A GENERALIZED TAYLOR-ARIS FORMULA AND SKEW DIFFUSION

JORGE M. RAMIREZ\textsuperscript{1}, ENRIQUE A. THOMANN\textsuperscript{2}, EDWARD C. WAYMIRE\textsuperscript{3}, ROY HAGGERTY\textsuperscript{4}, AND BRIAN WOOD\textsuperscript{5}

Abstract. This paper concerns the Taylor-Aris dispersion of a dilute solute concentration immersed in a highly heterogeneous fluid flow having possibly sharp interfaces (discontinuities) in the diffusion coefficient and flow velocity. The focus is two-fold: (i) Calculation of the longitudinal effective dispersion coefficient, and (ii) sample path analysis of the underlying stochastic process governing the motion of solute particles. Essentially complete solutions are obtained for both problems.

Key words. Skew-brownian motion, dispersion, heterogeneous media, effective dispersion.

AMS subject classifications. Primary: 60J70. Secondary: 35R05, 47D07, 86A05

1. Introduction. The seminal paper by Taylor [21] together with the refinements by Aris [1] describe the dispersion of a solute concentration immersed in Poiseuille flow directed along the horizontal axis of a cylindrical tube in terms of the tube radius $R$, molecular diffusion coefficient $D$, and the maximum flow $U$ (or cross-sectional average $U/4$) of the parabolic flow profile. In the case $U = 0$ the dispersion coincides with molecular diffusion, and when $D = 0$ the dispersion of solute is aligned with the parabolic profile of the flow. The relative contributions of these combined effects ($D > 0$, $U > 0$) are captured in Taylor’s remarkable insights leading to the formula

$$\bar{D} = 2D + \frac{R^2U^2}{96D}. \quad (1.1)$$

Motivated by considerations of the stability of a viscous liquid to two-dimensional disturbances in a porous-medium, Wooding [24] adapted their analysis to obtain the corresponding formula for dispersion of a solute in a unidirectional parabolic flow between two parallel planes separated by a distance $R$. Accordingly, in this setting

$$\bar{D} = 2D + \frac{8R^2U^2}{945D}. \quad (1.2)$$

In either case, the significance of such dispersion rates for the spread of solute concentration was demonstrated by the derivation of the time-asymptotic (homogenized) equation for the solute concentration averaged over the cross section.

In subsequent mathematical work, Fife and Nichols [8] used perturbation methods for this derivation, while later Bhattacharya and Gupta [5] obtained these results...
using central limit theory techniques from probability. While their approach can be applied to quite more general diffusion (or dispersion) and velocity coefficients than originally formulated by Taylor [21] and Aris [1], they restrict to the case of smoother coefficients than those considered here. Specifically we consider highly heterogeneous media allowing for characteristically \textit{sharp interfaces}, and analyze these effects on effective dispersion rates in the case of longitudinal flow in a cylindrical region. Moreover, we will see that new physical insights into the role of heterogeneities in dispersion emerge from these results, which should be useful in analyzing more complex heterogeneous flow problems. In addition to interest among geophysicists for solute transport in highly heterogeneous porous media (e.g. see Berensten et al. [2]), examples occur in micro-reactor engineering and liquid-liquid extraction (see Ueno et al. [22]), as well as in biological processes at cellular levels, see Saxton and Jacobsen [19].

![Figure 1.1](image)

**Fig. 1.1.** \textit{Three dimensional advection dispersion in a cylinder $G = G \times \mathbb{R}$ with heterogeneous diffusion coefficient. Coordinates are given by $x = (x,y) = (x,y_1,y_2)$}

Our focus is two-fold: (i) Calculation of the longitudinal effective dispersion coefficient, and (ii) determination of the probability laws governing the (stochastic) motion of the underlying solute particles. The model we consider has the following structure. Consider diffusion of a dilute solute on the region $G = \{x = (x,y) \in \mathbb{R}^d : x \in \mathbb{R}, y \in G\}$, where the constant cross section $G$ is a connected compact subset of $\mathbb{R}^{d-1}$ with smooth boundary and $d$ is equal to either 2 or 3 (see Figure 1.1). The velocity profile is a bounded measurable vector function of the form $U(x) = (U(y),0)$, and the diffusion coefficient is uniformly bounded and elliptic tensor $D(x) = D(y)$ in diagonal form with entries $D_x(y)$ and $D_y(y)$. According to Fick’s law, the concentration $c(x,t)$ at location $x = (x,y)$ at time $t > 0$ satisfies

\[
\begin{align*}
\frac{\partial c}{\partial t} &= \nabla \cdot (D \nabla c) - \nabla \cdot (cU) , \quad \text{in } G , \\
(D \nabla c)_{|G} \cdot n &= 0, \quad c|_{t=0} = c_0 .
\end{align*}
\]

An effective diffusion coefficient is derived through a general homogenization result for this problem. As a very interesting special case of the formulation, we consider
longitudinal flow between parallel plates through a layered medium. Namely, the diffusion coefficient $D$ is assumed to be a piecewise constant function of the transverse spatial variable. The presence of sharp interfaces between regions of different diffusivity, creates interesting effects on the motion of the solute particles that can be studied through the structure of the associated diffusion process governing the motion of individual solute particles. For this case, the effective diffusion coefficient is found to be a weighted average of the diagonal values of $D$ and its inverses.

Although our approach follows closely the probabilistic techniques developed by Bhattacharya [4], the mild conditions allowed for $U$ and $D$ lead to new mathematically interesting technical considerations in the identification of the underlying diffusion process associated with the solution to (1.3). This is further resolved through a combination of the theory of Dirichlet forms, partial differential equations, and the Sroock-Varadhan martingale problem and Itô-Tanaka stochastic calculus. The organization of the paper is as follows. In the next section the problem is re-cast in terms of stochastic processes associated with $L^2$-semigroups defined from the Dirichlet form corresponding to the operator $\nabla \cdot (D \nabla c) - \nabla \cdot (c U)$, $\nabla c \mid_{\partial G} \cdot \mathbf{n} = 0$. We then apply the central limit theorem and asymptotic variance formula of Bhattacharya [4] to obtain the generalized Taylor-Aris formula for dispersion. We include the special case of layered media with sharp interfaces as a corollary. The last section of the paper is devoted to a more detailed description of the stochastic motion (i.e. physical process) of the underlying solute particles in the presence of sharp interfaces.

