A Martingale Representation Theorem

§1 INTRODUCTION

We consider a finite stochastic base biult on the Finite probability space $(\Omega, \mathcal{F}_T, \mathbb{P})$ with the filtration of σ-fields $\mathcal{F}_0 = \{\Omega, \emptyset\} \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \ldots \subseteq \mathcal{F}_T$. From our work on {\em partitions}[1] we know that each of the σ-fields, $\mathcal{F}_i$, $0 \leq i \leq T$, are generated by partitions. Also, since $\mathcal{F}_{i-1} \subseteq \mathcal{F}_i$ it follows that the partition which generates $\mathcal{F}_i$ is finer than that which generates $\mathcal{F}_{i-1}$. The elements of a partition are sometimes called {\em cells} or {\em atoms} because they 'make up' larger objects in the σ-field that they generate, just as cells come together to create an organism or atoms a molecule[2]. Since the partition that creates $\mathcal{F}_i$ is finer than that which creates $\mathcal{F}_{i-1}$ it is clear that the sets of a partition at time '$i-1$' {\em split} to make the sets of the partition at time 'i'. The way in which sets in a partition split to form sets of a new partition turns out to be quite important, so we define the {\em splitting index} of a cell, E, in a partition to be the number of cells that it splits into in the new partition. Some examples follow.

<u>Notation</u>: For a cell, E, the splitting index of E " is denoted by S(E).

1. See: "Partitions and σ- fields on finite sets."

2. So if one is discussing sets in a σ-field and "the cell, E," is refered to, it is understood that E is a cell of the partition which generates the σ-field.

Examples

1. Let $\mathcal{F}_0 = \{\Omega, \phi\}$, $\mathcal{F}_1 = \{E, \Omega\setminus E, \Omega, \phi\}$, then $S(\Omega) = 2$, because $\mathcal{F}_0$ is generated by the partition, $\{\Omega\}$ and $\mathcal{F}_1$ by $\{E, \Omega\setminus E\}$ which is what Ω splits into.

2. Let $\mathcal{F}_0 = \{\Omega, \phi\}$ and $\Omega = \{\omega_1, \ldots, \omega_N\}$. Let $\mathcal{F}_1 = \mathcal{P}(\Omega)$, the power set of Ω. Then $\mathcal{F}_1$ is generated by the partition $\{\{\omega_1\}, \{\omega_2\}, \ldots, \{\omega_N\}\}$. So the splitting index of Ω is N in this case.

3. Let $\mathcal{F}_n = \sigma\{(\frac{i-1}{2^n}, \frac{i}{2^n}] : 1 \leqslant i \leqslant 2^n\}$, for $n = 1, 2, 3, \ldots$

A cell, E, in $\mathcal{F}_n$ splits into two cells of $\mathcal{F}_{n+1}$:

$$\left(\frac{i-1}{2^n}, \frac{i}{2^n}\right] = \left(\frac{2(i-1)}{2^{n+1}}, \frac{2i-1}{2^{n+1}}\right] \cup \left(\frac{2i-1}{2^{n+1}}, \frac{2i}{2^{n+1}}\right].$$

So $S(E) = 2$ for every E (and every n!)

Hereafter we are going to consider filtrations which are generated by partitions which have a uniform splitting index for their sets. Example 3 above is like this. So if we happen upon a cell at some time, t, and observe how it splits up at time t+1, then we would observe the same number of new cells spawned at time t+1 irrespective of {tem which} cell we observe or which time, t, we began our observations (bar the obvious exceptions).

Before we make our general arguments we assemble some facts and look at a special case which will suggest the approach in the general case.

First of all, if a σ-field $\mathcal{G}$ is generated by a partition, $\{E_1, \ldots, E_k\}$, say, then any $\mathcal{G}$-measurable random variable, g, takes a sole numerical value on each element of the partition. Formally,

$$g = \sum_{i=1}^{k} g_i I_{E_i} \qquad \text{where } g_i \in \mathbb{R}.$$

The point here is this; to specify a random variable we need only state its values on the cells of the generating partition.

The Theorem that we want to prove is as follows. Given a filtration of σ-fields, $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}_T$ on $(\Omega, \mathcal{F}_T, \mathbb{P})$ we wish to find a finite number of martingales, $W^1, W^2, \ldots, W^m$, such that

(i) $\langle W^i, W^j \rangle = 0$ if $i \neq j$, i.e. they are a strongly orthogonal set

(ii) For each martingale, X, of the filtration, there are predictable processes, $g^1_X, \ldots, g^m_X$, such that

$$X_t = X_0 + \sum_{i=1}^{m} \int_0^t g^i_X(s)\, dW^i_s$$

Moreover, we want to identify the natural number, m, appearing in the expressions above.

Before we pass to the general case we consider an example. First of all, recall that the splitting index is assumed to be <u>uniform</u> and that $\mathcal{F}_0$ is generated by the trivial partition, $\{\Omega\}$. If $S(\Omega) = 1$ then Ω "divides" into one set, obviously this one set is Ω itself and there is no change whatsoever as we pass from $\mathcal{F}_0$ to $\mathcal{F}_1$. For any martingale, X, of the Filtration this means that $X_0 = X_1$. Since our splitting index is uniform over all sets and all times this means that $\mathcal{F}_0 = \mathcal{F}_1 = \mathcal{F}_2 = \dots = \mathcal{F}_T$ and as a consequence the martingales on this Filtration are simply the constant sequences, $X = (X_t)$, where $X_t = X_0 \ \forall t = 1, 2, .., T$. In this degenerate situation we have

$$X_t = X_0$$

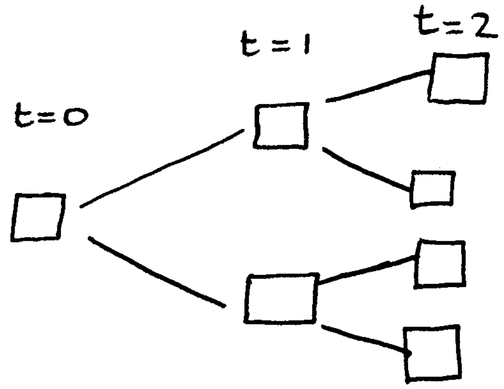
For every martingale, X, of the filtration. So, the set of orthogonal martingales, $\{W^1, .., W^m\}$, required for this situation is the {Item empty set}, $\emptyset$. So when the splitting index is 1, $m = 0$, one less than the splitting index.

