

## STOCHASTIC AREA FOR BROWNIAN MOTION ON THE SIERPINSKI GASKET

By B. M. HAMBLY<sup>1</sup> AND T. J. LYONS<sup>2</sup>

*University of Edinburgh and Imperial College, London*

We construct a Lévy stochastic area for Brownian motion on the Sierpinski gasket. The standard approach via Itô integrals fails because this diffusion has sample paths which are far rougher than those of semimartingales. We thus provide an example demonstrating the restrictions of the semimartingale framework. As a consequence of the existence of the area one has a stochastic calculus and can solve stochastic differential equations driven by Brownian motion on the Sierpinski gasket.

1. Introduction. The trajectories of Brownian motion on the Sierpinski gasket are rougher than their classical counterparts and do not constitute a semimartingale in the usual coordinate chart. Nonetheless we will prove here that the analogue of Lévy's area process still exists and that the Brownian path, together with this area, has finite homogeneous  $p$ -variation for any  $p > (\log 5)/(\log 2) > 2$ .

Our motivation for constructing the Lévy area for this class of processes and establishing its  $p$ -variation lies in the pivotal role an area process plays in the interpretation and solution of differential equations of the form

$$(1.1) \quad dY_t = \sum_{i=1}^K f^i(Y_t) dX_t^i, \quad Y_0 = y.$$

1.1. *Background.* If  $X$  is a smooth path and the vector fields  $f^i$  are Lipschitz, then (1.1) can be interpreted classically and possesses a unique smooth solution  $Y$ . We will refer to this functional  $X \rightarrow Y$  as the Itô functional.

For at least two generations, probabilists have been interested in widening the class of  $X$  for which the Itô functional is defined. The extension is trivial in the case where  $K = 1$  or the vector fields commute and can be dismissed as uninteresting, for in this special case the Itô functional is continuous in the uniform topology. Extension in the general case of noncommuting vector fields is a quite different sort of problem, and one must overcome some serious obstructions.

Itô provided the most celebrated and most important achievements in this direction by regarding (1.1) as an identity relating *random variables*. Exploiting his integration theory, he proved the existence and uniqueness of a random

---

Received June 1996; revised May 1997.

<sup>1</sup>Supported in part by EC Contract Stochastic Analysis CT92-0784.

<sup>2</sup>Supported by EPSRC Senior Fellowship B/93/SF/445, EPSRC Contract Dirichlet Processes in Analysis GR/J55946 and EC Contract Stochastic Analysis CT92-0784.

*AMS 1991 subject classifications.* Primary 60J60; secondary 60J65, 60J25.

*Key words and phrases.* Stochastic area, differential equations, fractals.

process  $Y$  solving the equation whenever  $X$  was a semimartingale (or with law a measure absolutely continuous with respect to a semimartingale measure). In this case one can readily show that  $Y$  is also a semimartingale.

Now, [6] and [7] prove that the classical solutions to (1.1), obtained by taking dyadic polygonal interpolations to the Brownian motion, converge almost surely to the Stratonovich solution (which usually requires construction using the Itô technology). This provides a reasonable approach to extending the Itô functional from smooth paths to Brownian and more generally semimartingale paths.

However, this Itô–Stratonovich extension is certainly not a continuous extension, and has certain problems associated with it:

1. one can prove that for any norm on path space which is small enough to carry Wiener measure the functional is discontinuous almost surely;
2. it is probabilistic and relies on Itô calculus and its extension to semimartingales.

In particular, the lack of pathwise definition creates problems of interpretation, and a priori the null set of problem paths may depend on the equations making the construction of flows difficult. There seemed to be little scope to extend the theory beyond the semimartingale setting.

In [3] one of the authors sets out a deterministic approach to the interpretation and solution of equations such as (1.1) in the case where  $X_t$  is a rough path. The main conclusion of that paper was that there were relatively simple metrics on the space of smooth paths for which (1.1) defines a uniformly continuous function and for which the completion can be described concretely. Loosely, the smoothness of a path is measured in terms of  $p$ -variation, and for every  $p \geq 1$  the Itô functional is continuous. For  $p > 2$  the completion contains in its support Wiener measure. The essential difference from earlier work is that the approach becomes *increasingly nonlinear* with increasing  $p$  and is already nonlinear when  $p = 2$ . The completion comprises enhanced paths. For  $3 > p \geq 2$  the nonlinearity is quadratic. The Itô–Stratonovich equation is a continuous function of a Brownian path and its Lévy area taken as a pair.

In this paper, we show that Brownian motion on the Sierpinski gasket is a second nontrivial stochastic model to fit the approach introduced in [3] so that there is always a natural solution to (1.1). Our approach is to construct a Lévy area for this process and to establish the appropriate Hölder bounds, after which it fits the hypotheses of [3]. To us it is an interesting example; it is definitely too rough to be a semimartingale and so classical approaches are not available, but still the deterministic approach yields a result. It is also provocative, as there are a number of variations on the theme of diffusions on fractals, and simply looking at area will not suffice to deal with all of them. (We do not explore this question further in the current paper.)

1.2. *Why Lévy area.* The technical aspects of this paper are devoted entirely toward constructing and analyzing Lévy area for Brownian motion on the Sierpinski gasket. However, in this section we explain briefly, but reason-

ably precisely, a simple special case of the theorem in [3] so the reader can understand how the parts fit together.

**DEFINITION 1.1.** Let  $X_t$  be a smooth path, and consider the chordal area process

$$A_{st}(X) = \frac{1}{2} \int \int_{s < u_1 < u_2 < t} dX_{u_1}^i dX_{u_2}^j - dX_{u_1}^j dX_{u_2}^i, \quad s, t \in [0, L].$$

The path  $X_t$  is said to have *finite  $p$ -variation* if for every  $[s, t] \subset [0, L]$  one has

$$\omega(s, t) := \sup_{\mathcal{D} = \{s \leq \dots \leq t_i < t_{i+1} \leq \dots \leq t\}} \sum_i |X_{t_{i+1}} - X_{t_i}|^p + \sum_i |A_{t_i, t_{i+1}}|^{p/2} < \infty.$$

It is said to be regular if  $\omega$  is continuous and zero on the diagonal.

**DEFINITION 1.2.** The path has *regular  $p$ -variation controlled by  $\nu(s, t)$*  if  $\nu$  is continuous, is zero on the diagonal and satisfies  $\nu(s, u) + \nu(u, t) \leq \nu(s, t)$  for  $u \in [s, t]$  and if

$$\omega < \nu.$$

The definition above gives a way to talk about paths being close.

