Some notes related to partial differential equations

Stephen Semmes
Rice University

Preface

These informal notes are intended to complement more detailed treatments, as in the references. Some familiarity with basic analysis and linear algebra would be helpful, and some definitions and results along these lines are reviewed here. Some familiarity with Lebesgue measure and integration could be helpful as well, but we shall normally not be getting into this too much here.

Of course, there are many connections between complex analysis and partial differential equations. The reader is not necessarily expected to be familiar with complex analysis here, although some familiarity would be helpful in some places.

The subject of partial differential equations is obviously closely related to that of ordinary differential equations. Often only basic facts about ordinary differential equations are used here, but some familiarity with standard results related to existence and uniqueness of solutions would be helpful in some places. More precisely, some familiarity with standard results concerning the dependence of solutions on initial conditions and other parameters would be helpful in some places.

There are many connections between partial differential equations, Fourier analysis, and functional analysis too. We shall not get into this too much here, but some of these connections will be mentioned a bit, or are fairly close.

A number of the texts in the bibliography include some aspects of the history of differential equations and related matters, such as Fourier analysis. In particular, one may be interested in [6, 20, 21, 38, 41, 42, 46, 48, 50, 51, 86, 88, 89, 90] in this regard.

Contents

Chapter 1

Some basic facts

Some very interesting introductory remarks about partial differential equations can be found in the first chapter of [29]. Another interesting overview with a somewhat different perspective is in Section A of Chapter 1 of [32]. Here we begin with some basic notions related to Euclidean spaces and functions on them, which are helpful for this.

1.1 Some preliminaries about $\mathbf{R}^n$

Let n be a positive integer, and let $\mathbf{R}^n$ be the usual space of n-tuples $x = (x_1, \ldots, x_n)$ of real numbers. If $x, y \in \mathbf{R}^n$ and $t \in \mathbf{R}$, then $x + y$ and $t\,x$ can be defined as elements of $\mathbf{R}^n$ using coordinatewise addition and scalar multiplication, as usual.

The *standard Euclidean norm* of $x \in \mathbf{R}^n$ is defined by

$$|x| = \Big(\sum_{j=1}^{n} x_j^2\Big)^{1/2}, \tag{1.1.1}$$

using the nonnegative square root on the right side. This reduces to the usual absolute value of a real number when $n = 1$. Observe that

$$|t\,x| = |t|\,|x| \tag{1.1.2}$$

for every $t \in \mathbf{R}$ and $x \in \mathbf{R}^n$. It is well known that

$$|x + y| \le |x| + |y| \tag{1.1.3}$$

for every $x, y \in \mathbf{R}^n$. This is called the *triangle inequality* for the standard Euclidean norm.

The *standard Euclidean metric* on $\mathbf{R}^n$ is defined by

$$d(x, y) = |x - y| \tag{1.1.4}$$

for every $x, y \in \mathbf{R}^n$. This may also be described as the distance between x and y, with respect to the standard Euclidean metric.

If $x \in \mathbf{R}^n$ and r is a positive real number, then the *open ball* in $\mathbf{R}^n$ centered at x with radius r is defined by

$$B(x, r) = \{y \in \mathbf{R}^n : |x - y| < r\}. \tag{1.1.5}$$

Similarly, the *closed ball* in $\mathbf{R}^n$ centered at x with radius r is defined by

$$\overline{B}(x, r) = \{y \in \mathbf{R}^n : |x - y| \le r\}. \tag{1.1.6}$$

A subset U of $\mathbf{R}^n$ is said to be an *open set* with respect to the standard Euclidean metric if for every $x \in U$ there is an $r > 0$ such that

$$B(x, r) \subseteq U. \tag{1.1.7}$$

It is well known and not too difficult to show that any open ball in $\mathbf{R}^n$ is an open set in this sense.

Let E be a subset of $\mathbf{R}^n$. The *closure* of E in $\mathbf{R}^n$ with respect to the standard Euclidean metric is defined to be the set $\overline{E}$ of all $x \in \mathbf{R}^n$ with the following property: for every $r > 0$ there is a $y \in E$ such that

$$|x - y| < r. \tag{1.1.8}$$

Equivalently, this means that for every $r > 0$,

$$E \cap B(x, r) \ne \emptyset. \tag{1.1.9}$$

Note that $E \subseteq \overline{E}$.

If

$$\overline{E} = E, \tag{1.1.10}$$

then E is said to be a *closed set* in $\mathbf{R}^n$ with respect to the standard Euclidean metric. It is well known and not too difficult to show that any closed ball in $\mathbf{R}^n$ is a closed set. If E is any subset of $\mathbf{R}^n$, then it is well known and not too hard to show that $\overline{E}$ is a closed set.

If $x \in \mathbf{R}^n$ and $r > 0$, then one can check that the closure of $B(x, r)$ in $\mathbf{R}^n$ is equal to $\overline{B}(x, r)$. However, this does not always work in arbitrary metric spaces.

If U is an open subset of $\mathbf{R}^n$, then the *boundary* may be defined as the set ∂U of points in the closure of U that are not in U,

$$\partial U = \overline{U} \setminus U. \tag{1.1.11}$$

If $x \in \mathbf{R}^n$ and $r > 0$, then

$$\partial B(x, r) = \{y \in \mathbf{R}^n : |x - y| = r\}, \tag{1.1.12}$$

but this does not always work in arbitrary metric spaces.

If E is any subset of $\mathbf{R}^n$, then the boundary of E is defined by

$$\partial E = \overline{E} \cap \overline{(\mathbf{R}^n \setminus E)}. \tag{1.1.13}$$

One can check that this is equivalent to the definition in the preceding paragraph when E is an open set.

1.2 Some spaces of functions

Let U be a nonempty open subset of $\mathbf{R}^n$, for some $n \geq 1$. The space of continuous real-valued functions on U may be denoted $C(U)$.

Let f be a real-valued function on U, let x be an element of U, and let l be a positive integer less than or equal to n. The *lth partial derivative* of f at x may be denoted

(1.2.1) $$\partial_l f(x) = D_l f(x) = \frac{\partial f}{\partial x_l}(x),$$

when it exists.

If (1.2.1) exists for every $x \in U$ and $l = 1, \ldots, n$, then f is said to be *continuously differentiable* on U. It is well known that this implies that f is continuous on U, although that may be included in the definition, for convenience. The space of continuously-differentiable real-valued functions on U may be denoted $C^1(U)$.

If k is any positive integer, then we may say that f is *k-times continuously differentiable* if f is continuous on U, and all derivatives of f up to order k of f exist at every point in U, and are continuous on U. The space of k-times continuously-differentiable real-valued functions on U may be denoted $C^k(U)$. More precisely, this may be defined recursively when $k \geq 2$, by saying that $C^k(U)$ consists of all continuously-differentiable real-valued functions f on U such that

(1.2.2) $$\frac{\partial f}{\partial x_l} \in C^{k-1}(U)$$

for each $l = 1, \ldots, n$. It is sometimes convenient to take $C^0(U) = C(U)$.

If derivatives of f of all orders exist everywhere on U and are continuous, then f is said to be *infinitely differentiable*, or *smooth*, on U. The space of infinitely-differentiable real-valued functions on U may be denoted $C^\infty(U)$.

An n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers is said to be a *multi-index*, of *order*

(1.2.3) $$|\alpha| = \sum_{j=1}^n \alpha_j.$$

Of course, this is not necessarily the same as the standard Euclidean norm of α, as an element of $\mathbf{R}^n$, and it should normally be clear which is intended. If $f \in C^k(U)$ for some $k \geq 1$ and $|\alpha| \leq k$, then the corresponding derivative of f of order $|\alpha|$ may be denoted

(1.2.4) $$\partial^\alpha f = D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}.$$

Note that this function is continuously differentiable of order

(1.2.5) $$k - |\alpha|$$

on U under these conditions.

If f is a twice continuously-differentiable function on U, then it is well known that

$$\frac{\partial^2 f}{\partial x_j \, \partial x_l} = \frac{\partial^2 f}{\partial x_l \, \partial x_j} \tag{1.2.6}$$

on U for every $j, l = 1, \ldots, n$. Similarly, if f is k-times continuously differentiable on U, then derivatives of f up to order k may be taken in any order.

Sometimes derivatives are expressed using subscripts to indicate the variables in which the derivative is taken. Thus one may put

$$f_{x_j} = \frac{\partial f}{\partial x_j}, \quad f_{x_j \, x_l} = \frac{\partial^2 f}{\partial x_j \, \partial x_l}, \tag{1.2.7}$$

and so on, where appropriate.

If $x \in \mathbf{R}^n$, then we may put

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \tag{1.2.8}$$

where $x_j^{\alpha_j}$ is interpreted as being equal to 1 when $\alpha_j = 0$, even when $x_j = 0$. This defines a real-valued function on $\mathbf{R}^n$, which is the *monomial* of degree $|\alpha|$ associated to α.

Similarly, (1.2.4) corresponds to

$$\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} \tag{1.2.9}$$

or

$$D^\alpha = D_1^{\alpha_1} \cdots D_n^{\alpha_n} \tag{1.2.10}$$

applied to f. More precisely, $\partial_j = D_j$ defines a linear mapping from $C^k(U)$ into $C^{k-1}(U)$ for each $k \geq 1$. Composition of these mappings can be considered as a type of multiplication, with $\partial_j^{\alpha_j} = D_j^{\alpha_j}$ interpreted as being the identity mapping when $\alpha_j = 0$.

1.3 Partial differential equations

Let k and n be positive integers, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let u be a k-times continuously-differentiable real-valued function on U. A kth-order partial differential equation for u on U can be expressed as

$$F(D^k u(x), D^{k-1} u(x), \ldots, Du(x), u(x), x) = 0, \tag{1.3.1}$$

as in Section 1.1 of [29], and Section A of Chapter 1 of [32]. Here $D^l u(x)$ is intended to represent the collection of all possible derivatives of u of order l at x, which may be identified with an element of $\mathbf{R}^{n^l}$. Thus F may be considered as a real-valued function on

$$\mathbf{R}^{n^k} \times \mathbf{R}^{n^{k-1}} \times \cdots \times \mathbf{R}^n \times \mathbf{R} \times U. \tag{1.3.2}$$

A linear kth-order partial differential equation for u on U can be expressed as

$$\sum_{|\alpha| \le k} a_\alpha(x)\, \partial^\alpha u(x) = f(x), \tag{1.3.3}$$

as in Section 1.1 of [29], and Section A of Chapter 1 of [32]. More precisely, the sum is taken over all multi-indices α with $|\alpha| \le k$, which is of course a finite set. Thus $a_\alpha(x)$ should be a function on U for each such α, as well as $f(x)$. If $f(x) = 0$ for every $x \in U$, then (1.3.3) is said to be *homogeneous*.

