Continuous Time Markov Chains
As with the PP, we first define what we would like a CTMC to be, and then figure out if we can construct it...

**Defn (Continuous Time Markov Chain)** Let \( X \) be a countable set (the state-space). A \( X \)-valued stochastic process \( X = \{X(t)\}_{t \in \mathbb{R}^+} \) is called a CTMC if \( \forall t, s \geq 0 \), and \( \forall k \in \mathbb{N} \), \( 0 \leq t_1 < t_2 < \ldots < t_k < t \), and \( \forall x_1, x_2, \ldots, x_k, x, y \in X^{k+1} \),

\[
P[X(t+s) = y | X(t) = x, X(t_1) = x_1, \ldots, X(t_k) = x_k] = P[X(t+s) = y | X(t) = x]
\]

Moreover, the chain is said to be homogeneous if the RHS is independent of \( t \), i.e.,

\[
P[X(t+s) = y | X(t) = x] = P_s(x, y)
\]

Before proceeding, some comments:

1) Does a CTMC exist? Well, we already saw one! Check that if \( N(t) \) is a PP \( (\lambda) \), then \( N(t) \) is a CTMC on \( \mathbb{N} \).

In particular, \( P_s(x, y) = e^{-\lambda s} \left( \frac{\lambda s}{y-x} \right)^{y-x} \) \( \forall y > x \), and \( 0 \) \( \forall y < x \).

2) We use the notation \( P_s(x, y) \) to maintain a similarity to the notation for DTMCs, where we wrote \( P(x,y) \) for the transition probabilities. Note though that \( P_s \) is not one matrix but a fn of \( s \), and that \( P_s \neq P_S \) (unlike in a DTMC).

3) As with PP, we will see 2 main ways to construct a CTMC

i) Analytic - Via the "transition semigroup"

ii) Probabilistic - Via an embedded discrete-time 'jumpchain'
Given the above defn of a CTMC, let $P(t) = \{P_t(x,y)\}_{x,y \in X}$, where $P_t(x,y) \equiv \mathbb{P}[X(s+t) = y | X(s) = x] \quad \forall s,t \geq 0$

As with DTMCs, such a $P$ must satisfy consistent eqns:

1. (Chapman-Kolmogorov Eqns) For $t,s \geq 0$, and $x,y \in X$, we have $P_{t+s}(x,y) = \sum_{z \in X} P_t(z,z) P_s(z,y)$, or compactly $P_{t+s} = P_t \circ P_s$, $P_0 = I$

2. Moreover, let $\Pi(t) = \{\Pi_x(t)\}_{x \in X}$ be the distribution of $X(t)$. Then we have $\forall t > 0$, $\Pi(t)^T = \Pi(0)^T P_t$

The problem is that there is no particular $t \neq s$ $P_t$ can be used to determine $P_s$ for any $s \neq t$ (unlike in a DTMC, where $P_1 = P$ and $P_n = P^n \forall n \in \mathbb{N}$).

Instead we need to define $P_t$ in terms of an ‘infinitesimal generator’. We outline this first for some simple cases.

**Example (Poisson Process)** First consider $N(t) \sim \text{P}(\lambda)$. By defn, we have $P_t(x,y) = \mathbb{P}[N(t+s) = y | N(s) = x] \quad \forall t,s \geq 0$

$= e^{-\lambda t} \frac{(\lambda t)^y}{y!} 1\{y \geq x\}$

On the other hand, recall we defined a PP($\lambda$) via the eqns $\mathbb{P}[N(t+s) - N(t) = 1] = 2\lambda + O(s^2)$, $\mathbb{P}[N(t+s) - N(t) = 0] = 1 - 2\lambda + O(s^2)$ and $\mathbb{P}[N(t+s) - N(t) > 1] = O(s^2)$
Using this, we can write
\[
\Pr [N(t+\delta) = y] = \sum_x \Pr [N(t) = x] \Pr [N(t+\delta) = y | N(t) = x] \\
= \lambda \Pr [N(t) = y-\delta] + (1-\lambda) \Pr [N(t) = y] + O(\delta^2)
\]

Let \( \Pi_t(x) = \Pr [N(t) = x] \). Then we can write
\[
\Pi_{t+\delta}(x) - \Pi_t(x) = \Pi_t(x-1) \delta - \Pi_t(x) \delta + O(\delta^2)
\]
\[
\Rightarrow \frac{\Pi_{t+\delta}(x) - \Pi_t(x)}{\delta} = \lambda (\Pi_t(x-1) - \Pi_t(x)) + O(\delta)
\]
Taking limit \( \delta \to 0 \), we get the differential equation
\[
\Pi_t'(x) = \lambda \left( \Pi_t(x-1) - \Pi_t(x) \right) \quad \forall x \geq 1, \quad \Pi_t(0) = -\lambda \Pi_t(1), \quad \Pi_0(0) = 1
\]
One way to solve these is to first solve for \( \Pi_t(0) \) as
\[
\int_0^t d\Pi_t(0) = \int -\lambda dt \Rightarrow \Pi_t(0) = e^{-\lambda t}
\]
Next, for \( \frac{d\Pi_t(1)}{dt} = \lambda e^{-\lambda t} - \lambda \Pi_t(1) \), we can write it as
\[
e^\lambda \frac{d\Pi_t(1)}{dt} + \lambda e^{\lambda t} \Pi_t(1) = \frac{d}{dt} \left( e^{\lambda t} \Pi_t(1) \right) = \lambda, \quad \text{and} \quad \Pi_0(1) = 0
\]
Solving we get \( \Pi_t(1) = 2te^{-\lambda t} \). Moreover, we can continue this via induction to show \( \Pi_t(x) = \frac{e^{-\lambda t} \Gamma(x+1)}{x!} \)
The differential eqns can be written concisely as \( \frac{d\Pi_t}{dt} = \Pi_t T \), \( \Theta(x,y) = \begin{cases} 1 & y = x+1 \\ -2 & y = x \\ 0 & \text{otherwise} \end{cases} \)
The system has a unique solution \( \Pi_t(x) = \frac{e^{-\lambda t} \Gamma(x+1)}{x!} \) if \( \Pi_0(n) = 1 \) for \( n \geq 0 \).
Eq. (Flip-Flop chain) Let \( N(t) \) be the Poisson process, and define \( X(t) \in \{ -1, 1 \} \) as \( X(t) = X(0) (-1)^{N(t)} \), \( X(0) \) is on \( \{ -1, 1 \} \).

