5.3 The Divergence Theorem

Again let $N \geq 2$. Perhaps the single most striking application of the construction made in the second part of §5.2.2 is to multidimensional integration by parts formulas, and this section is devoted to the derivation of one of the most useful of these, the one known as the divergence theorem.

5.3.1 Flows Generated by Vector Fields

Let $V : \mathbb{R}^N \rightarrow \mathbb{R}^N$, and think of $V$ as a vector field that at each point prescribes the velocity of a particle passing through that point. To describe mathematically the trajectory of such a particle, consider the ordinary differential equation

$$\dot{\Phi}(t, x) = V(\Phi(t, x)) \quad \text{with} \quad \Phi(0, x) = x. \quad (5.3.1)$$

Assuming that $V$ is uniformly Lipschitz continuous, one knows that, for each $x \in \mathbb{R}^N$, there is precisely one solution to (5.3.1) and that that solution exists for all time, both in the future, $t \in [0, \infty)$, and the past, $t \in (-\infty, 0]$.

As a consequence of uniqueness, one also knows that $\Phi$ satisfies the flow property

$$\Phi(s + t, x) = \Phi(t, \Phi(s, x)) \quad \text{for} \quad (s, t, x) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^N. \quad (5.3.2)$$

In particular, $\Phi(t, \Phi(-t, x)) = x = \Phi(-t, \Phi(t, x))$, and so $\Phi(t, \cdot)$ is one-to-one and onto. Furthermore, if, in addition, $V$ is twice continuously differentiable, then so is $\Phi(t, \cdot)$, and, by the chain rule,

$$\frac{\partial \Phi(-t, \cdot)}{\partial x} \circ \Phi(t, x) \frac{\partial \Phi(t, x)}{\partial x} = \mathbf{I}.$$

Hence $\Phi(t, \cdot)$ is a diffeomorphism,

$$\frac{1}{J\Phi(t, x)} = J\Phi(-t, \Phi(t, x)),$$

and so, by Theorem 5.2.2,

$$\int f \circ \Phi(t, x) \, dx = \int f(y) J\Phi(-t, y) \, dy$$

for any non-negative, $\mathcal{B}_{\mathbb{R}^N}$-measurable $f$. Finally, starting from (5.3.1), one has

$$\frac{d}{dt} \frac{\partial \Phi(t, x)}{\partial x} = \frac{\partial V}{\partial y}(\Phi(t, x)) \frac{\partial \Phi(t, x)}{\partial x}, \quad (5.3.3)$$
and from this one can derive

\[ J\Phi(t, x) = \exp \left( \int_0^t \text{div} V(\Phi(\tau, x)) \, d\tau \right), \tag{5.3.4} \]

where \( \text{div}(V) \equiv \sum_{i=1}^N \partial_{y_i} V_i \) is the divergence of \( V \). To see how to pass from (5.3.3) to (5.3.4), one can use Cramer’s rule to verify that, for any \( N \times N \) matrix \( A = (a_{ij}) \), \( \partial a_{ij} \det(A) = A_{ij} \), where \( A_{ij} \) is the \( (i,j) \)th cofactor of \( A \). Hence, from (5.3.3) and the fact that \( \sum_{j=1}^N a_{kj} A_{ij} = \delta_{k,i} \det(A) \),

\[
\frac{d}{dt} \det(\Phi(t, x)) \frac{\partial}{\partial x} = \sum_{i,j,k=1}^N \frac{\partial V}{\partial y}(\Phi(t, x)) \frac{\partial}{\partial y} \left( \frac{\partial \Phi(t, x)}{\partial x} \right) A_{ij} \frac{\partial \Phi(t, x)}{\partial x} \frac{\partial \Phi(t, x)}{\partial x} \frac{\partial \Phi(t, x)}{\partial x},
\]

from which (5.3.4) is obvious.

After combining (5.3.4) with the preceding, we now know that

\[
\int f(\circ \Phi(t, x)) \, dx = \int f(x) \exp \left( - \int_0^t \text{div}(V)(\Phi(-\tau, x)) \, d\tau \right) \, dx \tag{5.3.5}
\]

for all non-negative, \( \mathcal{B}_{\mathbb{R}^N} \)-measurable \( f \) on \( \mathbb{R}^N \).

