Topics in Integration

John McCuan

March 19, 2021

Some of this material is contained in the calculus of variations notes, but this gives a self-contained treatment which is somewhat different and includes some additional topics. Here is an outline of what is here:

I. Integration

- **a.** Evaluation
  - **i.** Fubini's Theorem/Principle
  - **ii.** Change of Variables
- **b.** Balls and Spheres
- **c.** Convolution
- **d.** Mollification
- **e.** Fundamental Lemma

II. The Divergence

- **a.** Definition
- **b.** Product Rule
- **c.** First Variation Formula for PDEs

II. Preliminary Results (PDE; Laplace's Equation)

- **a.** Mean Value Property
- **b.** Strong Maximum Principle
- **c.** Higher Regularity

The immediate objective of these notes is to describe/get the **divergence theorem** and the **product rule** for the divergence of a scaled field in order to obtain the Euler-Lagrange PDE for C^2 extremals in several variables. Starting on the preliminaries of integration, however, it appeared that several other "side topics" could/should be covered along the way.

1 Integration and Scaling Factors

We have noted that the integral of a real valued function $f : A \to \mathbb{R}$ defined on a set A can often be obtained as a limit

$$\int_A f = \lim_{\|\mathcal{P}\|\to 0} \sum_j f(p_j^*)\mu(A_j)$$

where μ is a **measure** on subsets of A and $\mathcal{P} = \{A_j\}_j$ is a **partition** of A for which $A = \cup_j A_j$ consisting of finitely many "pieces" A_j having negligible intersection, i.e., $\mu(A_i \cap A_j) = 0$ when $i \neq j$, and

$$\|\mathcal{P}\| = \max_j \operatorname{diam}(A_j).$$

The evaluation points p_j^* are required to satisfy $p_j^* \in A_j$ but are otherwise arbitrary in the sense that there is a number $L \in \mathbb{R}$ such that given any $\epsilon > 0$ if $\mathcal{P} = \{A_j\}_j$ is any partition as above with $\|\mathcal{P}\| < \epsilon$, then

$$\left| \sum_j f(p_j^*)\mu(A_j) - L \right| < \epsilon$$

for every possible choice of evaluation points $p_j^* \in A_j$. In this event, we call the number L the **integral** of f over A and write

$$\int_A f = L.$$

Such is our definition of an integral. Let us declare that in this definition, when it makes sense, the set A may be called a **domain of integration**. We also recall the alternative notation

$$\int_{p \in A} f(p)$$

which can be useful when there are many symbols/variables appearing in the expression for the function f; The use of the integration variable p signals which variable is under consideration for the integration—the rest being considered fixed.

The question immediately arises: How is one to compute the value of such an integral over a set? Ultimately, everything must be reduced to the familiar integration from 1-D calculus over an interval on which our abstract integration is obviously modeled (or perhaps the Lebesgue generalization of that integration). In any case, there are two initial principles which are often used to facilitate this reduction.

1.1 Evaluation Principles

In many instances, our abstract integration is applied to subsets of $\mathbb{R}^n$ when $n \geq 2$ which have "full dimension" in $\mathbb{R}^n$, so to speak. It is in this setting that we state our first evaluation principle known, roughly, as **Fubini's theorem:**

Let Π be a hyperplane in $\mathbb{R}^n$ passing through a point $\mathbf{x}_0$ and having (unit) normal $\mathbf{n}$, say

$$\Pi = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{n} = 0\}.$$

Let us also assume $\mathcal{U} \subset \mathbb{R}^n$ is a domain of integration and $\underline{\mathcal{U}}$ is a domain of integration in Π such that

$$\mathcal{U} = \bigcup_{\mathbf{x} \in \underline{\mathcal{U}}} I_{\mathbf{x}} \tag{1}$$

where $I_{\mathbf{x}}$ is a segment with

$$\overline{I_{\mathbf{x}}} = \{\mathbf{x} + t\mathbf{n} : a_{\mathbf{x}} \leq t \leq b_{\mathbf{x}}\}.$$

In this situation $\int_{\mathcal{U}} f$ can be written as an **iterated integral**

$$\int_{\mathcal{U}} f = \int_{\mathbf{x} \in \underline{\mathcal{U}}} \int_{I_{\mathbf{x}}} f\big|_{I_{\mathbf{x}}}. \tag{2}$$

In many applications of Fubini's principle the secondary domain of integration $\underline{\mathcal{U}}$ is a subset of a coordinate hyperplane.

Example 1 *(Boas 5.2.32)*

$$\mathcal{U} = \{(x, y) : y^2 < x < 1,\ 0 < y < 1\}$$

with $f : \mathcal{U} \to \mathbb{R}$ *by*

$$f(x, y) = \frac{e^x}{\sqrt{x}}.$$

In this case, we can set $I_y = \{x : y^2 < x < 1\}$ *for* $0 < y < 1$ *and write*

$$\int_{\mathcal{U}} f = \int_{\mathcal{U}} \frac{e^x}{\sqrt{x}} = \int_{y\in(0,1)} \int_{x\in I_y} \frac{e^x}{\sqrt{x}}$$
$$= \int_0^1 \left(\int_{y^2}^1 \frac{e^x}{\sqrt{x}}\, dx \right) dy.$$

The inner integral, it will be observed, cannot be evaluated in terms of elementary functions, but the entire area integral $\int_{\mathcal{U}} f$ *can. This is accomplished by using Fubini's theorem in a different way: Taking* $J_x = \{y : 0 < y < \sqrt{x}\}$ *for* $0 < x < 1$*, we have*

$$\begin{aligned}
\int_{\mathcal{U}} f &= \int_{x\in(0,1)} \int_{y\in J_x} \frac{e^x}{\sqrt{x}} \\
&= \int_0^1 \left(\int_0^{\sqrt{x}} \frac{e^x}{\sqrt{x}}\, dy \right) dx \\
&= \int_0^1 \frac{e^x}{\sqrt{x}} \left(\int_0^{\sqrt{x}} 1\, dy \right) dx \\
&= \int_0^1 e^x\, dx \\
&= e.
\end{aligned}$$

What is $\underline{\mathcal{U}}$*?*

Example 2 *(Boas 5.2.50) Here*

$$\mathcal{U} = \{(x, y, z) : y^2 + z^2 < x^2 < 4,\ x, y, z > 0\}$$

is part of the inside of a cone with $f : \mathcal{U} \to \mathbb{R}$ *by*

$$f(x, y, z) = z.$$

Setting $\underline{\mathcal{U}} = \{\zeta = (y, z) : y, z > 0,\ y^2 + z^2 \le 16\}$*, that is, a quarter disk in the* y, z*-coordinate plane, we have*

$$I_\zeta = \{x : \sqrt{y^2 + z^2} < x < 4\},$$

and from Fubini's principle

$$\begin{aligned}
\int_{\mathcal{U}} z &= \int_{\zeta=(y,z)\in\underline{\mathcal{U}}} \int_{x\in I_\zeta} z \\
&= \int_{\zeta=(y,z)\in\underline{\mathcal{U}}} \left(\int_{\sqrt{y^2+z^2}}^4 z\, dx \right) \\
&= \int_{\zeta=(y,z)\in\underline{\mathcal{U}}} z \left(\int_{\sqrt{y^2+z^2}}^4 1\, dx \right) \\
&= \int_{\zeta=(y,z)\in\underline{\mathcal{U}}} z \left(4 - \sqrt{y^2 + z^2} \right).
\end{aligned}$$

