Time-reversibility and nonvanishing Lévy area

Georg A Gottwald\(^1\) and Ian Melbourne\(^2\),\(^*\)

\(^1\) School of Mathematics and Statistics, University of Sydney, Sydney 2006, NSW, Australia
\(^2\) Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom

E-mail: i.melbourne@warwick.ac.uk

Received 21 October 2022; revised 4 April 2024
Accepted for publication 9 May 2024
Published 29 May 2024

Recommended by Dr Mark F Demers

Abstract

We give a complete description and clarification of the structure of the Lévy area correction to Itô/Stratonovich stochastic integrals arising as limits of time-reversible deterministic dynamical systems. In particular, we show that time-reversibility forces the Lévy area to vanish only in very specific situations that are easily classified. In the absence of such obstructions, we prove that there are no further restrictions on the Lévy area and that it is typically nonvanishing and far from negligible.

Keywords: stochastic integrals, Wong-Zakai approximation, homogenisation, deterministic fast-slow systems

Mathematics Subject Classification numbers: 60H05, 37A50, 37D25, 60L20

1. Introduction

The classical Wong-Zakai question [31] from 1965 is concerned with weak convergence of smooth processes \(W_n\) to Brownian motion \(W\) and the consequences for the interpretation of the corresponding stochastic integral \(\int WdW\).

\(^*\) Author to whom any correspondence should be addressed.

Original Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
In simple situations [17, 31], the limiting stochastic integrals are Stratonovich, denoted \( \int W \circ dW \), but numerous counterexamples exist in higher dimensions [27, 30]. In general, there is a correction, reminiscent of the Itô–Stratonovich correction, given by a deterministic quantity known as the Lévy area \( E \), whereby \( \int_0^t W_n \, dW_n \) converges weakly to \( \int_0^t W \circ dW + Et \) as \( n \to \infty \). Many references give closed-form formulas for the Lévy area \( E \), see [2, 6, 8, 16, 21, 23–26].

It is by now well-understood (if not always well-known) that the Lévy area is an important and nontrivial correction to the Wong–Zakai question. However, there are few investigations of whether this formula for \( E \) as displayed in section 2 (which looks like it may be nonzero) is actually nonzero in the presence of additional constraints such as time-reversibility.

The Lévy area is skew-symmetric (\( E^T = -E \)) and hence vanishes in the scalar case. In higher dimensions, Lévy area corrections vanish as a consequence of exactness, or commutativity of the defining vector fields, but such conditions are atypical outside the scalar case. On the other hand, an example where it is proved that \( E \neq 0 \) is given by Hairer, Pavliotis and Stuart, see [29, section 11.7.7].

From time to time, we have been asked whether time-reversibility can force \( E = 0 \). (The example of Hairer et al in [29] is not time-reversible.) In dispersing billiard examples considered by Chernov and Dolgopyat [4], the Lévy area is indeed zero as a consequence of time-reversibility and the structure of the equations. For Markov processes, there is a related condition, detailed balance, which forces \( E = 0 \), see for example [20, section 3.3.2] or [28, section 5.1]. See also [9, remark 3.4] and [25, section 1.4.2] for further comments on \( E \) being zero in certain time-reversible situations.

The recent work [10] gives numerical evidence that \( E \neq 0 \) in certain examples, but also questions whether this is typical or just occasional. On the other hand, it has recently been conjectured that in numerical simulations of certain stochastic systems which were obtained via some stochastic parametrization it is numerically advantageous to neglect the Lévy area and set \( E = 0 \) [3, 15].

In this paper, we offer a complete description and clarification of the structure of the Lévy area. In particular, we classify the constraints on \( E \) imposed by time-reversibility. The cases where \( E \) is forced to vanish are easily described. Outside of these rare situations, we find that \( E \) is typically far from negligible.

In terms of physical quantities such as position and momenta, there is the following simple description. Position is preserved by time-reversals and momenta are reflected, leading to three possibilities:

(i) all slow variables behave like position;
(ii) all slow variables behave like momenta;
(iii) some slow variables behave like position and some like momenta.

We show that \( E = 0 \) in cases (i) and (ii), but that in case (iii) — which seems natural in most physical applications — \( E \) is typically nonzero in the strongest possible sense.

**Remark 1.1.** A similar trichotomy arises in the theory of reversible Markov chains, see for example [1, 11]. Detailed balance corresponds to case (i) so we recover the fact that \( E = 0 \) when there is detailed balance. Case (ii) is called ‘modified’ detailed balance [13] or ‘skewed’ detailed balance [18]; again \( E = 0 \). Case (iii) is called ‘Yaglom reversibility’ [32] and it is here that our results say that \( E \) is typically nonzero.

The remainder of this paper is organised as follows. Section 2 contains the setup in this paper and gives a rough description of our main result. In section 3, we introduce time-reversibility.
and determine the constraints imposed by the time-reversibility on the Lévy area. In sections 4 and 5, we state and prove the main results described in section 2. Finally, in section 6, we verify that, in reasonable situations, nonvanishing Lévy area leads as anticipated to nontrivial corrections to limiting stochastic integrals.