**DEFINITION 1.3.** A sequence of smooth paths  $X_t^{(n)}$  is  *$\nu$ -Cauchy* if all the paths are controlled by  $\nu$ , and if for each  $\varepsilon > 0$  there is an  $N$  such that, for all  $m, n > N$ ,

$$\varepsilon > \sup_{\mathcal{D} = \{s \leq \dots \leq t_i < t_{i+1} \leq \dots \leq t\}} \left( \sum_i |(X_{t_{i+1}}^{(n)} - X_{t_i}^{(n)}) - (X_{t_{i+1}}^{(m)} - X_{t_i}^{(m)})|^p + \sum_i |A_{t_i, t_{i+1}}^{(n)} - A_{t_i, t_{i+1}}^{(m)}|^{p/2} \right)$$

**THEOREM 1.4.** *If  $3 > p \geq 2$ , the sequence  $X_t^{(n)}$  is  $\nu$ -cauchy, and if the vector fields are Lipschitz of smoothness  $\gamma > p$ , then the sequence of solutions  $Y^{(n)}$  to the equations driven by  $X^{(n)}$  is also  $C\nu$ -Cauchy for some constant  $C > 1$ .*

The approach is graded and for smoother vector fields can handle rougher paths. The full result shows that the solution can be defined as a uniformly continuous function in a  $p$ -variation norm involving *the first  $[p]$  iterated integrals of  $X$* . The classical Itô–Stratonovich equation can be recovered by applying this functional to the Brownian path together with its Lévy area. Consider the polygonal approximations used by Lévy and Cieselski in their construction of Brownian motion [2], the  $n$ th of which agrees with the path at the times  $k/2^n$  and interpolates linearly in between. With probability 1, this sequence of polygonal paths converges in the  $p$ -variation metric to Brownian motion, and moreover [5] the area processes associated with the polygonal paths converge to the Lévy area in the appropriate topology.

Therefore we have convergence of the solutions to (1.1) obtained when these polygonal paths successively take the role of the driving term  $X_t$  (and by the results of [6] and [7] we can identify this natural extension of (1.1) with the Stratonovich solution).

This paper establishes a result equivalent to that of Lévy for Brownian motion, and consequently fits into the framework in the same simple way.

**1.3. Conclusion.** This approach to defining and solving Stratonovich differential equations does not go via the conventional semimartingale approach [4]—and offers hope that one could study other classes of Markov processes in similar ways. Brownian motion on the Sierpinski gasket seems an appropriate simple example to study. It is rougher than classical Brownian motion but still has variation less than 3. It has natural polygonal approximations, obtained by considering its trace on the natural graph approximations to the gasket. One could consider a differential equation driven by these successive polygonal approximations and ask if the solutions converged as the level of the approximation increases. To prove this it suffices to show that, with probability 1, the polygonal approximations and their area processes both converge in the appropriate  $p$ -variation metric.

This paper establishes that this is the case. With probability 1 we construct a natural area process associated with Brownian motion on the Sierpinski gasket and show that it and the underlying path have finite  $p$ -variation. Hence we show that, almost surely, the pair satisfy the hypotheses for solving (1.1); the equation has meaning and a unique solution. In addition we study the tail behavior of this new Lévy area.

The construction does involve a semimartingale—but *with respect to the  $\sigma$ -field associated with revealing the level  $n$  trace of the path*, not the filtration which observes the path up to time  $t$ . The argument is a mixture of cancellation and careful accumulation of terms to avoid interference.

**2. Preliminaries for the Sierpinski gasket.** The Sierpinski gasket is a fractal subset of  $\mathbb{R}^2$  formed as the fixed point of a set of contraction maps. Let  $a_1 = (0, 0)$ ,  $a_2 = (1, 0)$ ,  $a_3 = (\frac{1}{2}, \sqrt{3}/2)$ . We write  $G_0$  for  $\{a_1, a_2, a_3\}$ , the set of three vertices, and  $A_0$  for the closed convex triangle with vertices  $G_0$ . Let  $\psi_i$ :  $i = 1, 2, 3$  be the contraction given by  $\psi_i(x) = \frac{1}{2}(x + a_i)$ . For any set  $A \subset \mathbb{R}^2$  we set

$$\Psi(A) = \bigcup_{i=1}^3 \psi_i(A)$$

and

$$\Psi_n(A) = \Psi(\Psi_{n-1}(A)).$$

Initially let  $G_n = \Psi_n(G_0)$  and  $A_n = \Psi_n(A_0)$ . The Sierpinski gasket  $G$  is the fixed point of the sequence of mappings  $\Psi(G) = G$ .

By inverting the contraction  $\psi_1$  the fractal can be extended to infinity. We now set  $G^{(n)} = \psi_1^{-n}(G_n)$  and form

$$G^{(\infty)} = \bigcup_{n=0}^{\infty} G^{(n)}.$$

We will use  $G^{(\infty)}$  for the union of this set with its reflection in the  $y$ -axis. We will now modify the definition of  $G_n$  by setting  $G_n = 2^{-n}G^{(\infty)}$ , so creating a  $\psi$ -invariant structure. Let  $G_{\infty} = \bigcup_n G_n$  and the infinite Sierpinski gasket  $G = \text{cl}(G_{\infty})$ .

Take  $G^0$  to be the graph with vertices  $G^{(0)}$ , where two vertices  $x, y$  are connected by an edge if  $|x - y| = 1$  and the line segment joining  $x, y$  is in  $G$ . Further define  $G^n = 2^{-n}G^0$  to be the  $n$ -level graph. In this way  $G^n$  is an infinite graph in which connected vertices on level  $n$  are  $2^{-n}$  apart and each vertex has four neighbors. Every edge in  $G^n$  is the edge of a unique triangle of edge length  $2^{-n}$  whose interior contains elements of  $G^{n+1}$ . It also forms part of the boundary of a unique triangular component of the compliment of  $G$  which we will refer to as the complimentary triangle associated with the edge.

We can orient the edges of the graph  $G^n$ . Define a function  $\sigma_{xy}$  on directed edges  $(x, y)$  taking only the values  $\pm 1$ , by setting  $\sigma_{xy} = 1$  if, traversing the edge  $(x, y)$  from  $x$  to  $y$ , one finds the neighboring complimentary triangle on the left, and  $\sigma_{xy} = -1$  if it is on the right.

