An Eulerian Perspective on Straight-Line Sampling

Panos Tsimpos
Operations Research Center
MIT, Cambridge, MA 02139, USA
ptsimpos@mit.edu

Youssef Marzouk
Laboratory of Information and Decision Systems
MIT, Cambridge, MA 02139, USA
ymarz@mit.edu

Abstract

We study dynamic measure transport for generative modeling: specifically, flows induced by stochastic processes that bridge a specified source and target distribution. The conditional expectation of the process' velocity defines an ODE whose flow map achieves the desired transport. We ask *which processes produce straight-line flows*—i.e., flows whose pointwise acceleration vanishes and thus are exactly integrable with a first-order method? We provide a concise PDE characterization of straightness as a balance between conditional acceleration and the divergence of a weighted covariance (Reynolds) tensor. Using this lens, we fully characterize affine-in-time interpolants and show that straightness occurs exactly under deterministic endpoint couplings. We also derive necessary conditions that constrain flow geometry for general processes, offering broad guidance for designing transports that are easier to integrate.

1 Introduction

Sampling from complex probability distributions is central to probabilistic inference and modern generative modeling. A recent line of work establishes *dynamic measure transport* as a unifying paradigm: construct a stochastic process $(X_t)_{t\in[0,1]}$ whose marginals interpolate from a tractable source distribution μ_0 to a target distribution μ_1, estimate the *conditional velocity* $v_t(x) := \mathbb{E}[\dot{X}_t \mid X_t = x]$ and generate samples by evaluating the ODE flow maps ϕ_t defined by $\partial_t\phi_t(x) = v_t(\phi_t(x))$ with initial condition $\phi_0(x) = x$. This perspective underlies methods such as *stochastic interpolants* [ABVE23, AVE22], *flow matching* [LCBH+23, TFM+24], and *score-based probability flow ODEs* [SSDK+21], as well as *rectified flows* [LGL22, Liu22, BRSR25, HCD25]—all of which have demonstrated strong empirical performance.

The computational efficiency of these methods hinges on the *geometry of the induced flow*. Generic flows demand many velocity-oracle evaluations because numerical integration error scales with the curvature (and higher derivatives) of ϕ_t. In contrast, if the flow is *straight*, meaning that the acceleration of the flow map vanishes,

$$\partial_{tt}\,\phi_t(x) \equiv 0 \quad \text{for all } (x,t),$$

then $\phi_t(x)$ is affine in t: $\phi_t(x) = (1-t)x + t\,\phi_1(x)$. Consequently, any first-order integrator is exact; one can traverse the entire path with *a single* velocity evaluation. This motivates a fundamental question:

Which stochastic processes $(X_t)_t$ with $X_0 \sim \mu_0$ and $X_1 \sim \mu_1$ induce straight flows?

Prior work (e.g., [LGL22, Liu22, BRSR25]) offers algorithmic frameworks that convert a given non-straight flow into a straight one, but a *structural characterization* of when straightness is intrinsic to the underlying process has remained open. This paper develops such a theory.

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Frontiers in Probabilistic Inference: Sampling Meets Learning.

Contributions. Our key contributions are as follows:

1. *PDE criterion for straightness.* We derive a new balance law, equation (2), that characterizes straight flows.
2. *Complete analysis of linear interpolations.* For processes of the form $X_t = (1-t)\, X_0 + t\, X_1$, with $X_0 \sim \mu_0$ and $X_1 \sim \mu_1$, we show that the resulting flow is straight if and only if $(X_t)_t$ is constructed from a deterministic coupling of μ_0 and μ_1.
3. *Necessary conditions for the general case.* We obtain geometric constraints that any straight line-inducing $(X_t)_{t\in[0,1]}$ must satisfy.

Scope and implications. This short paper is fully *theoretical* and applies to a broad class of stochastic processes. It offers a *new PDE lens* on flow models and their straightness. We expect this framework to aid the principled design of stochastic processes for sampling; here we extract only a few immediate consequences and highlight open directions for theoretical and algorithmic follow-up.