<u>Remark</u>: In the case above you could choose <u>any</u> martingale Y and <u>any</u> predictable process, g and take any martingale, X, and write

$$X_t = X_0 + \int_0^t g_s \, dY_s$$

because "dY_s" is zero for every s. This might tempt one to think that 'm' should be one in this situation. But we want m to be "least possible" in some sense which is not yet clear in this degenerate case. So, we move to consider the case where the splitting index is two.

When the splitting index is 2 we can represent the "information flow" by a diagram: The squares represent



etc!

cells of the partitions. Since we assume that $\mathcal{F}_0 = \{\Omega, \emptyset\}$ we see that an $\mathcal{F}_0$ random variable is uniquely specified by a single real number because it has the form, αI_Ω. An $\mathcal{F}_1$ random variable has the form, $\alpha I_E + \beta I_{\Omega \setminus E}$ and so this is uniquely specified by a pair of real numbers — an element of $\mathbb{R}^2$ — once we have specified a listing of the elements of the partition that generates $\mathcal{F}_1$. For example if we specify $E_1 = E$ and $E_2 = \Omega \setminus E$ then $(6,2)$ specifies $6I_{E_1} + 2I_{E_2}$. But if you decide that $E_1 = \Omega \setminus E$ and $E_2 = E$ then $(2,6)$ specifies this function. This identification of functions with elements of $\mathbb{R}^2$ — and later on elements of $\mathbb{R}^n$ for $n > 2$ — will be important in what follows.

We consider an arbitrary martingale, X, From what we have said above, we can identify X_0 with a single real number, x_0, i.e. $X_0 = x_0 I_\Omega$, and X_1 with an element of $\mathbb{R}^2$, (x_1', x_2'), i.e. $X_1 = x_1' I_{E_1} + x_2' I_{E_2}$. Now, as X is a martingale $M_0(X_1) = X_0$ and $\mathbb{E}(X_1) = \mathbb{E}(X_0)$. So

$$x_1' \mathbb{P}(E_1) + x_2' \mathbb{P}(E_2) = x_0 \quad \text{—} \quad (1)$$

Now let's think about the set of martingales, $\{W^1, .., W^m\}$ that we want to employ to represent {\em any} martingale X. Up to time $t=1$ our representation would look like,

$$X_1 = X_0 + \sum_{i=1}^{m}\int_0^1 g_s^i \, dW_s^i$$

$$= X_0 + \sum_{i=1}^{m} g_0^i (W_1^i - W_0^i) \qquad \text{———} \qquad (2)$$

but we don't know the value of m here let alone how to compute the various W^i's. A short investigation, which follows, gives us a clue.

Suppose that we can prove a representation theorem for martingales for which $X_0 = 0$. That is we can find a fixed set, $\{W^1, ., W^m\}$, such that for every such X, there exist $g^1, ., g^m$ predictable processes with

$$X_t = \sum_{i=1}^{m} \int_0^t g_s^i \, dW_s^i .$$

Then taking a martingale Y and forming $Y - Y_0$ gives us a process, $X = Y - Y_0$ and $X_0 = 0$, So applying the representation theorem, there are $g^1, ., g^m$ such that,

$$Y_t - Y_0 = \sum_{i=1}^{m} \int_0^t g_s^i \, dW_s^i$$

ie.

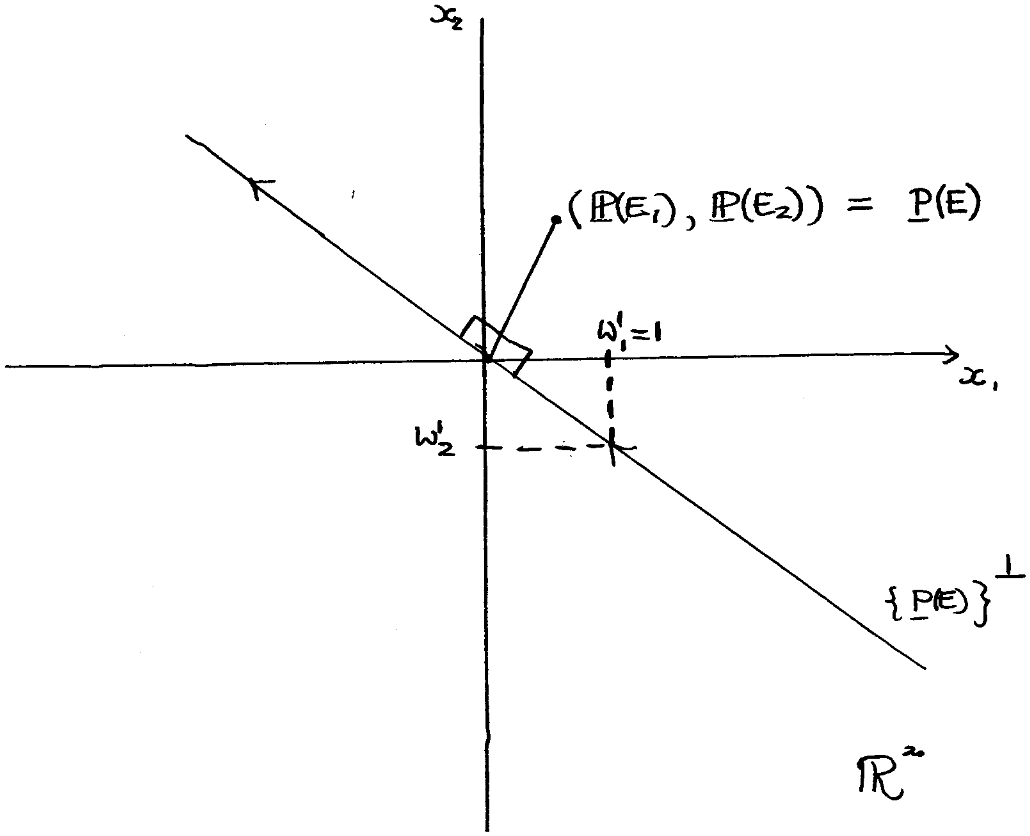
$$Y_t = Y_0 + \sum_{i=1}^{m} \int_0^t g_s^i \, dW_s^i$$