**LEMMA 2.1.** *If the edges of the graph  $G^n$  are assigned this orientation, then the sum of the orientations on each edge at any vertex will be 0.*

**PROOF.** By observing the reflection symmetry in the  $y$ -axis the sum must be zero at the origin. Using only rotations and translations, any vertex and its two neighboring triangles in the structure can be mapped to the origin and its neighboring triangles. This does not change any orientation, or the sum, and hence the sum must be zero at any vertex.  $\square$

In [1], Barlow and Perkins constructed a continuous strong Markov process  $X_t$  on  $G$ , and due to its uniqueness with respect to local isometries it is now called the Brownian motion on  $G$ . We will need to consider the path of this process started from an arbitrary point  $x$  and use a polygonal approximation formed as the trace of the path on  $G_n$ . In order to do this we define the following sequence of stopping times:

$$\begin{aligned} T_0^n &= \inf \{t \geq 0: X_t \in G_n\}, \\ T_i^n &= \inf \{t > T_{i-1}^n: X_t \in G_n \setminus \{X_{T_{i-1}^n}\}\}, \quad i > 0. \end{aligned}$$

The positions of the process at each of the stopping times are written  $x_i^n = X_{T_i^n}$ , and this allows us to define a filtration as the trace of the path on  $G_n$ ,

$$\mathcal{F}_n = \sigma(x_i^n: i \geq 0).$$

The path of the random walk approximation to the process on  $G_n$  is contained in this filtration. We may extend the trace to a continuous path using linear interpolation,

$$x_s^n = (i + 1 - s)x_i^n + (s - i)x_{i+1}^n, \quad i < s < i + 1, \quad 0 \leq i.$$

Clearly  $|x_i^n - X_{\tau(t)}| < 2^{-n}$  for some appropriate piecewise smooth reparameterization of time.

**REMARK 2.2.** The trace of the path can be constructed in a variety of apparently different ways. For example, one might consider the sequence of times corresponding to last exits:

$$\begin{aligned} \tilde{T}_0^n &= \sup\{t \leq T: X_t \in G_n\}, \\ \tilde{T}_i^n &= \sup\{t < \tilde{T}_{i-1}^n: X_t \in G_n \setminus \{X_{\tilde{T}_{i-1}^n}\}\}, \quad i > 0; \end{aligned}$$

then  $x_{\tilde{T}_i^n}^n = X_{\tilde{T}_i^n}$ , where  $I = \#\{i: \tilde{T}_i^n > -\infty\}$ .

The paths  $x^{n+1}$  and  $x^n$  agree at the vertices in the path  $x^n$ . In the filtration  $\mathcal{F}_n$  the polygonal path  $x^{n+1}$  consists of a sequence of independent copies of a rescaled path on a basic domain  $\mathcal{C}$  (see Figure 1), for each edge traversed by  $x^n$ .

As the paths are nested we let  $\phi_i^{n, n+r}$  be the index for the level  $n+r$  path corresponding to  $i$  on the level  $n$  path,

$$x_i^n = x_{\phi_i^{n, n+r}}^{n+r}, \quad i = 1, \dots$$

Then  $\phi_i^{n, n+r} - \phi_{i-1}^{n, n+r}$  are independent and identically distributed random variables and  $\phi_i^{n, n+1} - \phi_{i-1}^{n, n+1}$  has the distribution of the number of steps  $Z$  a simple random walk takes to exit the domain  $\mathcal{C}$  when started at  $p$  (see Figure 1).

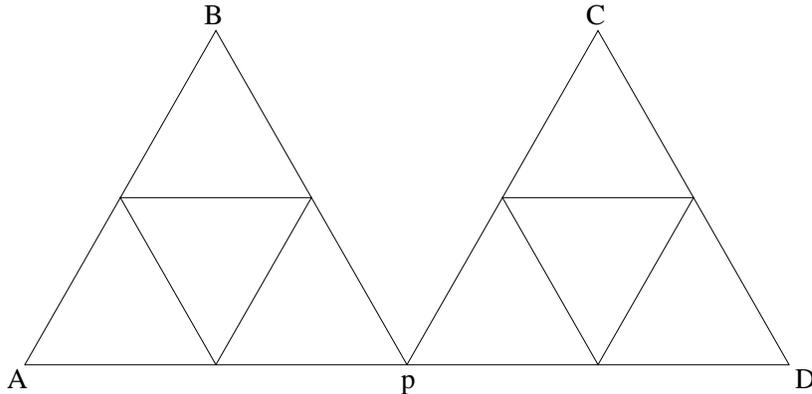


FIG. 1. The basic domain  $\mathcal{C}$ .

THEOREM 2.3 [1]. *The random variable  $Z$  has probability generating function*

$$f(u) := \mathbf{E}u^Z = \frac{u^2}{4 - 3u},$$

and, for a fixed  $m$ ,  $\{Z_n; n \geq 0\} = \{\phi_i^{m, m+n} - \phi_{i-1}^{m, m+n}; n \geq 0\}$  is a Galton–Watson process with offspring distribution given by  $Z$ .

Thus the number of steps in the level  $n$  path can be regarded as a supercritical Galton–Watson process. This branching process has an associated martingale.

THEOREM 2.4 [1].  *$Z_m 5^{-m}$  is a martingale and  $Z_m 5^{-m} \rightarrow W$  a.s. and in  $L^2$ .*

The random variable  $W$  represents the time taken by the limiting continuous Brownian path to traverse the unit triangle.

LEMMA 2.5. *The moments  $\mathbf{E}W^k < \infty$  for all  $k > 0$ .*

PROOF. We know  $Z$  has an exponential tail; precisely the explicit form of the generating function for the Galton–Watson process shows that  $\mathbf{E} \exp(\lambda Z) < \infty$  for  $\lambda < \ln(\frac{4}{3})$ . From the functional equation for  $\phi(\theta) = \mathbf{E}(\exp(-\theta W))$ ,

$$\phi(5\theta) = f(\phi(\theta)),$$

it is easy to see that  $W$  also has an exponential tail. Hence we have the existence of all the moments.  $\square$

3. Constructing a stochastic area. Suppose  $x_t, t \in [0, L]$ , is a piecewise smooth path in  $\mathbb{R}^2$ . Then, as in Definition 1.1, the two-parameter process

$$A_{st}(x) = \frac{1}{2} \int_{s < u_1 < u_2 < t} dx_{u_1}^1 dx_{u_2}^2 - dx_{u_1}^2 dx_{u_2}^1, \quad s, t \in [0, L],$$

is called the area process associated with  $x$ . Consider the closed directed curve defined by  $x_u, u \in [s, t]$ , and the chord  $\overrightarrow{x_t x_s}$ . The area it encloses is precisely  $A_{st}(x)$  providing orientation and multiplicity are taken into account.