2. Heterogenous dispersion in longitudinal flow.

2.1. The Model. Let $G \subset \mathbb{R}^{d-1}$, $(d = 2, 3)$ be a compact domain with smooth boundary in the sense of Davies [7], pp. 46-47. Consider the cylinder $G = \{x = (x, y) \in \mathbb{R}^d : x \in \mathbb{R}, y \in G\}$. In the case $d = 3$, $y$ denotes the planar vector $y = (y_1, y_2)$. For time $t > 0$, let $c(x, t)$ be the concentration of a solute at $x \in G$ which is diffusing in a fluid with velocity $U(x)$ and through a medium with diffusion tensor $D(x)$. We make the following assumptions on $U$ and $D$:

- the velocity in the cylinder is parallel to the $x$ coordinate, and

$$U(x) = (U(y), 0), \quad (2.1)$$

- $D(x)$ is bounded and uniformly positive definite on $G$, and has the form

$$D(x) = \begin{bmatrix} D_x(y) & 0 \\ 0 & D_y(y) \end{bmatrix}, \quad (2.2)$$

where in the case $d = 3$, $D_y(y)$ denotes the diagonal matrix with entries $D_{y_1}(y_1, y_2)$ and $D_{y_2}(y_1, y_2)$.

These assumptions cover the especially interesting case of diffusion in a medium with sharp interfaces as defined by a piecewise constant cross-sectional diffusion coefficient. Note also that the form (2.1) makes the velocity $U$ incompressible. This property is essential to many of the computations, though we do not always make special note when it occurs. Notice for example that one may write $\nabla \cdot (v U) \equiv U \cdot \nabla v$.

Modeling the flux of the concentration $c$ by Fick’s law leads to the conservation equation

$$\begin{cases}
\partial_t c = \nabla \cdot (D \nabla c) - \nabla \cdot (c U), & \text{in } G, \\
(D \nabla c) \mid_{\partial G} \cdot \mathbf{n} = 0, & c \mid_{t=0} = c_0,
\end{cases} \quad (2.3)$$
where spatial derivatives are to be understood in the weak sense. Specifically, we look for \( c(x, t) \in C^1([0, \infty), H^1(G)) \cap C^1([0, \infty), L^2(G)) \) such that \( c(x, 0) = c_0(x) \) and for \( t \geq 0 \),

\[
\partial_t (u, c(t, \cdot))_{L^2(G)} = -\mathcal{E}(u, c(t, \cdot))
\]

(2.4)

where \( \mathcal{E} \) is the bilinear form naturally associated with the differential equation (2.3),

\[
\mathcal{E}(u, v) = \int_G \nabla u \cdot \nabla v - (U \cdot \nabla u)v \, dx,
\]

(2.5)

and \((u, v)_{L^2(G)}\) denotes the usual inner product in \( L^2(G) \).

2.2. Homogenization Problem. Let \( c(x, t) \) be a solution to problem (2.3), and consider the cross-sectional average,

\[
\bar{C}(x, t) = \int_G c(x, y, t) \, dy.
\]

(2.6)

We seek large scale parameters \( \bar{U}, \bar{D} \) such that on space-time scales \( \lambda x, \lambda^2 t \), the weak limit

\[
\bar{C}(\hat{x}, \hat{t}) \, d\hat{x} = \lim_{\lambda \to \infty} C(\lambda \hat{x} + \bar{U} \lambda^2 \hat{t}, \lambda \hat{t}) \lambda \, d\hat{x}
\]

provides a solution

\[
\bar{C}(x, t) = \bar{C}(x - \bar{U} t, t)
\]

to the homogenized partial differential equation,

\[
\partial_t \bar{C} = \bar{D} \partial^2_x \bar{C} - \bar{U} \partial_x \bar{C}.
\]

(2.7)

Thus the goal is to prove the following theorem.

**Theorem 2.1.** (Generalized Taylor-Aris Formula) Let \( \pi(dy) \) be the uniform probability measure on \( G \), and let \( h \) be a solution in \( L^2(G, \pi) \) to the boundary value problem

\[
\begin{align*}
\nabla_y \cdot (D_y \nabla_y h) & = U(y) - \bar{U}, \quad \text{in} \ G, \\
(D_y \nabla_y h) \cdot n_y & = 0 \quad \text{in} \ \partial G.
\end{align*}
\]

(2.8)

Then, for any \( t > 0, x \in \mathbb{R} \), and Borel measurable \( A \subseteq \mathbb{R} \) with boundary \( \partial A \) such that \( \pi(\partial A) = 0 \),

\[
\lim_{\lambda \to \infty} \int_A C(\lambda x + \bar{U} \lambda^2 t, \lambda^2 t) \lambda \, dx = \int_A \bar{C}(x + \bar{U} t, t) \, dx
\]

(2.9)

with homogenized parameters

\[
\bar{U} = \int_G U(y) \pi(dy), \quad \bar{D} = \int_G \{D_z(y) + (D_y \nabla_y h) \cdot \nabla_y h(y)\} \, \pi(dy).
\]

(2.10)

In the case \( d = 2 \) and \( D_y \) is piecewise constant, we get (see Figure 2.1)
Corollary 2.2. (Diffusion in a layered medium) Assume $d = 2$, $G = [a,b]$ and $\mathbf{D}$ has the form
\[
\mathbf{D}(x) = \mathbf{D}(y) = \sum_{k=-m}^{M} \mathbf{D}^{(k)} 1_{[l_k, l_{k+1})}(y), \quad \mathbf{D}^{(k)} = \begin{bmatrix} D_x^{(k)} & 0 \\ 0 & D_y^{(k)} \end{bmatrix}
\]
where $a = l_{-m} < l_{-m+1} < \cdots < l_M < l_{M+1} = b$ is a collection of interfaces partitioning $[a,b]$. If $D_x^{(k)} > 0$ and $D_y^{(k)} > 0$ for all $k$, then the limit (2.9) of Theorem (2.1) holds with homogenized diffusion coefficient
\[
\bar{D} = \sum_{k=-m}^{M} \left\{ D^{(k)} \frac{l_{k+1} - l_k}{b - a} + \frac{1}{D_y^{(k)}} \int_{l_k}^{l_{k+1}} g(y)^2 \pi(dy) \right\}
\]
where $g$ is given by
\[
g(y) = \int_a^y (U(z) - \bar{U}) \pi(dz).
\]

Fig. 2.1. Two dimensional advection dispersion through a layered medium.