Now \( P_t(1,1) = \mathbb{P}[X(s+t) = 1 \mid X(0) = 1] = \mathbb{P}[N(t) \text{ is even}] = \sum_{k=0}^{\infty} \frac{e^{-2t}(2t)^{2k}}{(2k)!} = e^{2t}(\frac{e^{2t} + e^{-2t}}{2})^2 \)

Solving for \( P_t(1,1), P_t(1,-1) \) and \( P_t(-1,-1) \), we get that
\[
\begin{align*}
P_t &= \frac{1}{2} \begin{pmatrix} 1 + e^{-2at} & 1 - e^{-2at} \\ 1 - e^{-2at} & 1 + e^{-2at} \end{pmatrix}
\end{align*}
\]

Alternately, we can write for \( t, S > 0 \), and \( \Pi_t = \begin{pmatrix} \Pi_t(1,1) \\ \Pi_t(1,-1) \end{pmatrix} \)
\[
\begin{align*}
\Pi_{t+s}(-1) &= \Pi_t(-1)(1-2S) + \Pi_t(1) 2S + O(S^2) \\
\Pi_{t+s}(1) &= \Pi_t(-1)(2S) + \Pi_t(1)(1-2S) + O(S^2)
\end{align*}
\]

As before we can compute \((\Pi_{t+s}(x) - \Pi_t(x))/S\) and take \( \lim S \to 0 \) to get \( \frac{d\Pi_t}{dt} = \Pi_t^T \Theta \), where \( \Theta = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} \)

Solving this, we get \( \Pi_t^T = \Pi_0^T e^{\Theta t} \), where one can check via computing the eigenvalues of \( \Theta \) that \( e^{\Theta t} = \frac{1}{2} \begin{pmatrix} 1 + e^{2at} & e^{2at} \\ e^{2at} & 1 + e^{-2at} \end{pmatrix} \)

Thus in both cases, we managed to derive \( P_t \) by writing \( \frac{d\Pi_t}{dt} = \Pi_0^T \Theta \) and solving the system to get \( \Pi_t^T = \Pi_0^T P(t) \). Before formalizing this, we see another example, which generalizes the above 2.
Eg (Uniform CTMC) Let \( \{Y_n\}_{n \in \mathbb{N}} \) be a DTMC on countable state-space \( X \), with transition matrix \( K \), and let \( \{T_n\}_{n \in \mathbb{N}} \) be the arrival times of a PP(\( \lambda \)) process \( N(t) \) of rate \( \lambda \). Then the process \( X(t) = Y_{N(t)} \) is called a uniform MC with Poisson clock \( N(t) \) and subordinate chain \( Y_n \).

- Thus \( X(T_n) = Y_n \) for \( n \in \mathbb{N} \), and \( X(t) = X(t^-) \) if \( t \notin \{T_n\}_{n \in \mathbb{N}} \). Note also that \( T_n \) is not necessarily a discontinuity pt of \( X(t) \), since \( Y_n \) can equal \( Y_{n-1} \).

- Now we have \( P_t = \sum_{n=0}^{\infty} \left( e^{-\lambda t} (\lambda t)^n \right) . K^n \) 

- Note also that for \( t, s > 0 \), we have \( \forall y \in X \)

\[
\Pi_{t+s}(y) = 2s \left( \sum_{x \in X} \Pi_t(x) K(x,y) \right) + (1-2s) \Pi_t(y) + O(s^2)
\]

\[
\Rightarrow \frac{d}{dt} \Pi_t = \lim_{s \to 0} \frac{\Pi_{t+s} - \Pi_t}{s} = \lambda (K-I) \Pi_t
\]

Solving we get \( \Pi_t = \Pi_0 e^{\lambda t(K-I)} \), where we have \( e^{\lambda t(K-I)} = \sum_{n=0}^{\infty} e^{-\lambda t} (\lambda t K)^n = P_t \)

Now we try and formalize these ideas...
Defn (Stochastic Semigroup) \( \{P_t\}_{t \geq 0} \) is said to be a stochastic semigroup on \( X \) if \( \forall s,t \geq 0 \)

(i) \( P_t \) is stochastic matrix, i.e., \( \Sigma P_t(x,y) = 1 \ \forall x \in X \)

(ii) \( P_0 = I \), (iii) \( P_t+s = P_t P_s \ \forall s,t \geq 0 \)

* A stochastic semigroup \( P \) is called standard if it is continuous at the origin, i.e., \( \lim_{s \to 0} P_s = P_0 = I \) (pointwise convergence ?)

Then we have the following 2 properties:

1) \( P_t \) is continuous, i.e., \( \lim_{s \to 0} P_{t+s} = P_t \ \forall t \geq 0 \)

2) \( \forall x \in X, \mathbb{E} \Theta(x) = Q(x) \triangleq \lim_{s \to 0} \frac{1 - P_s(x,x)}{s}, \Theta(y) \triangleq \lim_{s \to 0} \frac{P_s(x,y)}{s} \)

The proof is purely analytical (see Brémaud, Ch8 Thm 2.1), and not crucial for our purposes, so we take it as a fact.

Defn (Infinitesimal Generator) - For a CTMC \( \{X(t)\} \) on \( X \) with stochastic semigroup \( P_t \), its infinitesimal generator is given by \( Q = \lim_{s \to 0} \frac{P_s - I}{s} \) (\( \therefore P_0 = I \))

* The generator \( Q \) is thus the derivative of \( P_t - I \) at \( t = 0 \), and can be found from \( P_t \). On the other hand, \( P_t \) can usually also be found from \( Q \) (as \( P_t = e^{tQ} \)) in most cases...
A birth-death process $X(t)$ is a CTMC taking values in $\mathbb{N}$ s.t. $\forall t, s > 0$ and $i \in \mathbb{N}$

$$P[X(t+s)=i+1|X(t)=i]=\lambda_i s + O(s^2)$$

$$P[X(t+s)=i-1|X(t)=i]=\mu_i s + O(s^2)$$

and all other transitions have probability $O(s^2)$

Given $\lambda_i, \mu_i, i \geq 1$, and $\lambda_0$, intuitively we would say $X(t)$ has a generator

$$Q(i,i) = \lambda_i, Q(i,i-1) = \mu_i, i \geq 1, Q(i,i) = 0 \forall i \neq i-1, i, i+1$$

However, if $\lambda_i, \mu_i \to \infty$ as $i \to \infty$, such a limit may not exist. We need to be somewhat careful dealing with such cases.

