Asymptotic Behavior. (Sec 6.6 in Durrett, 13.6 in Le Gall)

The theorem $E_x T_x \longrightarrow \frac{1}{\pi(y)}$ will do a lot of heavy lifting in this section.

We intend to understand $p^n(x,y)$ as $n \longrightarrow \infty$ for both recurrent and transient states y.

If y is transient

$$\sum_n p^n(x,y) < \infty \qquad p^n(x,y) \to 0 \text{ as } n \to \infty.$$

Recall we defined

transient : $\rho_{yy} < 1$ $\quad \mathbb{P}_y(T_y^1 < \infty) = \rho_{yy}$.

recurrent : $\rho_{yy} = 1$.

Thm 6.6.1 Suppose y is recurrent.

we need to include this for the case where you dont reach y from x

$$\frac{N_n(y)}{n} \to \frac{1}{\mathbb{E}_y T_y} \mathbb{1}_{\{T_y < \infty\}} \quad \mathbb{P}_x\text{-a.s.}$$

first return to y starting from x.

Pf.: OK. Lets start at y.

$T_y^k = R(k) =$ k^{th} return time to y.

$R(0) = 0$. $X_0 = y$ t_1, t_2 successive

Ex: return times are iid.

LLN $\quad \frac{R(k)}{k} \to \mathbb{E}[t_1] = \mathbb{E}_y T_y$.

(we have proved this before from Le Gall, it was Strong Markov)

If $N_n(y) = k$ then $T_y^k \leq n$ and $T_y^{k+1} > n$ (otherwise $N_y = k+1$)

# of returns up to time n.

$$\Rightarrow \quad \frac{R(N_n(y))}{N_n(y)} \leq \frac{n}{N_n(y)} < \frac{R(N_n(y)+1)}{N_n(y)+1} \cdot \frac{N_n(y)+1}{N_n(y)}$$

$$\Rightarrow \quad \lim_{n\to\infty} \frac{N_n(y)}{n} = \frac{1}{\mathbb{E}_y T_y}$$

if $x \neq y$. $N_n(y) = 0$ if $T_y^{(1)} = \infty$

If +ve recurrent, fraction of time spent at x is +ve. $\mathbb{E}_y T_y < \infty$

null recurrent, fraction of time spent at x is 0. $E_y T_y = +\infty$

Since

$$0 \leq \frac{N_n(y)}{n} \leq 1 \qquad \frac{(\#\text{ of visits to } y)}{n}$$

$$\Rightarrow \quad \frac{\mathbb{E} N_n(y)}{n} \longrightarrow \mathbb{E}_x\left[\frac{\mathbb{1}\{T_y < \infty\}}{E_y T_y}\right]$$

If $N_y^n = k$, then $T_y^k \leq n$ and $T_y^{k+1} > n$ (otherwise $N_y^n = k+1$)

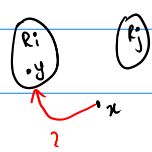
# of returns up to time n.

time of last return before n

$$\frac{R(N_y^n)}{N_y^n} \leq \frac{n}{N_y^n} \leq \frac{R(N_y^n+1)}{N_y^n} \cdot \frac{N_y^n+1}{N_y^n+1}$$

Use $\frac{R(k)}{k} \to E_y T_y$ To get $\lim_{n\to\infty} \frac{n}{N_y^n} \to E_y T_y$ (since $N_y^n \to \infty$ as y recurrent)

Suppose x is not in the same class as y. Then



$$P_x(T_y = +\infty) > 0 \Rightarrow E_x[T_y] = +\infty$$

and $N_y = 0$ on $T_y = +\infty$ Thus $\lim \frac{N_y^n}{n} \to 0 = \frac{1}{E_x[T_y]}$

But on the set $T_y < \infty$, we can apply SM again: For any bdd g,

check this pointwise equality

$$E_x\left[g(R(2)-R(1))\, 1_{\{T_y<\infty\}}\right] = E_x\left[g(R(1)\circ\theta_{T_y}\omega)\, 1_{\{T_y<\infty\}}\right]$$

$$= E_x\left[1_{\{T_y<\infty\}}\, E_{X_{T_y}}[g(R(1))]\right]$$

You can do this for $R(k+1)-R(k)$ as well.

