infty \le -\infty$). So inequality (76) always holds.

We note that $\log(\nu) \;=\; \log(\|\mathbf{v}-\mathbf{u}\|) \;=\; \vartheta(\mathbf{v})$, so without comment we shall use

$$\log(\nu) \;=\; \vartheta(\mathbf{v}) \quad .$$

(Of course if $\mathbf{v}=\mathbf{u}$ then both sides are $-\infty$.)

We now ask where the real position $\vartheta(\hat{\mathbf{x}}_\sigma)$ is in comparison to $\vartheta(\mathbf{v})$.

3.11 The $\vartheta(\mathbf{v}) < \hat{w}$ Case

Lemma 23
If $\vartheta(\mathbf{v}) \leq \hat{w}$ *then* $\nu \leq r$.

Proof:
If $\mathbf{v} = \mathbf{u}$ then $\nu = 0$, so $\nu \leq r$.
If $\mathbf{v} \neq \mathbf{u}$ then we have $\nu > 0$ and
$\log(\nu) = \vartheta(\mathbf{v}) \leq \hat{w}$.
Taking exponentials gives
$\nu \leq r$.
■

Lemma 24
If $\vartheta(\mathbf{v}) \leq \hat{w}$ *then*
$\vartheta(\hat{\mathbf{x}}_\sigma) \leq \hat{w} + \hat{n}$.

Proof:
From (76), lemma 23, and (75) we have
$\vartheta(\hat{\mathbf{x}}_\sigma) \leq \rho + 1 + \log(\nu + 2\theta) \leq \rho + 1 + \log(r + 2\theta) \leq \hat{w} + \hat{n}$.
■

Lemma 25
In any Good Era, if $\vartheta(\mathbf{v}) \leq \hat{w}$ *then*
$\vartheta(\hat{\mathbf{x}}_\sigma) \leq \hat{w}$.

Proof:
From lemma 22, lemma 23, and inequality (73) we have
$\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\| \leq 4\Delta(\nu + \theta) \leq 4\Delta(r + \theta) \leq r$.
Taking logs gives
$\vartheta(\hat{\mathbf{x}}_\sigma) \leq \hat{w}$.
(Of course it's possible that $\hat{\mathbf{x}}_\sigma = \mathbf{u}$, in which case $\vartheta(\hat{\mathbf{x}}_\sigma) = -\infty$.)
■

3.12 The $\vartheta(\mathbf{v}) \geq \hat{w}$ Case

Lemma 26
If $\vartheta(\mathbf{v}) \geq \hat{w}$, *then* $\nu > 0$ *and* $\frac{\theta}{\nu} \leq \frac{\theta}{r}$.

Proof:
If $\nu = 0$, then
$\vartheta(\mathbf{v}) = \log(\nu) = -\infty < \hat{w}$.
Since $\vartheta(\mathbf{v}) \geq \hat{w}$, we must have $\nu > 0$.
We now write
$\log(\nu) = \vartheta(\mathbf{v}) \geq \hat{w}$.
Taking exponentials of both sides gives us
$\nu \geq r$. Therefore,
$\frac{\theta}{\nu} \leq \frac{\theta}{r}$.
■

Lemma 27
If $\vartheta(\mathbf{v}) \geq \hat{w}$ *then*
$\vartheta(\hat{\mathbf{x}}_\sigma) \leq \vartheta(\mathbf{v}) + \hat{n}$.

Proof:
By lemma 26 we know $\nu > 0$.
Thus we can write
$\nu + 2\theta = \nu(1 + 2\frac{\theta}{\nu})$.
Taking logs of both sides gives us this.

$\log(\nu+2\theta) \ = \ \vartheta(\mathbf{v}) + \log(1+2\frac{\theta}{\nu})$
By the property of log we have
$\log(1+2\frac{\theta}{\nu}) \ \leq \ 2\frac{\theta}{\nu}$
and by lemma 26 we have
$\frac{\theta}{\nu} \ \leq \frac{\theta}{r}$.
Putting the last three items together gives us
$\log(\nu+2\theta) \ = \ \vartheta(\mathbf{v}) + \log(1+2\frac{\theta}{\nu}) \ \leq \ \vartheta(\mathbf{v}) + 2\frac{\theta}{\nu} \ \leq \ \vartheta(\mathbf{v}) + 2\frac{\theta}{r}$.
From this and (76), we get
$\vartheta(\hat{\mathbf{x}}_\sigma) \ \leq \ \rho + 1 + \vartheta(\mathbf{v}) + 2\frac{\theta}{r}$.
From this and (74), we obtain
$\vartheta(\hat{\mathbf{x}}_\sigma) \ \leq \ \vartheta(\mathbf{v}) + \hat{n}$.
■

Lemma 28
In any Good Era, if $\vartheta(\mathbf{v}) \geq \hat{w}$ *then we have*
$\vartheta(\hat{\mathbf{x}}_\sigma) \ \leq \ \vartheta(\mathbf{v}) \ - \ 1$.

Proof:
By lemma 26 we have
$\nu > 0$ and
$\frac{\theta}{\nu} \ \leq \frac{\theta}{r}$.
From this and (72) we have the following:
$\frac{\theta}{\nu} + 1 \ \leq \ \frac{\theta}{r} + 1 \ = \ d \ \leq \ -\log(4\Delta)$
$\log(4\Delta) + \frac{\theta}{\nu} \ \leq \ -1$
By the property of log we have $\log(1+\frac{\theta}{\nu}) \ \leq \frac{\theta}{\nu}$, so from this and the previous inequality we have

$$\log(4\Delta) + \log(1+\frac{\theta}{\nu}) \ \leq \ -1 \tag{77}$$

By lemma 26 we have $\nu > 0$, so we can write
$\nu + \theta \ = \ \nu(1+\frac{\theta}{\nu})$.
From this and lemma 22 we have
$\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\| \ \leq \ 4\Delta\nu(1+\frac{\theta}{\nu})$.
The right side is positive and the left side is non-negative. So we can take logs.
$\vartheta(\hat{\mathbf{x}}_\sigma) \ \leq \ \log(4\Delta) + \vartheta(\mathbf{v}) + \log(1+\frac{\theta}{\nu})$
(Of course the left side might be $-\infty$.)
From this last inequality and inequality (77), we have
$\vartheta(\hat{\mathbf{x}}_\sigma) \ \leq \ \vartheta(\mathbf{v}) \ - \ 1$.
■

3.13 Distribution of Real Positions

We imagine that a frog is hopping on the real line.
The distance the frog hops is always an integer. It's always $\hat{n}$ or 1 or 0.
I shall call the following set of real numbers the set of

frog-states.
$\{\hat{w}, \quad \hat{w}+1, \quad \hat{w}+2, \quad \hat{w}+3, \quad \hat{w}+4, \quad ,,,, \quad \}$

We place a frog on one of the frog-states and make him hop once every era.
(It won't be important to our argument, but you may remember that $\hat{w}$ is negative with a large magnitude.)

The frog hops according to the following rules.

The frog takes a look at the chain state i at the start of the era. This determines p_i . Remember that p_i is the probability that the next Era will be good. Remember too that $p_i \ \geq \ \dot{p}$. (See inequality (55).)

At the end of the era, the frog looks at the string σ that the chain actually traversed during the era, and so determines whether the era was indeed Good or Bad.

Hopping Rules:

If the era was Bad, then the frog hops distance $\hat{n}$ right.

If the era was Good then the frog does the following:

With probability $(p_i - \dot{p})\, p_i^{-1}$ he hops distance $\hat{n}$ right.

With probability $\dot{p}\, p_i^{-1}$ he hops distance 1 left,

unless he is at $\hat{w}$, in which case he stays put.

(By "stays put" I mean that the frog does hop, but he lands on the same number $\hat{w}$ that he hopped from.)
So the frog hops at the end of each era.
The first thing to note is that the frog is always on one of the frog-states. He never hops left of $\hat{w}$.
Let's look at one era. The frog has just hopped and the new era is about to start.

Let F_s be the position of the frog at the start of the era and let F_e be the position of the frog after he hops at the end of the era.

We already know that

$$F_s \;\geq\; \hat{w} \quad \text{and} \quad F_e \;\geq\; \hat{w} \quad . \tag{78}$$

We shall use these inequalities, often without comment, in the proof of the next lemma.

Lemma 29

If $F_s \;\geq\; \vartheta(\mathbf{v})$ *then* $F_e \;\geq\; \vartheta(\hat{\mathbf{x}}_\sigma)$.

Proof:

Suppose $F_s \;\geq\; \vartheta(\mathbf{v})$.

To prove $F_e \;\geq\; \vartheta(\hat{\mathbf{x}}_\sigma)$ we look at the direction the frog hopped.

If the frog hopped right then we do one of these two things:
If $\vartheta(\mathbf{v}) \;<\; \hat{w}$ then we use (78) and lemma 24 to obtain
$F_e \;=\; F_s + \hat{n} \;\geq\; \hat{w} + \hat{n} \;\geq\; \vartheta(\hat{\mathbf{x}}_\sigma)$.
If $\vartheta(\mathbf{v}) \;\geq\; \hat{w}$ then we use lemma 27 to obtain
$F_e \;=\; F_s + \hat{n} \;\geq\; \vartheta(\mathbf{v}) + \hat{n} \;\geq\; \vartheta(\hat{\mathbf{x}}_\sigma)$.
So if the frog hopped right then $F_e \;\geq\; \vartheta(\hat{\mathbf{x}}_\sigma)$.

Now suppose the frog hopped left.
So $F_e \;=\; F_s - 1$ and $F_s \;\geq \hat{w} + 1$.
In this case the era was Good.
If $\vartheta(\mathbf{v}) \;<\; \hat{w}$ then we use lemma 25 to obtain
$F_e \;\geq\; \hat{w} \;\geq \vartheta(\hat{\mathbf{x}}_\sigma)$.
If $\vartheta(\mathbf{v}) \;\geq\; \hat{w}$ then we use lemma 28 to obtain
$\vartheta(\hat{\mathbf{x}}_\sigma) \;\leq\; \vartheta(\mathbf{v}) - 1 \;\leq\; F_s - 1 \;=\; F_e$.
So if the frog hopped left then $F_e \;\geq\; \vartheta(\hat{\mathbf{x}}_\sigma)$.

Now suppose the frog stayed put.
So $F_e \;=\; F_s \;=\; \hat{w}$.
In this case the era was Good.
We have $\vartheta(\mathbf{v}) \;\leq\; F_s \;=\; \hat{w}$,
so we can use lemma 25 to obtain
$\vartheta(\hat{\mathbf{x}}_\sigma) \;\leq\; \hat{w} \;=\; F_e$.
So if the frog stayed put then $F_e \;\geq\; \vartheta(\hat{\mathbf{x}}_\sigma)$.

So in all three cases we have $F_e \;\geq\; \vartheta(\hat{\mathbf{x}}_\sigma)$.
■

Now suppose we are running the probabilistic bucket brigade with step size $\frac{\varepsilon}{k}$. We have our integer $\hat{w}$ and our probability $\dot{p}$. For each chain state i we have the probability p_i . We select a time and call it time 0. We then conceptually mark the times
$0,\ \ell k,\ 2\ell k,\ 3\ell k,\ 4\ell k,\ ....$.
I will call these the marked times.
An interval from one marked time to the next is of course an era. The chain states at the marked times are
$i_0,\ i_1,\ i_2,\ i_3,\ i_4,\ ....$.
The cash balance vectors at the marked times are
$\mathbf{x}_0,\ \mathbf{x}_1,\ \mathbf{x}_2,\ \mathbf{x}_3,\ \mathbf{x}_4,\ ....$.
The real position of a cash balance vector $\mathbf{x}$ is of course

$$\vartheta(\mathbf{x}) \;=\; \log(\|\mathbf{x} - \mathbf{u}\|) \quad . \tag{79}$$

The real positions at the marked times are
$\vartheta(\mathbf{x}_0),\ \vartheta(\mathbf{x}_1),\ \vartheta(\mathbf{x}_2),\ \vartheta(\mathbf{x}_3),\ \vartheta(\mathbf{x}_4),\ \ldots.$.

At time zero we place the frog on a frog state and let the frog hop at each marked time according to the Hopping Rules. The frog's positions at the marked times are
$F_0,\ F_1,\ F_2,\ F_3,\ F_4,\ \ldots.$.
Position F_0 is where we initially placed the frog. Position F_1 is where the frog lands after his first hop. That hop is made at time ℓk , and in this case the era referred to in the Hopping Rules is the era from time 0 to time ℓk . Real position $\vartheta(\mathbf{x}_1)$ is the real position at the end of that era, and F_1 is the frog's position after the hop he made at the end of that era.

Lemma 29 tells us that we have the following for any non-negative integer n .