### 5.3.2 Mass Transport

I now want to apply (5.3.5) to measure the rate at which the flow generated by \( V \) moves mass in and out of an open set \( G \). To be precise, if one interprets \( \int 1_G(x) \, dx - \int 1_G \circ \Phi(t, x) \, dx \) as the net loss or gain due to the flow at time \( t \), then (5.3.5) says that \( \int_G \text{div}(V(x)) \, dx \) is the rate of loss or gain at time \( t = 0 \). On the other hand, there is another way in which to think about this computation. Namely,

\[
\int 1_G(x) \, dx - \int 1_G \circ \Phi(t, x) \, dx = \int_G 1_{G^c}(\Phi(t, x)) \, dx - \int_{G^c} 1_G(\Phi(t, x)) \, dx,
\]

which indicates that one should be able to do the same calculation by observing how much mass has moved in each direction across the boundary of \( G \) during the time interval \([0, t]\). To carry out this approach, I will assume that \( G \) is a non-empty, bounded open set that is a smooth region in the sense that for each \( p \in \partial G \) there exist an open neighborhood \( W \ni p \) and an
$F \in C^3(W; \mathbb{R})$ such that $|\nabla F| > 0$ and $G \cap W = \{x \in W : F(x) < 0\}$. In particular, $\partial G$ is a compact hypersurface and, for $x \in W \cap \partial G$, $\nabla F(x)$ is the outward pointing unit normal to $\partial G$ at $x$.

Let $(\bar{U}, \Psi)$ be a coordinate chart for some subset of $\partial G$, and, referring to Lemma ??, take $n(u) = \frac{\nabla F \circ \Psi(u)}{|\nabla F \circ \Psi(u)|}$ to be the outward point normal, and define $\tilde{\Psi}$ on $\bar{U} \times (-\rho, \rho)$ accordingly. Given $(u, \xi) \in \bar{U}$, $x = \tilde{\Psi}(u, \xi) \in G$ if and only if $\xi < 0$. Since $|\Phi(t, x) - x| \leq t\|V\|_u$, $\tilde{\Psi}(t, x) \in \tilde{\Psi}(\bar{U})$ if $|t| < T = \frac{\rho}{\|V\|_u}$. We can therefore define the map

$$t \in [0, T] \mapsto (u(t), \xi(t)) = (\tilde{\Psi})^{-1}(\Phi(t, x)) \in \bar{U},$$

in which case $\Phi(t, x) \notin G$ if and only if $\xi(t) \geq 0$.

Observe that, because $|\Psi(u) - x| \leq t\|V\|_u$ if $\Phi(t, x) \notin G$, $\Phi(t, x) - x = tV(\Psi(u)) + E_0(t, x)$, where $|E_0(t, x)| \leq C_0 t^2$ for some $C_0 < \infty$, and using this and the fact that $x = \Psi(u) + \xi n(\Psi(u))$, one sees that

$$\xi = \left(n(\Psi(u)), \Phi(t, x) - \Psi(u)\right)_{\mathbb{R}^N} - t\left(n(\Psi(u)), V(\Psi(u))\right)_{\mathbb{R}^N} - E_0(t, u).$$

Next write $\left(n(\Psi(u)), \Phi(t, x) - \Psi(u)\right)_{\mathbb{R}^N}$ as the sum

$$\left(n(\Psi(u)), \Phi(t, x) - \Psi(u)\right)_{\mathbb{R}^N} = \left(n(\Psi(u)), \Phi(u(t)) - \Psi(u)\right)_{\mathbb{R}^N} + \left(n(\Psi(u)), \Phi(u(t)) - \Psi(u)\right)_{\mathbb{R}^N},$$

$$\left(n(\Psi(u)), \Phi(t, x) - \Psi(u(t)) + \left(n(\Psi(u) - n(\Psi(u(t))), \Phi(t, x) - \Psi(u(t))\right)_{\mathbb{R}^N}.$$