2. The setup

We consider fast-slow ordinary differential equations (ODEs) on $\mathbb{R}^d \times \mathbb{R}^m$ of the form

\begin{align}
\dot{x} &= a(x) + \epsilon^{-1} b(x) v(y), \quad x(0) = \xi \in \mathbb{R}^d, \\
\dot{y} &= \epsilon^{-2} g(y), \quad y(0) \in \Lambda
\end{align}

(2.1)

Here, $\Lambda \subset \mathbb{R}^m$ is a compact invariant set for the $\dot{y}$ equation and $y(0)$ is chosen randomly from $\Lambda$ according to an ergodic Borel probability measure $\mu$ on $\Lambda$. The functions

- $a : \mathbb{R}^d \to \mathbb{R}^d$, $b : \mathbb{R}^d \to \mathbb{R}^{d \times d}$, $v : \mathbb{R}^m \to \mathbb{R}^d$, $g : \mathbb{R}^m \to \mathbb{R}^m$, are assumed to be $C^r$ (for some $r \geq 1$) with $\int_\Lambda v \, d\mu = 0$. Let $g_i : \Lambda \to \Lambda$ denote the flow on $\Lambda$ generated by the ODE $\dot{y} = g(y)$. We assume moreover that $\mu$ is mixing, so that $\lim_{t \to \infty} \int_\Lambda \phi \circ g_t \, d\mu = \int_\Lambda \phi \, d\mu$ for all $\phi, \psi \in L^2(\Lambda)$.

The aim of homogenisation is to establish, as $\epsilon \to 0$ in (2.1), a limiting stochastic differential equation (SDE) of the form

\begin{equation}
\dot{X} = \tilde{a}(X) \, dt + b(X) \circ dW, \quad X(0) = \xi,
\end{equation}

(2.2)

such that $X = x^{(\epsilon)}$ converges weakly to $X$. Here, $W$ is $d$-dimensional Brownian motion, the stochastic integral $b(X) \circ dW$ has the Stratonovich interpretation, and $\tilde{a}$ is a modified drift term incorporating the correction (if any) to the Stratonovich integral.

In the deterministic setting of (2.1), convergence to an SDE of the form (2.2) was obtained in [21] under suitable chaoticity assumptions (subsequently optimised in [6, 14, 22]) on the fast dynamics $\dot{y} = g(y)$. Under these assumptions, we have convergent series of Green–Kubo-type:

- **Covariance** $\Sigma = \int_0^\infty \int_\Lambda \left\{ v \otimes (v \circ g_t) + (v \circ g_t) \otimes v \right\} \, d\mu \, dt$,
- **Lévy area** $E = \int_0^\infty \int_\Lambda \left\{ v \otimes (v \circ g_t) - (v \circ g_t) \otimes v \right\} \, d\mu \, dt$,

where $u \otimes v = u^T v \in \mathbb{R}^{d \times d}$ for $u, v \in \mathbb{R}^d$. These series define a positive semi-definite symmetric matrix $\Sigma \in \mathbb{R}^{d \times d}$ and a skew-symmetric matrix $E \in \mathbb{R}^{d \times d}$. Moreover (under smoothness assumptions on $a$ and $b$ which play no further role in this paper), the solutions $x^{(\epsilon)}$ converge weakly to solutions of the SDE (2.2) where the Brownian motion $W$ has covariance matrix $\Sigma$ and the modified drift term is given by

\begin{equation}
\tilde{a}(X) = a(X) + \frac{1}{2} \sum_{\alpha, \beta, \gamma = 1}^d E^{\alpha \beta} \partial_\alpha b^\beta (X) b^\gamma (X).
\end{equation}

(2.3)

Here, $Z^{ij}$ denotes the $(i, j)$th entry of a matrix $Z$ and $b^\beta$ denotes the $\beta$th column of $b$.

Generally, $\Sigma$ is positive definite in the setting of [21]. Indeed, the case $\det \Sigma = 0$ is infinitely unlikely in a sense that can be made precise, see for example [5, section 2.3]. Given its antisymmetry, a natural question is to ask whether the Lévy area $E$ may be forced to vanish in certain circumstances. Clearly, if $d = 1$ then $E = 0$. In addition, if $v$ transforms as $v \circ R = v$ or
\( v \circ R = -v \), where \( R \) is a time-reversal symmetry for the fast dynamics, then again it is easily verified (see remark 3.2) that \( E = 0 \). This situation can occur in simplified physical situations where \( v \) represents only position or velocity, such as in the dispersing billiards examples in [4] and is also the case for the special classes of reversible Markov chains satisfying detailed balance or skew detailed balance.

In this paper, we show that the cases \( v \circ R = v \) and \( v \circ R = -v \) are the only situations where time-reversal symmetry forces the Lévy area to vanish, and typically \( E \neq 0 \) for the remaining time-reversible systems. (These three cases are precisely the situations (i)–(iii) mentioned in the Introduction.)

In particular, we show that \( v \) transforms as \( v \circ R = Av \) where \( A = L_d^+ \oplus (-L_d^-) \) in appropriate coordinates (with \( d^+ + d^- = d \)), and that \( E = \begin{pmatrix} 0 & E_0 \\ -E_0^T & 0 \end{pmatrix} \) where \( E_0 \) is a \( d^+ \times d^- \) matrix. Our main results include that there are no further constraints on \( E \) (theorem 4.3) and that \( E_0 \) is typically of full rank \( \min\{d^+, d^-\} \) in both the topological sense (openness and density) and in the probabilistic sense (prevalence [7, 19]), see remark 5.2. Moreover, the map from the dynamical system to \( E_0 \in \mathbb{R}^{d^+ \times d^-} \) is locally surjective (theorem 5.4) in addition to the aforementioned surjectivity.

3. Time-reversal symmetry

In this section, we introduce time-reversal symmetry into the fast-slow ODE (2.1) and derive simplified formulas for the covariance \( \Sigma \) and Lévy area \( E \). Our only assumption in this section is that the Green-Kubo-type formulas for \( \Sigma \) and \( E \) converge.