Step Consistency Fact:
If $F_n \geq \vartheta(\mathbf{x}_n)$ then $F_{n+1} \geq \vartheta(\mathbf{x}_{n+1})$.

We arrange that the position in which we initially place the frog is
not left of $\vartheta(\mathbf{x}_0)$. That is,
$F_0 \geq \vartheta(\mathbf{x}_0)$.
We see from the Step Consistency Fact above that if we do this then the frog will never hop left of the current real position. In other words.

Lemma 30
For all non-negative integers n *we have*
$F_n \geq \vartheta(\mathbf{x}_n)$.

Now let's examine the frog's hops. His right-hops are always hops of distance $\hat{n}$. All his other hops I am going to call left-hops. A left-hop is of distance 1 unless the frog is sitting on $\hat{w}$, in which case it is of distance zero. (I know that a hop of distance zero is not really a left hop, but I'm calling it a left-hop.)

Now suppose the frog has just hopped, the current chain state is i , and we are at the beginning of an era. The frog's next hop will be at the end of that era. Will it be a right-hop or a left-hop? We see from the Hopping Rules that whether it is a right-hop or left-hop depends on whether the era turns out to be a Good Era or a Bad Era. That is, it depends on what the sequence σ of the next ℓk transitions turns out to be.

The probability that the era will be Good is p_i . The frog does a left-hop only if the era is Good, and then only with probability $\dot{p}\,p_i^{-1}$. So the probability that the frog does a left-hop is
$p_i\,[\dot{p}\,p_i^{-1}] = \dot{p}$.

The probability that the era will be Bad is $1-p_i$, so we can similarly use the Hopping Rules to calculate the probability that the frog does a right-hop. That probability is
$(1-p_i) + p_i\,[(p_i-\dot{p})\,p_i^{-1}] = 1-\dot{p}$.
But of course. The frog does a right-hop if and only if he doesn't do a left-hop, and the probability he doesn't do a left-hop is $1-\dot{p}$. We define $\dot{q} = 1-\dot{p}$, so we have $\dot{p}+\dot{q} = 1$.

So we see that each frog hop is a left-hop with probability $\dot{p}$ and a right-hop with probabity $\dot{q}$. A right-hop is always distance $\hat{n}$. A left-hop is distance 1, unless the frog is on $\hat{w}$, in which case it's distance 0.

So the frog has the Markov property. In fact, it's a Markov chain just like the one in section C. The frog follows the transition rules given at the start of section C. (The p and n there are $\dot{p}$ and $\hat{n}$ here.)

Of course in section C the frog was hoping over non-negative integers and here he is hopping over frog-states. But it's really the same notion. Integer i in section C corresponds to frog-state $\hat{w}+i$ here. So if we translate lemma 40 of section C into the language of frog-states, we have the following.

If the condition $\frac{1}{2} < \dot{p}^{\hat{n}}$ is met, then
the proportion of time the frog spends[10] at frog-state $\hat{w}$ is at least $2-\dot{p}^{-\hat{n}}$.

Inequality (36) tells us that the condition is indeed met, so the proportion of time the frog spends at $\hat{w}$ is indeed at least $2-\dot{p}^{-\hat{n}}$.

From (35) we have
$2-\dot{p}^{-\hat{n}} = 1-\hat{q}$,
so we conclude:

The proportion of time the frog spends at frog state $\hat{w}$ is at least $1-\hat{q}$.

[10]As I said in section C, my phrase "proportion of time the frog spends" means the number to which the proportion converges with probability 1.

Let's write F for the current frog position and $\vartheta(\mathbf{x})$ for the current real position. Lemma 30 tells us that if the frog has just hopped then $\vartheta(\mathbf{x}) \leq F$. And the proportion of hops that land on $\hat{w}$ is at least $1-\hat{q}$. The hops take place at the marked times, so if we look just at the marked times, we always have $\vartheta(\mathbf{x}) \leq F$, and in at least proportion $1-\hat{q}$ of these times we have $F = \hat{w}$. In that proportion of times we must have $\vartheta(\mathbf{x}) \leq \hat{w}$, though we could also have $\vartheta(\mathbf{x}) \leq \hat{w}$ at some times when $F > \hat{w}$. So the proportion of marked times in which $\vartheta(\mathbf{x}) \leq \hat{w}$ is at least $1-\hat{q}$, though it could be a larger proportion. To sum up.

Lemma 31
The proportion of marked times in which we have $\vartheta(\mathbf{x}) \leq \hat{w}$ *is at least* $1-\hat{q}$.

But what is a marked time? We ran the probabilistic bucket brigade producing a sequence of cash balance vectors. We then marked some of the times, spacing them at intervals of length ℓk. But where did we start marking? If we had started marking one time unit later, we would have marked entirely different times, and lemma 31 would hold on these times. And if we had started marking two time units later we would mark yet different times and the lemma would still hold. So we see that in fact the lemma holds for all times.

Lemma 32
The proportion of times in which we have $\vartheta(\mathbf{x}) \leq \hat{w}$ *is at least* $1-\hat{q}$.
In other words,

$$\wp[\,\vartheta(\mathbf{x}) \leq \hat{w}\,] \quad \geq \quad 1-\hat{q} \quad .$$

The lemma conclusion can be written
$\wp[\,\vartheta(\mathbf{x}) > \hat{w}\,] \leq \hat{q}$.
We can take exponentials inside the square brackets. Then our conclusion looks like this.

$$\wp[\,\|\mathbf{x}-\mathbf{u}\| > r\,] \leq \hat{q} \quad .$$

Recap:

We began (at the start of subsection 3.1) with a positive number r and a positive probability $\hat{q}$.
In subsection 3.1 we selected a real number $\hat{\varepsilon}$. At the end of subsection 3.3, we selected an integer $\mathcal{K}$.
The $\hat{\varepsilon}$ and $\mathcal{K}$ were appropriate for this r and $\hat{q}$.
Then in subsection 3.4, we chose a step size, which we called either $\bar{\varepsilon}$ or $\frac{\varepsilon}{k}$.
We said that the step size could be any positive real less than or equal to $(\mathcal{K}+1)^{-1}\hat{\varepsilon}$.
We showed that if the probabilistic bucket brigade uses this step size, then $\wp[\,\|\mathbf{x}-\mathbf{u}\| > r\,] \leq \hat{q}$.
But from the way we chose the step size, we see that any smaller step size would have given us the same inequality. So the inequality holds provided the step size is small enough.

But remember that the whole argument assumed that the payoff scaling factor was equal to the step size both in the probabilistic bucket brigade and the deterministic bucket brigade to which we compared it. So what we have proved is this

Theorem 1
Given any positive real numbers r *and* $\hat{q}$,
there is a positive real number $\mathcal{E}$ *such that for any real* $\bar{\varepsilon}$ *in the range* $0 < \bar{\varepsilon} \leq \mathcal{E}$ *we have the following.*
If the step size and payoff scaling factor are both $\bar{\varepsilon}$, *then*

$$\wp[\,\|\mathbf{x}-\mathbf{u}\| > r\,] \leq \hat{q} \quad . \tag{80}$$

3.14 How Did We Do That?

This subsection is motivational and informal and not necessary. It explains why we decided to do what we did.

Our argument all looks a bit like hocus-pocus. We needed some sort of bound on how far during an era the cash balance vector $\mathbf{x}$ can wander from $\mathbf{u}$ in the worst cases. That was tricky. These worst cases occurred in bad eras. Once we had a bound then we could increase k to lower the probability that the era would be bad. But that means the bound had to be independent of k. With a bound independent of k, we can reduce the effect of the bad cases as much as we like. But how did we get a bound independent of k?

So let's look at one of the worst cases. In many of the steps, $\mathbf{x}$ is moving away from $\mathbf{u}$. A big problem for us is that the size of the jump vector $\frac{\varepsilon}{k}\,\eta(\mathbf{x},i,j)$ is bigger if we are farther from the origin.

So if the vector $\mathbf{v}$ at the start of the era is large, the system can get a long way from $\mathbf{u}$ during the era. The larger $\mathbf{v}$ is, the farther the final vector $\hat{\mathbf{x}}_\sigma$ might be from $\mathbf{u}$. (I'm writing $\mathbf{v}$ for the vector at the start of the era and $\hat{\mathbf{x}}_\sigma$ for the vector at the end.)

So the problem is this. We want a bound on $\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\|$, but the larger $\|\mathbf{v}\|$ is the larger that bound will have to be.

But it's not really that bad. We don't really need a bound on $\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\|$. All we need is a bound on the ratio of $\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\|$ to $\|\mathbf{v} - \mathbf{u}\|$. (In our proof we used logs so the ratio became a difference, but here I won't do that because it introduces some hocus pocus.)

That's what we did. Our bound was on the ratio, so if $\mathbf{v}$ was far from $\mathbf{u}$ then $\hat{\mathbf{x}}_\sigma$ was allowed to be far from $\mathbf{u}$ too. A large $\mathbf{v}$ allowed bigger pathological steps, but our bound allowed that. A large $\mathbf{v}$ was no longer a problem.

The bound we got on the ratio of $\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\|$ to $\|\mathbf{v} - \mathbf{u}\|$ was very much like $e^{\rho+1}$. That's a fine bound and works very well. We have solved the problem of a large $\|\mathbf{v} - \mathbf{u}\|$.

The problem is that we don't really have that bound. The reason is this. The size of the jump vector is governed not by how far $\mathbf{v}$ is from $\mathbf{u}$, but by how far $\mathbf{v}$ is from the origin. We don't really have a bound on the ratio of $\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\|$ to $\|\mathbf{v} - \mathbf{u}\|$. Our bound is really on the ratio of $\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\|$ to $\|\mathbf{v}\|$.

The difference between the two ratios is not significant if $\|\mathbf{v} - \mathbf{u}\|$ is large. It is the distance between $\mathbf{u}$ and the origin that causes the problem, but if $\mathbf{v}$ is far away, the distance between $\mathbf{u}$ and the origin is comparatively insignificant. Then for calculating our ratio, $\|\mathbf{v} - \mathbf{u}\|$ and $\|\mathbf{v}\|$ are almost the same. As we said, we have solved the problem of a large $\|\mathbf{v} - \mathbf{u}\|$.

The problem here is at the other extreme. The problem is with an excessively *small* $\|\mathbf{v} - \mathbf{u}\|$, because that quantity is in the ratio's denominator. The denominator can be tiny, yet the numerator $\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\|$ can still be large because it depends on $\|\mathbf{v}\|$ rather than $\|\mathbf{v} - \mathbf{u}\|$. So when $\mathbf{v}$ gets close to $\mathbf{u}$, the ratio gets huge. We can't bound the ratio when $\mathbf{v}$ is close to $\mathbf{u}$.

So we didn't. We introduced the frog, and instead bounded the ratio of the frog's starting and finishing position. (The frog was actually hopping on a log scale, but in this discussion I'm not taking the log.) The frog always stayed farther from $\mathbf{u}$ than $\mathbf{v}$ was, but more important, the frog never got closer to $\mathbf{u}$ than the distance r. Thus the ratio's denominator could never get smaller than r, and so we could get a bound on the ratio.

When $\|\mathbf{v} - \mathbf{u}\|$ is smaller than r, our bound on the numerator $\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\|$ is

$$e^{\rho + 1 + \log(r + \|\mathbf{u}\| + \|\mathbf{a}\|)} .$$

It's huge, but it's independent of k. (Lemma 24 gives the logarithmic version of this bound.)

For bigger $\|\mathbf{v} - \mathbf{u}\|$, our bound on the numerator $\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\|$ is

$$\|\mathbf{v} - \mathbf{u}\| \, e^{\rho + 1 + \frac{1}{r}(\|\mathbf{u}\| + \|\mathbf{a}\|)} .$$

(Lemma 27 gives the logarithmic version of this bound, since the conclusion of that lemma is $\log(\|\hat{\mathbf{x}}_\sigma - \mathbf{u}\|) \le \log(\|\mathbf{v} - \mathbf{u}\|) + \rho + 1 + \frac{1}{r}(\|\mathbf{u}\| + \|\mathbf{a}\|)$.)

To get a bound on the ratio for the frog hop we use these numerators and actually increase them, but slightly. The trick is what we do with the denominators, because we want the denominators large. If $\|\mathbf{v} - \mathbf{u}\|$ is large, we use $\|\mathbf{v} - \mathbf{u}\|$ itself, but if it's small, we use r.

That's the idea. It's much tidier if we take logarithms, and that's what we did in our formal argument. This subsection is informal, but it gives the motivation for the approach we took. Our formal argument does not depend on anything in this subsection.

