Quasi-Stationary Distributions for Continuous-Time Markov Chains

David Sirl

AUSTRALIAN RESEARCH COUNCIL
Centre of Excellence for Mathematics
and Statistics of Complex Systems

Recall . . .

- A time-homogeneous CTMC $(X(t),\ t \geq 0)$ taking values in a countable set S ($\mathbb{Z}^+$) is completely described by its *transition function* $P(t) = (p_{ij}(t),\ i,j \in S,\ t \geq 0)$.
- In practice we usually know only the *transition rates*: $(q_{ij} = p'_{ij}(0^+),\ i,j \in S)$ is the *q-matrix*.
  - $q_{ij},\ i \neq j$, is the transition rate from state i to state j,
  - $-q_{ii} = q_i = \sum_{j \neq i} q_{ij}$ is the total rate out of state i.
- If we know P, we can in principle answer any question about the behaviour of the chain. The challenge is to try and answer these questions in terms of Q.

Recall . . .

A Birth-Death Process is a CTMC with state space $S = \{0, 1, 2, \ldots\}$ such that if the chain is in state i, transitions can only be made to state $i-1$ or $i+1$.

Its' q-matrix has non-zero entries

$$\begin{aligned} q_{i,i+1} &= \lambda_i, \\ q_{i,i-1} &= \mu_i, \\ q_{ii} &= -(\lambda_i + \mu_i), \end{aligned}$$

(put $\mu_0 = 0$) where $\lambda_i, \mu_i > 0 \; \forall i \in C$, and $\lambda_0 \geq 0$.

The Chemical Reaction

$$\mathsf{A} + \mathsf{X} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} 2\mathsf{X}$$

$$\mathsf{X} \xrightarrow{k_2} \mathsf{B}$$

The Chemical Reaction

$$\mathrm{A} + \mathrm{X} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} 2\mathrm{X}$$

$$\mathrm{X} \xrightarrow{k_2} \mathrm{B}$$

- Model the number of molecules of X with a CTMC — a birth-death process on $S = \{0\} \cup C$, where zero is absorbing and C is an irreducible transient class.
- The system can be either *closed* or *open* with respect to A & B. $C = \{1, 2, \ldots, N\}$ or $\{1, 2, \ldots\}$, respectively.

The Chemical Reaction

$$\mathsf{A} + \mathsf{X} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} 2\mathsf{X}$$

$$\mathsf{X} \xrightarrow{k_2} \mathsf{B}$$

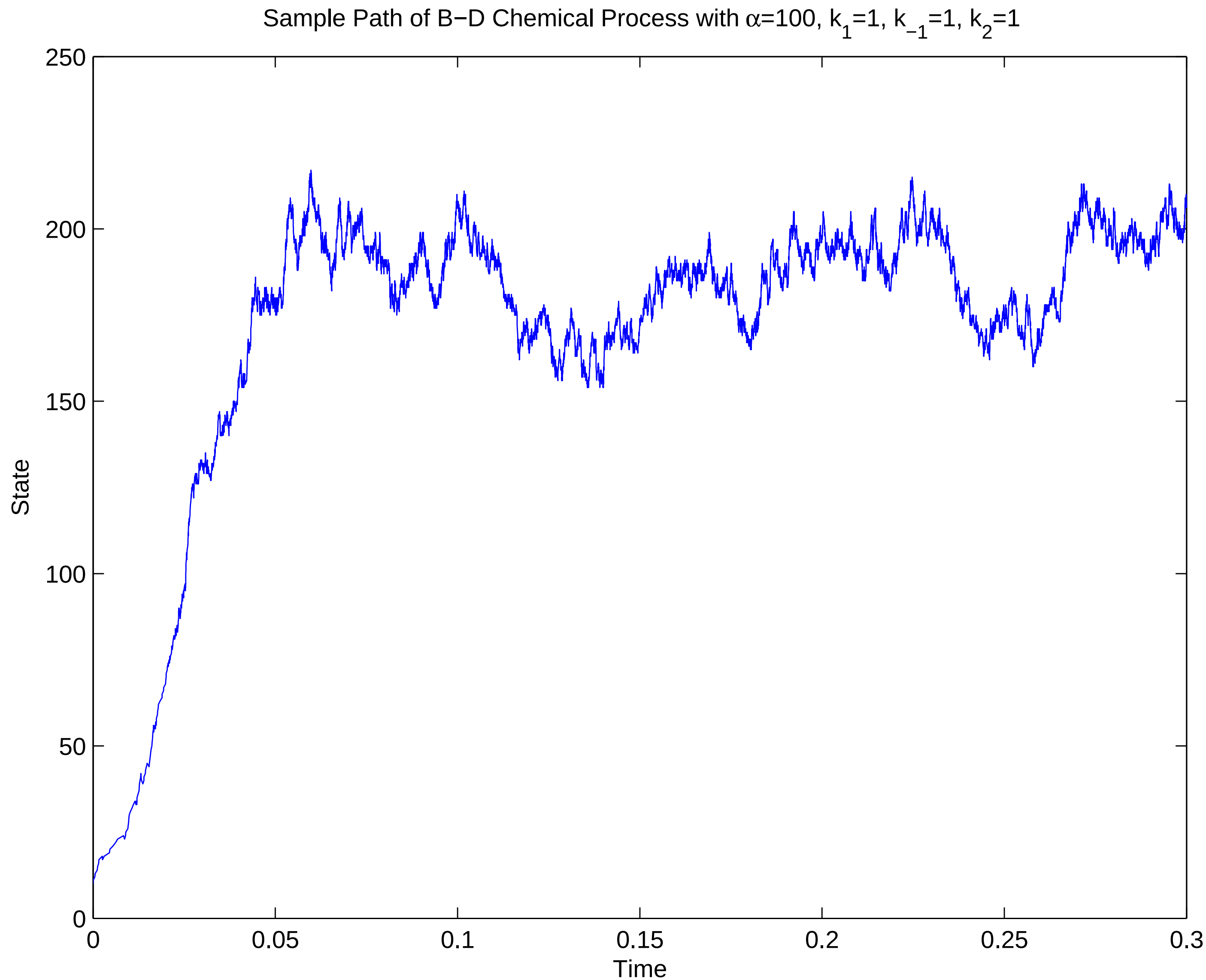
The birth and death rates are, respectively,

$$\lambda_i = \alpha k_1 i,$$

and

$$\mu_i = k_2 i + k_{-1} \frac{i(i-1)}{2}.$$

A Sample Path



A Stationary Distribution?

Q: Is this behaviour limiting?

- The state space is reducible — $S = \{0\} \cup C$.
- The class C is transient.
- The limiting distribution is $\pi = (1, 0, 0, \ldots)$.

A Stationary Distribution?

Q: Is this behaviour limiting?

- The state space is reducible — $S = \{0\} \cup C$.
- The class C is transient.
- The limiting distribution is $\pi = (1, 0, 0, \ldots)$.

A: No — looking at the limiting behaviour sees all of the probability going to the absorbing state.

A Stationary Distribution?

Q: Is this behaviour limiting?

- The state space is reducible — $S = \{0\} \cup C$.
- The class C is transient.
- The limiting distribution is $\pi = (1, 0, 0, \ldots)$.

A: No — looking at the limiting behaviour sees all of the probability going to the absorbing state.

So how can we explain this behaviour?

Instead

- We need to condition on the process having not been absorbed at time t.

Instead

- We need to condition on the process having not been absorbed at time t.
- Rather than the transition probabilities

$$p_{ij}(t) = \mathbb{P}(X(t) = j \mid X(0) = i),$$

we consider the conditional transition probabilities

$$\begin{aligned} m_{ij}(t) &\stackrel{\text{def}}{=} \frac{p_{ij}(t)}{1 - p_{i0}(t)} \\ &= \mathbb{P}(X(t) = j \mid X(t) \in C\,,\ X(0) = i). \end{aligned}$$

Definitions

- A distribution $a = (a_i,\ i \in C)$ is a QSD over C if when the initial distribution is a, the state probabilities $p_{aj}(t) = \sum_{i \in C} a_i p_{ij}(t)$ conditioned on non-absorption are time-invariant (and given by a):

$$\frac{p_{aj}(t)}{1 - p_{a0}(t)} = a_j, \quad j \in C.$$

- A distribution $b = (b_i,\ i \in C)$ is a a-LCD over C if when a is the initial distribution, b_j gives the limiting probability of the process being in state j, conditional on non-absorption:

$$\lim_{t \to \infty} \frac{p_{aj}(t)}{1 - p_{a0}(t)} = b_j, \quad j \in C.$$

Definitions

- A ν-invariant measure (over C) for P is a collection of numbers $m = (m_i,\ i \in C)$ which, for some $\nu > 0$, satisfy

$$\sum_{i \in C} m_i p_{ij}(t) = e^{-\nu t} m_j, \qquad j \in C,\ t \geq 0.$$

Definitions

- A ν-invariant measure (over C) for P is a collection of numbers $m = (m_i,\ i \in C)$ which, for some $\nu > 0$, satisfy

$$\sum_{i \in C} m_i p_{ij}(t) = e^{-\nu t} m_j, \qquad j \in C,\ t \geq 0.$$

- A ν-invariant measure (over C) for Q is a collection of numbers $m = (m_i,\ i \in C)$ which, for some $\nu > 0$, satisfy

$$\sum_{i \in C} m_i q_{ij} = -\nu m_j, \qquad j \in C.$$

Finite State Space

- Easy because of spectral decomposition (of P) and Perron-Frobenius theory.
- The δ_i-LCD and unique QSD is given by the probability measure m such that

$$mP_C(t) = e^{-\nu_1 t} m.$$

- This is equivalent to

$$mQ_C = -\nu_1 m,$$

where $-\nu_1$ is the eigenvalue with maximal real part (it is real and negative).