One may also consider *systems* of partial differential equations, as in [29]. In this case, one can think of u as taking values in $\mathbf{R}^m$ for some positive integer m. Continuous differentiability of u of order k on U means that each of the m components of u is k-times continuously differentiable as a real-valued function on U. One considers finitely many equations involving the components of u and their deriviatives of order up to k on U, as before.

Let us say that a partial differential equation as in (1.3.1) is *invariant under translations* if F does not depend on x in the last variable. This means that F may be considered as a real-valued function on

$$\mathbf{R}^{n^k} \times \mathbf{R}^{n^{k-1}} \times \cdots \times \mathbf{R}^n \times \mathbf{R}, \tag{1.3.4}$$

so that (1.3.1) becomes

$$F(D^k u(x), D^{k-1} u(x), \ldots, Du(x), u(x)) = 0. \tag{1.3.5}$$

If u satisfies this equation on U and $a \in \mathbf{R}^n$, then

$$u(x - a) \tag{1.3.6}$$

satisfies the same equation on

$$U + a = \{x + a : x \in U\}. \tag{1.3.7}$$

Note that this is also an open set in $\mathbf{R}^n$. Of course, there are analogous notions for systems.

The left side of (1.3.3) is said to have *constant coefficients* if $a_\alpha(x)$ is a constant for each multi-index α. If f is also a constant, then (1.3.3) is invariant under translations, as in the preceding paragraph. There are analogous notions for linear systems, as before.

Let v be a continuously-differentiable $\mathbf{R}^n$-valued function on U. The *divergence* of v is the real-valued function on U defined as usual by

$$\operatorname{div} v = \sum_{j=1}^n \frac{\partial v_j}{\partial x_j}, \tag{1.3.8}$$

where $v_j(x)$ is the jth coordinate of $v(x)$ for each $j = 1, \ldots, n$.

Let f be a real-valued function on U. The *directional derivative* of f at $x \in U$ in the direction $w \in \mathbf{R}^n$ is defined to be the derivative of

$$f(x + t\, w) \tag{1.3.9}$$

as a function of $t \in \mathbf{R}$ at $t = 0$, if it exists. If f is continuously differentiable on U, then it is well known that the directional derivative exists, and is equal to

$$\sum_{j=1}^{n} w_j \frac{\partial f}{\partial x_j}(x). \tag{1.3.10}$$

1.4 Complex numbers

A complex number z can be expressed in a unique way as

$$z = x + y\,i, \tag{1.4.1}$$

where $x, y \in \mathbf{R}$ and $i^2 = -1$. In this case, x and y are called the *real* and *imaginary parts* of z, and may be denoted $\operatorname{Re} z$, $\operatorname{Im} z$, respectively. The *complex conjugate* of z is the complex number

$$\overline{z} = x - y\,i, \tag{1.4.2}$$

and the *absolute value* or *modulus* of z is the nonnegative real number

$$|z| = (x^2 + y^2)^{1/2}. \tag{1.4.3}$$

In particular, the complex conjugate of $\overline{z}$ is z, and $|\overline{z}| = |z|$.

The real line $\mathbf{R}$ may be considered as a subset of the set $\mathbf{C}$ of complex numbers, and addition and multiplication of real numbers can be extended to complex numbers in a standard way. Note that

$$\overline{z + w} = \overline{z} + \overline{w}, \tag{1.4.4}$$

$$\overline{z\,w} = \overline{z}\,\overline{w} \tag{1.4.5}$$

and

$$z\,\overline{z} = |z|^2 \tag{1.4.6}$$

for every $z, w \in \mathbf{C}$. One can use this to get that

$$|z\,w| = |z|\,|w| \tag{1.4.7}$$

for every $z, w \in \mathbf{C}$. If $z \in \mathbf{C}$ and $z \neq 0$, then z has a multiplicative inverse in $\mathbf{C}$, namely,

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2}. \tag{1.4.8}$$

Of course, (1.4.3) is the same as the standard Euclidean norm of $(x, y) \in \mathbf{R}^2$. The triangle inequality for the standard Euclidean norm on $\mathbf{R}^2$ is the same as saying that

$$|z + w| \le |z| + |w| \tag{1.4.9}$$

for every $z, w \in \mathbf{C}$, which can also be verified more directly in this case. The standard metric on $\mathbf{C}$ is defined by

$$d(z, w) = |z - w|, \tag{1.4.10}$$

which corresponds exactly to the standard Euclidean metric on $\mathbf{R}^2$.

Let n be a positive integer, let U be an open subset of $\mathbf{R}^n$, and let f be a complex-valued function on U. Continuity of f on U can be defined in the same way as for real-valued functions, and is equivalent to continuity of the real and imaginary parts of f. Similarly, differentiability properties of f can be defined in the same way as for real-valued functions, and are equivalent to the corresponding differentiability properties of the real and imaginary parts of f. Complex analysis deals with different types of differentiability properties of complex-valued functions on open subsets of $\mathbf{C}$. This is related to the *Cauchy–Riemann equations* for the real and imaginary parts of such a function.

1.5 Complex exponentials

The *exponential* of a complex number z can be defined by

$$\exp z = \sum_{j=0}^{\infty} \frac{z^j}{j!}, \tag{1.5.1}$$

where the absolute convergence of the series can be obtained from the ratio test, for instance. This is equivalent to taking

$$\exp(x + y\,i) = (\exp x)\,(\cos y + i \sin y) \tag{1.5.2}$$

for every $x, y \in \mathbf{R}$.

It is well known that

$$\exp(z + w) = (\exp z)\,(\exp w) \tag{1.5.3}$$

for every $z, w \in \mathbf{C}$. This can be obtained using the binomial theorem, and standard results about products of absolutely convergent series.

In particular, if $z \in \mathbf{C}$, then one can take $w = -z$ in (1.5.3) to get that $\exp z \neq 0$, with

$$1/(\exp z) = \exp(-z). \tag{1.5.4}$$

Of course, if $x \in \mathbf{R}$, then $\exp x \in \mathbf{R}$, with $\exp x \geq 1$ when $x \geq 0$. If $x \leq 0$, then $0 < \exp x = 1/(\exp(-x)) \leq 1$.

It is easy to see that

$$\overline{(\exp z)} = \exp \overline{z} \tag{1.5.5}$$

for every $z \in \mathbf{C}$. One can use this to get that

$$|\exp(i\,y)| = 1 \tag{1.5.6}$$

for every $y \in \mathbf{R}$.

It is well known that $\exp z$ is complex-analytic, or equivalently holomorphic, as a complex-valued function of $z \in \mathbf{C}$. Here we shall be more concerned with related complex-valued functions of real variables. If $a \in \mathbf{C}$, then $\exp(a\,t)$ may be considered as a complex-valued function of $t \in \mathbf{R}$. It is well known that this function is differentiable, with

$$\frac{d}{dt}(\exp(a\,t)) = a\,(\exp(a\,t)). \tag{1.5.7}$$

Let n be a positive integer, and let $\mathbf{C}^n$ be the space of n-tuples $a = (a_1, \ldots, a_n)$ of complex numbers. If $a, b \in \mathbf{C}^n$, then put

$$a \cdot b = \sum_{j=1}^{n} a_j\, b_j. \tag{1.5.8}$$

If $a \in \mathbf{C}^n$ and $x \in \mathbf{R}^n$, then $\exp(a \cdot x)$ is a complex number, which defines a complex-valued function of x on $\mathbf{R}^n$. This function is continuously differentiable on $\mathbf{R}^n$, with

$$\frac{\partial}{\partial x_j} \exp(a \cdot x) = a_j\,(\exp(a \cdot x)) \tag{1.5.9}$$

for every $j = 1, \ldots, n$.

More precisely, $\exp(a \cdot x)$ is infinitely differentiable as a complex-valued function of x on $\mathbf{R}^n$. If α is a multi-index, then

$$\partial^\alpha \exp(a \cdot x) = a^\alpha \exp(a \cdot x). \tag{1.5.10}$$

Here $a^\alpha = a_1^{\alpha_1} \cdots a_n^{\alpha_n}$, as in Section 1.2, which is now a complex number.

1.6 Complex-valued functions

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. The space of continuous complex-valued functions on U may be denoted $C(U, \mathbf{C})$, and we may use $C(U, \mathbf{R})$ for the space of continuous real-valued functions on U, to be more precise. Note that a complex-valued function on U is continuous if and only if its real and imaginary parts are continuous.

Similarly, if k is a positive integer, then we let $C^k(U, \mathbf{C})$ be the space of k-times continuously-differentiable complex-valued functions on U. Equivalently, these are the complex-valued functions on U whose real and imaginary parts are k-times continuously differentiable. We may use $C^k(U, \mathbf{R})$ for the space of k-times continuously-differentiable real-valued functions on U. As before, we may use the same notation with $k = 0$ for the corresponding spaces of real and complex-valued continuous functions. The space of infinitely-differentiable complex-valued functions on U may be denoted $C^\infty(U, \mathbf{C})$, and we may use $C^\infty(U, \mathbf{R})$ for the space of smooth real-valued functions on U.

Note that $C(U, \mathbf{R})$ and $C(U, \mathbf{C})$ are *vector spaces* over the real and complex numbers, respectively, with respect to pointwise addition and scalar multiplication of functions. We may consider $C^k(U, \mathbf{R})$, $C^k(U, \mathbf{C})$ as linear subspaces of $C(U, \mathbf{R})$, $C(U, \mathbf{C})$, respectively, for each $k \geq 1$. Similarly, $C^\infty(U, \mathbf{R})$, $C^\infty(U, \mathbf{C})$ are linear subspaces of $C^k(U, \mathbf{R})$, $C^k(U, \mathbf{C})$, respectively, for each k.

If α is a multi-index with $|\alpha| \le k$, then ∂^α defines a linear mapping from each of $C^k(U, \mathbf{R})$, $C^k(U, \mathbf{C})$ into $C^{k-|\alpha|}(U, \mathbf{R})$, $C^{k-|\alpha|}(U, \mathbf{C})$, respectively. Similarly, ∂^α defines a linear mapping from each of $C^\infty(U, \mathbf{R})$, $C^\infty(U, \mathbf{C})$ into itself.

Let $a \in \mathbf{C}^n$ be given, so that $\exp(a \cdot x)$ is a smooth complex-valued function on $\mathbf{R}^n$, as in the previous section. This function is an eigenvector for $\partial/\partial x_j$ for each $j = 1, \ldots, n$, as a linear mapping from $C^\infty(\mathbf{R}^n, \mathbf{C})$ into itself, as before. Similarly, $\exp(a \cdot x)$ is an eigenvector for ∂^α for each muli-index α, as a linear mapping from $C^\infty(\mathbf{R}^n, \mathbf{C})$ into itself.

