$$
\begin{aligned}
g(y_1,\dots,y_K) &= f(p_1,\dots,p_K|Y)f_Y(y)|J| \\
&= f(p_1,\dots,p_K)\frac{1}{\Gamma(\theta)}y^{\theta-1}e^{-y}\frac{1}{y^{K-1}} \\
&= \frac{\Gamma(\theta)}{\Gamma(\theta_1)\dots\Gamma(\theta_K)} \\
&\cdot\left(\frac{y_1}{\sum y_i}\right)^{\theta_1-1}\cdots\left(\frac{y_K}{\sum y_i}\right)^{\theta_K-1}\frac{1}{\Gamma(\theta)}(\sum y_i)^{(\theta-1)}e^{-\sum y_i}.(\sum y_i)^{-(K-1)} \\
&= \prod_{i=1}^{K}\frac{1}{\Gamma(\theta_i)}y_i^{\theta_i-1}e^{-y_i}
\end{aligned}
$$

Note that this coincides with the Dirichlet$(\theta_1,\dots,\theta_K)$ distribution. ■

Corollary 6.12 *The invariant measure of the infinitely many alleles model can be represented by the random probability measure*

$$(6.26)\quad Y(A)=\frac{X_\infty(A)}{X_\infty([0,1])},\qquad ,A\in\mathcal{B}([0,1]).$$

where $X_\infty(\cdot)$ is the equilibrium of the above Jirina process and $Y(\cdot)$ and $X_\infty([0,1])$ are independent.

Reversibility

Recall that the Dirichlet distribution is a *reversible* stationary measure for the $K-type$ Wright-Fisher model with house of cards mutation (Theorem 5.9). From this and the projective limit construction it can be verified that $\mathcal{L}(Y(\cdot))$ is a reversible stationary measure for the infinitely many alleles process. Note that reversibility actually characterizes the IMA model among neutral Fleming-Viot processes with mutation, that is, any mutation mechanism other than the "type-independent" or "house of cards" mutation leads to a stationary measure that is not reversible (see Li-Shiga-Yau (1999) [431]).

6.6.3 The Poisson-Dirichlet Distribution

Without loss of generality we can assume that ν_0 is Lebesgue measure on $[0,1]$. This implies that the IMA equilibrium is given by a random probability measure which is pure atomic

$$(6.27)\quad p_\infty=\sum_{i=1}^{\infty}a_i\delta_{x_i},\qquad \sum_{i=1}^{\infty}a_i=1,\ x_i\in[0,1]$$

in which the $\{x_i\}$ are i.i.d. $U([0,1])$ and the atom sizes $\{a_i\}$ correspond to the normalized jumps of the Moran subordinator. Let $(\xi_1, \xi_2, \dots)$ denote the reordering of the atom sizes $\{a_i\}$ in decreasing order.

The Poisson-Dirichlet $PD(\theta)$ distribution is defined to be the distribution of the infinite sequence $\xi = (\xi_1, \xi_2, \dots)$ which satisfies

$$\xi_1 \geq \xi_2 \geq \dots, \ \sum_k \xi_k = 1.$$

This sequence is given by

$$(6.28)\quad \xi_k = \frac{\eta_k(\theta)}{G(\theta)} = \frac{\eta_k}{\sum \eta_\ell}, \quad k = 1, 2, \dots,$$

where $\eta_k = \eta_k(\theta)$ is the height of the kth largest jump in $[0,\theta]$ of the Moran process (subordinator), G and $G(\theta) = \sum_{\ell=1}^\infty \eta_\ell$.

Properties of the Poisson-Dirichlet Distribution

Recalling (6.19), (6.21) we note that the set of heights of the jumps of $G(\cdot)$ in $[0,\theta]$ form a Poisson random field Π_θ on $(0,\infty)$ with intensity measure

$$\theta \frac{e^{-u}}{u} du.$$

We can then give a direct description of $PD(\theta)$ in terms of such a Poisson random field. If $\eta_1 \geq \eta_2 \geq \eta_3 \geq \dots$ are the points of such a random field ordered by size then

$$\xi_k = \frac{\eta_k}{\sum_{\ell=1}^\infty \eta_\ell}$$

defines a sequence ξ having the distribution $PD(\theta)$.