A Simple Example

Lets look at the CTMC with the following q-matrix:

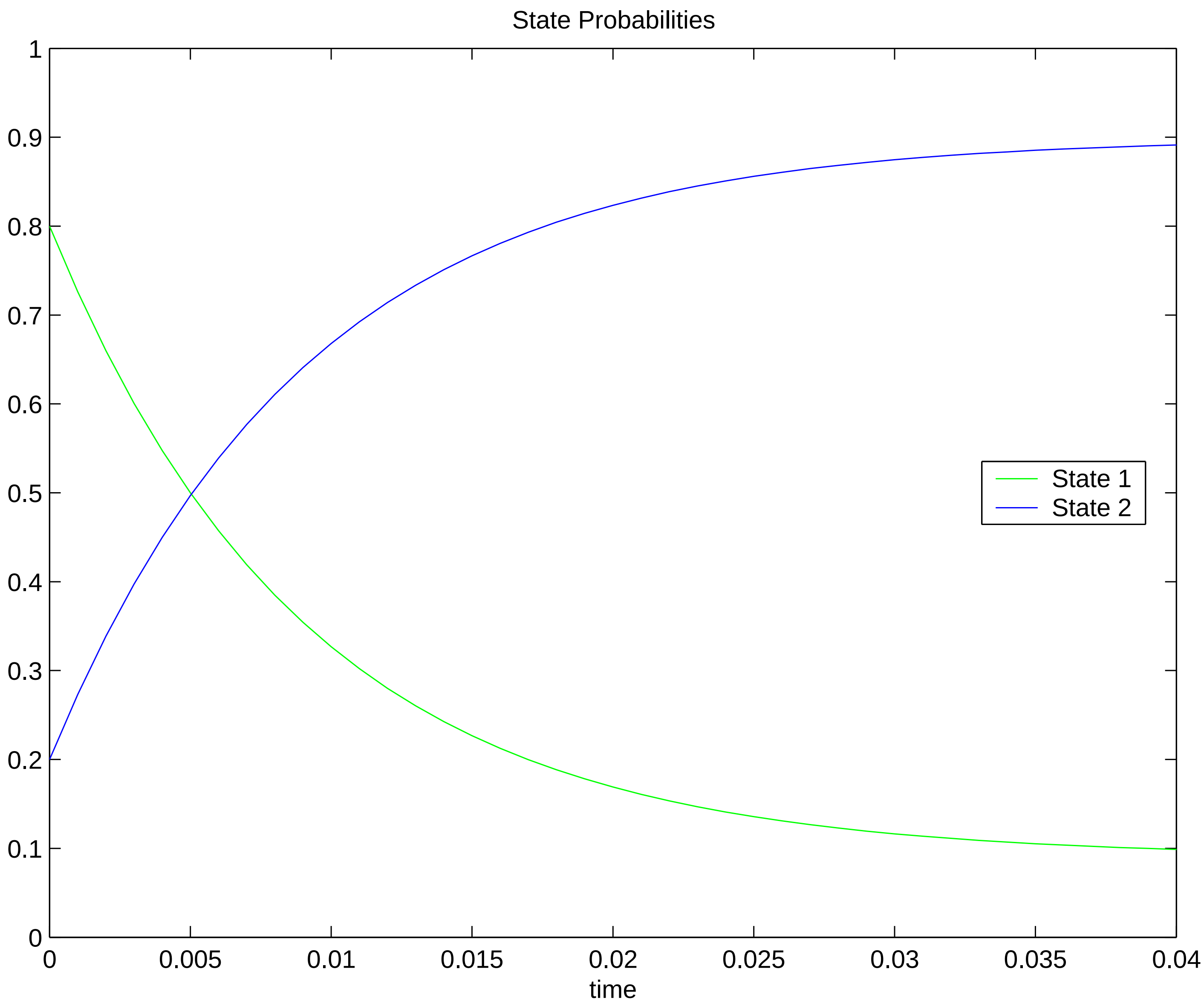
$$Q = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -100 & 99 \\ 0 & 10 & -10 \end{pmatrix}$$

- We know that the transition function is $P(t) = \exp(tQ)$.
- We can get Maple to calculate $P(t)$, and indeed the state probabilities

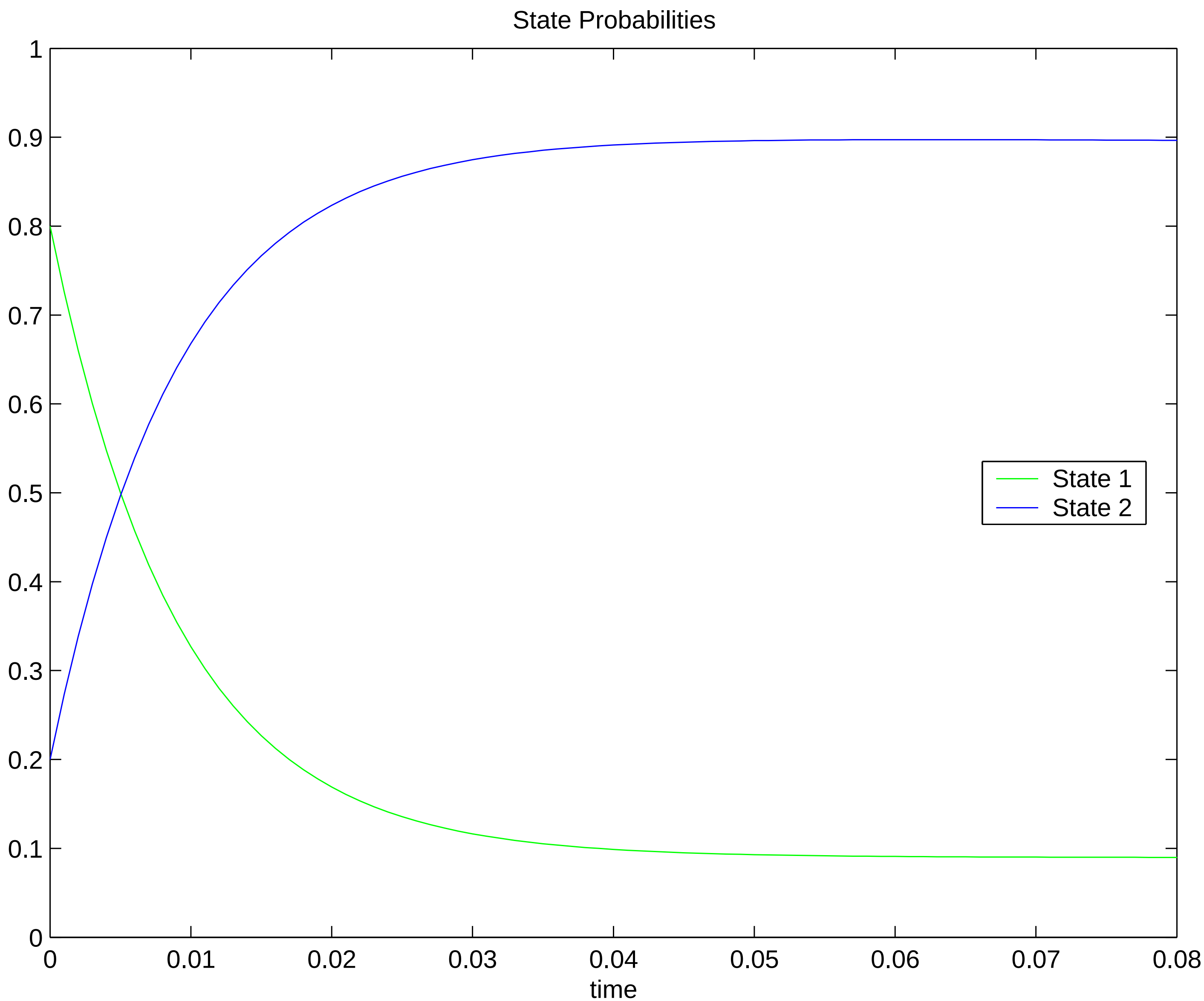
$$p_j(t) = \sum_{i \in S} a_i p_{ij}(t), \quad j \in S,$$

where a is an initial distribution, say $\left(0 \ \frac{4}{5} \ \frac{1}{5}\right)$.

