Aspects of the Karlin-McGregor Theorem

Ryan McCorvie[*]

UC Berkeley

December 1st, 2016

Contents

1 Overview 1

2 Main Theorem 1

3 Asymptotics for non-colliding Brownian motion 4

4 Doob's h-transform 6

References 9

1 Overview

In this survey, we examine some consequences of the Karlin-McGregor theorem first proved in [KM59]. The Karlin-McGregor theorem is to provide an expression for the transition kernel of a set of Markov processes which stop in the case they are ever coincident. We use the formula to analyze the asymptotic probabilities of no collision in the case of Brownian motion, and we analyze the related process which is conditional on no collisions.

2 Main Theorem

Consider a time-homogeneous Markov process on Ω with transition kernel $p_t(x, dy)$. For $n \in \mathbb{N}$, let $X_t = (X_t^1, \ldots, X_t^n)$ be the Markov process on Ω^n formed by taking n independent copies. The kernel for X_t is given by $\pi_t(x, dy) = \prod_{i=1}^{n} p_t(x_i, dy_i)$.

We say X_t is *coincident* at time t if two of the independent process components have the same value, that is if $X_t^i = X_t^j$ for some $i \neq j$. Our ultimate

[*]mccorvie@gmail.com

aim is to analyze the paths which are never coincident for any time t, which is informally to say they never collide. Let S_n be the symmetric group on n letters. Any permutation $\sigma \in S_n$, acts on Ω^n by permuting the coordinates $\sigma : (x_1, \ldots, x_n) \hookrightarrow (x_{\sigma 1}, \ldots, x_{\sigma n})$. Let $C \subset \Omega^n$ be the set of x with at least one pair of coincident coordinates, $C = \{x \in S \mid x_i = x_j \text{ for some } i \neq j\}$. Clearly C is closed, so we can define the stopping time $\tau = \inf\{t \mid X_t \in C\}$.

2.1 proposition. *Let $A^*(x, E)$ be the event that $X_0 = x$, $X_t \in E$ and X_s is never coincident for $0 \le s \le t$. Consider the quantity*

$$\pi^*_t(x, dy) = \begin{vmatrix} p_t(x_1, dy_1) & \ldots & p_t(x_n, dy_1) \\ \vdots & \ddots & \vdots \\ p_t(x_1, dy_n) & \ldots & p_t(x_n, dy_n) \end{vmatrix} = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \; \pi_t(x, \sigma\, dy) \qquad (1)$$

The following relation holds

$$\sum_{\sigma \in S_n} \operatorname{sgn} \sigma \Pr(A^*(x, \sigma E)) = \int_E \pi^*_t(x, dy) \qquad (2)$$

Proof. The idea of the proof is to use the strong Markov property, the reflection principle on paths with collisions, and the antisymmetry of $\pi^*_t(x, dy)$. Using linearity and the fact $\Pr(A(x, E)) = \int_E \pi(x, dy)$

$$\sum_{\sigma \in S_n} \operatorname{sgn} \sigma \Pr(A(x, \sigma E)) = \sum_{\sigma \in S_n} \operatorname{sgn} \sigma \int_E \pi_t(x, \sigma\, dy) = \int_E \pi^*_t(x, dy) \qquad (3)$$

Since $A^*(x, E) = A(x, E) \cap \{\tau > t\}$, consider $B(x, E) = A(x, E) \setminus A^*(x, E) = A(x, E) \cap \{\tau \le t\}$, which is the set of paths which stat at x, end up in E, but have at least one collision. The proposition is equivalent to the fact

$$\sum_{\sigma \in S_n} \operatorname{sgn} \sigma \Pr(B(x, \sigma E)) = 0 \qquad (4)$$

since $\Pr(A(x, E)) = \Pr(A^*(x, E)) + \Pr(B(x, E))$

For each coincident point $x \in C$ we can associate a unique transposition $\lambda = (i, j)$ with $i < j$ such that all of the $x_1, \ldots, x_{j-1}$ are distinct but $x_i = x_j$. This allows us to decompose C into disjoint sets $C = \coprod_{\lambda \in \Lambda} C_\lambda$ where C_λ consists of all the points associated with λ and Λ is the set of all transpositions. Paths starting at $x \in C_\lambda$ can be reflected via λ to give a one-to-one correspondence. Thus the kernel satisfies $\pi_t(x, E) = \pi_t(x, \lambda E)$ for all $x \in C_\lambda$