2 Main results

2.1 Preliminaries

For the standard notation used in this paper, please refer to Section A.1. Below, we review less typical notation used in our work. Fix a stochastic process $X := (X_t)_{t\in[0,1]}$ with sample paths in $W^{2,2}([0,1];\mathbb{R}^d)$. We define the *conditional velocity* and *conditional acceleration* fields, also referred to as the *ensemble velocity* and *ensemble acceleration*, by $v_t(x) := \mathbb{E}\left[\dot{X}_t \mid X_t = x\right]$ and $a_t(x) := \mathbb{E}\left[\ddot{X}_t \mid X_t = x\right]$. We define the *second moment velocity tensor* and the *covariance* or *Reynolds stress tensor* as $\Sigma_t(x) := \mathbb{E}\left[\dot{X}_t \otimes \dot{X}_t \mid X_t = x\right]$ and $\Pi_t(x) := \Sigma_t(x) - v_t(x) \otimes v_t(x)$, respectively. We let $\mu_t = \mathrm{Law}(X_t)$ be the marginal law of the process X at time t and write ρ_t for the density of μ_t with respect to the Lebesgue measure, if it exists. Furthermore, given a velocity field $v := (v_t)_t$, where $t \in [0,1]$ and $v_t : D \subset \mathbb{R}^d \to \mathbb{R}^d$, we define the induced *flow* $\phi_t : \mathbb{R}^d \to \mathbb{R}^d$ to be the solution map to the ODE $\partial_t\, \phi_t(x) = v_t(\phi_t(x))$ with initial condition $\phi_0(x) = x$. We call a flow *straight* if it is of the form $\phi_t(x) = (1-t)\, x + t\, \phi_1(x)$, which is equivalent to $\partial_{tt}\, \phi_t(x) = 0$. We define the *material derivative* of v_t at $x \in \mathbb{R}^d$ by $D_t\, v_t(x) = \partial_t\, v_t(x) + (v_t(x) \cdot \nabla)\, v_t(x)$. Finally, for matrices $A, B \in \mathbb{R}^{d_1 \times d_2}$ we denote the *Frobenius inner product* by $A : B := \mathrm{Tr}(A^\top B) = \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} A_{ij} B_{ij}$, and for a matrix field $T : D \subset \mathbb{R}^d \to \mathbb{R}^{d\times d}$ consisting of differentiable entries we define its *divergence* as the vector field $\nabla \cdot T : D \to \mathbb{R}^d$ with components $(\nabla \cdot T)_j(x) = \sum_{i=1}^d \partial_i T_{ij}(x)$.

2.2 PDE characterization of straight-line flows.

Fix $D \subset \mathbb{R}^d$ compact. We are concerned with the following problem:

Problem: Given $\mu_0, \mu_1 \in \mathcal{P}(D)$, characterize all stochastic processes $(X_t)_{t\in[0,1]}$, with $X_i \sim \mu_i$ for $i \in \{0,1\}$, such that the flow $(\phi_t)_t$ generated by the ensemble velocity $(v_t)_t$ satisfies:

$$\partial_{tt}\, \phi_t(x) = 0\,, \quad \forall\, t \in [0,1]\,.$$

Before we start, we make some regularity assumptions that simplify the analysis:

Assumption 1. *The marginal densities ρ_t of X_t exist and are positive ($\rho_t > 0$) on the compact set $D \subset \mathbb{R}^d$, and vanish on D^c. Also, sample paths of X_t are in $W^{2,2}([0,1];\mathbb{R}^d)$ and the induced flow maps $t \mapsto \phi_t(x)$ are in $C^2([0,1];\mathbb{R}^d)$ for each $x \in D$.*

First, an elementary reformulation allows us to move from the flow ϕ_t to the ensemble velocity.

Proposition 1. *For all $(t,x) \in [0,1] \times D$, one has*

$$\partial_{tt}\, \phi_t(x) = 0 \quad \iff \quad D_t v_t(x) = 0\,.$$

Please see Appendix A for the proof. This result essentially constitutes a passage from the Lagrangian to the Eulerian perspective. Rather than tracking the motion of a single particle via the flow map, the Eulerian perspective considers the *total* change of the ensemble velocity field.

The second step is to relate the material derivative to other statistical quantities of interest, in particular the second order tensors Σ_t and Π_t. Here, a *momentum balance* identity allows us to make progress.

Lemma 2. *The following equation holds for all $t \in [0,1]$ and $x \in D \subseteq \mathbb{R}^d$:*

$$\partial_t \left(\rho_t \, v_t\right) + \nabla \cdot \left(\rho_t \, \Sigma_t\right) = \rho_t \, a_t \, . \tag{1}$$

For the proof, refer to Appendix A. This is a fundamental identity with an elegant physical interpretation. Fix a control volume dV. From left to right in (1): the first term is the rate of change of the momentum in dV, the second term is the net momentum flux out of dV, and the third is the net body force acting on dV. Put differently, this is a manifestation of the $F = ma$ and $\dot{p}_t = F_t$ relations, where p is momentum, a is acceleration, m is mass, and F is force.