Applying Fubini's theorem to the domain of integration $\underline{\mathcal{U}} \subset \mathbb{R}^2$, we continue our evaluation:

$$\begin{aligned}
\int_{\mathcal{U}} z &= \int_0^4 \int_0^{\sqrt{16-y^2}} \left(4z - z\sqrt{y^2+z^2}\right) dz\, dy \\
&= \int_0^4 \left(2(16-y^2) - \frac{1}{3}(y^2+z^2)^{3/2} \Big|_{z=0}^{\sqrt{16-y^2}} \right) dy \\
&= \int_0^4 \left(32 - 2y^2 - \frac{1}{3}(16)^{3/2} + \frac{1}{3}y^3 \right) dy \\
&= \int_0^4 \left(\frac{32}{3} - 2y^2 + \frac{1}{3}y^3 \right) dy.
\end{aligned}$$

Another common special case of Fubini's theorem (some instances of which you may not have seen) is when $\mathcal{U}$ is a cross-product:

Example 3 *Say $h > 0$ and*

$$\mathcal{C} = \{(x,y,z) : x^2 + y^2 = 1,\ 0 < z < h\} = \partial B_1(\mathbf{0}) \times (0,h)$$

is a cylinder where $B_1(\mathbf{0}) = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. The cylindrical surface $\mathcal{C}$ is a valid domain of integration having the form required by Fubini's theorem with $\underline{\mathcal{U}} = \partial B_1(\mathbf{0})$, and

$$\int_{\mathcal{C}} f = \int_{(x,y)\in\partial B_1(\mathbf{0})} \left(\int_0^h f(x,y,z)\, dz \right).$$

In order to complete the evaluation in the last example, one needs to know how to integrate on a circle.

Change of Variables and Scaling Factors

In the last example we reduced an integral over a cylinder to an iterated integral involving an integral over a circle. Say we want to compute

$$\int_{\partial B_r(\mathbf{0})} f$$

where $\partial B_r(\mathbf{0}) = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = r^2\}$. One way to evaluate such an integral is to use the **parameterization**

$$\psi : [0, 2\pi] \to \partial B_r(\mathbf{0}) \qquad \text{by} \qquad \psi(\theta) = (r\cos\theta, r\sin\theta).$$

Then we can write

$$\int_{\partial B_r(\mathbf{0})} f = \int_0^{2\pi} f \circ \psi(\theta)\, \sigma\, d\theta$$

where σ is an appropriate **scaling factor**. This construction applies in rather great generality.

The Change of Variables Principle: Assume $\psi : \mathcal{A} \to \mathcal{B}$ is a parameterization of one domain of integration $\mathcal{B}$ with measure μ on another domain of integration $\mathcal{A}$ with measure ν. Assume also a **scaling factor**

$$\sigma : \mathcal{A} \to [0, \infty) \qquad \text{exists with} \qquad \sigma(p) = \lim_{\mathcal{A}_j \to \{p\}} \frac{\mu(\psi(\mathcal{A}_j))}{\nu(\mathcal{A}_j)}$$

where $\mathcal{A}_j$ is a typical "piece" of the domain of integration $\mathcal{A}$ and

$$\psi(\mathcal{A}_j) = \{\psi(x) : x \in \mathcal{A}_j\}$$

is the image of the piece $\mathcal{A}_j$ of $\mathcal{A}$. Then

$$\int_{\mathcal{B}} f = \int_{\mathcal{A}} f \circ \psi \, \sigma. \tag{3}$$

In order to use the change of variables principle one must determine the scaling factor in each particular case, but two common situations give us a good start.

1. If $\mathcal{B}, \mathcal{A} \subset \mathbb{R}^n$ are both "full dimension" sets in $\mathbb{R}^n$ and ψ is differentiable, then

   $$\sigma = |\det D\psi|$$

   where $D\psi$ is the (square $n \times n$) matrix of first partial derivatives of the components of ψ:

   $$D\psi = \left(\frac{\partial \psi_i}{\partial x_j} \right)_{i,j}.$$

   In this case, σ is often called the **Jacobian scaling factor**.

2. If $\mathcal{A} \subset \mathbb{R}^k$ and $\mathcal{B} \subset \mathbb{R}^n$ with $n > k$ and $D\psi$ is the $n \times k$ matrix of first partials of the components of ψ, then

   $$\sigma = \sqrt{\det \left(D\psi^T D\psi \right)}$$

   where $D\psi^T$ is the transpose of $D\psi$. Note that $D\psi^T D\psi$ will be a square ($k \times k$ symmetric) matrix with non-negative determinant.

The second scaling factor is what we need for integration on a circle:

$$\psi(\theta) = (r\cos\theta, r\sin\theta) \qquad (k = 1, n = 2).$$

$$D\psi = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \end{pmatrix} \qquad \text{and} \qquad D\psi^T = (-r\sin\theta, r\cos\theta),$$

so

$$D\psi^T D\psi = r^2 \qquad \text{and} \qquad \sigma = r.$$

Exercise 1 *Use the first scaling factor rule to derive the scaling factor for polar coordinates in $\mathbb{R}^2$, and cylindrical and spherical coordinates in $\mathbb{R}^3$.*

As an example of the first scaling factor formula, consider $\psi : B_a(\mathbf{p}) \to B_b(\mathbf{q})$ by

$$\psi(\mathbf{x}) = \mathbf{q} + b(\mathbf{x} - \mathbf{p})/a \tag{4}$$

where $\mathbf{p}, \mathbf{q} \in \mathbb{R}^2$ and $a, b > 0$. Then the scaling factor associated with ψ is

$$\sigma = \lim_{\mathcal{A}_j \to \{\mathbf{x}\}} \frac{\text{area}(\psi(\mathcal{A}_j))}{\text{area}(\mathcal{A}_j)} = \left(\frac{b}{a}\right)^2$$

and

$$\int_{B_b(\mathbf{q})} f = \int_{B_a(\mathbf{p})} f \circ \psi \, \left(\frac{b}{a}\right)^2.$$

Exercise 2 *Verify the scaling factor we have given for the mapping (4) in $\mathbb{R}^2$ and determine the scaling factor for the mapping with the same formula mapping one ball to another in $\mathbb{R}^n$.*

1.2 Balls and Spheres

An important and useful situation in which the scaling factor for a change of variables does not arise immediately from the formulas above is when changing variables from one sphere to another. By **sphere** here we mean the boundary of a ball which may occur in any dimension. The boundary of $B_r(\mathbf{p}) \subset \mathbb{R}^2$ is a circle; the boundary of $B_r(\mathbf{p}) \subset \mathbb{R}^3$ is what we usually call a sphere, and the boundary of $B_r(\mathbf{p}) \subset \mathbb{R}^n$ when $n \geq 4$ is often called a hypersphere, but we are calling all of them spheres. In general, we say $B_r(\mathbf{p}) \subset \mathbb{R}^n$ is an n-dimensional submanifold of $\mathbb{R}^n$, or that it has full dimension. The boundary $\partial B_r(\mathbf{p}) \subset \mathbb{R}^n$ is an $(n-1)$-dimensional submanifold of $\mathbb{R}^n$ or a hypersurface.

Consider $\psi : \partial B_a(\mathbf{p}) \to \partial B_b(\mathbf{q})$ given by formula (4) with $B_a(\mathbf{p}), B_b(\mathbf{q}) \subset \mathbb{R}^2$. The scaling factor for this parameterization of one circle over another is

$$\sigma = \frac{b}{a} \qquad \text{and} \qquad \int_{\partial B_b(\mathbf{q})} f = \int_{\partial B_a(\mathbf{p})} f \circ \psi \, \frac{b}{a}.$$

To see this, we can parameterize each domain of integration on $[0, 2\pi]$ and apply the **principle of composition for scaling factors**:

If $\psi_1 : [0, 2\pi] \to \partial B_a(\mathbf{p})$ by $\psi_1(\theta) = a(\cos\theta, \sin\theta)$ with scaling factor $\sigma_1 = a$, and $\psi_2 : [0, 2\pi] \to \partial B_b(\mathbf{q})$ by $\psi_2(\theta) = b(\cos\theta, \sin\theta)$ with scaling factor $\sigma_2 = b$, then

$$\sigma_1 \sigma = \sigma_2.$$

That is, $a\sigma = b$ or $\sigma = b/a$.