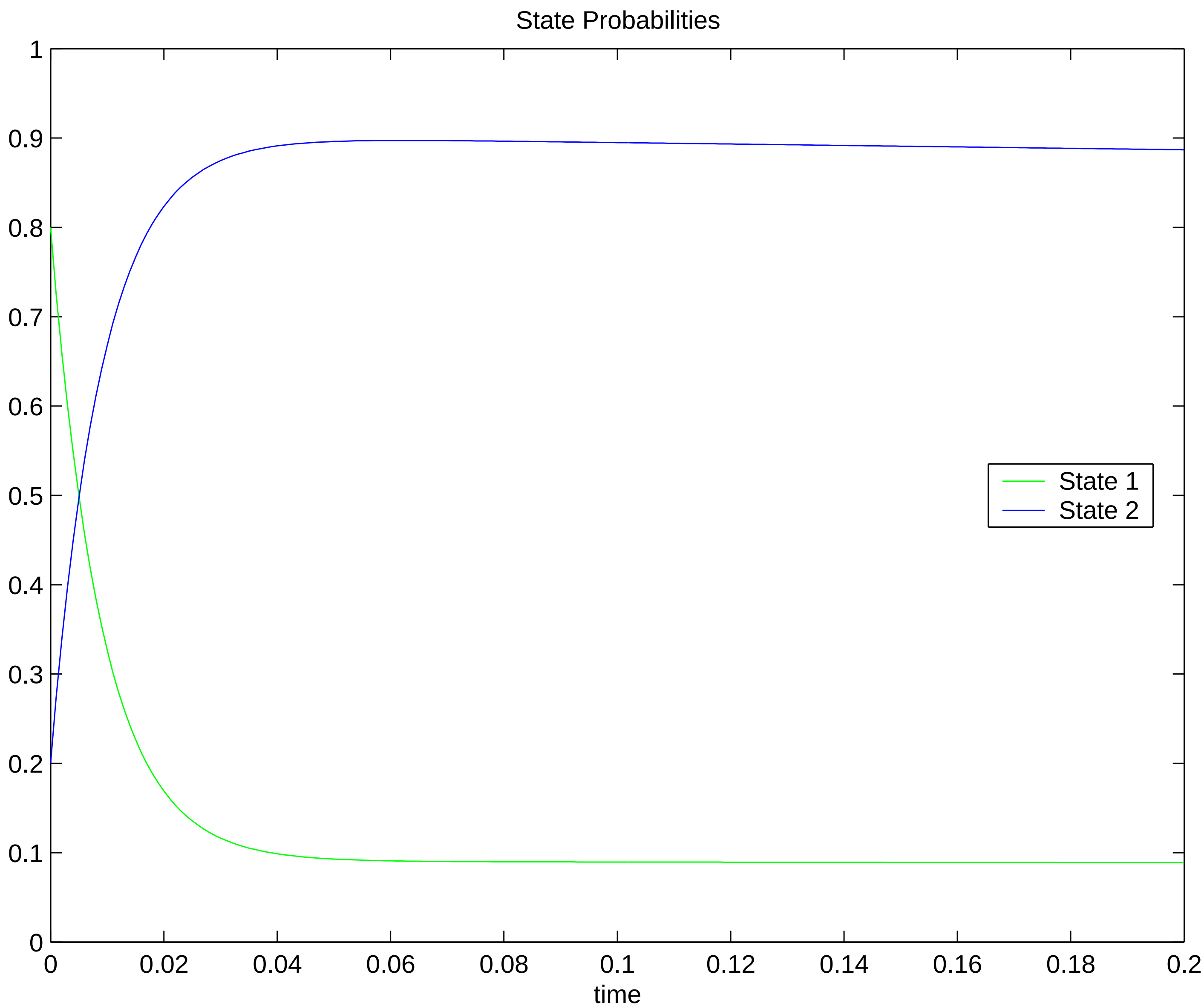
A Simple Example



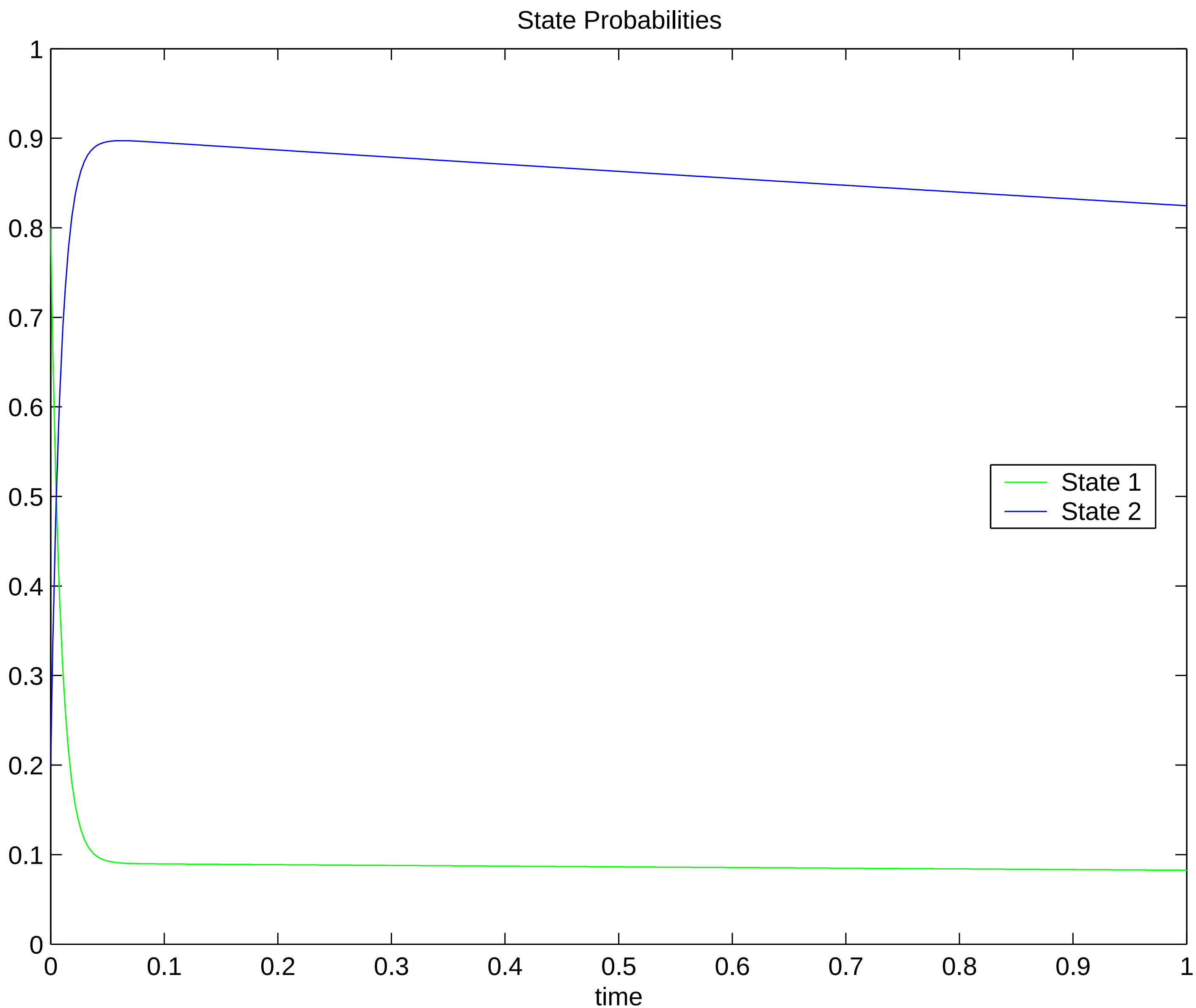
A Simple Example



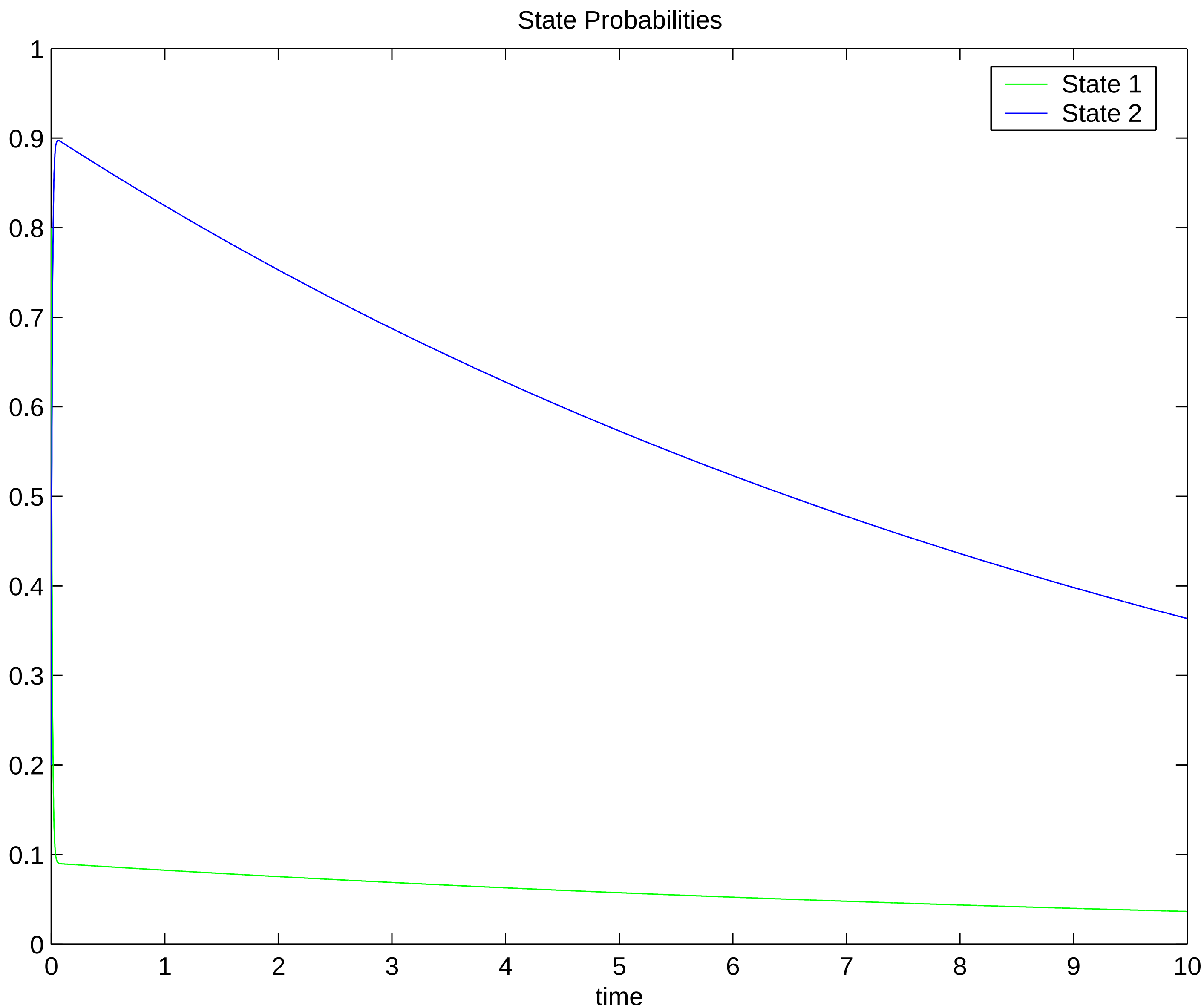
A Simple Example



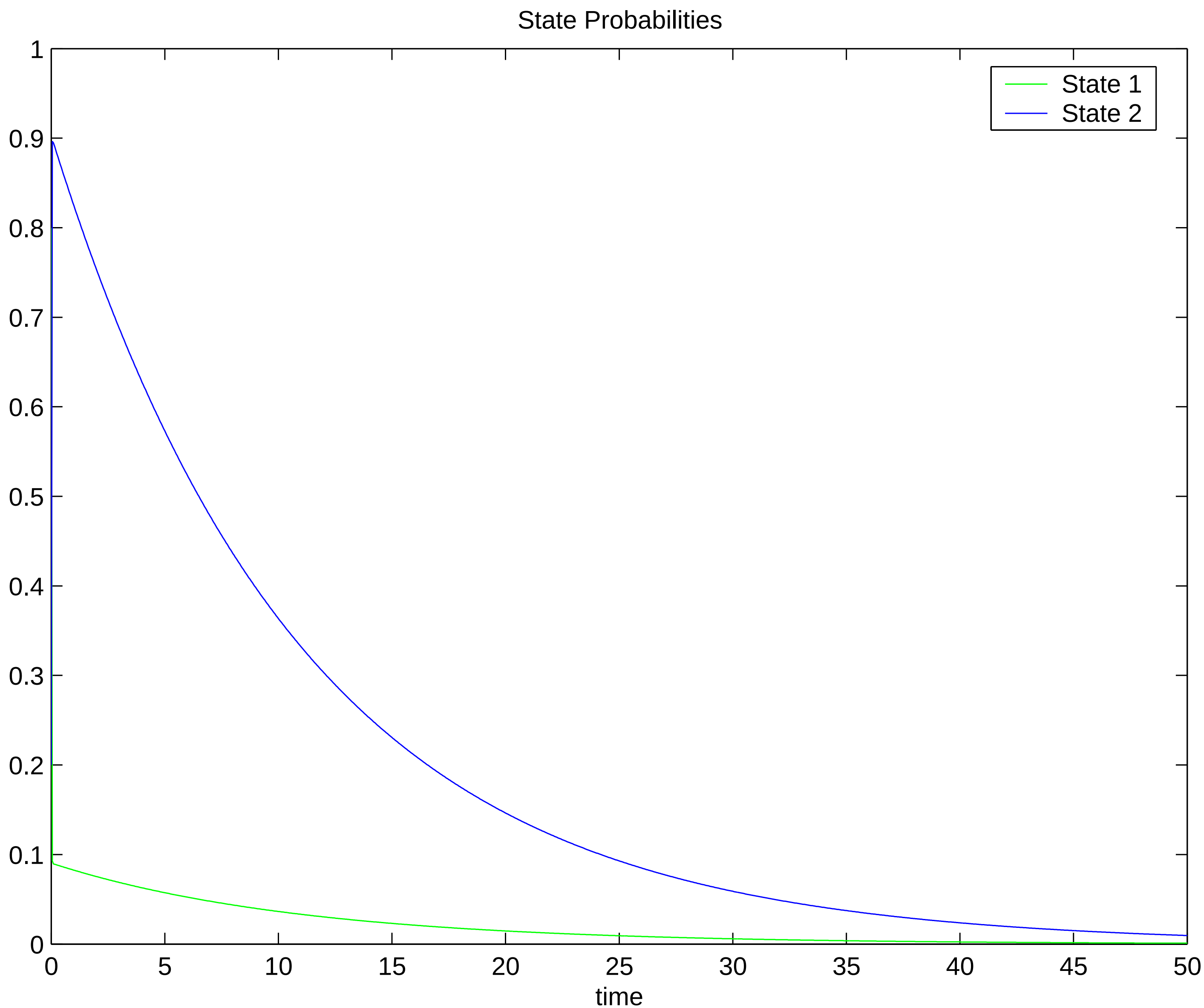
A Simple Example



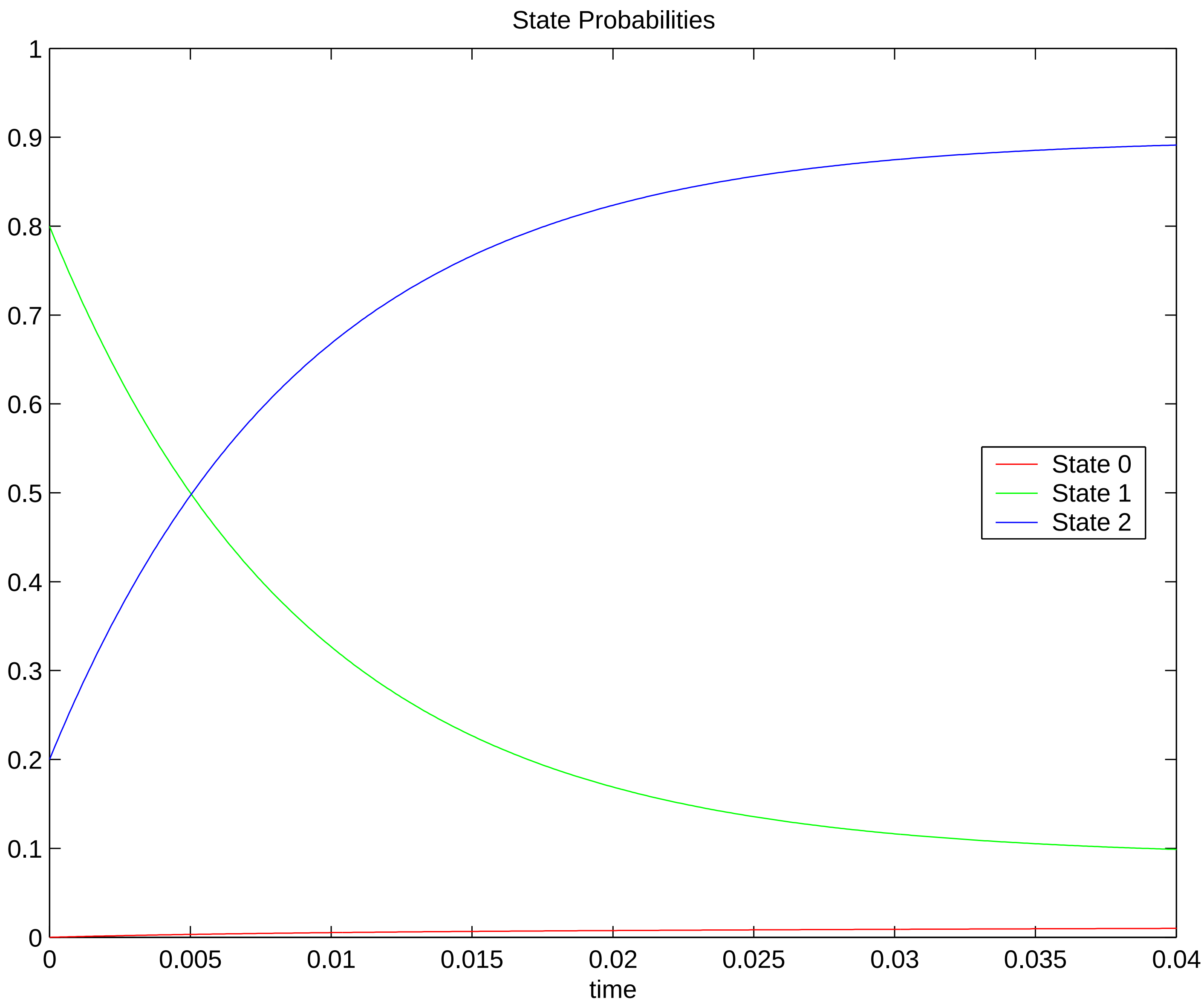
A Simple Example



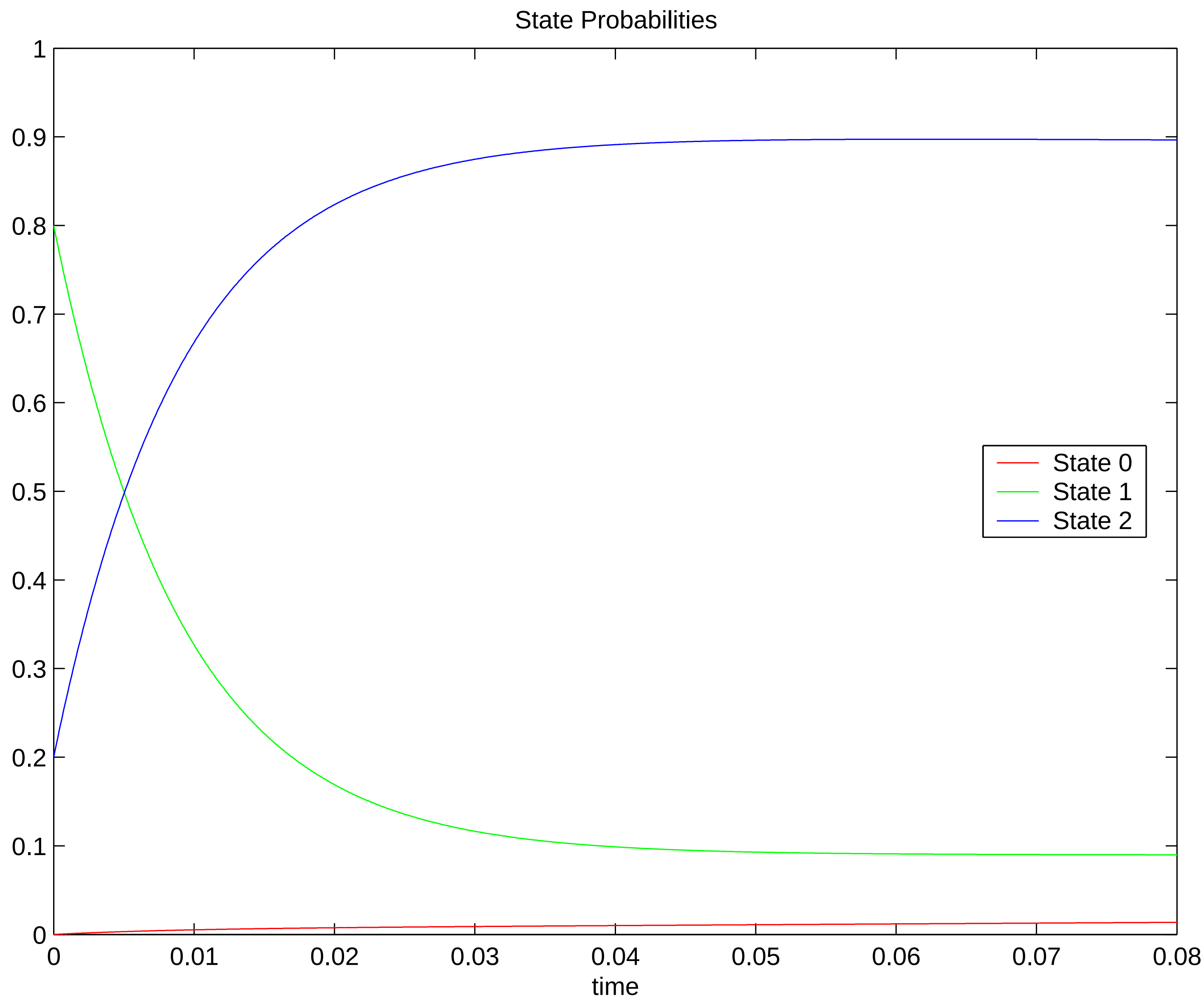
A Simple Example



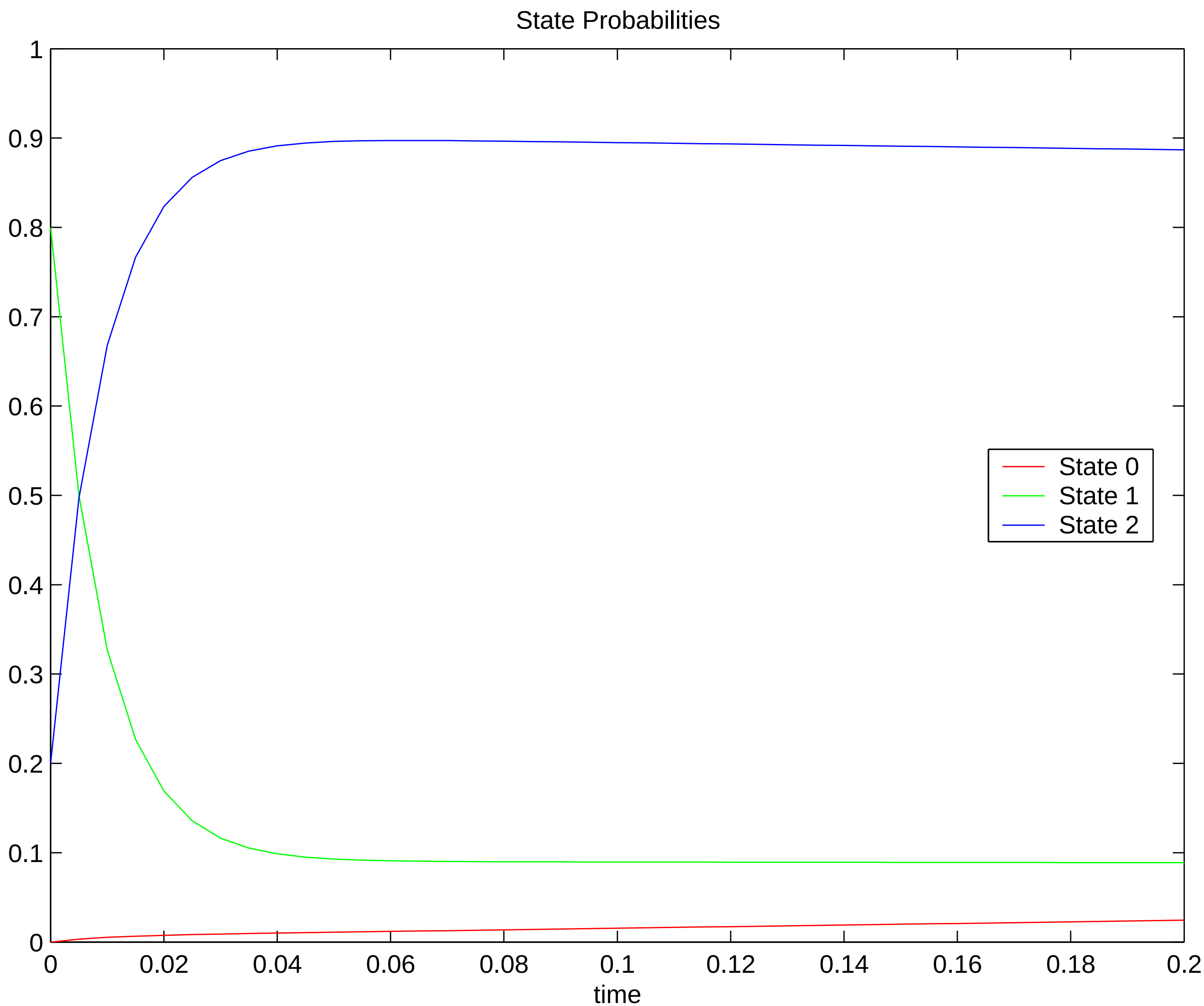