2.3. Probabilistic reformulation. In order to relate the solution of the analytic problem (2.3) to a stochastic process, some special attention is needed to cover the case when the diffusion tensor $\mathbf{D}$ is discontinuous. The following lemma provides the technical basis for the construction of the stochastic process associated with the Dirichlet form (2.5). Let $D[0,\infty)$ be the space of right-continuous functions on $[0,\infty)$ with left-hand limits.

Lemma 2.3. For each path $y \in D[0,\infty)$, define a process $X^{(y)} \equiv \{X_t^{(y)} : t \geq 0\}$ by
\[
X_t^{(y)} = x_0 + \int_0^t \mu(y(s)) \, ds + \int_0^t \sigma(y(s)) \, dB_s,
\]
where $\mu$ and $\sigma$ are bounded measurable functions with $\sigma(z) \geq \delta > 0$ for all $z \in \mathbb{R}$. Let $Y = \{Y_t : t \geq 0\}$ be a Markov process with sample paths in $D[0,\infty)$ having stationary transition probabilities with infinitesimal generator $(A, \mathcal{D}_A)$. Also assume $Y$ is independent of $\{B_t : t \geq 0\}$. Define $X_t = (X_t^{(y)}, Y_t), t \geq 0$. Then $\{X_t : t \geq 0\}$, is a Markov process with stationary transition probabilities whose generator $A$ acts on a core of functions of the form $f(x_1, x_2) = g(x_1)h(x_2)$, where $g$ is twice continuously differentiable, $h \in \mathcal{D}_A$, by
\[
Af(x_1, x_2) = g(x_1)Ah(x_2) + h(x_2) \left\{ \mu(x_2)g'(x_1) + \frac{1}{2} \sigma^2(x_2)g''(x_1) \right\}
\]
Proof. The Markov property follows immediately from the decomposition for \(0 \leq t' < t\)

\[
X_t^{(Y)} = X_{t'}^{(Y)} + \int_{t'}^t \mu(Y_s) \, ds + \int_{t'}^t \sigma(Y_s) \, dB_s,
\]

the independence of \(Y\), \(B\), and the Markov property for \(Y\). Consider the following \(\sigma\)-fields,

\[
\mathcal{G}_t := \sigma(B_s, Y_s, s \leq t), \quad \mathcal{G}_t := \sigma(Y, X_s^{(Y)}, s \leq t), \quad \mathcal{H}_t := \sigma(Y, s \leq t).
\]

By an application of Itô’s lemma one may first write

\[
g(X_t^{(Y)}) = g(x_1) + \int_0^t \left\{ \mu(Y_s) g'(X_s^{(Y)}) + \frac{1}{2} \sigma^2(Y_s) g''(X_s^{(Y)}) \right\} \, ds
\]

\[
+ \int_0^t \sigma(Y_s) g'(X_s^{(Y)}) \, dB_s.
\]

In particular,

\[
G_t := g(X_t^{(Y)}) - \int_0^t \left\{ \mu(Y_s) g'(X_s^{(Y)}) + \frac{1}{2} \sigma^2(Y_s) g''(X_s^{(Y)}) \right\} \, ds
\]

is the martingale \(\int_0^t \sigma(Y_s) g'(X_s^{(Y)}) \, dB_s, \ t \geq 0\), with respect to \(\mathcal{G}_t\). Similarly,

\[
H_t := h(Y_t) - \int_0^t Ah(Y_s) \, ds, \quad t \geq 0,
\]

is a martingale with respect to \(\mathcal{H}_t\). Combining this martingale structure with projective properties of conditional expectations based on \(\mathcal{H}_t \subseteq \mathcal{G}_t \subseteq \mathcal{G}_t\), one may check that

\[
M_t := g(X_t^{(Y)}) h(Y_t) - \int_0^t h(Y_s) \left\{ \mu(Y_s) g'(X_s^{(Y)}) + \frac{1}{2} \sigma^2(Y_s) h''(X_s^{(Y)}) \right\} \, ds
\]

\[
- \int_0^t g(X_s^{(Y)}) Ah(Y_s) \, ds
\]

is a martingale with respect to the filtration \(\mathcal{G}_t\). The result now follows by an application of Stroock-Varadahn theory; see \([20]\). \(\Box\)

Remark 2.4. It is interesting to note that a direct calculation of the generator formally involves the computation of the limit as \(t \downarrow 0\) in the following expression:

\[
\frac{1}{t} \left\{ \mathbb{E}(g(X_t^{(Y)}) h(Y_t) - g(x_1) h(x_2)) \right\} = g(x_1) \frac{1}{t} \mathbb{E} \{ h(Y_t) - h(x_2) \}
\]

\[
+ \frac{1}{t} \int_0^t \mathbb{E} \{ \mu(Y_s) g'(X_s^{(Y)}) h(Y_t) \} \, ds
\]

\[
+ \frac{1}{2t} \int_0^t \mathbb{E} \{ \sigma^2(Y_s) g''(X_s^{(Y)}) h(Y_t) \} \, ds
\]

\[
\rightarrow g(x_1) Ah(x_2) + \mu(x_2) g'(x_1) h(x_2) + \frac{1}{2} \sigma^2(x_2) g''(x_1) h(x_2).
\]