If $B_a(\mathbf{p}), B_b(\mathbf{q}) \subset \mathbb{R}^n$, then formula (4) still gives a parameterization $\psi : \partial B_a(\mathbf{p}) \to \partial B_b(\mathbf{p})$ and

$$\int_{\partial B_b(\mathbf{q})} f = \int_{\partial B_a(\mathbf{p})} f \circ \psi \, \left(\frac{b}{a}\right)^{n-1}. \tag{5}$$

Exercise 3 *Give a careful justification for (5) using the principle of composition in the case* $n = 3$.

There is one more integration principle I want to discuss. A ball $B_a(\mathbf{p})$, as a domain of integration,[1] can be decomposed into a disjoint union of (hyper)spheres $\partial B_r(\mathbf{p})$ for $0 < r < a$. In particular,

$$\int_{B_a(\mathbf{p})} f = \int_0^a \left(\int_{\partial B_r(\mathbf{p})} f \right) dr. \tag{6}$$

Again, this formula does not follow immediately from the principles above, but we can use Fubini's theorem together with the principles of scaling/change of variables to verify (6). First parameterize $B_a(\mathbf{p})$ on the product $(0, a) \times \partial B_1(\mathbf{0})$ by

$$\psi_1(r, \mathbf{x}) = \mathbf{p} + r\mathbf{x}.$$

Note that

$$D\psi_1 = \begin{pmatrix} 1 & \mathbf{0}_{n-1} \\ \mathbf{0}_{n-1}^T & rI_{(n-1)\times(n-1)} \end{pmatrix}.$$

Here, we have used block matrix notation and denoted the zero (row) vector in $\mathbb{R}_k$ by $\mathbf{0}_k$ and the $k \times k$ identity matrix by $I_{k\times k}$. It follows that the scaling factor for ψ_1 is

$$\sigma_1 = |\det D\psi_1| = r^{n-1}.$$

[1] Notice I've left out the values $r = 0$ and $r = a$ giving the ball as a union of subsets. This is a valid decomposition *as a domain of integration* because the (full dimension) measures of the center point $\{\mathbf{p}\}$ and the outer boundary $\partial B_a(\mathbf{p})$ are zero. We could omit any finite or countable number of such measure zero sets and not effect the integral. The same idea has been used above in regard to the negligible intersection of pieces $\mathcal{A}_j$ in the definition of integration and in the appearance of the closure of the intervals $I_{\mathbf{x}}$ in Fubini's theorem.

Therefore, by Fubini's theorem

$$\int_{B_a(\mathbf{p})} f = \int_0^a \left(\int_{\partial B_1(\mathbf{0})} f \circ \psi_1 \; r^{n-1} \right) dr = \int_0^a r^{n-1} \left(\int_{\partial B_1(\mathbf{0})} f \circ \psi_1 \right) dr.$$

In the inner integral $\psi_2 : \partial B_r(\mathbf{p}) \to \partial B_1(\mathbf{0})$ by

$$\psi_2(\mathbf{x}) = (\mathbf{x} - \mathbf{p})/r$$

is a valid change of variables, as discussed above, with scaling factor $\sigma_2 = 1/r^{n-1}$. Thus,

$$\int_{\partial B_1(\mathbf{0})} f \circ \psi_1 = \int_{\partial B_r(\mathbf{p})} f \circ \psi(r, \psi_2) \, \frac{1}{r^{n-1}} = \frac{1}{r^{n-1}} \int_{\mathbf{x} \in \partial B_r(\mathbf{p})} f \circ \psi(r, \psi_2(\mathbf{x})).$$

Since $\psi_1(r, \psi_2(\mathbf{x})) = \mathbf{p} + r(\mathbf{x} - \mathbf{p})/r = \mathbf{x}$,

$$\int_{\partial B_1(\mathbf{0})} f \circ \psi_1 = \frac{1}{r^{n-1}} \int_{\partial B_r(\mathbf{p})} f,$$

and we have established (6).

Measures and Averages

This is a convenient place to give some notation for the measures of balls and spheres. We write

$$\omega_n = \int_{B_1(\mathbf{0})} 1$$

for the n-dimensional measure of the unit ball $B_1(\mathbf{0}) \subset \mathbb{R}^n$. That is, $\omega_1 = 2$, $\omega_2 = \pi$, $\omega_3 = 4\pi/3$, and so on.

Exercise 4 *Show the $n-1$-dimensional measure of $\partial B_1(\mathbf{0}) \subset \mathbb{R}^n$ is $n\omega_n$.*

Scaling then shows that for $B_r(\mathbf{p}) \subset \mathbb{R}^n$,

$$\mu(B_r(\mathbf{p})) = \omega_n r^n \qquad \text{and} \qquad \nu(\partial B_r(\mathbf{p})) = n\omega_n r^{n-1}.$$

Here μ denotes the n-dimensional measure on $\mathbb{R}^n$ denoted by either $\mathcal{L}^n$ (full dimension **Lebesgue measure**) or $\mathcal{H}^n$ (n-dimensional **Hausdorff measure** in $\mathbb{R}^n$). The measure ν is $\mathcal{H}^{n-1}$, the $(n-1)$-dimensional Hausdorff measure on subsets of $\mathbb{R}^n$. For $n = 1$, $\mathcal{L}^1 = \mathcal{L} = \mathcal{H}^1$ is **length measure** and $\mathcal{H}^0$ is **counting measure**. For $n = 2$, $\mathcal{L}^2 = \mathcal{H}^2 = \text{area}$ is **area measure** and $\mathcal{H}^1$ is **length measure** in $\mathbb{R}^2$. For $n = 3$, $\mathcal{L}^3 = \mathcal{H}^3 = \text{vol}$ is **volume measure** and $\mathcal{H}^2$ is **area measure** in $\mathbb{R}^3$. This description continues for all dimensions and, in short, there are a lot of different measures.

The average value of a function $f : B_r(\mathbf{p}) \to \mathbb{R}$ is given by

$$\frac{1}{\omega_n r^n} \int_{B_r(\mathbf{p})} f,$$

and more generally the average of any function f over a domain of integration $\mathcal{A}$ with measure μ is

$$\frac{1}{\mu \mathcal{A}} \int_{\mathcal{A}} f.$$

One simple version of the **intermediate value theorem for integrals of continuous functions** is the following:

Theorem 1 *If $u \in C^0(\overline{\mathcal{U}})$ where $\mathcal{U}$ is a bounded open subset of $\mathbb{R}^n$, then the numbers*

$$m = \min\{u(\mathbf{x}) : \mathbf{x} \in \overline{\mathcal{U}}\} \qquad \text{and} \qquad M = \max\{u(\mathbf{x}) : \mathbf{x} \in \overline{\mathcal{U}}\}$$

are well-defined and there exists some $\mathbf{x}_* \in \mathcal{U}$ with

$$m \le u(\mathbf{x}_*) = \frac{1}{\mu\mathcal{U}} \int_{\mathcal{U}} u \le M$$

with strict inequality unless $u \equiv m = M$ is constant.

We will prove a special case of this result below.

1.3 Convolution

If $u, v \in C_c^0(\mathbb{R}^n)$, then $u * v : \mathbb{R}^n \to \mathbb{R}$ by

$$(u * v)(\mathbf{p}) = \int_{\mathbf{x}\in\mathbb{R}^n} u(\mathbf{x})\, v(\mathbf{p} - \mathbf{x})$$

is called the **convolution** of u and v.

Exercise 5 *Show that $u * v \in C_c^0(\mathbb{R}^n)$.*

Exercise 6 *Convolution is commutative in the sense that $u * v = v * u$.*

Exercise 7 *A function $u : \mathcal{U} \to \mathbb{R}$ defined on an open subset of $\mathbb{R}^n$ is said to be* **locally integrable** *if*

$$\int_K |u| < \infty$$

whenever $K \subset\subset \mathcal{U}$. The collection of all locally integrable functions on $\mathcal{U}$ is denoted by $L^1_{loc}(\mathcal{U})$.