By the law of large numbers (for the Poisson) we get

$$\lim_{t\to 0} \frac{\#\{k : \eta_k > t\}}{L(t)} = 1$$

with probability one where

$$L(t) = \int_t^\infty \theta \frac{e^{-u}}{u} du \sim -\theta \log t.$$

Thus

$$\begin{aligned} \#\{k : \eta_k > t\} &\sim -\theta \log t \\ \eta_{\theta \log \frac{1}{t}} &\approx t \quad \text{as } t \to 0. \\ \eta_k &\approx e^{-k/\theta} \end{aligned}$$

Thus ξ_k decays exponentially fast

$$-\log \xi_k \sim \frac{k}{\theta} \quad \text{as } k \to \infty.$$

The Distribution of Atom Sizes

We now introduce the random measure on $(0, \infty)$,

$$Z_\theta((a,b)) = \frac{\Xi_\theta([0,1] \times (a,b))}{G(\theta)}$$
$$\int_0^\infty u Z_\theta(du) = 1.$$

This is the *distribution of normalized atom sizes* and this just depends on the normalized ordered atoms and hence is independent of $X_\infty([0,1])$. Intuitively, as $\theta \to \infty$, $Z_\theta(\theta du)$ converges in some sense to

$$\frac{e^{-u}}{u} du.$$

To give a precise formulation of this we first note that

$$\int_0^\infty u^k \left(\frac{e^{-u}}{u}\right) du = \Gamma(k) = (k-1)!$$

Then one can show (see Griffiths (1979), [280]) that

$$(6.29) \quad \lim_{\theta \to \infty} \theta^{k-1} \int_0^\infty u^k Z_\theta(du) = (k-1)!$$

and there is an associated CLT

$$(6.30) \quad \sqrt{\theta}\, \frac{\theta^{k-1} \int u^k Z_\theta(du) - (k-1)!}{(k-1)!} \Rightarrow N(0, \sigma_k^2),$$

with $\sigma_k^2 = \frac{(2k-1)! - (k!)^2}{((k-1)!)^2}$ Joyce, Krone and Kurtz (2002) [358]. Also see Dawson and Feng (2006) [146] for the related large deviation behaviour.

6.6.4 The GEM Representation

Without loss of generality we can assume $\nu_0 = U[0,1]$. Consider a partition of $[0,1]$ into K intervals of equal length. Then the random probability

$$\vec{p}_K = (p_1, \ldots, p_K)$$

has the symmetric Dirichlet distribution $D(\alpha, \ldots, \alpha)$ with $\alpha = \frac{\theta}{K}$.

Randomized Ordering via Size-biased sampling

Let $\mathcal{N}$ be a random variable having values in $\{1, 2, \ldots, K\}$ in such a way that

$$P(\mathcal{N} = k | \vec{p}_K) = p_k, \quad (1 \leq k \leq K)$$

Then a standard calculation shows that the vector

$$\vec{p'} = (p_{\mathcal{N}}, p_1, \ldots, p_{\mathcal{N}-1}, p_{\mathcal{N}+1}, \ldots, p_K)$$

has distribution (cf. (5.55))

$$D(\alpha + 1, \alpha, \ldots, \alpha)$$

It follows that $(p_{\mathcal{N}}, 1 - p_{\mathcal{N}})$ has the Dirichlet distribution (Beta distribution)

$$D(\alpha + 1, (K - 1)\alpha)$$

so that it has probability density function

$$\frac{\Gamma(K\alpha + 1)}{\Gamma(\alpha + 1)\Gamma(K\alpha - \alpha)} p^{\alpha}(1 - p)^{(K-1)\alpha - 1}.$$

Given $v_1 = p_{\mathcal{N}}$, the conditional distribution of the remaining components of $\vec{p}$ is the same as that of $(1 - p_{\mathcal{N}})p^{(1)}$, where the $(K - 1)$-vector $\vec{p}^{(1)}$ has the symmetric distribution $D(\alpha, \ldots, \alpha)$.

We say that $p_{\mathcal{N}}$ is obtained from $\vec{p}$ by size-biased sampling. This process may now be applied to $\vec{p}^{(1)}$ to produce a component, v_2 with distribution

$$\frac{\Gamma((K - 1)\alpha + 1)}{\Gamma(\alpha + 1)\Gamma((K - 1)\alpha - \alpha)} p^{\alpha}(1 - p)^{(K-2)\alpha - 1}$$

and a $(K - 2)$ vector $\vec{p}^{(2)}$ with distribution $D(\alpha, \ldots, \alpha)$. This is an example of Kolmogorov's *stick breaking* process.