However direct justification of such an approach does not seem possible in the absence of (spatial) continuity of the coefficients \(\mu\) and \(\sigma\). The Stroock-Varadahn theory beautifully exploits the construction of the underlying Markov processes in replacing the calculation of derivatives by calculations of integrals.
In accordance with Lemma 2.3 and as suggested by the differential equation (2.3) in the transversal directions, consider the differential operator \( A \) with domain \( D_A \) given by
\[
Au = \nabla_y \cdot (D_y \nabla_y u), \quad u \in D_A, \\
D_A = \{ u \in H^1(G) : D_y \nabla_y u \in H^1(G), \, D_y \nabla_y u|_{\partial G} \cdot n_y = 0 \}.
\] (2.13)
The bilinear form associated to \( A \) is given by
\[
\mathcal{E}(u, v) = \int_G D_y \nabla_y u \cdot \nabla_y v \, dy, \quad u, v \in D_\mathcal{E} = H^1(G),
\] (2.14)
Standard considerations, see e.g. Example 2.b, page 45 in Ma and Röckner [13], or Nash [15], show that there exists a (strong) Markov process \( Y = \{ Y_t : t \geq 0 \} \) with right continuous paths with left limits (Hunt process) defined on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and associated with \( \mathcal{E} \). Observe that \( \pi(y) = 1/|G| \) is the unique invariant probability for \( Y \), uniquely solving the differential equation \( A\pi = 0 \) (up to non-normalized constants), In particular
\[
T_t f(y) = \mathbb{E}_y f(Y_t), \quad t \geq 0, \, f \in L^2(G).
\] (2.15)
defines an \( L^2 \)-semigroup. Some of the technical properties of this process and its infinitesimal generator are summarized in the following lemma

**Lemma 2.5.** Let \( A \) be the operator with domain \( D_A \) defined in (2.13). Then the corresponding Markov process \( Y \) is an ergodic process. Furthermore, if \( f \in L^2(G) \) and \( \int_G f(y) \, dy = 0 \), then \( f \in \text{Ran}(A) \) := \{ \{ A u : u \in D_A \} \}.

**Proof.** The first property follows immediately since, as noted above, \( \pi(y) = 1/|G| \) is the unique solution, up to non-normalized constants, of the differential equation \( A\pi = 0 \). The identification of the range follows from Bhattacharya [4], Proposition 2.3 (c) since, in view of the smooth Neumann boundary \( \partial G \), \( (\lambda - A)^{-1} \) is a compact operator for any \( \lambda > 0 \) (Davies [7], p. 48, Theorem 1.7.12), and 0 is a simple eigenvalue of \( A \). \( \Box \)

Next, let \( B = \{ B_t : t \geq 0 \} \) be a standard Brownian motion independent of \( Y \) and define the Itô process \( X \) by
\[
dX_t = U(Y_t) \, dt + \sqrt{2D_y(Y_t)} \, dB_t.
\] (2.16)
Then, to complete the construction of the stochastic process associated with the differential equation (2.3) via Lemma 2.3, we consider the process defined by
\[
X = \{ X_t = (X_t, Y_t) : t \geq 0 \}.
\] (2.17)

**Remark 2.6.** The decomposition \( X = (X, Y) \) implies that the lifetime of \( X \) is a.s. infinite; i.e. the process is non-explosive. Specifically, \( \mathbb{P}_x \) almost surely one has,
\[
|X_t - x| \leq \|U\|_{L^\infty(G)} t + \sqrt{2\|D_x\|_{L^\infty(G)}} |B_t|, \quad t \geq 0,
\]
and the process in the right hand side does not explode in finite time. Also explosion does not occur for \( Y \) by ergodicity.

The stochastic process defined by (2.17) defines a semigroup whose infinitesimal generator coincides with that defining the bilinear form \( \mathcal{E} \). In particular the asymptotic homogenization problem may be viewed as the asymptotic (marginal) distribution of the rescaled longitudinal process \( X^{(Y)} \). This is the topic of the next section.
2.4. Proof of main results. The homogenization result will follow from an application of the following central limit theorem and formula for the variance for ergodic Markov processes due to Bhattacharya [4].

**Theorem 2.7.** Let \( Y = \{ Y_t : t \geq 0 \} \) be a progressively measurable stationary Markov process on \( G \), having invariant measure \( \pi \) and infinitesimal generator \( A : D_A \to L^2(G, \pi) \). Let \( U_0 \in \text{Ran}(A) \) and consider the process \( Z_t = \int_0^t U_0(Y_s) \, ds, t \geq 0 \). Then, as \( n \to \infty \), the distribution of the scaled process \( \{ n^{-\frac{1}{2}} Z_{nt} : t \geq 0 \} \) converges weakly to the Wiener measure with zero drift and variance parameter
\[
-(U_0, h)_{L^2(G, \pi)} = -\int_G U_0(y) \, h(y) \, \pi(\, dy), \tag{2.18}
\]
where \( h \) is a solution to \( Ah = U_0 \).

The result in Theorem 2.7 now follows easily. By Lemma (2.3), the longitudinal component of \( X \) is given by
\[
X_t = x + \int_0^t U(Y_s) \, ds + \int_0^t \sqrt{2D_x(Y_s)} \, dB_s.
\]

Let \( \bar{U} \) be the mean of \( U \) with respect to \( \pi \) (see equation 2.10), and denote the centered drift of \( Y \) by
\[
U_0(y) = U(y) - \bar{U}. \tag{2.19}
\]

Let \( \bar{X}_t = (X_t - \bar{U}t - x) \), and for \( n \geq 0 \), consider the scaled process \( \bar{X}^{(n)} = \{ n^{-\frac{1}{2}} \bar{X}_{nt} : t \geq 0 \} \),
\[
\bar{X}_t^{(n)} = n^{-\frac{1}{2}} \int_0^{nt} U_0(Y_s) \, ds + n^{-\frac{1}{2}} \int_0^{nt} \sqrt{2D_x(Y_s)} \, dB_s \equiv Z_t^{(n)} + W_t^{(n)}. \tag{2.20}
\]

For the process \( W^{(n)} \), note that if \( W_0^{(n)} \) has distribution equal to the invariant probability \( \pi \), then for each \( n \) and \( t > 0 \), \( \mathbb{E} W_t^{(n)} = 0 \). The Itô isometry gives
\[
\mathbb{E}_\pi [W_t^{(n)}]^2 = n^{-1} \int_0^{nt} \mathbb{E}_\pi [2D_x(Y_s)] \, ds = t \int_G D_x(y) \pi(\, dy). \tag{2.21}
\]