Show that if $u \in C_c^0(\mathbb{R}^n)$ and $v \in L^1_{loc}(\mathbb{R}^n)$, then $u * v$ and $v * u$ given by the formula above are both well-defined and

1. *$u * v = v * u$, but*
2. *it may not be the case that $u * v$ has compact support. (Show this by example.)*

Exercise 8 *If $u \in C^0(\mathbb{R}^n)$ and $\phi \in C_c^\infty(\mathbb{R}^n)$, then $\phi * u \in C^\infty(\mathbb{R}^n)$.*

1.4 Mollification

We begin with some special functions in $C_c^\infty(\mathbb{R}^n)$. Let $\phi_0 \in C_c^\infty(\mathbb{R}^n)$ by

$$\phi_0(\mathbf{x}) = \begin{cases} e^{-1/(1-|\mathbf{x}|^2)}, & |\mathbf{x}| < 1 \\ 0, & |\mathbf{x}| \ge 1. \end{cases}$$

Next we define $\phi_1 \in C_c^\infty(\mathbb{R}^n)$ by

$$\phi_1(\mathbf{x}) = \frac{\phi_0(\mathbf{x})}{\int \phi_0}.$$

For each $\epsilon > 0$ define $\mu_\epsilon \in C_c^\infty(\mathbb{R}^n)$ by

$$\mu_\epsilon(\mathbf{x}) = \frac{1}{\epsilon^n}\phi_1\left(\frac{\mathbf{x}}{\epsilon}\right).$$

Exercise 9 *Show the following:*

(a) $\operatorname{supp} \mu_\epsilon = \overline{B_\epsilon(\mathbf{0})}$,

(b) $\int \mu_\epsilon = 1$, *and*

(c) *If we denote by* $\psi \in C_c^\infty(\mathbb{R})$ *the even function* ϕ_1 *when* $n = 1$*, then what is the relation between* μ_ϵ *and*

$$\nu_\epsilon(\mathbf{x}) = \frac{1}{\epsilon^n}\psi\left(\frac{|\mathbf{x}|}{\epsilon}\right)?$$

The functions $\{\mu_\epsilon\}_{\epsilon>0}$ are called an **approximate identity** and each function μ_ϵ is called a **standard mollifier**.

Theorem 2 *If* $v \in L^1_{loc}(\mathbb{R}^n)$*, then* $\mu_\epsilon * v : \mathbb{R}^n \to \mathbb{R}$ *by*

$$(\mu_\epsilon * v)(\mathbf{q}) = \int_{\mathbf{x}\in B_\epsilon(\mathbf{q})} \mu_\epsilon(\mathbf{q}-\mathbf{x})\, v(\mathbf{x}) = \int_{\mathbf{x}\in B_\epsilon(\mathbf{0})} \mu_\epsilon(\mathbf{x})\, v(\mathbf{q}-\mathbf{x})$$

satisfies $\mu_\epsilon * v \in C^\infty(\mathbb{R}^n)$.

Proof: For any multi-index β,

$$D^\beta(\mu_\epsilon * v)(\mathbf{q}) = \int D^\beta \mu_\epsilon(\mathbf{q}-\mathbf{x})\, v(\mathbf{x}).$$

That is, $D^\beta(\mu_\epsilon * v) = D^\beta \mu_\epsilon * v$. $\square$

If one wanted to fill in the details of this proof (by integrating under the integral sign) one could use induction and something like the **dominated convergence theorem**. Here is roughly how the first step of that would go: We consider a difference quotient

$$\frac{(\mu_\epsilon * v)(\mathbf{q}+h\mathbf{e}_j) - (\mu_\epsilon * v)(\mathbf{q})}{h} = \frac{1}{h}\int_{\mathbf{x}\in B_\epsilon(\mathbf{q})} [\mu_\epsilon(\mathbf{q}-\mathbf{x}+h\mathbf{e}_j) - \mu_\epsilon(\mathbf{q}-\mathbf{x})]\, v(\mathbf{x})$$
$$= \int_{\mathbf{x}\in B_\epsilon(\mathbf{q})} \left[\frac{\mu_\epsilon(\mathbf{q}-\mathbf{x}+h\mathbf{e}_j) - \mu_\epsilon(\mathbf{q}-\mathbf{x})}{h}\right] v(\mathbf{x}).$$

Notice that the pointwise limits as $h \to 0$ of the inner difference quotients are

$$\lim_{h\to 0} \frac{\mu_\epsilon(\mathbf{q}-\mathbf{x}+h\mathbf{e}_j) - \mu_\epsilon(\mathbf{q}-\mathbf{x})}{h} = \frac{\partial \mu_\epsilon}{\partial x_j}(\mathbf{q}-\mathbf{x}).$$

It is not difficult to believe, or even prove, that the convergence in this limit is **uniform** in $\mathbf{x} \in \mathbb{R}^n$. In particular, one can find a uniform bound C for which the functions $w : \mathbb{R}^n \to \mathbb{R}$ by

$$w(\mathbf{x}) = w(\mathbf{x}; h) = \left[\frac{\mu_\epsilon(\mathbf{q}-\mathbf{x}+h\mathbf{e}_j) - \mu_\epsilon(\mathbf{q}-\mathbf{x})}{h}\right] v(\mathbf{x})$$

are uniformly integrable on the compact set $K = \overline{B_\epsilon(\mathbf{q})}$ with

$$\int_K |w| \le C \int_K |v|.$$

This is essentially what is needed to apply the dominated convergence theorem to conclude

$$\lim_{h\to 0}\int_{\mathbf{x}\in K} w(\mathbf{x}; h) = \int_{\mathbf{x}\in K} \lim_{h\to 0} w(\mathbf{x}; h).$$

It can be shown that the function $\mu_\epsilon * v$ converges to the function v in various senses, depending on the regularity/integrability of the function v, as $\epsilon \searrow 0$. This is the origin of the term **approximate identity**: $\mu_\epsilon * v$ is a function which approximates v and converges to v as $\epsilon \searrow 0$.

Exercise 10 *If $v \in C^0(\mathbb{R}^n)$, then $\mu_\epsilon * v$ converges to v uniformly on compact subsets, that is given any $K \subset\subset \mathbb{R}^n$*

$$lim_{\epsilon \searrow 0} \max\{|(\mu_\epsilon * v(\mathbf{x}) - v(\mathbf{x})| : \mathbf{x} \in K\} = 0.$$

1.5 Fundamental Lemma

This is a convenient place to state the fundamental lemma of the calculus of variations for functions of several variables:

Lemma 1 *If $\mathcal{U} \subset \mathbb{R}^n$ is an open set, $f \in C^0(\mathcal{U})$, and*

$$\int_{\mathcal{U}} f\phi = 0 \qquad \text{for every } \phi \in C_c^\infty(\mathcal{U}),$$

then $f(\mathbf{x}) = 0$ for every $\mathbf{x} \in \mathcal{U}$.

This should seem quite natural, and you should be able to prove it without too much trouble. What is rather more difficult, but perhaps interesting for you to know, is the following version:

Lemma 2 *If $\mathcal{U} \subset \mathbb{R}^n$ is an open set, $v \in L^1_{loc}(\mathcal{U})$, and*

$$\int_{\mathcal{U}} v\phi = 0 \qquad \text{for every } \phi \in C_c^\infty(\mathcal{U}),$$

then $\{\mathbf{x} \in v(\mathbf{x}) \neq 0\}$ has **measure zero**.