We assume that there is a time-reversal symmetry \((x, y) \mapsto (Sx, Ry)\) where \( S \in \mathbb{R}^{d \times d} \) and \( R \in \mathbb{R}^{m \times m} \) satisfy \( S^2 = I \) and \( R^2 = I \). As usual, this means that \((Sx(-t), Ry(-t))\) is a solution of (2.1) whenever \((x(t), y(t))\) is a solution. We suppose also that \( \mu \) is \( R \)-invariant.

For the fast dynamics, time-reversibility means that

\[
\begin{align*}
g_t(Ry) &= Rg_{-t}(y) \quad \text{for all } y \in \Lambda, t \in \mathbb{R}. 
\end{align*}
\]

Equivalently, \( g(Ry) = -Rg(y) \) for all \( y \in \Lambda \).

**Proposition 3.1.**

\[
\begin{align*}
\Sigma &= \int_0^\infty \int_\Lambda \left\{ v \otimes (v \circ g_t) + (v \circ R) \otimes (v \circ R \circ g_t) \right\} \, d\mu \, dt, \\
E &= \int_0^\infty \int_\Lambda \left\{ v \otimes (v \circ g_t) - (v \circ R) \otimes (v \circ R \circ g_t) \right\} \, d\mu \, dt.
\end{align*}
\]

**Proof.** By invariance of \( \mu \) under the fast flow \( g_t \) and \( R \), and (3.1),

\[
\begin{align*}
\int_\Lambda (v \circ g_t) \otimes v \, d\mu &= \int_\Lambda (v \circ R) \otimes (v \circ g_{-t} \circ R) \, d\mu \\
&= \int_\Lambda (v \circ R) \otimes (v \circ R \circ g_t) \, d\mu.
\end{align*}
\]

Substituting this into the covariance and Lévy area formulas from section 2 yields the result. \( \square \)

**Remark 3.2.** It can be seen already from proposition 3.1 that \( E = 0 \) if either \( v \circ R = v \) or \( v \circ R = -v \). (This was the case in [4].) Our main results imply in particular that typically \( E \neq 0 \) outside of these cases.
For the slow dynamics, time-reversibility means that
\[
a(Sx) + \epsilon^{-1} b(Sx)v(Ry) = -S \left\{ a(x) + \epsilon^{-1} b(x)v(y) \right\} \quad \text{for all } x \in \mathbb{R}^d, y \in \Lambda, \epsilon > 0.
\]

This simplifies to the requirement that
\[
a(Sx) = -Sa(x) \quad \text{and} \quad b(Sx)v(Ry) = -Sb(x)v(y) \quad \text{for all } x \in \mathbb{R}^d, y \in \Lambda. \tag{3.2}
\]

To avoid pathologies, from now on we suppose that \( b : \mathbb{R}^d \to \mathbb{R}^{d \times d} \) defines a nonsingular matrix in \( \mathbb{R}^{d \times d} \) on a dense subset of \( \mathbb{R}^d \) and that \( \{v(y) : y \in \Lambda\} \) spans \( \mathbb{R}^d \). Then the second condition in (3.2) simplifies further:

**Lemma 3.3.** Condition (3.2) holds if and only if there exists \( A \in \mathbb{R}^{d \times d} \) with \( A^2 = I \) such that
\[
a(Sx) = -Sa(x), \quad b(Sx)v(Ry) = -Sb(x)v(y), \quad v(Ry) = Av(y),
\]
for all \( x \in \mathbb{R}^d, y \in \Lambda \).

**Proof.** It is immediate that if \( a, b \) and \( v \) satisfy these restrictions for some \( A \) with \( A^2 = I \), then condition (3.2) holds.

Conversely, suppose that condition (3.2) holds. Let \( X = \{x \in \mathbb{R}^d : \det b(x) \neq 0\} \). Define
\[
A : X \to \mathbb{R}^{d \times d}, \quad A(x) = -b(x)^{-1}Sb(Sx).
\]

Then
\[
v(y) = A(x_1)v(Ry) \quad \text{and} \quad v(Ry) = A(x_2)v(y) \quad \text{for all } x_1, x_2 \in X, y \in \Lambda.
\]

Hence \( v(y) = A(x_1)A(x_2)v(y) \) for all \( y \) and it follows from the spanning assumption on \( v \) that
\[
A(x_1)A(x_2) = I \quad \text{for all } x_1, x_2 \in X.
\]

Taking \( x_1 = x_2 \), we obtain that \( A(x) \equiv A \) is constant on \( X \) with \( A^2 = I \). We immediately obtain that \( v \circ R = Av \). Also \( b \circ S = -SbA \) on the dense set \( X \subset \mathbb{R}^d \) and hence on the whole of \( \mathbb{R}^d \) by continuity of \( b \).

Let \( \pi^\pm : \mathbb{R}^d \to \mathbb{R}^d \) be the projections onto the \( \pm 1 \) eigenspaces of \( A \). Then we can write \( v \) in (2.1) uniquely as \( v = v^+ + v^- \) where \( v^\pm = \pi^\pm v \). Note that \( \int_\Lambda v^\pm d\mu = 0 \).