4 Extensions

4.1 Estimating c — different payoff scaling factors

The proof of (80) made the assumption that the payoff scaling factor equals the step size, both in the probabilistic bucket brigade and in the deterministic bucket brigade with which we are comparing it. Thus in equation (23), we have $\delta = \varepsilon$, and consequently $\mathbf{c}^\top = \mathbf{u}^\top - \bar{u}\,\mathbf{e}^\top$, where of course $\bar{u} = \tilde{\mathbf{p}}\mathbf{u}^\top$.

We are using the bucket brigade to estimate $\mathbf{c}$. Theorem 1 says that the cash balance vector $\mathbf{x}$ is close to $\mathbf{u}$, so a reasonable estimate of $\mathbf{c}$ is the vector $\boldsymbol{\gamma}$, defined by
$\boldsymbol{\gamma}^\top = \mathbf{x}^\top - \bar{x}\,\mathbf{e}^\top$,
where $\bar{x} = \tilde{\mathbf{p}}\mathbf{x}^\top$.
We now ask how good this estimate is.

First we note that $\tilde{\mathbf{p}}(\mathbf{x}^\top - \mathbf{u}^\top)$ can't have modulus greater than $\|\mathbf{x} - \mathbf{u}\|$. So we have
$|\bar{x} - \bar{u}| = |\tilde{\mathbf{p}}(\mathbf{x}^\top - \mathbf{u}^\top)| \le \|\mathbf{x} - \mathbf{u}\|$.
$\boldsymbol{\gamma}^\top - \mathbf{c}^\top = (\mathbf{x}^\top - \mathbf{u}^\top) + (\bar{x} - \bar{u})\,\mathbf{e}^\top$
$\|\boldsymbol{\gamma}^\top - \mathbf{c}^\top\| \le \|\mathbf{x}^\top - \mathbf{u}^\top\| + |\bar{x} - \bar{u}|\,\|\mathbf{e}^\top\| = \|\mathbf{x}^\top - \mathbf{u}^\top\| + |\bar{x} - \bar{u}| \le 2\,\|\mathbf{x}^\top - \mathbf{u}^\top\|$
So we have
$\|\boldsymbol{\gamma} - \mathbf{c}\| \le 2\,\|\mathbf{x} - \mathbf{u}\|$.
Therefore we have the following.
$\|\mathbf{x} - \mathbf{u}\| \le r \implies \|\boldsymbol{\gamma} - \mathbf{c}\| \le 2r$
$\|\boldsymbol{\gamma} - \mathbf{c}\| > 2r \implies \|\mathbf{x} - \mathbf{u}\| > r$
$\wp[\,\|\boldsymbol{\gamma} - \mathbf{c}\| > 2r\,] \le \wp[\,\|\mathbf{x} - \mathbf{u}\| > r\,]$
Then by (80) we have

$$\wp[\,\|\boldsymbol{\gamma} - \mathbf{c}\| > 2r\,] \le \hat{q} \quad . \tag{81}$$

This holds provided the step size is small enough, assuming always that the payoff scaling factor equals the step size.

But what if it doesn't? The step size we decided on was $\bar{\varepsilon}$ (or $\frac{\varepsilon}{k}$ if you prefer) so the payoff scaling factor is $\bar{\varepsilon}$. What if we change the payoff scaling factor? For example, what if we change it to 1 ? Then when we visit state i , instead of adding $\bar{\varepsilon} a_i$ to the cash balance of state i , we add a_i . The amount of added cash has been multiplied by $\bar{\varepsilon}^{-1}$. And of course the keep the amount of cash constant we subtracted $w_k \bar{\varepsilon} a_i$ from the cash balance of each state k , and now we are subtracting $w_k a_i$.

This isn't really a problem. It's really just an exchange rate issue. It's as if we had been figuring all our cash in terms of dollars and now we switched and figured it in terms of drachmas. If a drachma is worth only $\bar{\varepsilon}$ of a dollar (that is, there are $\bar{\varepsilon}^{-1}$ drachmas to the dollar) then all the cash figures get multiplied by $\bar{\varepsilon}^{-1}$. All the cash balances are $\bar{\varepsilon}^{-1}$ times what they were before. The step size is still $\bar{\varepsilon}$, so in each step, proportion $\bar{\varepsilon}$ of the cash is passed back. Everything works just as before, but now in drachmas. This all works because the bucket brigade has what Chris Watkins calls the "cash conservation property". (Q-learning doesn't have this property.)

So now all we need to do is convert back to dollars when estimating $\mathbf{c}$. Our estimate now is
$\boldsymbol{\gamma}^\top = \bar{\varepsilon}\,(\mathbf{x}^\top - \bar{x}\,\mathbf{e}^\top)$.
Then (81) still holds.

Of course we can use the same trick with any payoff scaling factor we like. If we want to use payoff scaling factor δ , we simply use the following estimate of $\mathbf{c}$.

$$\boldsymbol{\gamma}^\top = \delta^{-1}\bar{\varepsilon}\,(\mathbf{x}^\top - \bar{x}\,\mathbf{e}^\top) \quad . \tag{82}$$

So we have shown the following.[11]

Corollary 1
Given any positive real numbers δ, r, and $\hat{q}$,
there is a positive number $\mathcal{E}$ such that for any real $\bar{\varepsilon}$ in the range $0 < \bar{\varepsilon} \le \mathcal{E}$ we have the following.
If the step size is $\bar{\varepsilon}$ and the payoff scaling factor is δ , and if we define
$\boldsymbol{\gamma}^\top = \delta^{-1}\bar{\varepsilon}\,(\mathbf{x}^\top - \bar{x}\,\mathbf{e}^\top)$ *, then*
$\wp[\,\|\boldsymbol{\gamma} - \mathbf{c}\| > r\,] \le \hat{q}$.

This is satisfactory. The vector $\boldsymbol{\gamma}$ can be used as an estimate of $\mathbf{c}$ in adaptation schemes (schemes for changing P). The vector $\boldsymbol{\gamma}$ is reasonably easy to calculate. The $\bar{\varepsilon}$ is the proportion of cash passed back in the bucket brigade so it is known by the system, as is δ and $\mathbf{x}$. The $\bar{x}$ is a bit annoying. We see that
$\boldsymbol{\gamma} = \delta^{-1}\bar{\varepsilon}\mathbf{x} - \delta^{-1}\bar{\varepsilon}\bar{x}\mathbf{e}$.
It would be simpler if we could dispense with the second term on the right. This essay is not the place to discuss adaptation, but I will mentiion that in many adaptation schemes the second term is automatically normalized out, so it is possible to simply use $\boldsymbol{\gamma} = \delta^{-1}\bar{\varepsilon}\mathbf{x}$. The reason has to do with the fact that the vector described by the second term has every entry the same.

Suppose we re-write the last corollary, but change the last line to
$\wp[\,\|\mathbf{x} - \mathbf{u}\| > r\,] \le \hat{q}$.
Is it still true? The short answer is that I don't know. As we look at smaller and smaller $\bar{\varepsilon}$, with δ held constant, the vectors $\mathbf{x}$ and $\mathbf{u}$ get bigger and bigger, so it seems possible that the error $\|\mathbf{x} - \mathbf{u}\|$ will become more likely to be greater than r . But I haven't investigated this matter.

[11] For simplicity I'm writing $2r$ as r .

4.2 Using $\mathbf{m}$ Instead of $\mathbf{a}$

In component (2) of the Cash Change Description (section 2) the bucket brigade adds δa_j to the cash balance of state j . Now $a_j = m_j - \bar{m}$, so if the bucket brigade is really going to add δa_j , it needs to know what $\bar{m}$ is. It is not too difficult to estimate $\bar{m}$, but the whole business is rather annoying. It would be simpler just to add δm_j . Then in component (3) it can compensate by subtracting $\delta m_j w_k$ from the cash balance of every state k .

I had assumed that if we made this change it would make no difference to our results. I was wrong. That's true only if $\mathbf{w} = \tilde{\mathbf{p}}$, which is annoying. Let's see why.

There are really two questions.

Question 1:
If we make the change from $\mathbf{a}$ to $\mathbf{m}$ does the deterministic bucket brigade still converge to some vector $\mathbf{u}$, and if so, is that $\mathbf{u}$ the same as the $\mathbf{u}$ we obtained when we were using $\mathbf{a}$?
Question 2:
Assuming the deterministic bucket brigade converges to a vector $\mathbf{u}$, does the probabilistic bucket brigade converge to the same vector in the sense of theorem 1?

Let's call the bucket brigade with this change made the *new bucket brigade* and call the original bucket brigade the *old bucket brigade*. So the new bucket brigade uses $\mathbf{m}$ in components (2) and (3), whereas the old bucket brigade uses $\mathbf{a}$.

Let's look at **Question 1**. Suppose we take the argument in section 2 and simply read the argument with $\mathbf{m}$ in place of $\mathbf{a}$ as we have suggested. Does the argument still work?

Things go fine up to formula (16). This is the formula for the change in the cash balance vector $\mathbf{v}^{(n)}$ in one step of the deterministic bucket brigade. With $\mathbf{m}$ in place of $\mathbf{a}$, it now reads as follows.
$\sum_{ij} F_{ij}\varepsilon v_j^{(n)}(\mathbf{e}_i - \mathbf{e}_j) + \sum_{ij} F_{ij}\delta m_j \mathbf{e}_j - \delta\sum_{ij} F_{ij} m_j \mathbf{w}$
And that's where things start to go wrong. What we next did was say that the third term is zero. But now the third term is not zero. It was zero because $\sum_{ij} F_{ij} a_j = 0$. But $\sum_{ij} F_{ij} m_j = \sum_j \tilde{p}_j m_j = \bar{m}$, so now the third term is $-\delta\bar{m}\mathbf{w}$. Things are more complicated. The formula now becomes
$\sum_{ij} F_{ij}\varepsilon v_j^{(n)}\mathbf{e}_i - \sum_{ij} F_{ij}\varepsilon v_j^{(n)}\mathbf{e}_j + \sum_{ij} F_{ij}\delta m_j \mathbf{e}_j - \delta\bar{m}\mathbf{w}$,
which simplifies to
$\sum_{ij} F_{ij}\varepsilon v_j^{(n)}\mathbf{e}_i - \sum_j \tilde{p}_j\varepsilon v_j^{(n)}\mathbf{e}_j + \sum_j \tilde{p}_j \delta m_j \mathbf{e}_j - \delta\bar{m}\mathbf{w}$.

What we did in section 2 was obtain formula (18) for the k'th entry in our vector. Let's do that now. Let's decide what the k'th entry in our vector is. We multiply the last formula on the right by $\mathbf{e}_k^\top$.
$\sum_j F_{kj}\varepsilon v_j^{(n)} - \tilde{p}_k\varepsilon v_k^{(n)} + \tilde{p}_k\delta m_k - \delta\bar{m}w_k$.

To simplify this formula, it will be convenient to define the vector $\hat{\mathbf{a}}$ by

$$\hat{\mathbf{a}} = \mathbf{m}^\top - \bar{m}D^{-1}\mathbf{w}^\top \quad . \tag{83}$$

Then $\hat{a}_k = m_k - \bar{m}\,\tilde{p}_k^{-1} w_k$ and
$\tilde{p}_k\hat{a}_k = \tilde{p}_k m_k - \bar{m}w_k$.
Using this, we see that our formula can be written
$\sum_j F_{kj}\varepsilon v_j^{(n)} - \tilde{p}_k\varepsilon v_k^{(n)} + \delta\tilde{p}_k\hat{a}_k$.
Since $\sum_j F_{kj} = \tilde{p}_k$, our formula can be written

$$\sum_j F_{kj}(\varepsilon v_j^{(n)} - \varepsilon v_k^{(n)} + \delta\hat{a}_k) \tag{84}$$

Now that is the formula for the change in the cash balance of state k in the new bucket brigade. The formula for the old bucket brigade is (18).

Now consider a different Markov chain, which I shall call the *strange chain.* Its state diagram is the same as that of our original chain, and the transition probabilities P are the same, but the payoffs on the states are different. Instead of $\mathbf{m}$, the payoff vector is $\hat{\mathbf{a}}$. We see that in the strange chain, the average payoff per time unit is
$\tilde{\mathbf{p}}\hat{\mathbf{a}}^\top = \tilde{\mathbf{p}}\mathbf{m}^\top - \bar{m}\tilde{\mathbf{p}}D^{-1}\mathbf{w}^\top = \bar{m} - \bar{m}\mathbf{e}\mathbf{w}^\top = \bar{m} - \bar{m} = 0$.
Therefore, the excess payoff vector in the strange chain is also $\hat{\mathbf{a}}$.
We define the vector $\hat{\mathbf{c}}$ of payoff values in the strange chain just as in the original chain. That is

$$\hat{\mathbf{c}} = \hat{P}^{-1}\hat{\mathbf{a}}^\top \quad . \tag{85}$$

(see (6))
And just as in the original chain we have $\tilde{\mathbf{p}}\hat{\mathbf{c}}^\top = 0$.