A set $A \subset \mathbb{R}^n$ has **measure zero** if for any $\epsilon > 0$, there exists a sequence of open sets $U_1, U_2, U_3, \ldots$ with

$$A \subset \bigcup_{j=1}^{\infty} U_j \tag{7}$$

measures satisfying

$$\sum_{j=1}^{\infty} \mu U_j < \epsilon.$$

Incidentally, any time you have a collection of open sets containing a set A in their union, as in (7), the collection of open sets is called an **open cover** of A. You do not have to have countably many (a sequence) to have an open cover. For example, the set of all balls

$$\{B_r(\mathbf{p}) : r > 0 \text{ and } \mathbf{p} \in \mathbb{R}^n\}$$

is an open cover of $\mathbb{R}^n$.

Also, a function $v \in L^1_{loc}(\mathcal{U})$ satisfying the conclusion of the fundamental lemma is considered to be "the constant zero function" in $L^1_{loc}(\mathcal{U})$. In particular, if $\{\mathbf{x} : v(\mathbf{x}) \neq 0\}$ has measure zero, then

$$\int vw = 0$$

for every function w (assuming integrating the product vw makes sense). Thus, you can't tell any difference between v and the zero function in terms of integration. For example, $v : \mathbb{R} \to \mathbb{R}$ by

$$v(\mathbf{x}) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & \text{otherwise,} \end{cases}$$

is[2] the zero function in $L^1(\mathbb{R})$, even though it technicaly differs from the constant zero function at infinitely many points.

Exercise 11 *Show $\mathbb{Q} \subset \mathbb{R}$ has measure zero.*

[2]Note that $\mathbb{Q} = \{p/q : p \in \mathbb{Z}, q \in \mathbb{N}\}$ is the set of rational numbers; $\mathbb{N} = \{1, 2, 3, \ldots\}$ is the set of natural numbers and $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$ is the set of integers.

2 The Divergence

You may be familiar with the divergence operator div : $C^1(\mathcal{U} \to \mathbb{R}^n) \to C^0(\mathcal{U})$ given by

$$\operatorname{div} \mathbf{v} = \sum_{j=1}^{n} \frac{\partial v_j}{\partial x_j} = \nabla \cdot \mathbf{v}$$

where $\mathcal{U}$ is an open subset of $\mathbb{R}^n$ and $C^1(\mathcal{U} \to \mathbb{R}^n)$ denotes the collection of C^1 **vector fields** on $\mathcal{U}$. This is the definition of the divergence with respect to standard rectangular coordinates; notice the appearance of partial derivatives. It will be useful to understand the divergence in a coordinate free manner. For one thing, this will allow us to consider the divergence on more general domains of integration like surfaces.

2.1 Definition

Let $\mathcal{A}$ be a domain of integration with measure μ. We consider a subdomain $\mathcal{A}_0$ of $\mathcal{A}$ having the properties of a "piece" in the definition but let us also require $\mathcal{A}_0$ to be a particularly nice piece of $\mathcal{A}$ in the sense that $\partial\mathcal{A}_0$ is also a domain of integration having a well-defined unit normal $\mathbf{n} = \mathbf{n}(\mathbf{x})$ pointing out of $\mathcal{A}_0$ at $\mathbf{x} \in \partial\mathcal{A}_0$. You can imagine this situation with a ball $B_a(\mathbf{p}) \subset \mathbb{R}^n$, but you could also imagine this with a well-behaved piece of a surface. (On a surface, the outward pointing normal $\mathbf{n}$ would be tangent to the surface.)

We also need the domain of integration $\mathcal{A}$ to have enough structure to make sense of the idea of a C^1 vector field on $\mathcal{A}$. If $\mathcal{A}$ happens to be a full dimension open subset of $\mathbb{R}^n$, then this is easy, we just mean $C^1(\mathcal{A} \to \mathbb{R}^n)$ is the set of C^1 vector fields. If $\mathcal{A}$ is a surface, or a k-dimensional submanifold of $\mathbb{R}^n$, then there is also such a collection. In all cases where it makes sense, let us denote the collection of C^1 vector fields on $\mathcal{A}$ by $C^1(\mathcal{A} \to \mathcal{V})$ where $\mathcal{V}$ is some appropriate co-domain to "catch" the collection of all vectors tangent to $\mathcal{A}$.

Taking a vector field $\mathbf{v} \in C^1(\mathcal{A} \to \mathcal{V})$, we can form the **flux integral**

$$\int_{\partial\mathcal{A}_0} \mathbf{v} \cdot \mathbf{n}.$$

This measures, in a certain sense, the total "flow" out of $\mathcal{A}_0$ determined by $\mathbf{v}$. The divergence is an infinitesimal measure of this flow:

$$\operatorname{div}_{\mathcal{A}} \mathbf{v}(\mathbf{x}) = \lim_{\mathcal{A}_0 \to \{\mathbf{x}\}} \frac{1}{\mu(\mathcal{A}_0)} \int_{\partial\mathcal{A}_0} \mathbf{v} \cdot \mathbf{n}.$$

Notice the scaling of the integral here: $\mu(\mathcal{A}_0)$ is a full-dimension measure of the set $\mathcal{A}_0$, though the measure used to compute the integral is a lower dimensional measure defined with respect to $\partial\mathcal{A}_0$.

Exercise 12 *Let $\mathcal{U}$ be an open subset of $\mathbb{R}^3$ with $\mathbf{p} = (p_1, p_2, p_3) \in \mathcal{U}$. Consider the "epsilon cube" with center $\mathbf{p}$ given by*

$$\mathcal{U}_\epsilon = \mathcal{U}_\epsilon(\mathbf{p}) = \{\mathbf{x} = (x_1, x_2, x_3) : |x_j - p_j| < \epsilon \text{ for } j = 1, 2, 3\}.$$

Note that $\partial\mathcal{U}_\epsilon$ has a well-defined outward normal $\mathbf{n}$ except on a (negligible) set of surface measure zero (the edges and corners). Thus, given a vector field $\mathbf{v} = (v_1, v_2, v_3) \in C^1(\mathcal{U} \to \mathbb{R}^3)$, the integral

$$\int_{\partial\mathcal{U}_\epsilon} \mathbf{v} \cdot \mathbf{n}$$

makes sense.

(a) *Let* $F = \{(p_1 + \epsilon, x_2, x_3) : |x_j - p_j| < \epsilon \text{ for } j = 2, 3\}$ *be the front face of* $\partial\mathcal{U}_\epsilon$*. Define the back face* B *directly opposite* F *on* $\partial\mathcal{U}_\epsilon$ *and use the mean value theorem to show*

$$\int_{F\cup B} \mathbf{v}\cdot\mathbf{n} = 2\epsilon \int_{-\epsilon}^{\epsilon}\int_{-\epsilon}^{\epsilon} \frac{\partial v_1}{\partial x_1}(x_*, p_2 + y, p_3 + z)\, dy\, dz$$

where $x_* = x_*(y, z)$ *is some real number (depending on* y *and* z*) with* $p_1 - \epsilon < x_* < p_1 + \epsilon$*.*

(b) *Write down similar expressions for the other two pairs of faces (top and bottom, left and right).*

(c) *Compute the limit*

$$\lim_{\epsilon\searrow 0} \frac{1}{\mu(\mathcal{U}_\epsilon)} \int_{\partial\mathcal{U}_\epsilon} \mathbf{v}\cdot\mathbf{n}$$

to obtain the usual formula for $\operatorname{div}\mathbf{v}$ *in standard rectangular coordinates.*

