Time-reversibility and nonvanishing Lévy area

Georg Gottwald * Ian Melbourne †

20 October, 2022. Updated 25 March, 2024

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.

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 $\int W\,dW$.

In simple situations [31, 17], 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, 24, 25, 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

*School of Mathematics and Statistics, University of Sydney, Sydney 2006 NSW, Australia
†Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK. i.melbourne@warwick.ac.uk

outside the scalar case. On the other hand, an example where it is proved that $E \neq 0$ is given by Hairer, Pavliotis & 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 & 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 parametrisation it is numerically advantageous to neglect the Lévy area and set $E = 0$ [15, 3].

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

$$\begin{aligned} \dot{x} &= a(x) + \epsilon^{-1} b(x) v(y), \quad && x(0) = \xi \in \mathbb{R}^d \\ \dot{y} &= \epsilon^{-2} g(y), && y(0) \in \Lambda \end{aligned} \tag{2.1}$$

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

$$a : \mathbb{R}^d \to \mathbb{R}^d, \quad b : \mathbb{R}^d \to \mathbb{R}^{d \times d}, \quad v : \mathbb{R}^m \to \mathbb{R}^d, \quad g : \mathbb{R}^m \to \mathbb{R}^m,$$

are assumed to be C^r (for some $r \geq 1$) with $\int_\Lambda v \, d\mu = 0$. Let $g_t : \Lambda \to \Lambda$ denote the flow on Λ generated by the ODE $\dot{y} = g(y)$. We assume moreover that μ is mixing, so that $\lim_{t\to\infty} \int_\Lambda \phi \, \psi \circ g_t \, d\mu = \int_\Lambda \phi \, d\mu \int_\Lambda \psi \, d\mu$ 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

$$dX = \tilde{a}(X) \, dt + b(X) \circ dW, \quad X(0) = \xi, \tag{2.2}$$

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 $\Sigma = \int_0^\infty \int_\Lambda \big\{ v \otimes (v \circ g_t) + (v \circ g_t) \otimes v \big\} \, d\mu \, dt$,

Lévy area $E = \int_0^\infty \int_\Lambda \big\{ v \otimes (v \circ g_t) - (v \circ g_t) \otimes v \big\} \, d\mu \, dt$,

where $u \otimes v = uv^T \in \mathbb{R}^{d \times d}$ 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 Σ and the modified drift term is given by

$$\tilde{a}(X) = a(X) + \tfrac{1}{2} \sum_{\alpha,\beta,\gamma=1}^{d} E^{\gamma\beta} \partial_\alpha b^\beta(X) b^{\alpha\gamma}(X). \tag{2.3}$$

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

Generally, Σ 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), (ii), (iii) mentioned in the Introduction.)

In particular, we show that v transforms as $v\circ R = Av$ where $A = I_{d^+} \oplus (-I_{d^-})$ 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 Σ and Lévy area E. Our only assumption in this section is that the Green-Kubo-type formulas for Σ 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 μ is R-invariant.

For the fast dynamics, time-reversibility means that

$$g_t(Ry) = Rg_{-t}(y) \quad \text{for all } y \in \Lambda,\ t \in \mathbb{R}. \tag{3.1}$$

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

Proposition 3.1

$$\Sigma = \int_0^\infty \int_\Lambda \big\{v \otimes (v \circ g_t) + (v \circ R) \otimes (v \circ R \circ g_t)\big\}\, d\mu\, dt,$$
$$E = \int_0^\infty \int_\Lambda \big\{v \otimes (v \circ g_t) - (v \circ R) \otimes (v \circ R \circ g_t)\big\}\, d\mu\, dt.$$

Proof By invariance of μ under the fast flow g_t and R, and (3.1),

$$\int_\Lambda (v\circ g_t)\otimes v\,d\mu = \int_\Lambda (v\circ R)\otimes(v\circ g_{-t}\circ R)\,d\mu$$
$$= \int_\Lambda (v\circ R)\otimes(v\circ R\circ g_t)\,d\mu.$$

Substituting this into the covariance and Lévy area formulas from Section 2 yields the result. ∎

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

$$a(Sx)+\epsilon^{-1}b(Sx)v(Ry) = -S\big\{a(x)+\epsilon^{-1}b(x)v(y)\big\}\quad \text{for all } x\in\mathbb{R}^d,\ y\in\Lambda,\ \epsilon>0.$$