**Corollary 3.4.** In the \((\pi^+, \pi^-)\) coordinates,
\[
\Sigma = \begin{pmatrix} \Sigma^+ & 0 \\ 0 & \Sigma^- \end{pmatrix}, \quad E = \begin{pmatrix} 0 & E_0 \\ -E_0 & 0 \end{pmatrix},
\]
where
\[
\Sigma^\pm = 2 \int_0^\infty \int_\Lambda v^\pm \otimes (v^\pm \circ g_t) \, d\mu dt, \quad E_0 = 2 \int_0^\infty \int_\Lambda v^+ \otimes (v^- \circ g_t) \, d\mu dt.
\]

**Proof.** Since \( v \circ R = Av = v^+ - v^- \), we obtain
\[
\Sigma = 2 \int_0^\infty \int_\Lambda \left\{ v^+ \otimes (v^+ \circ g_t) + v^- \otimes (v^- \circ g_t) \right\} \, d\mu dt,
\]
\[
E = 2 \int_0^\infty \int_\Lambda \left\{ v^+ \otimes (v^- \circ g_t) + v^- \otimes (v^+ \circ g_t) \right\} \, d\mu dt.
\]
The result follows.
4. Generality of $E_0$

In this section, we show that there are no further restrictions on the Lévy area $E$ beyond those in corollary 3.4. Recall that the vector field $g$ in (2.1) is assumed to be $C^r$ for some $r \geq 1$. We suppose that the fast dynamics defined by $g$ is mixing sufficiently quickly, so that the series for $E$ in section 2 and hence $E_0$ in corollary 3.4 converge for all $C^r$ functions $v : \Lambda \to \mathbb{R}^d$ with $
abla^2 v = 0$. We exclude the uninteresting case where $\Lambda$ is a fixed point. (For then $v|_{\Lambda} \equiv 0$, so $\Sigma = E = 0$ and there is no stochasticity in the limit.)

**Remark 4.1.** In the setting of [6, 14, 21, 22], the matrices $\Sigma$ and $E$ depend continuously on $v \in C^r$.

Throughout this section, the functions $a$ and $b$ play no role and neither does the involution $S$. The fast vector field $g$ is fixed, as are the involutions $A$ and $R$. Our focus is purely on the dependence of $E_0$ (and hence $E$) on the function $v$. Recall from section 2 that $A = I_{d^+} \oplus (-I_{d^-})$ (with $d^+ + d^- = d$), and that $E = \begin{pmatrix} 0 & E_0 \\ -E_0^T & 0 \end{pmatrix}$ where $E_0$ is a $d^+ \times d^-$ matrix.

We say that a function $f$ defined on $\Lambda$ is $R$-invariant if $f \circ R = f$.

**Proposition 4.2.** For any $R$-invariant functions

$$f \in C^1 \left( \mathbb{R}^m, \mathbb{R}^{d^+} \right), \quad h \in C^{r+1} \left( \mathbb{R}^m, \mathbb{R}^{d^-} \right),$$

with $\int_\Lambda f \, d\mu = 0$, there exists

$$v \in C^1 \left( \mathbb{R}^m, \mathbb{R}^d \right) \text{ with } v \circ R = Av \quad \text{and} \quad \int_\Lambda v \, d\mu = 0,$$

such that $E_0 = \int_\Lambda f \otimes h \, d\mu$.

**Proof.** Recall that $g_t$ denotes the fast flow generated by the vector field $g$ in (2.1). We define $v = v^+ + v^- : \Lambda \to \mathbb{R}^d$ by setting

$$v^+ = -f, \quad v^- = \nabla h \cdot g = \sum_j \frac{\partial h}{\partial y_j} g_j.$$

Clearly, $v^+ \circ R = v^+ = Av^+$. By the chain rule, $(\nabla h)_R = (\nabla h) R$. This combined with the identity $g \circ R = -Rg$ ensures that $v^- \circ R = -v^- = Av^-$. Hence $v \circ R = Av$. Also, $\int_\Lambda v^- \, d\mu = \int_\Lambda v^\circ R \, d\mu = -\int_\Lambda v^- \, d\mu$. Hence $\int_\Lambda v^- \, d\mu = 0$. But $\int_\Lambda v^+ \, d\mu = 0$ by construction, so $\int_\Lambda v^- \, d\mu = 0$.

Along solutions $y(t)$ to $\dot{y} = g(y),$

$$v^-(y(t)) = \sum_j \frac{\partial h}{\partial y_j} (y(t)) \, \dot{y}_j (t) = \frac{d}{dt} h(y(t)).$$

Hence $\int_0^T v^-(y(t)) \, dt = h(y(T)) - h(y(0))$. In other words,

$$\int_0^T v^- \, g_t \, dt = h \circ g_T - h.$$
Since the flow is mixing and \( \int_{\Lambda} v^+ \, d\mu = 0 \),
\[
\int_0^T \int_{\Lambda} v^+ \otimes (v^- \circ g_t) \, d\mu \, dt = \int_{\Lambda} v^+ \otimes (h \circ g_T - h) \, d\mu \\
\to - \int_{\Lambda} v^+ \otimes h \, d\mu = \int_{\Lambda} f \otimes h \, d\mu
\]
as \( T \to \infty \). By the definition of \( E_0 \) in corollary 3.4 (up to a factor of 2 which can be incorporated into \( f \)) we have proved the result. \( \square \)

**Theorem 4.3.** For any \( d^+ \times d^- \) matrix \( F \), there exists a \( C^1 \) function \( v : \mathbb{R}^m \to \mathbb{R}^d \) with \( v \circ R = Av \) and \( \int_{\Lambda} v \, d\mu = 0 \) such that \( E_0 = F \).

**Proof.** We claim that it is possible to choose \( v \) so that \( E_0 \) has full rank, namely \( \min\{d^+, d^-\} \).

Assuming this is the case, let \( L^\Lambda \in \mathbb{R}^{d^+ \times d^-} \). By the definition of \( E_0 \) in corollary 3.4, transforming \( v \) to \( (L^+ \cdot v^+ \cdot L^- \cdot v^-) \) changes \( E_0 \) to \( L^+ E_0 (L^-)^T \). By standard linear algebra, this results in any desired matrix in \( \mathbb{R}^{d^+ \times d^-} \).