**Defn -** Consider semigroup $P_t$ with generator $Q = \lim_{s \to 0} \frac{P_s - I}{s}$

- $P_t$ is stable iff $(-Q(2)) = Q(2) = \lim_{s \to 0} \frac{1 - P_s(2)}{s} < \infty \forall x \in \mathcal{X}$

- $P_t$ is conservative iff $(-Q(2)) = Q(2) = \sum_{y \neq x} Q(x, y) \forall x \in \mathcal{X}$

- Note that for any $S$, by defn of the stock semigroup

$$\sum_{y \in \mathcal{X}} P_s(x,y) = 1 \Rightarrow \frac{1 - P_s(x,x)}{s} = \sum_{y \neq x} P_s(x,y)$$

Thus if $\lim_{s \to 0} \sum_{y \neq x} P_s(x,y) = \sum_{y \neq x} \lim_{s \to 0} P_s(x,y)$, then $P_t$ is stable and conservative. We assume henceforth that $Q$ is stable and conservative - note though that checking this for a CTMC (for example, a birth-death chain) is non-trivial.
Kolmogorov's Differential Equations

- Given a standard stochastic semigroup $P_t$, we can write

$$
\frac{P_{t+s} - P_t}{s} = P_t \frac{P_s - I}{s} = \frac{P_s - I}{s} P_t
$$

Assuming the limit $s \to 0$ exists, we get two systems of differential equations:

i) \( \frac{dP_t}{dt} = P_t \Theta \) (Forward diff system)

ii) \( \frac{dP_t}{dt} = \Theta P_t \) (Backward diff system)

In more detail, for any $x, y \in \mathcal{X}$, we have the differential equations:

Forward Eqs i) \( \frac{dP_t(x, y)}{dt} = \sum_{z \in \mathcal{X}} P_t(x, z) Q(z, y) = -P_t(x, y) Q(y) + \sum_{z \in \mathcal{X}} P_t(x, z) Q(z, y) \)

Backward Eqs ii) \( \frac{dP_t(x, y)}{dt} = \sum_{z \in \mathcal{X}} Q(z, x) P_t(z, y) = -Q(x) P_t(x, y) + \sum_{z \in \mathcal{X}} Q(z, x) P_t(z, y) \)

- If $\mathcal{X}$ is finite, then subject to $P(0) = I$, the above systems have a unique solution $P(t) = e^{t\Theta} = \sum_{k=0}^{\infty} \frac{t^k Q^k}{k!}$.

- For verifying this, the main thing to check is that $e^{t\Theta}$ is defined. For this, we have the following:

**Lemma** - For any $n \times n$ matrix $A$ with $A(i, j) \in \mathbb{R}$, and for all $t > 0$, the series $\sum_{k=0}^{\infty} \frac{t^k A^k}{k!}$ converges component-wise (i.e., for all $i, j \in [n]^2$)

**Proof** - Let $A_k(i, j) = (A^k)_{i, j}$, and define $\Delta = \max_{i, j} |A_1(i, j)|$. Check via induction that $|A_k(i, j)| \leq \Delta^k n^{-1}$. Hence $\forall i, j \in [n]^2$, we have $A_k(i, j) \frac{t^k}{k!} \leq \frac{1}{n} (n \Delta t)^k \frac{t^k}{k!} \Rightarrow (e^{At})_{i, j} \leq e^{n \Delta t} \frac{t^k}{k!}$.
What about when \( X \) is countable? This gets more technical, so we state the main results without proof.

**Theorem:** Let \( P_t \) be a standard stochastic semigroup.

i) If \( P_t \) is stable and conservative, then \( \frac{dP_t}{dt} = \Omega P_t \)

(i.e., we can take limits to get Kolmogorov’s backward system)

ii) If in addition \( \sum_{x \in \mathcal{X}} P_t(x,k) \mathcal{Q}(k) < \alpha \) for all \( x \in \mathcal{X} \), then also \( \frac{dP_t}{dt} = P_t \mathcal{Q} \) (i.e., Kolmogorov’s forward system is satisfied)

iii) Finally let \( T_t \) denote the distribution of \( X(t) \) at any \( t \geq 0 \). Assuming the above conditions, and also, that \( \forall t \geq 0 \) we have \( \sum_{x \in \mathcal{X}} \mathcal{Q}(x) T_t(x) < \alpha \). Then, we have

\[
\frac{dT_t}{dt} = T_t^{T} \mathcal{Q}, \text{ i.e., } \frac{dT_t(x)}{dt} = -T_t(x) \mathcal{Q}(x) + \sum_{y \neq x} T_t(y) \mathcal{Q}(y) \]

To summarize, assuming \( Q(x) < \alpha \) and \( Q(x) = \sum_{y \neq x} Q(x,y) \) (i.e., \( Q \) is stable and conservative), we can solve the backward equations to obtain \( P_t = e^{Qt} \). Thus \( Q \) is in a sense completely defines the CTMC.

- \( Q(x,y) \) is sometimes referred to as the transition rate from \( x \) to \( y \) (for \( x \neq y \)), as it represents the rate of probability flowing from \( x \) to \( y \) (i.e., \( P_0(x,y) = Q(x,y) s + o(s^2) \)). This can be represented by a transition rate diagram:

```
\[ \xrightarrow{Q_{xy}} \]
```

- \( Q_{xy} \)
**Defn (Irreducibility)** - A CTMC $X(t)$ with generator $Q$ is irreducible if $P_t(x,y) > 0$ for any $t > 0$, and all $x, y \in \mathcal{X}$.