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*

$$a(Sx)=-Sa(x),\qquad b(Sx)=-Sb(x)A,\qquad v(Ry)=Av(y),$$

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

$$A:X\to\mathbb{R}^{d\times d},\qquad A(x)=-b(x)^{-1}Sb(Sx).$$

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 ± 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 (π^+, π^-) coordinates,*

$$\Sigma = \begin{pmatrix} \Sigma^+ & 0 \\ 0 & \Sigma^- \end{pmatrix}, \qquad E = \begin{pmatrix} 0 & E_0 \\ -E_0^T & 0 \end{pmatrix},$$

where

$$\Sigma^\pm = 2\int_0^\infty \int_\Lambda v^\pm \otimes (v^\pm \circ g_t) \, d\mu \, dt, \qquad E_0 = 2\int_0^\infty \int_\Lambda v^+ \otimes (v^- \circ g_t) \, d\mu \, dt.$$

Proof Since $v \circ R = Av = v^+ - v^-$, we obtain

$$\Sigma = 2\int_0^\infty \int_\Lambda \big\{ v^+ \otimes (v^+ \circ g_t) + v^- \otimes (v^- \circ g_t) \big\} \, d\mu \, dt,$$
$$E = 2\int_0^\infty \int_\Lambda \big\{ v^+ \otimes (v^- \circ g_t) + v^- \otimes (v^+ \circ g_t) \big\} \, d\mu \, dt.$$

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 $\int_\Lambda v \, d\mu = 0$. We exclude the uninteresting case where Λ 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 Σ 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 Λ is *R-invariant* if $f \circ R = f$.

Proposition 4.2 *For any R-invariant functions*

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

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

$$v \in C^r(\mathbb{R}^m, \mathbb{R}^d) \quad \textit{with} \quad v \circ R = Av \quad \textit{and} \quad \textstyle\int_\Lambda v\, d\mu = 0,$$

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

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

Clearly, $v^+ \circ R = v^+ = Av^+$. By the chain rule, $(\nabla h)_{Ry} = (\nabla h)_y R$. 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,

$$\int_0^T v^- \circ g_t\, dt = h \circ g_T - h.$$

Since the flow is mixing and $\int_\Lambda v^+\, d\mu = 0$,

$$\begin{aligned} \int_0^T \int_\Lambda v^+ \otimes (v^- \circ g_t)\, d\mu\, dt &= \int_\Lambda v^+ \otimes \Big(h \circ g_T - h\Big)\, d\mu \\ &\to -\int_\Lambda v^+ \otimes h\, d\mu = \int_\Lambda f \otimes h\, d\mu \end{aligned}$$

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. ∎

Theorem 4.3 *For any* $d^+ \times d^-$ *matrix* F*, there exists a* C^r *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 $L^\pm \in \mathbb{R}^{d^\pm \times d^\pm}$. By the definition of E_0 in Corollary 3.4, transforming v to (L^+v^+, L^-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 $f_1 \in L^2(\Lambda, \mathbb{R}^{d^+})$ and $h_1 \in L^2(\Lambda, \mathbb{R}^{d^-})$ with $\int_\Lambda f_1\, d\mu = 0$ such that $\int_\Lambda f_1 \otimes h_1\, d\mu$ has full rank. The second step is to approximate f_1 and h_1 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 Λ is not a fixed point, we can choose infinitely many orthonormal R-invariant functions $\phi_j \in L^2(\Lambda)$ with $\int_\Lambda \phi_j\, d\mu = 0$ and $\int_\Lambda \phi_i\phi_j\, d\mu = \delta_{ij}$. Let $f_1 = \sum_{i=1}^{d^-} \alpha_i\phi_i$ and $h_1 = \sum_{j=1}^{d^-} \beta_j\phi_j$ where $\alpha_i \in \mathbb{R}^{d^+}$, $\beta_j \in \mathbb{R}^{d^-}$. Then

$$\int_\Lambda f_1 \otimes h_1\, d\mu = \sum_{i,j} (\alpha_i \otimes \beta_j) \int_\Lambda \phi_i\phi_j\, d\mu = \sum_{j=1}^{d^-} \alpha_j \otimes \beta_j.$$

Let $F_1 = \int_\Lambda f_1 \otimes h_1\, d\mu$. Taking β_j to be the j'th canonical unit vector, F_1 is the $d^+ \times d^-$ matrix with columns $\alpha_1, \dots, \alpha_{d^-}$. In particular, F_1 is arbitrary and we can choose the α_i so that F_1 is a matrix of full rank.