It remains to prove the claim. The first step is to construct suitable \( R \)-invariant functions \( f_i \in L^2(\Lambda) \) and \( h_i \in L^2(\Lambda) \) with \( \int_{\Lambda} f_i \, d\mu = 0 \) such that \( \int_{\Lambda} f_i \otimes h_i \, d\mu \) has full rank. The second step is to approximate \( f_i \) and \( h_i \) by smooth \( R \)-invariant functions \( f : \mathbb{R}^m \to \mathbb{R}^{d^+} \) and \( h : \mathbb{R}^m \to \mathbb{R}^{d^-} \) with \( \int_{\Lambda} f \, d\mu = 0 \) so that \( \int_{\Lambda} f \otimes h \, d\mu \) has full rank. The third step is to apply proposition 4.2.

**Step 1.** Since \( \Lambda \) is not a fixed point, we can choose infinitely many orthonormal \( R \)-invariant functions \( \tilde{\phi}_j \in L^2(\Lambda) \) with \( \int_{\Lambda} \tilde{\phi}_j \, d\mu = 0 \) and \( \int_{\Lambda} \tilde{\phi}_i \tilde{\phi}_j \, d\mu = \delta_{ij} \). Let \( f_i = \sum_{j=1}^{d^-} \alpha_{ij} \tilde{\phi}_j \) and \( h_i = \sum_{j=1}^{d^-} \beta_{ij} \tilde{\phi}_j \) where \( \alpha_{ij} \in \mathbb{R}^{d^+} \) and \( \beta_{ij} \in \mathbb{R}^{d^-} \). Then
\[
\int_{\Lambda} f_i \otimes h_i \, d\mu = \sum_{i,j} (\alpha_{ij} \otimes \beta_{ij}) \int_{\Lambda} \tilde{\phi}_i \tilde{\phi}_j \, d\mu = \sum_{i,j} \alpha_{ij} \otimes \beta_{ij}.
\]
Let \( F_i = \int_{\Lambda} f_i \otimes h_i \, d\mu \). Taking \( \beta_{ij} \) to be the \( j \)th canonical unit vector, \( F_i \) is the \( d^+ \times d^- \) matrix with columns \( \alpha_{1i} \ldots \alpha_{di} \). In particular, \( F_i \) is arbitrary and we can choose the \( \alpha_{ij} \) so that \( F_i \) is a matrix of full rank.

**Step 2.** Now choose \( R \)-invariant functions \( \hat{\phi}_j \in C^\infty(\mathbb{R}^m) \) with \( \int_{\Lambda} \hat{\phi}_j \, d\mu = 0 \) and \( \int_{\Lambda} |\hat{\phi}_j - \tilde{\phi}_j|^2 \, d\mu \) small\(^3\) and define \( f, h \) using \( \hat{\phi}_j \) in place of \( \tilde{\phi}_j \) with \( \alpha_{ij}, \beta_{ij} \) unchanged. This results in a matrix \( F = \int_{\Lambda} f \otimes h \, d\mu \) close to \( F_i \). In particular, taking the approximation close enough ensures that \( F \) is still of full rank.

**Step 3.** It follows from proposition 4.2 that we can choose \( v \in C^1(\mathbb{R}^m, \mathbb{R}^d) \) with \( v \circ R = Av \) and \( \int_{\Lambda} v \, d\mu = 0 \) so that \( E_0 = \int_{\Lambda} f \otimes h \, d\mu \). In particular, such \( E_0 \) has full rank, proving the claim. \( \square \)

---

\(^3\) The existence of such functions is standard. For instance, to approximate \( \hat{\phi}_j \) by a \( C^\infty \) function, first approximate \( \phi_i \) in \( L^2 \) by a simple function \( \sum_{k=1}^N c_k \chi_{E_k} \) (with \( c_k \in \mathbb{R} \) and \( E_k \subset \Lambda \) measurable). By outer regularity of the Borel probability measure \( \mu \), there exist open neighbourhoods \( U_k \subset \mathbb{R}^m \) of \( E_k \) with \( \mu(U_k \setminus E_k) \) small. Hence \( \sum_{k=1}^N c_k \chi_{U_k} \) is \( L^2 \)-close to \( \phi_i \). Choose \( V_k \subset \mathbb{R}^m \) closed such that \( \overline{U_k \setminus E_k} \subset \text{Int} V_k \) with \( \mu(V_k \setminus U_k) \) small, by Urysohn’s Lemma, there exists a continuous function \( \psi_k : \mathbb{R}^m \to [0,1] \) supported on \( V_k \) with \( \psi_k \chi_{V_k} = 1 \). In this way we obtain a continuous function \( \sum_{k=1}^N c_k \psi_k \) that is \( L^2 \)-close to \( \phi_i \). Each \( \psi_k \) can be uniformly approximated by a \( C^\infty \) function \( \xi_k \) resulting in a \( C^\infty \) function \( \phi_i = \sum_{k=1}^N c_k \xi_k \) that is \( L^2 \)-close to \( \phi_i \). Finally, replace \( \phi_i \) by \( \phi_i - \int_{\Lambda} \phi_i \, d\mu \).
5. Local surjectivity for perturbations of $E_0$

Let $\chi : \mathcal{C}(\mathbb{R}^m, \mathbb{R}^d) \to \mathbb{R}^{d^+ \times d^-}$ be the mapping defining $E_0 = \chi(v)$ in corollary 3.4.

In section 4, we showed that $E_0$ was a general matrix in the sense that $\chi$ is surjective. As mentioned in remark 4.1, $\chi$ is continuous in reasonable situations. In this section, we complete the picture by establishing a local surjectivity result for $\chi$; namely that if $\chi(v_0) = E_0$ and $F_0$ is close to $E_0$ of full rank, then there exists $v_1$ close to $v_0$ with $\chi(v_1) = F_0$.