Now suppose we run our old bucket brigade on the strange chain. Then everything is just as in section 2 except that the excess payoffs are everywhere $\hat{\mathbf{a}}$ instead of $\mathbf{a}$. We are however using excess payoffs and not payoffs, so the entire argument in the section holds. The deterministic bucket brigade converges to an equilibrium cash balance vector, which we can call $\hat{\mathbf{u}}$ to distinguish it from the equilibrium cash balance vector $\mathbf{u}$ of the original chain.

The convergence is just as in the original chain except that the convergence is to $\hat{\mathbf{u}}$ rather than $\mathbf{u}$. Equation (27) becomes

$$(\mathbf{v}^{(n)})^\top - \hat{\mathbf{u}}^\top = Q^n(\mathbf{v}^\top - \hat{\mathbf{u}}^\top) \quad . \tag{86}$$

Just as we used equation (23) to derive $\mathbf{c}$ from $\mathbf{u}$ in the original chain,[12] we can use the analogous formula to calculate $\hat{\mathbf{c}}$ from $\hat{\mathbf{u}}$ in the strange chain.

$$\delta\hat{\mathbf{c}}^\top = \varepsilon\,(\hat{\mathbf{u}}^\top - (\tilde{\mathbf{p}}\hat{\mathbf{u}}^\top)\,\mathbf{e}^\top) \tag{87}$$

Equation (26) holds for the original chain. The analogous formula for the strange chain is

$$\hat{\mathbf{u}}^\top = \tfrac{\delta}{\varepsilon}\,(I-M)\,\hat{\mathbf{c}}^\top + \tfrac{\tau}{N}\,\mathbf{e}^\top \quad . \tag{88}$$

(Remember, τ is the total amount of cash, which is constant.)

Let's look at the change in the cash balance of state k in one step of the deterministic bucket brigade. Remember that for the new bucket brigade on the original chain that change was formula (84). Of course for the old bucket brigade on the original chain it was (18). The analogous formula for the old bucket brigade on the strange chain we obtain by changing $\mathbf{a}$ to $\hat{\mathbf{a}}$. This gives us formula (84).

So we see that the change in the cash balance of state k in one step of the deterministic bucket brigade is (84) both in the new bucket brigade on the original chain and in the old bucket brigade on the strange chain. Those two bucket brigades have the same formula for cash change. The two deterministic bucket brigades are exactly the same, with exactly the same cash changes. Of course the corresponding probabilistic bucket brigades are very different. Cash converges to $\hat{\mathbf{u}}$ in the second deterministic bucket brigade, and hence it does so in the first.

So now we are in a position to answer question 1. If we have the deterministic bucket brigade running on the original chain and we make the change from $\mathbf{a}$ to $\mathbf{m}$, the new bucket brigade does indeed converge to some vector, namely $\hat{\mathbf{u}}$. And equation (86) gives its rate.

We can then go ahead and use equation (87) to obtain $\hat{\mathbf{c}}$.

So let me say it again. On the original chain, we can run the deterministic bucket brigade using $\mathbf{m}$ to obtain $\hat{\mathbf{u}}$, and from that we can calculate $\hat{\mathbf{c}}$.

But of course what we want is $\mathbf{c}$. The hope is that the change from using $\mathbf{a}$ to using $\mathbf{m}$ has changed nothing important. That is, the hope is that $\hat{\mathbf{u}}$ is the same as $\mathbf{u}$, or at least that $\hat{\mathbf{c}}$ is the same as $\mathbf{c}$. Unfortunately this turns out not to be the case in general. Let's look at the differences $\hat{\mathbf{c}}-\mathbf{c}$ and $\hat{\mathbf{u}}-\mathbf{u}$.

By definition,
$\hat{\mathbf{a}}^\top = \mathbf{m}^\top - \bar{m}D^{-1}\mathbf{w}^\top$.
Subtracting $\mathbf{a}^\top$ from both sides and using $\mathbf{m}^\top - \mathbf{a}^\top = \bar{m}\mathbf{e}^\top$ gives us
$\hat{\mathbf{a}}^\top - \mathbf{a}^\top = \bar{m}\mathbf{e}^\top - \bar{m}D^{-1}\mathbf{w}^\top$.
$\hat{\mathbf{a}}^\top - \mathbf{a}^\top = \bar{m}D^{-1}(\tilde{\mathbf{p}}-\mathbf{w})^\top$
Since $\mathbf{c}^\top = \hat{P}^{-1}\mathbf{a}^\top$ and $\hat{\mathbf{c}}^\top = \hat{P}^{-1}\hat{\mathbf{a}}^\top$, we have
$\hat{\mathbf{c}}^\top - \mathbf{c}^\top = \hat{P}^{-1}(\hat{\mathbf{a}}^\top - \mathbf{a}^\top)$

$$\hat{\mathbf{c}}^\top - \mathbf{c}^\top = \bar{m}\hat{P}^{-1}D^{-1}(\tilde{\mathbf{p}}-\mathbf{w})^\top \tag{89}$$

Now suppose $\bar{m}\neq 0$. Then $\hat{\mathbf{c}}^\top - \mathbf{c}^\top$ is zero if and only if $\hat{P}^{-1}D^{-1}(\tilde{\mathbf{p}}-\mathbf{w})^\top$ is zero. And since $\hat{P}^{-1}D^{-1}$ is a nonsingular matrix, that occurrs if and only if $(\tilde{\mathbf{p}}-\mathbf{w})^\top$ is zero. So we see that $\hat{\mathbf{c}}^\top = \mathbf{c}^\top$ if and only if either $\bar{m}=0$ or $\mathbf{w}=\tilde{\mathbf{p}}$.

Now suppose we subtract (26) from (88).

$$\hat{\mathbf{u}}^\top - \mathbf{u}^\top = \tfrac{\delta}{\varepsilon}\,(I-M)\,(\hat{\mathbf{c}}^\top - \mathbf{c}^\top) \tag{90}$$

The matrix M is trivially row stochastic and strongly connected, so the Frobenius Facts say that the set of right eigenvectors of eigenvalue 1 form a one dimensional subspace. This is the kernel of the transformation

[12] Remember that the $\bar{u}$ in (23) is $\tilde{\mathbf{p}}\mathbf{u}^\top$.

$I - M$, so that kernel is a one dimensional subspace. The vector $\mathbf{e}^\top$ is clearly in the kernel, so for any vector $\mathbf{y}$ we see that if $(I - M)\mathbf{y}^\top = 0$ then $\mathbf{y} = \chi \mathbf{e}^\top$ for some scalar χ .

Suppose the vector $\mathbf{y}$ is $\hat{\mathbf{c}}^\top - \mathbf{c}^\top$. Then equation (90) tells us that if $\hat{\mathbf{u}}^\top - \mathbf{u}^\top = 0$ then $\hat{\mathbf{c}}^\top - \mathbf{c}^\top = \chi \mathbf{e}^\top$
for some scalar χ . In that case we can multiply the last equation by $\tilde{\mathbf{p}}$ on the left and obtain $0 = \chi$, so $\hat{\mathbf{c}}^\top = \mathbf{c}^\top$. Thus we see that if $\hat{\mathbf{u}} = \mathbf{u}$ then $\hat{\mathbf{c}} = \mathbf{c}$.

Of course by (90) the implication also goes the other way. In conclusion,
$\hat{\mathbf{u}} = \mathbf{u}$ if and only if $\hat{\mathbf{c}} = \mathbf{c}$.

So now we have three equivalent conditions.

(1) $\hat{\mathbf{u}}^\top = \mathbf{u}^\top$
(2) $\hat{\mathbf{c}}^\top = \mathbf{c}^\top$
(3) either $\bar{m} = 0$ or $\mathbf{w} = \tilde{\mathbf{p}}$

Either all three conditions hold or none of them do.

So if we really want to use $\mathbf{m}$ instead of $\mathbf{a}$ in the deterministic bucket brigade, we can use it to calculate $\hat{\mathbf{c}}$, but this will not be what we want (we want $\mathbf{c}$) unless $\mathbf{w} = \tilde{\mathbf{p}}$, that is, unless the $\mathbf{w}$ we use in component (3) of the Cash Change Description is $\tilde{\mathbf{p}}$.

This is a shame. The probabilities $\tilde{\mathbf{p}}$ are at least as hard for the system to estimate as is $\bar{m}$. I was hoping the system could use a simple $\mathbf{w}$ like $\mathbf{w} = \frac{1}{N}\mathbf{e}$, but it appears this causes problems. The problem is that $\hat{\mathbf{c}}$ reflects not only where cash enters the system in component (2), but also where it leaves in component (3), unless $\mathbf{w} = \tilde{\mathbf{p}}$.

So let's summarize the answer to Question 1. The old bucket brigade uses $\mathbf{a}$ and the new bucket brigade uses $\mathbf{m}$. The deterministic old bucket brigade converges to $\mathbf{u}$; the deterministic new bucket brigade converges to $\hat{\mathbf{u}}$. The respective convergence speeds are given by equations (27) and (86). The values of $\mathbf{u}$ and $\mathbf{c}$ are related by equations (23) and (26). Similarly, the values of $\hat{\mathbf{u}}$ and $\hat{\mathbf{c}}$ are related by equations (87) and (88). The difference between $\hat{\mathbf{c}}$ and $\mathbf{c}$ is given by equation (89). The difference between $\hat{\mathbf{u}}$ and $\mathbf{u}$ is given by equation (90). In these two equations, the $\mathbf{w}$ is the $\mathbf{w}$ used in the new bucket brigade. The vectors $\mathbf{u}$ and $\mathbf{c}$ are independent of $\mathbf{w}$, as is almost everything in the deterministic old bucket brigade.

Now let's look at **Question 2**. We know the deterministic new bucket brigade using $\mathbf{m}$ instead of $\mathbf{a}$ does converge to $\hat{\mathbf{u}}$. Does the probabilistic new bucket brigade converge ot $\hat{\mathbf{u}}$ also? As it stands, the proof in section 3 doesn't tell us.

But we can modify that proof. We simply replace $\mathbf{a}$ with $\mathbf{m}$ throughout and replace $\mathbf{u}$ with $\hat{\mathbf{u}}$ throughout. The first place we do this is in the definition of θ , where we replace $\mathbf{a}$ with $\mathbf{m}$. So in most cases θ is larger, Δ is smaller, $\hat{\ell}$ is larger, and the step sizes are smaller. Subsections 3.2 and 3.3 are unaffected by the change.

No further changes are needed until the last paragraph of subsection 3.4. This is the first reference in the proof to the deterministic bucket brigade, which now needs to be the deterministic new bucket brigade. The vectors $\mathbf{v}^{(n)}$ are cash balance vectors in that bucket brigade. So here of course, $\mathbf{u}$ is replaced in the discussion by $\hat{\mathbf{u}}$, and the appeal to equation (27) needs to be replaced by an appeal to equation (86). So in inequality (53) at the end of that paragraph, the $\mathbf{u}$ is replaced by $\hat{\mathbf{u}}$ as it is everywhere in the proof.

After the definition of θ , the vector $\mathbf{a}$ does not appear explicitly in the argument until subsection 3.6. In this section we replace $\mathbf{a}$ with $\mathbf{m}$ throughout. This changes the definitions of η , β , and S_n .
$\eta(\mathbf{x}, i, j) = x_i(\mathbf{e}_i - \mathbf{e}_j) + m_j(\mathbf{e}_j - \mathbf{w})$
$\beta = e^\rho(\|\mathbf{v}\| + \|\mathbf{m}\|)$
S_n is the set of cash balance vectors whose norm is less than or equal to
$(1 + \frac{\varepsilon}{k})^{-n}\beta - \|\mathbf{m}\|$.
Nevertheless, the entire argument in the section works fine with $\mathbf{m}$ in place of $\mathbf{a}$. This is because nowhere in the section did we use $\tilde{\mathbf{p}}\mathbf{a}^\top = 0$.

So the Norm Bound Fact (lemma 18) holds with $\mathbf{m}$ in place of $\mathbf{a}$. Every cash balance vector in the Era has norm less than or equal to
$\beta - \|\mathbf{m}\|$,
where β is now defined as
$\beta = e^\rho(\|\mathbf{v}\| + \|\mathbf{m}\|)$.

Subsection 3.7 works too with $\mathbf{m}$ in place of $\mathbf{a}$, since here too we nowhere use $\tilde{\mathbf{p}}\mathbf{a}^\top = 0$. The definition of η here is of course the new definition.