Theorem 6.13 *(a) As $K \to \infty$, with $K\alpha = \theta$ constant, the distribution of the vector $\vec{q}_K = (q_1, q_2, \ldots, q_K)$ converges weakly to the GEM distribution with parameter θ, that is the distribution of the random probability vector $\vec{q} = (q_1, q_2, \ldots)$ where*

$$q_1 = v_1, \; q_2 = (1 - v_1)v_2, \; q_3 = (1 - v_1)(1 - v_2)\nu_3, \; \ldots$$

with $\{v_k\}$ are i.i.d. with Beta density (Beta$(1, \theta)$)

$$\theta(1 - p)^{\theta - 1}, \; 0 \leq p \leq 1$$

(b) If $\vec{q} = (q_1, q_2, \ldots)$ is reordered (by size) as $\vec{p} = (p_1, p_2, \ldots)$, that is i.e. p_k is the kth largest of the $\{q_j\}$,then $\vec{p}$ has the Poisson-Dirichlet distribution, $PD(\theta)$.

Proof. (a) Let $(p_1^K,\ldots,p_k^K)\in\Delta_{K-1}$ be a random probability vector obtained by decreasing size reordering of a probability vector sampled from the distribution $D(\alpha,\ldots,\alpha)$ with $\alpha=\frac{\theta}{K}$. Then let $(q_1^K,\ldots,q_k^K)$ be the size-biased reordering of $(p_1^K,\ldots,p_k^K)$. Then as shown above we can rewrite this as

$$(6.31)\quad q_1^K=v_1^K,\; q_2^K=(1-v_1^K)v_2^K),\; q_3^K=(1-v_1^K)(1-v_2^K)v_3^K,\ldots$$

where $v_1^K,\ldots,v_{K-1}^K$ are independent and v_r^K has pdf

$$(6.32)\quad \frac{\Gamma((K-r)\alpha+1)}{\Gamma(\alpha+1)\Gamma((K-r)\alpha-\alpha)}u^{\alpha}(1-u)^{(K-r-1)\alpha-1},\quad 0\le u\le 1.$$

Now let $K\to\infty$ with $K\alpha=\theta$. Then

$$\frac{\Gamma(K\alpha+1)}{\Gamma(\alpha+1)\Gamma(K\alpha-\alpha)}p^{\alpha}(1-p)^{(K-1)\alpha-1}\to\theta(1-p)^{\theta-1}.$$

and

$$\begin{aligned}&\frac{\Gamma((K-r)\alpha+1)}{\Gamma(\alpha+1)\Gamma((K-r)\alpha-\alpha)}p^{\alpha}(1-p)^{(K-r-1)\alpha-1}\\ &=\frac{\Gamma((K-r)\theta/K+1)}{\Gamma(\theta/K+1)\Gamma((K-r)\theta/K-\theta/K)}p^{\theta/K}(1-p)^{(K-r-1)\theta/K-1}\\ &\to\theta(1-p)^{\theta}\end{aligned}$$

Thus the distributions of the first m components of the vector $\vec{q}_K$ converge weakly to the distribution of the first m components of the random (infinite) probability vector $\vec{q}$ defined by

$$q_1=v_1,\; q_2=(1-v_1)v_2,\; q_3=(1-v_1)(1-v_2)\nu_3,\;\ldots$$

where $\{v_k\}$ are i.i.d. with Beta density (Beta$(1,\theta)$)

$$\theta(1-p)^{\theta-1},\; 0\le p\le 1.$$

(b) By the projective limit construction, the PD(θ) distribution arises as the limit in distribution of the ordered probability vectors $(p_1^K,\ldots,p_k^K)$. Then the size-biased reorderings converge in distribution to the size-biased reordering $q_1,q_2,\ldots$ of the probability vector $p_1,p_2,p_3,\ldots$. Clearly the decreasing-size reordering of $q_1,q_2,\ldots$ reproduces $p_1,p_2,\ldots$. ■