Now, conditionally given \( Y \) up to time \( t \), \( W^{(n)} \) is a linear functional of the Brownian motion \( B \). The characteristic function of \( W^{(n)} \) can be then computed for each \( n \), and the limit is taken using the ergodicity of \( Y \),
\[
\mathbb{E}_\pi \exp \left( i \xi W_t^{(n)} \right) = \mathbb{E}_\pi \mathbb{E} \left[ \exp \left( i \xi W_t^{(n)} \right) \left| \{ Y_s : 0 \leq s \leq t \} \right. \right]
= \mathbb{E}_\pi \exp \left( -\frac{\xi^2}{nt} \int_0^{nt} D_x(Y_s) \, ds \right)
\to \exp \left( -\xi^2 \int_G D_x(y) \pi(\, dy) \right), \quad \text{as} \ n \to \infty.
\]

Therefore \( W^{(n)} \) is asymptotically a Brownian motion with diffusion coefficient
\[
\bar{D}_W = \int_G D_x(y) \pi(\, dy). \tag{2.22}
\]
To establish the convergence of process $Z^{(n)}$, note first that by Lemma (2.5), $U_0 \in \text{Ran}(A)$. Now apply Lemma (2.7) to get that for any initial distribution of $W_0^{(n)}$, as $n \to \infty$, $Z^{(n)}$ converges weakly to a Brownian motion with zero drift and diffusion coefficient given by

$$
\bar{D}_Z = -\int_G U_0(y) h(y) \pi(dy) = -\int_G A h(y) \pi(dy) = \int_G (D_y \nabla_y h) \cdot \nabla_y h(y) \pi(dy).
$$ (2.23)

From here we obtain the convergence of $\bar{X}^{(n)}$ to a Brownian motion with the asserted dispersion coefficient as follows.

**Proposition 2.8.** $\bar{X}^{(n)}$ converges weakly to a Brownian motion with drift $\bar{U}$ and diffusion coefficient

$$
\bar{D} = \bar{D}_W + \bar{D}_Z.
$$

**Proof.** Since we have shown that each of the two processes $Z^{(n)}$ and $W^{(n)}$ converge weakly to Brownian motions, it suffices to show that for each pair of times $0 \leq s \leq t$, $(Z_s^{(n)}, W_t^{(n)})$ converges weakly to a pair of independent Gaussian random variables. For this consider the bivariate moment generating function along the above lines. Specifically,

$$
\mathbb{E} \exp\left(\lambda_1 Z_s^{(n)} + \lambda_2 W_t^{(n)}\right) = \mathbb{E} \left\{ \exp\left(\lambda_1 Z_s^{(n)}\right) \mathbb{E}\left[\exp\left(\lambda_2 W_t^{(n)}\right) | \{Y_r : r \leq t\}\right]\right\}
$$

$$
= \mathbb{E} \left\{ \exp\left(\lambda_1 Z_s^{(n)}\right) \exp\left(\frac{\lambda_2^2}{\lambda_1^2} \int_0^t D_x(Y_r) d\tau\right)\right\}
$$

$$
= \mathbb{E} \exp\left(\lambda_1 \left(Z_s^{(n)} + \frac{\lambda_2^2}{\lambda_1^2} \int_0^t D_x(Y_r) d\tau\right)\right). \quad (2.24)
$$

Now, since by the Ergodic Theorem the random variable $\lambda_2^2 \int_0^t D_x(Y_r) d\tau$ converges a.s. to a constant (for fixed $t$), namely $\gamma := \frac{\lambda_2^2}{\lambda_1^2} \int_G D_x(y) \pi(dy)$, it follows that the sum $Z_s^{(n)} + \frac{\lambda_2^2}{\lambda_1^2} \int_0^t D_x(Y_r) d\tau$ converges in distribution to the sum of the limit distributions; see Billingsley [6], Chapter 1, Section 4. In particular, therefore, the moment generating function of the sum (as a function of $\lambda_1$) has the following asymptotic factorization

$$
\lim_{n \to \infty} \mathbb{E} \exp\left(\lambda_1 Z_s^{(n)} + \lambda_2 W_t^{(n)}\right) = \lim_{n \to \infty} \mathbb{E} \exp\left(\lambda_1 \left(\frac{\lambda_2^2}{\lambda_1^2} \int_0^t D_x(Y_r) d\tau\right)\right)
$$

$$
= \exp\left(\frac{\lambda_2^2}{2} \bar{D}_Z + \lambda_1 \gamma\right)
$$

$$
= \exp\left(\frac{\lambda_2^2}{2} \bar{D}_Z\right) \exp\left(\frac{\lambda_2^2}{2} \bar{D}_W\right). \quad (2.25)
$$

Since the two limit Brownian motions are independent at any two time points, the processes are independent. □

To relate the convergence of $X$ to Theorem 2.1, consider an initial longitudinal concentration $C_0(x)$ and define

$$
v_0(x) = C_0(x) \frac{1}{|G|}, \quad x \in G.
$$
This amounts to $Y_0$ having $v_0$ as its distribution. Let $C(x, t)$ be the cross-sectional concentration defined in (2.6) and $I$ an interval. Since

$$\int_I C(x, t) \, dx = \int_{G \times I} v(x, t) \, dx = \mathbb{P}(X_t \in G \times I) = \mathbb{P}(X_t \in I),$$

the homogenization result follows.

For the particular case treated in Corollary (2.2) consider the one dimensional diffusion with $G = [a, b]$ and

$$D(y) = \begin{bmatrix} D_x(y) & 0 \\ 0 & D_y(y) \end{bmatrix},$$

where $D_x$ and $D_y$ are positive and bounded away from zero and infinity. Let $g$ be as in (2.12), and define

$$h(y) = \int_a^y \frac{g(z)}{D_y(z)} \pi(\, dy \,).$$

Then $h$ solves the two dimensional version of problem (2.8). Using this solution in (2.10), gives

$$\bar{D} = \int_G \{ D_x(y) + \frac{g(y)^2}{D_y(y)} \} \pi(\, dy \,). \quad (2.26)$$

The result of Corollary (2.2) now follows by applying (2.26) to the case of $D$ given by (2.2).