Let $\Phi(s) = \Pr(\tau \le s)$. Using the strong Markov property we can condition paths on τ and X_τ to write

$$\Pr(B(x,E)) = \int\limits_{[0,t]} d\Phi(s) \int\limits_{C} \pi_s(x,y)\pi_{t-s}(y,E)\,dy \tag{5}$$

Therefore

$$\begin{aligned}
\sum_{\sigma\in S_n} \operatorname{sgn}\sigma \Pr(B(x,\sigma E)) &= \sum_{\sigma\in S_n}\sum_{\lambda\in\lambda} \operatorname{sgn}\sigma \int d\Phi(s) \int\limits_{C_\lambda} \pi_s(x,y)\pi_{t-s}(y,\sigma E)\,dy \\
&= -\sum_{\sigma\in S_n}\sum_{\lambda\in\lambda} \operatorname{sgn}\lambda\sigma \int d\Phi(s) \int\limits_{C_\lambda} \pi_s(x,y)\pi_{t-s}(y,\lambda\sigma E)\,dy \\
&= -\sum_{\sigma\in S_n} \operatorname{sgn}\sigma \Pr(B(x,\sigma E))
\end{aligned}$$

Since the expression on the left is equal to its negative, its zero. □

Now specialize to the case that $\Omega \subset \mathbb{R}$ and suppose whenever $X_s = a$ and $X_t = c$ then for every $b \in [a,c] \cap \Omega$ there is a $u \in [s,t]$ such that $X_u = b$ almost surely. We call this the *intermediate value property*. Any process, such as a diffusion, which is pathwise continuous satisfies this property. However, this property also shared when Ω is discrete but X_t transitions only to neighboring sites, such as is the case for a birth and death process.

Consider the *Weyl chamber* $W = \{x \in \mathbb{R} \mid x_1 > x_2 > \cdots > x_n\}$. If X_t satisfies the intermediate value property and it starts in W but ends outside, then it must cross C at be coincident. Note that copies of the Weyl chamber partition the whole space and $\mathbb{R}^n \setminus C = \coprod_{\sigma\in S_n} \sigma W$, so each non-coincident process is restricted to its copy of the Weyl chamber.

2.2 COROLLARY (Karlin-MacGregor). *Suppose X_t satisfies the intermediate value property, and that $x \in W$*

$$\Pr(A^*(x,E)) = \int\limits_{E\cap W} \pi^*_t(x,dy) \tag{6}$$

Proof. First $\Pr(A^*(x,E)) = \Pr(A^*(x,E\cap W))$ since, by the intermediate value property, a path starting inside W and ending outside W must have a coincident point. Assuming $E \subset W$, similar reasoning shows $\Pr(A^*(x,\sigma E)) = 0$ for $\sigma \neq 1$ since σE is disjoint from W. Therefore all the terms on the left hand side of (2) are zero except for $\sigma = 1$. □

This corollary shows that π^*_t is the kernel for the stopped Markov chain $X^*_t = X_{t\wedge\tau}$ with $x \in W$. Its trivial to extend this kernel to the boundary $\partial W \subset C$ since points there remain constant for all time with probability 1.

In the case of a diffusion, let the function $f(t,y) = p_t(x,y)$ be the fundamental solution of the generator L subject to $f(0,y) = \delta_x$. In this case, equation (6) can be viewed as an application of the method of images where the solution $f^*(t,y) = \sum \operatorname{sgn}\sigma\, f(t,\sigma y)$ satisfies $f(t,y) = 0$ on the boundary ∂W.