And we have a general representation theorem. So we consider the case $X_0 = 0$. Refering back to Eq 1 this states that

$$x_1^1 \mathbb{P}(E_1) + x_2^1 \mathbb{P}(E_2) = 0 \quad - \quad (3)$$

Put another way (x_1^1, x_2^1) is orthogonal to the vector $(\mathbb{P}(E_1), \mathbb{P}(E_2))$. This is an important observation.

What we have seen here is that the time 1 random variable of every martingale, X,[(†)] corresponds to a vector which is orthogonal to $(\mathbb{P}(E_1), \mathbb{P}(E_2))$. Looking at the situation geometrically we have;



The set of vectors orthogonal to $(\mathbb{P}(E_1), \mathbb{P}(E_2)) \triangleq \underline{P}(E)$ is denoted by $\{\underline{P}(E)\}^{\perp}$. In $\mathbb{R}^2$ the set of vectors orthog- -onal to a (non-zero) vector is a linear subspace of dimension 1[(*)]. In $\mathbb{R}^n$ the set of vectors orthogonal to to a vector, $\underline{v} \neq 0$, is a subspace of dimension, $n-1$. This is an important observation, remember it. Returning to our present example; we want to find m martingales, $W^1, .., W^m$, so that we can write, from equation 2,

$$X_1 = \sum_{i=1}^{m} g_0^i (W_1^i - W_0^i).$$

If we specify that $\{W^1, .., W^m\}$ are not only a strongly

(†) Well, every X for which $X_0 = 0$.
(*) A straight line through the origin.

orthogonal set but also, $W_0^i = 0$, $1 \leq i \leq m$, then the values that determine W_1^i, $1 \leq i \leq m$ correspond to points lying on the line $\{\underline{P}(E)\}^\perp$. Now, strong orthogonality of martingales entails orthogonality of their random variable at each point in time if the martingales are zero at time 0. This follows because: M and N are strongly orthogonal $\Leftrightarrow$ the process $\langle M, N \rangle \equiv 0$. But $MN - \langle M, N \rangle$ is an L^1 martingale and $M_0 N_0 - \langle M, N \rangle_0 = 0$ if $M_0 = N_0 = 0$.(+) So $\mathbb{E}(MN) = \mathbb{E}(\langle M, N \rangle) = 0$ at all times. This means that our random variables, $W_1^1, \ldots, W_m^1$ must

ⓐ Lie in the one dimensional space $\{\underline{P}(E)\}^\perp$ and

ⓑ $W_1^1 \ldots, W_m^1$ must be orthogonal to each other.

This forces the conclusion that $m = 1$, if the set of martingales exists at all. On the other hand it is obvious from the geometry of the situation that $\{\underline{P}(E)\}^\perp$ is spanned by a(ny) single vector. We can choose this vector as we please so long as it resides in $\{\underline{P}(E)\}^\perp$. For example we can specify $W_1^1 \equiv (\omega_1^1, \omega_2^1)^{(*)}$ by

$$\omega_1^1 \mathbb{P}(E_1) + \omega_2^1 \mathbb{P}(E_2) = 0$$

$$\text{and} \quad \omega_1^1 = 1$$

a quick look at the diagram above shows that this will give us a vector in $\{\underline{P}(E)\}^\perp$, indeed we could set $\omega_1^1 = \alpha \in \mathbb{R} \setminus \{0\}$, as we pleased: it doesn't matter, the vector we obtain would simply be a scaling

(+) In fact you dont need $M_0 = N_0 = 0$, $\mathbb{E}(MN) = \mathbb{E}(\langle M, N \rangle)$ and $\langle M, N \rangle = 0$ is enough.

(*) I'm writing $W_1^1 \equiv (\omega_1^1, \omega_2^1)$ to indicate that W_1^1 is identified with the pair (ω_1^1, ω_2^1). I might forget and write $W_1^1 = (\omega_1^1, \omega_2^1)$ occasionally.

of $W'_1 \equiv \left(1, \frac{-\mathbb{P}(E_1)}{\mathbb{P}(E_2)}\right)$ (we've solved for w'_2). Certainly

$$X_1 = g_0(W'_1 - W'_0) = g_0 W'_1 \quad (\text{since } W'_0 = 0)$$

for a suitable choice of g_0. Indeed if $X_1 \equiv (x'_1, x'_2)$ and $x'_1 \mathbb{P}(E_1) + x'_2 \mathbb{P}(E_2) = 0$ then $g_0 = -\frac{x'_2 \mathbb{P}(E_2)}{\mathbb{P}(E_1)}$

and $X_1 = g_0 W'_1$ ie. $(x_1, x_2) = -x'_2 \frac{\mathbb{P}(E_2)}{\mathbb{P}(E_1)} \left(1, -\frac{\mathbb{P}(E_1)}{\mathbb{P}(E_2)}\right)$.