- In fact, for any $x, y \in \mathcal{X}$, $P_t(x,y) > 0 \forall t$, or $P_t(x,y) = 0 \forall t$.

**Defn (Stationary Distribution)** - A stochastic vector $\pi$ (i.e., with $\pi(x) > 0 \forall x \in \mathcal{X}$ and $\sum_x \pi(x) = 1$) is a stationary distr of a CTMC $(Q, P_t)$ if $\pi^T P_t = \pi^T \forall t \geq 0$. Moreover, if $Q$ is stable and conservative, then $\pi$ satisfies the global balance eqn $\pi^T Q = 0$.

Given the above defn, we can state a convergence theorem for CTMCs.

**Thm (CTMC Convergence Theorem)** - For an irreducible CTMC $(P_t, Q)$,

1. If stationary dist $\pi$ exists, then it is unique, and moreover $P_t(x,y) \xrightarrow{t \to \infty} \pi(y) \forall x, y \in \mathcal{X}$.
2. If no stationary $\pi$ exists, then $P_t(x,y) \xrightarrow{t \to \infty} 0 \forall x, y \in \mathcal{X}$.

**Pf Sketch** (G&S, Ch6, Thm 21) - For any $k > 0$, define skeleton DTMC $Y_n = X(hk)$.

Note $Y_n$ is irreducible, positive recurrent (i.e., $X(t)$ irreducible and $P_t(x,y) \xrightarrow{t \to \infty} \pi(y)$) and aperiodic (i.e., $P_t(x,n) > 0$) $\Rightarrow Y_n$ has unique stationary distr $\pi(y)$, and $P_n(x,y) \xrightarrow{t \to \infty} \pi(y)$. Now consider $k, k' \in \mathbb{Q}$, since $k n = k' n'$ infinitely often $\Rightarrow \pi(n) = \pi(n')$. For any other $t \in \mathbb{Q}$, we can complete the proof via continuity arguments.

**Thm (CTMC Ergodic Theorem)** - For irreducible CTMC $X(t)$ with stationary distr $\pi$, we have $\lim_{t \to \infty} \frac{1}{t} \int_0^t f(x(s))ds = \sum_{x \in \mathcal{X}} f(x) \pi(x)$ a.s. $\forall f$ s.t $\mathbb{E}[|f(x)|] < \infty$.

**Pf Sketch** - Similar to DTMC (via renewal cycles).
CTMCs via Embedded Chains

• An alternate approach to constructing CTMCs is by constructing them from DTMCs. There are two ways to do this:

  i) The Jump Chain - This exists for any CTMC
  ii) The Uniformized Chain - This exists when $\sup_{x \in X} Q(x) < x$

The Jump Chain

• We construct a process $X(t)$ on some countable $X$, for $t \in \mathbb{R}_+$, as follows -
  - Start with a DTMC $\{Y_n\}_{n \in \mathbb{N}}$ on $X$, with $Y_0 \sim \text{Exp}$, and transition prob matrix $A = \{A(x, y)\}$. We assume that $A(x, x) = 0 \forall x \in X$, in other words, $Y_n$ has no self-loops, but always "jumps" to a new state.
  - Next suppose we are given a sequence $\{E_n\}_{n \in \mathbb{N}}$ of iid $\text{Exp}(1)$ r.v.s (inf of $Y_n$), and a function $\{T(x); x \in X\}$ of inverse holding times for each state. Essentially, whenever $X(t)$ reaches a state $x \in X$, we want it to stay there for time $W \sim \text{Exp}(\delta(x))$ before jumping to some $y \neq x$. 
We now construct the chain as follows.

- Let $X(0) = Y_0 \sim \Pi_0$ and $T_0 = 0$.
- Define $W_0 = E_0 / \lambda(Y_0) \sim \exp(\lambda(Y_0))$.
- Set $T_1 = T_0 + W_0$, and $X(T_1) = Y_1$.

Subsequently, for any $k \geq 1$, we define $W_k = E_k / \lambda(Y_k)$, $T_{k+1} = T_k + W_k$, and $X(T_{k+1}) = Y_{k+1}$.

Define $T_{\infty} = \lim_{k \to \infty} T_k$. Then we can write

$$X(t) = \sum_{k=0}^{\infty} Y_k \mathbb{1}_{\{t \in [T_k, T_{k+1})\}} \quad \forall t \in [0, T_{\infty})$$

- It is not hard to check that the above process is indeed Markovian. Moreover, we could also allow $\lambda(x) = 0$ to model absorbing states, or $\lambda(x) = \infty$ to model states visited instantaneously. For the following, however, we restrict to $\lambda(x) \in (0, \infty) \quad \forall x \in X$.

- One potential problem still is that $T_{\infty}$ could be finite (and hence $X(t)$ is only defined on $t \in [0, T_{\infty})$).

**Defn** - The process $X(t)$ is said to be explosive

if $\mathbb{P}_x[T_{\infty} < \infty] > 0$ for some $X(0) = x$ and regular

if $\mathbb{P}_x[T_{\infty} < \infty] = 0$ for all $X(0) = x \in X$.

As an example, consider a birth process with $\lambda(x) = x^2$. 
Thm - For any $x \in X$, given $\{Y_n\}_{n \in \mathbb{N}}$ as above

$$P_x[T_x < \infty] = P_x \left[ \sum \frac{1}{\lambda(Y_n)} < \infty \right]$$

In other words, $X(t)$ is regular iff $\sum \lambda(Y_n)^{-1} = \infty$ as.

Moreover, this holds whenever one of the following hold

i) $X$ is finite, ii) $X(x) \leq \rho < \infty \ \forall x \in X$

iii) Given $A \subset X$, the transient states of $Y_n$, we have $\forall x \in X, \quad P_x[Y_n \in A \forall n \in \mathbb{N}] = 0$

We first need a property of Exponential r.v.