Finally, inserting the definition of the Reynolds tensor Π_t together with the continuity equation into the above identity, we obtain yet another identity, elucidating the connection between the material derivative and the Reynolds (i.e., covariance) tensor.

Lemma 3. *The following equation holds for all $(x,t) \in [0,1] \times D$:*

$$\rho_t \, D_t v_t + \nabla \cdot \left(\rho_t \, \Pi_t\right) = \rho_t \, a_t \, .$$

This immediately leads to a corollary and a reformulation of our problem.

Corollary 4. *For all $(t,x) \in [0,1] \times D$ we have:*

$$D_t v_t = 0 \iff \nabla \cdot \left(\rho_t \, \Pi_t\right) = \rho_t \, a_t \, .$$

Problem reformulation: Given $\mu_0, \mu_1 \in \mathcal{P}(D)$, characterize all stochastic processes $(X_t)_{t\in[0,1]}$, with $X_i \sim \mu_i$ for $i \in \{0,1\}$, such that

$$\nabla \cdot \left(\rho_t \, \Pi_t\right) = \rho_t \, a_t \, , \quad \forall \, t \in [0,1] \, . \tag{2}$$

2.3 Linear characteristics

In this section we demonstrate the usefulness of reformulation (2) to obtain a complete characterization of affine processes that give rise to straight flows. We define an *affine process* with marginals $\mu_0, \mu_1 \in \mathcal{P}(\mathbb{R}^d)$ to be a process of the form

$$X_t = (1-t) \, X + t \, Y \, ,$$

where $(X,Y) \sim \gamma$ and γ is a coupling of the measures μ_0 and μ_1. Although seemingly restrictive, this process is widespread in computational statistics and generative modelling, e.g., optimal-transport displacement interpolation and modern flow-based generative modeling where linear sample-to-sample paths or linear noise–data blends define the training trajectory; see [McC97, LCBH+23, Liu22, ABVE23].

The main result of this section is that the flow generated by the ensemble velocity of an affine process can be of straight-line type if and only if the coupling γ between the endpoints is deterministic, i.e., iff there is some measurable map $T : \mathbb{R}^d \to \mathbb{R}^d$ such that $T(X) = Y$ almost surely, or equivalently $\gamma = (\mathrm{id} \times T)_\sharp \mu_0$, where $\mathrm{id} \times T : D \to \mathbb{R}^d \times \mathbb{R}^d$ acts as $x \mapsto (x, T(x))$.

The key input, here, is that an affine process has vanishing ensemble acceleration $a_t(x) = \mathbb{E}[\ddot{X}_t \mid X_t = x] \equiv 0$ since, of course, it has no acceleration at the "particle" level, $\ddot{X}_t \equiv 0$. Thus, our straight line flow characterization given by equation (2) becomes

$$\nabla \cdot \left(\rho_t \, \Pi_t\right) = 0 \, .$$

Theorem 5. *Under the additional assumptions that $\mathbb{E} \, \|X_t\|^2 < \infty$ and $\mathbb{E} \, \|\Pi_t(X_t)\|^2 < \infty$ for all $t \in [0,1]$, it holds that $\nabla \cdot \left(\rho_t \, \Pi_t\right) = 0$ if and only if there exists a measurable map $T : D \to D$ satisfying $T_\sharp \mu_0 = \mu_1$ such that $(\mathrm{id} \times T)_\sharp \, \mu_0 = \gamma$, where $\mathrm{id} : \mathbb{R}^d \to \mathbb{R}^d$ is the identity map.*

Please find the proof in Appendix A. This result characterizes all affine processes inducing straight flows as those with deterministically coupled endpoints.

2.4 Geometric constraints on arbitrary processes.

In this section, we fix some process $(X_t)_{t\in[0,1]}$ satisfying (2) and derive some necessary conditions constraining the geometry of the sample paths of X. Namely, by repeating the argument of Theorem 5 we obtain:

Theorem 6. *Under the assumptions* $\mathbb{E}\,\|X_t\|^2 < \infty$, $\mathbb{E}\,\|\Pi_t(X_t)\|^2 < \infty$, *and* $\mathbb{E}\,\|a_t(X_t)\|^2 < \infty$, *any process* X *satisfying* (2) *satisfies* $-\mathbb{E}\Big[\operatorname{Tr}\Pi_t(X_t)\Big] = \mathbb{E}\left[X_t \cdot \ddot{X}_t\right]$

This result links the expected (rescaled) *radial acceleration* with the integral of the trace of the covariance tensor. Since this trace, however, consists of non-negative quantities, and combining with the identity $\partial_{tt}\,\|X_t\|^2 = 2\,X_t \cdot \ddot{X}_t + 2\,\|\dot{X}_t\|^2$ we obtain:

Corollary 7. *Any process satisfying the integrability assumptions of Theorem 6 together with equation* (2) *satisfies the identities:*

$$
\begin{aligned}
&(1) \quad \mathbb{E}\left[X_t \cdot \ddot{X}_t\right] \leq 0\,, \\
&(2) \quad \mathbb{E}\Big[\partial_{tt}\,\|X_t\|^2\Big] \leq 2\,\mathbb{E}\left[\|\dot{X}_t\|^2\right]\,,
\end{aligned}
$$

Though perhaps opaque at first glance, we believe these identities can provide meaningful insight into the structure of non-trivial solutions to the PDE (2).

3 Discussion

We have developed a novel PDE characterization of straight-line flows generated by the ensemble velocity of a stochastic process indexed on the unit time interval, with given marginals μ_0 and μ_1. Our main insight is a characterization of straight line flows in terms of a balance law that links the conditional variance tensor, i.e., the Reynolds tensor, to the ensemble acceleration field. Using this characterization, we (i) showed that affine processes yielding straight-line flows must have deterministically coupled endpoints; and (ii) derived necessary conditions constraining the geometry of arbitrary processes that induce straight flows.

Limitations. Our analysis is fully theoretical and operates under regularity assumptions, both on the marginals of the process X, e.g., absolute continuity and positivity of the density, as well on the sample paths of X, e.g., that they are at least in $W^{2,2}$, to make sense of the velocity and acceleration fields. While applicable to many modern flow-based models, this rules out other important processes, such as diffusions.

Open directions.

- *Non-trivial solutions to* (2)*.* Construct processes with non-trivial acceleration fields $a_t \not\equiv 0$ that satisfy (2). Such processes might give rise to novel sampling dynamics, given the empirical effectiveness of other processes inducing straight line flows.
- *Full characterization of straightness.* Derive necessary and sufficient conditions on the process X under which the PDE (2) holds. Of particular interest is the case of $X_t = F(t, X_0, X_1)$ for a sufficiently regular $F : [0,1] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and a random vector $(X_0, X_1) \sim \mu_0 \otimes \mu_1$. This is precisely the setting of *stochastic interpolants* [AVE22, ABVE23] and readily leads to algorithmic insights.
- *Process classes.* Extend the characterization to SDEs (drift–diffusion pairs), non-Markovian processes, or manifold-constrained processes.

Outlook. We view the balance law $\nabla\cdot(\rho_t\Pi_t) = \rho_t a_t$ and its consequences as a compact organizing principle for *geometry-aware* transport design. While we extract only a few implications here, we expect this perspective to guide principled constructions of stochastic processes for *sampling*, and to catalyze empirical investigations into when—and how—straightness can be achieved in practice.

Acknowledgments and Disclosure of Funding

PT and YM acknowledge support from the US Air Force Office of Scientific Research (AFOSR) MURI program, under award number FA9550-20-1-0397.

References

[ABVE23] Michael S Albergo, Nicholas M Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *arXiv preprint arXiv:2303.08797*, 2023.

[AVE22] Michael S Albergo and Eric Vanden-Eijnden. Building normalizing flows with stochastic interpolants. *arXiv preprint arXiv:2209.15571*, 2022.

[Bog07] Vladimir I Bogachev. *Measure theory*. Springer, 2007.

[BRSR25] Vansh Bansal, Saptarshi Roy, Purnamrita Sarkar, and Alessandro Rinaldo. On the wasserstein convergence and straightness of rectified flow, 2025.

[HCD25] Johannes Hertrich, Antonin Chambolle, and Julie Delon. On the relation between rectified flows and optimal transport. *arXiv preprint arXiv:2505.19712*, 2025.

[LCBH+23] Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling, 2023.

[LGL22] Xingchao Liu, Chengyue Gong, and Qiang Liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. *arXiv preprint arXiv:2209.03003*, 2022.

[Liu22] Qiang Liu. Rectified flow: A marginal preserving approach to optimal transport, 2022.

[McC97] Robert J. McCann. A convexity principle for interacting gases. *Advances in Mathematics*, 128(1):153–179, 1997.

[SSDK+21] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations, 2021.

[TFM+24] Alexander Tong, Kilian Fatras, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport, 2024.