A Simple Example



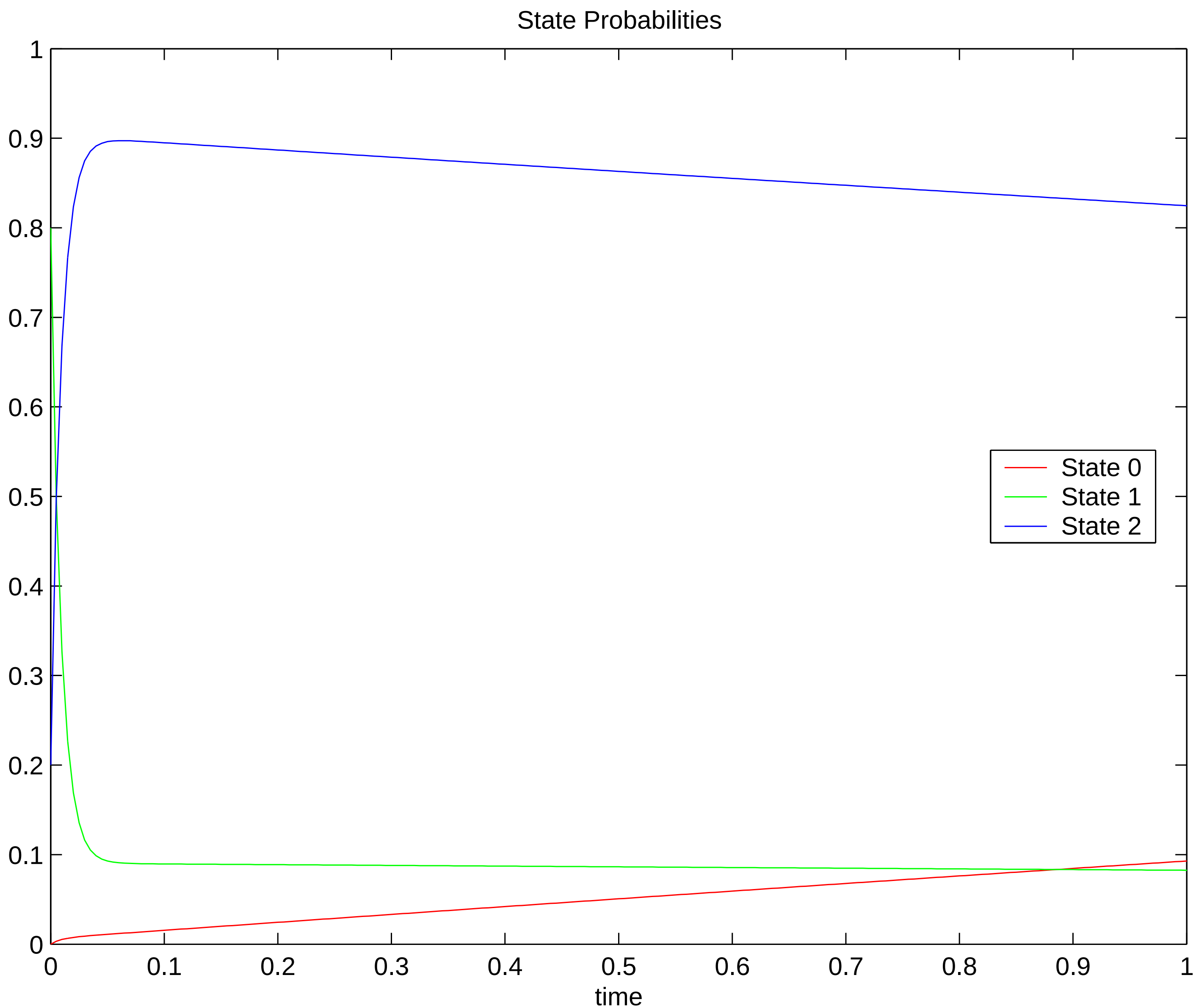
A Simple Example



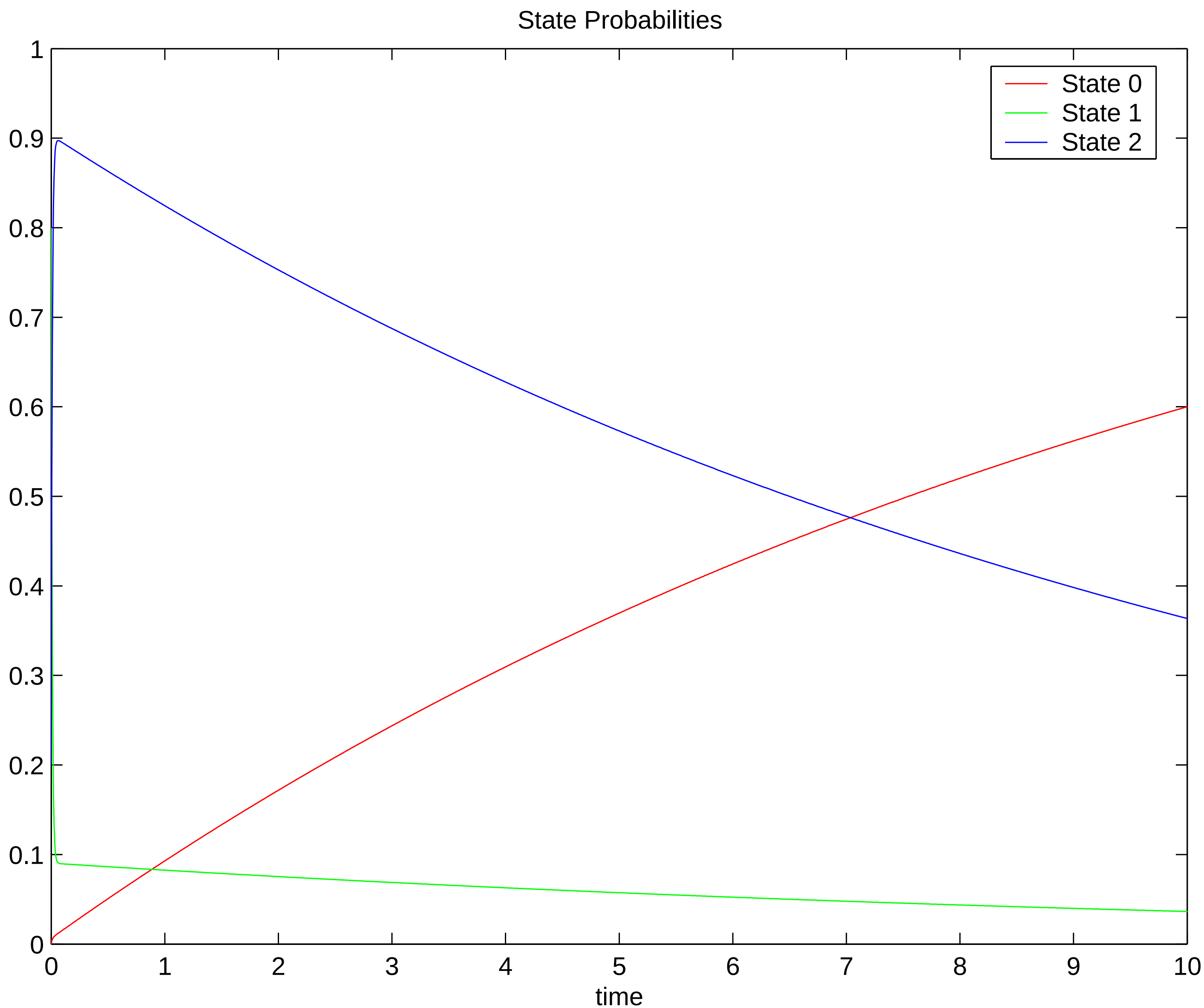
A Simple Example



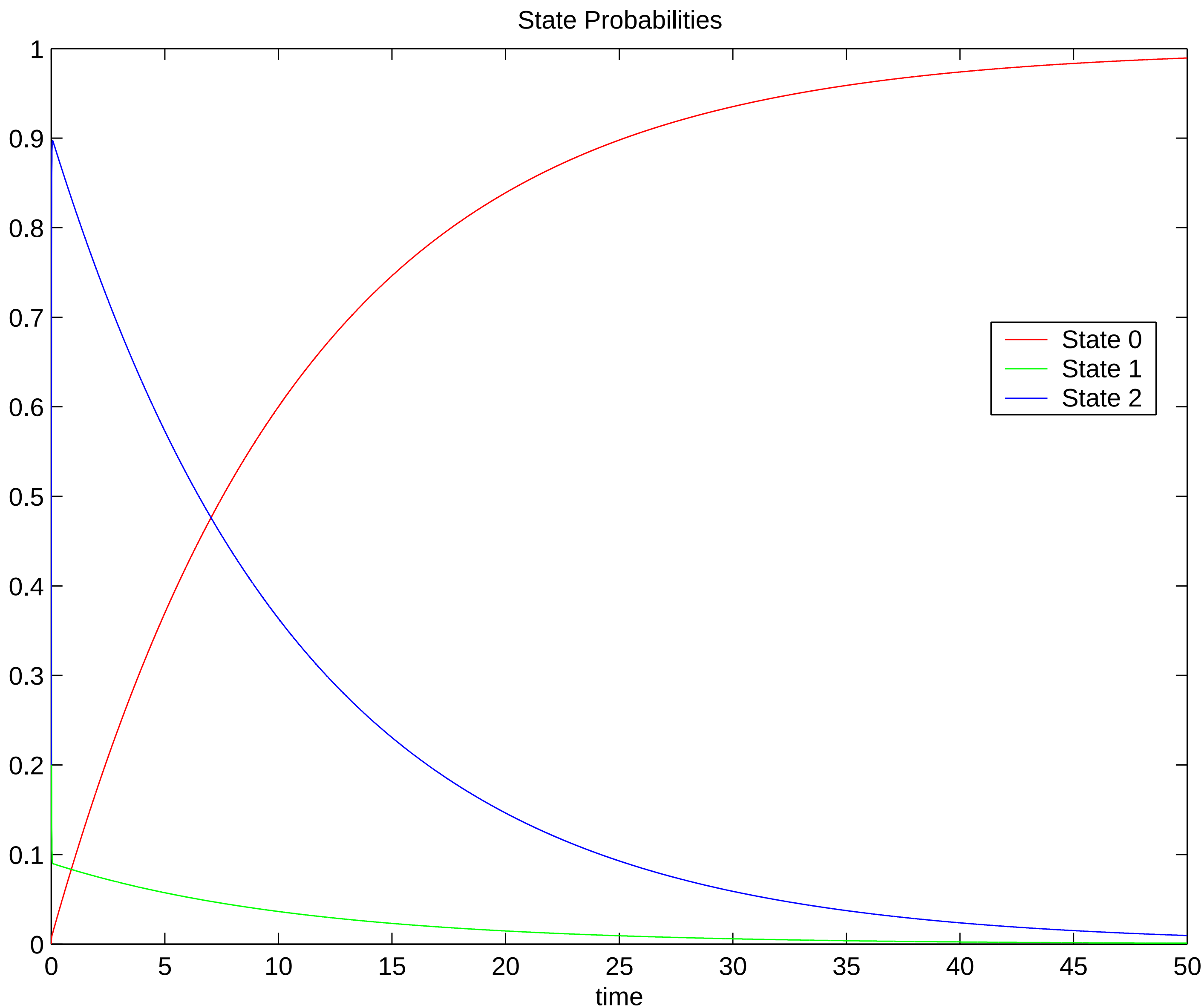
A Simple Example



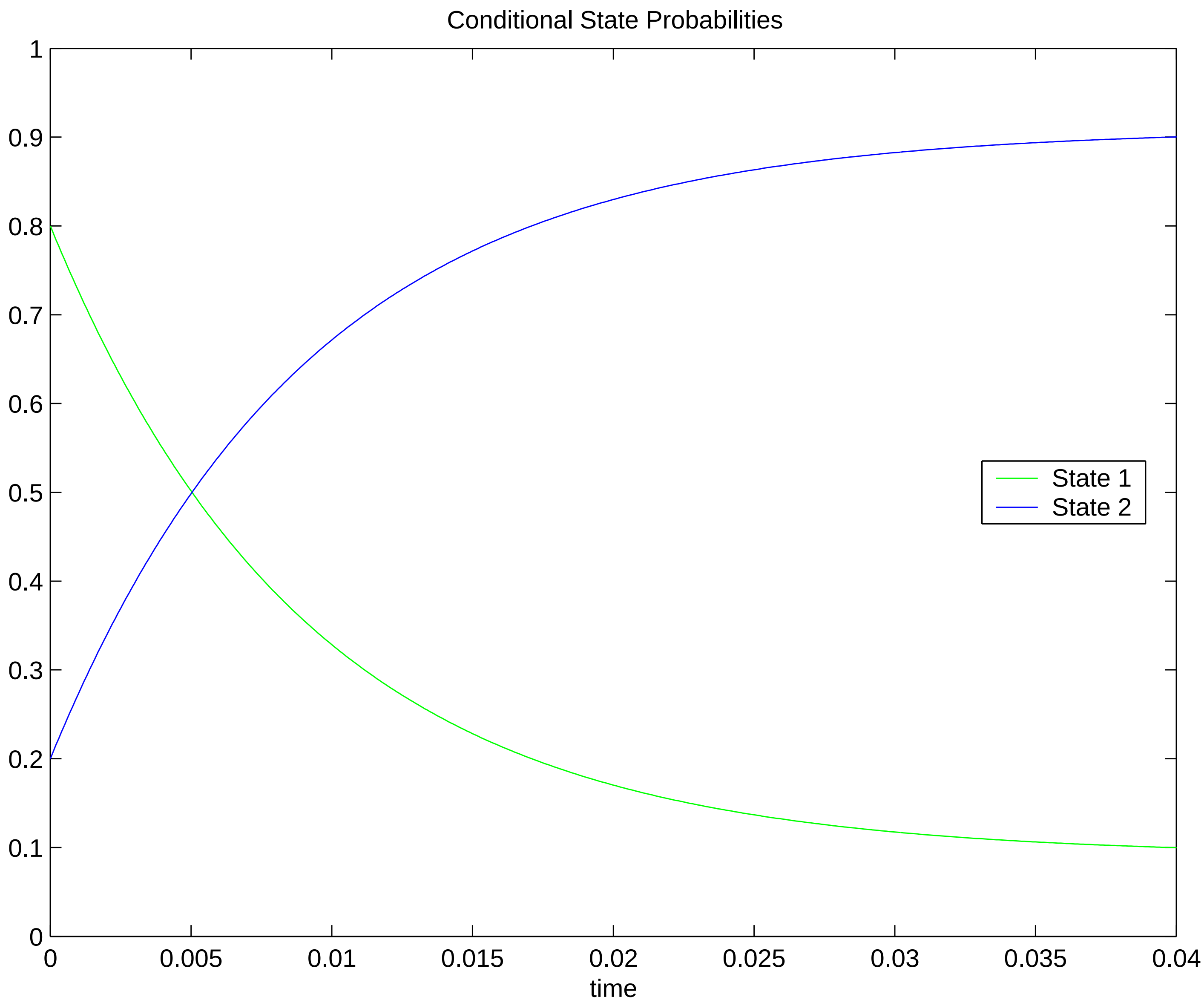
A Simple Example



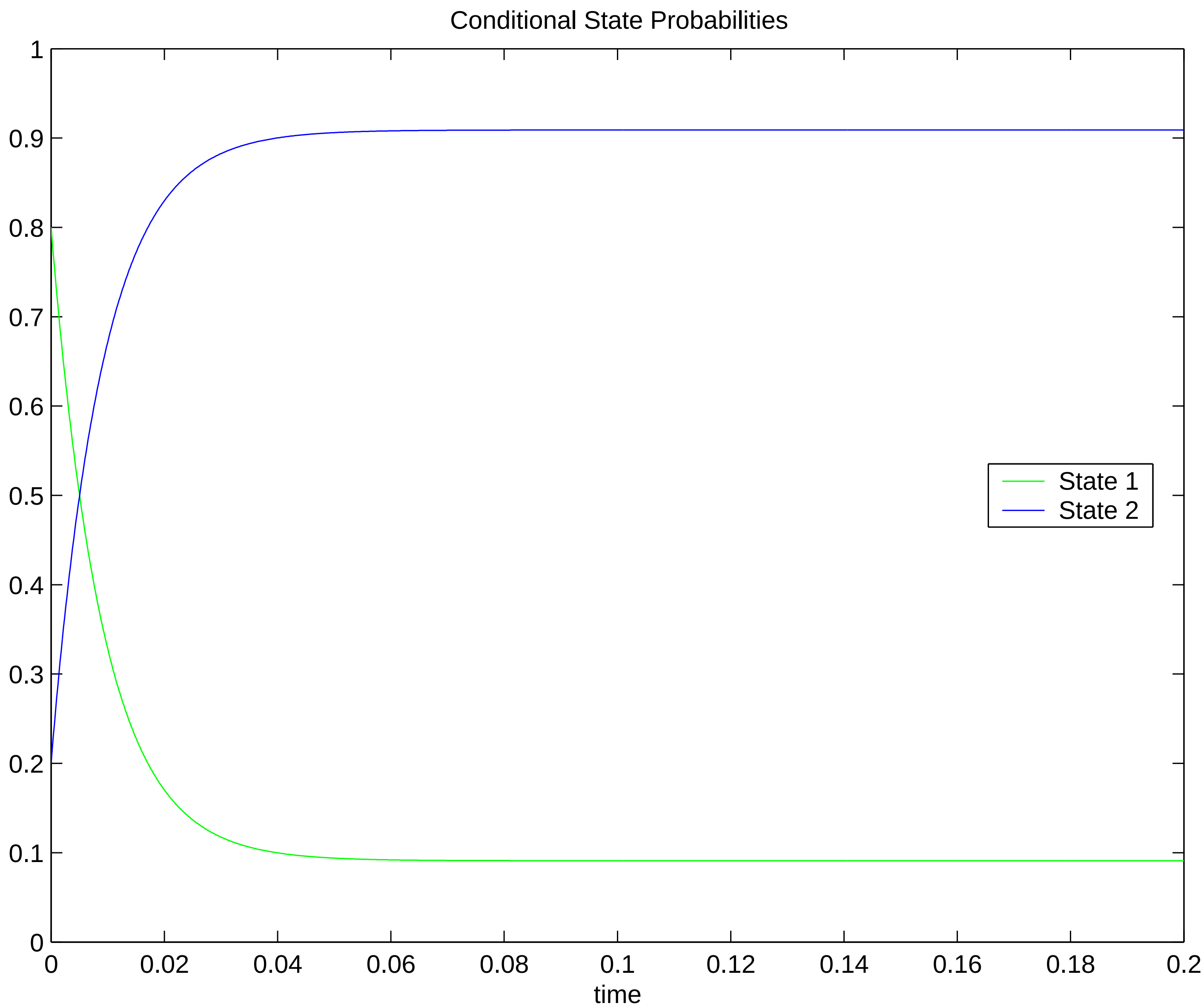
A Simple Example



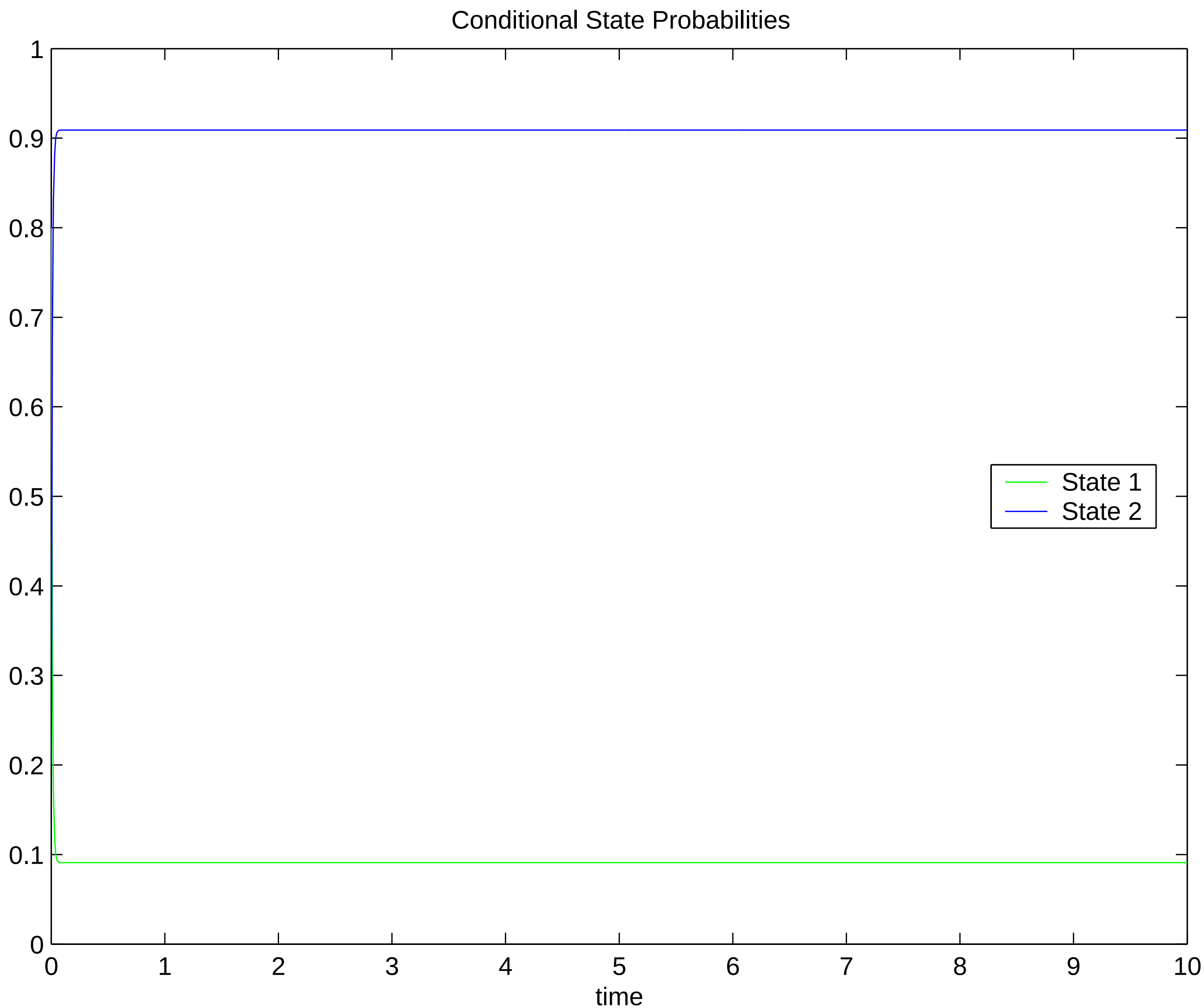
A Simple Example