Proposition - If $\{E_i\}$ are independent Exponential r.v.s.s.t $E_i \sim \text{Exp}(\lambda_i) \ \forall i \in \mathbb{N}$. Then

$$\sum_{n \in \mathbb{N}} E_n < \infty \ a.s. \iff \sum_{n \in \mathbb{N}} \lambda_i^{-1} < \infty$$

Pf of theorem - By construction, we have $T_x = \sum_{n \in \mathbb{N}} E_n / \lambda(Y_n)$

This is a sum of independent Exponential r.v.s and by the above prop,

$$P[T_x < \alpha | \{Y_n\}] = \begin{cases} 1 & \text{if } \sum \lambda(Y_n)^{-1} < \infty \\ 0 & \text{if } \sum \lambda(Y_n)^{-1} = \infty \end{cases}$$

Thus $P[T_x < \alpha] = P_x \left[ \sum \frac{1}{\lambda(Y_n)} < \infty \right]$

Now we want to verify the sufficient conditions
- For (i), note that $X$ finite means $\forall x \in X, \exists \tau < \infty$.
Thus it's enough to verify (ii).
- For (ii), we have $\sum_{n} P(\tau_n^{-1}) \geq \sum_{n} \tau_n^{-1} = \infty$.
- For (iii), suppose $\text{IP}[X_n \in A \forall n] = 0$ implies that
  $\exists$ some $x_0 \in X \mid A$ s.t. $x_0$ is hit infinitely often.
Suppose $y_n = x$ for some set $\mathcal{N}_j, j \in \{1, 2, \ldots\}$. Then
  $\sum_{\mathcal{N}_j} f(\tau_n^{-1}) \geq \sum_{\mathcal{N}_j} \mathcal{P}(\tau_n^{-1}) = \sum_{\mathcal{N}_j} \mathcal{P}(x_0) = \infty$.

Proposition: $\forall x, y \in X, t \geq 0$, we have
  $\mathcal{P}_t(x, y) = e^{-\lambda x t} \mathbb{I}_{x = y} + \int_0^t \lambda e^{-\lambda s} \sum_{z \neq z} A(x, z) \mathcal{P}_{s-t}(z, y) ds$.

Proof: $\mathcal{P}_t(x, y) = \text{IP}[X(t) = y \mid X(0) = x]$
  $= \text{IP}[X(t) = y, W_0 > t \mid X(0) = x] + \text{IP}[X(t) = y, W_0 \leq t \mid X(0) = x]$
Moreover, by construction, we have
  $\text{IP}[X(t) = y, W_0 > t \mid X(0) = x] = e^{-\lambda x t} \mathbb{I}_{x = y}$
and $\text{IP}[X(t) = y, W_0 \leq t \mid X(0) = x] = \sum_{z \neq x} \text{IP}[X(t) = y, Y_1 = z, W_0 \leq t \mid X(0) = x]$
  $= \sum_{z \neq x} \int_0^t e^{-\lambda z s} \mathcal{P}(z) A(x, z) \mathcal{P}_{s-t}(z, y) ds$. 
Now given the expression for \( P_t(x,y) \), we get the following:

i) \( \lim_{t \to 0} P_t(x,y) = \prod_{\omega \in \Omega} \mathbb{1}(\omega = y \circ \omega) \Rightarrow \lim_{t \to 0} P_t = I \Rightarrow P_t \) is standard two.

ii) \( P_t(x,y) = \exp\left[-tW(x) \right] \left[ \prod_{s} \mathbb{1}(\omega = y \circ \omega) \exp\left( \int_{s}^{t} W(x) \right) \exp\left( \sum_{\omega \in \Omega} A(\omega) P_t(\omega,y) \right) ds \right] \)

\[ \text{(substituting } u = t - s) \]

\[ \Rightarrow \frac{dP_t(x,y)}{dt} = -W(x) P_t(x,y) + \exp\left(-tW(x)\right) \exp\left(\sum_{\omega \in \Omega} A(\omega) P_t(\omega,y) \right) \]

\[ = \mathbb{1}(x)(-P_t(x,y) + \sum_{\omega \in \Omega} A(\omega) P_t(\omega,y)) \]

\[ \Rightarrow \frac{dP_t}{dt} = \Theta P_t, \quad \text{where } \Theta(x,z) = \begin{cases} -W(x), & z = x \\ \sum_{\omega \in \Omega} A(\omega) P_t(\omega,z), & z \neq x \end{cases} \]

\[ = \text{diag}(\mathbb{1}(x)) \left( \begin{array}{cccc} 0 & \cdots & 0 & -W(x) \\ 0 & \cdots & 0 & \sum_{\omega \in \Omega} A(\omega) P_t(\omega,x) \end{array} \right) \]

Moreover, \( \Theta(x,x) = -W(x) \neq 0 \), and \( \sum_{\omega \in \Omega} A(\omega) P_t(\omega,x) = -W(x) = -\Theta(x,x) \),

Thus, \( \Theta \) is stable and conservative.

iii) The formula for \( P_t(x,y) \) is based on conditioning on the first jump (and hence yields the backward DE). To get the forward DE, we need to condition on the last jump — for this to exist, we need the system to not be explosive. Now conditioning on the last jump, we get \( P_t(x,y) = \exp\left(-tW(x)\right) \mathbb{1}(x,y) + \int_{0}^{t} \exp\left(-sW(x)\right) P_{s}(x,y) \Theta(y,y) ds \)

Differentiate to derive the Kolmogorov forward equation.
iv) Given a (stable + conservative) infinitesimal generator $Q$, we can derive the jump chain parameters $(A, Y)$ as
\[
Y(x) = -Q(x, y) \quad \forall x \in X, \quad A(x, y) = \Theta(x, y) / \pi(x) \quad \forall x, y \in X
\]
Similarly, given a sample path $X(t)$ of a CTMC, we can obtain the subordinate jump DTMC by tracking the sequence of unique states of $X(t)$ (so $Y_0 = X(0)$, $T_1 = \inf\{t > 0 | X(t) \neq X(0)\}$, $Y_1 = X_{T_1}$, $T_2 = \inf\{t > T_1 | X(t) \neq X(T_1)\}$, $Y_2 = X_{T_2}$, ...). Moreover, the holding times $W_i = T_i - T_{i-1}$ can be used to give the underlying driving clock process as $\bar{E}_i = W_i$. $X(T_i)$