In subsection 3.8 we refer back to the deterministic bucket brigade and section 2. We justify formula (65) by using the definition of η to produce a formula identical to (15) in section 2, which gives the change in the cash balance vector in the deterministic bucket brigade. But now we want to use the new definition of η . This changes the a_j to m_j in the formula we produce.

And this is good because it now is the same as formula (15) for the deterministic new bucket brigade that uses $\mathbf{m}$ instead of $\mathbf{a}$. So the whole argument in subsection 3.8 works with $\mathbf{m}$ in place of $\mathbf{a}$.

The vector $\mathbf{a}$ is not referred to explicitly in the remaining subsections, but of course the various quantities are now different because $\mathbf{a}$ has been replaced by $\mathbf{m}$ in the definitions.

The vector $\mathbf{u}$ however does occur explicitly. Of course we replace it with $\hat{\mathbf{u}}$ throughout. This replacement occurs in the definition of ν in subsection 3.9 and in the definition of $\vartheta(\mathbf{x})$ in subsection 3.10.

The first reference to $\mathbf{u}$ in subsection 3.9 is an appeal to inequality (53) in deriving inequality (69). The $\mathbf{u}$ occurs both in (53) and (69) and in both it is replaced by $\hat{\mathbf{u}}$ so everything works.

Nowhere in the entire proof in section 3 did we use $\tilde{\mathbf{p}}\mathbf{a}^\top = 0$, so the entire argument works with $\mathbf{m}$ in place of $\mathbf{a}$ and $\hat{\mathbf{u}}$ in place of $\mathbf{u}$. The result is theorem 1 with $\hat{\mathbf{u}}$ in place of $\mathbf{u}$.

Now suppose we define $\boldsymbol{\gamma}$ as in equation (82). The entire argument in subsection 4.1 holds with $\hat{\mathbf{u}}$ in place of $\mathbf{u}$ and $\hat{\mathbf{c}}$ in place of $\mathbf{c}$, and with the appeal to equation (23) replaced by an appeal to equation (87). So corollary 1 holds with $\hat{\mathbf{c}}$ in place of $\mathbf{c}$. This means that $\boldsymbol{\gamma}$ is a good estimate of $\hat{\mathbf{c}}$ (but not of $\mathbf{c}$).

So we see that theorem 1 and corollary 1 hold for a bucket brigade that uses $\mathbf{m}$ instead of $\mathbf{a}$ in components (2) and (3) of the Cash Change Description.

So where does this leave us? We can go ahead and use $\mathbf{m}$ instead of $\mathbf{a}$ in components (2) and (3) of the Cash Change Description and the probabilistic bucket brigade will converge nicely to the equilibrium cash balance vector $\hat{\mathbf{u}}$. But this may not be the vector $\mathbf{u}$ that we want. Using equations (90) and (89), we see that the error in our cash balance vector is this.

$$\hat{\mathbf{u}}^\top - \mathbf{u}^\top = \bar{m}\tfrac{\delta}{\varepsilon}(I - M)\hat{P}^{-1}D^{-1}(\tilde{\mathbf{p}} - \mathbf{w})^\top \tag{91}$$

In order to get rid of the error, we need either to set $\bar{m}$ to 0, which amounts to using $\mathbf{a}$ instead of $\mathbf{m}$, or set $\mathbf{w}$ to $\tilde{\mathbf{p}}$. Either of these is difficult. In the first case the system needs to estimate $\bar{m}$. It might do this using a running average, but any estimate is bound to be faulty. In the second case the system needs to estimate $\tilde{\mathbf{p}}$, which is if anything more problematic.

I have always assumed that we could go ahead and use $\mathbf{m}$ instead of $\mathbf{a}$, and simply use some simple method of stabilizing the total cash, in effect, some simple $\mathbf{w}$, maybe $\mathbf{w} = \frac{1}{N}\mathbf{e}$. I was wrong. I thought that whatever error we might be stuck with could be reduced by decreasing the step size. Again I was wrong, and this was the really big shock. The error $\hat{\mathbf{c}} - \mathbf{c}$ is independent of $\bar{\varepsilon}$. (see (89)) Using $\mathbf{m}$ instead of $\mathbf{a}$ introduces an error into the system of which I was unaware. This source of error needs to be faced and dealt with appropriately.

We have not discussed the impact of this error on adaptation. I think the error causes genuine problems in adaptation.

I think many bucket brigade implementations ignore this source of error, to their cost. But there are many other issues in bucket brigade implementations, so the problem may be masked or implicitly dealt with. Of course discounting (which we have not discussed) masks the problem and complicates it.

There is, however, some good news. Suppose the system uses a running average to estimate $\bar{m}$ and suppose, as will inevitably be the case, that the estimate is wrong. Let's call the estimate $\dot{m}$. The system subtracts that from each m_i to obtain what it thinks is a_i . But it's not. It's $m_i - \dot{m}$. The system thinks it's using $\mathbf{a}$, but it's really using a new payoff vector $\mathbf{m} - \dot{m}\mathbf{e}$. Using those payoffs, what is the average payoff per time unit? It's $\bar{m} - \dot{m}$. Now of course when the system uses these payoffs, the probabilistic bucket brigade converges to a vector $\hat{\mathbf{u}}$ that is not $\mathbf{u}$. The error is given by (91), but with $\bar{m}$ replaced by the new average payoff $\bar{m} - \dot{m}$.

So if $\dot{m}$ is close to $\bar{m}$ the error is small. The closer it is, the smaller the error. The system can use any $\mathbf{w}$ it likes (even $\mathbf{e}_1$) and if the estimate is close the error will be small. The probabilistic bucket brigade does not become unstable. the cash balances converge nicely to $\hat{\mathbf{u}}$, a vector close to what we want. And any sampling error noise in the system is covered by our discussion of question 2. Decreasing step size decreases noise, just as when we were using $\mathbf{a}$.

4.3 Can We Omit Component (3) ?

Let us now examine a probabilistic bucket brigade that uses $\mathbf{m}$ instead of $\mathbf{a}$ in components (2) and (3) of the Cash Change Description and in which $\mathbf{w} = \frac{1}{N}\mathbf{e}$. The chain runs through a sequence of states $i_0,\ i_1,\ i_2,\ i_3,\ i_4,\ \ldots.$.

At the same time, the probabilistic bucket brigade runs through a sequence of cash balance vectors
$\mathbf{x}^{(0)},\ \mathbf{x}^{(1)},\ \mathbf{x}^{(2)},\ \mathbf{x}^{(3)},\ \mathbf{x}^{(4)},\ ....$.
Each cash balance vector $\mathbf{x}$ gives us a vector $\boldsymbol{\gamma}$ according to equation (82), that is,
$\boldsymbol{\gamma} \ = \ \delta^{-1}\bar{\varepsilon}\,(\mathbf{x} - \bar{x}\,\mathbf{e})$
(where $\bar{x} \ = \ \mathbf{x}\tilde{\mathbf{p}}^\top$) .
So we get a sequence of vectors
$\boldsymbol{\gamma}^{(0)},\ \boldsymbol{\gamma}^{(1)},\ \boldsymbol{\gamma}^{(2)},\ \boldsymbol{\gamma}^{(3)},\ \boldsymbol{\gamma}^{(4)},\ ....$,
from the sequence of cash balance vectors. That is,

$$\boldsymbol{\gamma}^{(n)} \ = \ \delta^{-1}\bar{\varepsilon}\,(\mathbf{x}^{(n)} - (\mathbf{x}^{(n)}\tilde{\mathbf{p}}^\top)\,\mathbf{e}) \tag{92}$$

for each n .

This bucket brigade uses $\mathbf{m}$ instead of $\mathbf{a}$, so we can make the usual definitions of $\hat{\mathbf{a}}$, $\hat{\mathbf{c}}$, and $\hat{\mathbf{u}}$. ((83), (85), and (88)) In these definitions we will have $\mathbf{w} \ = \ \frac{1}{N}\mathbf{e}$, and of course τ is the (constant) total amount of cash.

We are interested in estimating $\hat{\mathbf{c}}$. We have seen that theorem 1 holds with $\hat{\mathbf{u}}$ in place of $\mathbf{u}$, so the cash balance vectors $\mathbf{x}^{(n)}$ are close to $\hat{\mathbf{u}}$. And corollary 1 holds with $\hat{\mathbf{c}}$ in place of $\mathbf{c}$, so the vectors $\boldsymbol{\gamma}^{(n)}$ are close to $\hat{\mathbf{c}}$. The point is of course that $\boldsymbol{\gamma}^{(n)}$ is a good estimate of $\hat{\mathbf{c}}$.

Now suppose we re-run the bucket brigade but this time without component (3) of the Cash Change Description. So now the total amount of cash is no longer a constant. Suppose the chain goes through exactly the same sequence of states. Again we get a sequence of cash balance vectors $\dot{\mathbf{x}}^{(n)}$ and a sequence of vectors $\dot{\boldsymbol{\gamma}}^{(n)}$. I've written a dot over the letter to indicate that component (3) is not being used.

How do the new vectors differ from the old ones. For example, how does $\dot{\mathbf{x}}^{(n)}$ differ from $\mathbf{x}^{(n)}$?

Well, for one thing, the total amount of cash $(\dot{\mathbf{x}}^{(n)}\mathbf{e}^\top)$ is now not constant. It's doing a random walk. The bias of the walk is $\bar{m}$. If $\bar{m}$ is 0 then the walk is unbiased, but still it's a walk and might go anywhere. There can be no convergence. Nothing like theorem 1 can hold.

We introduced the Cash Change Description in section 2. There we said that component (3) adds $-\delta a_j\mathbf{w}$ to the cash balance vector. But here the bucket brigade is using $\mathbf{m}$ instead of $\mathbf{a}$, and the $\mathbf{w}$ it is using is $\frac{1}{N}\,\mathbf{e}$. So here, component (3) is adding
$-\delta\, m_i\frac{1}{N}\,\mathbf{e}$
to the cash balance vector. The point I want to make is that what it is adding is a scalar multiple of $\mathbf{e}$.

Suppose $\dot{\mathbf{x}}^{(0)}$ is the same as $\mathbf{x}^{(0)}$. How does $\dot{\mathbf{x}}^{(n)}$ differ from $\mathbf{x}^{(n)}$? The difference is that in each of the n steps, component (3) has added something in the sequence leading to $\mathbf{x}^{(n)}$, whereas it hasn't in the sequence leading to $\dot{\mathbf{x}}^{(n)}$. Each of the things it added is a scalar multiple of $\mathbf{e}$. These additions are the only difference between $\mathbf{x}^{(n)}$ and $\dot{\mathbf{x}}^{(n)}$. Therefore,
the difference $\dot{\mathbf{x}}^{(n)} - \mathbf{x}^{(n)}$ is a scalar multiple of $\mathbf{e}$.

Now let's see how $\dot{\boldsymbol{\gamma}}^{(n)}$ differs from $\boldsymbol{\gamma}^{(n)}$.
By (92) we have
$\delta\,\bar{\varepsilon}^{-1}\boldsymbol{\gamma}^{(n)} \ = \ \mathbf{x}^{(n)} - (\mathbf{x}^{(n)}\tilde{\mathbf{p}}^\top)\,\mathbf{e}$.
The corresppponding equation for $\dot{\boldsymbol{\gamma}}^{(n)}$ is
$\delta\,\bar{\varepsilon}^{-1}\dot{\boldsymbol{\gamma}}^{(n)} \ = \ \dot{\mathbf{x}}^{(n)} - (\dot{\mathbf{x}}^{(n)}\tilde{\mathbf{p}}^\top)\,\mathbf{e}$.
Subtracting the previous equation from this gives
$\delta\,\bar{\varepsilon}^{-1}(\dot{\boldsymbol{\gamma}}^{(n)} - \boldsymbol{\gamma}^{(n)}) \ = \ (\dot{\mathbf{x}}^{(n)} - \mathbf{x}^{(n)}) - ((\dot{\mathbf{x}}^{(n)} - \mathbf{x}^{(n)})\,\tilde{\mathbf{p}}^\top)\,\mathbf{e}$.
Since $\dot{\mathbf{x}}^{(n)} - \mathbf{x}^{(n)}$ is a scalar multiple of $\mathbf{e}$, there is a scalar χ such that
$(\dot{\mathbf{x}}^{(n)} - \mathbf{x}^{(n)}) \ = \ (\chi\mathbf{e})$.
It follows that
$(\dot{\mathbf{x}}^{(n)} - \mathbf{x}^{(n)})\,\tilde{\mathbf{p}}^\top \ = \ \chi\mathbf{e}\tilde{\mathbf{p}}^\top \ = \ \chi$.
We substitute these two items into the previous equation.
$\delta\,\bar{\varepsilon}^{-1}(\dot{\boldsymbol{\gamma}}^{(n)} - \boldsymbol{\gamma}^{(n)}) \ = \ (\chi\mathbf{e}) - (\chi)\,\mathbf{e}$.
So we see that
$\dot{\boldsymbol{\gamma}}^{(n)} = \ \boldsymbol{\gamma}^{(n)}$.