1.7 Polynomials in n variables

Let n be a positive integer, and let us consider polynomials in the n variables $w_1, \ldots, w_n$ with coefficients in $\mathbf{R}$ or $\mathbf{C}$. Such a polynomial can be expressed as

$$p(w) = \sum_{|\alpha| \le N} a_\alpha \, w^\alpha, \tag{1.7.1}$$

where N is a nonnegative integer, and the sum is taken over all multi-indices α with $|\alpha| \le N$. The coefficients a_α may be real or complex numbers for each such α, and the monomial w^α is as defined in Section 1.2. More precisely, p is said to have degree less than or equal to N in this case. Note that $p(w) \in \mathbf{C}$ when $w \in \mathbf{C}^n$, and $p(w) \in \mathbf{R}$ when $w \in \mathbf{R}^n$ and the coefficients a_α are real numbers.

If p is as in (1.7.1), then put

$$p(\partial) = \sum_{|\alpha| \le N} a_\alpha \, \partial^\alpha, \tag{1.7.2}$$

or equivalently

$$p(D) = \sum_{|\alpha| \le N} a_\alpha \, D^\alpha. \tag{1.7.3}$$

This defines a *differential operator* on $\mathbf{R}^n$ with constant coefficients in $\mathbf{R}$ or $\mathbf{C}$, as appropriate, of order less than or equal to N.

Let U be a nonempty open subset of $\mathbf{R}^n$, and suppose that f is a k-times continuously-differentiable real or complex-valued function on U, with $N \le k$. Under these conditions,

$$p(\partial)(f) = \sum_{|\alpha| \le N} a_\alpha \, \partial^\alpha f \tag{1.7.4}$$

defines a $(k - N)$-times continuously-differentiable real or complex-valued function on U, as appropriate. More precisely, this defines a linear mapping from $C^k(U, \mathbf{R})$ or $C^k(U, \mathbf{C})$ into $C^{k-N}(U, \mathbf{R})$ or $C^{k-N}(U, \mathbf{C})$, respectively, as appropriate. Similarly, this defines a linear mapping from $C^\infty(U, \mathbf{R})$ or $C^\infty(U, \mathbf{C})$ into itself, as appropriate.

If $b \in \mathbf{C}^n$, then $\exp(b \cdot x)$ defines an infinitely-differentiable complex-valued function of x on $\mathbf{R}^n$, as in Section 1.5. Observe that

$$p(\partial)(\exp(b \cdot x)) = p(b)\, \exp(b \cdot x). \tag{1.7.5}$$

Thus $\exp(b \cdot x)$ is an eigenvector for $p(\partial)$ as a linear mapping from $C^\infty(\mathbf{R}^n, \mathbf{C})$ into itself, with eigenvalue $p(b)$.

If α, β are multi-indices, then $\alpha + \beta$ can be defined by coordinatewise addition, as usual, and is another multi-index. Clearly

$$|\alpha + \beta| = |\alpha| + |\beta|, \tag{1.7.6}$$

where $|\cdot|$ refers to the order of the multi-index, as in Section 1.2. Observe that

$$w^\alpha\, w^\beta = w^{\alpha+\beta}. \tag{1.7.7}$$

Similarly,

$$\partial^\alpha\, \partial^\beta = \partial^{\alpha+\beta}, \tag{1.7.8}$$

because of the commutativity of derivatives under suitable conditions, as in Section 1.2.

Let $p_1(w)$, $p_2(w)$ be polynomials in $w_1, \ldots, w_n$ with real or complex coefficients, and of degrees less than or equal to nonnegative integers N_1, N_2. The product

$$p(w) = p_1(w)\, p_2(w) \tag{1.7.9}$$

can be defined as a polynomial of degree less than or equal to $N_1 + N_2$ in the usual way, using (1.7.7). Similarly,

$$p(\partial) = p_1(\partial)\, p_2(\partial), \tag{1.7.10}$$

because of (1.7.8).

More precisely, let f be a k-times continuously-differentiable real or complex-valued function on a nonempty open subset U of $\mathbf{R}^n$ again. If α, β are multi-indices with $|\alpha| + |\beta| \le k$, then $\partial^\beta f$ is $(k - |\beta|)$-times continuously differentiable on U, and

$$\partial^\alpha(\partial^\beta f) = \partial^{\alpha+\beta} f \tag{1.7.11}$$

on U. If p_1, p_2, and p are as in the preceding paragraph and $N_1 + N_2 \le k$, then $p_2(\partial)(f)$ is $(k - N_2)$-times continuously differentiable on U, and

$$p_1(\partial)(p_2(\partial)(f)) = p(\partial)(f) \tag{1.7.12}$$

on U.

1.8 Connectedness and convexity

Let n be a positive integer, and let E be a subset of $\mathbf{R}^n$. We say that E is *convex* if for every $x, y \in E$ and $t \in \mathbf{R}$ with $0 \le t \le 1$, we have that

$$(1 - t)\, x + t\, y \in E. \tag{1.8.1}$$

It is well known and not too difficult to show that open and closed balls in $\mathbf{R}^n$, with respect to the standard Euclidean metric (or the metric associated to any norm), are convex.

We say that E is *path connected* if for every $x, y \in E$ there is a continuous path in E connecting x and y. More precisely, this means that there is a continuous mapping f from the closed unit interval $[0, 1]$ in the real line into $\mathbf{R}^n$ such that

$$f(0) = x, \quad f(1) = y, \tag{1.8.2}$$

and $f(t) \in E$ for every $t \in [0, 1]$. If $f_j(t)$ is the jth coordinate of $f(t)$ for every $j = 1, \ldots, n$ and $t \in [0, 1]$, then the continuity of f as a mapping from $[0, 1]$ into $\mathbf{R}^n$ is equivalent to the continuity of f_j as a real-valued function on $[0, 1]$ for each j. If E is convex, then E is clearly path connected.

The precise definition of *connectedness* of subsets of $\mathbf{R}^n$ is a bit complicated, and although we shall not discuss it here, we shall mention some of its properties. It is well known and not too difficult to show that

(1.8.3) path-connected sets are connected.

It is also well known that

(1.8.4) a subset of the real line is connected if and only if it is convex.

Another well-known theorem states that

(1.8.5) connected open subsets of $\mathbf{R}^n$ are path connected.

Let U be an open subset of $\mathbf{R}^n$. In this case,

(1.8.6) U is not connected

if and only if

(1.8.7) U can be expressed as the union of two nonempty disjoint open subsets of $\mathbf{R}^n$.

This is close to the definition of connectedness, depending on how it is formulated.

If $U \neq \emptyset, \mathbf{R}^n$, then

$$\partial U \neq \emptyset. \tag{1.8.8}$$

This is the same as saying that U is not a closed set, because U is an open set, by hypothesis. This can be obtained from the connectedness of $\mathbf{R}^n$. Alternatively, if $x \in U$ and $z \in \mathbf{R}^n \setminus U$, then one can show that there is a $t_0 \in \mathbf{R}$ such that $0 < t_0 \leq 1$ and

$$(1 - t_0)\, x + t_0\, z \in \partial U. \tag{1.8.9}$$

More precisely, one can take t_0 to be the infimum of the set of $t > 0$ such that

$$(1 - t)\, x + t\, z \in \mathbf{R}^n \setminus U. \tag{1.8.10}$$

Let E be a nonempty subset of $\mathbf{R}^n$, and let f be a function on E with values in any set. Let us say that f is *locally constant* at a point $x \in E$ if there is an $r > 0$ such that

$$f(x) = f(y) \tag{1.8.11}$$

for every $y \in E$ with $|x - y| < r$. If E is connected, and f is locally constant at every point in E, then one can show that

$$f \text{ is constant on } E. \tag{1.8.12}$$

One can also show that connectedness is characterized by this property.

Let U be a nonempty open subset of $\mathbf{R}^n$, and let f be a real or complex-valued function on U. Observe that f is locally constant on U if and only if f is continuously-dfferentiable on U, with all of its first partial derivatives equal to 0 on U. The remarks in the preceding paragraph are also a bit simpler in this case.

1.9 Compactness in $\mathbf{R}^n$

Let n be a positive integer, and let E be a subset of $\mathbf{R}^n$. We say that E is *bounded* if there is a nonnegative real number C such that

$$|x| \le C \tag{1.9.1}$$

for every $x \in E$. It is easy to see that open and closed balls in $\mathbf{R}^n$ with respect to the standard Euclidean metric are bounded sets.

The precise definition of compactness of a subset of $\mathbf{R}^n$, or of an arbitrary metric space, is a bit complicated, and we shall not discuss it here. However, we would like to mention the following two well-known results about compactness. The first is that a subset E of $\mathbf{R}^n$ is *compact* if and only if it is closed and bounded. The second is the *extreme value theorem*, which states that if f is a continuous real-valued function on a nonempty compact set E, then the maximum and minimum of f on E are attained.

Let U be an open subset of $\mathbf{R}^n$. The *relative closure* of a subset E of U may be defined to be the intersection of the closure of E in $\mathbf{R}^n$ with U,

$$\overline{E} \cap U. \tag{1.9.2}$$

In particular, E is said to be *relatively closed* in U if

$$E = \overline{E} \cap U. \tag{1.9.3}$$

If E is closed as a subset of $\mathbf{R}^n$, then it follows that E is relatively closed in U. Note that U is automatically relatively closed as a subset of itself.

There is a notion of compactness of a subset E of U relative to U, with respect to the restriction of the standard Euclidean metric on $\mathbf{R}^n$. However, it is well known that this holds if and only if E is compact as a subset of $\mathbf{R}^n$.

Let f be a real or complex-valued function on $\mathbf{R}^n$, or a function with values in $\mathbf{R}^m$ for some positive integer m. The *support* of f is the subset of $\mathbf{R}^n$ defined by

$$\operatorname{supp} f = \overline{\{x \in \mathbf{R}^n : f(x) \neq 0\}}. \tag{1.9.4}$$

Of course, this is a closed set in $\mathbf{R}^n$, by construction.

Thus the support of f is compact exactly when it is bounded. This is the same as saying that $f(x) = 0$ when $|x|$ is sufficiently large.

Suppose now that f is a function defined on an open set $U \subseteq \mathbf{R}^n$. We say that f has *compact support* in U if there is a compact set $E \subseteq \mathbf{R}^n$ such that $E \subseteq U$ and

$$\{x \in U : f(x) \neq 0\} \subseteq E. \tag{1.9.5}$$

1.10 Some derivatives

Let n be a positive integer, and let α be a multi-index. It is customary to put

$$\alpha! = \alpha_1! \, \alpha_2! \cdots \alpha_n!, \tag{1.10.1}$$

which is a positive integer. Observe that

$$\partial^\alpha x^\alpha = \alpha!. \tag{1.10.2}$$

Let β be another multi-index. If

$$\beta_j < \alpha_j \text{ for some } j, \tag{1.10.3}$$

then

$$\partial^\alpha x^\beta = 0. \tag{1.10.4}$$

In particular, this holds when $|\alpha| = |\beta|$ and $\alpha \neq \beta$.