Ex: Use the Strong Markov property to show that conditional on $\{T_y < \infty\}$, t_2, t_3 are iid and $P_x(t_k = n) = P_y(T_y = n)$ $k \geq 2$ $\quad t_k = R(k) - R(k-1)$

$$\frac{R(k)}{k} = \frac{t_1 + t_2 \cdots t_k}{k} \longrightarrow E_y T_y \quad P_x\text{-a.s.}$$

Thus $\frac{N_y^n}{n} \longrightarrow \frac{1}{E_y T_y}$ P_x a.s. on $\{T_y < \infty\}$

Since $0 \le \frac{N_y^n}{n} \le 1$, by BDD have

$$E_x \frac{N_y^n}{n} \longrightarrow \frac{P_x(T_y<\infty)}{E_y T_y}$$

$$\Rightarrow \quad \frac{1}{n}\sum_{m=1}^{n} p_m(x,y) \longrightarrow \frac{\rho_{xy}}{E_y T_y} \qquad — \star 1$$

Recall we proved early on $E_x N_y = \frac{\rho_{xy}}{1-\rho_{yy}}$.

$\star 1$ holds for recurrent y, but also for transient y since $E_y T_y = +\infty$

$\star 1$ shows $p_m(x,y)$ always converges in the Cesaro sense.

But $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ shows $P^{2k} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ $P^{2k+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

which does not converge.

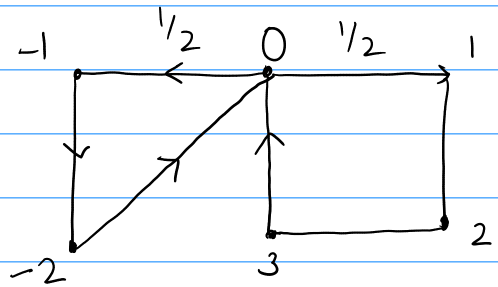
So there is convergence along a subsequence.
This is of course true more generally.

Define: A period.

$$I_x = \{n \geq 1 : p^n(x,x) > 0\}$$

$$d_x = \gcd(I_x)$$

Ex:



What is the period of 0 ?

$p^3(0,0) > 0$ and $p^4(0,0) > 0$

So $3,4 \in I_3$ and hence $d_0 = 1$.

<u>Define a period</u> :

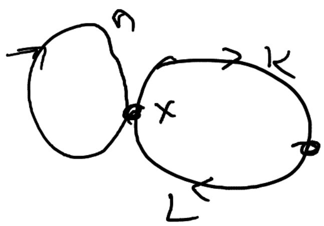
$$I_x = \{ n \geq 1 : p^n(x,x) > 0 \}.$$

$$d_x = \gcd(I_x).$$

<u>Lemma</u> 6.62. If $\rho_{xy} > 0$ then
$d_y = d_x$ (If $n \in d_x$, then $d_y \mid n$)

<u>Remark</u> : This shows that periods are again a class property.

<u>Pf</u> : $p^k(x,y) > 0 \quad p^L(y,x) > 0$



form a loop. (Irred.)

$\Rightarrow \; k+L \in I_y \Rightarrow d_y \mid k+L.$

if $n \in I_x \Rightarrow n+k+L \in I_y \Rightarrow$

$d_y \mid n+K+L \Rightarrow d_y \mid n$ since $d_y \mid K+L$

<u>Lemma 6.63</u> If $d_x = 1$ then $p^{(m)}(x,x) > 0$ for $m \geqslant m_0$.

Enough to show that if I_x contains 2 consecutive #s, $n, n+1$ then I_x contains n^2+k $\forall k$.