Note that this equivalence holds only if $\mathbf{w} \ = \ \frac{1}{N}\,\mathbf{e}$.

This is nice. Corollary 1 told us that the vectors $\boldsymbol{\gamma}^{(n)}$ are good estimates of $\hat{\mathbf{c}}$. Now we see that we can obtain these vectors by running the bucket brigade without component (3).

In this sense, running the bucket brigade without component (3) is like running it with a component (3) that uses $\mathbf{w} \ = \ \frac{1}{N}\mathbf{e}$.

Many, perhaps most, bucket brigade implementations omit component (3). This is like running the bucket brigade with $\mathbf{w} \ = \ \frac{1}{N}\mathbf{e}$.

This all looks really nice. The bucket brigade without component (3) is very simple. It gives us a nice estimeate $\boldsymbol{\gamma}$ of $\hat{\mathbf{c}}$.

The bad news is that the estimate is of $\hat{\mathbf{c}}$, not of $\mathbf{c}$. Equation (89) gives us a formula for $\hat{\mathbf{c}}$. (Remember, $\mathbf{w} = \frac{1}{N}\mathbf{e}$.)

$$\hat{\mathbf{c}}^\top = \mathbf{c}^\top + \bar{m}\hat{P}^{-1}D^{-1}(\tilde{\mathbf{p}} - \mathbf{w})^\top \quad (93)$$

We see the problem. Estimating $\hat{\mathbf{c}}$ without using component (3) gives us $\mathbf{w} = \frac{1}{N}\mathbf{e}$. But for $\hat{\mathbf{c}}$ to be equal to $\mathbf{c}$ we want $\mathbf{w} = \tilde{\mathbf{p}}$, as formula (93) makes clear. We can't usually have it both ways. Omitting component (3), as many implementations do, leads to something that is not a good estimate of $\mathbf{c}$.

But it is theoretically possible to have it both ways. The idea is to set $\mathbf{w} = \frac{1}{N}\mathbf{e}$ and $\bar{m} = 0$. Then we can omit component (3) and equation (93) tells us that $\hat{\mathbf{c}} = \mathbf{c}$. What I am really saying is that we can have it both ways if we go back to using $\mathbf{a}$ instead of $\mathbf{m}$. So here is what we do.

We use the bucket brigade as described in the Cash Change Description in section 2 but without component (3). (It uses $\mathbf{a}$, not $\mathbf{m}$.) We derive $\boldsymbol{\gamma}$ from the current $\mathbf{x}$ by formula (82) and use it as an estimate of $\mathbf{c}$.

That works. Corollary 1 holds.

Of course the cash balances might wander together way up positive or way down negative, but this doesn't affect the estimate.

If they wander too far, we can subtract an appropriate scalar multiple of $\mathbf{e}$ from the current cash balance vector $\mathbf{x}$. That has no effect on $\boldsymbol{\gamma}$.

Of course this brings us back to the problem of estimating $\bar{m}$ so we can obtain the excess payoffs to use in component (2). But remember that if the estimate of $\bar{m}$ is wrong but close, then $\hat{\mathbf{c}}$ is close to $\mathbf{c}$, so $\boldsymbol{\gamma}$ is still a pretty good estimate of $\mathbf{c}$.

4.4 Intuitive Discussion of $\bar{m}$ Estimation

As we said, $\bar{m}$ can be estimated using a running average, but there are other ways. If our estimate $\dot{m}$ of $\bar{m}$ is too low, then the total amount of cash will gradually climb. (The total amount of cash does a random walk whose bias is positive.) If the estimate is too high then the total amount of cash will gradually fall. By monitoring the total amount of cash, the estimate $\dot{m}$ can be appropriately adjusted. We want an unbiased random walk.

In the Cash Change Description, we can think of component (2) as adding both m_j and $-\dot{m}$ to the cash of the current state j . If the total cash is rising, then the $-\dot{m}$ cash is too much and we need to reduce it. When we reduce it, we are not reducing cash inflow equally across states; we are reducing cash inflow more in states that are visited more often. This is correct.

But we said that if the total amount of cash is inconveniently high we can subtract a scalar multiple of $\mathbf{e}$ from the cash balance vector. This subtracts the same amount of cash from every state. This is also correct. But it looks like a contradiction. Let me give an intuitive explanation of why it's not.

One nice thing about the bucket brigade is that it is additive. Looking at component (2), suppose we color the added m_j cash blue and the added $-\dot{m}$ cash red. The bucket brigade moves both red and blue cash, but in an additive fashion. (Of course red cash can be negative and so can blue cash.) We can run the bucket brigade adding only blue cash and see what the cash balances are. We can run it (putting the chain through the same state sequence) adding only red cash and see what the cash balances are. And we can run the bucket brigade adding both blue and red cash. The blue cash and red cash each travel without regard to the other, just as when the other wasn't present. And a state's cash balance is the sum of its blue cash and red cash.

This is general. If we have two payoff vectors, $\mathbf{m}$ and $\dot{\mathbf{m}}$, and we run the bucket brigade with each, obtaining two different cash balance vectors,[13] then the sum of the vectors is what we get if the payoff vector is $\mathbf{m} + \dot{\mathbf{m}}$. And the state values also add. Using the dot to indicate quantities when payoff is $\dot{\mathbf{m}}$, we have $(\mathbf{c} + \dot{\mathbf{c}}) = \hat{P}^{-1}(\mathbf{a} + \dot{\mathbf{a}})$. Both cash balances and values are additive, which they must be if cash balances give us values the way they do.

So we see that the real issue is what is happening to the red cash. The red cash component (2) input to every state is $-\dot{m}$. It's just as if every state had the same payoff, and hence the same value. So clearly the cash balances of the different states should be the same. Okay, not really the same because everything is probabilistic, but there should be no tendency for one state to have more cash than another.

But there is. Suppose $-\dot{m}$ is positive, so the total amount of cash is climbing. There is more cash input on states visited more often. These states have more cash. Of course the bucket brigade distributes

[13] assuming the sequence of states is the same in the two cases

the cash, so the state cash balances climb together, but the more prevalent states are usually slightly in the lead.

So a positive $-\dot{m}$ has two effects. The big effect is that the cash balances are all big and climbing together. The small effect is that there is a vector of cash balance differences across the states, fluctuating because of probabilistic effects, but averaging out to a nonzero vector.

It is only the small effect that is relevant to the derived state value estimates. It needs to match and cancel the similar effect in the blue cash. It does this when $-\dot{m}$ is adjusted to make it equal to $-\bar{m}$. We know the adjsutment is correct when the big effect climbing stops.

There is no contradiction.

So we have a bucket brigade that works. We monitor total cash to obtain an estimate $-\dot{m}$ of $-\bar{m}$. We use $m_j - \dot{m}$ for a_j in component (2) of the Cash Change Description. We omit component (3). We use component (1) as given.

5 So What Have We learned?

But it is not the purpose of this essay to make recommendations. The purpose is to raise questions and to provide formal answers where we can.

The message is not that this or that solution is the correct one. The message is that we should be aware of each possible source of error and bear it in mind in our design or our analysis. I confess again that the errors discussed here were a surprise to me, even after decades of thinking about the bucket brigade.

We have just examined some of the sources of error in the bucket brigade. But this is really a small part of the story.

In practice the bucket brigade contains many sources of error. The whole idea of the bucket brigade is to use the estimate of $\mathbf{c}$ in the adaptation process.[14] Adaptation changes P . But if P is changing, that introduces error into the estimate of $\mathbf{c}$. We do not discuss this problem or any other aspects of adaptation in this essay.

And of course it is not in Markov chains where the bucket brigade is usually used. It is used in rule based systems, where the special propeties of the bucket brigade are manifest. But rule based systems introduce a whole family of interesting errors. This essay does not discuss rule based systems.

In this larger context, the error discussed here begins to look overshadowed. I found the error surprising, but it looks managable. There may be a way to get around it entirely, but if there is I have not found it.

To me, the really interesting question is whether there is an analog of this error in the mathematics of evolving populations. But that question can be addressed only in the context of adaptation, which we are not discussing here.[15]

All these matters are more interesting than what is discussed in this essay. Discussion of them assumes that bucket brigade cash balances can be a reasonable representation of values. But can they? Only if they converge in some sense to something related to values.

This essay shows that they do in the simple fundamental case of a finite state Markov chain. It thus puts beyond doubt the convergence question in this case.[16] This provides a foundation for a discussion of the more interesting questions.

Appendix

A Comment on the Exponential as a Limit

We know that for real x and integer variable k we have
$\lim_{k\to\infty} (1 + \frac{x}{k})^k = e^x$,
but is the limit approached from above or below or what? To answer this question we pretend k is a real variable and take some derivatives with respect to k .

[14]The use can be implicit.

[15]Learning systems that model nets of neurons must prevent synapse strengths from increasing indefinitely. Presumably there is an analog of our error in these systems, but given the current state of our knowledge, discussion of this would probably be premature.

[16]unless of course I have made an error somewhere

The variable k will vary over all numbers greater than max(0, -x) .
So we always have $k > 0$ and $k > -x$.
Since $k > -x$ we have $k + x > 0$.
Dividing by our positive k gives $1 + \frac{x}{k} > 0$, so we always have $1 + \frac{x}{k} > 0$, and $\log(1 + \frac{x}{k})$ is a real number.

$$\left(1 + \frac{x}{k}\right)' = -\frac{x}{k^2}$$

$$\left(\frac{x}{k+x}\right)' = -\frac{x}{(k+x)^2}$$

$$\left(\log\left(1 + \frac{x}{k}\right)\right)' = \frac{1}{1+\frac{x}{k}}\left(-\frac{x}{k^2}\right) = \frac{1}{k+x}\left(-\frac{x}{k}\right) = -\frac{x}{k(k+x)}$$

We define the real variables z and y , both functions of k .

$$z = k \log\left(1 + \frac{x}{k}\right) \quad .$$

$$y = \log\left(1 + \frac{x}{k}\right) - \frac{x}{k+x} \quad .$$

$$y' = -\frac{x}{k(k+x)} + \frac{x}{(k+x)^2} = \frac{-x^2}{k(k+x)^2}$$

Now we see that y must be positive for all k greater than max(0, -x) , since y' is negative and $\lim_{k\to\infty} y = 0$.

Now z is a real variable, so e^z is positive. Since y and e^z are both positive,

$$e^z y > 0$$

for all k greater than max(0, -x) .

$$z' = \log\left(1 + \frac{x}{k}\right) + k\left(-\frac{x}{k(k+x)}\right) = \log\left(1 + \frac{x}{k}\right) - \frac{x}{k+x} = y$$

$$(e^z)' = e^z z' = e^z y > 0$$

So e^z is an increasing function of real k for all k greater than max(0, -x) .
Now let's look just at when k is an integer. Then $e^z = \left(1 + \frac{x}{k}\right)^k$.
So $(1 + \frac{x}{k})^k$ is an increasing function of integer k for all k greater than max(0, -x) .

Thus if $k > 0$ and $k > -x$, we have

$$\left(1 + \frac{x}{k}\right)^k < e^x \quad . \tag{94}$$

B Matrix Prerequsites

Lemma 33 *If λ is a simple zero (multiplicity 1) of the characteristic polynomial of a square complex matrix, then the set of eigenvectors of eigenvalue λ is a one dimensional subspace of the vector space.*[17]

Proof:
Suppose B is a complex square matrix and λ is a simple zero of its characteristic polynomial. Suppose we convert B into Jordan normal form. Then the characteristic polynomial is unchanged. The eigenvectors are also unchanged, except that now they are written using the Jordan basis. So we see that it is sufficient to prove the lemma for the case when B is in Jordan normal form.

[17] This lemma is used in the proof of lemma 38.

So suppose B is in Jordan normal form, and suppose its diagonal entries are $w_1, w_2, w_3, ..., w_N$. Then the characteristic polynomial is

$$\prod_{i=1}^{N}(x - w_i) \quad .$$

so if λ is a simple zero then exactly one of the diagonal entries equals λ .

Let's suppose for example that in the upper left hand corner of B is a 4×4 Jordan block whose diagonal elements are all ω . Of course $\omega \neq \lambda$. Suppose $\mathbf{v}$ and $\mathbf{u}$ are vectors such that $\mathbf{u} = \mathbf{v}B$. Then

$u_1 = \omega v_1$,
$u_2 = \omega v_2 + v_1$,
$u_3 = \omega v_3 + v_2$, and
$u_4 = \omega v_4 + v_3$.