Suppose now that $\alpha_j \leq \beta_j$ for each $j = 1, \ldots, n$, so that $\beta - \alpha$ is a multi-index. Of course, $\partial^\alpha x^\beta$ is a multiple of $x^{\beta - \alpha}$ in this case. If $\alpha \neq \beta$, then we get that $\partial^\alpha \, x^\beta$ is equal to 0 at 0.

Let U be a nonempty open subset of $\mathbf{R}^n$, let k be a positive integer, and let f be a k-times continuously-differentiable real-valued function on U. The degree k Taylor polynomial of f at a point $w \in U$ may be expressed as

$$P(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} \, \partial^\alpha f(w) \, x^\alpha, \tag{1.10.5}$$

where the sum is taken over all multi-indices α with $|\alpha| \leq k$. Using the remarks in the previous paragraphs, we get that

$$\partial^\beta P(0) = \partial^\beta f(w) \tag{1.10.6}$$

for every multi-index β with $|\beta| \leq k$.

Put

$$g(x) = f(w + x) - P(x) \tag{1.10.7}$$

for $x \in U - w$. This is a k-times continuously-differentiable function on $U - w$, with

$$\partial^\beta g(0) = \partial^\beta P(0) - \partial^\beta f(w) = 0 \tag{1.10.8}$$

for every multi-index β with $|\beta| \le k$.

If $x \in \mathbf{R}^n$ and $|x|$ is sufficiently small, then

$$t\,x \in U - w \tag{1.10.9}$$

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

$$g(t\,x) \tag{1.10.10}$$

may be considered as a k-times continuously-differentiable function of t on an open set in the real line that contains $[0, 1]$. The derivatives of $g(t\,x)$ in t up to order k can be expressed in terms of derivatives of g, as a function on $U - w$, of the same order. These derivatives are equal to 0 at $t = 0$, because of (1.10.8).

One can use this to show that

$$\lim_{x \to 0} |x|^{-k}\, g(x) = 0, \tag{1.10.11}$$

which is Taylor's theorem in n dimensions. This uses the fact that $\partial^\beta g$ is small near 0 when $|\beta| = k$, because of (1.10.8) and the continuity of $\partial^\beta g$ on $U - w$. More precisely, this implies that the kth derivative of (1.10.10) in t is small when $|x|$ is small and $t \in [0, 1]$.

1.11 Some smooth functions

Consider the real-valued function defined on $\mathbf{R}$ by

$$\begin{aligned} \psi(t) &= \exp(-1/t) \quad \text{when } t > 0 \\ &= 0 \qquad\qquad\quad\; \text{when } t \le 0. \end{aligned} \tag{1.11.1}$$

It is well known and not too difficult to show that ψ is infinitely differentiable on $\mathbf{R}$, with all of its derivatives at 0 equal to 0.

Let a, b be real numbers with $a < b$, and put

$$\psi_{a,b}(t) = \psi(t - a)\, \psi(b - t). \tag{1.11.2}$$

This is an infinitely-differentiable function on $\mathbf{R}$ that is positive on (a, b), and equal to 0 otherwise.

One can integrate $\psi_{a,b}$ to get an infinitely-differentiable function on $\mathbf{R}$ that is equal to 0 when $t \le a$, is a positive constant when $t \ge b$, and strictly increasing on (a, b). Using this, one can get infinitely-differentiable nonnegative real-valued functions on $\mathbf{R}$ that are equal to 1 on any given closed interval, and equal to 0 on the complement of a slightly larger open interval.

Alternatively,

$$\psi(t-a)+\psi(b-t) \tag{1.11.3}$$

is a positive smooth function on $\mathbf{R}$, so that

$$\frac{\psi(t-a)}{\psi(t-a)+\psi(b-t)} \tag{1.11.4}$$

and

$$\frac{\psi(b-t)}{\psi(t-a)+\psi(b-t)} \tag{1.11.5}$$

are nonnegative smooth functions on $\mathbf{R}$. It is easy to see that (1.11.4) is equal to 0 when $t \le a$, and to 1 when $t \ge b$. Similarly, (1.11.5) is equal to 0 when $t \ge b$, and to 1 when $t \le a$. Note that the sum of (1.11.4) and (1.11.5) is equal to 1 for every $t \in \mathbf{R}$.

If n is any positive integer, then one can use functions like these to get a lot of infinitely-differentiable nonnegative real-valued functions on $\mathbf{R}^n$ with compact support. One can take products of smooth functions on $\mathbf{R}$ with compact support in each variable, for instance. If $a \in \mathbf{R}^n$, then

$$|x-a|^2 = \sum_{j=1}^{n}(x_j-a_j)^2 \tag{1.11.6}$$

is a polynomial in x, and infinitely differentiable on $\mathbf{R}^n$ in particular. If ϕ is a smooth real-valued function on $\mathbf{R}$, then

$$\phi(|x-a|^2) \tag{1.11.7}$$

is a smooth function on $\mathbf{R}^n$. If $\phi(t) = 0$ when $t \in \mathbf{R}$ is sufficiently large, then (1.11.7) has compact support in $\mathbf{R}^n$.

1.12 Semilinearity and quasilinearity

Let k and n be positive integers, let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a k-times continuously-differentiable real-valued function on U. One may be interested in kth-order partial differential equations for u on U that have some linearity properties, without being linear in u and its derivatives. Such a differential equation is said to be *semilinear* if it can be expressed as

$$\sum_{|\alpha|=k} a_\alpha(x)\,\partial^\alpha u(x) + a_0(D^{k-1}u(x),\ldots,Du(x),u(x),x) = 0, \tag{1.12.1}$$

as in Section 1.1 of [29]. Here the sum is taken over all multi-indices α with $|\alpha| = k$, and $a_\alpha(x)$ should be a real-valued function on U for each such α. As before, a_0 may be considered as a real-valued function on

$$\mathbf{R}^{n^{k-1}} \times \cdots \times \mathbf{R}^n \times \mathbf{R} \times U. \tag{1.12.2}$$

Similarly, a kth order partial differential equation for u on U is said to be *quasilinear* if it can be expressed as

$$\sum_{|\alpha|=k} a_\alpha(D^{k-1}u(x), \ldots, Du(x), u(x), x)\, \partial^\alpha u(x) \\ + a_0(D^{k-1}u(x), \ldots, Du(x), u(x), x) = 0, \tag{1.12.3}$$

as in Section 1.1 of [29], and Section A of Chapter 1 of [32]. In this case, the coefficients a_α as well as a_0 may be considered as real-valued functions on (1.12.2).

A kth-order partial differential equation for u on U is said to be *fully nonlinear* if it depends nonlinearly on at least some of the kth-order derivatives of u, as in [29]. Of course, there are analogous notions for systems of partial differential equations.

As in Section 1.3, one may be interested in partial differential equations that are invariant under translations. In the case of a semilinear equation as in (1.12.1), this means that a_α is a constant for each multi-index α with $|\alpha| = k$, and that a_0 does not depend on x in the last variable. Thus a_0 may considered as a real-valued function on

$$\mathbf{R}^{n^{k-1}} \times \cdots \times \mathbf{R}^n \times \mathbf{R}. \tag{1.12.4}$$

Similarly, a quasilinear equation as in (1.12.3) is invariant under translations when the a_α's and a_0 do not depend on x in the last variable, so that they may be considered as real-valued functions on (1.12.4). There are analogous statements for systems of partial differential equations, as usual.

1.13 More on $\mathbf{R}^n$

Let n be a positive integer, and let U be an open subset of $\mathbf{R}^n$. Suppose that K is a compact subset of $\mathbf{R}^n$ such that

$$K \subseteq U. \tag{1.13.1}$$

Under these conditions, it is well known that there is a positive real number t such that for every $x \in K$, we have that

$$B(x, t) \subseteq U. \tag{1.13.2}$$

Suppose now that w is an element of U and r is a positive real number such that

$$\overline{B}(w, r) \subseteq U. \tag{1.13.3}$$

Remember that closed balls in $\mathbf{R}^n$ are closed and bounded, as in Sections 1.1 and 1.9, and thus compact. It follows that there is a positive real number ϵ such that

$$B(w, r + \epsilon) \subseteq U, \tag{1.13.4}$$

by the remarks in the preceding paragraph.

Let $y \in U$ be given, and let A be the set of positive real numbers r such that

$$B(y, r) \subseteq U. \tag{1.13.5}$$

Note that A is nonempty, because U is an open set, by hypothesis. Suppose that

$$U \neq \mathbf{R}^n, \tag{1.13.6}$$

so that there is a point z in the complement of U in $\mathbf{R}^n$. If $r \in A$, then we get that

$$r \leq |y - z|, \tag{1.13.7}$$

because $z \notin B(y, r)$. This means that $|y - z|$ is an upper bound for A in $\mathbf{R}$.

It is well known that A has a least upper bound or supremum ρ in $\mathbf{R}$ under these conditions. One can check that

$$B(y, \rho) \subseteq U, \tag{1.13.8}$$

because otherwise A would have an upper bound strictly less that ρ. We also have that

$$B(y, \rho + \epsilon) \not\subseteq U \tag{1.13.9}$$

for every $\epsilon > 0$, because ρ is an upper bound for A.

Using (1.13.9), we obtain that

$$\overline{B}(y, \rho) \not\subseteq U, \tag{1.13.10}$$

because of the earlier remarks. This means that

$$\partial B(y, \rho) \not\subseteq U, \tag{1.13.11}$$

because of (1.13.8).

It is easy to see that

$$\overline{B}(y, \rho) \subseteq \overline{U}, \tag{1.13.12}$$

using (1.13.8). Combining this with (1.13.11), we get that

$$\partial B(y, \rho) \cap \partial U \neq \emptyset. \tag{1.13.13}$$

In particular, $\partial U \neq \emptyset$, as mentioned in Section 1.8.

1.14 More on complex exponentials

Let a be a complex number. Suppose that f is a differentiable complex-valued function on the real line such that

$$f' = a\, f \tag{1.14.1}$$

on $\mathbf{R}$. This implies that

$$\frac{d}{dt}(\exp(-a\,t)\,f(t)) = 0 \tag{1.14.2}$$

on $\mathbf{R}$. Of course, this means that $\exp(-a\,t)\,f(t)$ is constant on $\mathbf{R}$. It follows that

$$f(t) = f(0)\,\exp(a\,t) \tag{1.14.3}$$

for every $t \in \mathbf{R}$.