Theorem 3 *(The Divergence Theorem) If* $\mathcal{A}$ *is a domain of integration with an outward normal on* $\partial\mathcal{A}$ *and well-defined* C^1 *vector fields* $C^1(\overline{\mathcal{A}} \to \mathcal{V})$*, then*

$$\int_{\mathcal{A}} \operatorname{div}\mathbf{v} = \int_{\partial\mathcal{A}} \mathbf{v}\cdot\mathbf{n}$$

for every $\mathbf{v} \in C^1(\overline{\mathcal{A}} \to \mathcal{V})$*.*

(pseudo)Proof: We assume we can partition $\mathcal{A}$ using well-behaved pieces $\mathcal{A}_j$, $j = 1, 2, \dots, k$ (each admitting flux integrals around $\partial\mathcal{A}_j$). We also assume the pieces can be made small so that we obtain a family of these partitions with

$$\|\mathcal{P}\| = \max\{\operatorname{diam}(\mathcal{A}_j) : j = 1, 2, \dots, k\} \to 0.$$

Observing that when $A = \partial\mathcal{A}_i \cap \partial\mathcal{A}_j$ with $i \neq j$, the outward normal to $\partial\mathcal{A}_i$ points in the opposite direction to the outward normal to $\partial\mathcal{A}_j$, so that

$$\int_{A\cap\partial\mathcal{A}_i} \mathbf{v}\cdot\mathbf{n} + \int_{A\cap\partial\mathcal{A}_j} \mathbf{v}\cdot\mathbf{n} = 0$$

and

$$\sum_{j=1}^{k} \int_{\partial\mathcal{A}_j} \mathbf{v}\cdot\mathbf{n} = \int_{\partial\mathcal{A}} \mathbf{v}\cdot\mathbf{n}.$$

Taking the pieces $\mathcal{A}_j$ small then we have by the definition of the divergence

$$\begin{aligned}\int_{\partial\mathcal{A}} \mathbf{v}\cdot\mathbf{n} &= \sum_{j=1}^{k} \int_{\partial\mathcal{A}_i} \mathbf{v}\cdot\mathbf{n} \\ &\approx \sum_{j=1}^{k} \mu(\mathcal{A}_j) \operatorname{div}\mathbf{v}(x_j^*)\end{aligned}$$

where x_j^* is some point in $\mathcal{A}_j$. But note that this last expression is precisely a Riemann sum for

$$\int_{\mathcal{A}} \operatorname{div}\mathbf{v}.$$

Thus, taking the limit as $\|\mathcal{P}\| \to 0$, we should obtain the result. The only weak point in this argument, with respect to making it rigorous, is that the number of terms in the approximation

$$\sum_{j=1}^{k} \mu(\mathcal{A}_j) \operatorname{div}\mathbf{v}(x_j^*)$$

tends to infinity, so some uniformity is required to conclude

$$\lim_{\|\mathcal{P}\|\to 0} \left| \sum_{j=1}^{k} \left(\int_{\partial \mathcal{A}_i} \mathbf{v} \cdot \mathbf{n} - \mu(\mathcal{A}_j) \operatorname{div} \mathbf{v}(x_j^*) \right) \right| = 0.$$

We can call a domain of integration $\mathcal{A}$ on which the divergence of C^1 vector fields is well-defined on on which the divergence theorem holds a **domain of integration for the divergence theorem**.

2.2 Product Rule for the Divergence

If $\phi \in C^1(\overline{\mathcal{A}})$ where $\mathcal{A}$ is a domain of integration for the divergence theorem, then $\phi\mathbf{v} \in C^1(\mathcal{A} \to \mathcal{V})$ for each $\mathbf{v} \in C^1(\mathcal{A} \to \mathcal{V})$. Let us assume $\mathcal{A}$ also admits directional differentiation and an inner product, so that a gradient operator grad : $C^1(\mathcal{A}) \to C^0(\mathcal{A} \to \mathcal{V})$ is defined by

$$df_{\mathbf{x}}(\mathbf{v}) = \langle \operatorname{grad} f(\mathbf{x}), \mathbf{v} \rangle \qquad \text{for all } \mathbf{v} \in \mathcal{V} \text{ at } x \in \mathcal{A}.$$

In this situation and

$$\operatorname{div}(\phi\mathbf{v}) = \langle \operatorname{grad} \phi, \mathbf{v} \rangle + \phi \operatorname{div} \mathbf{v}.$$

This is called the product rule for the divergence.

If $\mathcal{A} = \mathcal{U}$ is a full-dimension open subset of $\mathbb{R}^n$, then this product rule holds in the (coordinate) form

$$\operatorname{div}(\phi\mathbf{v}) = D\phi \cdot \mathbf{v} + \phi \operatorname{div} \mathbf{v}. \tag{8}$$

Unfortunately, I do not know a nice coordinate free proof of the product rule for the divergence. In the special case of (8), the result can be obtained easily using the product rule for partial derivatives.

2.3 First Variation Formula

We now address our main objective. Let $\mathcal{F} : \mathcal{A} \to \mathbb{R}$ be a Lagrangian integral functional of the form

$$\mathcal{F}[u] = \int_{\mathcal{U}} F(\mathbf{x}, u, Du)$$

where $\mathcal{U}$ is a bounded open subset of $\mathbb{R}^n$, the admissible class $\mathcal{A} \subset C^1(\overline{\mathcal{U}})$, and $F \in C^1(\mathcal{U} \times \mathbb{R}^1 \times \mathbb{R}^n$ is the Lagrangian with $F = F(\mathbf{x}, z, \mathbf{p})$. Given $u \in \mathcal{A}$ and $\phi \in C_c^\infty(\mathcal{U})$, or more generally an admissible variation $\phi \in \mathcal{V}$, we define the first variation as usual by

$$\delta\mathcal{F}_u[\phi] = \frac{d}{d\epsilon} \mathcal{F}[u + \epsilon\phi]\Big|_{\epsilon=0}.$$

Using the chain rule, since $F(\mathbf{x}, u + \epsilon\phi, D(u + \epsilon\phi)) = F(\mathbf{x}, u + \epsilon\phi, Du + \epsilon D\phi)$, we have

$$\delta\mathcal{F}_u[\phi] = \int_{\mathcal{U}} \left(\frac{\partial F}{\partial z} \phi + \sum_{j=1}^{n} \frac{\partial F}{\partial p_j} \frac{\partial \phi}{\partial x_j} \right). \tag{9}$$

Looking at the first variation formula (9) we recognize in the last term a Euclidean dot product of two fields. One is the gradient of the test function ϕ. The other field is given by the partials of the Lagrangian with respect to the "velocity" variables $p_1, p_2, \ldots, p_n$:

$$D_{\mathbf{p}}F = \left(\frac{\partial F}{\partial p_1}, \frac{\partial F}{\partial p_2}, \ldots, \frac{\partial F}{\partial p_n} \right).$$

Let us call this field the **Lagrangian gradient field** or the **kinetic gradient field**. If we now apply the product rule (8) to the scaled field $\phi D_{\mathbf{p}} F$ we find

$$\operatorname{div}(\phi D_{\mathbf{p}} F) = D_{\mathbf{p}} F \cdot D\phi + \operatorname{div} D_{\mathbf{p}} F \ \phi.$$

In particular, assuming $u \in C^2(\mathcal{U})$ so that the second derivatives in

$$\operatorname{div} D_{\mathbf{p}} F(x, u, Du)$$

make sense, we can write the first variation formula for $u \in \mathcal{A} \cap C^2(\mathcal{U})$ as

$$\delta \mathcal{F}_u[\phi] = \int_{\mathcal{U}} \left(-\operatorname{div} D_{\mathbf{p}} F \ + \frac{\partial F}{\partial z} \right) \phi + \int_{\mathcal{U}} \operatorname{div}(\phi D_{\mathbf{p}} F).$$

By the divergence theorem, the last term is

$$\int_{\mathcal{U}} \operatorname{div}(\phi D_{\mathbf{p}} F) = \int_{\partial \mathcal{U}} \phi D_{\mathbf{p}} F \cdot \mathbf{n} = 0 \tag{10}$$

where $\mathbf{n}$ is the outward normal to $\mathcal{U}$ along $\partial \mathcal{U}$. It may reasonably be objected at this point that $\mathcal{U}$ may be an open set with extremely irregular boundary so that an outward unit normal $\mathbf{n}$ is not defined. In fact, it does require some care to show the vanishing of the expression on the left in (10). Recall, however, that $\operatorname{supp}(\phi) \subset\subset \mathcal{U}$. Thus,

$$\operatorname{dist}(\operatorname{supp}(\phi), \mathcal{U}^c) = \inf\{|\mathbf{x} - \mathbf{y}| : \mathbf{x} \operatorname{supp}(\phi), \ \mathbf{y} \in \mathcal{U}\} > 0.$$