Note that  $A(x)$  is dependent on the geometry of the path  $x$  and not on the speed of passage, so that if  $\tau(t)$  is a continuous time change,

$$A_{s,t}(x \circ \tau) = A_{\tau(s), \tau(t)}(x).$$

Let  $X_t, t < \infty$ , be Brownian motion on  $G$ . Recall that  $\{x_i^n; i \geq 0\}$  is the trace of  $X_t$  on  $G^n$  and that we may produce a piecewise linear function by interpolation. From [1] it is known that, with probability 1,  $x_{[5^n t]}^n \rightarrow X_t$  locally uniformly, and so one approach to defining  $A_{st}(X)$  would be to prove that

$A_{5^s, 5^t}(x^n)$  converges. However, constructing  $A_{st}(x)$  directly is a little awkward because the coupling of  $x^n, x^{n+1}$  is far better understood on a certain random time interval.

Recall that  $x_i^n$  is a simple random walk on  $G^n$  and that  $\phi_i^{n, n+m}$  embeds  $x_i^n$  as the  $G^n$  trace of  $x_i^{n+m}$ . Our initial intention is to prove that

$$\lim_{m \rightarrow \infty} A_{\phi_i^{n, n+m} \phi_j^{n, n+m}}(x^m)$$

exists in a well-behaved way. In other words we compute the “area” of  $X_t$  between the  $i$ th and  $j$ th visit to the set  $G^n$ . By doing this simultaneously for all  $i, j$  and  $n$  we are quite close to constructing  $A_{st}(X)$  for all  $s, t$ . If we prove that the paths  $(x^n, A(x^n))$  converge in  $p$ -variation norm, then we will have successfully constructed  $A_{st}$  for all  $s, t$ .

Now fix attention on the path  $X_t$  during its first passage from one site in  $G^0$  to another. We have the successive approximations  $x_u^n$ , where  $\phi_0^{0, n} < u < \phi_1^{0, n}$ . Let  $A^n = A_{\phi_0^{0, n} \phi_1^{0, n}}(x_u^n)$ . We study the convergence of  $A^n$ ; rescaling will give us the behavior of the areas on other intervals.

Now  $A^0 = 0$ , so  $A^n = \sum_{j=1}^n A^j - A^{j-1}$ . Ignoring the special shape of  $G^n$  for the moment, one has the schematic picture shown in Figure 2.

At each stage the difference between the areas of the two polygons can be expressed as a sum:

$$A^{n+1} - A^n = \sum_{i=0}^{Z_n-1} A_{\phi_i^{n, n+1} \phi_{i+1}^{n, n+1}}(x^{n+1}),$$

where  $Z_n$  is the number of intervals between points of the trace  $x^n$  (and 1 less than the number of points). Each of the terms in the right-hand side represents the area of one of the shaded regions in the figure.

Recall that  $\sigma_{xy}$  is an anticlockwise orientation for the graph  $G^n$ .

LEMMA 3.1. *The random variables*

$$\eta_{i, n} := \sigma_{x_i^n x_{i+1}^n} 2^{2n} A_{\phi_i^{n, n+1} \phi_{i+1}^{n, n+1}}(x^{n+1})$$

are independent over  $i$  and identically distributed for all  $i, n$ .

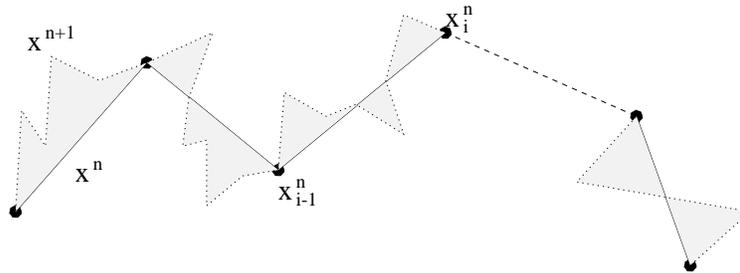


FIG. 2. *The increase in area due to the addition of rescaled paths.*

PROOF. The points  $x_i^n, x_{i+1}^n$  are adjacent points in  $G^n$ ; fix the connected component  $\mathcal{C}$  of  $x_i^n$  in  $G^{n+1} \setminus (G^n \setminus \{x_i^n\})$ . Up to a rotation by a multiple of  $60^\circ$  one has the situation of Figure 1 where  $x_{i+1}^n$  is one of the four vertices in  $G^n$  labelled A, B, C or D in Figure 1. We have already remarked that the random variables  $Z(i+1) = \phi_{i+1}^{n, n+1} - \phi_i^{n, n+1}$  are independent and identically distributed given  $\mathcal{F}_n$ , for  $X_0 \in G^0$ . Also that  $x_j^{n+1}, \phi_i^{n, n+1} < j < \phi_{i+1}^{n, n+1}$ , given  $\mathcal{F}_n$ , is a simple random walk  $\{y_j: j \geq 0\}$  on  $\mathcal{C}$  conditioned to leave  $\mathcal{C}$  at  $x_{i+1}^n$ , and independent of the process interpolating  $\{x_k^n, x_{k+1}^n\}, k \neq i$ .

From these observations we just need to determine the area associated with the walk  $y$ . Now the rotation of  $\mathcal{C}$  does not change  $A_{0Z}(y)$  and rescaling  $\mathcal{C}$  by  $2^n$ , to be of unit edge size, does not affect the process at all, but involves a scaling of the area by  $2^{2n}$ .

Thus  $\sigma_{x_i^n x_{i+1}^n} 2^{2n} A_{\phi_i^n \phi_{i+1}^n}(x^{n+1})$  are clearly independent given  $\mathcal{F}_n$  and have one of at most four apparently distinct distributions according to whether  $x_{i+1}^n$  is A, B, C or D after the rotation.

The path starting from  $p$  makes a number of loops back to  $p$ , each one contained in one of the triangles, followed by a single terminal excursion. The area generated by this path is the sum of contributions from each loop and from the final excursion. If we regard  $\mathcal{C}$  as a pair of rigid triangles hinged at  $p$ , then we can rotate each triangle through  $30^\circ$  so as to achieve a figure with horizontal and vertical reflection symmetry. As each loop is wholly contained within a single triangle, this transformation leaves the area process unchanged. Exploiting this symmetry it is clear that if  $A^x$  is the law of the area (on the unit fractal) given the process  $\{y_j: j \geq 0\}$  quits  $\mathcal{C}$  at  $x \in \{A, B, C, D\}$ , then  $A^A = -A^D$  in distribution, and similarly that  $A^B = -A^C, A^A = -A^B, A^C = -A^D$ .