Now suppose $\mathbf{v}$ is an eigenvector of eigenvalue λ . Then $u_i = \lambda v_i$ for all i . So we have

$\lambda v_1 = \omega v_1$,
$\lambda v_2 = \omega v_2 + v_1$,
$\lambda v_3 = \omega v_3 + v_2$, and
$\lambda v_4 = \omega v_4 + v_3$.

So we see that $v_1 = 0$. Consequently $v_2 = 0$, and $v_3 = 0$, and $v_4 = 0$.

What this shows us is that the only non-zero entry in the eigenvector $\mathbf{v}$ is the entry corresponding to the single λ on the diagonal of B . Any eigenvector of eigenvalue λ has just one non-zero entry, the entry corresponding to the λ on the diagonal.

■

Lemma 34

Let x be any positive real number.
Let B be any complex square matrix for which $\|B\| \leq 1$.
Let n be an integer variable.
As $n \to \infty$, the quantity $\|e^{tB} - (I + \frac{t}{n}B)^n\|$ goes to zero uniformly for all t in the interval $[0, x]$.

Proof:

We transform part of the expression in the lemma as follows.
(The part in square brackets is the product of k factors.)

$$\begin{aligned}
(I + \tfrac{t}{n}B)^n &= \sum_{k=0}^{n} \binom{n}{k} \frac{t^k}{n^k} B^k \\
&= I + \sum_{k=1}^{n} \binom{n}{k} \frac{t^k}{n^k} B^k \\
&= I + \sum_{k=1}^{n} \frac{1}{k!} \left[(n)(n-1)(n-2)(n-3)\cdots(n-k+1)\right] \frac{t^k}{n^k} B^k \\
&= I + \sum_{k=1}^{n} \frac{t^k}{k!} B^k \left[\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\left(\frac{n-3}{n}\right)\cdots\left(\frac{n-k+1}{n}\right)\right] \\
&= I + \sum_{k=1}^{n} \frac{t^k}{k!} B^k \left[\left(1\right)\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\left(1 - \frac{3}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right)\right]
\end{aligned}$$

By the definition of the exponential. we have

$$e^{tB} = \sum_{k=0}^{\infty} \frac{t^k}{k!} B^k = I + \sum_{k=1}^{\infty} \frac{t^k}{k!} B^k \quad .$$

Therefore, the difference is the following

$$\begin{aligned}
&e^{tB} - (I + \tfrac{t}{n}B)^n \\
&= \left(\sum_{k=1}^{n} \frac{t^k}{k!} B^k \left[1 - \left(1\right)\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\left(1 - \frac{3}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right)\right]\right) + \sum_{k=n+1}^{\infty} \frac{t^k}{k!} B^k
\end{aligned}$$

Notice that the value of the part is square brackets is always in the interval $[0,1]$, whatever the value of n .

Since $\|B\| \leq 1$, equation (11) gives us $\|B^k\| \leq 1$. So we obtain the following.

$$\|e^{tB} - (I + \tfrac{t}{n}B)^n\| \leq \left(\sum_{k=1}^{n} \tfrac{t^k}{k!} \left[1 - \left(1\right)\left(1 - \tfrac{1}{n}\right)\left(1 - \tfrac{2}{n}\right)\left(1 - \tfrac{3}{n}\right)\cdots\left(1 - \tfrac{k-1}{n}\right)\right]\right) + \sum_{k=n+1}^{\infty} \tfrac{t^k}{k!}$$

Now we see that for any t in the interval $[0,x]$ we have the following.

$$\|e^{tB} - (I + \tfrac{t}{n}B)^n\| \leq \left(\sum_{k=1}^{n} \tfrac{x^k}{k!} \left[1 - \left(1\right)\left(1 - \tfrac{1}{n}\right)\left(1 - \tfrac{2}{n}\right)\left(1 - \tfrac{3}{n}\right)\cdots\left(1 - \tfrac{k-1}{n}\right)\right]\right) + \sum_{k=n+1}^{\infty} \tfrac{x^k}{k!}$$

The series on the right side of the inequality is written in two parts, which I shall call the first part and the second part.

We are going to increase n and watch what happens to the individual terms. Let's watch the 25'th term, the term where $k = 25$. Let's call it our term. As long as n is less than 25, our term is in the second part. But once n passes 25, our term is in the first part, and its value keeps changing as n increases.

The bit of our term in square brackets is always 1 minus the product of 25 factors, since for our term $k = 25$. But notice that in our term, the value of the bit in square brackets is always less than or equal to 1, so the value of our term is always less than or equal to $\frac{x^k}{k!}$, where in this case k of course means 25.

Notice further that as $n \to \infty$, the bit of our term in square brackets goes to zero. So the value of our term goes to zero.

So we see in general that as $n \to \infty$, each term goes to zero, each term remains non-negative and bounded above by a bound independent of n , and the sum of the term bounds is finite.[18] Therefore, the whole sum goes to zero as $n \to \infty$.

So if we look at the last inequality, the right side goes to zero as $n \to \infty$. The variable t does not occur on the right side, so the convergence of the right side has nothing to do with t . Hence the left side converges to zero uniformly for all t in the interval $[0,x]$.

■

Now we define the matrix

$$A \;=\; D(I{-}P) \tag{12}$$

Now $\|D\| \;\leq\; 1$ and $\|-DP\| \;=\; \|DP\| \;\leq\; 1$. Therefore we have the following.
$\|A\| \;=\; \|D(I-P)\| \;=\; \|D - DP\| \;\leq\; \|D\| + \|-DP\| \;\leq\; 2$

$$\left\|\tfrac{1}{2}A\right\| \;\leq\; 1 \quad . \tag{95}$$

In the following discussion, I am going to use the letter α for a real number in the range $0 < \alpha \leq 1$.

We use lemma 34 in the trivial special case when $x = 2$, $t = 2$, and $B \;=\; -\frac{1}{2}\alpha A$.
So now we have $\|B\| \;=\; \|-\frac{1}{2}\alpha A\| \;=\; |-\alpha|\,\|\frac{1}{2}A\| \;\leq\; \|\frac{1}{2}A\| \;\leq\; 1$, so the lemma condition is satisfied.
Then $tB \;=\; -\alpha A$, and the lemma gives us this.
$\lim_{n\to\infty} \| e^{-\alpha A} - (I - \frac{1}{n}\alpha A)^n \| \;=\; 0$.
This gives us the following.

$$\lim_{n\to\infty} (I - \tfrac{1}{n}\alpha A)^n \;=\; e^{-\alpha A} \tag{96}$$

Lemma 35 *If* $0 < t \leq 1$ *then the matrix* $I - tA$ *is bistochastic, strongly connected, and primitive.*

Proof:
Every entry in $\tilde{\mathbf{p}}$ is positive, there is more than one entry, and the entries sum to 1. Therefore, every entry in $\tilde{\mathbf{p}}$ is strictly between 0 and 1, and the same is true for the diagonal entries in D . Suppose $0 < t \leq 1$. Then $I - tD$ is a diagonal matrix with all its diagonal entries positive. Then if we define $B = I - tD + tDP$, we see that B is a non-negative matrix with all its diagonal entries positive. We note that $\mathbf{e}B = \mathbf{e}$ and $B\mathbf{e}^\top = \mathbf{e}^\top$, so B is bistochastic. The matrix B has positive entries everywhere P does, so B is strongly connected (irreducible). Since the diagonal entries of B are positive, lemma 5

[18] The sum of bounds is of course e^x .

tells us that B is primitive. And $B = I - tA$.

■

Now if n is a positive integer then lemma 35 tells us that $I - A$ and $I - \frac{1}{n}\alpha A$ are both bistochastic, strongly connected, and primitive.

Lemma 36
If $0 \leq \alpha \leq 1$ *then* $e^{-\alpha A}$ *is bistochastic.*

Proof:
If $\alpha = 0$ then by the power series definition, $e^{-\alpha A} = I$, which is bistochastic. If $0 < \alpha \leq 1$, then lemma 35 tells us that for any positive integer n , the matrix $I - \frac{1}{n}\alpha A$ is bistochastic. Hence so is $(I - \frac{1}{n}\alpha A)^n$. Now use (96).

■

Since $I - A$ is primitive, it is irreducible, so by the Frobenius facts it has an eigenvalue 1 of multiplicity one. By lemma 3 applied to $I - A$, every other eigenvalue of $I - A$ is in the open unit disk.

Suppose B is a square matrix. We want to compare the eigenvalues of matrices B and e^B .
Consider the definition
$e^B = \sum_{n=0}^{\infty} \frac{1}{n!} B^n$.
If we write the matrices using a Jordan basis for B , then each B^n is upper triangular and so is e^B . By looking at the diagonal elements we can set up a one to one correspondence between the eigenvalues of B and the eigenvalues of e^B . To each eigenvalue λ of B there corresponds the eigenvalue e^λ of e^B .

Now we see that to each eigenvalue λ of $I - A$ there corresponds the eigenvalue $\lambda - 1$ of $-A$ and the eigenvalue $e^{\lambda-1}$ of e^{-A} .

First suppose the eigenvalue λ of $I - A$ is 1 . Then $e^{\lambda-1} = 1$. This is the corresponding eigenvalue of e^{-A} .

Now suppose the eigenvalue λ of $I - A$ is one of those in the open unit disk. Then $\lambda - 1$ is in the negative half plane, and $e^{\lambda-1}$ is in the open unit disk. This last is the corresponding eigenvalue of e^{-A} .

Exactly one of the eigenvalues of $I - A$ is 1 (multiplicity one) and the others are all in the open unit disk. So we see that exactly one of the eigenvalues of e^{-A} is 1 and its other eigenvalues are all in the open unit disk. The eigenvalue 1 has multiplicity one.

Let's define
$\hat{A} = e^{-A}$.
Lets review what we know about $\hat{A}$.
It's bistochastic,
exactly one of its eigenvalues is 1 (multiplicity one),
and the other eigenvalues are all in the open unit disk.

Lemma 37 *Every non-zero eigenvalue of* $\hat{A} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}$ *is also an eigenvalue of* $\hat{A}$ *.*

Proof:
Suppose λ is a non-zero eigenvalue of $\hat{A} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}$. Then there is a non-zero vector $\mathbf{w}$ such that
$\mathbf{w}(\hat{A} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}) = \lambda\mathbf{w}$.

$$\mathbf{w}\hat{A} - \frac{1}{N}\mathbf{w}\,\mathbf{e}^\top\mathbf{e} = \lambda\mathbf{w} \tag{97}$$

We multiply on the right by $\mathbf{e}^\top$ and use $\hat{A}\mathbf{e}^\top = \mathbf{e}^\top$.
$\mathbf{w}\mathbf{e}^\top - \mathbf{w}\mathbf{e}^\top = \lambda\mathbf{w}\mathbf{e}^\top$
$0 = \lambda(\mathbf{w}\mathbf{e}^\top)$
Since λ is non-zero, we have $\mathbf{w}\mathbf{e}^\top = 0$.
Thus equation (97) becomes
$\mathbf{w}\hat{A} = \lambda\mathbf{w}$,
and we see that λ is an eigenvalue of $\hat{A}$.

■

Lemma 38 *The number 1 is not an eigenvalue of* $\hat{A} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}$.

Proof:
The number 1 is an eigenvalue of $\hat{A}$ with multiplicity one. That is, 1 is a simple root of the characteristic polynomial of $\hat{A}$. Hence by lemma 33, the subspace of eigenvectors of $\hat{A}$ with eigenvalue 1 is a one

dimensional subspace. Since $\mathbf{e}$ is an eigenvector of $\hat{A}$ with eigenvalue 1, every eigenvector of $\hat{A}$ with eigenvalue 1 is a scalar multiple of $\mathbf{e}$.

Now suppose 1 is an eigenvalue of $\hat{A} - \frac{1}{N}\,\mathbf{e}^\top\mathbf{e}$. Then there is a non-zero vector $\mathbf{v}$ such that
$\mathbf{v}(\hat{A} - \frac{1}{N}\,\mathbf{e}^\top\mathbf{e}) = \mathbf{v}$.
$\mathbf{v}\hat{A} - \frac{1}{N}\,\mathbf{v}\,\mathbf{e}^\top\mathbf{e} = \mathbf{v}$.
Multiplying on the right by $\mathbf{e}^\top$ and using $\hat{A}\mathbf{e}^\top = \mathbf{e}^\top$ yields
$0 = \mathbf{v}\mathbf{e}^\top$.
Substituting 0 for $\mathbf{v}\mathbf{e}^\top$ in the previous equation gives us
$\mathbf{v}\hat{A} = \mathbf{v}$.
So $\mathbf{v}$ is an eigenvector of $\hat{A}$ with eigenvalue 1.
As we have seen, that means $\mathbf{v}$ is a scalar multiple of $\mathbf{e}$. So there is a scalar χ such that
$\mathbf{v} = \chi\mathbf{e}$.
Multiplying on the right by $\mathbf{e}^\top$ gives
$\mathbf{v}\mathbf{e}^\top = \chi N$.
The left side is zero so $\chi = 0$.
So $\mathbf{v} = \chi\mathbf{e} = 0$.
But $\mathbf{v}$ was a non-zero vector.
Contradiction.
■

Taken together, the last two lemmas tell us that every non-zero eigenvalue of $\hat{A} - \frac{1}{N}\,\mathbf{e}^\top\mathbf{e}$ is an eigenvalue of $\hat{A}$ other than 1. Hence it is in the open unit disk.