From this observation, it is easy to believe a suitable open set $\mathcal{V}$ with smooth boundary and with $\operatorname{supp} \phi \subset\subset \mathcal{V} \subset\subset \mathcal{U}$ can be constructed to replace $\mathcal{U}$ in the middle expression of (10). In any case, we conclude that for $u \in \mathcal{A} \cap C^2(\mathcal{U})$ there holds

$$\delta \mathcal{F}_u[\phi] = \int_{\mathcal{U}} \left(-\sum_{j=1}^{n} \frac{\partial}{\partial x_j} \frac{\partial F}{\partial p_j} + \frac{\partial F}{\partial z} \right) \phi.$$

In particular, we can apply the fundamental lemma of the calculus of variations to conclude

$$\operatorname{grad} \mathcal{F}[u] = -\sum_{j=1}^{n} \frac{\partial}{\partial x_j} \left(\frac{\partial F}{\partial p_j}(x, u, Du) \right) + \frac{\partial F}{\partial z}(x, u, Du) = 0 \tag{11}$$

whenever $u \in \mathcal{A} \cap C^2(\mathcal{U})$ is an extremal, i.e., satisfies $\delta \mathcal{F}_u[\phi] \equiv 0$ for all $\phi \in C_c^\infty(\mathcal{U})$. The condition (11) is called the **Euler-Lagrange PDE** for the function u of several variables.

Exercise 13 *Find the Euler-Lagrange equation for the Dirichlet energy*

$$\mathcal{D}[u] = \frac{1}{2} \int_{\mathcal{U}} |Du|^2.$$

The equation you find is called **Laplace's equation**.

Exercise 14 *Convince yourself that the area of the graph*

$$\mathcal{G} = \{(x, y, u(x, y)) : (x, y) \in \mathcal{U}\}$$

of a function $u \in C^1(\mathcal{U})$ where $\mathcal{U}$ is an open subset of $\mathbb{R}^2$ is given by

$$\operatorname{area}[u] = \int_{\mathcal{U}} \sqrt{1 + |Du|^2}.$$

Hint: Find a parameterization of $\mathcal{G}$ on $\mathcal{U}$ and then change variables in the area integral

$$\int_{\mathcal{G}} 1.$$

The Euler-Lagrange equation associated with the area functional area : $\mathcal{A} \to \mathbb{R}$ *is called the* **mean curvature equation***. Find the mean curvature equation and explain (1) why the Euler-Lagrange operator in that case represents a curvature, (2) why all affine functions on $\mathcal{U}$ are solutions, and (3) why all solutions are not given by affine functions.*

3 Preliminary Results for Laplace's Equation

We consider here functions $u \in C^2(\mathcal{U})$ defined on a bounded open subset $\mathcal{U}$ of $\mathbb{R}^2$. Such a function is said to satisfy **Laplace's equation** (classically) if

$$\Delta u = \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_j^2} = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} = 0.$$

A solution of Laplace's equation is also called a **harmonic function**.

3.1 Mean Value Property

If u satisfies Laplace's equation and $B_a(\mathbf{p}) \subset\subset \mathcal{U}$, then

$$u(\mathbf{p}) = \frac{1}{\omega_n} a^n \int_{B_a(\mathbf{p})} u$$

and

$$u(\mathbf{p}) = \frac{1}{n\omega_n r^{n-1}} \int_{\partial B_a(\mathbf{p})} u.$$

Either of these expressions is called the **mean value property** of harmonic functions. The proof that these expressions hold is given as Problem 3 of Assignment 4 (MATH 6702 Spring 2001).

3.2 Strong Maximum Principle

It is relatively easy to see, just using elementary calculus, that if $u \in C^2(\mathcal{U}) \cap C^0(\overline{\mathcal{U}})$ where $\mathcal{U}$ is a bounded open subset of $\mathbb{R}^n$ and

$$\Delta u = \sum_{j=1}^{m} \frac{\partial^2 u}{\partial x_j^2} = 0 \qquad \text{on } \mathcal{U},$$

then

$$u(\mathbf{p}) \le \max_{\mathbf{x} \in \partial\mathcal{U}} u(\mathbf{x}) \qquad \text{for } \mathbf{p} \in \mathcal{U}.$$

This is called the maximum principle (or sometimes the weak maximum principle). The proof of the weak maximum principle was given as Problem 9 of Assignment 3 (MATH 6702 Spring 2001). The classical strong maximum principle is the following:

Theorem 4 *If $\mathcal{U} \subset \mathbb{R}^n$ is open, bounded, and connected and $u \in C^2(\mathcal{U})$ is harmonic on $\mathcal{U}$, then either*

$$u(\mathbf{p}) < \max_{\mathbf{x} \in \partial\mathcal{U}} u(\mathbf{x}) \qquad \textit{for } \mathbf{p} \in \mathcal{U}$$

or $u(\mathbf{p}) \equiv \max_{\mathbf{x} \in \partial\mathcal{U}} u(\mathbf{x})$ for all $\mathbf{p} \in \overline{\mathcal{U}}$.

Proof: If $\mathbf{p} \in \mathcal{U}$, there is some $r > 0$ with $\overline{B_r(\mathbf{p})} \subset \mathcal{U}$. Let $M = \max_{\partial\mathcal{U}} u$. Then by the weak maximum principle and the mean value property

$$u(\mathbf{p}) = \frac{1}{\omega_n r^n} \int_{B_r(\mathbf{p})} u \le M.$$

If $u(\mathbf{p}) = M$, then we claim $u(\mathbf{x}) \equiv M$ for every $\mathbf{x} \in B_r(\mathbf{p})$. To see this, assume there is some $\epsilon > 0$ for which

$$A = \{\mathbf{x} \in B_r(\mathbf{p}) : u(\mathbf{x}) \le M - \epsilon\} \ne \phi$$

By continuity A will have positive measure. Then we can write $B_r(\mathbf{p}) = A \cup B$ where

$$B = \{\mathbf{x} \in B_r(\mathbf{p}) : M - \epsilon \le u(\mathbf{x}) \le M\}.$$

Since A has positive measure, then $\int_A u \le (M - \epsilon)\mu(A)$, and

$$\begin{aligned} M = u(\mathbf{p}) &= \frac{1}{\omega_n r^n} \int_{B_r(\mathbf{p})} u \\ &= \frac{1}{\omega_n r^n} \left(\int_A u + \int_B u \right) \\ &\le \frac{1}{\omega_n r^n} \left((M - \epsilon)\mu(A) + M\mu(B) \right) \\ &= \frac{1}{\omega_n r^n} \left(M(\mu(A) + \mu(B)) - \epsilon\mu(A) \right) \\ &= M - \frac{\epsilon\mu(A)}{\omega_n r^n} \\ &< M. \end{aligned}$$

This contradiction tells us $u \equiv M$ on $B_r(\mathbf{p})$. Since this applies to any point $\mathbf{p}$ with $u(\mathbf{p}) = M$, we have shown

$$U_1 = \{\mathbf{x} \in \mathcal{U} : u(\mathbf{x}) = M\}$$

is open. It is quite easy to see from continuity that

$$U_2 = \mathcal{U} \backslash U_1 = \{\mathbf{x} \in \mathcal{U} : u(\mathbf{x}) < M\}$$

is also open. This essentially completes the proof if we know the definition of **connected**, which is the following: A set C is connected if C cannot be written as a disjoint union of nonempty open sets. That is, if $C = U_1 \cup U_2$ where U_1 and U_2 are open sets and C is connected, then one of the following must hold

1. $U_1 \cap U_2 \ne \phi$,
2. $U_1 = \phi$, or
3. $U_2 = \phi$.