Now  $\sigma_{x_i^n A} = \sigma_{x_i^n C} = 1$  and  $\sigma_{x_i^n B} = \sigma_{x_i^n D} = -1$ ; the result follows.  $\square$

Our idea for constructing  $A^\infty$  is to prove that  $A^n$  is a semimartingale with respect to  $\mathcal{F}_n$  and, by controlling the mean and variance of the increments  $A^{n+1} - A^n$ , to establish that the semimartingale is convergent.

**THEOREM 3.2.** *The area  $\lim A^n = A$  exists; the convergence is in every  $L^p$ .*

PROOF. Consider the Doob–Meyer decomposition of the semimartingale  $A^n$ :

$$(3.1) \quad \lim_{n \rightarrow \infty} A^n = \sum_0^\infty (A^{n+1} - A^n - \mathbb{E}(A^{n+1} - A^n \mid \mathcal{F}_n)) + \sum_0^\infty \mathbb{E}(A^{n+1} - A^n \mid \mathcal{F}_n).$$

For the martingale part we are able to characterize the appropriate bracket process precisely in terms of the Galton–Watson process. Writing  $\sigma_i$  for  $\sigma_{x_i x_{i+1}}$  and  $\eta_i$  for  $\eta_{i, n}$  so that  $\sigma_i \eta_i 2^{-2n} = A_{\phi_i^{n, n+1} \phi_{i+1}^{n, n+1}}(x^{n+1})$ , recall that

$$A^{n+1} - A^n = \sum_{j=0}^{Z_n-1} \sigma_j \eta_j 2^{-2n}.$$

Hence

$$\mathbf{E}(A^{n+1} - A^n | \mathcal{F}_n) = \sum_{j=0}^{Z_n-1} \sigma_j \mathbf{E} \eta_j 2^{-2n},$$

and we can write the bracket process as

$$\sum_{n=0}^{\infty} [A^{n+1} - A^n - \mathbf{E}(A^{n+1} - A^n | \mathcal{F}_n)]^2 = \sum_{n=0}^{\infty} \left( \sum_{j=0}^{Z_n-1} \sigma_j (\eta_j - \mathbf{E}(\eta_j)) \right)^2 2^{-2n}.$$

Observing that  $\sigma_j = \pm 1$  and writing  $Y_j = \eta_j - \mathbf{E}(\eta_j)$ , we have that

$$N_k = \sum_{j=0}^{k-1} \sigma_j Y_j$$

is a discrete stochastic integral. We can write, using the Burkholder and triangle inequalities, that, for  $p > 2$ ,

$$\begin{aligned} \left\| \sum_m^{\infty} (A^{n+1} - A^n - \mathbb{E}(A^{n+1} - A^n | \mathcal{F}_n)) \right\|_p &\leq c_p \left\| \left( \sum_m^{\infty} 16^{-i} N_{Z_i}^2 \right)^{1/2} \right\|_p \\ &\leq c_p \left( \sum_m^{\infty} 16^{-i} \|N_{Z_i}\|_p^2 \right)^{1/2}. \end{aligned}$$

Controlling  $N$  by its bracket we have, by the inequality of Burkholder,

$$\begin{aligned} \|N_{Z_i}\|_p^2 &\leq c_p \| [N]_{Z_i} \|_{p/2} \\ &\leq c_p \left( \mathbb{E} \left( \sum_{j=0}^{Z_i-1} Y_j^2 \right)^{p/2} \right)^{2/p} \\ &\leq c_p 5^i \left\| W^* \frac{1}{Z_i} \sum_{j=0}^{Z_i-1} Y_j^2 \right\|_{p/2} \end{aligned}$$

for all  $i \geq m$ , where  $W^* = \sup_n Z_n 5^{-n}$  is the maximal random variable. Finally, by the  $L^p$ -version of the strong law of large numbers and the  $L^p$ -maximum inequality,

$$\left\| \sum_m^{\infty} (A^{n+1} - A^n - \mathbb{E}(A^{n+1} - A^n | \mathcal{F}_n)) \right\|_p \leq C_p [\mathbb{E}(W^{p/2})]^{1/p} \left( \sum_m^{\infty} (2^{-4} 5)^n \right)^{1/2}.$$

For the bounded variation part, as  $Z_n/5^n$  is a convergent martingale with a nonzero limit  $W$  there is an apparent danger that the second term in (3.1) might diverge. However, if we look at the sequence  $S_j = \sum_0^j \sigma_i$ , then  $S_j$  is a simple random walk (however, be cautious, this is not true if, e.g., one conditions by  $\mathcal{F}_1$ ). Thus it can be controlled,  $|S_j| \leq K j^{1/2} \psi(j)$ , where  $\psi(j) = \sqrt{\log \log(j+2)}$  and  $K$  is a positive random variable with an exponential tail independent of  $j$ .

Using this the bounded variation part can be controlled by

$$(3.2) \quad \left| \sum_m^\infty \left( \mathbb{E}(A^{n+1} - A^n) \mid \mathcal{F}_n \right) \right| \leq E(\eta) \sum_m^\infty 4^{-n} K_n \sqrt{Z_n} \psi(Z_n),$$

where the  $K_n$  are distributed like  $K$  but need not be independent. Since  $Z_n \leq 5^n W^*$  and  $\psi$  is slowly varying, the right-hand side of (3.2) converges strongly and is of order

$$\left[ \frac{5}{16} \right]^{-m/2} \tilde{K} \sqrt{W^*} \psi(W^*),$$

where  $\tilde{K} = \sum_n^\infty 4^{-n} K_n 5^{n/2} \sqrt{\log n}$ .

As all moments of  $W$  are finite by Lemma 2.5 it follows easily that our area process converges, and in fact there are rather strong controls.  $\square$

We close this section with a lemma concerning the behavior of the area in a basic domain  $\mathcal{L}$ .