A Proofs

A.1 Notation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an abstract probability space. Random variables are measurable functions $X : \Omega \to \mathbb{R}^d$, [1] and the expectation $\mathbb{E}[X]$ is defined as the integral $\int_\Omega X \, d\mathbb{P}$. For random variables $X, Y : \Omega \to \mathbb{R}^d$ the conditional expectation $\mathbb{E}[X \mid Y]$ is defined at the conditional expectation $\mathbb{E}[X \mid \sigma(Y)]$ where $\sigma(Y)$ is the σ-algebra generated by Y and for $y \in \mathbb{R}$ the conditional expectation $\mathbb{E}[X \mid Y = y]$ is the integral $\int_\Omega X \, d\mathbb{P}_y$ and $\mathbb{P}_y$ is the *disintegration* [Bog07, Section 10.6] of $\mathbb{P}$ on the level sets of Y. Finally, given another measurable space $(\Omega', \mathcal{F}')$ we denote the set of measures on Ω' by $\mathcal{P}(\Omega')$ and the subset of absolutely continuous measures by $\mathcal{P}_{\text{a.c.}}(\Omega')$.

Let $C^k([0,1]; \mathbb{R}^d)$ be the space of component-wise k-times continuously differentiable functions from $[0,1]$ to $\mathbb{R}^d$ and $W^{k,p}([0,1] \to \mathbb{R}^d)$ to be the space of component-wise k-times weakly differentiable functions $[0,1] \to \mathbb{R}^d$ with weak derivatives in $L^p([0,1])$. An $\mathbb{R}^d$-valued stochastic process is a collection $\{X_t\}_{t \in I}$ indexed by some set I, where $X : \Omega \to \mathbb{R}^d$ is measurable. In this paper, we take $I = [0,1]$ and write $X \coloneqq (X_t)_t \coloneqq (X_t)_{t \in [0,1]}$. A sample path of a stochastic process is the function $t \mapsto X_t(\omega)$ for a fixed realization $\omega \in \Omega$. We note that a stochastic process with sample paths in $W^{k,p} \coloneqq W^{k,p}([0,1]; \mathbb{R}^d)$ can equivalently be viewed as a random variable $X : \Omega \to W^{k,p}$ and its law is thus in $\mathcal{P}(W^{k,p})$.

A.2 Proofs

Proposition 1. *The following are equivalent:*

1. *For all $(t, x) \in [0,1] \times D$, one has*

$$\frac{d^2}{dt^2} \phi_t(x) = 0 \, .$$

[1] In this paper, subsets of $\mathbb{R}^d$ are always equipped with the Borel σ-algebra.

2. *For all* $(t, x) \in [0, 1] \times D$*, one has*

$$\phi_t(x) = (1 - t)\, x + t\, \phi_1(x)\,,$$

and $(\phi_1)_\sharp\, \mu_0 = \mu_1$.

3. *For all* $(t, x) \in [0, 1] \times D$*, one has*

$$D_t v_t(x) = 0\,.$$

Proof. It is clear that $(2) \implies (1)$ Let us show that $(1) \implies (2)$. We have

$$\frac{d^2}{dt^2}\phi_t(x) = 0 \qquad \implies \qquad \exists c \in \mathbb{R}^d \,:\, \partial_t \phi_t(x) = c\ ,$$

so using the boundary condition $\phi_t(x) = x$ we get

$$\phi_t(x) = x + t\, c\,.$$

This further implies

$$c = \phi_1(x) - x\,,$$

and doing some algebra we get

$$\phi_t(x) = (1 - t)\, x + t\, \phi_1(x)\,.$$

The fact that $(\phi_1)_\sharp\, \mu_0 = \mu_1$ follows from the boundary condition $\phi_1(X_0) \sim \mu_1$.
Finally, let us show that $(1) \iff (3)$. By definition we have

$$\partial_t \phi_t(x) = v_t(\phi_t(x))\,,$$

and, therefore,

$$\begin{aligned}\frac{d^2}{dt^2}\phi_t(x) &= \partial_t v_t(\phi_t(x)) \\ &= \partial_t v_t(\phi_t(x)) + (v_t(x) \cdot \nabla)\, v_t(\phi_t(x)) \\ &= D_{v_t(\phi_t(x))} v_t(\phi_t(x))\,.\end{aligned}$$