3. One dimensional diffusion in heterogenous media..

3.1. Skew Brownian Motion and Skew Diffusion.. Consider first the case of a solute immersed in a medium without boundary in which $G = \mathbb{R}$. Assume zero drift, and a diffusion (dispersion) coefficient having a single interfacial point of discontinuity at $l_0 = 0$, namely, $D(y) = D^{(-1)}1_{(-\infty, 0)}(y) + D^{(0)}1_{[0, \infty)}(y)$, $y \in \mathbb{R}$. To simplify notation write

$$D^{(-1)} \equiv D^- \quad D^{(0)} \equiv D^+.$$ 

In this case the equations governing solute concentration $c(y, t)$ are given by

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial y} \left( D(y) \frac{\partial c}{\partial y} \right), \quad t > 0,$$

$$c(0^-, t) = c(0^+, t), \quad D^- \frac{\partial c}{\partial x}(0^-, t) = D^+ \frac{\partial c}{\partial x}(0^+, t), \quad t > 0,$$

$$c(y, 0^+) = c_0(y). \quad (3.1)$$

For this problem one may explicitly compute the fundamental solution $p(t, x, y)$ by solving the half-space problems on $[0, \infty)$ and $(-\infty, 0]$ with (apriori unknown) Neumann boundary fluxes at 0. One then may use the continuity of concentration and fluxes to match the values at the interface and determine $c(y, t)$. In particular, it follows that

$$p(x, y; t) = \begin{cases} \frac{1}{\sqrt{4\pi D^- t}} \exp \left\{ -\frac{(y-x)^2}{4D^- t} \right\} + \frac{\sqrt{D^- - \sqrt{D^+ D^-}}}{\sqrt{D^- + \sqrt{D^+ D^-}}} \exp \left\{ -\frac{(y-x)^2}{4D^- t} \right\}, & \text{if } x > 0, y > 0 \\ \frac{1}{\sqrt{4\pi D^- t}} \exp \left\{ -\frac{(y-x)^2}{4D^- t} \right\} - \frac{\sqrt{D^- + \sqrt{D^+ D^-}}}{\sqrt{D^- + \sqrt{D^+ D^-}}} \exp \left\{ -\frac{(y-x)^2}{4D^- t} \right\}, & \text{if } x < 0, y < 0 \end{cases}$$

$$p(x, y; t) = \begin{cases} \frac{1}{\sqrt{4\pi D^+ t}} \exp \left\{ \frac{(y-x)^2}{4D^+ t} \right\} - \frac{\sqrt{D^+ - \sqrt{D^- D^+}}}{\sqrt{D^+ + \sqrt{D^- D^+}}} \exp \left\{ \frac{(y-x)^2}{4D^+ t} \right\}, & \text{if } x \leq 0, y > 0 \\ \frac{1}{\sqrt{4\pi D^+ t}} \exp \left\{ \frac{(y-x)^2}{4D^+ t} \right\} + \frac{\sqrt{D^+ + \sqrt{D^- D^+}}}{\sqrt{D^+ + \sqrt{D^- D^+}}} \exp \left\{ \frac{(y-x)^2}{4D^+ t} \right\}, & \text{if } x \geq 0, y < 0. \end{cases} \quad (3.2)$$
One may identify the underlying stochastic process as a function of the skew-Brownian motion with parameter \( \alpha \in [0, 1] \) introduced by Itô and McKean [10], and here with the particular skew parameter

\[
\alpha^* = \frac{\sqrt{D^+}}{\sqrt{D^+} + \sqrt{D^-}}.
\]

Specifically,

**Proposition 3.1.** Let \( B^{(\alpha)} = \{B^{(\alpha)}_t : t \geq 0\} \) denote skew Brownian motion with parameter \( \alpha \in (0, 1) \), and let \( Y^{(\alpha^*)} = \{Y^{(\alpha^*)}_t : t \geq 0\} \) be the diffusion on \( \mathbb{R} \) defined by

\[
Y^{(\alpha^*)}_t := \sqrt{2D(B^{(\alpha^*)}_t)} B^{(\alpha^*)}_t,
\]

\[
\alpha^* = \frac{\sqrt{D^+}}{\sqrt{D^+} + \sqrt{D^-}}.
\]

Then \( Y^{(\alpha^*)} \) is the diffusion with transition probabilities given by (3.2). In particular, the solution to (3.1) is given by

\[
c(y, t) = \mathbb{E}_y c_0(Y^{(\alpha^*)}_t), \quad t \geq 0.
\]

**Proof.** Since the function \( \varphi(y) = \sqrt{2D(y)}y, y \in \mathbb{R} \), is continuous and one-to-one, it follows that for any \( \alpha \in (0, 1) \), \( \{\varphi(B^{(\alpha)}_t) : t \geq 0\} \) is a diffusion, i.e., strong Markov process with continuous sample paths, as this is directly inherited from \( B^{(\alpha)} \). One may directly compute the transition probabilities of \( \{\varphi(B^{(\alpha)}_t) : t \geq 0\} \) in terms of those for skew Brownian motion given by Walsh [23]. A comparison with (3.2) then yields the determination of \( \alpha = \alpha^* \) as asserted. \( \Box \)

**Remark 3.2.** We refer to the stochastic process \( Y^{(\alpha^*)} \) as the physical skew diffusion corresponding to the problem (3.1). There are now a number of alternative ways in which to describe the motion of solute particles. For example, Walsh [23] characterized skew Brownian motion by the property

\[
\text{Law}(|B^{(\alpha)}_t|) = \text{Law}(|B_t|), \quad \mathbb{P}_0(B^{(\alpha)}_t > 0) = \alpha,
\]

and used this to calculate the transition probability densities. Harrison and Shepp [9] obtained skew Brownian motion as a weak limit of rescaled birth-death processes on the integer lattice with transition probabilities \( p_{i,j} = \frac{1}{2}, |i - j| = 1, i \neq 0 \), and \( p_{0,1} = \alpha, p_{0,-1} = 1 - \alpha \). They also showed that skew Brownian motion is a solution to the stochastic differential equation

\[
dB^{(\alpha)}_t = (2\alpha - 1) dI^{(\alpha)}_0(t) + dB_t,
\]

where \( B \equiv B^{(\frac{1}{2})} \) is standard Brownian motion and \( \{I^{(\alpha)}_0(t) : t \geq 0\} \) is the local-time (at 0) for the (unknown) solution process \( B^{(\alpha)} \). The stochastic processes \( B^{(\alpha)} \) were later shown by Le Gall [12] to be a strong solutions to (3.4). In any case, it follows from this and the Itô-Tanaka lemma that \( Y^{(\alpha^*)} \) satisfies the stochastic differential equation

\[
dY^{(\alpha^*)}_t = K dI^{(\alpha^*)}_0 + \varphi(Y^{(\alpha^*)}_t) dB_t,
\]