**LEMMA 3.3.** *The Laplace transform  $L(\theta) = E \exp(-\theta A^1)$  of the area  $A^1$  generated by the random walk on  $G_1$  is given by*

$$L(\theta) = -2 \frac{e^{-\sqrt{3}\theta/8} (4e^{-3\sqrt{3}\theta/16} + 3e^{-\sqrt{3}\theta/8} + 8e^{-\sqrt{3}\theta/16} + 15)}{e^{-\sqrt{3}\theta/4} + 6e^{-3\sqrt{3}\theta/16} - 74e^{-\sqrt{3}\theta/8} + 6e^{-\sqrt{3}\theta/16} + 1}.$$

**PROOF.** The area  $A^1$  is generated by a random walk on  $G_1$ . The area increment at each step is a multiple of  $\sqrt{3}/16$ , the area of a triangle in  $G_1$ . By the symmetry of the four boundary points we need only consider the exit process from any one of them. We can derive a set of equations by conditioning on the first step of the random walk on  $G_1$ . Let  $L_a(\theta) = E(\exp(-\theta A^1) \mid X_0 = a)$  and write  $N(a)$  for the vertices of  $G_1$  which are neighbors of  $a$  in the graph. Then

$$L_a(\theta) = \sum_{b \in N(a)} p_{ab} L_b(\theta) \exp(-\theta A_{01}) \quad \forall a \in \mathcal{L}.$$

For the first level gasket we obtain seven equations in seven unknowns which we solve to give the result.  $\square$

We can write  $\tilde{L}(\theta)$  for the Laplace transform of  $\eta_1 = \sigma_1 4A^1$ . This is equivalent to the symmetrized version of the transform which can be written as

$$\tilde{L}(\theta) = - \frac{8 \cosh(\sqrt{3}\theta/4)^3 + 3 \cosh(\sqrt{3}\theta/4)^2 - 2 \cosh(\sqrt{3}\theta/4) + 6}{\cosh(\sqrt{3}\theta/4)^2 + 3 \cosh(\sqrt{3}\theta/4) - 19},$$

for  $|\theta| \leq \theta_0 = 4.1587\dots$ . The form of the Laplace transform shows that the scaled and symmetrized area random variable  $\eta_1$  has an exponential tail.

4. Area between an arbitrary pair of times and finite  $p$ -variation. Consider a path of our underlying continuous process  $\{X_t: t \in [0, T]\}$ . We wish to establish the existence of a real-valued process  $A_{s,t}$ , which we will call the area process, and show that

$$(4.1) \quad A_{s,u} = A_{s,t} + A_{t,u} + \text{Area}(\overrightarrow{X_s X_t X_u}),$$

where  $\text{Area}(\overrightarrow{X_s X_t X_u}) = \frac{1}{2}[X_t - X_s, X_u - X_t]$  is the area of the oriented triangle defined by the three points in the path, and that  $(X_t, A_{s,t})$  has finite  $p$ -variation. That is to say if  $\mathcal{D} = \{s \leq \dots \leq t_i < t_{i+1} \leq \dots \leq t\}$  is a partition of  $[s, t]$ , then

$$\sup_{\mathcal{D}} \sum_i |X_{t_{i+1}} - X_{t_i}|^p + \sum_i |A_{t_i, t_{i+1}}|^{p/2} < \infty.$$

The  $p$ -variation of a path over an interval is independent of the parameterization of the path within the interval and so we can be flexible about the naming of our intervals. If  $x_t^n$  are the coupled random walks we defined previously, linearly interpolated so as to be piecewise linear paths progressing at speed  $5^{-n}$  along the graph  $G^n$ , the functions  $\phi_i^{n, n+m}$  identify the corresponding times at the different levels so that

$$x_i^n = x_{\phi_i^{n, n+m}}^{n+m}$$

and the limiting path  $X_t = \lim x_{5^n t}^n$ . The connection between the time of the  $i$ th move on the piecewise linear path at level  $n$  and the time scale of the limiting path  $X_t$  is given by

$$t_{i,n} = \lim_{m \rightarrow \infty} 5^{-n-m} \phi_i^{n, n+m}.$$

The partitions  $(t_{i,n})_{i=0}^\infty$  are successive refinements of the interval  $[0, T]$  and grow in total number like  $5^n T$ . We will be concerned with the maximum number of time intervals  $\{t_{j, n+m}, j = 1, 2, \dots\}$ , into which an interval  $t_{i,n}$  splits. An easy Borel–Cantelli argument shows that there is a random constant  $C$  such that for all  $n$  one has  $\max_i \phi_i^{n, n+1} < nC$ ; of course the mean number  $\phi_i^{n, n+1}$  is 5.

We have already established the existence of an area process  $A_{t_{i,n}, t_{i+1,n}}$  defined separately for each  $n$  and satisfying the algebraic relationship (4.1) to  $X_t$  of an area process at the times that the process is defined.

We can now define  $A_{s,t}$  for any interval  $J = [s, t]$  in  $[0, t_{1,0}]$ . Partition the interval into disjoint intervals of the type  $[t_{i,n}, t_{i+1,n}]$  as follows: vary  $i, n$  and include in the partition exactly those intervals of the type  $[t_{i,n}, t_{i+1,n}]$  which are contained in  $J$  and which are maximal with respect to inclusion. An interval is said to be maximal if it is not contained in any other interval  $[t_{j,m}, t_{j+1,m}] \subset J$ . Label the intervals  $J_{r,n}$  in the partition sequentially according to their level. For simplicity let  $(t_k)_{k=-\infty}^{k=+\infty}$  denote the partition of  $[s, t]$  defined by the endpoints of the intervals.

We can now make a rigorous definition of the area process for all times, and establish  $p$ -variation estimates. Let  $z(u)$  be the piecewise linear function on

$[s, t]$  that agrees with  $X$  at the times  $t_k$ . Then the algebraic property of the area process forces the relation

$$\begin{aligned} A_{t_k, t_l} &= \sum_{j=k}^{j=l-1} A_{t_j, t_{j+1}} + \frac{1}{2} \int_{t_k < u_1 < u_2 < t_l} [dz(u_1), dz(u_2)] \\ &= \sum_{j=k}^{j=l-1} A_{t_j, t_{j+1}} + \frac{1}{2} \sum_{k \leq j_1 < j_2 < l} [X_{t_{j_1+1}} - X_{t_{j_1}}, X_{t_{j_2+1}} - X_{t_{j_2}}], \end{aligned}$$

and providing we are able to take the limits then continuity forces us to the conclusion that

$$(4.2) \quad A_{s, t} = \sum_{j=-\infty}^{j=+\infty} A_{t_j, t_{j+1}} + \frac{1}{2} \sum_{-\infty < j_1 < j_2 < \infty} [X_{t_{j_1+1}} - X_{t_{j_1}}, X_{t_{j_2+1}} - X_{t_{j_2}}],$$

where in general it would be unreasonable to expect the sums in the right-hand expression to be well defined. However, in our case, the series will be absolutely convergent because the time intervals are carefully chosen. Observe that

$$|[X_{t_{j_1+1}} - X_{t_{j_1}}, X_{t_{j_2+1}} - X_{t_{j_2}}]| \leq |X_{t_{j_1+1}} - X_{t_{j_1}}| |X_{t_{j_2+1}} - X_{t_{j_2}}|$$

and hence

$$\frac{1}{2} \sum_{-\infty < j_1 < j_2 < \infty} [X_{t_{j_1+1}} - X_{t_{j_1}}, X_{t_{j_2+1}} - X_{t_{j_2}}] \leq \frac{1}{2} \left( \sum_{-\infty < j < \infty} |X_{t_{j+1}} - X_{t_j}| \right)^2.$$