Remark 6.14 *The distribution of sizes of the* age ordered alleles *in the infinitely many alleles model is given by the GEM distribution. The intuitive idea is as follows. By exchangeability at the individual level the probability that the kth allele at a given time survives the longest (time to extinction) among those present at that time is proportional to p_k the frequency of that allele in the population at that time. Observing that the ordered survival times correspond to ages under time reversal, the result follows from reversibility. See Ethier [222] for justification of this argument and a second proof of the result.*

Remark 6.15 *There is a two parameter analogue of the Poisson-Dirichlet introduced by Perman, Pitman and Yor (1992) [509] that shares some features with the PD distribution. See Feng (2009) [246] for a recent detailed exposition.*

6.6.5 Application of the Poisson-Dirichlet distrbution: The Ewens Sampling Formula

In analyzing population genetics data under the neutral hypothesis it is important to know the probabilities of the distribution of types obtained in taking a random sample of size n. For example, this is used to test for neutral mutation in a population.

Consider a random sample of size n chosen from the random vector $\xi = (\xi_1, \xi_2, \dots)$ chosen from the distribution $PD(\theta)$.

We first compute the probability that they are all of the same type. Conditioned on ξ this is $\sum_{k=1}^{\infty} \xi_k^n$ and hence the unconditional probability is

$$h_n = E\left\{\sum_{k=1}^{\infty} \xi_k^n\right\}$$

Using Campbell's formula we get

$$\begin{aligned} &E\left\{\sum_{k=1}^{\infty} \eta_k^n\right\} \\ &= \frac{d}{ds}\left(E(e^{\int_0^1\int_0^\infty (e^{sz^n}-1)\frac{\theta e^{-z}}{z}dz}\right)|_{s=0} \\ &= \int z^n \frac{\theta e^{-z}}{z} dz = \theta(n-1)! \end{aligned}$$

Also

$$E((\sum_{k=1}^{\infty} \eta_k)^n) = \frac{\Gamma(n+\theta)}{\Gamma(\theta)}.$$

By the Gamma representation (6.28) for the ordered jumps of the Gamma subordinator we get

$$\begin{aligned} E\left\{\sum_{k=1}^{\infty} \eta_k^n\right\} &= E\left\{(\sum_{k=1}^{\infty} \eta_k)^n \sum_{k=1}^{\infty} \xi_k^n\right\} \\ &= E((\sum_{k=1}^{\infty} \eta_k)^n) E\left\{\sum_{k=1}^{\infty} \xi_k^n\right\} \quad \text{by independence} \end{aligned}$$

Therefore

$$h_n = \frac{\theta\Gamma(\theta)(n-1)!}{\Gamma(n+\theta)} = \frac{(n-1)!}{(1+\theta)(2+\theta)\dots(n+\theta-1)}$$

In general in a sample of size n let

$$
\begin{aligned}
a_1 &= \text{ number of types with 1representative} \\
a_2 &= \text{ number of types with 2 representatives} \\
&\dots \\
a_n &= \text{number of types with n representatives}
\end{aligned}
$$

Of course,

$$a_i \geq 0, \quad a_1 + 2a_2 + \cdots + na_n = n.$$

We can also think of this as a partition

$$\mathfrak{a} = 1^{a_1} 2^{a_2} \dots n^{a_n}$$

Let $P_n(\mathfrak{a})$ denote the probability that the sample exhibits the partition $\mathfrak{a}$. Note that $P_n(n^1) = h_n$.

Proposition 6.16 *(Ewens sampling formula)*

$$P_n(\mathfrak{a}) = P_n(a_1, \dots, a_n) = \frac{n!\Gamma(\theta)}{\Gamma(n+\theta)} \prod_{j=1}^{n} \left(\frac{\theta^{a_j}}{j^{a_j} a_j!} \right). \tag{6.33}$$

Proof. We can select the partition of $\{1, \dots, n\}$ into subsets of sizes $(a_1, \dots, a_n)$ as follows. Consider a_1 boxes of size 1, ... a_n boxes of size n. Then the number of ways we can distribute $\{1, \dots, n\}$ is $n!$ but we can reorder the a_i boxes in $a_i!$ ways and and there are $(j!)$ permutations of the indices in each of the a_j partition elements with j elements. Hence the total number of ways we can do the partitioning to $\{1, \dots, n\}$ is $\frac{n!}{\prod (j!)^{a_j} a_j!}$.