The Uniformized Chain

• Earlier we saw a uniform $X(t) = Y_{N(t)}$, where $Y_n$ is a DTMC with transition matrix $K$ (where $K(\cdot, x)$ can be $\geq 0$, i.e., self loops are allowed), and $N(t) \sim \text{PP}(\lambda)$. The associated stochastic semi-group is $P_t(x, y) = \sum_{n=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^n}{n!} K^n(x, y)$

• We can differentiate to check that
\[
\frac{dP_t(x, y)}{dt} = -\lambda e^{-\lambda t} 1_{\{x = y\}} + \sum_{n=1}^{\infty} \frac{(\lambda^2 t + n\lambda)(\lambda t)^{n-1}}{n!} e^{-\lambda t} K^n(x, y)
\]
Setting $t=0$, we get $\frac{dP_0(x, y)}{dt} = \lambda (K_{xy} - 1_{\{x = y\}})$, so $Q = \lambda (K - I)$

• To determine the jump chain $(A, \Theta)$ associated with the Uniform CTMC $(K, \lambda)$, we have $\Theta(x) = -Q(x, x) = \lambda (1 - K_{xx})$, $A(x, x) = 0$ and $A(x, y) = \Theta(x, y) / \Theta(x) = K_{xy} / \lambda (1 - K_{xx})$. Note that $A(x, y) = \mathbb{P}[Y_{t+1} = y | Y_t = x, Y_t+1 \neq x]$. 

\[\text{Note that } A(x, y) = \mathbb{P}[Y_{t+1} = y | Y_t = x, Y_t+1 \neq x] \]
Note that $-\Theta(x,y) = \chi(1-K(x,y)) \leq 2 \Rightarrow \sup_{x \in \mathcal{X}} (-\Theta(x,y)) \leq 2 < \infty$. Moreover, given any $\Theta$ sat $\sup_{x \in \mathcal{X}} (-\Theta(x,y)) < \infty$, we can obtain a uniform MC with $\lambda = \sup_{x \in \mathcal{X}} (-\Theta(x,y))$, $K(y|x) = 1 + \Theta(x,y)$, $K(y|x) = O(\lambda x^y) / \lambda$.

Thus CTMC $\Theta$ is uniformizable iff $\sup_{x \in \mathcal{X}} (-\Theta(x,y)) < \infty$.

Examples

i) Poisson process $PP(\lambda) = \mathcal{Q}(\lambda) = \lambda \Xi y = y+1 - \lambda \Xi y = y$

rate diagram

jump chain

uniformized chain

ii) Generalized flip-flop

rate diagram

jump chain

uniformized chain

iii) Generalized birth-death process - $Q(\lambda i, i+1) = \lambda_i$, $Q(\lambda i, i-1) = \mu_i$

rate diagram

jump chain

uniformized chain

$\lambda = \sup (\lambda_i + \mu_i) < \infty$
Finally we want to understand stationary distributions.

- Recall: Given CTMC \((P_t, Q)\), \(\Pi^T\) is a stationary distribution if
  i) \(\Pi^T P_t = \Pi^T \quad \forall \ t \geq 0\)

  ii) (Assuming \(Q\) is stable, conservative) \(\Pi^T Q = 0\)

The CTMC is ergodic if it is irreducible, and \(\Pi(x)>0\) for some \(x \in \mathcal{X}\).

We now want to relate this to the jump chain \((A, \lambda)\) and uniform chain \((K, \lambda)\).

Recall \(Q^j = D_0 (A-I)\) for the jump chain, \(Q^u = \lambda (K-I)\) for the uniformized chain. Thus the stationary distribution satisfies

- **Uniformized Chain** -

  \[
  \Pi^T \lambda (K-I) = 0 \Rightarrow \Pi^T K = \Pi^T
  \]

  In other words, stationary distribution of CTMC \(Q^u = \text{stationary dist of DTMKC} K\).

  Moreover CTMC \(X(t)\) with generator \(Q\) is ergodic \(\iff\ X(t)\) is uniformizable, with uniform chain \((\lambda, K)\), and \(K\) is ergodic.

- **Jump Chain** -

  \[
  \Pi^T D_0 (A-I) = 0 \Rightarrow \Pi(x) \delta(y) = \sum_{y \neq x} \Pi(y) \delta(y) A(y, z) \delta(z) \]

Moreover, the chain is ergodic if \(\sum_{x \in \mathcal{X}} \Pi(x) < \infty\).
To summarize the formulations of CTMCs -

<table>
<thead>
<tr>
<th>Generator</th>
<th>Jump Chain</th>
<th>Uniformized Chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q(x,y) = \lim_{s \to 0} P_s(x,y)/s$</td>
<td>$Q(x,y) = -\sum_{y \neq x} Q(y,x) &lt; \infty$</td>
<td>$(\text{Uniform}) \text{ rate - } \lambda$</td>
</tr>
<tr>
<td>$Q(x,x) = -\sum_{y \neq x} Q(y,x)$ (stable + conservative)</td>
<td>$\text{Holding rates - } P_i(x), x \in \mathbb{X}$</td>
<td>Subordinate DTMC - $A$</td>
</tr>
<tr>
<td>$dP_t = QP_t = P_t \Theta$</td>
<td>$\text{Subordinate DTMC - } A$</td>
<td>$\text{Subordinate DTMC - } K$</td>
</tr>
<tr>
<td>$\Pi^T \Theta = 0 \Leftrightarrow \Pi^T P_t = \Pi^T$</td>
<td>$D_f = \text{diag}(\varphi(x))$</td>
<td>$\Theta = \lambda (K - I)$</td>
</tr>
<tr>
<td>$\Pi^T \theta = \Pi^T_0 P_t$, $\frac{d}{dt} \Pi^T P_t = \Pi^T \Theta$</td>
<td>$Q = D_f (A-I)$</td>
<td>Only valid for $Q$ st.</td>
</tr>
<tr>
<td>$\Pi^T \Theta = 0$</td>
<td>$\text{Non-explosive} = \sup_{x \in \mathbb{X}}</td>
<td>Q(y,x)</td>
</tr>
<tr>
<td></td>
<td>i) Finite, ii) $\sup</td>
<td>Q(x,y)</td>
</tr>
<tr>
<td></td>
<td>iii) $A$ is ergodic</td>
<td>(i.e., $\Pi^T K = \Pi^T$)</td>
</tr>
</tbody>
</table>