We have

$$|A_{s, t}|^{p/2} \leq 2^{p/2-1} \left( \sum_{j=-\infty}^{j=+\infty} |A_{t_j, t_{j+1}}| \right)^{p/2} + \frac{1}{2} \left( \sum_{j=-\infty}^{j=+\infty} |X_{t_{j+1}} - X_{t_j}| \right)^p.$$

For  $\{e_i : i \geq 0\}$ ,  $e_i \in \mathbb{R}_+$ , we have a reverse Hölder inequality:

$$\left( \sum_{i=1}^{\infty} i^{-\gamma/(q-1)} \right)^{(q-1)} \sum_{i=1}^{\infty} i^\gamma (e_i)^q \geq \left( \sum_{i=1}^{\infty} e_i \right)^q.$$

Now grouping our intervals according to the level and applying the reverse Hölder inequality with  $q = p/2$  and  $e_n = \sum_r |A_{J_{r,n}}|$ ,  $\sum_r |X_{J_{r,n}}|$  we have that, providing  $p > 2$  and  $\gamma > p - 1$ ,

$$\begin{aligned} |A_J|^{p/2} &\leq 2^{p/2-1} \left( \sum_{n=1}^{\infty} n^{-\gamma/(p/2-1)} \right)^{(p/2-1)} \sum_{n=1}^{\infty} n^\gamma \left( \sum_r |A_{J_{r,n}}| \right)^{p/2} \\ &\quad + \frac{1}{2} \left( \sum_{n=1}^{\infty} n^{-\gamma/(p-1)} \right)^{(p-1)} \sum_{n=1}^{\infty} n^\gamma \left( \sum_r |X_{J_{r,n}}| \right)^p. \end{aligned}$$

As we have noted the sum  $\sum_r |X_{J_{r,n}}|$  is over at most  $nC$  terms, so

$$\left( \sum_r |X_{J_{r,n}}| \right)^p \leq c_1 \sum_r |X_{J_{r,n}}|^p n^{p-1},$$

similarly for  $\sum_r |A_{J_{r,n}}|$ , and hence

$$\begin{aligned} |A_J|^{p/2} &\leq c(\gamma, p, C) \sum_{n=1}^{\infty} \sum_r n^\gamma n^{p/2-1} |A_{J_{r,n}}|^{p/2} \\ &\quad + c(\gamma, p, C) \sum_{n=1}^{\infty} \sum_r n^\gamma n^{p-1} |X_{J_{r,n}}|^p. \end{aligned}$$

Our estimate for  $|A_J|^{p/2}$  is a sum over terms associated with basic intervals contained in  $J$ . We would like to estimate  $\sum_J |A_J|^{p/2}$ , where  $J$  ranges over the intervals in some arbitrary partition of  $[0, t_{1,0}]$ . We can obviously do this by summing over all the basic intervals:

$$\sum_J |A_J|^{p/2} \leq c(\gamma, p, C) \sum_{n=1}^{\infty} n^\gamma n^{p-1} \sum_i (|A_{t_{i,n}, t_{i+1,n}}|^{p/2} + |X_{t_{i,n}, t_{i+1,n}}|^p),$$

where the right-hand side is independent of the partition. The right-hand side is easily seen to be finite almost surely,  $c(\gamma, p, C) < \infty$  almost surely, as we mentioned earlier, and the sum has finite expectation given that  $p > \log 5 / \log 2$ , as

$$\begin{aligned} &\mathbb{E} \left[ \sum_{n=1}^{\infty} n^\gamma n^{p-1} \sum_i (|A_{t_{i,n}, t_{i+1,n}}|^{p/2} + |X_{t_{i,n}, t_{i+1,n}}|^p) \right] \\ &< \mathbb{E} [ (|A_{0, t_{1,0}}|^{p/2} + |X_{0, t_{1,0}}|^p) ] \sum_{n=1}^{\infty} n^\gamma n^{p-1} 5^n 2^{-np}. \end{aligned}$$

We have simultaneously proved the existence and finite  $p$ -variation of the path  $(X, A)$ . It is clear that it satisfies the algebraic relation of a multiplicative functional since it must do so at every basic time pair and they are dense on the line. We have completed the proof of the following theorem.

**THEOREM 4.1.** *The pair  $(X_t, A_{s,t})$  of the Brownian motion and its area process have finite  $p$ -variation for all  $p > \log 5 / \log 2$ .*

An obvious modification of this argument shows that the polygonal paths at level  $n$  and their areas converge in  $p$ -variation to the limiting path (if given the consistent time scales).

We conclude with estimates on the tail of the area between successive visits to different points on  $G^{(\infty)}$  and for all times.

**THEOREM 4.2.** *For the area  $A_{0, t_{1,0}}$  there exist constants  $c, a$  such that*

$$P(A_{0, t_{1,0}} > \lambda) \leq c \exp(-a\lambda).$$

**PROOF.** From the results in Section 3 we can estimate the area  $A_{0, t_{1,0}}$  by

$$A_{0, t_{1,0}} = \sum_{n=0}^{\infty} \sum_{i=0}^{Z_n-1} \sigma_i \eta_i 2^{-2n}.$$

Thus

$$\begin{aligned} P(A_{0,t_{1,0}} > \lambda) &= P\left(\sum_{n=0}^{\infty} \sum_{i=0}^{Z_n-1} \sigma_i \eta_i 2^{-2n} > \lambda\right) \\ &= P\left(\sum_{n=0}^{\infty} \sum_{i=0}^{Z_n-1} \sigma_i \eta_i 2^{-2n} > \sum_{n=0}^{\infty} b_n \lambda\right), \end{aligned}$$

where  $\sum_n b_n = 1$  will be chosen later. Now

$$\begin{aligned} P(A_{0,t_{1,0}} > \lambda) &\leq P\left(\sum_{i=0}^{Z_n-1} \sigma_i \eta_i 2^{-2n} > \lambda b_n \text{ for at least one } n\right) \\ &\leq \sum_{n=0}^{\infty} P\left(\sum_{i=1}^{Z_n} \sigma_i \eta_i 2^{-2n} > b_n \lambda\right). \end{aligned}$$