Let n be a positive integer, and let b be an element of $\mathbf{C}^n$. Suppose that u is a complex-valued function on $\mathbf{R}^n$ such that for each $j = 1, \ldots, n$, the jth partial derivative of u exists at every point in $\mathbf{R}^n$, with

$$\frac{\partial u}{\partial x_j} = b_j\,u. \tag{1.14.4}$$

Under these conditions, one can check that

$$u(x) = u(0)\,\exp(b \cdot x) \tag{1.14.5}$$

for every $x \in \mathbf{R}^n$, using the remarks in the preceding paragraph.

Let a be a complex number again. If t is a positive real number, then put

$$t^a = \exp(a\,\log t). \tag{1.14.6}$$

This is a smooth complex-valued function of t on the set $\mathbf{R}_+$ of positive real numbers, with

$$\frac{d}{dt}(t^a) = a\,t^{a-1} \tag{1.14.7}$$

for every $t > 0$.

Let g be a differentiable complex-valued function on $\mathbf{R}_+$ such that

$$g'(t) = a\,t^{-1}\,g(t) \tag{1.14.8}$$

for every $t > 0$. Using this, we get that

$$\frac{d}{dt}(t^{-a}\,g(t)) = 0 \tag{1.14.9}$$

on $\mathbf{R}_+$. This implies that $t^{-a}\,g(t)$ is constant on $\mathbf{R}_+$, so that

$$g(t) = g(1)\,t^a \tag{1.14.10}$$

for every $t > 0$.

1.15 The dot product on $\mathbf{R}^n$

If $x, y \in \mathbf{R}^n$ for some positive integer n, then their *dot product* is defined by

$$x \cdot y = \sum_{j=1}^{n} x_j\, y_j, \tag{1.15.1}$$

which is consistent with the notation in Section 1.5. This is also known as the *standard inner product* on $\mathbf{R}^n$. Clearly

$$x \cdot y = y \cdot x \tag{1.15.2}$$

for every $x, y \in \mathbf{R}^n$.

Note that

$$x \cdot x = \sum_{j=1}^{n} x_j^2 = |x|^2 \tag{1.15.3}$$

for every $x \in \mathbf{R}^n$. This means that the standard Euclidean norm on $\mathbf{R}^n$ is the same as the norm associated to the standard inner product.

It is well known that

$$|x \cdot y| \le |x|\,|y| \tag{1.15.4}$$

for every $x, y \in \mathbf{R}^n$, which is a version of the *Cauchy–Schwarz inequality.* This can be used to obtain the triangle inequality for the standard Euclidean norm on $\mathbf{R}^n$, by a standard argument.

If $x, y \in \mathbf{R}^n$, then

$$\begin{aligned} |x+y|^2 = (x+y)\cdot(x+y) &= x \cdot x + x \cdot y + y \cdot x + y \cdot y \\ &= |x|^2 + 2\, x \cdot y + |y|^2. \end{aligned} \tag{1.15.5}$$

Thus

$$x \cdot y = (1/2)\,(|x+y|^2 - |x|^2 - |y|^2), \tag{1.15.6}$$

which is known as a *polarization identity.*

Let T be a linear mapping from $\mathbf{R}^n$ into itself. It is easy to see that

$$\ker T = \{x \in \mathbf{R}^n : T(x) = 0\} \tag{1.15.7}$$

is a linear subspace of $\mathbf{R}^n$, which is called the *kernel* of T.

One can check that T is one-to-one on $\mathbf{R}^n$ if and only if $\ker T = \{0\}$, using linearity. It is well known that T is one-to-one on $\mathbf{R}^n$ if and only if T maps $\mathbf{R}^n$ onto itself, which is to say that $T(\mathbf{R}^n) = \mathbf{R}^n$. In this case, the inverse mapping T^{-1} is linear on $\mathbf{R}^n$ too.

A one-to-one linear mapping T from $\mathbf{R}^n$ onto itself is said to be an *orthogonal transformation* if T preserves the standard inner product on $\mathbf{R}^n$. This means that

$$T(x) \cdot T(y) = x \cdot y \tag{1.15.8}$$

for every $x, y \in \mathbf{R}^n$. Under these conditions, the inverse mapping T^{-1} is an orthogonal transformation on $\mathbf{R}^n$ as well.

If we take $x = y$ in (1.15.8), then we get that

$$|T(x)| = |x|. \tag{1.15.9}$$

Conversely, if (1.15.9) holds for every $x \in \mathbf{R}^n$, then (1.15.8) holds for every $x, y \in \mathbf{R}^n$. This uses the linearity of T and the polarization identity (1.15.6).

Of course, if (1.15.9) holds for every $x \in \mathbf{R}^n$, then $\ker T = \{0\}$. This implies that T is one-to-one on $\mathbf{R}^n$, and thus that T maps $\mathbf{R}^n$ onto itself, as before.

If T is any linear mapping from $\mathbf{R}^n$ into itself, then it is well known that there is a unique linear mapping T' from $\mathbf{R}^n$ into itself such that

$$T(x) \cdot y = x \cdot T'(y) \tag{1.15.10}$$

for every $x, y \in \mathbf{R}^n$. More precisely, every linear mapping from $\mathbf{R}^n$ into itself corresponds to an $n \times n$ matrix of real numbers in a standard way. The matrix associated to T' in this way is the transpose of the matrix associated to T.

If T is an orthogonal transformation on $\mathbf{R}^n$, then one can check that T' is the same as the inverse of T. Conversely, if T is an invertible linear mapping on $\mathbf{R}^n$, with inverse equal to T', then one can verify that T is an orthogonal transformation on $\mathbf{R}^n$.

Chapter 2

Some related notions

2.1 The Laplacian

Let n be a positive integer. The *Laplacian* on $\mathbf{R}^n$ defined by

$$\Delta = \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2}. \tag{2.1.1}$$

Let $p(w)$ be the polynomial in n variables $w_1, \ldots, w_n$ with real coefficients defined by

$$p(w) = \sum_{j=1}^{n} w_j^2. \tag{2.1.2}$$

Observe that

$$p(\partial) = \sum_{j=1}^{n} \partial_j^2 = \Delta, \tag{2.1.3}$$

using the notation in Section 1.7.

If $w \in \mathbf{C}^n$, then

$$p(w) = w \cdot w, \tag{2.1.4}$$

using the notation in Section 1.5. If $w \in \mathbf{R}^n$, then

$$p(w) = |w|^2. \tag{2.1.5}$$

If $b \in \mathbf{C}^n$, then

$$\Delta(\exp(b \cdot x)) = (b \cdot b)\, \exp(b \cdot x), \tag{2.1.6}$$

as in Section 1.7. In particular,

$$\Delta(\exp(b \cdot x)) = 0 \tag{2.1.7}$$

when $b \cdot b = 0$.

Let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a twice continuuously-differentiable real or complex-valued function on U. We say that u is *harmonic* on U if it satisfies *Laplace's equation*

$$\Delta u = 0 \tag{2.1.8}$$

on U.

Let T be a linear mapping from $\mathbf{R}^n$ into itself. It is easy to see that T is continuous, so that the inverse image $T^{-1}(U)$ of U under T is an open subset of $\mathbf{R}^n$ too. If u is any twice continuously-differentiable function on U, then the composition $u \circ T$ of T and u is twice continuously differentiable on $T^{-1}(U)$.

If T is an orthogonal transformation on $\mathbf{R}^n$, then one can check that

$$\Delta(u \circ T) = (\Delta u) \circ T \tag{2.1.9}$$

on $T^{-1}(U)$. In particular, if u is harmonic on U, then $u \circ T$ is harmonic on $T^{-1}(U)$.

2.2 Two differential operators on $\mathbf{R}^2$

Consider the differential operators

$$L = \frac{1}{2}\Big(\frac{\partial}{\partial x_1} - i\,\frac{\partial}{\partial x_2}\Big) \tag{2.2.1}$$

and

$$\overline{L} = \frac{1}{2}\Big(\frac{\partial}{\partial x_1} + i\,\frac{\partial}{\partial x_2}\Big) \tag{2.2.2}$$

on $\mathbf{R}^2$. Observe that

$$L(x_1 + i\,x_2) = \overline{L}(x_1 - i\,x_2) = 1 \tag{2.2.3}$$

and

$$L(x_1 - i\,x_2) = \overline{L}(x_1 + i\,x_2) = 0. \tag{2.2.4}$$

If $z = x_1 + i\,x_2$ is considered as a complex variable, then L and $\overline{L}$ may be denoted $\partial/\partial z$ and $\partial/\partial\overline{z}$, respectively.

Let U be a nonempty open subset of $\mathbf{R}^2$, and let f be a continuously-differentiable complex-valued function on U. If

$$\overline{L}(f) = 0 \tag{2.2.5}$$

on U, then f is said to be *complex analytic* or *holomorphic* on U, as a function of the complex variable z. More precisely, (2.2.5) is equivalent to the usual *Cauchy–Riemann equations* for the real and imaginary parts of f. In this case,

$$f' = L(f) \tag{2.2.6}$$

is the usual complex derivative of f.

Note that

$$L\,\overline{L} = \overline{L}\,L = \frac{1}{4}\,\Delta. \tag{2.2.7}$$

Let u be a twice continuously-differentiable complex-valued function on U. This implies that $L(u)$ and $\overline{L}(u)$ are continuously differentiable on U, and we have that

$$L(\overline{L}(u)) = \overline{L}(L(u)) = \frac{1}{4}\,\Delta(u) \tag{2.2.8}$$

on U. If u is harmonic on U, then it follows that $L(u)$ is holomorphic on U.

If f is holomorphic on U, then it is well known that f is smooth on U, and twice continuously differentiable in particular. It follows that

$$\Delta(f) = 4\,L(\overline{L}(f)) = 0, \tag{2.2.9}$$

so that f is harmonic on U.

If f, g are any continuously-differentiable complex-valued functions on U, then

$$L(f\,g) = L(f)\,g + f\,L(g) \tag{2.2.10}$$

and

$$\overline{L}(f\,g) = \overline{L}(f)\,g + f\,\overline{L}(g) \tag{2.2.11}$$

on U, by the product rule. In particular, if f and g are holomorphic on U, then their product $f\,g$ is holomorphic on U.

If f is any continuously-differentiable complex-valued function on U again, then it is easy to see that

$$\overline{L(f)} = \overline{L}(\overline{f}) \tag{2.2.12}$$

on U. It follows that $L(f) = 0$ on U if and only if $\overline{f}$ is holomorphic on U.

Observe that

$$V = \{(x_1, -x_2) : (x_1, x_2) \in U\} \tag{2.2.13}$$

is an open subset of $\mathbf{R}^2$ as well. If f is a continuously-differentiable complex-valued function on U again, then

$$\widetilde{f}(x_1, x_2) = f(x_1, -x_2) \tag{2.2.14}$$

is a continuously-differentiable complex-valued function on V. One can check that f is holomorphic on U if and only if

$$\overline{\widetilde{f}(x_1, x_2)} = \overline{f(x_1, -x_2)} \tag{2.2.15}$$

is holomorphic on V.