**Example (General Birth-Death Chain)**

$X = \text{IN}_0$, in state $i$ new arrival after time $\text{Exp}(\lambda_i)$, departure after time $\text{Exp}(\mu_i)$

**Rate Diagram**

![Rate Diagram](null)

**Jump Chain**

$\delta(i) = \mu_i + \lambda_i \forall i$

**Uniformized Chain**

$\lambda = \sup (\lambda_i + \mu_i) < \infty$
Assuming this has a stationary dist \( \pi \), we can solve for it in 3 ways:

i) Using the rate matrix - \( \pi \) stationary if \( \sum_{x=0}^{\infty} \pi(x) = 1 \) and \( \pi^T Q = 0 \)

\[
\begin{pmatrix}
\pi(0) & \pi(1) & \cdots \\
\mu_1 & -\mu_1 & 0 \\
0 & \mu_2 & -\mu_2 & 0 \\
& & & \ddots
\end{pmatrix}
\]

Thus \( \pi(2) \mu_2 = (\lambda_i + \mu_i) \pi(i) - \lambda_i \pi(i) = \lambda_i \pi(i+1) + \mu_i \pi(i-1) \) \( \forall i \geq 1 \)

Now by induction, we can show \( \pi(i) \mu_i = \pi(i+1) \), \( \forall i \geq 1 \)

Let \( \alpha_0 = 1 \), \( \alpha_x = \frac{\lambda_x}{\mu_x} \) \( \Rightarrow \pi(x) = \pi(0) \cdot (\alpha_0 \alpha_1 \cdots \alpha_x) \forall x \in \mathbb{N}_0 \)

Thus \( \pi(x) = \frac{\alpha_0 \alpha_1 \cdots \alpha_x}{Z} \), where partition fn (renormalization) \( Z = \sum_{x=0}^{\infty} \beta_x \)

ii) Using the jump chain - \( \pi(x) \mathbb{P}(x) = \sum_{y} \pi(y) \mathbb{P}(y|x) \) \( \forall x \geq 0 \)

\( \pi(0) \alpha_0 = \mu_1 \pi(1) \), \( \pi(i) \) = \( \lambda_{i-1} \pi(i-1) + \mu_{i+1} \pi(i+1) \) \( \forall i \geq 1 \)

Now we solve as before to get \( \pi(x) = \frac{\beta_x}{Z} \), \( Z = \sum_{x=0}^{\infty} \beta_x \)

iii) Using the uniform chain - \( x \geq \max_{i \geq 0} \{ \mu_i + \lambda_i \} \) (assume \( \alpha < \infty \))

Moreover \( \pi(x) \mu_x/\lambda = \pi(x-1) \mu_{x-1}/\lambda \Rightarrow \pi(x) = \alpha_x \cdot \pi(x-1) = \prod_{i=0}^{x} (\alpha_i) \pi(0) \)

This follows from the balance eqns for DTMC \( K \) - Note also \( K \) is reversible

Again we get \( \pi(x) = \frac{\beta_x}{Z} = \frac{\prod_{i=0}^{x} (\alpha_i)}{Z} \), \( Z = \sum_{x=0}^{\infty} \beta_x \)

Thus in all cases we get \( \pi(x) = \frac{\beta_x}{Z} \), \( \beta_x = \prod_{i=0}^{x} (\frac{\alpha_i}{\mu_{i+1}}) \), \( Z = \sum_{x=0}^{\infty} \beta_x \)

\[
\begin{vmatrix}
1 & \frac{\alpha_0}{\mu_1} & \frac{\alpha_0 \alpha_1}{\mu_1 \mu_2} & \frac{\alpha_0 \alpha_1 \alpha_2}{\mu_1 \mu_2 \mu_3} + \ldots
\end{vmatrix} < \infty
\]
Finally, to guarantee that \( T \) exists, we need \( \nu + \text{non-explosiveness of the MC.} \) We can ensure this in two ways.

i) \( \sup_i (\lambda_i + \mu_i) < \alpha \) (ie, uniformizable chain), or

ii) The jump chain DMC A is positive recurrent: for this, we can check balance equations as above to get the cdf

\[
1 + \frac{2 \nu_1}{\nu_1} + \frac{\nu_1 \nu_2 \nu_3}{\nu_1 \nu_2 \nu_3} + \cdots < \alpha \implies \left(1 + \frac{2 \nu_1}{\nu_1} + \frac{\nu_1 \nu_2 \nu_3}{\nu_1 \nu_2 \nu_3} + \cdots \right) < \alpha
\]

This has many special cases

i) \( \lambda_i = \lambda, \mu_i = \mu \ \forall i > 0 \) \( \forall i \)

\[ M/M/1 \text{ queue} \]

\[
\delta_0 \xrightarrow{\mu} \delta_1 \xrightarrow{\lambda} 0 \xrightarrow{\mu} 1 \xrightarrow{\lambda} 2 \xrightarrow{\mu} \cdots \\
Z = \sum_{i=0}^{\infty} \frac{\mu^i}{\lambda^i} \quad \text{if } \rho < 1 \quad \text{and } \quad \prod(i) = (1 - \rho) \rho^i
\]

ii) \( \lambda_i = \lambda, \mu_i = (i \wedge k) \mu \ \forall i \)

\[ M/M/k \text{ queue} \]

\[
Z = \sum_{i=0}^{k-1} \frac{\mu^i}{\lambda^i} + \frac{\mu^k}{\lambda^k} \sum_{i=0}^{\infty} \frac{\rho^i}{k!} < \alpha \quad \text{if } \rho < k \mu \\
\prod(z) = \frac{\rho^x}{x!} \quad \text{if } z = k, \quad \prod(z) = \frac{\mu^k}{z!} \rho^x
\]

iii) \( \lambda_i = \lambda I \{i \leq k^2\}, \mu_i = i \mu I \{i \leq k^3\} \) \( \rho = 2 \)

\[ M/M/k/k \text{ queue} \]

\[
Z = \sum_{i=0}^{k^2} \frac{\mu^i}{\lambda^i} < \alpha \quad \forall \rho, \quad \prod(z) = \frac{\rho^x}{x!} \quad \forall x \leq k
\]