\(^3\alpha = 0 \text{ or } 1 \text{ is permitted but corresponds to reflecting boundary at } 0. \text{ This is a somewhat degenerate case and will be excluded from the general discussion.} \)
for a constant $K$ depending on $\alpha, D^+, D^-$, and where $t_0^{(\alpha)}$ is the local time at 0 of the process $Y^{(\alpha)}$. Ouknine [16] obtained the class of processes $Y^{(\alpha)}$, for $\alpha \in (0, 1)$, in an extension of Le Gall [12] to a theory of strong solutions to the corresponding stochastic differential equation. The transition semigroup of $Y^{(\alpha)}$ was also shown by Ouknine [16] to have a functional-valued (generalized) infinitesimal generator, namely

$$
\lim_{t \to 0} \frac{1}{t} \mathbb{E}_y \left\{ u(Y_t^{(\alpha)}) - u(y) \right\} = \{ \partial(D \partial) + [D \partial]_0 \delta_0 \} u, \quad u \in H^1(\mathbb{R}).
$$

(3.6)

Here $\delta_0$ and $[\cdot]_0$ denote the Dirac delta distribution and the jump at point zero respectively, and the term $[D \partial]_0 \delta_0$ gives explicitly the propagation of the discontinuity of $D$ as a generalized drift coefficient. The idea of a infinitesimal generator with generalized coefficients has been extended in Portenko [17], Mastrangelo and Mouloud [14], Kopytko [11] and others, to analytically describe diffusion processes.

In their original construction Itô and McKean [10] defined the skew Brownian motion process $B^\alpha$ by independently assigning $\pm$ signs to the excursions around zero of a Brownian motion with respective probabilities $\alpha, 1-\alpha$. The probabilistic behavior of the skew Brownian motion process was then described in terms of the probabilities and rates of escape from arbitrary intervals via computation of its scale function and speed measure. From this one may easily obtain the corresponding scale function and speed measure for the physical skew diffusion since the transformation $y \mapsto \varphi(y) = \sqrt{2D(y)y}$ is invertible. This is the approach that we will follow in the next subsection to obtain a probabilistic description of the underlying solute particles in the generalized Taylor-Aris problem in a layered medium. In particular the picture that emerges is that particles behave locally as in a regular diffusion, but are perturbed at the interfaces. The strength of this perturbations depend on the size of the jump of the diffusion coefficient at the interface, and have the net effect of directing the particles to the regions of higher diffusivity.

**3.2. Diffusion in a layered medium.** We now turn our attention to the layered medium described in Corollary (2.2). The process $Y$ is a one-dimensional diffusion in the interval $G = [a, b]$, experiencing sharp discontinuities of the diffusion coefficient at discrete interfaces $l_k$ in the interior of $G$, and reflecting boundaries at the endpoints $^2$ By Lemma (2.3), $Y$ is the diffusion process associated with the Dirichlet form

$$
\mathcal{E}(u, v) = \int_G D(y) \partial u(y) \partial v(y) \, dy, \quad u, v \in D_\mathcal{E} = H^1(G),
$$

where $D$ is the piecewise constant function,

$$
D(y) = \sum_{k=-m}^{M} D^{(k)} 1_{[l_k, l_{k+1})}(y)
$$

for some positive local diffusion coefficients $D^{(k)}$ and interfaces $a = l_{-m} < \cdots < l_{M+1} = b$. We now seek to obtain the basic probabilistic structure of $Y$ in terms of its scale function and speed measure as computed below. Our approach is to first “guess” the scale function and speed measure of $Y$ by informal consideratons of the expected relationship to skew Brownian motion, and then give proofs (see Theorem 3.3 below)

$^2$Although in general interfaces are not boundaries, as noted earlier reflecting boundaries may be regarded as cases of degenerate interfaces.
using the said relationship and known properties of skew-Brownian motion. The basis for this guess is that since the speed measure and scale function measure the time to escape a sufficiently small interval and the probabilities to escape to the right or left of the interval respectively, it is enough to consider the process started in one of two types of intervals: an interval containing a single interface and an interval in which there is no interface. For these one may apply the results noted in the previous subsection on the speed measure and scale function of skew Brownian motion.

First, for $c \in \mathbb{G} = [a, b]$, let $\tau(c)$ be the hitting time of $c$ by the process $Y$, if $c < d$, denote $\tau(c, d) = \min\{\tau(c), \tau(d)\}$, the exit time of the interval $(c, d)$. The scale function (or scale measure) of $Y$ is a continuous strictly increasing function $s : (a, b) \to \mathbb{R}$ such that for each $y, c, d$ with $a < c < y < d < b$,

$$
\mathbb{P}_y(\tau(c, d) = \tau(d)) = \frac{s(y) - s(c)}{s(d) - s(c)}
$$

The scale function is unique up to an additive constant and is characterized by its associated positive measure $s(dx)$ (see, e.g. Revuz and Yor, [18], Sec. VII.3). If $(c, d) \subset (l_k, l_{k+1})$, then up to $\tau(c, d)$ the process behaves as if it was in natural scale, namely

$$
\mathbb{P}_y(\tau(c, d) = \tau(d)) = \frac{y - c}{d - c}.
$$

This gives

$$
s(dy) = s_k \, dy, \quad y \in (l_k, l_{k+1}),
$$

for some positive constants $s_k$, $k = -m, \ldots, M$. Now fix an interface $l_k$ and set $c = l_k - \delta \sqrt{2D^{(k-1)}}$, $d = l_k + \delta \sqrt{2D^{(k)}}$ for some $\delta > 0$ such that $(c, d) \subset (l_{k-1}, l_{k+1})$. In view of (3.3), we must have

$$
\alpha_k := \frac{\sqrt{D^{(k)}}}{\sqrt{D^{(k)}} + \sqrt{D^{(k-1)}}} = \mathbb{P}_{l_k}(\tau(c, d) = \tau(d)) = \frac{s_{k-1}(\delta \sqrt{2D^{(k-1)}})}{s_{k-1}(\delta \sqrt{2D^{(k-1)}}) + s_k(\delta \sqrt{2D^{(k)}})}.
$$