2.3 Some complex first-order operators

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Suppose that $a_1, \ldots, a_n$ are n complex-valued functions on U. Thus

$$a(x) = (a_1(x), \ldots, a_n(x)) \tag{2.3.1}$$

may be considered as a mapping from U into $\mathbf{C}^n$. If u is a continuously-differentiable complex-valued function on U, then

$$L_a(u) = \sum_{j=1}^{n} a_j \, \frac{\partial u}{\partial x_j} \tag{2.3.2}$$

defines a complex-valued function on U.

Let v be another continuously-differentiable complex-valued function on U, so that the product of u and v is continuously-differentiable on U as well. Observe that

$$L_a(u\,v) = L_a(u)\,v + u\,L_a(v) \tag{2.3.3}$$

on U, by the product rule. If $L_a(u) = 0$ on U, then

$$L_a(u\,v) = u\,L_a(v) \tag{2.3.4}$$

on U. If $L_a(v) = 0$ on U too, then

$$L_a(u\,v) = 0 \tag{2.3.5}$$

on U.

Suppose now that $a_1, \ldots, a_n$ are continuously differentiable on U, and let $b_1, \ldots, b_n$ be another n continuously-differentiable complex-valued functions on U. Let b and L_b be as before, and put

$$c_j = L_a(b_j) - L_b(a_j) \tag{2.3.6}$$

for $j = 1, \ldots, n$. These are continuous complex-valued functions on U, and we let c and L_c be as before again.

Suppose that u is twice continuously differentiable on U. This implies that $L_a(u)$ and $L_b(u)$ are continuously differentiable on U, because the a_j's and b_j's are continuously differentiable on U, by hypothesis. It is easy to see that

$$L_a(L_b(u)) - L_b(L_a(u)) = L_c(u) \tag{2.3.7}$$

on U.

Put

$$\operatorname{Re} a(x) = (\operatorname{Re} a_1(x), \ldots, \operatorname{Re} a_n(x)) \tag{2.3.8}$$

and

$$\operatorname{Im} a(x) = (\operatorname{Im} a_1(x), \ldots, \operatorname{Im} a_n(x)) \tag{2.3.9}$$

for each $x \in U$, which define mappings from U into $\mathbf{R}^n$. If u is any continuously-differentiable complex-valued function on U, then $L_{\mathrm{Re}\, a}(u)$ and $L_{\mathrm{Im}\, a}(u)$ can be defined on U as before, and we have that

$$L_a(u) = L_{\mathrm{Re}\, a}(u) + i\, L_{\mathrm{Im}\, a}(u). \tag{2.3.10}$$

Of course, if u is real-valued on U, then $L_{\mathrm{Re}\, a}(u)$ and $L_{\mathrm{Im}\, a}(u)$ are real-valued on U as well. Otherwise,

$$\mathrm{Re}\, L_a(u) = L_{\mathrm{Re}\, a}(\mathrm{Re}\, u) - L_{\mathrm{Im}\, a}(\mathrm{Im}\, u) \tag{2.3.11}$$

and

$$\mathrm{Im}\, L_a(u) = L_{\mathrm{Re}\, a}(\mathrm{Im}\, u) + L_{\mathrm{Im}\, a}(\mathrm{Re}\, u) \tag{2.3.12}$$

on U. In particular, $L_a(u) = 0$ may be considered as a system of first-order homogeneous linear partial differential equations in the real and imaginary parts of u, with real coefficients.

2.4 Linear differential operators

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$ again. Also let N be a nonnegative integer, and for each multi-index α with order $|\alpha| \le N$, let a_α be a real or complex-valued function on U. If u is an N-times continuously-differentiable real or complex-valued function on U, then put

$$L(u) = \sum_{|\alpha| \le N} a_\alpha\, \partial^\alpha u \tag{2.4.1}$$

on U, where the sum is taken over all multi-indices α with $|\alpha| \le N$, as usual. This defines a differential operator on U, which can have variable coefficients.

Let r be a nonnegative integer, and suppose that a_α is r-times continuously differentiable on U for each multi-index α with $|\alpha| \le N$. If u is $(N + r)$-times continuously differentiable on U, then $L(u)$ is r-times continuously differentiable on U. In this case, L defines a linear mapping from $C^{N+r}(U, \mathbf{C})$ into $C^r(U, \mathbf{C})$. If a_α is real-valued on U for each α, then L defines a linear mapping from $C^{N+r}(U, \mathbf{R})$ into $C^r(U, \mathbf{R})$.

Similarly, suppose that a_α is infinitely differentiable on U for every multi-index α with $|\alpha| \le N$. If u is infinitely differentiable on U, then $L(u)$ is infinitely differentiable on U too. This means that L defines a linear mapping from $C^\infty(U, \mathbf{C})$ into itself. If a_α is real-valued on U for each α, then L defines a linear mapping from $C^\infty(U, \mathbf{R})$ into itself.

One can check that the coefficients a_α are uniquely determined by $L(u)$ for polynomials u of degree less than or equal to N. More precisely, a_0 is the same as $L(u)$ when $u(x) \equiv 1$ on U. If $\alpha \ne 0$, then a_α can be obtained from $L(x^\alpha)$ and the coefficients a_β with $|\beta| < |\alpha|$.

Let $\widetilde{N}$ be another nonnegative integer, and let b_β be a real or complex-valued function on U for each multi-index β with $|\beta| \le \widetilde{N}$. If u is an $\widetilde{N}$-times continuously-differentiable real or complex-valued function on U, then

$$\widetilde{L}(u) = \sum_{|\beta| \le \widetilde{N}} b_\beta \, \partial^\beta u \tag{2.4.2}$$

defines a real or complex-valued function on U, as appropriate.

Suppose that b_β is N-times continuously differentiable on U for each multi-index β with $|\beta| \le \widetilde{N}$. If u is $(N + \widetilde{N})$-times continuously differentiable on U, then $\widetilde{L}(u)$ is N-times continuously differentiable on U. This means that

$$L(\widetilde{L}(u)) \tag{2.4.3}$$

is defined as a real or complex-valued function on U, as appropriate.

Under these conditions, (2.4.3) may be expressed as

$$\widehat{L}(u) = \sum_{|\gamma| \le N + \widetilde{N}} c_\gamma \, \partial^\gamma u, \tag{2.4.4}$$

where c_γ is a real or complex-valued function on U for every multi-index γ with $|\gamma| \le N + \widetilde{N}$. More precisely, the c_γ's can be expressed as sums of products of the α_α's with the b_β's and their derivatives of order less than or equal to N.

Let r be a nonnegative integer again, and suppose that a_α is r-times continuously differentiable on U for every α with $|\alpha| \le N$. If the b_β's are $(N+r)$-times continuously differentiable on U for every β with $|\beta| \le \widetilde{N}$, then the c_γ's are r-times continuously differentiable on U for every γ with $|\gamma| \le N + \widetilde{N}$. If u is also $(N + \widetilde{N} + r)$-times continuously differentiable on U, then $\widetilde{L}(u)$ is $(N + r)$-times continuously differentiable on U, and $\widehat{L}(u)$ is r-times continuously differentiable on U.

Similarly, if the a_α's and b_β's are infinitely differentiable on U, then the c_γ's are infinitely differentiable on U. If u is infinitely differentiable on U too, then $\widetilde{L}(u)$ and $\widehat{L}(u)$ are infinitely differentiable on U as well.

2.5 Some remarks about polynomials

Let n be a positive integer, and let

$$p(x) = \sum_{|\alpha| \le N} a_\alpha \, x^\alpha \tag{2.5.1}$$

be a polynomial in the n variables $x_1, \ldots, x_n$ with complex coefficients, as in Section 1.7. Thus N is a nonnegative integer, $a_\alpha \in \mathbf{C}$ for each multi-index α with order $|\alpha| \le N$, and the sum is taken over all such multi-indices, as before.

If

$$p(x) = 0 \text{ for every } x \in \mathbf{R}^n, \tag{2.5.2}$$

then $\partial^\beta p(x) = 0$ for every $x \in \mathbf{R}^n$ and multi-index β. In particular, this implies that

$$\partial^\beta p(0) = 0 \text{ for every multi-index } \beta. \tag{2.5.3}$$

In this case, this means that

$$a_\alpha = 0 \text{ for every multi-index } \alpha,\ |\alpha| \le N. \tag{2.5.4}$$

If $x \in \mathbf{C}^n$, then $p(x)$ can be defined as a complex number as in (2.5.1). If (2.5.4) holds, then we get that

$$p(x) = 0 \text{ for every } x \in \mathbf{C}^n. \tag{2.5.5}$$

Let r be a positive real number, and suppose that

$$p(x) = 0 \text{ for every } x \in \mathbf{R}^n \text{ with } |x| < r. \tag{2.5.6}$$

This implies that $\partial^\beta p(x) = 0$ for every $x \in \mathbf{R}^n$ with $|x| < r$, and every multi-index β. It follows that (2.5.3) holds in particular under these conditions.

If $b \in \mathbf{R}^n$ or $\mathbf{C}^n$, then $p(x + b)$ can be expressed as a polynomial in x with complex coefficients too. If

$$p(x + b) = 0 \text{ for every } x \in \mathbf{R}^n \text{ with } |x| < r, \tag{2.5.7}$$

then the previous remarks imply that $p(x + b) = 0$ for every $x \in \mathbf{C}^n$. This is the same as saying that (2.5.5) holds.

Note that

$$\{x \in \mathbf{R}^n : p(x) = 0\} \tag{2.5.8}$$

is a closed set in $\mathbf{R}^n$, because p is continuous on $\mathbf{R}^n$. If this set contains a ball of positive radius, then (2.5.8) is equal to $\mathbf{R}^n$, as in the preceding paragraph.

Equivalently,

$$\{x \in \mathbf{R}^n : p(x) \neq 0\} \tag{2.5.9}$$

is an open set in $\mathbf{R}^n$. If this set is nonempty, then its intersection with any ball in $\mathbf{R}^n$ of positive radius is nonempty, as in the previous paragraph. This means that the closure of (2.5.9) in $\mathbf{R}^n$ is equal to $\mathbf{R}^n$ in this case, which is the same as saying that (2.5.9) is dense in $\mathbf{R}^n$, with respect to the standard Euclidean metric.

Of course, if $n = 1$ and $a_\alpha \neq 0$ for some α, then it is well known that $p(x) = 0$ for at most N points $x \in \mathbf{C}$.

