The first-order ordinary differential equation (ODE)

$$\frac{d}{dt}x_t = v_t(x_t) \tag{1}$$

where $x_t \in \mathbb{R}^d$ for time $t \in [0, T]$, and where $v : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a time-dependent vector field that is generally non-linear. If we sample a random initial condition $x_0 \sim p_0$ from some distribution p_0, then the time-evolved random vector $x_t \sim p_t$ will also be a random variable with some probability density function (PDF) p_t, for all $t \in [0, T]$. By invoking the change of variable formula for PDFs on infinitesimal time steps of Eq. (1), one can show that p_t satisfies the *continuity equation*

$$\frac{\partial}{\partial t}p_t(x) = -\nabla \cdot \big[v_t(x)p_t(x)\big] \tag{2}$$

with respect to the velocity field v_t, for time $t \in [0, T]$.

Proof. Fix t and consider a small time increment $\varepsilon > 0$. By a first-order Taylor expansion of the ODE solution,

$$x_{t+\varepsilon} = x_t + \varepsilon v_t(x_t) + o(\varepsilon). \tag{3}$$

Define the deterministic map

$$T_\varepsilon(x) := x + \varepsilon v_t(x), \tag{4}$$

so that, to first order in ε, we have $x_{t+\varepsilon} = T_\varepsilon(x_t)$.
Since $x_{t+\varepsilon}$ is obtained by pushing forward x_t through T_ε, the change-of-variables formula for probability densities yields

$$p_{t+\varepsilon}(y) = p_t\big(T_\varepsilon^{-1}(y)\big)\left|\det \nabla T_\varepsilon^{-1}(y)\right|. \tag{5}$$

From the definition of T_ε, its inverse admits the expansion

$$T_\varepsilon^{-1}(y) = y - \varepsilon v_t(y) + o(\varepsilon). \tag{6}$$

Therefore (Taylor expansion):

$$p_t\big(T_\varepsilon^{-1}(y)\big) = p_t(y) - \varepsilon\nabla p_t(y) \cdot v_t(y) + o(\varepsilon). \tag{7}$$

The Jacobian matrix of T_ε is

$$\nabla T_\varepsilon(x) = I + \varepsilon\nabla v_t(x), \tag{8}$$

so that

$$\det \nabla T_\varepsilon(x) = 1 + \varepsilon\nabla \cdot v_t(x) + o(\varepsilon). \tag{9}$$

Consequently,

$$\det \nabla T_\varepsilon^{-1}(y) = 1 - \varepsilon\nabla \cdot v_t(y) + o(\varepsilon). \tag{10}$$

Substituting the above expansions into the change-of-variables formula, we obtain

$$\begin{aligned} p_{t+\varepsilon}(y) &= \Big(p_t(y) - \varepsilon\nabla p_t(y) \cdot v_t(y)\Big)\Big(1 - \varepsilon\nabla \cdot v_t(y)\Big) + o(\varepsilon) && (11)\\ &= p_t(y) - \varepsilon\Big[\nabla p_t(y) \cdot v_t(y) + p_t(y)\nabla \cdot v_t(y)\Big] + o(\varepsilon). && (12) \end{aligned}$$

Noting that

$$\nabla \cdot \big(v_t(y)p_t(y)\big) = \nabla p_t(y) \cdot v_t(y) + p_t(y)\nabla \cdot v_t(y), \tag{13}$$

we conclude that

$$p_{t+\varepsilon}(y) = p_t(y) - \varepsilon\nabla \cdot \big(v_t(y)p_t(y)\big) + o(\varepsilon). \tag{14}$$

Dividing both sides by ε and taking the limit $\varepsilon \to 0$ yields

$$\frac{\partial}{\partial t}p_t(y) = -\nabla \cdot \big(v_t(y)p_t(y)\big), \tag{15}$$

which is precisely the continuity equation. $\square$