Since the middle term equals $\xi(t)$, $\Phi(t, x) \notin G$ if and only if this term is non-negative. Further, it is clear that if $\Phi(t, x) \notin G$, then the absolute value of the last term is dominated by a constant times $t^2$. Finally, the first term vanishes at $t = 0$, and its derivative at $t = 0$ equals

$$\left(n(\Psi(u), v\right)_{\mathbb{R}^N} = 0$$

since $v = \left.\frac{d}{dt}\Psi(u(t))\right|_{t=0} \in T_{\Psi(u)}\partial G$. Hence, this term is also of order $t^2$, which means that

$$x \in G \& \Phi(t, x) \notin G \iff 0 \leq \xi \geq -t\left(n(\Psi(u)), V(\Psi(u))\right)_{\mathbb{R}^N} - E(t, u),$$

where $|E(t, u)| \leq Ct^2$ for some $C < \infty$.

Now let $\Gamma(t) = \{(u, \xi) \in \bar{U} : \xi < 0 \& \xi(t) \geq 0\}$. Then, since

$$(u, \xi) \in \Gamma(t) \iff u \in U \& 0 \geq \xi \geq -t\left(n(\Psi(u)), V(\Psi(u))\right)_{\mathbb{R}^N} - E(t, u),$$
\[
\frac{1}{t} \int_{G \setminus \hat{\Psi}(\Gamma(t))} 1_G \Phi(t, x) \lambda_{\mathbb{R}^N} \, dx = \frac{1}{t} \int \left( \int 1_{\Gamma(t)}(u, \xi) J\hat{\Psi}(u, \xi) \, d\xi \right) \, du
\]
\[
= \int_U \left( \mathbf{n}(\Psi(u)), V(\Psi(u)) \right)_{\mathbb{R}^N}^+ J\hat{\Psi}(u) \, du + O(t).
\]

Hence, after using an elementary covering argument, we see that
\[
\lim_{t \searrow 0} \frac{1}{t} \int_{G^c} 1_G \Phi(x) \, dx = \int_{\partial G} \left( \mathbf{n}(x), V(x) \right)_{\mathbb{R}^N}^+ \lambda_{\partial G} \, dx.
\]

By essentially the same argument, one can show that
\[
\lim_{t \searrow 0} \frac{1}{t} \int_{G^c} 1_G(x) \, dx = \int_{\partial G} \left( \mathbf{n}(x), V(x) \right)_{\mathbb{R}^N}^- \lambda_{\partial G} \, dx.
\]

Therefore, by combining these calculations with the earlier one, the one that was based on (5.3.5), we arrive at the following statement.

**Theorem 5.3.1 (Divergence Theorem)** Let \( G \) be a bounded, smooth region in \( \mathbb{R}^N \) and \( V : \mathbb{R}^N \rightarrow \mathbb{R}^N \) a twice continuously differentiable vector field with uniformly bounded first derivative. Then
\[
\int_G \text{div}(V)(x) \, dx = \int_{\partial G} \left( \mathbf{n}(x), V(x) \right)_{\mathbb{R}^N} \lambda_{\partial G} \, dx,
\]
where \( \mathbf{n}(x) \) is the outward pointing unit normal to \( \partial G \) at \( x \in \partial G \).

There are so many applications of Theorem 5.3.1 that it is hard to choose among them. However, here is one that is particularly useful. In its statement, \( L_V \) is the directional derivative operator \( \sum_{i=1}^N V_i \partial x_i \) determined by \( V \), and \( L_V^\top \) is the corresponding formal adjoint operator given by
\[
L_V^\top f = -\sum_{i=1}^N \partial x_i (fV_i) = -L_V f - f \text{div}(V).
\]

**Corollary 5.3.2** Referring to the preceding, one has
\[
\int_G fL_V g \, d\lambda_{\mathbb{R}^N} = \int_G gL_V^\top f \, d\lambda_{\mathbb{R}^N} + \int_{\partial G} f(g(\mathbf{n}, V))_{\mathbb{R}^N} \, d\lambda_{\partial G}
\]
for all \( f, g \in C^2_b(\mathbb{R}^N; \mathbb{R}) \).

**Proof.** Simply observe that \( \text{div}(fgV) = gL_V f - fL_V^\top g \), and apply Theorem 5.3.1 to the vector field \( fgV \). \( \square \)