Step 2 Now choose R-invariant functions $\tilde\phi_j \in C^\infty(\mathbb{R}^m)$ with $\int_\Lambda \tilde\phi_j\, d\mu = 0$ and $\int_\Lambda |\phi_j - \tilde\phi_j|^2\, d\mu$ small[1] and define f, h using $\tilde\phi_j$ in place of ϕ_j with α_i, β_j unchanged. This results in a matrix $F = \int_\Lambda f \otimes h\, d\mu$ close to F_1. 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 $v \in C^r(\mathbb{R}^m, \mathbb{R}^d)$ 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. ∎

5 Local surjectivity for perturbations of E_0

Let $\chi : C^r(\mathbb{R}^m, \mathbb{R}^d) \to \mathbb{R}^{d^+\times d^-}$ 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 χ is surjective. As mentioned in Remark 4.1, χ is continuous in reasonable situations. In

[1] The existence of such functions is standard. For instance, to approximate ϕ_1 by a C^∞ function, first approximate ϕ_1 in L^2 by a simple function $\sum_{k=1}^\ell c_k 1_{E_k}$ (with $c_k \in \mathbb{R}$ and $E_k \subset \Lambda$ measurable). By outer regularity of the Borel probability measure μ, there exist open neighbourhoods $U_k \subset \mathbb{R}^m$ of E_k with $\mu(U_k \setminus E_k)$ small. Hence $\sum_{k=1}^\ell c_k 1_{U_k}$ is L^2-close to ϕ_1. Choose $V_k \subset \mathbb{R}^m$ closed such that $\overline{U_k} \subset \operatorname{Int} V_k$ with $\mu(V_k - U_k)$ small. By Urysohn's Lemma, there exists a continuous function $\psi_k : \mathbb{R}^m \to [0,1]$ supported on V_k with $\psi_k|_{U_k} \equiv 1$. In this way we obtain a continuous function $\sum_{k=1}^\ell c_k\psi_k$ that is L^2-close to ϕ_1. Each ψ_k can be uniformly approximated by a C^∞ function ζ_k resulting in a C^∞ function $\tilde\phi_1 = \sum_{k=1}^\ell c_k\zeta_k$ that is L^2-close to ϕ_1. Finally, replace $\tilde\phi_1$ by $\tilde\phi_1 - \int_\Lambda \tilde\phi_1\, d\mu$.

this section, we complete the picture by establishing a local surjectivity result for χ; 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 $v_0 \in C^r(\mathbb{R}^m, \mathbb{R}^d)$, there exists v arbitrarily C^r-close to v_0 such that* $\operatorname{rank}\chi(v) = \min\{d^+, d^-\}$.

Proof Suppose without loss of generality that $d^+ \le d^-$. By Theorem 4.3, we can choose $v_* \in C^r(\mathbb{R}^m, \mathbb{R}^d)$ such that $\chi(v_*) = \left(\begin{array}{c|c} I_{d^+} & 0 \end{array} \right)$. Let $v_t = v_0 + tv_*$ and define $E_t = \chi(v_t)$. Then $E_t = \left(\begin{array}{c|c} A_t & B_t \end{array} \right)$ where $A_t = A_0 + tA_1 + t^2 I_{d^+}$ for some matrices $A_0, A_1 \in \mathbb{R}^{d^+ \times d^+}$, $B_t \in \mathbb{R}^{d^+ \times d^-}$. Define the polynomial $p(t) = \det A_t$ of degree $2d^+$. Note that $p(t) = t^{2d^+}(1 + O(t^{-1}))$ so $p(t) \ne 0$ for large t. Since p is a polynomial, it follows that there exists $\epsilon > 0$ such that $p(t) \ne 0$ for all $t \in (0, \epsilon)$. Hence A_t is invertible, and so $\operatorname{rank} E_t = d^+$, for all $t \in (0, \epsilon)$. ∎

Remark 5.2 When Remark 4.1 applies, E (and hence E_0) depends continuously on $v \in C^r$. Hence E_0 has full rank for an open set of $v \in C^r$. By Lemma 5.1, the full rank property for v holds on an open and dense subset of C^r.

In addition, the proof of Lemma 5.1 shows that E_0 having full rank is typical in the probabilistic sense of *prevalence* [7, 19]. In particular, the map $v_0 \mapsto v_t = v_0 + tv_*$ is a "probe" in the terminology of [19].

Next, we require a basic result from linear algebra.