$k = gn + r$ where $0 \leq r \leq n$ (remainder thm.)

and $g \geqslant 0$

$$n^2+k = n^2+gn+r = r(n+1) + (n-r+g)n \quad !$$

add and subtract nr

n^2, $n(n-1)+n+1 = n^2+1$, ... and so on.

do n n times

$\emptyset^n$

do n, $n-1$ times and then an $n+1$

Its easiest to see this way:

$k, k+1 \in I_x$

$\Rightarrow \quad 2k, 2k+1, 2k+2 \in I_x$

$\Rightarrow \quad 3k, 3k+1, 3k+2, 3k+3$

$\vdots$

$\Rightarrow \quad k^2, k^2+1, \dots\dots, k^2+k-1 \in I_x$

$+k \quad +k \quad \dots\dots$

adding k will keep increasing

Next, to show that $\exists$ two consecutive integers in I_x.

Suppose $k = \inf\{m : \exists n_0, \text{ st } n_0, n_0+m \in I_x\}$

Assume $k \neq 1$ for contradiction.

<u>Since</u> $d_y = 1$, we must have an $n_1 \in I_x$

so $k \nmid n_1$. Write $n_1 = gk + r$ as before.

$0 < r < k$

$(g+1)(n_0+k) - [(g+1)n_0 + n_1]$ both numbers in I_x

$= (g+1)k - n_1$ of course

$= k - r < k$ (recall $r \neq 0$).

But then $(g+1)n_0 + n_1$ and $(g+1)(n_0+k) \in I_x$.

(Did I write this proof myself?)

Thm 13.27 (Le Gall): This is a nice theorem, gives the a.s. version of $p_n(x,y) \longrightarrow \pi(x)$. The advantage is that it just needs recurrence not +ve recurr.

Let P be recurrent irreducible and let ν be stationary. Let $\int f\,d\nu < \infty$ where $f : S \to \mathbb{R}_+$ be nonnegative. Then $\forall x \in S$

$$\frac{1}{n}\sum_{k=0}^{n} f(X_k) \longrightarrow \frac{1}{n}\frac{\int f\,d\nu}{\nu(x)} \qquad \mathbb{P}_x \text{ a.s}$$

(He states it as $\dfrac{\sum f(X_k)}{\sum g(X_k)} \longrightarrow \dfrac{\int f\,d\nu}{\int g\,d\nu}$)

Pf: Let $Z_0 = \displaystyle\sum_{i=0}^{T_x^{(1)}-1} f(X_i)$ $\qquad Z_k = \displaystyle\sum_{i=T_x^{(k-1)}}^{T_x^{(k)}-1} f(X_i)$

Then by S-Markov, Z_k are iid and

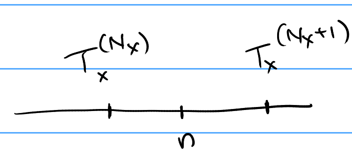
$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} Z_k \longrightarrow E_x[Z_0] = E_x\left[\sum_{n=1}^{T_x-1} f(X_n)\right]$$

$$= E_x\left[\sum_y \sum_{n=1}^{T_x-1} f(y)\, 1_{\{X_n=y\}}\right] = \sum_y f(y)\, E_x\left[\sum_{n=1}^{T_x-1} 1_{\{X_n=y\}}\right] \quad \text{(by definition)}$$

$$= \sum_y f(y)\,\mu_x(y) = \sum_y f(y)\, c\,\nu(y) = \frac{\int f\,d\nu}{\nu(x)}$$

Since $\mu_x(x) = 1 = c\,\nu(x) \Rightarrow c = \dfrac{1}{\nu(x)}$

Then, $$\frac{\sum_{i=0}^{N_x^n} Z_i}{N_x^n} \leq \frac{\sum_{i=0}^{n} f(X_i)}{N_x^n} \leq \frac{\sum_{i=0}^{N_x^n+1} Z_i}{N_x^n+1}\,\frac{N_x+1}{N_x}$$



Since $N_x \to \infty$ by recurrence, $$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} f(X_i)}{N_x^n} \longrightarrow \frac{\int f\,d\nu}{\nu(x)}$$

Cor: If x is positive recurrent, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(X_i) \to \frac{1}{E_x T_x} \frac{\int f \, d\nu}{\nu(x)}$$

$\sum_y \nu(y) = c$

$\Rightarrow E_x T_x = \frac{\sum_y \nu(y)}{\nu(x)}$

$$= \int f \, d\pi$$, where π is the stationary meas.