We now know W'_0 and W'_1. To see how we can successively 'biuld up' W' for each time step we now consider things up to time step $t = 2$. What we want to choose W'_2 so that

ⓐ $M_1(W'_2) = W'_1$

ⓑ For each martingale, X, for which $X_0 = 0$, we have

$$X_2 = \int_0^2 g_s \, dW'_s \equiv g_0(W'_1 - W'_0) + g_1(W'_2 - W'_1)$$

for some suitable process, (g_0, g_1). Here g_0 is $\mathcal{F}_0$ measurable and g_1 is $\mathcal{F}_1$ measurable.

We run ahead of ourselves a little here: How do we know that it will still be possible to represent X with a single martingale, W'? The answer, which is affirmative, reveals itself when we consider what happens on each cell of the partition which generates $\mathcal{F}_1$. Luckily there are only two sets, E_1 and E_2. To make the notation simple choose one of E_1 and E_2 and call it E for the remainder of this section. Now the splitting index is 2 so E divides into two sets, we might just as well call these E_1 and E_2 — don't confuse these will the E_1 and E_2 that preceded them.[(†)] Consider any martingale, X, at time $t=2$. On the sets E_1, E_2, the random variable X_2 takes the numerical values x_1, x_2, respectively. Since X is a martingale, $M_1(X_2) = X_1$ and because we are dealing with sets in a partition this relationship has a nice form.

Lemma 1

Let X_1 take the value x' on the cell E. The cell E splits into E_1 and E_2 on which X_2 takes the values x_1^2 and x_2^2, respectively. We have

$$(x_1^2 - x')\,\mathbb{P}(E_1) + (x_2^2 - x')\,\mathbb{P}(E_2) = 0$$

so that $(x_1^2 - x',\ x_2^2 - x')$ is orthogonal to the vector $(\mathbb{P}(E_1), \mathbb{P}(E_2))$.

Pf For any $H \in \mathcal{F}_1$, $\langle M_1(X_2), I_H \rangle = \langle X_2, I_H \rangle$. In particular, if H runs over the sets of the partition that generates $\mathcal{F}_1$, writing $X_2 = \sum_E (x_1^2 I_{E_1} + x_2^2 I_{E_2})$ we see that,

$$\langle M_1(X_2), I_E \rangle = \langle x_1^2 I_{E_1} + x_2^2 I_{E_2}, I_E \rangle .$$

But also $M_1(X_2)$ is $\mathcal{F}_1$ measurable and so it takes a sole numerical value on the cell E. Now, $M_1(X_2) = x'$ on E. Then

$$\langle M_1(X_2), I_E \rangle = \int_\Omega M_1(X_2) I_E \, d\mathbb{P} = \int_\Omega x' I_E \, d\mathbb{P} = x'\,\mathbb{P}(E)$$

and

$$= \langle x_1^2 I_{E_1} + x_2^2 I_{E_2}, I_E \rangle = x_1^2\,\mathbb{P}(E_1) + x_2^2\,\mathbb{P}(E_2).$$

(†) I'm doing this for a reason; the general argument considers a cell, E, at time t. The cell E splits into $s(E)$ cells at time $t+1$. I want to apply a <u>single</u> argument across <u>all</u> cells at <u>all</u> times. So <u>which</u> cell and <u>which</u> time are unimportant. My notation introduces and reflects this fact.

So, $x' = \dfrac{x_1^2 \mathbb{P}(E_1) + x_2^2 \mathbb{P}(E_2)}{\mathbb{P}(E)}$ ie, you arrive at the expectation by averaging X_2 over E, for each cell E of $\mathcal{F}_1$. It follows that $(x_1^2 - x')\mathbb{P}(E_1) + (x_2^2 - x')\mathbb{P}(E_2) = 0$ because $\mathbb{P}(E) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$.

$\square$

<u>Corollary</u>

If X is a martingale and E is a cell of $\mathcal{F}_1$ splitting into cells E_1, E_2 of $\mathcal{F}_2$ and we identify the values of X_2 on the cells E_1, E_2 with an element (x_1^2, x_2^2) of $\mathbb{R}^2$ then writing $\underline{x} = (x_1^2, x_2^2)$, $\underline{P}(E) = (\mathbb{P}(E_1), \mathbb{P}(E_2))$, $\underline{I} = (1,1)$ and indicating scalar products with "$\cdot$", we have: The vector $\underline{x}$ is characterised by

$$\underline{x} - \frac{\underline{x}\cdot\underline{P}(E)}{\underline{P}(E)\cdot\underline{I}}\,\underline{I} \text{ is orthogonal to } \underline{P}(E)$$

<u>Pf</u> Obvious; $x' = \dfrac{\underline{x}\cdot\underline{P}(E)}{\underline{P}(E)\cdot\underline{I}}$.

<u>Corollary</u> The subspace of $\mathbb{R}^2$ consisting of vectors orthogonal to $\underline{P}(E)$ is $\left\{ \underline{x} - \dfrac{\underline{x}\cdot\underline{P}(E)}{\underline{P}(E)\cdot\underline{I}}\,\underline{I} \; : \; \underline{x} \in \mathbb{R}^2 \right\}$.

<u>Pf</u> Exercise

Once again we have arrived at the space of vectors which are orthogonal to a fixed vector ($\underline{P}(E)$). This is a <u>one dimensional</u> subspace and is spanned by any one of its non-zero vectors. We will now choose this vector so that it is $W_2^1 - W_1^1$ that is to say, so that $M_1(W_2^1) = W_1^1$ on E. But is this possible?

In fact we have already demonstrated that this is so at time $t = 1$. The argument is identical to that with 1, only some of the terms change. From Lemma 1, if W_2^1 exists then we know that

$$(w_1^2 - w^1)\mathbb{P}(E_1) + (w_2^2 - w^1)\mathbb{P}(E_2) = 0$$

that is, $w_1^2\mathbb{P}(E_1) + w_2^2\mathbb{P}(E_2) = w^1\mathbb{P}(E)$. This is a linear equation in the variables w_1^2 and w_2^2. So we <u>set</u> $w_1^2 = 1$ and then $w_2^2 = \dfrac{w^1\mathbb{P}(E) - \mathbb{P}(E_1)}{\mathbb{P}(E_2)}$.