3 Asymptotics for non-colliding Brownian motion

Now let's specialize to the case that X_n is Brownian motion in $\mathbb{R}^n$ and where the transition of each component is given by

$$p_t(x,y) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}\right) \tag{7}$$

Thus, in the Karlin-MacGregor determinant (6), its possible to factor $e^{-y_i^2/2t}$ from each row and $e^{-x_j^2/2t}$ from each column to get an expression

$$\pi_t^*(x_i,y_j) = (2\pi t)^{-n/2} e^{-\|x\|^2/2t} e^{-\|y\|^2/2t} \det e^{x_i y_j/t} \tag{8}$$

In what follows we'll make use of the *Vandermonde determinant*, given by

$$\Delta(x) = \begin{vmatrix} x_1^{n-1} & x_1^{n-2} & \dots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \dots & x_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n^{n-1} & x_n^{n-2} & \dots & x_n & 1 \end{vmatrix} = \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \prod_{k=1}^{n} x_k^{n-\sigma(k)} = \prod_{i<j}(x_i - x_j) \tag{9}$$

Also, with a slight abuse of notation, let

$$\Delta_n = \Delta(n-1, n-2, \dots, 1, 0) = \prod_{0 \le j < i \le n-1} (i-j) = \prod_{k=1}^{n-1} k! \tag{10}$$

3.1 proposition. *As $t \to \infty$, the law for noncolliding Brownian motion X_t^* converges to*

$$\Pr(X^*(t) \in dy) \sim \frac{1}{(2\pi)^{n/2}\Delta_n}\, t^{-n^2/2} \Delta(x)\Delta(y) e^{-\|y\|^2/2t} \tag{11}$$

Furthermore

$$\Pr(\tau > t) \sim C\,\Delta(x)\, t^{-n(n-1)/4} \tag{12}$$

Proof. First we show the probability that any $y_i > t^{1/2+\epsilon}$ is negligible for any small $\epsilon > 0$. The probability that the stopped diffusion satisfies $y_i > t^{1/2+\epsilon}$

is certainly less than the probability the normal diffusion does. And, for fixed x and large enough t, the probability a diffusion starting at x ends up at $y > t^{1/2+\epsilon}$ is less than the probability for a diffusion starting at 0 ends at $y > t^{1/2+\epsilon'}$ for $0 < \epsilon_0 < \epsilon$. So we can use the sum bound to find

$$\Pr(\bigcup_i \{X_t^{*i} > t^{1/2+\epsilon}\}) \leq \frac{n}{\sqrt{2\pi t}} \int_{t^{1+\epsilon_0}}^{\infty} \exp(-z^2/2t)\, dz \leq \frac{n \exp(-\frac{1}{2} t^{2\epsilon_0})}{3t^{\epsilon_0}} \tag{13}$$

its clear this expression goes to 0 much faster than $t^{-n(n-1)/4}$.

Now analyze the asymptotics of each term in (8). For fixed x, $e^{-\|x\|^2/2t} \to 0$ as $t \to \infty$. So the key is to approximate the last term $\det e^{x_i y_j/t}$. Consider an entry-by-entry approximation by the Taylor series

$$M_N(t,x,y) = \left(\sum_{k=0}^{N-1} \frac{x_i^k y_j^k}{k! t^k} \right)_{i,j} \tag{14}$$

Using multilinearity row-by-row, its not hard to see $\det M_N(t,x,y) = \det e^{x_j y_i/t} + O((y/t)^N)$. It stands to reason that for N large enough, $\det M_N$ has the right asymptotics when $\max y_i < t^{1/2+\epsilon}$.

In fact we'll take $N = n$ which is the smallest N with $\det M_N \neq 0$ since for any smaller N, the columns are collinear. Observe first that

$$M_n = \begin{pmatrix} \frac{x_1^{n-1}}{(n-1)!} & \frac{x_1^{n-2}}{(n-2)!} & \cdots & x_1 & 1 \\ \frac{x_2^{n-1}}{(n-1)!} & \frac{x_2^{n-2}}{(n-2)!} & \cdots & x_2 & 1 \\ \vdots & & \ddots & & \vdots \\ \frac{x_n^{n-1}}{(n-1)!} & \frac{x_n^{n-2}}{(n-2)!} & \cdots & x_n & 1 \end{pmatrix} \begin{pmatrix} \frac{y_1^{n-1}}{t^{n-1}} & \frac{y_2^{n-1}}{t^{n-1}} & \cdots & \frac{y_n^{n-1}}{t^{n-1}} \\ \frac{y_1^{n-2}}{t^{n-2}} & \frac{y_2^{n-2}}{t^{n-2}} & \cdots & \frac{y_n^{n-2}}{t^{n-2}} \\ \vdots & & \ddots & \vdots \\ y_1 & y_2 & \cdots & y_n \\ 1 & 1 & \cdots & 1 \end{pmatrix} \tag{15}$$