Now condition on the vector $\xi = (\xi(1), \xi(2), \dots)$. The probability that we select the types (ordered by their frequencies is then given by)

$$P_n(\mathfrak{a}|\xi) = \frac{n!}{\prod (j!)^{a_j} a_j!} \sum_{I_{a_1, \dots an}} \xi(k_{11}) \xi(k_{12}) \dots \xi(k_{1a_1}) \xi(k_{21})^2 \dots \xi(k_{2a_2})^2 \xi(k_{31})^3 \dots$$

where the summation is over

$$I_{a_1, \dots an} := \{k_{ij} : i = 1, 2, \dots; j = 1, 2, \dots, a_i\}$$

Hence using the Gamma representation we get

$$
\begin{aligned}
&\frac{\Gamma(n+\theta)}{\Gamma(\theta)} P_n(\mathfrak{a}) \\
&= \frac{n!}{\prod (j!)^{a_j} a_j!} E \left\{ \sum \eta(k_{11}) \eta(k_{12}) \dots \eta(k_{1a_1}) \eta(k_{21})^2 \dots \eta(k_{2a_2})^2 \eta(k_{31})^3 \dots \right\}
\end{aligned}
$$

But

$$E\left\{\sum_{i,j}\eta\left(k_{11}\right)\eta(k_{12})\ldots\eta(k_{1a_1})\eta(k_{21})^2\ldots\eta(k_{2a_2})^2\eta(k_{31})^3\ldots\right\}$$

$$=\prod_{j=1}^{n}E\left\{\sum_{k=1}^{\infty}\eta(k)^j\right\}^{a_j}=\prod_{j=1}^{n}\left\{\int_0^{\infty}z^j\theta\frac{e^{-z}}{z}dz\right\}^{a_j}\quad\text{by Campbell's thm}$$

$$=\prod_{j=1}^{n}\{\theta(j-1)!\}^{a_j}$$

Therefore substituting we get

$$P_n(\mathfrak{a})=\frac{n!\Gamma(\theta)}{\Gamma(n+\theta)}\prod_{j=1}^{n}\left(\frac{\theta^{a_j}}{j^{a_j}a_j!}\right).$$

■

Chapter 7

Martingale Problems and Dual Representations

7.1 Introduction

We have introduced above the basic mechanisms of branching, resampling and mutation mainly using generating functions and semigroup methods. However these methods have limitations and in order to work with a wider class of mechanisms we will introduce some this additional tools of stochastic analysis in this chapter. The martingale method which we use has proved to be a natural framework for studying a wide range of problems including those of population systems. The general framework is as follows:

- the object is to specify a Markov process on a Polish space E in terms of its probability laws $\{P_x\}_{x\in E}$ where P_x is a probability measure on $C_E([0,\infty))$ or $D_E([0,\infty))$ satisfying $P_x(X(0)=x)=1$.
- the probabilities $\{P_x\}\in\mathcal{P}(C_E([0,\infty)))$ satisfy a *martingale problem* (MP). One class of martingale problems is defined by the set of conditions of the form

  $$(7.1)\quad F(X(t))-\int_0^t GF(X(s))ds,\quad F\in\mathcal{D}\qquad (\mathcal{D},G)-\text{martingale problem}$$

  is a P_x martingale where G is a linear map from $\mathcal{D}$ to $C(E)$, and $\mathcal{D}\subset C(E)$ is measure-determining.
- the martingale problem MP has one and only one solution.

Two martingale problems MP_1, MP_2 are said to be *equivalent* if a solution to MP_1 problem is a solution to MP_2 and vice versa.

In our setting the existence of a solution is often obtained as the limit of a sequence of probability laws of approximating processes. The question of uniqueness is often the more challenging part. We introduce the method of dual representation which can be used to establish uniqueness for a number of basic population processes. However the method of duality is applicable only for special classes of models. We introduce a second method, the Cameron-Martin-Girsanov type change of measure which is applicable to some basic problems of stochastic population systems. Beyond the domain of applicability of these methods, things are much more challenging. Some recent progress has been made in a series of papers of Athreya, Barlow, Bass, Perkins [11], Bass-Perkins [28], [29] but open problems remain.

We begin by reformulating the Jirina and neutral IMA Fleming-Viot in the martingale problem setting. We then develop the Girsanov and duality methods in the framework of measure-valued processes and apply them to the Fleming-Viot process with selection.