Here w^1 is the value W^1 took on the (general) cell E. Looking back, the values for w^1 were "1 and $-\dfrac{\mathbb{P}(E_1)}{\mathbb{P}(E_2)}$" I've put this in quote marks because it refers back to a time when E_1 and E_2 meant something different from what they do now; the cell E we speak of above is either "E_1" or "E_2" in this formula for w^1. The point is that you can calculate w_2^2 explicitly in terms of the values $w^1, \mathbb{P}(E), \mathbb{P}(E_1), \mathbb{P}(E_2)$.

But is $M_1(W_2^1) = W^1$ on E?

<u>Lemma 2</u> With everything as above; $M_1(I_{E_i}) = \dfrac{\mathbb{P}(E_i)}{\mathbb{P}(E)} I_E$

<u>Pf</u> From the proof to Lemma 1, for any random variable $X_2 = x_1^2 I_{E_1} + x_2^2 I_{E_2}$,

$$\langle M_1(X_2), I_E \rangle = x^1\mathbb{P}(E) = x_1^2\mathbb{P}(E_1) + x_2^2\mathbb{P}(E_2)$$

and $M_1(X_2)$ takes a single numerical value on E, x^1, so

$$x^1 = \frac{x_1^2\mathbb{P}(E_1) + x_2^2\mathbb{P}(E_2)}{\mathbb{P}(E)}$$

On $\Omega\setminus E$ we have $\langle M_1(X_2), I_{\Omega\setminus E}\rangle = \langle X_2, I_{\Omega\setminus E}\rangle = 0$ because E_1 and E_2 lie in E — not $\Omega\setminus E$. It follows that the single numerical value that $M_1(X_2)$ takes on $\Omega\setminus E$ is 0. By taking $x_1^2 = 1$, $x_2^2 = 0$ and then $x_1^2 = 0$, $x_2^2 = 1$ we arrive at,

$$M_1(I_{E_i}) = \frac{\mathbb{P}(E_i)\, I_E}{\mathbb{P}(E)}, \quad i = 1,2.$$

This says that you obtain the expectation of a cell by multiplying its 'parent' by its relative strength (w.r.t. $\mathbb{P}$) in the parent(†).

$\square$

Returning to $M_1(W_2^1)$ on E: We want

$$M_1\left(w_1^2 I_{E_1} + w_2^2 I_{E_2}\right)$$

by the Lemma

$$= \left(w_1^2 \frac{\mathbb{P}(E_1)}{\mathbb{P}(E)} + w_2^2 \frac{\mathbb{P}(E_2)}{\mathbb{P}(E)}\right) I_E$$

$$= w' I_E, \quad \text{i.e.} \quad M_1(W_2^1) = W_1^1$$

because E is any cell of $\mathcal{F}_1$.

(†) A parent in the bacteriological sense! And $\mathbb{P}(E_i)/\mathbb{P}(E)$ is your 'relative strength'.

To prove that we can represent any martingale, X, up to time $t=2$ as a stochastic integral w.r.t. W' i.e. is enough to show that we can find a predictable g_X such that

$$g_X(W^1_2 - W^1_1) = X_2 - X_1$$

because this is,

$$\int_1^2 g'_X(s)\,dW^1_s = X_2 - X_1$$

and we already know that there is g such that

$$\int_0^1 g(s)\,dW^1_s = X_1 - X_0$$

(remember we're doing the case $X_0 = 0$). Putting the parts together gives a g_X (patched together from $g(s)$ and $g'_X(s)$) such that

$$X_2 = X_0 + \int_0^2 g(s)\,dW^1_s .$$

So let us consider the problem of finding g_X:

<u>Lemma 3</u>

Identify W^1 with vectors in $\mathbb{R}^2$ on each cell, E of $\mathcal{F}_1$. The 'vector' $W^1_2 - \dfrac{W^1_2 \cdot P(E)\, I}{P(E) \cdot I}$ spans the set of vectors $\underline{x} - \dfrac{\underline{x} \cdot P(E)\, I}{P(E) \cdot I}$.

Pf $\quad W^1_2 - \dfrac{W^1_2 \cdot P(E)\, I}{P(E) \cdot I} = \left(\omega^2_1 - \omega^1,\ \omega^2_2 - \omega^1\right)$

Putting in the values for w_1^2 and w_2^2 gives

$$(w_1^2 - w^1,\ w_2^2 - w^1) = \left(1 - w^1,\ \frac{w^1 \mathbb{P}(E) - \mathbb{P}(E_1)}{\mathbb{P}(E_2)} - w^1\right)$$

$$= \left(1 - w^1,\ (w^1 - 1)\frac{\mathbb{P}(E_1)}{\mathbb{P}(E_2)}\right)$$

Taking on $\underline{x} = (x_1^2, x_2^2)$ the vector $\underline{x} - \dfrac{\underline{x} \cdot \mathbb{P}(E)\, I}{\mathbb{P}(E) \cdot I}$

has the form $(x_1^2 - x^1,\ x_2^2 - x^1)$ with

$$x^1 = \frac{x_1^2\, \mathbb{P}(E_1) + x_2^2\, \mathbb{P}(E_2)}{\mathbb{P}(E)}$$

So $x_2^2 - x^1 = (x^1 - x_1^2)\dfrac{\mathbb{P}(E_1)}{\mathbb{P}(E_2)}$ and

$$(x_1^2 - x^1,\ x_2^2 - x^1) = \left(x_1^2 - x^1,\ (x^1 - x_1^2)\frac{\mathbb{P}(E_1)}{\mathbb{P}(E_2)}\right)$$