The first step is to show that if $E_0 = \chi(v_0)$ does not have full rank, then the rank can be increased under small perturbations.

**Lemma 5.1.** For any $v_0 \in \mathcal{C}(\mathbb{R}^m, \mathbb{R}^d)$, there exists $v$ arbitrarily $\mathcal{C}$-close to $v_0$ such that \[ \text{rank } \chi(v) = \text{min}\{d^+, d^-\}. \]

**Proof.** Suppose without loss of generality that $d^+ \leq d^-$. By theorem 4.3, we can choose $v_+ \in \mathcal{C}(\mathbb{R}^m, \mathbb{R}^d)$ such that $\chi(v_+) = \left( \begin{array}{c} I_{d^+} \\ 0 \end{array} \right)$. Let $v_1 = v_0 + tv_+$ and define $E_1 = \chi(v_1)$. Then $E_1 = \left( \begin{array}{c|c} A_t & B_t \end{array} \right)$, where $A_t = A_0 + tA_1 + t^2I_{d^+}$ for some matrices $A_0, A_1 \in \mathbb{R}^{d^+ \times d^+}, B_t \in \mathbb{R}^{d^+ \times d^-}$.

First, we require a basic result from linear algebra.

**Proposition 5.3.** Let $E_0, F_0 \in \mathbb{R}^{m \times d}$ be matrices of full rank and suppose that $F_0$ is close to $E_0$. Then there exist near identity matrices $P \in \mathbb{R}^{m \times m}, Q \in \mathbb{R}^{d \times d}$ such that $PE_0Q^T = F_0$.

**Proof.** Suppose without loss of generality that $m \leq n$. Since rank $E_0 = m$, there exist invertible matrices $P_0 \in \mathbb{R}^{m \times m}, Q_0 \in \mathbb{R}^{d \times d}$ such that $P_0E_0Q_0^T = \left( \begin{array}{c} I_m \\ 0 \end{array} \right)$. Then $P_0F_0Q_0^T$ is close to $\left( \begin{array}{c} I_m \\ 0 \end{array} \right)$ and it is easily seen that there exist near identity matrices $P_1 \in \mathbb{R}^{m \times m}, Q_1 \in \mathbb{R}^{d \times d}$ corresponding to near identity row and column operations such that $(P_1P_0)F_0(Q_1Q_0)^T = \left( \begin{array}{c} I_m \\ 0 \end{array} \right)$. Moreover, $(P_0^{-1}P_1P_0)F_0(Q_0^{-1}Q_1Q_0)^T = E_0$. Hence the result holds with $P = P_0^{-1}P_1P_0, Q = Q_0^{-1}Q_1Q_0$.

We can now state and prove the main result of this section.

**Theorem 5.4.** Suppose that $E_0 = \chi(v_0)$ and that $F_0$ is of full rank and close to $E_0$. Then there exists $v$ that is $\mathcal{C}$-close to $v_0$ such that $\chi(v) = F_0$.

**Proof.** Suppose without loss of generality that $d^+ \leq d^-$. By lemma 5.1, we can make an initial perturbation so that rank $E_0 = d^+$. By proposition 5.3, there exist near identity matrices $P \in \mathbb{R}^{d^+ \times d^+}, Q \in \mathbb{R}^{d^- \times d^-}$ such that $PE_0Q^T = F_0$. Hence we can take $v = (P \oplus Q)v_0$. 

8
6. Corrections to limiting stochastic integrals

In the previous sections, we gave a complete description of the Lévy area, namely the skew symmetric matrix $E$ whose entries determine the correction

$$\frac{1}{2} \sum_{\alpha, \beta, \gamma = 1}^{d} E_{\alpha \beta} \partial_{\alpha} b^{\beta} (X) b^{\alpha \gamma} (X)$$ \hspace{1cm} (6.1)

to the drift term $a(X)$ as given in (2.3). This is not the complete story since even when $E$ is nonzero, it might be the case that the correction is forced to vanish due to the structure of $b$. Indeed, this happens when $b : \mathbb{R}^{d} \to \mathbb{R}^{d \times d}$ satisfies an exactness condition, namely that $b^{-1} = dh$ for some $h : \mathbb{R}^{d} \to \mathbb{R}^{d}$, see for example [17].

In general, $b$ satisfies the time-reversibility constraint

$$b (Sx) = - S b (x) A,$$ \hspace{1cm} (6.2)

from lemma 3.3 which places restrictions on the correction (6.1).

In this section, we consider some simple time-reversible situations with $E$ nonvanishing and show that in these cases the correction (6.1) is typically nonzero. Given the bilinearity of this term, it suffices to exhibit a single $b$ for which the term is nonzero$^4$.

Recall that $A^t = S^t = I$ and that we can choose coordinates such that $A = I_{d^+} \oplus I_{d^-}$. Recall also that $E_{\beta \gamma} = 0$ if $1 \leq \beta, \gamma \leq d^+$ and if $d^+ + 1 \leq \beta, \gamma \leq d$, while the remaining entries of $E$ are general. Since $E$ is assumed to be nonvanishing, we have $d^+ \geq 1$.