The first term has common factors in each row, the second has common factors in each colmumn. Pulling those factors out leaves a Vandermonde matrix $V(x)$ in x and the transpose $V(y)^\intercal$ in y.

$$\begin{aligned} \det M_n &= t^{-n(n-1)/2} \prod_{k=1}^{n-1} \frac{1}{k!} \det V(x) \det V(y)^\intercal \\ &= t^{-n(n-1)/2} \Delta(x) \Delta(y) / \Delta_n \end{aligned} \tag{16}$$

Thus combining (8) with the approximation (16) yields (11). Note that $\det e^{x_i y_j/t}$ is an analytic function which also vanishes whenever $\Delta(y) = 0$, since in that case two columns of the determinant are collinear. Therefore the Taylor series in y/t must match the polynomial $\det M_n$ up to the degree of M_n. Thus

$\det e^{x_i y_j / t} = \det M_n + o(t^{-n(n+1)/2})$, which heuristically justifies taking $N = n$.

Now integrate (11) over W to get

$$\begin{aligned} \Pr(\tau > t) &= C \int\limits_W t^{-n^2/2} \Delta(x)\Delta(y) e^{-\|y\|^2/2t}\, dy \\ &= C t^{-n^2/2} \Delta(x) \int\limits_W \Delta(z\sqrt{t}) e^{-\|z\|^2/2}\, t^{-n/2}\, dz \\ &= C t^{-n(n-1)/4} \Delta(x) \int\limits_W \Delta(z) e^{-\|z\|^2/2} dz \end{aligned} \tag{17}$$

This is the same as (12), since the last integral is a constant. □

A proof which uses Shur functions to justify the asymptotic approximation rather than handwaving can be found in [Gra99], which also provides the values of the constants.

3.2 COROLLARY. *Conditional on no collision up to time t, the distribution of $r = \frac{\|X_t\|}{\sqrt{t}}$ converges in measure to a Bessel process at time one in $n(n+1)/2$ dimensions starting at 0.*

Proof. Integrate the radial part of (11) over the set $\|y\|/\sqrt{t} = r$ to get, after a change of variables,

$$\Pr\left(\|X_t^*\| = r\sqrt{t},\ \tau > t\right) = t^{-n(n-1)/4} \Delta(x) e^{-r^2/2} r^{n(n+1)/2} \int\limits_{\|z\|=1} \Delta(z)\, dz \tag{18}$$

The integral on the right is constant, so the conditional distribution formed by dividing by (12) is

$$\Pr\left(\|X_t^*\|/\sqrt{t} = r \;\middle|\; \tau > t\right) = C\, e^{-r^2/2} r^{n(n+1)/2} \tag{19}$$

This matches the distribution of a Bessel process at time 1 □

4 Doob's h-transform

A function h is *harmonic* for a Markov process M_t if $h(x) = \mathrm{E}_x\, h(M_t)$ for all t, which is to say $h(M_t)$ is a martingale. Note that if $Lf = \frac{d}{dt} \mathrm{E}[f(X_t)]\Big|_{t=0}$ is the infinitesimal generator of M_t, then $Lh = 0$. We can use harmonic functions to create new Markov processes.