Now set $\lambda = \dfrac{x_1^2 - x^1}{1 - w^1}$ and we have

$$\lambda(w_1^2 - w^1,\ w_2^2 - w^1) = \left(\frac{x_1^2 - x^1}{1 - w^1}\right)\left(1 - w^1,\ (w^1 - 1)\frac{\mathbb{P}(E_1)}{\mathbb{P}(E_2)}\right)$$

$$= \left(x_1^2 - x^1,\ (x^1 - x_1^2)\frac{\mathbb{P}(E_1)}{\mathbb{P}(E_2)}\right)$$

$$= (x_1^2 - x^1,\ x_2^2 - x^1)$$

$\square$

In the last Lemma it is implicitly assumed that $\omega' \neq 1$. In fact this is one of the values chosen earlier for W_1' when determining this random variable — see pp8. However, this was only one of a continuum of possible values for ω_1' and ω_2' and so it is possible to choose values which make the Lemma valid. What this demostrates is that choices at time $t+1$ are not independent of the choices at time t. We could get round the problem by redefining ω_1^2, we set $\omega_1^2 = 1$ on pp 12. It turns out that this is a possible value of ω', so if we specify that ω_1^2 is a value not equal to ω' then the argument of the Lemma holds. I had originally chosen the values for ω_1' to coincide with those given in a treatment by M. Dothan of the Martingale Representation Theorem. She seems to be unaware of the restrictions required on later choices, however her treatment is descriptive and not a formal proof so perhaps I am asking too much.

We are now able to verify that; in the case that the splitting index is uniform over sets and time and the index is equal to two then, up to time $t = 2$, any martingale, X(+), has a representation with respect to W'. The last bit is easy; on each cell, E, of $\mathcal{F}_1$ we find the scalar, g_E, such that (identifying $X \upharpoonright E$ with elements of $\mathbb{R}^2$...)

$$X_2 - \frac{X_2 \cdot \underline{P}(E)\, \underline{I}}{\underline{P}(E)\cdot \underline{I}} = g_E \left(W_2' - \frac{W_2' \cdot \underline{P}(E)\, \underline{I}}{\underline{P}(E)\cdot \underline{I}} \right)$$

This defines an $\mathcal{F}_1$ measurable g_1 ($= g_E$ on the cell E)

and we have $X_2 = g_0(W_1' - W_0') + g_1(W_2' - W_1')$.

(+) We stick with the case $X_0 = 0$.

We can now turn to a more general case. We suppose the splitting index $S(E)$ is uniform over sets and time. We investigate the changes occurring between times t and $t+1$. Any cell, E, in $\mathcal{F}_t$ splits into $S(E)$ cells of $\mathcal{F}_{t+1}$, and we consider an $(\mathcal{F}_t)$ martingale, $(X_t) \equiv X$. On the $\mathcal{F}_t$ cell E the martingale X takes a single numerical value which we denote by x^1. As E splits into $S(E)$ cells at time $t+1$ then X_{t+1} take values $x_1^{t+1}, \dots, x_s^{t+1}$, and here we have introduce a contraction of the notation, $s = S(E)$ (this is partly hindsight and partly instinct).
Now the martingale property: $M_t(X_{t+1}) = X_t$, can be restricted to the cell, E; $M_t(X_{t+1})I_E = M_t(X_{t+1}I_E) = X_t I_E$. Which, if we say that E splits into $\mathcal{F}_{t+1}$ cells, $E_1, E_2, \dots, E_s$, can be written,

$$M_t\left(\sum_{i=1}^{s} x_i^{t+1} I_{E_i}\right) = x^1 I_E$$

or

$$M_t\left(\sum_{i=1}^{s} (x_i^{t+1} - x^1) I_{E_i}\right) = 0$$

since $I_E = \sum_{i=1}^{s} I_{E_i}$, and, therefore

$$\mathbb{E}\left(M_t\left(\sum_{i=1}^{s} (x_i^{t+1} - x^1) I_{E_i}\right)\right) = \sum_{i=1}^{s} (x_i^{t+1} - x^1)\mathbb{P}(E_i) = 0.$$

This states that the vector of "martingale differences on E", $X_{t+1}(E) - X_t(E) \triangleq (x_i^{t+1} - x^1) \in \mathbb{R}^s$ is orthogonal to the single vector $\underline{P}(E) = (\mathbb{P}(E_1), \mathbb{P}(E_2), \dots, \mathbb{P}(E_s)) \in \mathbb{R}^s$. Indeed we have:

<u>Lemma 4</u>

Let $\underline{P}(E) \in \mathbb{R}^s$ be the vector defined above. The subspace of $\mathbb{R}^s$ consisting of vectors orthogonal to $\underline{P}(E)$ is

$$\{\underline{P}(E)\}^{\perp} = \left\{ \underline{x} - \frac{\underline{x}\cdot\underline{P}(E)\,\underline{I}}{\underline{P}(E)\cdot\underline{I}} \;:\; \underline{x}\in\mathbb{R}^{s} \right\}.$$

Here $\underline{I} = (1,1,1,\ldots,1)\in\mathbb{R}^{s}$ and "$\cdot$" denotes the scalar product in $\mathbb{R}^{s}$. Moreover every element of $\{\underline{P}(E)\}^{\perp}$ defines a "martingale difference on E".