Our simplifying assumption, which is natural in term-reversible situations, is that $S$ is also diagonal in these coordinates, so $S = \text{diag} \{ S_1, S_2, \ldots, S_d \}$ where $S_\beta \in \{ \pm 1 \}$, and that $S \neq \pm I$. Define $B \subset \{ 1, \ldots, d \}$ so that $S_\beta = 1$ if and only if $\beta \in B$. Then the constraint (6.2) reduces to the constraints

$$b^\beta \gamma (Sx) = b^\beta \gamma (x) \quad \text{for} \ \beta \in B, \gamma > d^+ \text{ and for } \beta \notin B, \gamma \leq d^+,$$

and

$$b^\beta \gamma (Sx) = - b^\beta \gamma (x) \quad \text{for} \ \beta \in B, \gamma \leq d^+ \text{ and for } \beta \notin B, \gamma > d^+.$$

Fix $i \in B, j \notin B$. Then an allowable choice of $b$ is obtained by setting

$$b^{id} (x) = b^{i1} (x) = x^i$$

and setting the remaining entries to zero. Substituting into the sum in (6.1), we see immediately that nonzero terms require $\alpha = i$. Then the factor $b^{\alpha \gamma} (x)$ is nonzero only for $\gamma = d$. The column vector $b^\beta (x)$ has nonzero entries only for $\beta = 1, d$. Since $E_{id} = 0$, the correction (6.1) reduces to a single term

$$\frac{1}{2} E_{id} \partial_{i} b^1 (x) b^{id} (x) = \frac{1}{2} E_{id} x^i e_j$$

where $e_j$ is the canonical unit vector with 1 in the $j$’th entry. We know that typically $E_{ii} \neq 0$, so this yields a nontrivial correction to the drift term $a$ as required.

$^4$ A similar argument was used in the more difficult proof of lemma 5.1. Given one $b_\epsilon$ that succeeds, we can use it to perturb any $b_0$ that fails: the perturbation $b_0 + \epsilon b_\epsilon$ succeeds for all small $\epsilon > 0$. 

Nonlinearity 37 (2024) 075018
G Gottwald and I Melbourne

9
6.1. Explicit example

A classical example of a time-reversible mixing flow is the geodesic flow for a negatively curved compact orientable surface $M$. The unit tangent bundle $\Lambda = T^1 M$ is a compact three-dimensional manifold and the geodesic flow $g_t : \Lambda \to \Lambda$ is given by $g_t(x) = \gamma_t(x)$ where $\gamma_t : \mathbb{R} \to M$ is the unique geodesic such that $\gamma_t(0) = x$. An ergodic probability measure is given by the normalised Liouville measure $\mu$. Writing $y \in \Lambda$ as $y = (q,p)$ where $q \in M$ and $p \in T^1_q M$, the geodesic flow is time-reversible with $R(q,p) = (q,-p)$. By [12], the geodesic flow is exponentially mixing and hence the formula for $E_0$ in corollary 3.4 converges and depends continuously on Hölder observables $v$.

To incorporate the slow variables, we take $d = 2$ and $A = S = R = \text{diag}\{1,-1\}$. In particular, $d^+ = d^- = 1$ and $E_0$ is scalar.

Let $g : T^1 M \to T(T^1 M)$ be the vector field for the geodesic flow and let $f,h : T^1 M \to \mathbb{R}$ be any $R$-invariant functions with $\int_\Lambda f \, d\mu = 0$. Following the proof of proposition 4.2, we define $v = (v^+,v^-) : \Lambda \to \mathbb{R}^2$ where $v^+ = -f$ and $v^- = \nabla h \cdot g$. Then $E_0 = \int_\Lambda fh \, d\mu$.

Since $M$ is orientable, we can view it as a unit vector in $\mathbb{R}^2$. Let $\kappa$ denote curvature. Explicit choices of $f$ and $h$ are $f = \psi - \int \psi$ and $h = \psi$ where $\psi(q,p) = \kappa(q)p_2$, resulting in

$$E_0 = \int_\Lambda \left( \kappa(q)p_2^2 - \int_\Lambda \kappa(q)p_2^2 \, d\mu \right)^2 \, d\mu > 0.$$

The constraints $a(Sx) = -Sa(x)$ and $b(Sx) = -Sb(x)A$ reduce to $a^1$, $b^{11}$, $b^{22}$ being odd in $x_2$ and $a^2$, $b^{12}$, $b^{21}$ being even in $x_2$. We can take

$$a(x_1,x_2) = (x_2 \cos x_1, \cos x_2), \quad b(x_1,x_2) = \begin{pmatrix} 0 & x_1 \\ x_1 & 0 \end{pmatrix}.$$  

By the calculation at the beginning of this section, we obtain the nontrivial drift correction

$$\tilde{a}(x) = a(x) + \left(0, \frac{1}{2}E_0x_1\right) = (x_2 \cos x_1, \cos x_2 + \frac{1}{2}E_0x_1)$$

with $E_0 > 0$ as above.

**Remark 6.1.** The slow dynamics in this example is not very physical, but the reader is invited to take their favourite choice of $a$, $b$ and $v$ subject to the constraints arising from $R$, $A$ and $S$. If $\tilde{a}(x) = a(x)$, then perturb your choice slightly.

**Data availability statement**

No new data were created or analysed in this study.

**Acknowledgments**

This research was supported in part by the visitor program at the Sydney Mathematical Research Institute (SMRI) in early 2020. I M is very grateful for the hospitality of the University of Sydney, and in particular SMRI, at both ends of the pandemic.