4.1 PROPOSITION. *Let M_t be a Markov process on Ω. Let $\widetilde{\Omega}$ be the set where $h(x) > 0$ and let $q_t(x, dy) = \frac{h(y)}{h(x)} p_t(x, dy)$.*

1. *q_t defines a kernel for a time-homogenous Markov process M_t^h*

2. *The process M_t^h has generator $L^h = h^{-1}Lh$*

3. *Suppose h grows only polynomially fast at infinity, and that M_t is a diffusion specified by*

$$dM_t = \sigma(M_t)dB_t + b(M_t)dt$$

then M_t^h is a diffusion with modified drift term

$$dM_t^h = \sigma(M_t)dB_t + \left(b(M_t) + \sigma(M_t)\sigma^\intercal(M_t)\nabla \log h(M_t)\right)dt$$

Proof. Consider $q_t(x,B) = h^{-1}(x)\int_B h(y)p_t(x,dy)$. Since its just the product of measurable functions, the map $x \hookrightarrow q_t(x,B)$ is measurable. Treating the expression $h(y)/h(x)$ as a Radon-Nikodym derivative, the map $B \hookrightarrow q_t(x,B)$ is a measure which is absolutely continuous to p_t on $\widetilde{\Omega}$. Furthermore q_t is a probability measure since $q_t(x,\widetilde{\Omega}) = h^{-1}(x)\,\mathrm{E}_x[h(M_t)] = 1$. Let p_t act on bounded functions f by $P_t f = \mathrm{E}_x[f(M_t)] = \int_\Omega f(y)p_t(x,dy)$. The Markov property is equivalent to the fact that P_t satisfies the semigroup property $P_{t+s} = P_t P_s$. The operator Q_t is given by

$$Q_t f = h(x)^{-1}\int\limits_\Omega f(y)h(y)p_t(x,dy) = h^{-1}P_t h \tag{20}$$

and therefore $Q_t Q_s = (h^{-1}P_t h)(h^{-1}P_s h) = h^{-1}P_{t+s}h = Q_{t+s}$ also satisfies the semigroup property. This shows (1).

From the relation $Lf = \frac{d}{dt}E_x f(M_t)\Big|_{t=0} = \frac{d}{dt}P_t\Big|_{t=0}$, it follows that $L^h f = h^{-1}Lh$, which shows (2). For a diffusion $dM_t = \sigma(M_t)dB_t + b(M_t)dt$ the generator is given by

$$L = \frac{1}{2}\sum_{ij} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_i b_i \frac{\partial}{\partial x_i} \tag{21}$$

where the matrix $(a_{ij}) = \sigma\sigma^\intercal$. The growth condition on h permits us to take certain derivatives under the integral sign

$$\begin{aligned} L\,hf =& \frac{f}{2}\sum_{ij} a_{ij}\frac{\partial^2 h}{\partial x_i \partial x_j} + \frac{h}{2}\sum_{ij} a_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{ij} a_{ij}\frac{\partial h}{\partial x_i}\frac{\partial f}{\partial x_j} + h\sum_i b_i\frac{\partial f}{\partial x_i} + f\sum_i b_i \frac{\partial h}{\partial x_i} \\ =& fLh + hLf + a\nabla h \nabla f \end{aligned}$$

Since h is harmonic, this is equivalent to the operator equation

$$L^h = h^{-1}Lh = L + a\frac{\nabla h}{h} \cdot \nabla \tag{22}$$

This is the generator for a diffusion with augmented drift term, which justifies (3). ☐

One source of harmonic functions h are events A which are invariant under the time-shift operator $A = \theta_t^{-1}A$. Then we can let $h(x) = \mathrm{E}_x \mathbb{1}_A(M_t)$. This expression on the right is independent of t by the strong Markov property and the fact A is time-shift invariant. The function h is harmonic by the tower property of conditional expectations. In this case, q_t has a natural interpretation as the transition kernel for the Markov process conditional on A.

$$\begin{aligned}\Pr(M_t = y \mid M_0 = x, A) &= \frac{\Pr(A \mid M_t = y, M_0 = x)\Pr(M_t = y \mid M_0 = x)}{\Pr(A \mid M_0 = x)} \\ &= \frac{h(y)p_t(x, dy)}{h(x)}\end{aligned}$$

The equivalence $\Pr(A \mid M_t = y, M_0 = x) = \Pr(A \mid M_0 = y) = h(y)$ is justified by the Markov property and time-shift invariance.

For a closed set K let τ be the hitting time $\tau = \inf\{t | M_t \in K\}$. Then for $U \subset K$ the events $M_\tau \in U$ are time-shift invariant. Thus if we can find harmonic functions h for the stopped Markov chain which are 0 on U and 1 on $K \setminus U$, the h-transform gives the stopped Markov process which conditionally never hits U.