2.6 Some remarks about $\mathbf{C}^n$

Let n be a positive integer, and consider the space $\mathbf{C}^n$ of n-tuples of complex numbers. If $z = (z_1, \ldots, z_n) \in \mathbf{C}^n$, then put

$$|z| = \Big(\sum_{j=1}^n |z_j|^2\Big)^{1/2}, \tag{2.6.1}$$

using the nonnegative square root on the right side, as usual. Here $|z_j|$ is the modulus of $z_j \in \mathbf{C}$ for each $j = 1, \ldots, n$, as in Section 1.4. We may call (2.6.1) the *standard Euclidean norm* on $\mathbf{C}^n$.

If $z, w \in \mathbf{C}^n$ and $t \in \mathbf{C}$, then $z + w$ and $t\,z$ may be defined as elements of $\mathbf{C}^n$ using coordinatewise addition and scalar multiplication. It is easy to see that

$$|t\,z| = |t|\,|z| \tag{2.6.2}$$

for every $z \in \mathbf{C}^n$ and $t \in \mathbf{C}$. One can check that

$$|z + w| \le |z| + |w| \tag{2.6.3}$$

for every $z, w \in \mathbf{C}^n$, using the analogous statements for the modulus of a complex number and the standard Euclidean norm on $\mathbf{R}^n$, as in Sections 1.1 and 1.4. The *standard Euclidean metric* on $\mathbf{C}^n$ is defined by

$$d(z, w) = |z - w| \tag{2.6.4}$$

for every $z, w \in \mathbf{C}^n$.

If $z, w \in \mathbf{C}^n$, then we put

$$\langle z, w \rangle = \langle z, w \rangle_{\mathbf{C}^n} = \sum_{j=1}^{n} z_j\,\overline{w_j}. \tag{2.6.5}$$

This is the *standard inner product* on $\mathbf{C}^n$. Observe that (2.6.5) is *Hermitian symmetric*, in the sense that

$$\langle z, w \rangle = \overline{\langle w, z \rangle} \tag{2.6.6}$$

for every $z, w \in \mathbf{C}^n$.

Of course,

$$\langle z, z \rangle = \sum_{j=1}^{n} |z_j|^2 = |z|^2 \tag{2.6.7}$$

for every $z \in \mathbf{C}^n$. This means that the standard Euclidean norm on $\mathbf{C}^n$ is the same as the norm associated to the standard inner product. It is well known that

$$|\langle z, w \rangle| \le |z|\,|w| \tag{2.6.8}$$

for every $z, w \in \mathbf{C}^n$, which is another version of the *Cauchy–Schwarz inequality.* This can also be used to obtain the triangle inequality for the standard Euclidean norm on $\mathbf{C}^n$.

Every $z \in \mathbf{C}^n$ can be expressed in a unique way as

$$z = x + i\,y, \tag{2.6.9}$$

with $x, y \in \mathbf{R}^n$. One can use this to identify $\mathbf{C}^n$ with $\mathbf{R}^{2n}$. Using this identification, the standard Euclidean norm and metric on $\mathbf{C}^n$ correspond exactly to their analogues on $\mathbf{R}^{2n}$. Similarly, one can check that the real part of the standard inner product on $\mathbf{C}^n$ corresponds to the standard inner product on $\mathbf{R}^{2n}$.

Consider the differential operators

$$L_j = \frac{\partial}{\partial z_j} = \frac{1}{2}\Big(\frac{\partial}{\partial x_j} - i\,\frac{\partial}{\partial y_j}\Big) \tag{2.6.10}$$

and

$$\overline{L}_j = \frac{\partial}{\partial \overline{z_j}} = \frac{1}{2}\Big(\frac{\partial}{\partial x_j} + i\,\frac{\partial}{\partial y_j}\Big) \tag{2.6.11}$$

on $\mathbf{C}^n$, as identified with $\mathbf{R}^{2n}$, for each $j = 1, \ldots, n$. Let U be a nonempty open subset of $\mathbf{C}^n$, which may be identified with an open subset of $\mathbf{R}^{2n}$. Also let f be a continuously-differentiable complex valued function on U, as an open subset of $\mathbf{R}^{2n}$. This means that the partial derivatives of f in x_j and y_j exist and are continuous on U for each $j = 1, \ldots, n$. If

$$\overline{L}_j(f) = 0 \tag{2.6.12}$$

on U for each $j = 1, \ldots, n$, then f is said to be *holomorphic* on U.

It is easy to see that products of holomorphic functions on U are holomorphic. The coordinate functions z_l are holomorphic on $\mathbf{C}^n$ for each $l = 1, \ldots, n$. It follows that polynomials in $z_1, \ldots, z_n$ with complex coefficients are holomorphic on $\mathbf{C}^n$.

2.7 Polynomials on $\mathbf{C}^n$

Let n be a positive integer, and let $p(z)$ be a polynomial with complex coefficients on $\mathbf{C}^n$. If $n = 1$, and $p(z)$ is not constant, then it is well known that $p(z) = 0$ for some $z \in \mathbf{C}$. More precisely, the number of zeros of p, counted with their multiplicities, is equal to the degree of p.

Suppose now that $n \geq 2$, and let us identify $\mathbf{C}^n$ with $\mathbf{C}^{n-1} \times \mathbf{C}$. If $z = (z_1, \ldots, z_n)$ is an element of $\mathbf{C}^n$, then $z' = (z_1, \ldots, z_{n-1}) \in \mathbf{C}^{n-1}$, and we identify z with $(z', z_n) \in \mathbf{C}^{n-1} \times \mathbf{C}$. Using this, we may express $p(z)$ as

$$p(z) = p(z', z_n) = \sum_{l=0}^{r} p_l(z')\, z_n^l, \tag{2.7.1}$$

where r is a nonnegative integer, and $p_l(z')$ is a polynomial on $\mathbf{C}^{n-1}$ for each $l = 0, \ldots, r$.

Suppose that $r \geq 1$, and that $p_r(z')$ is not identically 0 on $\mathbf{C}^{n-1}$. Otherwise, if $p_l(z')$ is identically 0 on $\mathbf{C}^{n-1}$ for each $l \geq 1$, then $p(z)$ would not depend on z_n, and we could consider it as a polynomial in a smaller number of variables.

Let $z' \in \mathbf{C}^{n-1}$ be given, and suppose that $p_r(z') \neq 0$. Under these conditions, (2.7.1) may be considered as a polynomial of degree r in z_n, which has r roots, with multiplicities, as before. There is an analogous statement as long as $p_l(z') \neq 0$ for some $l \geq 1$.

2.8 The Euler operator

Let n be a positive integer, and put

$$a_j(x) = x_j \tag{2.8.1}$$

for each $j = 1, \ldots, n$ and $x \in \mathbf{R}^n$. In this case,

$$a(x) = (a_1(x), \ldots, a_n(x)) = (x_1, \ldots, x_n) \tag{2.8.2}$$

is the identity mapping on $\mathbf{R}^n$.

Let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a continuously-differentiable real or complex-valued function on U. Let $L_a(u)$ be the continuous real or complex-valued function on U, as appropriate, defined by

$$(L_a(u))(x) = \sum_{j=1}^{n} x_j \, \frac{\partial u}{\partial x_j}(x) \tag{2.8.3}$$

for every $x \in U$, as in Section 2.3. The differential operator L_a is known as the *Euler operator.*

In this case, $(L_a(u))(x)$ is equal to the directional derivative of u at x in the direction x. Alternatively, if $x \in U$, then $u(t\,x)$ may be considered as a continuously-differentiable real or complex-valued function of t in an open subset of $\mathbf{R}$ that contains 1. The derivative of $u(t\,x)$ in t at 1 is equal to $(L_a(u))(x)$.

Let b be a complex number. A real or complex-valued function u on $\mathbf{R}^n \setminus \{0\}$ is said to be *homogeneous of degree* b if

$$u(t\,x) = t^b \, u(x) \tag{2.8.4}$$

for every $x \in \mathbf{R}^n \setminus \{0\}$ and $t \in \mathbf{R}_+$. If u is continuously differentiable on $\mathbf{R}^n \setminus \{0\}$, then (2.8.4) implies that

$$L_a(u) = b\,u \tag{2.8.5}$$

on $\mathbf{R}^n \setminus \{0\}$.

Let $x \in \mathbf{R}^n \setminus \{0\}$ be given. If u is continuously differentiable on $\mathbf{R}^n \setminus \{0\}$, then $u(t\,x)$ is continuously differentiable as a function of $t \in \mathbf{R}_+$. In this case,

$$\frac{d}{dt}(u(t\,x)) = \sum_{j=1}^{n} x_j \, (\partial_j u)(t\,x) = t^{-1} \, (L_a(u))(t\,x) \tag{2.8.6}$$

for every $t > 0$. If (2.8.5) holds, then we get that

$$\frac{d}{dt}(u(t\,x)) = b\,t^{-1} \, u(t\,x) \tag{2.8.7}$$

for every $t > 0$. This implies that (2.8.4) holds, as in Section 1.14.

Let U be a nonempty open subset of $\mathbf{R}^n$, and let t be a positive real number. Observe that

$$t^{-1}\, U = \{t^{-1}\, x : x \in U\} \tag{2.8.8}$$

is an open set in $\mathbf{R}^n$ too. If u is a continuously-differentiable real or complex-valued function on U, then $u(t\,x)$ is continuously differentiable as a function of x on $t^{-1}\,U$. The partial derivatives of $u(tx)$ are equal to

$$\frac{\partial}{\partial x_j}(u(t\,x)) = t\,(\partial_j u)(t\,x) \tag{2.8.9}$$

for each $j = 1, \ldots, n$ and $x \in t^{-1}\,U$.

Let us now take $U = \mathbf{R}^n \setminus \{0\}$, so that $t^{-1}\,U = U$ for every $t > 0$. If u is homogeneous of degree $b \in \mathbf{C}$ on $\mathbf{R}^n \setminus \{0\}$, then

$$\frac{\partial}{\partial x_j}(u(t\,x)) = t^b\,(\partial_j u)(x) \tag{2.8.10}$$

for each $j = 1, \ldots, n$. It follows that $\partial_j u$ is homogeneous of degree $b - 1$ on $\mathbf{R}^n \setminus \{0\}$ under these conditions.

One can check that

$$|t^b| = t^{\mathrm{Re}\, b} \tag{2.8.11}$$

for every $t > 0$ and $b \in \mathbf{C}$. Suppose that $\mathrm{Re}\, b > 0$, and let us interpret t^b as being equal to 0 when $t = 0$. Let us say that a real or complex-valued function u on $\mathbf{R}^n$ is *homogeneous of degree* b if (2.8.4) holds for every $x \in \mathbf{R}^n$ and nonnegative real number t. This means that $u(0) = 0$, and that u is homogeneous of degree b on $\mathbf{R}^n \setminus \{0\}$.

Let u, v be real or complex-valued functions on $\mathbf{R}^n \setminus \{0\}$ that are homogeneous of degrees $b, c \in \mathbf{C}$, respectively. It is easy to see that their product $u\,v$ is homogeneous of degree $b + c$ on $\mathbf{R}^n \setminus \{0\}$. Of course, there is an analogous statement for homogeneous functions on $\mathbf{R}^n$.