We would like to thank Greg Pavliotis for his comments on detailed balance and the referees for many helpful comments and suggestions. We are particularly grateful to Sam Livingstone for explaining to us how reversibility in Markov chains correlates with reversibility in physical systems.
ORCID ID

Georg A Gottwald  https://orcid.org/0000-0002-5046-8520

References

[1] Andrieu C and Livingstone S 2021 Peskun-Tierney ordering for Markovian Monte Carlo: beyond the reversible scenario Ann. Stat. 49 1958–81
[2] Bass R F, Hambly B M and Lyons T J 2002 Extending the Wong-Zakai theorem to reversible Markov processes J. Eur. Math. Soc. 4 237–69
[3] Boulard P-M and Mémin E 2024 Diagnostic of the Lévy area for geophysical flow models in view of defining high order stochastic discrete-time schemes Found. Data Sci. 6 1–21
[4] Chernov N and Dolgopyat D 2009 Brownian Brownian Motion. I vol 198 (Memoirs of the American Mathematical Society) p viii + 193
[5] Chevyrev I et al 2019 Multiscale systems, homogenization and rough paths Probability and Analysis in Interacting Physical Systems: In Honor of S.R.S. Varadhan (Springer Proc. in Mathematics & Statistics) (Berlin, August, 2016) vol 283, ed P Friz (Springer) pp 17–48
[6] Chevyrev I, Friz P K, Korepanov A, Melbourne I and Zhang H 2022 Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 2 Ann. Inst. Henri Poincaré Probab. Stat. 58 1328–50
[7] Christensen J P R 1973 Measure theoretic zero sets in infinite dimensional spaces and applications to differentiability of Lipschitz mappings Publ. Dép. Math. 10 29–39
[8] Delong S, Sun Y, Griffith B E, Vanden-Eijnden E and Donev A 2014 Multiscale temporal integrators for fluctuating hydrodynamics Phys. Rev. E 90 063312
[9] Deuschel J-D, Orenshtein T and Perkowski N 2021 Additive functionals as rough paths Ann. Probab. 49 1450–79
[10] Diamantakis T, Holm D D and Pavliotis G A 2023 Variational principles on geometric rough paths and the Lévy area correction SIAM J. Appl. Dyn. Syst. 22 1182–218
[11] Dobrushin R L, Yu M S and Fritts I 1988 AN Kolmogorov—founder of the theory of reversible Markov processes Usp. Mat. Nauk 43 167–88
[12] Dolgopyat D 1998 On the decay of correlations in Anosov flows Ann. Math. 147 357–90
[13] Fang Y, Sanz-Serna J-M and Skeel R D 2014 Compressible generalized hybrid Monte Carlo J. Chem. Phys. 140 174108
[14] Fleming-Vázquez N 2022 Functional correlation bounds and optimal iterated moment bounds for slowly-mixing nonuniformly hyperbolic maps Commun. Math. Phys. 391 173–98
[15] Fiorini C, Boulard P-M, Li L and Mémin E 2023 A two-step numerical scheme in time for surface quasi geostrophic equations under location uncertainty Stochastic Transport in Upper Ocean Dynamics ed B Chapron, D Crisan, D Holm, E Mémin and A Radomska (Springer) pp 57–67
[16] Friz P and Oberhauser H 2009 Rough path limits of the Wong-Zakai type with a modified drift term J. Funct. Anal. 256 3236–56
[17] Gottwald G A and Melbourne I 2013 Homogenization for deterministic maps and multiplicative noise Proc. R. Soc. A 469 20130201
[18] Hukushima K and Sakai Y 2013 An irreversible Markov-chain Monte Carlo method with skew detailed balance conditions J. Phys.: Conf. Ser. 473 012012
[19] Hunt B R, Sauer T and Yorke J A 1992 Prevalence: a translation-invariant “almost every” on infinite-dimensional spaces Bull. Am. Math. Soc. 27 217–38
[20] Jiang D-Q, Qian M and Qian M-P 2004 Mathematical Theory of Nonequilibrium Steady States (Lecture Notes in Mathematics vol 1833) (Springer)
[21] Kelly D and Melbourne I 2016 Smooth approximation of Stochastic differential equations Ann. Probab. 44 479–520
[22] Korepanov A, Koslolf Z and Melbourne I 2022 Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 1 Ann. Inst. Henri Poincaré Probab. Stat. 58 1305–27
[23] Lejay A and Lyons T 2005 On the importance of the Lévy area for studying the limits of functions of converging Stochastic processes. Application to homogenization Current Trends in Potential Theory (Theta Ser. Adv. Math. vol 4) (Theta) pp 63–84
[24] Lim S H 2021 Anomalous thermodynamics in homogenized generalized Langevin systems *J. Phys. A: Math. Theor.* **54** 155001

[25] Lopusanschi O and Simon D 2018 Lévy area with a drift as a renormalization limit of Markov chains on periodic graphs *Stoch. Process. Appl.* **128** 2404–26

[26] MacKay R S 2010 Langevin equation for slow degrees of freedom of Hamiltonian systems *Nonlinear Dynamics and Chaos: Advances and Perspectives (Underst. Complex Syst.)* (Springer) pp 89–102

[27] McShane E J 1972 Stochastic differential equations and models of random processes Proc. 6th Berkeley Symp. on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971) (Probability Theory vol III) (University of California Press) pp 263–94

[28] Pavliotis G A 2014 Diffusion processes, the Fokker-Planck and Langevin equations *Stochastic Processes and Applications (Texts in Applied Mathematics)* (Springer)

[29] Pavliotis G A and Stuart A M 2008 *Multiscale Methods: Averaging and Homogenization (Texts in Applied Mathematics)* (Springer)

[30] Sussmann H J 1991 Limits of the Wong-Zakai type with a modified drift term *Stochastic Analysis* (Academic) pp 475–93

[31] Wong E and Zakai M 1965 On the convergence of ordinary integrals to stochastic integrals *Ann. Math. Statist.* **36** 1560–4

[32] Yaglom A M 1949 On the statistical reversibility of Brownian motion *Mat. Sbornik N.S.* **24** 457–92