Neural fundamentals of optimal transport and its operator learning

Andrew Gracyk*

Fall 2025

Contents

1 Deriving the continuity equation

1.1 First derivation

Let $V \subseteq \mathbb{R}^d$ be a closed, connected, and bounded volume in Euclidean space and $\rho : \mathcal{V} \times [0,T] \to \mathbb{R}^+$ be a mass density. Let $v : \mathcal{V} \times [0,T] \to \mathbb{R}^d$ be a vector field. We have

$$\text{total mass} = \int_V \rho dV. \tag{1}$$

It is well known via flux theory in calculus that the change of mass through the surface is given by

$$\text{change of mass in the volume} = \int_{\partial V} \rho v \cdot n dS. \tag{2}$$

Since the mass is not changing in the volume, we get

$$\partial_t \int_V \rho dV = -\int_{\partial V} \rho v \cdot n dS. \tag{3}$$

Here, our surface integral is outward-oriented. We will assume basic regularity properties such as smoothness on ρ, and since our domain is sufficiently nice, we can exchange differentiation and integration. Hence,

$$\int_V \partial_t \rho dV + \int_{\partial V} \rho v \cdot n dS = 0. \tag{4}$$

We cannot combine the integrals into one equation yet due to the first being an integral over Euclidean space and the second a surface integral. Notice the criteria of the divergence theorem are satisfied (we will allow v to be sufficiently smooth), hence

$$\int_V \partial_t \rho dV + \int_V \nabla \cdot (\rho v) dV = \int_V (\partial_t \rho + \nabla \cdot (\rho v)) dV = 0. \tag{5}$$

Just because an integral is zero, that does not imply the integrand is zero. Moreover, ρ is not a (compactly supported) test function.

There are some ways to conclude the proof. For example, we have $\partial_t \rho + \nabla \cdot (\rho v)$ if we permit V to be arbitrary, and if the equation must hold for all V.

In the next slide, we will establish an argument in the sense of distributions.

*Purdue Mathematics; a thanks to Rognjie Lai for supervising this reading course

1.2 Second derivation

Let m be a vector-valued measure such that $dm = \rho dx$. Here, ρ is also vector-valued. Let ϕ be a C^∞ compactly supported test function. It is known by integration by parts

$$\int \phi d(\nabla \cdot m) = -\int \nabla\phi \cdot dm. \tag{6}$$

There is an equivalence

$$\int \phi(\nabla \cdot \rho)dx = -\int \nabla\phi \cdot \rho dx. \tag{7}$$

By the conservation of mass equation we saw on the previous slide, but in the weak form

$$\partial_t \int \phi\rho_s dx = -\int \phi(\nabla \cdot \rho_s v)dx. \tag{8}$$

Here, $\rho = \rho_s v$ is the vector field multiplied by scalar density.

It is a known fact from PDE theory

$$\partial_t \int \phi d\alpha - \int \nabla\phi \cdot v d\alpha = 0 = \partial_t \int \phi d\alpha - \int \phi\nabla \cdot v d\alpha = 0 \implies \partial_t \alpha + \nabla \cdot (v\alpha) = 0. \tag{9}$$

The second equality is due to integration by parts. This is because ϕ is any (smooth compactly supported) function. We will denote $d\alpha = p_s dx$. Thus,

$$\begin{aligned} 0 &= -\int \phi(\nabla \cdot \rho_s v)dx - \int \nabla\phi \cdot \rho dx \qquad (10)\\ &= \partial_t \int \phi\rho_s dx - \int \nabla\phi \cdot v\rho_s dx \qquad (11)\\ &\implies \partial_t \rho_s + \nabla \cdot (v\rho_s) = 0, \qquad (12)\end{aligned}$$

and we have the result.

2 JKO conditions with continuity equation equivalence

We show the JKO conditions as in [1] are equivalent to a continuity equation.

Consider the density under the flow map $\rho(t, T(t,x))$. By the chain rule,

$$\frac{d}{dt}\rho(t, T(t,x)) = \partial_t \rho + \nabla\rho \cdot \partial_t T(t,x). \tag{13}$$

We are aware that $\partial_t T = v$, so that gives

$$\frac{d}{dt}\rho = \partial_t \rho + \nabla\rho \cdot v. \tag{14}$$

Now, consider the following:

$$\begin{aligned} \frac{d}{dt}\rho(t, T(t,x)) &= \frac{d}{dt}\frac{\rho^k(x)}{\det|J(t,x)|} = -\frac{\rho^k(x)}{(\det|J(t,x)|)^2} \cdot \partial_t \det|J(t,x)| \qquad (15)\\ &= -\frac{\rho^k(x)}{\det|J(t,x)|} \cdot \frac{\partial_t \det|J(t,x)|}{\det|J(t,x)|} = -\frac{\rho^k(x)}{\det|J(t,x)|}\partial_t \log\det|J(t,x)| \qquad (16)\\ &= -\rho(t, T(t,x))\partial_t \log\det|J(t,x)| = -\rho(t, T(t,x))\mathrm{div}(v). \qquad (17)\end{aligned}$$

Thus,

$$-\rho(t, T(t,x))\mathrm{div}(v) = \partial_t \rho + \nabla\rho \cdot v \implies \partial_t \rho + \nabla \cdot (\rho v) = 0, \tag{18}$$

as desired.

References

[1] Wonjun Lee, Li Wang, and Wuchen Li. Deep jko: time-implicit particle methods for general nonlinear gradient flows, 2023. URL https://arxiv.org/abs/2311.06700.