And moreover if $f = 1_y(x)$ Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i = y\}}(\omega) \to \pi(y) \qquad P_x \text{ a.s.}$$

If x is null recurrent then $E_x T_x = +\infty \Rightarrow \frac{N_x^n}{n} \to 0$

$$\Rightarrow \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(X_i) = \lim_{n\to\infty} \frac{N_x^n}{n} \lim_{n\to\infty} \frac{1}{N_x^n} \sum_{i=1}^{n} f(X_i) = 0$$

This is IMPORTANT. It means you can find rates of convergence and the like using standard MC techniques.

Thm: Assume chain is irreducible, positive recurrent and aperiodic. Let π be the unique invariant prob. meas. Then $\forall\ x \in S$

$$\sum_{y \in E} |P_x(X_n = y) - \pi(y)| \longrightarrow 0$$

(Durrett states $\quad |P_x(X_n = y) - \pi(y)| \longrightarrow 0$)

Also stated as $\quad \| P_n - \pi \|_{TV} = \sup_A |P_x(X_n \in A) - \pi(A)| \longrightarrow 0$

Durrett's proof in version 5a is fine (The old proof seemed wrong)

It is easy to see that $\pi \times \pi$ is a stationary distribution for $S \times S$.

But if $\bar{p}$ is not irreducible, then things can go wrong for us, even though every site is recurrent

Ex: Tell me what goes wrong if $\bar{p}$ is not irreducible even though $\pi(x)\pi(y) > 0 \; \forall x, y$

Lemma: $\bar{p}$ is irreducible.

Fix (x,y) and (a,b). We certainly have K, L st
(a,b) $\quad p^K(x,a) > 0$ and $p^L(y,b) > 0$

$\boxed{\text{Since } d_x = 1,}$ $\quad p^{L+M}(a,a) > 0$ and $p^{K+M}(b,b) > 0$
$\forall M > m_0$.

$$\bar{p}^2((x,y),(a,b)) = \sum_{(m,n)} \bar{p}((x,y),(m,n))\,\bar{p}((m,n),(a,b))$$

$$= \sum_{m,n} p(x,m)\,p(y,n)\;p(m,a)\,p(n,b)$$

$$= p^2(x,a)\,p^2(y,b)$$

Ex: Generalize to $\bar{p}^J((x,y),(a,b)) = p^J(x,a)\;p^J(y,b)$

Then $\bar{p}^{K+L+M}((x,y),(a,b))$

$$= p^{K+L+M}(x,a)\, p^{K+L+M}(y,b)$$

$$> p^{K}(x,a)\, p^{L+M}(a,a)\, p^{L}(y,b)\, p^{K+M}(b,b) > 0$$

Step: Notice $\bar{\pi}(x,y) = \pi(x)\pi(y)$ is a stationary distribution.

$$\bar{\pi}(x,y) = \sum \bar{p}((a,b),(x,y))\,\bar{\pi}(a,b)$$ To check.

$$= \sum_{a,b} p(a,x)\,p(b,y)\,\pi(a)\,\pi(b)$$

$$= \pi(x)\,\pi(y)$$ ✓

let $T =$ first hitting time of diagonal $\{(y,y) : y \in S\}$

Start (X_n) at x. and (Y_n) with stationary measure. Then $\bar{P} = P_{\delta_{x_0}, \pi}$ is a Markov process that is irreducible and recurrent (since $\bar{\pi}$ exists)

with transition $\bar{p}$

On $\{T \leq n\}$ X_n and Y_n have the same distribution:

$$\bar{P}(X_n = y, T \leq n) = \sum_{m=1}^{n} \sum_{x} \bar{P}(T = m, X_m = x, X_n = y)$$

$$= \sum_{m=1}^{n} \sum_{x} \bar{P}(T = m, X_m = x)\, \bar{P}(X_n = y \mid X_m = x)$$