Pf

If $\underline{v}\in\{\underline{P}(E)\}^{\perp}$ then $\underline{v}\cdot\underline{P}(E) = 0$ and so

$\underline{v} = \underline{v} - \dfrac{\underline{v}\cdot\underline{P}(E)\,\underline{I}}{\underline{P}(E)\cdot\underline{I}}$. On the other hand

$$\left(\underline{x} - \frac{\underline{x}\cdot\underline{P}(E)\,\underline{I}}{\underline{P}(E)\cdot\underline{I}}\right)\cdot\underline{P}(E) = \underline{x}\cdot\underline{P}(E) - \left(\frac{\underline{x}\cdot\underline{P}(E)}{\cancel{\underline{P}(E)\cdot\underline{I}}}\right)\cancel{\underline{P}(E)\cdot\underline{I}}$$

$$= 0.$$

Take any element of $\{\underline{P}(E)\}^{\perp}$, $\underline{x} - \dfrac{\underline{x}\cdot\underline{P}(E)}{\underline{P}(E)\cdot\underline{I}}\underline{I}$, and

set $X_{t+1} = \begin{cases} x_i \text{ on the cell } E_i, & 1\le i\le s \\ 0 \text{ on every other cell in } \mathcal{F}_{t+1} \end{cases}$ and set X_t

equal to $x' = \dfrac{\underline{x}\cdot\underline{P}(E)}{\underline{P}(E)\cdot\underline{I}}$ on the cell E and zero

elsewhere in $\mathcal{F}_t$. Recall the action of the conditional expectation M_t on cells in $\mathcal{F}_{t+1}$ — you obtain the conditional expectation by multiplying the 'parent' in $\mathcal{F}_t$ by the cells 'relative strength', i.e. $M_t(I_{E_i}) = \dfrac{\mathbb{P}(E_i)}{\mathbb{P}(E)} I_E$

So $M_t(X_{t+1} I_E) = M_t\left(\sum_{i=1}^{s} x_i I_{E_i}\right) = \left(\sum_{i=1}^{s} \dfrac{x_i\,\mathbb{P}(E_i)}{\mathbb{P}(E)}\right) I_E$

$$= \frac{\underline{x} \cdot \underline{P}(E)}{\underline{P}(E) \cdot \underline{I}} I_E = x' I_E .$$ On any other cell, F of $\mathcal{F}_t$ we have $M_t(X_{t+1} I_F) = M_t(0) = 0$ because X_{t+1} is only non-zero on the cells spawned from E. So $M_t(X_{t+1}) = X_t$ and our element of $\{\underline{P}(E)\}^\perp$ does indeed define a martingale difference on E.

$\square$

Lemma 4 provides us with some vital information: The set of martingale differences on E can be identified with a subspace of $\mathbb{R}^s$ consisting of vectors orthogonal to $\underline{P}(E)$. This subspace has dimension $s-1$ because $\langle \underline{P}(E) \rangle$, the subspace generated by $\underline{P}(E)$ has dimension 1 and

$$\langle \underline{P}(E) \rangle \oplus \langle \underline{P}(E) \rangle^\perp = \mathbb{R}^s .$$

Accordingly, we will be able to find a 'basis' for the martingale differences consisting of $s-1$ <u>orthogonal</u>(†) martingale differences whose linear span(†) is the space of all martingale differences. Let us write $W^i_{t+1}(E) - W^i_t(E)$, $1 \leq i \leq s-1$, for this basis, on the cell E. For any martingale difference on E, $X_{t+1}(E) - X_t(E)$, say, we can find scalars $g^i(E)$ such that:

(†) Once we have identified the martingale differences with vectors in $\mathbb{R}^s$ orthogonal will mean "at right angles" but will it imply <u>strong</u> orthogonality of the martingales we construct? See later. Also, we need a <u>minimum</u> of $s-1$ of these basis elements and any more than $s-1$ is unnecessary!

$$X_{t+1}(E) - X_t(E) = \sum_{i=1}^{s-1} g^i(E)\left(W^i_{t+1}(E) - W^i_t(E)\right)$$

If we write $\sum_E$ to indicate the sum across all cells of the partition that generates $\mathcal{F}_t$ then,

$$X_{t+1} - X_t = \sum_E \left(X_{t+1}(E) - X_t(E)\right) = \sum_E \sum_{i=1}^{s-1} g^i(E)\left(W^i_{t+1}(E) - W^i_t(E)\right)$$

$$= \sum_{i=1}^{s-1} \sum_E g^i(E)\left(W^i_{t+1}(E) - W^i_t(E)\right)$$

$$= \sum_{i=1}^{s-1} g^i_t \left(W^i_{t+1} - W^i_t\right)$$

Here g^i_t is the $\mathcal{F}_t$ random variable that takes the value $g^i(E)$ on the $\mathcal{F}_t$-cell E.

Remark The choice of 'basis' martingale differences here is quite arbitrary. We know a basis exists and consists of 'orthogonal' differences which are $s-1$ in number. When we are applying this theory to Finance it could be very useful to pick out a particular kind of basis — which is appropriate to the problem being considered — a bit like choosing a frame of reference for a problem in mechanics which reflects some symmetry of the situation. But, I speculate .... .

The way in which we have put together our basic martingales, (W^i_t), is also somewhat arbitrary. We have written them as a list on the cell E

i.e., $W^1_{t+1}-W^1_t$, $W^2_{t+1}-W^2_t$, ..., $W^{s-1}_{t+1}-W^{s-1}_t$. If we move to the time period $t+1$ to $t+2$ exactly the same kind of thing happens on each of the s cells $E_1, \ldots, E_s$, we get a 'list' of $s-1$ differences — one for each cell — and call the first $W^1_{t+2}(E_i)-W^1_{t+1}(E_i)$, the second $W^2_{t+2}(E_i)-W^2_{t+1}(E_i)$, etc. But the labels 1, 2, 3, etc are arbitrary, one could rearrange the actual choices made for the differences and it would not affect our basic result. As a final remark; we would define $W^i_0 \equiv 0$, $1 \leq i \leq s$, and W^i_t is just the sum of the differences up to time t.

We turn to the final requirement for our basis martingales. Are they strongly orthogonal already or can they be chosen to be strongly orthogonal? Notice first of all that one we had identified the martingale differences with an $s-1$ dimensional subspace of $\mathbb{R}^s$ we went on to choose a basis for this subspace consisting of orthogonal vectors in this subspace(+). As we remarked in a footnote to pp 19 we do not yet know if this "Euclidean orthogonality" entails strong orthogonality of the associated martingales. There are clearly occasions when this last proposition is true. In order to demonstrate this and to facilitate a clear understanding of the issues for our situation we quote a "well know" result concerning the strong orthogonality of martingales.