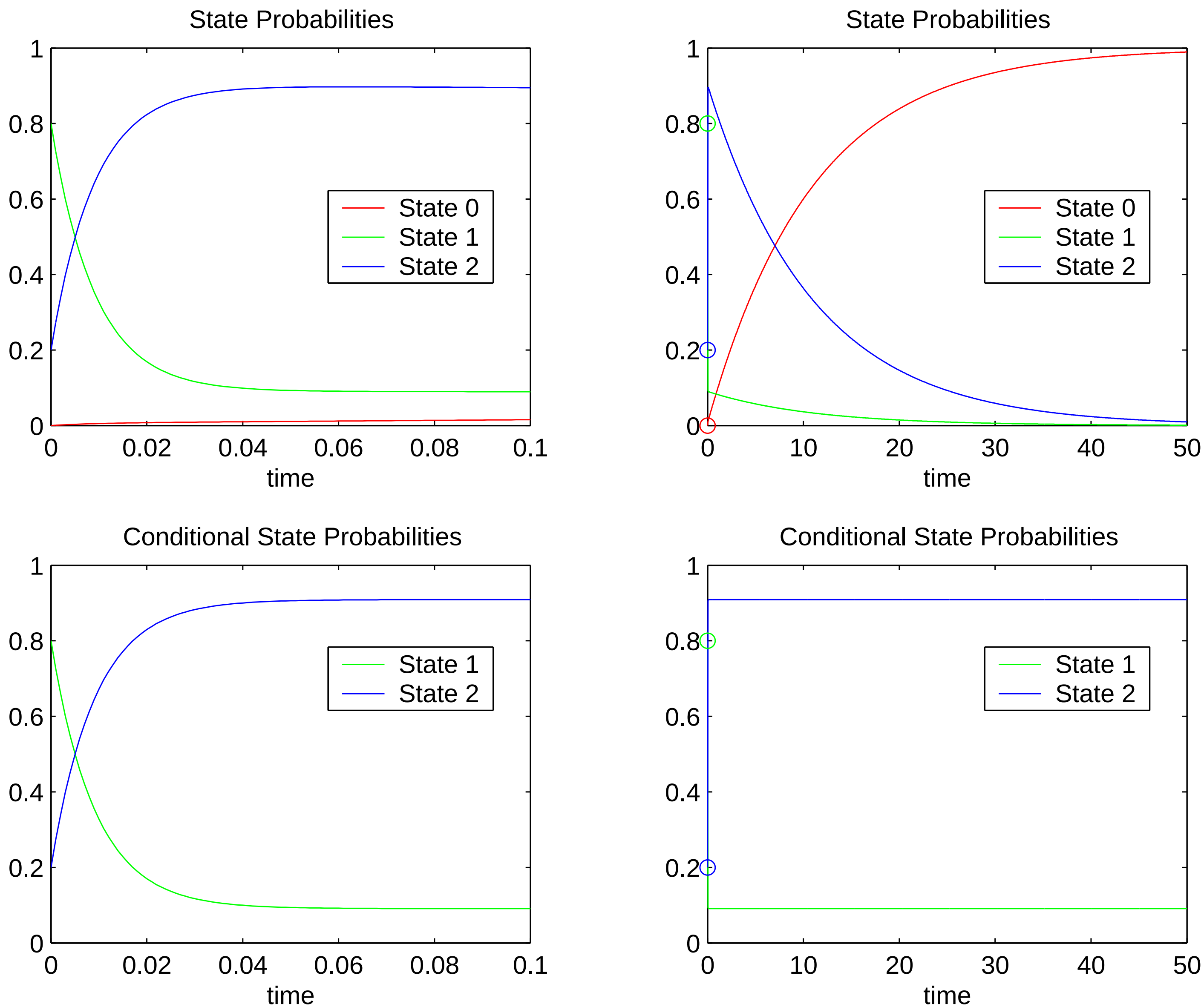
A Simple Example



A Simple Example



A Simple Example



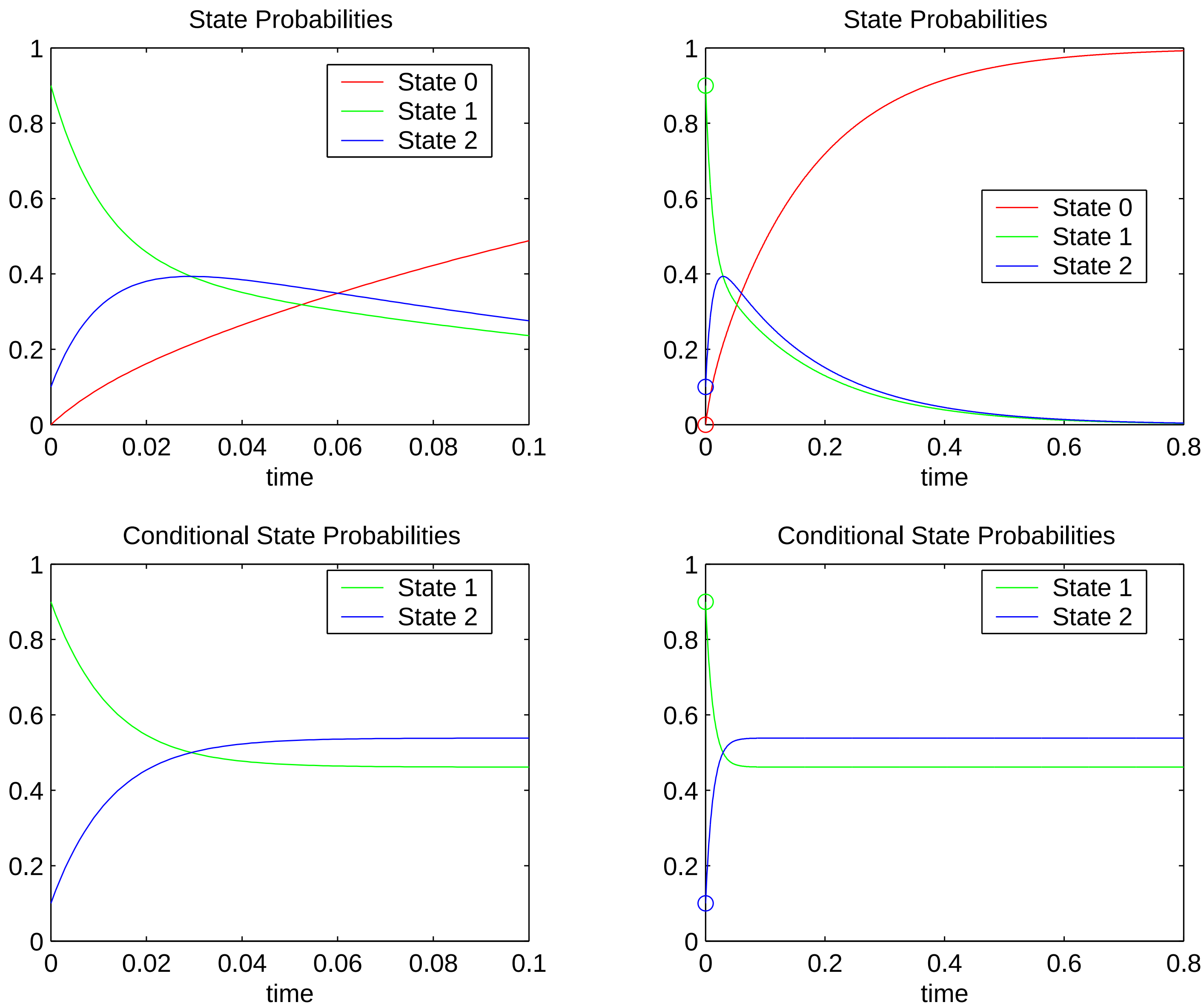
A Simple Example II

Lets look at the CTMC with the following q-matrix:

$$Q = \begin{pmatrix} 0 & 0 & 0 \\ 13 & -55 & 42 \\ 0 & 42 & -42 \end{pmatrix}$$

Again we can get Maple to evaluate P and p, and then use Matlab to plot them:

A Simple Example II



A Simple Example II

$$P(t) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + e^{-6t} \begin{pmatrix} 0 & 0 & 0 \\ \frac{-78}{85} & \frac{36}{85} & \frac{42}{85} \\ \frac{-91}{85} & \frac{42}{85} & \frac{49}{85} \end{pmatrix} + e^{-91t} \begin{pmatrix} 0 & 0 & 0 \\ \frac{-7}{85} & \frac{49}{85} & \frac{-42}{85} \\ \frac{6}{85} & \frac{-42}{85} & \frac{36}{85} \end{pmatrix}$$

A Simple Example II

$$P(t) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + e^{-6t} \begin{pmatrix} 0 & 0 & 0 \\ \frac{-78}{85} & \frac{36}{85} & \frac{42}{85} \\ \frac{-91}{85} & \frac{42}{85} & \frac{49}{85} \end{pmatrix} + e^{-91t} \begin{pmatrix} 0 & 0 & 0 \\ \frac{-7}{85} & \frac{49}{85} & \frac{-42}{85} \\ \frac{6}{85} & \frac{-42}{85} & \frac{36}{85} \end{pmatrix}$$