So every eigenvalue of $\hat{A} - \frac{1}{N}\,\mathbf{e}^\top\mathbf{e}$ is in the open unit disk.
Therefore, if n is an integer variable, we have
$\lim_{n\to\infty} (\hat{A} - \frac{1}{N}\,\mathbf{e}^\top\mathbf{e})^n = 0.$
Since $\hat{A}$ is bistochastic, we have
$(\hat{A} - \frac{1}{N}\,\mathbf{e}^\top\mathbf{e})^n = \hat{A}^n - \frac{1}{N}\,\mathbf{e}^\top\mathbf{e}$.
Therefore, writing $\hat{A}$ as e^{-A} , we have

$$\lim_{n\to\infty} (e^{-A})^n = \tfrac{1}{N}\,\mathbf{e}^\top\mathbf{e} \tag{98}$$

Now suppose t is a positive real number.
$e^{-tA} = \sum_{n=0}^{\infty} \frac{1}{n!}(-tA)^n = \sum_{n=0}^{\infty} \frac{1}{n!}(-1)^n(2t)^n(\frac{1}{2}A)^n$
Using inequality (95), we obtain
$\|e^{-tA}\| \le \sum_{n=0}^{\infty} \frac{1}{n!}(2t)^n(\|\frac{1}{2}A\|)^n \le \sum_{n=0}^{\infty} \frac{1}{n!}(2t)^n = e^{2t}$.
So if t is a positive real, we have

$$\|e^{-tA}\| \le e^{2t} \quad . \tag{99}$$

We know that if two $N\times N$ matrices B and C commute, then
$e^{B+C} = e^B e^C$.
Two useful things follow. The first is that
$e^{-A-A} = e^{-A}e^{-A}$, so $e^{-2A} = (e^{-A})^2$.
Indeed for any integer n , we have

$$e^{-nA} = (e^{-A})^n \quad . \tag{100}$$

The second useful thing is that for any two real numbers s and t , we have

$$e^{(s+t)A} = e^{sA}e^{tA} \quad . \tag{101}$$

From equations (100) and (98), we see that for integer variable n , we have

$$\lim_{n\to\infty} e^{-nA} = \tfrac{1}{N}\,\mathbf{e}^\top\mathbf{e} \quad . \tag{102}$$

Now suppose ρ is any positive real number. Let n be the greatest integer that doesn't exceed ρ . (So $n \le \rho$.)
Define $t = \rho - n$, so we have $0 \le t < 1$.

Consider the following expression.
$(e^{-nA} - \frac{1}{N}\mathbf{e}^\top\mathbf{e})\,e^{-tA}$
Lemma 36 tells us that the matrix e^{-tA} is bistochastic. (If $t = 0$ then $e^{-tA} = I$.)
So multiplying out gives us this.

$e^{-nA}e^{-tA} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}$
Now using equation (101) gives us this.
$e^{-\rho A} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}$
From the above and from equations (11) and (99) and $e^{2t} < e^2$ we obtain the following.
$\|e^{-\rho A} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}\| \;=\; \|(e^{-nA} - \frac{1}{N}\mathbf{e}^\top\mathbf{e})\, e^{-tA}\| \;\le\; \|e^{-nA} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}\| \; \|e^{-tA}\| \;\le\; \|e^{-nA} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}\|\, e^2$.
Let's write out this inequality.
$\|e^{-\rho A} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}\| \;\le\; e^2\, \|e^{-nA} - \frac{1}{N}\mathbf{e}^\top\mathbf{e}\|$.
The inequality holds for any positive real ρ , provided n is the largest integer that doesn't exceed ρ . We ask what happens to the two sides of the inequality if we let $\rho \to \infty$. Both sides change, but the right side changes in jerks. Every time ρ passes an integral value, the n on the right side suddenly jumps up by 1. As ρ goes to ∞ , n also jerks its way to ∞ . Equation (102) tells us that as n jerks its way to ∞ , the right side of the above inequality jerks its way to zero. So as $\rho \to \infty$, the right side goes to zero. So the left side goes to zero too. So for positive real variable ρ , we have

$$\lim_{\rho\to\infty} \|e^{-\rho A} - \tfrac{1}{N}\,\mathbf{e}^\top\mathbf{e}\| \;=\; 0 \quad . \tag{13}$$

It follows that for real variable ρ we have

$$\lim_{\rho\to\infty} e^{-\rho A} \;=\; \tfrac{1}{N}\,\mathbf{e}^\top\mathbf{e} \quad . \tag{103}$$

Equations (13) and (103) are really eqivalent. Each one follows directly from the other.

C Asymmetric Markov Chain on Integers

We consider a Markov chain whose state set is the non-negative integers. The random variable X_t is the state at time t . (t is an integer.) I will think of the non-negative integers as strung out from left to right with zero at the left.

The number n is a positive integer. The numbers p and q are probabilities, with $p+q=1$. The transition rules are as follows:

Transition Rules
With probability q we move right n steps.
With probability p we move left one step,
except that if the current state is 0 then we stay at 0.

In other words, if $X_t = i$ then

$$X_{t+1} \;=\; \begin{cases} i+n & \text{with probability } q \\ \max(0,\ i-1) & \text{with probability } p \end{cases}$$

We note that the chain is irreducible (strongly connected) in the sense that given an ordered pair of states there is a legal transition sequence from the first state to the second.

Now suppose $n=1$ and $q<p$.

We are now going to use (37) from section 3.2. To use it we need an invariant distribution $\boldsymbol{\pi}$.
Define $r \;=\; \frac{q}{p}$.
For each state i , define
$\pi_i \;=\; (1-r)\, r^i$.
We see that $\boldsymbol{\pi}$ is an invariant distribution since for any positive i we have
$\pi_i \;=\; p\,\pi_{i+1} + q\,\pi_{i-1}$,
and we also have
$\pi_0 \;=\; p\,\pi_1 + p\,\pi_0$, and
$\sum_{i=0}^{\infty} \pi_i \;=\; 1$.
So we see that $\boldsymbol{\pi}$ is an invariant distribution and we can use (37).

This tells us that for each state i , the proportion of time the chain is in state i converges (with probability 1) to π_i . Over time, the proportion of time the chain spends in state i is π_i .

In particular, the proportion of time the chain spends in state 0 is π_0 , which is
$1 - \frac{q}{p}$.
So we have shown

Lemma 39 *If* $n = 1$ *and* $q < p$ *then the proportion of time the chain spends in state 0 is* $1 - \frac{q}{p}$.

Now let's consider the general case, where n can be any positive number. We can think of a frog hopping from state to state. The position of the frog at time t is the random variable X_t . In any time unit, the frog either hops left or hops right. To simplify the terminology, I shall say that if the frog hops from state 0 to state 0, I shall also call that a left hop. So the length of a right hop is always n . The length of a left hop is 1, except that if the frog is hopping from state 0 then the length of a left hop is zero. The frog hops left with probability p and hops right with probability q .

We suppose there is also a toad hopping on the same states. When the frog hops left so does the toad. When the frog hops right so does the toad. Toad left hops are length 1 unless the toad is hopping from state 0 in which case a left hop has length zero, just like frog left hops.

But when a toad hops right, it makes sure it lands on a state (number) that is divisible by n . It hops at least n states, but might have to hop more. For example, suppose $n = 7$ and the frog and toad are both sitting on state 9. If they both do a right hop, the frog will land on state 16 and the toad will land on state 21.

Again suppposing $n = 7$, suppose the frog is on state 0 and the toad is on state 4. If they both do a left hop, the frog lands on state 0 and the toad lands on state 3. If they both do a right hop, the frog lands on state 7 and the toad lands on state 14.

The toad begins either on the same state as the frog or on a state somewhere to the right of the frog. We see that there is no way the toad can ever get to the left of the frog.

The clock ticks once each time unit. Attached to the clock is a bell that dings on some time units and not on others. I will now say when the bell dings.

If the frog and toad hop right, then the bell dings when they land. If the frog and toad hop left then the bell doesn't usually ding, but sometimes it does. Let me explain. If there have been n hops since the last ding, then the bell dings even if the frog and toad hop left.

Let me illustrate. Suppose $n = 5$. Suppose the sequence of left (ℓ) and right (r) hops is this, beginning with time 0.

$r\,\ell\,\ell\,\ell\,\ell\,\ell\,\ell\,\ell\,r\,\ell\,\ell\,\ell\,\ell\,\ell\,\ell\,\ell$

Then there is a right hop at times 0 and 8. The bell dings at times 0, 5, 8, and 13.

Times on which the bell dings I shall call marked times. (We complete the story by saying that the bell always dings at time 0.)

The hops from one ding to the next I will call an *interval*. There are two possibilities. In what I call a *left interval*, the interval is a sequence of n left hops. In what I call a *right interval*, it is a sequence of less than n left hops, followed by a single right hop.

At the end of a right interval, the toad is on a state divisible by n . If a left interval begins with the toad on a state divisible by n , then it ends with him on a state divisible by n . So we see that once a right hop is made, the toad will at all marked times be on a state divisible by n . An interval takes the toad from a state divisible by n to a state divisible by n .

So suppose at the start of an interval the toad is on state ni . Where will the toad be at the end of the interval? If it is a right interval, the toad will be at $n(i + 1)$. If it is a left interval, the toad will be at $n(i - 1)$ if $i > 0$, and will be at 0 if $i = 0$.

The probability of a left interval is p^n and the probability of a right interval is $1 - p^n$. The distance moved in an interval is n . (Or 0 if it is a left interval starting at 0.) So if we look just at the marked time units and states $0, n, 2n, 3n, ...$ we have a Markov chain and can use lemma 39. The p and q there are here p^n and $1 - p^n$ respectively. So the condition $q < p$ becomes $1 - p^n < p^n$ or $\frac{1}{2} < p^n$. The limiting probability of state 0 is

$$1 - \frac{1 - p^n}{p^n}$$

or

$$2 - \frac{1}{p^n} \quad . \tag{104}$$

How long is an interval? The length of a left interval is n . The average length of a right interval is $(1 - p^n)^{-1} \sum_{k=0}^{n-1} (k + 1) p^k q$. This quantity times $1 - p^n$ plus n times p^n gives us the average length of an interval. Right intervals are of course shorter than left intervals.

Average interval length turns out to be $(1 - p^n)\, q^{-1}$, but the exact value is not the important point. The important point is that the interval length calculations are independent of the state the inteval starts in. If the starting state is ni , then all the states during the interval (not including the starting and finishing state) are all strictly between $n(i - 1)$ and ni . Or rather, they are if $i > 0$. If $i = 0$, then all those states in the interval are 0.

This discussion of interval lengths makes it clear that the proportion of time the toad spends in an interval that starts at 0 is simply the proportion of marked times in which the toad is at 0, and this proportion is $2-p^{-n}$. That's the proportion of time the toad spends in intervals that start at 0. During those intervals, the toad is in state 0, so the proportion of time the toad spends in state 0 is at least $2-p^{-n}$. Or rather, That's the proportion of time if the limiting probability exists. And the limiting probability exists if we meet the condition $\frac{1}{2} < p^n$. If the condition is not met then perhaps the toad heads off to the right and the proportion of time the toad is on 0 is zero.

The toad is never to the left of the frog, so we have this lemma.

Lemma 40
For a frog following the Transition Rules, we have this result.
If the condition $\frac{1}{2} < p^n$ *is met, then*
the proportion of time the frog spends in state 0 is at least $2-p^{-n}$.

Alternative Wording:

Lemma 40
For a Markov chain described by the Transition Rules, we have this result.
If the condition $\frac{1}{2} < p^n$ *is met, then*
the proportion of time the chain spends in state 0 is at least $2-p^{-n}$.

References

[1] F. R. Gantmacher. *Applications of the Theory of Matrices.* Interscience Publishers, New York, 1959.

[2] Michel Loeve. *Probability Theory, second edition.* Van Nostrand, Princeton, 1955.

[3] J. R. Norris. *Markov Chains.* Cambridge University Press, Cambridge, 1997.