<u>Theorem</u>

Let M, N be martingales on $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. Write

(+) The Gram-Schmidt Theorem states that they exist.

$M \perp\!\!\!\perp N$ to indicate that M and N are strongly orthogonal.
Then,

$$M \perp\!\!\!\perp N \iff MN \text{ is a martingale with } M_0N_0 = 0 \iff \langle M, N\rangle = 0$$

$$\iff M_0N_0 = 0 \text{ and } \mathbb{E}(M_\tau N_\tau) = 0 \quad \forall \text{ stop time}, \tau.$$

Pf See P. Kopp, Martingales and Stochastic Integrals. CUP 1984.

So consider now the martingale differences $W^i_{t+1} - W^i_t$, $W^j_{t+1} - W^j_t$, for $i \neq j$. As usual, it is enough to look at things on a cell, E, in $\mathcal{F}_t$. Each of these differences can be identified with vectors in $\mathbb{R}^s$, which are orthogonal to $\underline{P}(E) = (\mathbb{P}(E_1), \ldots, \mathbb{P}(E_s))$. The vectors have the form, $\underline{x} - \dfrac{\underline{x} \cdot \underline{P}(E)\,\underline{I}}{\underline{P}(E)\cdot\underline{I}}$, (Lemma 4),
So we could think of $W^j_{t+1} - W^j_t$ as $(\omega^j_{t+1}(\ell) - \omega'_j)$ and $W^i_{t+1} - W^i_t$ as $(\omega^i_{t+1}(\ell) - \omega'_i)$, on the cell E. They were chosen to be orthogonal in the sense that $\sum_{\ell=1}^{s} (\omega^i_{t+1}(\ell) - \omega'_i)(\omega^j_{t+1}(\ell) - \omega'_j) = 0$. The product, $(W^i_t W^j_t)$ will be a martingale iff for each t,

$$M_t(W^i_{t+1} W^j_{t+1}) = W^i_t W^j_t \iff$$

$$M_t(W^i_{t+1} W^j_{t+1} - W^i_t W^j_t) = 0 \iff$$

$$M_t(W^i_{t+1} W^j_{t+1} - W^i_t W^j_{t+1} - W^i_{t+1} W^j_t + W^i_t W^j_t) = 0 \iff$$

here we have used the fact that W^i and W^j are martingales

$\Leftrightarrow$

$$M_t\left((W^i_{t+1} - W^i_t)(W^j_{t+1} - W^j_t) \right) = 0 .$$

Now $(W^i_{t+1} - W^i_t)(W^j_{t+1} - W^j_t) I_E = \sum_{l=1}^{s} (w^i_{t+1}(l) - w'_i)(w^j_{t+1}(l) - w'_j) I_{E_l}$

and so if $M_t\left((W^i_{t+1} - W^i_t)(W^j_{t+1} - W^j_t) I_E \right) = 0$ we must have

$$\sum_{l=1}^{s} (w^i_{t+1}(l) - w'_i)(w^j_{t+1}(l) - w'_j) \frac{\mathbb{P}(E_l)}{\mathbb{P}(E)} I_E = 0 ,$$ that is,

$$\sum_{l=1}^{s} (w^i_{t+1}(l) - w'_i)(w^j_{t+1}(l) - w'_j) \mathbb{P}(E_l) = 0 ,$$

for every cell E ~~spawning~~ spawning cells $E_1, \ldots, E_s$. Moreover this is sufficent to ensure W^i and W^j are strongly orthogonal. We can see now that if $\mathbb{P}(E_r) = \mathbb{P}(E_1)$ for $1 \leqslant r \leqslant s$ then the 'Euclidean orthogonality' that we employed in our choice of vectors will entail strong orthogonality of W^i, W^j (for $i \neq j$ of course). But can we always choose a strongly orthogonal set? The answer is yes, and here is how.

We start with our identification of the martingale differences with the subspace of $\mathbb{R}^s$ given by,

$$\{\mathbb{P}(E)\}^{\perp} = \left\{ \underline{x} - \frac{\mathbb{P}(E) \cdot \underline{x}}{\mathbb{P}(E) \cdot \underline{I}} \underline{I} \; : \; \underline{x} \in \mathbb{R}^s \right\}$$

We will write a typical vector in this subspace as $(x_i - x')$, we've seen this before! Now this space is $s-1$ dimensional — this is an algebraic fact. So if we introduce a 'new' scalar product on $\{\underline{P}(E)\}^\perp$ by means of

$$\langle (x_i - x'), (y_i - y') \rangle_{P(E)} \triangleq \sum_{i=1}^{s} (x_i - x')(y_i - y') \mathbb{P}(E_i)$$

then one can verify that this is a positive definite symmetric bilinear form on $\{\underline{P}(E)\}^\perp$ — it is an inner product. This inner product doesn't 'destroy' any non zero vectors in $\{\underline{P}(E)\}^\perp$: If

$$\langle (x_i - x'), (x_i - x') \rangle = 0 \text{ then } \sum_{i=1}^{n} (x_i - x')^2 \mathbb{P}(E_i) = 0$$

each $P(E_i) > 0$ (for the E_i's are cells of a partition)

So $|x_i - x'|^2 = 0 \ \forall i$, So the inner product just assigns different lengths to vectors in $\{\underline{P}(E)\}^\perp$ and different <u>angles</u> between vectors. What we do now is to choose $s-1$ vectors in $\{\underline{P}(E)\}^\perp$ which are;

ⓐ non-zero,

ⓑ Orthogonal with respect to $\langle \cdot , \cdot \rangle_{P(E)}$.

It is then true that these $s-1$ vectors are linearly independent and span $\{\underline{P}(E)\}^\perp$.

We will choose our martingale differences, $W^i_{t+1} - W^i_t$, $1 \leq i \leq s-1$ in this fashion. All of what we have said before holds true with the additional property that W^i and W^j are strongly orthogonal when $i \neq j$.