7.2 The Jirina martingale problem

By our projective limit construction of the Jirina process (with $\nu_0 =$ Lebesgue), we have a probability space $(\Omega, \mathcal{F}, \{X^\infty : [0,\infty) \times \mathcal{B}([0,1]) \to [0,\infty)\}, P)$ such that a.s. $t \to X^\infty_t(A)$ is continuous and $A \to X^\infty_\cdot(A)$ is finitely additive. We can take a modification, X, of X^∞ such that a.s. $X : [0,\infty) \to M_F([0,1])$ is continuous where $M_F([0,1])$ is the space of (countably additive) finite measures on $[0,1]$ with the weak topology. We then define the filtration

$$\mathcal{F}_t : \sigma\{X_s(A) : 0 \le s \le t,\ A \in \mathcal{B}([0,1])\}$$

and $\mathcal{P}$, the σ-field of predictable sets in $\mathbb{R}_+ \times \Omega$ (ie the σ-algebra generated by the class of $\mathcal{F}_t$-adapted, left continuous processes).

Recall that for a fixed set A the Feller CSB with immigration satisfies

$$(7.2)\quad X_t(A) - X_0(A) - \int_0^t c(\nu_0(A) - X_s(A))ds = \int_0^t \sqrt{2\gamma X_t(A)}dw^A_t$$

which is an L^2-martingale.

Moreover, by polarization

$$(7.3)\quad \langle M(A_1), M(A_2)\rangle_t = \gamma \int_0^t X_s(A_1 \cap A_2)ds$$

and if $A_1 \cap A_2 = \emptyset$, then the martingales $M(A_1)_t$ and $M(A_2)_t$ are orthogonal. This is an example of an *orthogonal martingale measure.*

Therefore for any Borel set A

$$(7.4)\quad M_t(A) := X_t(A) - X_0(A) - \int_0^t c(\nu_0(A) - X_s(A))ds$$

is a martingale with increasing process

$$(7.5)\ \ \langle M(A)\rangle_t = \gamma \int_0^t X_s(A)ds.$$

We note that we can define integrals with respect to an orthogonal martingale measure (see next subsection) and show that (letting $X_t(f) = \int f(x)X_t(dx)$ for $f \in \mathcal{B}([0,1])$)

$$(7.6)\ \ M_t(f) := X_t(f) - X_0(f) - \int_0^t c(\nu_0(f) - X_s(f))ds = \int f(x)M_t(dx)$$

which is a martingale with increasing process

$$(7.7)\ \ \langle M(f)\rangle_t = \gamma \int_0^t f^2(x)X_s(dx)ds.$$

This suggests the martingale problem for the Jirina process which we state in subsection 7.2.2.

7.2.1 Stochastic Integrals wrt Martingale Measures

A general approach to martingale measures and stochastic integrals with respect to martingale measures was developed by Walsh [596]. We briefly review some basic results.

Let

$$\mathcal{L}^2_{loc} = \left\{\psi : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R} : \psi \text{ is } \mathcal{P}\times\mathcal{E}\text{-measurable, } \int_0^t X_s(\psi_s^2)ds < \infty,\ \forall t > 0\right\}$$

A $\mathcal{P}\times\mathcal{E}$-measurable function ψ is *simple* ($\psi \in \mathcal{S}$) iff

$$\psi(t,\omega,x) = \sum_{i=1}^K \psi_{i-1}(\omega)\phi_i(x)1_{(t_{i-1},t_i]}(t)$$

for some $\phi_i \in b\mathcal{B}([0,1])$, $\psi \in b\mathcal{F}_{t_{i-1}}$, $0 = t_0 < t_1 \cdots < t_K \le \infty$. For such a ψ, define

$$M_t(\psi) := \int_0^t \int \psi(s,x)dM(s,x) = \sum_{i=1}^K \psi_{i-1}(M_{t\wedge t_i}(\phi_i) - M_{t\wedge t_{i-1}}(\phi_i))$$

Then $M_t(\psi) \in \mathcal{M}_{\text{loc}}$ (the space of $\mathcal{F}_t$ local martingales) and

$$\langle M(\psi)_t\rangle = \int_0^t X_s(\gamma\psi_s^2)ds.$$

Lemma 7.1 *For any $\psi \in \mathcal{L}^2_{loc}$ there is a sequence $\{\psi_n\}$ in $\mathcal{S}$ such that*

$$P\left(\int_0^n \int (\psi_n - \psi)^2(s,\omega,x)\gamma(x)X_s(dx)ds > 2^{-n}\right) < 2^{-n}.$$