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

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

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 C^r-close to v_0 such that $\chi(v) = F_0$.*

Proof Suppose without loss of generality that $d^+ \le d^-$. By Lemma 5.1, we can make an initial perturbation so that $\operatorname{rank} E_0 = d^+$. 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 $v = (P \oplus Q)v_0$. ∎

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

$$\tfrac{1}{2}\sum_{\alpha,\beta,\gamma=1}^{d} E^{\gamma\beta}\partial_\alpha b^\beta(X)b^{\alpha\gamma}(X) \tag{6.1}$$

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

$$b(Sx) = -Sb(x)A, \tag{6.2}$$

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.[2]

Recall that $A^2 = S^2 = I$ and that we can choose coordinates such that $A = I_{d^+} \oplus I_{d^-}$. Recall also that $E^{\gamma\beta} = 0$ if $1 \le \beta, \gamma \le d^+$ and if $d^+ + 1 \le \beta, \gamma \le d$, while the remaining entries of E are general. Since E is assumed to be nonvanishing, we have $d^\pm \ge 1$.

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

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

and

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

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

$$b^{id}(x) = b^{j1}(x) = x^i$$

[2]A similar argument was used in the more difficult proof of Lemma 5.1. Given one b_* that succeeds, we can use it to perturb any b_0 that fails: the perturbation $b_0 + tb_*$ succeeds for all small $t > 0$.

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 $E^{dd} = 0$, the correction (6.1) reduces to a single term

$$\tfrac{1}{2} E^{d1} \partial_i b^1(x) b^{id}(x) = \tfrac{1}{2} E^{d1} x^i e_j$$

where e_j is the canonical unit vector with 1 in the j'th entry. We know that typically $E^{d1} \neq 0$, so this yields a nontrivial correction to the drift term a as required.

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 $\Lambda = T^1M$ is a compact three-dimensional manifold and the geodesic flow $g_t : \Lambda \to \Lambda$ is given by $g_t(x) = \gamma_x'(t)$ where $\gamma_x : \mathbb{R} \to M$ is the unique geodesic such that $\gamma_x'(0) = x$. An ergodic probability measure is given by the normalised Liouville measure μ. Writing $y \in \Lambda$ as $y = (q, p)$ where $q \in M$ and $p \in T_q^1 M$, 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 $A = S = R = \operatorname{diag}\{1, -1\}$. In particular, $d^+ = d^- = 1$ and E_0 is scalar.

Let $g : T^1M \to T(T^1M)$ be the vector field for the geodesic flow and let $f, h : T^1M \to \mathbb{R}$ 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 p as a unit vector in $\mathbb{R}^2$. Let κ denote curvature. Explicit choices of f and h are $f = \psi - \int \psi$ and $h = \psi$ where $\psi(q, p) = \kappa(q) p_2^2$, resulting in

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

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

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

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

$$\tilde{a}(x) = a(x) + (0, \tfrac{1}{2} E_0 x_1) = (x_2 \cos x_1, \cos x_2 + \tfrac{1}{2} E_0 x_1)$$

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.

Acknowledgements This research was supported in part by the visitor program at the Sydney Mathematical Research Institute (SMRI) in early 2020. IM 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.

References

[1] C. Andrieu and S. Livingstone. Peskun-Tierney ordering for Markovian Monte Carlo: beyond the reversible scenario. *Ann. Statist.* **49** (2021) 1958–1981.

[2] R. F. Bass, B. M. Hambly and T. J. Lyons. Extending the Wong-Zakai theorem to reversible Markov processes. *J. Eur. Math. Soc. (JEMS)* **4** (2002) 237–269.

[3] P.-M. Boulvard and E. Mémin. Diagnostic of the Lévy area for geophysical flow models in view of defining high order stochastic discrete-time schemes. *Foundations of Data Science* (2023).

[4] N. Chernov and D. Dolgopyat. Brownian Brownian motion. I. *Mem. Amer. Math. Soc.* **198** (2009) no. 927, viii+193.

[5] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne and H. Zhang. Multiscale systems, homogenization, and rough paths. *Probability and Analysis in Interacting Physical Systems: In Honor of S.R.S. Varadhan, Berlin, August, 2016"* (P. Friz et al., ed.), Springer Proceedings in Mathematics & Statistics **283**, Springer, 2019, pp. 17–48.

[6] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne and H. Zhang. Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 2. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** (2022) 1328–1350.

[7] J. P. R. Christensen. Measure theoretic zero sets in infinite dimensional spaces and applications to differentiability of Lipschitz mappings. *Publ. Dép. Math. (Lyon)* **10** (1973) 29–39.