$\{T = m, X_m = x\} = \{T_{(x,x)} = m\}$ $= \{T = m, Y_m = x\}$ (surprisingly indep of initial distribution) since $\bar{p}$ same

$$= \sum_{m=1}^{n} \sum_{x} \bar{P}(T = m, Y_m = x)\, P(Y_n = y \mid Y_m = x)$$

$$= \sum_{m=1}^{n} \sum_{x} \bar{P}(T = m, Y_m = x, Y_n = y)$$

$$= \bar{P}(Y_n = y, T \leq n)$$

Thus $|\bar{P}(X_n = y) - \bar{P}(Y_n = y)| = |p_n(x,y) - P_\pi(Y_n = y)|$ marginals.

$= |p_n(x,y) - \pi(y)|$. But also from previous

$= |\bar{P}(X_n = y, T > n) - \bar{P}(Y_n = y, T > n)|$.

$\leq \quad 2\, \bar{P}(T > n)$

But $\overline{P}$ is an irreducible, recurrent chain, $\Rightarrow$ $\overline{P}(T>n) \to 0$ as $n \to \infty$.

This is a very nice proof. Le Gall's proof is the same, more-or-less.

Durrett mentions the true coupling proof.

$$q((x_1, y_1), (x_2, y_2)) = \begin{cases} p(x_1,x_2)\,p(y_1,y_2) & \text{if } x_1 \neq y_1 \\ p(x_1,x_2) & \text{if } x_1 = y_1 \quad x_2 = y_2 \\ 0 & \text{otherwise.} \end{cases}$$

Once the processes meet, then they're COUPLED. Once again

$$Q(X_n = y) = p_n(x,y) \qquad Q(Y_n = y) = P_\pi(Y_n = y)$$

$$|Q(X_n = y) - Q(Y_n = y)| \leq 2\,Q(T' > n)$$

But $T' \stackrel{d}{=} T$ and thus this $\to 0$ by recurrence.

How does one upgrade this to $\|P_x(X_n \in \cdot) - \pi\|_{TV} \leq \sum_y |P_x(X_n = y) - \pi(y)|$

Recall we had said: $\leq 2\,\overline{P}^{x \times \pi}_{\delta}(T > n)$

$$\sum_y P_x(X_n = y) - \pi(y) = \sum_y \overline{P}(T>n, X_n = y) - \overline{P}(T>n, Y_n = y)$$

$$= \sum_y \overline{E}\left[1_{\{T>n\}}\left(1_{\{X_n = y\}} - 1_{\{Y_n = y\}}\right)\right]$$

$$\Rightarrow \sum_y |P_x(X_n = y) - \pi(y)| \leq \sum_y E\left[1_{\{T>n\}}\left(1_{\{X_n = y\}} + 1_{\{Y_n = y\}}\right)\right]$$

But $\sum_y \left(\mathbb{1}_{\{X_n = y\}} + \mathbb{1}_{\{Y_n = y\}} \right) \leq 2$ They're each bounded by 1

$$\Rightarrow \quad \| P_x (X_n \in \cdot) - \pi(\cdot) \|_{TV} \leq 2 \, \overline{P} (\tau > n)$$

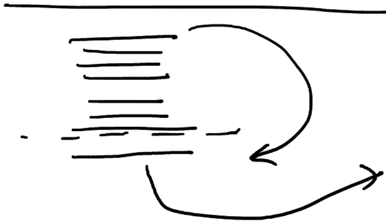
This is a beautiful way of doing it.