Proof. Let $\bar{\mathcal{S}}$ denote the set of bounded $\mathcal{P}\times\mathcal{E}$-measurable functions which can be approximated as above. $\bar{\mathcal{S}}$ is closed under $\rightarrow^{bp}$. Using $\mathcal{H}_0 = \{f_{i-1}(\omega)\phi_i(x), \phi \in b\mathcal{E},\ f_{i-1} \in b\mathcal{F}_{t_{i-1}}, \phi_i \in b\mathcal{E}\}$, we see that $\psi(t,\omega,x) = \sum_{i=1}^K \psi_{i-1}(\omega,x)1_{(t_{i-1},t_i]}(t)$ is in $\bar{\mathcal{S}}$ for any $\psi_{i-1} \in b(\mathcal{F}_{t_{i-1}}\times\mathcal{E})$. If $\psi \in b(\mathcal{P}\times\mathcal{E})$, then

$$\psi_n(s,\omega,x) = 2^n \int_{(i-1)2^{-n}}^{i2^{-n}} \psi(r,\omega,x)dr \text{ is } s \in (i2^{-n},(i+1)2^{-n}],\ i = 1,2,\dots$$

satisfies $\psi_n \in \bar{\mathcal{S}}$ by the above. For each $(\omega,x), \psi_n(s,\omega,x) \rightarrow \psi(s,\omega,x)$ for Lebesgue a.a. s by Lebesgue's differentiation theorem and it follows easily that $\psi \in \bar{\mathcal{S}}$. Finally if $\psi \in \mathcal{L}^2_{\text{loc}}$, the obvious truncation argument and dominated convergence (set $\psi_n = (\psi \wedge n) \vee (-n)$ completes the proof. ■

Proposition 7.2 *There is a unique linear extension of $M : \mathcal{S} \rightarrow \mathcal{M}_{loc}$ (the space of local martingales) to a map $M : \mathcal{L}^2_{loc} \rightarrow \mathcal{M}_{loc}$ such that $M_t(\psi)$ is a local martingale with increasing process $\langle M(\psi)\rangle_t$ given by*

$$\langle M(\psi)\rangle_t := \int_0^t \gamma X_s(\psi_s^2)ds\ \forall\, t \geq 0\ \ a.s. \forall\, \psi \in \mathcal{L}^2_{loc}.$$

Proof. We can choose $\psi_n \in \mathcal{S}$ as in the Lemma. Then

$$\langle M(\psi) - M(\psi_n)\rangle_n = \langle M(\psi - \psi_n)\rangle_n = \gamma \int_0^n X_s(\gamma(\psi(s) - \psi_n(s))^2 ds$$

$$P\left(\langle M(\psi) - M(\psi_n)\rangle_n > 2^{-n}\right) < 2^{-n}$$

The $\{M_t(\psi_n)\}_{t\geq 0}$ is Cauchy and using Doob's inequality and the Borel-Cantelli Lemma we can define $\{M_t(\psi)\}_{t\geq 0}$ such that

$$\sup_{t\leq n} |M_t(\psi) - M_t(\psi_n)| \rightarrow 0 \text{ a.s. as } n \rightarrow \dot{\infty}.$$

This yields the required extension and its uniqueness. ■

Note that it immediately follows by polarization that if $\psi, \phi \in \mathcal{L}^2_{\text{loc}}$,

$$\langle M(\phi), M(\psi)\rangle_t = \gamma \int_0^t X_s(\phi_s\psi_s)ds$$

Moreover, in this case $M_t(\psi)$ is a L^2-martingale, that is,

$$E(\langle M(\psi)\rangle_t) = \gamma \int_0^t E(X_s(\psi_s^2))ds < \infty$$

provided that

$$\psi \in \mathcal{L}^2 = \{\psi \in \mathcal{L}^2_{\text{loc}} : E(\int_0^t X_s(\psi_s^2)ds < \infty,\ \forall t > 0\}.$$

Remark 7.3 *Walsh (1986) [596] defined a more general class of martingale measures on a measurable space* $(E, \mathcal{E})$ *for which the above construction of stochastic integrals can be extended.* $\{M_t(A) : t \geq 0,\ A \in \mathcal{E}\}$ *is an* L^2*-martingale measure wrt* $\mathcal{F}_t$ *iff*