[8] S. Delong, Y. Sun, B. E. Griffith, E. Vanden-Eijnden and A. Donev. Multiscale temporal integrators for fluctuating hydrodynamics. *Phys. Rev. E* **90** (2014) 063312.

[9] J.-D. Deuschel, T. Orenshtein and N. Perkowski. Additive functionals as rough paths. *Ann. Probab.* **49** (2021) 1450–1479.

[10] T. Diamantakis, D. D. Holm and G. A. Pavliotis. Variational principles on geometric rough paths and the Lévy area correction. *SIAM J. Appl. Dyn. Syst.* **22** (2023) 1182–1218.

[11] R. L. Dobrushin, Yu. M. Sukhov and Ĭ. Fritts. A. N. Kolmogorov—founder of the theory of reversible Markov processes. *Uspekhi Mat. Nauk* **43** (1988) 167–188.

[12] D. Dolgopyat. On the decay of correlations in Anosov flows. *Ann. of Math.* **147** (1998) 357–390.

[13] Y. Fang, J.-M. Sanz-Serna and R. D. Skeel. Compressible generalized hybrid Monte Carlo. *J. Chem. Phys.* **140** (2014) 174108.

[14] N. Fleming-Vázquez. Functional correlation bounds and optimal iterated moment bounds for slowly-mixing nonuniformly hyperbolic maps. *Comm. Math. Phys.* **391** (2022) 173–198.

[15] C. Fiorini, P.-M. Boulvard, L. Li and E. Mémin. A two-step numerical scheme in time for surface quasi geostrophic equations under location uncertainty. *Stochastic Transport in Upper Ocean Dynamics* (B. Chapron, D. Crisan, D. Holm, E. Mémin and A. Radomska, eds.), Springer International Publishing, Cham, 2023, pp. 57–67.

[16] P. Friz and H. Oberhauser. Rough path limits of the Wong-Zakai type with a modified drift term. *J. Funct. Anal.* **256** (2009) 3236–3256.

[17] G. A. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proc. R. Soc. London A* **469** (2013) 20130201.

[18] K. Hukushima and Y. Sakai. An irreversible Markov-chain Monte Carlo method with skew detailed balance conditions. *J. Phys., Conf. Ser.* **473** (2013) 012012.

[19] B. R. Hunt, T. Sauer and J. A. Yorke. Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces. *Bull. Amer. Math. Soc. (N.S.)* **27** (1992) 217–238.

[20] D.-Q. Jiang, M. Qian and M.-P. Qian. *Mathematical Theory of Nonequilibrium Steady States*. Lecture Notes In Mathematics **1833**, Springer, New York, 2004.

[21] D. Kelly and I. Melbourne. Smooth approximation of stochastic differential equations. *Ann. Probab.* **44** (2016) 479–520.

[22] A. Korepanov, Z. Kosloff and I. Melbourne. Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 1. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** (2022) 1305–1327.

[23] A. Lejay and T. Lyons. 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. **4**, Theta, Bucharest, 2005, pp. 63–84.

[24] S. H. Lim. Anomalous thermodynamics in homogenized generalized Langevin systems. *Journal of Physics A: Mathematical and Theoretical* **54** (2021) 155001.

[25] O. Lopusanschi and D. Simon. Lévy area with a drift as a renormalization limit of Markov chains on periodic graphs. *Stochastic Process. Appl.* **128** (2018) 2404–2426.

[26] R. S. MacKay. Langevin equation for slow degrees of freedom of Hamiltonian systems. *Nonlinear dynamics and chaos: advances and perspectives*, Underst. Complex Syst., Springer, Berlin, 2010, pp. 89–102.

[27] E. J. McShane. Stochastic differential equations and models of random processes. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory* (Berkeley, Calif.), Univ. California Press, 1972, pp. 263–294.

[28] G. A. Pavliotis. *Stochastic processes and applications*. Texts in Applied Mathematics **60**, Springer, New York, 2014. Diffusion processes, the Fokker-Planck and Langevin equations.

[29] G. A. Pavliotis and A. M. Stuart. *Multiscale methods: Averaging and homogenization*. Texts in Applied Mathematics **53**, Springer, New York, 2008.

[30] H. J. Sussmann. Limits of the Wong-Zakai type with a modified drift term. *Stochastic analysis*, Academic Press, Boston, MA, 1991, pp. 475–493.

[31] E. Wong and M. Zakai. On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Statist.* **36** (1965) 1560–1564.

[32] A. M. Yaglom. On the statistical reversibility of Brownian motion. *Mat. Sbornik N.S.* **24/66** (1949) 457–492.