4.2 proposition. *Let X_t be the n-dimensional Brownian motion. Up to a positive multiplicative constant, the function $h(x) = \Delta(x)$ is the unique positive harmonic function for the semigroup X_t^* on the Weyl chamber W which vanishes on the boundary.*

Proof. Clearly h is positive on W since its the product of positive terms, and $h = 0$ on ∂Wsince since $\delta(x)$ vanishes whenever two components of x are coincident. The infinitesimal generator for X_t^* is the same as for X_t, and is equal to $\frac{1}{2}\nabla^2 = \frac{1}{2}\sum_i \frac{\partial^2}{\partial x_i^2}$. Thus to show that $h(x)$ is harmonic with respect to X_t^* its sufficient to show $\Delta(x)$ is harmonic in the classical sense.

Calculating, we get the expression

$$\nabla^2\Delta(x) = \sum_{j=1}^{n} \sum_{\sigma \in S_n} \operatorname{sgn}\sigma \, (\sigma(j) - 1)(\sigma(j) - 2)\frac{1}{x_j^2} \prod_{k=1}^{n} x_k^{\sigma(k)-1} \tag{23}$$

The terms with $\sigma(j) = 1$ or $\sigma(j) = 2$ are zero. Otherwise, the terms cancel in pairs. Note that the lowered exponent of x_j is $l = \sigma(j) - 3$. Also l is the exponent $\sigma(i) - 1$ of x_i for some unique i. This can be seen by solving $i = \sigma^{-1}(\sigma(j) - 2)$. So consider the term associated with $\sigma' = \lambda\sigma$ and i where $\lambda = (i, j)$ is the transposition. For $k \notin \{i, j\}$, the exponents of x_k are the same since $\sigma(k) = \sigma'(k)$. However the lowered exponent $\sigma'(i) - 3$ of x_i and the exponent $\sigma'(j) - 1$ of x_j both equal to l. Since $\operatorname{sgn}\sigma' = -\operatorname{sgn}\sigma$, one term cancels the other. □

The h-transform of the stopped diffusion X_t^* with respect to this function yields a process $Y_t = X_t^{*h}$. The transition kernel given by the Johansson formula

$$q_t(x, dy) = \frac{1}{(2\pi t)^{n/2}} \frac{\Delta(y)}{\Delta(x)} e^{-\|x\|^2/2t} e^{-\|y\|^2/2t} \det e^{x_i y_j / t} \tag{24}$$

If we apply the same approximation as in (11), and consider the distribution of $z = y/\sqrt{t}$ we get the Ginibre formula

$$\rho(z) \sim \frac{\Delta(z)^2 e^{-\|z\|^2/2}}{(2\pi)^{n/2}\Delta_n} \tag{25}$$

Its tempting to interpret this as the process formed from n Brownian motions conditional on never having a collision. However some care must be taken since it follows from (12) that the event $\{\tau = \infty\}$ has probability 0. We may instead condition on the event that $\|X_t\| = r$ occurs before a collision happens, and then take the limit as $r \to \infty$.

We can now write down Y_t equations as a stochastic process using proposition 4.1, using the fact that $\frac{\partial \Delta(x)}{\partial x_i} = \sum_j \frac{\Delta(x)}{x_i - x_j}$

$$dY_t^i = \sum_j \frac{1}{Y_t^i - Y_t^j} dt + dB_t \tag{26}$$

The drift term adds a repulsion between between the ith and jth component whose strength increases as $1/r$ as they get closer. This is called Dyson's Brownian motion. Its an interesting coincidence that the eigenvalues for a Hermitian matrix whose entries are independent Brownian motion also follows the same process.

References

[Gra99] David J Grabiner. Brownian motion in a weyl chamber, non-colliding particles, and random matrices. *Annales de l'IHP Probabilités et statistiques*, 35(2):177–204, 1999.

[KM59] Samuel Karlin and James McGregor. Coincidence probabilities. *Pacific J. Math*, 9(4):1141–1164, 1959.