Here, we note that by Assumption 1 the map $\phi_t : D \to D$ is surjective. Indeed, since the domain of ϕ_t is compact and ϕ_t is continuous, the image $\phi_t(D)$ is compact and hence closed. Now assume there is $x \in D^\circ \cap \phi(D)^c$, where D° is the topological interior of ϕ_t. As a finite intersection of open sets, this set is open; hence there is an open neighborhood $x \in U \subset D^\circ$ with $U \cap \phi(D)^c = \emptyset$. But this means that $U \cap \text{supp}\rho_t = \emptyset$, since $\rho_t = (\phi_t)_\sharp\, \rho_0$, contradicting the assumed positivity of ρ_t in D. Thus, we have shown that $\phi_t : D \to D^\circ$ is surjective and since ϕ_t is continuous we can extend it uniquely to ∂D, hence, w.l.o.g. we have that $\phi_t : D \to D$ is surjective. Thus, we can conclude that

$$D_t v_t(x) = 0 \qquad \iff \qquad \frac{d^2}{dt^2}\phi_t(x) = 0\,.$$

□

Lemma 2. *The following equation holds for all* $t \in [0, 1]$ *and* $x \in D \subseteq \mathbb{R}^d$*:*

$$\partial_t(\rho_t\, v_t) + \nabla \cdot (\rho_t\, \Sigma_t) = \rho_t\, a_t\,.$$

Proof. For a vector valued test function $\Phi \in C_c^\infty(D;\mathbb{R}^d)$, we have

$$\begin{aligned}
\int \Phi(x)\cdot\partial_t(\rho_t(x)\,v_t(x))\,dx &= \partial_t \int \Phi(x)\cdot(\rho_t(x)\,v_t(x))\,dx \\
&= \partial_t \int \Phi(x)\cdot v_t(x)\,d\rho_t(x) \\
&= \partial_t \mathbb{E}\left[\,\Phi(X_t)\cdot\mathbb{E}\left[\dot{X}_t \mid X_t\right]\right] \\
&= \partial_t \mathbb{E}\left[\,\Phi(X_t)\cdot\dot{X}_t\right] \\
&= \mathbb{E}\left[\,\Phi(X_t)\cdot\ddot{X}_t\right] + \mathbb{E}\left[\left(\nabla\Phi(X_t)\,\dot{X}_t\right)\cdot\dot{X}_t\right] \\
&= \mathbb{E}\left[\,\Phi(X_t)\cdot a_t(X_t)\,\right] + \mathbb{E}\left[\,\nabla\Phi(X_t) : \dot{X}_t\otimes\dot{X}_t\right] \\
&= \mathbb{E}\left[\,\Phi(X_t)\cdot a_t(X_t)\,\right] + \mathbb{E}\left[\,\nabla\Phi(X_t) : \Sigma_t(X_t)\right] \\
&= \int \Phi(x)\cdot(\rho_t(x)\,a_t(x))\;dx + \int \nabla\Phi(x) : (\rho_t(x)\,\Sigma_t(x))\;dx \\
&= \int \Phi(x)\cdot(\rho_t(x)\,a_t(x))\;dx - \int \Phi(x)\cdot\left[\nabla\cdot(\rho_t(x)\,\Sigma_t(x))\right]\;dx
\end{aligned}$$

where we have only used definitions and the integration by parts formula. Since the above holds for all $\Phi \in C_c^\infty(D;\mathbb{R}^d)$, we can conclude that

$$\partial_t(\rho_t\,v_t) + \nabla\cdot(\rho_t\,\Sigma_t) = \rho_t\,a_t\,,$$

concluding the proof. □

Lemma 3. *The following equation holds for all $(x,t)\in[0,1]\times D$:*

$$\rho_t\,D_t v_t + \nabla\cdot(\rho_t\,\Pi_t) = \rho_t\,a_t\,.$$

Proof. By the definition of the Reynolds stress tensor we have

$$\Sigma_t(x) = v_t\otimes v_t + \Pi_t(X_t)\,.$$

Now write $\pi_t^{(i)}\in\mathbb{R}^d$ for the i-th row of Π_t and $v_t^i, a_t^i\in\mathbb{R}\times\mathbb{R}$ for the i-th components of v_t and a_t, respectively. Plugging the above display into (2) and looking at the i-th component of the resulting vector we obtain

$$\partial_t\left(\rho_t\,v_t^i\right) + \nabla\cdot\left(\rho_t\,v_t^i\,v_t\right) + \nabla\cdot\left(\rho_t\,\pi_t^{(i)}\right) = \rho_t\,a_t^i\,.$$

Now expanding the differential operators we have

$$v_t^i\,\partial_t\rho_t + \rho_t\,\partial_t v_t^i + v_t^i\,\nabla\cdot(\rho_t\,v_t) + \rho_t\,v_t\cdot\nabla v_t^i + \nabla\cdot(\rho_t\,\pi_t^{(i)}) = \rho_t\,a_t^i\,.$$

and rearranging

$$v_t^i\left(\partial_t\rho_t + \nabla\cdot(\rho_t\,v_t)\right) + \rho_t\left(\partial_t v_t^i + v_t\cdot\nabla v_t^i\right) + \nabla\cdot(\rho_t\,\pi_t^{(i)}) = \rho_t\,a_t^i\,.$$