iv) \( \lambda_i = \lambda, \mu_i = i \mu I \{i > 0\} \) \( \rho = 2 \chi \)

\[ M/M/\infty \text{ queue} \]

\[
Z = \sum_{i=0}^{\infty} \frac{\mu^i}{\lambda^i} = e^{-\rho} < \alpha \quad \forall \rho
\]

\[
\prod(z) = \frac{e^{-\rho} \rho^x}{x!} = \text{Poi}(\rho)
\]
Reversibility

Birth-death chains are a special case of reversible CTMCs. To define these, consider any $T > 0$, and CTMC $X(t)$. The time reversed process $\tilde{X}(t) = X(t-T)$ is a CTMC on $[0, T]$ with semigroup $\tilde{P}$ satisfying $\forall x, y \in X$, the detailed balance eqn $\tilde{T}(x) \tilde{P}(x,y) = \tilde{T}(y) \tilde{P}(y,x)$. To avoid dependence on $T$, we extend $X(t)$ to negative time by defining $\{X(-t); t \geq 0\}$ as a CTMC with semigroup $\tilde{P}$.

Thm (CTMC Kelly's Lemma) - Let $X(t)$ be a regular DTMC with generator $Q$, and consider any dist $\Pi$. Let $\tilde{Q}$ be defined such that $\tilde{T}(x) \tilde{Q}(x,y) = \tilde{T}(y) \tilde{Q}(y,x)$ $\forall x, y \in X$, and $\tilde{Q}(x,x) = - \sum \tilde{Q}(x,y)$. If $\tilde{Q}(x,x) = Q(x,x)$ $\forall x \in X$, then $\Pi$ is the stationary distr of $Q$, and $\tilde{Q}$ generates the reverse-time process $\tilde{X}(t)$.

- The proof is similar to the DTMC case. Moreover, if $\tilde{Q} = Q$, then $X(t)$ is reversible (i.e., $X(t) \overset{\text{rev}}{=} \tilde{X}(t)$)
- Corr: A stationary birth-death process is reversible.

Recall $\Pi(x) = \frac{1}{2} \frac{\lambda_x}{\mu_x} \frac{2!}{\mu_x} \cdots \frac{x!}{\mu_x}$. Thus for any $x \geq 1$, we have $\Pi(x) p_x = \Pi(x-1) \lambda_{x-1}$, and thus it's reversible.
A more surprising and useful consequence of reversibility occurs when we consider birth-death chains where $\lambda_x = \lambda \forall x \geq 0$, i.e., all the birth rates are the same. This can be interpreted as saying that the births follow a $PP(\lambda)$ process, independent of the state. Assume also that the chain is ergodic, i.e., $\sum_{i=1}^{\infty} 2^i \rho_i < \infty$ (from $\bullet$)

**Thm (Burke’s Thm)** - Let $X(t)$ be a birth-death process with birth-rate $\lambda_x = \lambda \forall x \geq 0$, and let $A(s,t]$ and $D(s,t]$ denote the number of births (i.e., arrivals) and deaths (i.e., departures) in any interval $(s,t]$. Then

i) $\forall t$, $\{D(s,t], s \leq t\} \parallel X(t) \parallel \{A(t,u], u > t\}$

ii) The departure process is $PP(\lambda)

**Pf** - By reversibility, $X(t) \overset{d}{=} X(-t)$. Also the upward jumps in $X(t)$ (i.e., arrivals) form a $PP(\lambda)$, and are equal in distribution to the upward jumps in $X(-t)$, which correspond to departures in $X(t)$.

- Also by construction, $A(t,u] \parallel X(t)$ for any $0 \leq t < u$. These however are departures $D(-u,-t]$ for $X(-t)$. Thus the past departures $\parallel X(t)$
This now allows us to build complex networks of queues!

\[ \text{Eg (Tandem Queues)} - \ 2 \ M/M/1 \text{ queue is series} \]

\[ \begin{array}{c}
    2 \\
    \downarrow \scriptstyle{\mu_1} \\
    \square \\
    \downarrow \scriptstyle{\mu_2} \\
    X = N_0 X \ N_0
\end{array} \]

- Suppose \( \rho_1 = \frac{\lambda}{\mu_1} < 1 \Rightarrow X_1(t) \sim M/M/1, \lim_{t \to \infty} P[X_1(t) = n] = \rho_1^n (1 - \rho_1) \)

- By Burke's Thm, departures from 1=arrivals from 2 \( \sim \text{PP}(\lambda) \)

Now if \( \rho_2 = \frac{\lambda}{\mu_2} \), then \( X_2(t) \sim M/M/1, \lim_{t \to \infty} P[X_2(t) = n] = \rho_2^n (1 - \rho_2) \)

- Claim - Stationary dist of \( (X_1, X_2) \equiv \Pi((x, y), (x', y')) = \rho_1 x \rho_2 y (1 - \rho_1)(1 - \rho_2) \)

- Pf - Kelly's Lemma! Check \( \Pi((x, y), (x', y')) = \Pi((y, y'), (y', y')) \)

- Note - This does not mean \( X_1(t) \perp X_2(t) \). Rather, what it says is that they are independent under the stationary dist \( \Pi \).

Such a distribution is said to be product form.

\[ \text{Eg (Queue with Feedback)} - \ 2 \ M/M/1 \text{ queue returns w.p. } p. \]

\[ \begin{array}{c}
    \text{PP}(\lambda) \\
    \downarrow \scriptstyle{\lambda} \\
    X(t) \downarrow \scriptstyle{\mu} \\
    \downarrow \scriptstyle{1-p} \\
    M
\end{array} \]

Suppose \( X(t) \) converges to a stationary dist.

Then the 'steady-state' rate of arrivals \( \lambda \) must obey \( \lambda = \lambda + p \lambda \Rightarrow \lambda = \frac{\lambda}{1-p} \)

Now assume \( p = \frac{\lambda}{\mu(1-\rho)} < 1 \). Intuitively, \( \Pi((x, y)) = (1-p) x^y \quad \forall x \geq 0 \)

Again this is true - check by verifying reversibility!