We can calculate the Laplace transform of the sum by composition of generating functions. Let

$$S_n(\theta) = E \exp\left(\theta \sum_{i=0}^{Z_n-1} \sigma_i \eta_i\right).$$

Then it is easy to see that

$$S_n(\theta) = f_n(\tilde{L}(\theta)),$$

where  $f_n$  is the  $n$ -fold composition of  $f$ , the generating function for the Galton–Watson process. As  $\tilde{L}(\theta)$  is symmetric we have that by Taylor's theorem, for  $\theta < \theta_1 = \{\theta: \tilde{L}(\theta) = \frac{4}{3}\} \leq \theta_0$ ,

$$\begin{aligned} P\left(\sum_{i=0}^{Z_n-1} \sigma_i \eta_i 2^{-2n} > b_n \lambda\right) &\leq \exp(-\theta 4^n b_n \lambda) S_n(\theta) \\ &\leq \exp(-\theta 4^n b_n \lambda) f_n(1 + c_0 \theta^2) \\ &\leq \exp(-\theta 4^n b_n \lambda) (1 + c_1 5^n \theta^2). \end{aligned}$$

Thus for an  $\varepsilon > 0$ , by choosing  $\theta = 5^{-n/2}$  and  $b_n = (1 - \sqrt{5}/4 - \varepsilon)(\sqrt{5}/4 + \varepsilon)^n$ , we have

$$\begin{aligned} P(A_{0,t_{1,0}} > \lambda) &\leq \sum_{n=0}^{\infty} P\left(\sum_{i=0}^{Z_n-1} \sigma_i \eta_i 2^{-2n} > b^n \lambda\right) \\ &\leq \sum_n (1 + c_1) \exp(-(1 - \sqrt{5}/4 - \varepsilon) \lambda \rho^n), \end{aligned}$$

where  $\rho = 1 + \varepsilon(4/\sqrt{5}) > 1$ . As the sum is finite we obtain the required exponential tail.  $\square$

For the area between two arbitrary times we can apply the same argument using the representation given as (4.2) and prove that the tail of the distribution for  $\sqrt{A_{s,t}}$  is no worse than exponential.

Consider the sum grouped according to the intervals

$$\begin{aligned} A_{s,t} &= \sum_{j=-\infty}^{j=+\infty} A_{t_j, t_{j+1}} + \frac{1}{2} \sum_{-\infty < j_1 < j_2 < \infty} [X_{t_{j_1+1}} - X_{t_{j_1}}, X_{t_{j_2+1}} - X_{t_{j_2}}] \\ &\leq \sum_{n=0}^{\infty} \sum_r A_{J_{r,n}} + \frac{1}{2} \left( \sum_{n=0}^{\infty} \sum_r |X_{J_{r,n}}| \right)^2. \end{aligned}$$

Now we consider the two pieces separately. First, the sum of the areas can be controlled by the estimate obtained in Theorem 4.2 using the same technique of proof,

$$P\left(\sum_{n=0}^{\infty} \sum_r A_{J_{r,n}} > \lambda\right) \leq \sum_{n=0}^{\infty} P\left(\sum_r A_{J_{r,n}} > b_n \lambda\right).$$

We can estimate the number of terms at each level. Let  $M_n = \max_i \phi_i^{n, n+1}$ . Then as the terms on level  $n$  in the partition can occur at both ends of the time interval we have

$$\sum_r A_{J_{r,n}} \leq \sum_{i=0}^{2M_n-1} |\eta_i| 2^{-2n}.$$

We know that both  $M_n$  and  $|\eta_i|$  have exponential tails and thus it is easy to show the sum has the indicated tail behavior.

For the second sum, which involves the square of the path increments, we have

$$P\left(\left(\sum_{n=0}^{\infty} \sum_r |X_{J_{r,n}}|\right)^2 > \lambda\right) = P\left(\sum_{n=0}^{\infty} \sum_r |X_{J_{r,n}}| > \sqrt{\lambda}\right).$$

As the increments are fixed by our choice of path dissection, we have

$$P\left(\sum_{n=0}^{\infty} \sum_r |X_{J_{r,n}}| > \sqrt{\lambda}\right) \leq P\left(\sum_{n=0}^{\infty} 2M_n 2^{-n} > \sqrt{\lambda}\right).$$

The random variables  $M_n$  each have an exponential tail, as will their weighted sum and hence we have the following result.

**THEOREM 4.3.** *For the area  $A_{s,t}$  for  $s, t \in [0, t_{1,0}]$  there exist constants  $c_1, \alpha_1$  such that*

$$P(A_{s,t} > \lambda) \leq c_1 \exp(-\alpha_1 \sqrt{\lambda}).$$

## REFERENCES

- [1] BARLOW, M. T. and PERKINS, E. A. (1988). Brownian motion on the Sierpinski gasket. *Probab. Theory Related Fields* 79 543–624.
- [2] LÉVY, P. (1948). *Processus Stochastiques et Mouvement Brownien*. Gauthier-Villars, Paris.
- [3] LYONS, T. J. (1995). Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*. To appear.

- [4] PROTTER, P. (1977). On the existence, uniqueness, convergence and explosions of solutions of systems of stochastic differential equations. *Ann. Probab.* 5 243–261.
- [5] SIPILÄINEN, E.-M. (1993). A pathwise view of solutions of stochastic differential equations Ph.D. dissertation, Univ. Edinburgh.
- [6] WONG, E. and ZAKAI, M. (1965). On the relationship between ordinary and stochastic differential equations. *Internat. J. Engrg. Sci.* 3 213–229.
- [7] WONG, E. and ZAKAI, M. (1965). On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Statist.* 36 1560–1564.

DEPARTMENT OF MATHEMATICS AND STATISTICS  
UNIVERSITY OF EDINBURGH  
THE KING'S BUILDINGS, MAYFIELD ROAD  
EDINBURGH, EH9 3JZ  
UNITED KINGDOM  
E-MAIL: bmh@maths.ed.ac.uk

DEPARTMENT OF MATHEMATICS  
IMPERIAL COLLEGE  
180 QUEEN'S GATE  
LONDON, SW7 2BZ  
UNITED KINGDOM