Solving gives $s_k D^{(k)} = s_{k-1} D^{(k-1)}$. Set $s_{-m} = 1$ to get

$$
s_k = \frac{D^{(k-1)}}{D^{(k)}}, \quad k = -m - 1, \ldots, M.
$$

The speed measure of $A$ is the unique Radon measure $m$ on $(a, b)$ such that for any $y, c, d$ with $y \in (c, d) \subset \mathbb{G}$,

$$
\mathbb{E}_y \tau(c, d) = \int_{G} G_{c,d}(y, y') \, m(dy'),
$$

where $G_{c,d}$ is the Green’s function of $Y$ for the interval $(c, d)$ (see Revuz and Yor, [18], Sec. VII.3).

$$
G_{c,d}(y, y') = \begin{cases} 
\frac{(s(y) - s(c))(s(d) - s(y'))}{s(d) - s(c)} & \text{if } c \leq y \leq y' \leq d, \\
\frac{(s(y') - s(c))(s(d) - s(y))}{s(d) - s(c)} & \text{if } c \leq y' \leq y \leq d, \\
0 & \text{otherwise.}
\end{cases}
$$
Assume a speed measure of the form $m(dy) = m_k dy$ for $y \in (l_k, l_{k+1})$ and consider an interval $(y - \delta, y + \delta) \subset (l_k, l_{k+1})$. On such an interval, the diffusion coefficient is constant and equal to $D^{(k)}$, therefore

$$\frac{\delta^2}{2D^{(k)}} = E_y \tau(y - \delta, y + \delta) = \int_{y - \delta}^{y + \delta} G_{y - \delta, y + \delta}(y, y') m(dy')$$

$$= s_k m_k \left\{ \int_y^{y + \delta} \frac{\delta(y + \delta - y')}{2\delta} dy' + \int_{y - \delta}^{y} \frac{(y' - \delta + y)\delta}{2\delta} dy' \right\}$$

$$= s_k m_k \frac{\delta^2}{2}$$

Solving for $m_k$, gives that $m(dy)$ is constant and equal to

$$m(dy) = \frac{1}{D^{(k)} s_k} dy, \quad y \in (l_k, l_{k+1}). \quad (3.11)$$

The behavior at the boundaries is determined by the value of the speed measure at $a$ and $b$. In particular, for instantaneously reflecting endpoints, $m$ must satisfy (see Revuz and Yor, [18], Def. 3.3.11)

$$m\{a\} = m\{b\} = 0. \quad (3.12)$$

**Theorem 3.3.** $Y$ is a Feller process on $G = [a, b]$ with scale function $s$ given by (3.8) and (3.9), and speed measure $m$ given by (3.11) and (3.12).

**Proof.** The Feller property follows from noting that the process $Y$ can be written as $Y = F(\tilde{Y})$ where $F$ is a continuous function that folds $\mathbb{R}$ onto $[a, b]$, and $\tilde{Y}$ is a diffusion process on $\mathbb{R}$ with diffusion coefficient function given by the appropriate periodic extension of $D$ to all $\mathbb{R}$; see Bhattacharya and Waymire [3]. The Feller property of $\tilde{Y}$ can be obtained by the classical bounds on solutions to parabolic partial differential equations developed in Nash [15]. To check that $s$ and $m$ are indeed the scale and speed measures of $Y$, we use the representation of the infinitesimal generator as a second derivative with respect to this measures (see, e.g., Revuz and Yor, [18], Theorem 3.3.12.) Let $\mathcal{D}_{A_0} = \{ u \in \mathcal{D}_A : D\partial_y u \text{ is absolutely continuous} \}$ and $A_0$ the restriction of $A$ to $\mathcal{D}_{A_0}$. One has to check first that for $u \in \mathcal{D}_{A_0}$ and $x$ in the interior of $G$, the $s$-derivative of $u$ exists. Note that

$$\frac{du}{ds}(y) = \frac{1}{s_k} \partial_y u(y), \quad \text{for } x \in (l_k, l_{k+1}),$$

which clearly exists at each point away from the interfaces and boundaries of $G$. At an interface $l_k$ this condition is equivalent to

$$\frac{1}{s_{k-1}} \partial_y u(l_k^-) = \frac{1}{s_k} \partial_y u(l_k^+),$$

which by (3.9), is satisfied since $D\partial_y u$ is continuous. Now let $y, y'$ be such that $l_{k-1} < y < l_k < y' < l_{k+1}$ for some $k$, and use (3.9) and the absolute continuity of $D\partial_y u$ to compute,

$$\frac{du}{ds}(y) - \frac{du}{ds}(y') = \frac{1}{s_{k-1}} \partial_y u(y) - \frac{1}{s_k} \partial_y u(y') = \frac{1}{s_k D^{(k)}} \left( D^{(k-1)} \partial_y u(y) - D^{(k)} \partial_y u(y') \right)$$

$$= \frac{1}{s_k D^{(k)}} \int_y^{y'} \partial_y (D \partial_y u) dy = \int_y^{y'} A_0 u(y) \ m(dy).$$
Lastly, the conditions \( \frac{du}{ds}(a) = m\{a\}A_0u(a) \) and \( \frac{du}{ds}(b) = m\{b\}A_0u(b) \) hold trivially by the boundary condition imposed to functions in \( D_A \). We then have that \( m \) and \( s \) satisfy \( A = \frac{d^2}{ds^2} + \frac{du}{ds} \) on \( D_A \), and therefore these are the speed and scale measures of \( Y \). \( \square \)

REFERENCES


[24] R. A. Wooding, *Rayleigh instability of a thermal boundary layer in flow through a porous...