Using the continuity equation (7) the first parenthesis vanishes and vectorizing the equation we obtain

$$\rho_t\,\left(\partial_t v_t + v_t\cdot\nabla v_t\right) + \nabla\cdot(\rho_t\,\Pi_t) = \rho_t\,a_t\,.$$

Finally, using the definition of the material derivative we conclude. □

Theorem 4. *Under the additional assumptions that*

$$\mathbb{E}\,\|X_t\|^2 < \infty \text{ and}$$
$$\mathbb{E}\,\|\Pi_t(X_t)\|^2 < \infty\,,$$

for all $t\in[0,1]$ one has

$$\nabla\cdot(\rho_t\,\Pi_t) = 0\,,$$

if and only if there exists a measurable map $T:\mathbb{R}^d\to\mathbb{R}^d$ satisfying $T_\sharp\mu_0 = \mu_1$ such that

$$(\mathrm{id}\times T)_\sharp\,\mu_0 = \gamma\,,$$

where $\mathrm{id}:\mathbb{R}^d\to\mathbb{R}^d$ *is the identity map.*

Proof. For $D^\circ \subset \mathbb{R}^d$, fix $R > 0$ and take a cut-off function $\eta_R \in C_c^\infty(D)$ such that $\text{supp}(\eta_R) \subset B_{2R}(0)$ and $\text{supp}(1-\eta_R) \subset B_R(0)^c$. Now consider the test function $\Phi_R(x) \in C_c^\infty(D;\mathbb{R}^d)$ given by $\Phi_R(x) = \eta_R(x)\, x$ and compute

$$\begin{aligned}
&\int \Phi_R(x) \cdot (\nabla \cdot (\rho_t\, \Pi_t))\; dx = \\
&- \int \nabla \Phi_R(x) : (\rho_t\, \Pi_t)\; dx = \\
&- \int \nabla \Phi_R(x) : \Pi_t(x)\, d\rho_t(x) = \\
&- \int_{B_R(0)} I_d : \Pi_t(x)\, d\rho_t(x) + \int_{B_R(0)^c} \nabla \Phi_R(x) : \Pi_t(x)\, d\rho_t(x) = \\
&- \int_{B_R(0)} \operatorname{Tr} \Pi_t(x)\, d\rho_t(x) + \int_{B_R(0)^c} \nabla \Phi_R(x) : \Pi_t(x)\, d\rho_t(x) =
\end{aligned}$$

Taking the limit $R \to \infty$ and using Lemma 5 together with dominated convergence we obtain

$$\int_{\mathbb{R}^d} \operatorname{Tr} \Pi_t(x)\, d\rho_t(x) = 0$$

which is equivalent to

$$\sum_{i=1}^d \int_{\mathbb{R}^d} \text{Var}\left(\dot{X}_t^i \mid X_t = x \right)\; d\rho_t(x) = 0\,.$$

Now since $\text{Var}(\dot{X}_t^i | X_t = x) \geq 0$ we have that

$$\text{Var}\left(\dot{X}_t^i \mid X_t = x \right) = 0 \text{ for } \rho_t\text{-almost every } x \in \mathbb{R}^d.$$

Thus, there is a Borel measurable function $G : \mathbb{R}^d \to \mathbb{R}^d$ such that for all $t \in [0,1]$

$$\dot{X}_t = G(X_t) \text{ almost surely.}$$

Taking $t = 0$ this reads $Y = X + G(X)$ almost surely, thus after setting $T = \text{id} + G$ we have

$$Y = T(X) \text{ almost surely.}$$

This concludes the proof of the "only if" part of the theorem. For the "if" part, notice that under the assumptions of the theorem we have $Y = T(X)$ almost surely and hence

$$X_t = (1-t)\, X + t\, T(X)$$

implying that $\text{Var}(\dot{X}_t | X_t = x) = 0$ for ρ_t-almost every $x \in \mathbb{R}^d$. This finishes the proof. □

Lemma 5. *Assume that*

$$\begin{aligned}
&\mathbb{E}\, \|X_t\|^2 < \infty \text{ and} \\
&\mathbb{E}\, \|\Pi_t(X_t)\|^2 < \infty\,,
\end{aligned}$$

for all $t \in [0,1]$. Then, for all $t, x \in [0,1] \times \mathbb{R}^d$ there is a collection of random variables $Y_t : \Omega \to \mathbb{R}$ such that

$$Y_t \in L^1\,,$$

and

$$\left| \nabla \Phi_R(X_t) : \Pi_t(X_t) \right| \leq Y_t\,,$$

uniformly in $R > 0$.