Finite chains

Exer: One can show
$$\mathbb{P}(\tau > n) \leq C e^{-kn}$$
(geometric decay).

HOWEVER, not true for ∞ chains.

Example: n cards, shuffle by taking top card and pushing it to above one of the $n-1$ cards underneath it.



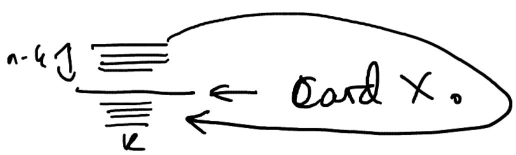
Durrett claims: you can tell that the chain has reached stationarity when the original bottom card has had about n

This is a distinguished bottom card

cards inserted underneath it.

Durrett claims that underneath the distinguished bottom card, if there are k cards, then all k! arrangements equally likely.

So you want to see how long it's going to take for the bottom card to bubble up to the top



chance is $\frac{k}{n-1}$ of picking card X.

Let T_1 = 1 card inserted under card X.

T_k = k^{th} card inserted under card X

$T_k - T_{k-1} = \tau_k$ are independent

$\tau_k \sim \text{Geom}\left(\frac{k}{n-1}\right)$.

$$E\, T_n = E \sum_{k=1}^{n-1} \tau_k = \sum_{k=1}^{n} \frac{n-1}{k} \sim n \log n$$

So $\frac{T_n}{n \log n} \to 1$ in probability.

Note that this is a transposition walk on permutations.

$\pi \in S_n$

Let $R = \{ \tau_{(0,k)} \}$ ← $k = \cancel{0}1, \dots n-1$

Each time we hit π with a transposition.

Nice!

Maybe its a way to show universality.

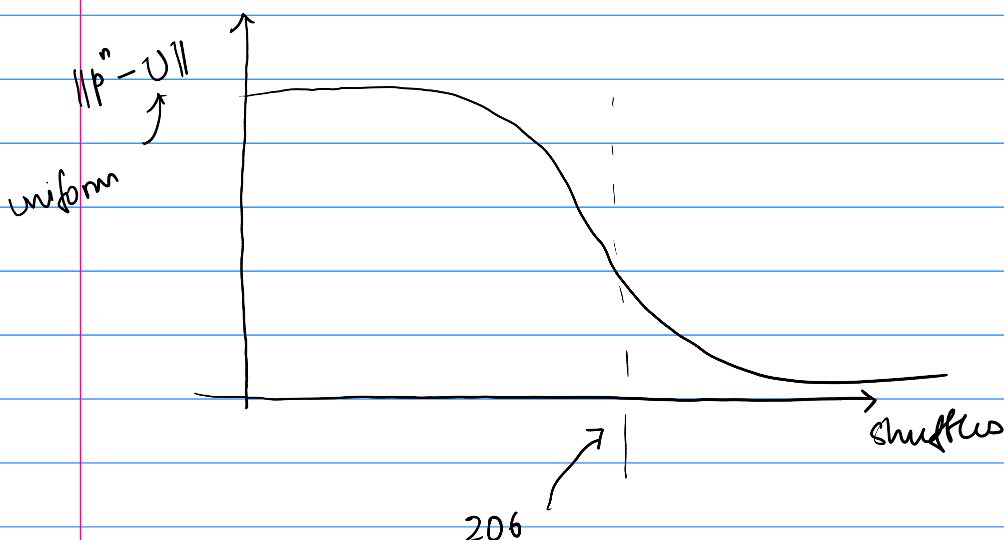
This argument and more can be found in Aldous-Diaconis, 1986

In the trivial shuffle, you would need

$$52 \log 52$$

shuffles to get close to uniform. This is about 206 shuffles.

These shuffles display a 'cutoff phenomenon'



What if we used a more elaborate shuffle, the common riffle shuffle:

1) Cut the deck of cards approximately in half
2) Interleave the cards together

Aldous - Diaconis show that $T_n \sim \frac{3}{2} \log n$

which shows $T_{52} \sim 6$