In our case, $\mathcal{U} = U_1 \cup U_2$ and U_1 and U_2 are clearly disjoint. We have also assume $U_1 \ne \phi$. Therefore, it must be that $U_2 = \phi$ and $u \equiv M$. $\square$

3.3 Higher Regularity

We will prove a **converse** of the mean value property and a rather amazing higher regularity result. Note that if $\Delta u = 0$, then this says some particular linear combination of the homogeneous second partial derivatives vanishes. On the face of it, there is really no reason to think even the second mixed partial derivatives of u exist much less that third and higher order partials are well-defined.

Theorem 5 *If $\mathcal{U}$ is an open subset of $\mathbb{R}^n$ and*

1. $u \in C^0(\mathcal{U})$, *and*
2. *For each $\mathbf{p} \in \mathcal{U}$, we have*

$$u(\mathbf{p}) = \frac{1}{n\omega_n r^{n-1}} \int_{\partial B_r(\mathbf{p})} u \qquad \text{whenever} \quad B_r(\mathbf{p}) \subset\subset \mathcal{U},$$

then

(a) $u \in C^\infty(\mathcal{U})$, *and*

(b) $\Delta u \equiv 0$.

Proof: Let $\mathbf{p} \in \mathcal{U}$. Then there is some $a > 0$ with $B_a(\mathbf{p}) \subset\subset \mathcal{U}$. Let μ_ϵ be the standard mollifier with $\epsilon < a/2$. Finally, define the extension $v : \mathbb{R}^n \to \mathbb{R}$ of u by

$$v(\mathbf{x}) = \begin{cases} u(\mathbf{x}), & \text{if } \mathbf{x} \in \overline{B_a(\mathbf{p})} \\ 0, & \text{if } \mathbf{x} \in \mathbb{R}^n \backslash \overline{B_a(\mathbf{p})}. \end{cases}$$

Observe that $\mu_\epsilon * v : \mathbb{R}^n \to \mathbb{R}$ defines a function $\mu_\epsilon * v \in C^\infty(B_{a/2}(\mathbf{p}))$. On the other hand, for $\mathbf{x} \in B_{a/2}(\mathbf{p})$, we have $B_\epsilon(\mathbf{x}) \subset B_a(\mathbf{p})$ and

$$\begin{aligned} \mu_\epsilon * v(\mathbf{x}) &= \int_{\mathbf{q} \in B_\epsilon(\mathbf{x})} \mu_\epsilon(\mathbf{x} - \mathbf{q})\ v(\mathbf{q}) \\ &= \int_{\mathbf{q} \in B_\epsilon(\mathbf{x})} \mu_\epsilon(\mathbf{x} - \mathbf{q})\ u(\mathbf{q}) \\ &= \int_0^\epsilon \left(\int_{\mathbf{q} \in \partial B_r(\mathbf{x})} \mu_\epsilon(\mathbf{x} - \mathbf{q}) u(\mathbf{q}) \right) dr. \end{aligned}$$

However, recall that by part (c) of Exercise 9

$$\mu_\epsilon(\mathbf{x} - \mathbf{q}) = \frac{1}{\epsilon^n} \psi\left(\frac{|\mathbf{x} - \mathbf{q}|}{\epsilon} \right) = \frac{1}{\epsilon^n} \psi\left(\frac{r}{\epsilon} \right)$$

is constant for $\mathbf{q} \in B_r(\mathbf{x})$. Therefore,

$$\begin{aligned} \mu_\epsilon * v(\mathbf{x}) &= \int_0^\epsilon \left(\frac{1}{\epsilon^n} \psi\left(\frac{r}{\epsilon}\right) \right) \left(\int_{\mathbf{q} \in \partial B_r(\mathbf{x})} u(\mathbf{q}) \right) dr. \\ &= \int_0^\epsilon \left(\frac{1}{\epsilon^n} \psi\left(\frac{r}{\epsilon}\right) \right) \left(n\omega_n r^{n-1} u(\mathbf{x}) \right) dr. \\ &= u(\mathbf{x}) \int_0^\epsilon \left(\frac{1}{\epsilon^n} \psi\left(\frac{r}{\epsilon}\right) \right) \left(\int_{\partial B_r(\mathbf{x})} 1 \right) dr. \\ &= u(\mathbf{x}) \int_0^\epsilon \left(\int_{\mathbf{q} \in \partial B_r(\mathbf{x})} \mu_\epsilon(\mathbf{x} - \mathbf{q}) \right) dr. \\ &= u(\mathbf{x}) \int \mu_\epsilon \\ &= u(\mathbf{x}). \end{aligned}$$

We have shown u agrees with a smooth function $\mu_\epsilon * v$ on $B_{a/2}(\mathbf{p})$. Since $\mathbf{p}$ was arbitrary, this means $u \in C^\infty(\mathcal{U})$.

Having established the remarkable fact that u is smooth, merely by virtue of the fact that its values are given by the avarages of its values on spheres, we begin again with a ball $B_\epsilon(\mathbf{p}) \subset\subset \mathcal{U}$ and compute using the divergence theorem first based on the observation that div $Du = \Delta u$:

$$\begin{aligned}
\int_{B_\epsilon(\mathbf{p})} \Delta u &= \int_{\mathbf{x}\in\partial B_\epsilon(\mathbf{p})} Du(\mathbf{x}) \cdot \mathbf{n} \\
&= \int_{\mathbf{x}\in\partial B_\epsilon(\mathbf{0})} Du(\mathbf{p}+\mathbf{x}) \cdot \frac{\mathbf{x}}{\epsilon} \\
&= \epsilon^{n-1} \int_{\mathbf{x}\in\partial B_1(\mathbf{0})} Du(\mathbf{p}+\epsilon\mathbf{x}) \cdot \mathbf{x} \\
&= \epsilon^{n-1} \frac{d}{d\epsilon} \int_{\mathbf{x}\in\partial B_1(\mathbf{0})} u(\mathbf{p}+\epsilon\mathbf{x}) \\
&= \epsilon^{n-1} \frac{d}{d\epsilon} \left(\frac{1}{\epsilon^{n-1}} \int_{\mathbf{x}\in\partial B_\epsilon(\mathbf{p})} u(\mathbf{x}) \right) \\
&= n\omega_n \epsilon^{n-1} \frac{d}{d\epsilon} \left(\frac{1}{n\omega_n \epsilon^{n-1}} \int_{\mathbf{x}\in\partial B_\epsilon(\mathbf{p})} u(\mathbf{x}) \right) \\
&= n\omega_n \epsilon^{n-1} \frac{d}{d\epsilon} [u(\mathbf{p})] \\
&= 0.
\end{aligned}$$

In view of this calculation, we obtain by a minor variation of the argument to prove the fundamental lemma of the calculus of variations that

$$\Delta u \equiv 0 \qquad \text{on } \mathcal{U}. \qquad \square$$

Corollary 3 *If $\mathcal{U}$ is open in $\mathbb{R}^n$ and $u \in C^2(\mathcal{U})$ is harmonic, then $u \in C^\infty(\mathcal{U})$.*

Given a bounded open set $\mathcal{U} \subset \mathbb{R}^n$, a function $u \in C^1(\overline{\mathcal{U}})$ is said to be a **weak extremal** for the Dirichlet energy if

$$\int_{\mathcal{U}} Du \cdot D\phi = 0 \qquad \text{for all } \phi \in C_c^\infty(\mathcal{U}).$$

Such a function is also called C^1 weakly harmonic, and there are even weaker versions of weakly harmonic functions. They all turn out to be in $C^\infty(\mathcal{U})$.