Proof. Start by noticing that the cut-off function η_R can be chosen such that

$$\|\, \eta_R \,\|_{W^{1,\infty}} \leq C$$

for some constant $C > 0$ independent of R and so for all $x \in \mathbb{R}^d$

$$|\nabla \Phi_R(x)| \lesssim \|x\| \,.$$

Thus, for any $x \in \mathbb{R}^d$ we can write

$$\begin{aligned}
\Big|\nabla \Phi_R(x) : \Pi_t(x)\Big| &\leq \sum_{ij} \Big|\nabla \Phi_R(x)\Big| \, \Big|\, \Pi_t^{ij}(x) \,\Big| \\
&\lesssim \|x\| \sum_{ij} \Big|\, \Pi_t^{ij}(x) \,\Big| \\
&\lesssim \|x\|^2 + \Big(\sum_{ij} \Big|\, \Pi_t^{ij}(x) \,\Big| \Big)^2 \\
&\lesssim \|x\|^2 + \sum_{ij} \Big|\, \Pi_t^{ij}(x) \,\Big|^2
\end{aligned}$$

where in the third line we used the Cauchy-Schwartz inequality and in the third line we used Jensen's inequality, suppressing constants depending on the dimension d. Now we can set

$$Y_t = \|X_t\|^2 + \sum_{ij} \Big|\, \Pi_t^{ij}(X_t) \,\Big|^2$$

which is clearly in L^1 by the assumptions in the statement of the lemma. □

Theorem 6. *Under the assumptions*

$$\begin{aligned}
&\mathbb{E}\, \|X_t\|^2 < \infty \text{ and} \,, \\
&\mathbb{E}\, \|\Pi_t(X_t)\|^2 < \infty \,, \\
&\mathbb{E}\, \|a_t(X_t)\|^2 < \infty \,.
\end{aligned}$$

any process X satisfying (2) *satisfies*

$$-\mathbb{E}\Big[\, \mathrm{Tr}\, \Pi_t(X_t) \,\Big] = \mathbb{E}\, \Big[X_t \cdot \ddot{X}_t \Big] \tag{A.1}$$

Proof. The proof follows exactly the same steps as the proof of Theorem 5 with the additional application of the Lebesgue dominated convergence theorem in the quantity

$$\mathbb{E}\Big[\eta_R(X_t) \cdot a(X_t)\Big] = \mathbb{E}\Big[\eta_R(X_t) \cdot \ddot{X}_t\Big] \,,$$

which is justified since by the Cauchy-Schwartz inequality we have

$$\Big|\mathbb{E}\Big[\eta_R(X_t) \cdot a(X_t)\Big]\Big|^2 \leq \mathbb{E}\Big[\|a(X_t)\|^2\Big]\, \mathbb{E}\Big[\|X_t\|^2\Big] < \infty \,,$$

using a Cauchy-Schwartz inequality and $|\eta_R(X_t)| \leq |X_t|$ by the definition of η_R. □

Lemma 7 (Continuity Equation)**.** *We have the identity*

$$\partial_t\, \rho_t + \nabla \cdot (\rho_t\, v_t) = 0$$

for all $t \in [0, 1]$.

Proof. For $D^\circ \subset \mathbb{R}^d$, fix a test function $\varphi \in C_c^\infty(D)$ and compute

$$\begin{aligned}
\int_{\mathbb{R}^d} \varphi(x)\, \partial_t\, \rho_t(x)\, dx &= \partial_t \int_{\mathbb{R}^d} \varphi(x)\, d\rho_t(x) \\
&= \partial_t\, \mathbb{E}\, [\varphi(X_t)] \\
&= \mathbb{E}\, \Big[\nabla \varphi(X_t) \cdot \dot{X}_t\Big] \\
&= \mathbb{E}\Big[\nabla \varphi(X_t) \cdot v_t(X_t)\Big] \\
&= \int_{\mathbb{R}^d} \nabla \varphi(x) \cdot v_t(x)\, d\rho_t(x) \\
&= -\int_{\mathbb{R}^d} \varphi(x)\, \nabla \cdot (\rho_t(x)\, v_t(x))\; dx
\end{aligned}$$

where we used integration by parts, the fact that φ has compact support as well as properties of the conditional expectation. Since the above holds for all $\varphi \in C_c^\infty(\mathbb{R}^d)$ we get the desired result. $\square$