Now, solving $mQ = -\nu m$ gives

$$\nu_1 = 6, \quad \nu_2 = 91$$

and

$$m_1 = \begin{pmatrix} \frac{6}{13} & \frac{7}{13} \end{pmatrix}.$$

Definitions

- A distribution $a = (a_i,\ i \in C)$ is a QSD over C if when the initial distribution is a, the state probabilities $p_{aj}(t) = \sum_{i \in C} a_i p_{ij}(t)$ conditioned on non-absorption are time-invariant (and given by a):

$$\frac{p_{aj}(t)}{1 - p_{a0}(t)} = a_j, \quad j \in C.$$

- A distribution $b = (b_i,\ i \in C)$ is a a-LCD over C if when a is the initial distribution, b_j gives the limiting probability of the process being in state j, conditional on non-absorption:

$$\lim_{t \to \infty} \frac{p_{aj}(t)}{1 - p_{a0}(t)} = b_j, \quad j \in C.$$

Definitions

- A ν-invariant measure (over C) for P is a collection of numbers $m = (m_i,\ i \in C)$ which, for some $\nu > 0$, satisfy

$$\sum_{i \in C} m_i p_{ij}(t) = e^{-\nu t} m_j, \qquad j \in C,\ t \geq 0.$$

- A ν-invariant measure (over C) for Q is a collection of numbers $m = (m_i,\ i \in C)$ which, for some $\nu > 0$, satisfy

$$\sum_{i \in C} m_i q_{ij} = -\nu m_j, \qquad j \in C.$$

The Decay Parameter

- The quantity

$$\lambda_C := \lim_{t \to \infty} \frac{-\log(p_{ij}(t))}{t}$$

exists and is independent of $i, j \in C$.

- Called the decay parameter because

$$p_{ij}(t) \leq M_{ij} e^{-\lambda_C t}, \qquad 0 < M_{ij} < \infty.$$

- Can show that for a ν-invariant measure for P over C to exist, it is necessary that $(0 <)\nu \leq \lambda_C$.

Infinite State Space

- P is now an infinite matrix, so Perron-Frobenius theory no longer applies.

Infinite State Space

- P is now an infinite matrix, so Perron-Frobenius theory no longer applies.
- QSDs (if they exist) are given by *probability measures* m such that

$$mP_C(t) = e^{-\nu t} m.$$

Infinite State Space

- P is now an infinite matrix, so Perron-Frobenius theory no longer applies.
- QSDs (if they exist) are given by *probability measures* m such that
$$mP_C(t) = e^{-\nu t} m.$$
- The δ_i-LCD (if it exists) is given by m such that
$$mP_C(t) = e^{-\lambda_C t} m.$$

Infinite State Space

- P is now an infinite matrix, so Perron-Frobenius theory no longer applies.
- QSDs (if they exist) are given by *probability measures* m such that

$$mP_C(t) = e^{-\nu t} m.$$

- The δ_i-LCD (if it exists) is given by m such that

$$mP_C(t) = e^{-\lambda_C t} m.$$

- These expressions involve P and λ_C, which are not known and impossible (or at best very difficult) to find analytically — we need conditions in terms of the q-matrix.

Infinite State Space

Theorem: If m is a ν-invariant probability measure for Q, then

$$\nu = \sum_{i \in C} m_i q_{i0}$$

is neccesary and sufficient for m to be a QSD.

Infinite State Space

Theorem: If m is a ν-invariant probability measure for Q, then

$$\nu = \sum_{i \in C} m_i q_{i0}$$

is neccesary and sufficient for m to be a QSD.

- This allows us to find *all* ν-invariant probability measures for Q which are QSDs.
- Another result tells us that a QSD must be ν-invariant for Q.

Infinite State Space

In order to find these ν-invariant measures for Q we must solve the system

$$\sum_{i \in C} m_i q_{ij} = -\nu m_j, \quad j \in C.$$

Infinite State Space

In order to find these ν-invariant measures for Q we must solve the system

$$\sum_{i \in C} m_i q_{ij} = -\left(\sum_{i \in C} m_i q_{i0}\right) m_j, \quad j \in C.$$

We can eliminate ν explicitly from the system we need to solve, however this renders the system non-linear in m.

Infinite State Space

In order to find these ν-invariant measures for Q we must solve the system

$$\sum_{i \in C} m_i q_{ij} = -\left(\sum_{i \in C} m_i q_{i0} \right) m_j, \quad j \in C.$$

We can eliminate ν explicitly from the system we need to solve, however this renders the system non-linear in m.

- Finding explicit expressions for QSDs is rarely possible.

Infinite State Space

Theorem: If the equations

$$\sum_{i \in C} y_i q_{ij} = \kappa y_j, \qquad y_i \geq 0,\ j \in C, \quad \sum_{i \in C} y_i < \infty$$

have only the trivial solution for some (all) $\kappa > 0$, then all ν-invariant probability measures for Q are also ν-invariant for P and are therefore QSDs.

Infinite State Space

Theorem: If the equations

$$\sum_{i \in C} y_i q_{ij} = \kappa y_j, \qquad y_i \geq 0,\ j \in C, \quad \sum_{i \in C} y_i < \infty$$

have only the trivial solution for some (all) $\kappa > 0$, then all ν-invariant probability measures for Q are also ν-invariant for P and are therefore QSDs.

- Call this condition the "Reuter FE Condition"

Infinite State Space

Theorem: If the equations

$$\sum_{i \in C} y_i q_{ij} = \kappa y_j, \qquad y_i \geq 0,\ j \in C, \quad \sum_{i \in C} y_i < \infty$$

have only the trivial solution for some (all) $\kappa > 0$, then all ν-invariant probability measures for Q are also ν-invariant for P and are therefore QSDs.

- Call this condition the “Reuter FE Condition”
- If this condition holds, all we have to do is find a ν-invariant measure for Q and this is a QSD.

A Minor Problem

Recall that for a ν-invariant measure for Q to exist, it is necessary that $\nu \in (0, \lambda_C]$. However, depending on the process, there are two situations that arise:

- There are finite ν-invariant measures for all $\nu \in (0, \lambda_C]$.
- There is only one finite ν-invariant measure; for $\nu = \lambda_C$.

This gives rise to some important questions:

Some Interesting Questions

When there is more than one QSD,

- For a given initial distribution a, which QSD is the a-LCD?
- For each QSD m, which initial distributions a have m as the a-LCD?

When there is only one QSD,

- Is it the a-LCD for *all* initial distributions a?
- or are there initial distributions for which there is no LCD?

Birth-Death Processes

Theorem: Suppose a birth-death process is absorbed with probability one. Then

- If $\mathcal{D} < \infty$ then there is a unique finite ν-invariant measure (QSD), corresponding to $\nu = \lambda_C$.
- If $\mathcal{D} = \infty$ then either
  - $\lambda_C = 0$ and there are no QSDs, or
  - $\lambda_C > 0$ and there is a one-parameter family of finite ν-invariant measures (QSDs), for $0 < \nu \leq \lambda_C$.

Here

$$\mathcal{D} = \sum_{n=1}^{\infty} \frac{1}{\mu_n \pi_n} \sum_{m=n}^{\infty} \pi_m, \qquad \pi_n = \frac{\lambda_1 \cdots \lambda_{n-1}}{\mu_2 \cdots \mu_n}.$$

The Chemical Reaction

$$\mathsf{A} + \mathsf{X} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} 2\mathsf{X}$$

$$\mathsf{X} \xrightarrow{k_2} \mathsf{B}$$

The birth and death rates are, respectively,

$$\lambda_i = \alpha k_1 i,$$

and

$$\mu_i = k_2 i + k_{-1} \frac{i(i-1)}{2}.$$

The Chemical Reaction

- One can show that this process is absorbed with probability one (and is therefore regular).

The Chemical Reaction

- One can show that this process is absorbed with probability one (and is therefore regular).
- We can also show that

$$\mathcal{D} = \sum_{n=1}^{\infty} \frac{n\Gamma(n+r)}{[nk_2 + n(n-1)\frac{k_{-1}}{2}](\alpha s)^{n-1}} \sum_{m=n}^{\infty} \frac{(\alpha s)^{m-1}}{m\Gamma(m+r)},$$

The Chemical Reaction

- One can show that this process is absorbed with probability one (and is therefore regular).
- We can also show that

$$\mathcal{D} = \sum_{n=1}^{\infty} \frac{n\Gamma(n+r)}{[nk_2 + n(n-1)\frac{k_{-1}}{2}](\alpha s)^{n-1}} \sum_{m=n}^{\infty} \frac{(\alpha s)^{m-1}}{m\Gamma(m+r)},$$

and that this is finite.

The Chemical Reaction

- One can show that this process is absorbed with probability one (and is therefore regular).
- We can also show that

$$\mathcal{D} = \sum_{n=1}^{\infty} \frac{n\Gamma(n+r)}{[nk_2 + n(n-1)\frac{k_{-1}}{2}](\alpha s)^{n-1}} \sum_{m=n}^{\infty} \frac{(\alpha s)^{m-1}}{m\Gamma(m+r)},$$

  and that this is finite.
- So there is a unique quasistationary distribution, which is limiting conditional (at least whenever the initial distribution has finite support).

A Connection

- For a Birth-Death process, the Reuter FE conditions hold iff $\mathcal{D} = \infty$.

A Connection

- For a Birth-Death process, the Reuter FE conditions hold iff $\mathcal{D} = \infty$.
- So, let's replace

$\mathcal{D}$ diverges (converges)

in van Doorns' result with

the Reuter FE condition holds (fails).

Infinite State Space

Conjecture: Suppose a process is absorbed with probability one. Then

- If the Reuter FE conditions fail then there is only one QSD.
- If the Reuter FE conditions hold, either
  - $\lambda_C = 0$ and there are no QSDs, or
  - $\lambda_C > 0$ and there is a one-parameter family of finite ν-invariant measures (QSDs), $0 < \nu \leq \lambda_C$.

Another Chemical Reaction

$$\mathrm{A} + \mathrm{X} \xrightarrow{k_1} 2\mathrm{X}$$

$$2\mathrm{X} \xrightarrow{k_2} \mathrm{B}$$

- This is not a Birth-Death process: it has jumps up of size 1, but jumps down of size 2:

$$q_{i,i+1} = \alpha i k_1,$$
$$q_{i,i-2} = k_2 \frac{i(i-1)}{2}.$$

- Hopefully my conjecture can deal with this!!

Further work

- Domain of attraction problem for LCDs.
- Conjecture — is it true? if not, what can we learn from a counterexample?
- Approximation methods: does $m^{(n)} \to m$ in some sense if we solve

$$\sum_{i=1}^{n} m_i^{(n)} q_{ij} = -\nu_1^{(n)} m_j^{(n)}, \quad j = 1, \ldots, n$$

  with $\nu_1^{(n)}$ the P-F maximal eigenvalue of $Q^{(n)} = (q_{ij},\ i, j = 1, \ldots, n)$, for succesively larger n?
- The 'renewal dynamical' approach.

The Quasi-Stationary Distribution