(a) $M_0(A) = 0\ \ \forall A \in \mathcal{E}$,

(b) $\{M_t(A), t \geq 0\}$ *is an* $\mathcal{F}_t$*-martingale for every* $A \in \mathcal{E}$,

(c) for all $t > 0,\ M_t$ *is an* L^2*-valued* σ*-finite measure.*

The martingale measure is worthy if there exists a σ*-finite "dominating measure"* $K(\cdot,\cdot,\cdot,\omega)$, *on* $\mathcal{E} \times \mathcal{E} \times \mathcal{B}(\mathbb{R}_+),\ \omega \in \Omega$ *such that*

(a) K *is symmetric and positive definite, i.e. for any* $f \in b\mathcal{E} \times \mathcal{B}(\mathbb{R}_+)$,

$$\int\int\int f(x,s)f(y,s)K(dx,dy,ds) \geq 0$$

(b) for fixed $A, B,\ \{K(A \times B \times (0,t]), t \geq 0\}$ *is* $\mathcal{F}_t$*-predictable*

(c) $\exists\ E_n \uparrow E$ *such that* $E\{K(E_n \times E_n \times [0,T]\} < \infty\ \forall n$,

(d) $|\ \langle M(A), M(A)\rangle_t\ | \leq K(A \times A \times [0,t])$.

7.2.2 Uniqueness and stationary measures for the Jirina Martingale Problem

A probability law, $\mathbb{P}_\mu \in C_{M_F([0,1])}([0,\infty))$, is a solution of the *Jirina martingale problem*, if under P_μ, $X_0 = \mu$ and

$$\begin{aligned} (7.8)\qquad & M_t(\phi) := X_t(\phi) - X_0(\phi) - \int_0^t c(\nu_0(\phi) - X_s(\phi))ds, \\ & \quad \text{is a } L^2,\ \mathcal{F}_t\text{-martingale } \forall\ \phi \in b\mathcal{B}([0,1]) \text{ with increasing process} \\ & \langle M(\phi)\rangle_t = \gamma \int_0^t X_s(\phi^2)ds, \text{ that is,} \\ & M_t^2(\phi) - \langle M(\phi)\rangle_t \text{ is a martingale.} \end{aligned}$$

Remark 7.4 *This is* equivalent *to the martingale problem*

$$(7.9)\quad M_F(t) = F(X_t) - \int_0^t GF(X(s))ds \quad \textit{is a martingale}$$

for all $F \in \mathcal{D} \subset C(M_F([0,1])$ *where*

$$\mathcal{D} = \{F : F(\mu) = \prod_{i=1}^n \mu(f_i),\ \ f_i \in C([0,1]),\ i = 1, \dots, , n,\ n \in \mathbb{N}\}$$

and

$$\begin{aligned} GF(\mu) = \quad & c\int \left[\int \frac{\partial F(\mu)}{\partial \mu(x)}\nu_0(dx) - \frac{\partial F(\mu)}{\partial \mu(x)}\right]\mu(dx) \\ & + \frac{\gamma}{2}\int\int \frac{\partial^2 F(\mu)}{\partial\mu(x)\partial\mu(y)}(\delta_x(dy)\mu(dx) - \mu(dx)\mu(dy)) \end{aligned}$$

Theorem 7.5 *There exists one and only one solution* $P_\mu \in \mathcal{P}(C_{M_F([0,1])}([0,\infty))$ *to the martingale problem (7.8). This defines a continuous* $M_F(0,1])$ *continuous strong Markov process.*
(b) (Ergodic Theorem) Given any initial condition, X_0*, the law of* X_t *converges weakly to a limiting distribution as* $t \to \infty$ *with Laplace functional*

$$(7.10)\quad E(e^{-\int_0^1 f(x)X_\infty(dx)}) = \exp\left(-\frac{2c}{\gamma}\int_0^1 \log(1 + \frac{f(x)}{\theta})\nu_0(dx)\right).$$

This can be represented by

$$(7.11)\quad X_\infty(A) = \frac{1}{\theta}\int_0^1 f(x)G(\theta ds)$$

where G *is the Gamma (Moran) subordinator (recall (6.17)).*

Exercise for Lecture 10

Prove the equivalence of the martingale problems (7.8) and (7.9).