2.9 Some spaces of polynomials

Let n be a positive integer, and let $\mathcal{P}(\mathbf{R}^n, \mathbf{R})$ and $\mathcal{P}(\mathbf{R}^n, \mathbf{C})$ be the spaces of polynomials on $\mathbf{R}^n$ with real and complex coefficients, respectively. These are linear subspaces of the spaces $C^\infty(\mathbf{R}^n, \mathbf{R})$ and $C^\infty(\mathbf{R}^n, \mathbf{C})$ of smooth real and complex-valued functions on $\mathbf{R}^n$.

Let N be a nonnegative integer, and suppose that a_α is a polynomial on $\mathbf{R}^n$ for each multi-index α of order $|\alpha| \le N$. Under these conditions,

$$L = \sum_{|\alpha| \le N} a_\alpha\,\partial^\alpha \tag{2.9.1}$$

defines a differential operator on $\mathbf{R}^n$ with polynomial coefficients. Of course, the sum is taken over all mutli-indices α with $|\alpha| \le N$, as usual.

It is easy to see that L defines a linear mapping from $\mathcal{P}(\mathbf{R}^n, \mathbf{C})$ into itself. If a_α is a polynomial with real coefficients for each α, then L maps $\mathcal{P}(\mathbf{R}^n, \mathbf{R})$ into itself.

The composition of two differential operators on $\mathbf{R}^n$ with polynomial coefficients is a differential operator with polynomial coefficients too, as in Section 2.4.

Let k be a nonnegative integer. If α is a multi-index of order $|\alpha| = k$, then the monomial x^α is homogeneous of degree k as a real-valued function on $\mathbf{R}^n$. If a polynomial p on $\mathbf{R}^n$ can be expressed as a finite linear combination of monomials x^α with $|\alpha| = k$, then it follows that p is homogeneous of degree k on $\mathbf{R}^n$. Conversely, if a polynomial on $\mathbf{R}^n$ is homogeneous of degree k on $\mathbf{R}^n$, then one can check that it is of this form.

Let $\mathcal{P}_k(\mathbf{R}^n, \mathbf{R})$ and $\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$ be the space of polynomials on $\mathbf{R}^n$ with real and complex coefficients, respectively, that are homogeneous of degree k. These are linear subspaces of $\mathcal{P}(\mathbf{R}^n, \mathbf{R})$ and $\mathcal{P}(\mathbf{R}^n, \mathbf{C})$, respectively.

Let N be a nonnegative integer, and let p be a polynomial on $\mathbf{R}^n$ with real or complex coefficients of degree less than or equal to N. It is easy to see that p can be expressed in a unique way as a sum of homogeneous polynomials of degrees from 0 to N.

If a real or complex-valued function u on $\mathbf{R}^n$ is k-times continuously differentiable and homogeneous of degree k, then one can check that u is equal to its degree k Taylor approximation at the origin.

2.10 Polynomials on $\mathbf{R}^2$

Of course,

$$z = x_1 + i\,x_2,\ \overline{z} = x_1 - i\,x_2 \tag{2.10.1}$$

are homogeneous polynomials of degree 1 with complex coefficients on $\mathbf{R}^2$. We also have that

$$x_1 = (1/2)\,(z + \overline{z}),\ x_2 = (-i/2)\,(z - \overline{z}). \tag{2.10.2}$$

This means that every polynomial on $\mathbf{R}^2$ with complex coefficients corresponds to a polynomial in z and $\overline{z}$ with complex coefficients, and that every polynomial in z and $\overline{z}$ with complex coefficients determines a polynomial in x_1, x_2 with complex coefficients. More precisely, homogeneous polynomials in x_1, x_2 correspond to homogeneous polynomials in z, $\overline{z}$ of the same degree in this way.

Let $\partial/\partial z$ and $\partial/\partial \overline{z}$ be as in Section 2.2, and remember that $(\partial/\partial z)(z) = (\partial/\partial \overline{z})(\overline{z}) = 1$ and $(\partial/\partial z)(\overline{z}) = (\partial/\partial \overline{z})(z) = 0$. If j is a positive integer, then it follows that

$$\frac{\partial}{\partial z}(\overline{z}^j) = \frac{\partial}{\partial \overline{z}}(z^j) = 0, \tag{2.10.3}$$

by the product rules for these operators. Similarly,

$$\frac{\partial}{\partial z}(z^j) = j\,z^{j-1} \tag{2.10.4}$$

and

$$\frac{\partial}{\partial \overline{z}}(\overline{z}^j) = j\,\overline{z}^{j-1}. \tag{2.10.5}$$

If l is another positive integer, then we get that

$$\begin{aligned}\Delta(z^j\,\overline{z}^l) &= 4\,\frac{\partial}{\partial z}\frac{\partial}{\partial \overline{z}}(z^j\,\overline{z}^l) \\ &= 4\Big(\frac{\partial}{\partial z}(z^j)\Big)\Big(\frac{\partial}{\partial \overline{z}}(\overline{z}^l)\Big) = 4\,j\,l\,z^{j-1}\,\overline{z}^{l-1}.\end{aligned} \tag{2.10.6}$$

Note that $z^j\,\overline{z}^l$ is harmonic when j or l is equal to 0.

If k is a nonnegative integer, then a homogeneous polynomial of degree k on $\mathbf{R}^2$ with complex coefficients may be expressed as

$$\sum_{j=0}^{k} c_j\,z^j\,\overline{z}^{k-j} \tag{2.10.7}$$

for some complex coefficients c_j, $0 \le j \le k$. If $j \le k/2$, then

$$z^j\,\overline{z}^{k-j} = |z|^{2j}\,\overline{z}^{k-2\,j}. \tag{2.10.8}$$

If $j \ge k/2$, then

$$z^j\,\overline{z}^{k-j} = z^{2\,j-k}\,|z|^{2j} \tag{2.10.9}$$

If

$$|z|^2 = x_1^2 + x_2^2 = 1, \tag{2.10.10}$$

then we get that

$$z^j\,\overline{z}^{k-j} = \overline{z}^{k-2\,j} \tag{2.10.11}$$

when $j \le k/2$, and that

$$z^j\,\overline{z}^{k-j} = z^{2\,j-k} \tag{2.10.12}$$

when $j \ge k/2$. It follows that there is a harmonic polynomial on $\mathbf{R}^2$ of degree less than or equal to k that is equal to (2.10.7) on the unit circle. Using this, it is easy to see that every polynomial on $\mathbf{R}^2$ agrees with a harmonic polynomial on the unit circle. This corresponds to some remarks on p138 of [125].

Let U be a nonempty bounded open subset of $\mathbf{R}^n$ for some positive integer n. If f is a continuous real or complex-valued function on ∂U, then the *Dirichlet problem* asks one to find a continuous real or complex-valued function u on $\overline{U}$, as appropriate, such that

$$u = f \text{ on } \partial U, \tag{2.10.13}$$

and u is harmonic on U. The remarks in the preceding paragraph show that if $n = 2$, U is the open unit disk in $\mathbf{R}^2$, and f is the restriction to the unit circle of a polynomial on $\mathbf{R}^2$, then one can take u to be the restriction to $\overline{U}$ of a harmonic polynomial on $\mathbf{R}^2$.

2.11 Poisson's equation

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. If f is a real or complex-valued function on U, then one might like to find a real or complex-valued function u on U, as appropriate, such that

$$\Delta u = f \tag{2.11.1}$$

on U. This is *Poisson's equation*, as on p193 of [10], and p20 of [29].

Of course, one might like u to be twice continuously-differentiable on U, which would mean that f should be continuous on U. There are extended formulations of the equation, which allow for less regularity. One may also be interested in additional boundary conditions on u.

If f is a homogeneous polynomial of degree $k \geq 0$ on $\mathbf{R}^2$, then one can find a homogeneous polynomial u of degree $k+2$ on $\mathbf{R}^2$ that satisfies (2.11.1) on $\mathbf{R}^2$, as in the previous section. It follows that if f is any polynomial on $\mathbf{R}^2$, then one can find a polynomial u on $\mathbf{R}^2$ that satisfies (2.11.1).

Let g be a real or complex-valued function on ∂U. Another version of the *Dirichlet problem* asks one to find a real or complex-valued function u on $\overline{U}$, as appropriate, such that (2.11.1) holds on U and

$$u = g \text{ on } \partial U, \tag{2.11.2}$$

as in Section C of Chapter 2 of [32]. One might like u to be continuous on $\overline{U}$, so that g should be continuous on ∂U. There are extended versions of this too.

The case where

$$u = 0 \text{ on } \partial U \tag{2.11.3}$$

is known as *Dirichlet boundary conditions.* If one can solve Poisson's equation (2.11.1) without restrictions on u on ∂U, and if one can solve the Dirchlet problem for harmonic functions on U with arbitrary boundary values, then one can get a solution to Poisson's equation on U with prescribed boundary values. Similarly, if one can solve Poisson's equation U with Dirichlet boundary conditions, then one can try to use that to solve the Dirichlet problem for harmonic functions on U.

2.12 An interesting inner product

Let n be a positive integer, and let p, q be polynomials on $\mathbf{R}^n$ with complex coefficients. Note that the complex conjugate $\overline{q}$ of q is a polynomial on $\mathbf{R}^n$ too. Put

$$\langle p, q \rangle = \langle p, q \rangle_{\mathcal{P}(\mathbf{R}^n, \mathbf{C})} = (p(\partial)(\overline{q}))(0), \tag{2.12.1}$$

where $p(\partial)$ is as in Section 1.7.

If p and q are homogeneous polynomials of the same degree k, then $p(\partial)(\overline{q})$ is a constant, and (2.12.1) is the same as

$$\langle p, q \rangle = \langle p, q \rangle_{\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})} = p(\partial)(\overline{q}). \tag{2.12.2}$$

This is the definition that is used in the proof of Proposition 2.47 in Section G of Chapter 2 of [32], on p175 of [76], on p69 of [119], and on p139 of [125]. If p and q are homogeneous polynomials of different degrees, then it is easy to see that

$$\langle p, q \rangle = 0. \tag{2.12.3}$$

More precisely, if α, β are multi-indices, then

$$\begin{aligned} \langle x^\alpha, x^\beta \rangle &= \alpha! \quad \text{when } \alpha = \beta \\ &= 0 \quad \text{when } \alpha \neq \beta. \end{aligned} \tag{2.12.4}$$

Using (2.12.4), one can check that

$$\langle p, q \rangle = \overline{\langle q, p \rangle} \tag{2.12.5}$$

for all polynomials p, q on $\mathbf{R}^n$ with complex coefficients. Of course, (2.12.1) is linear in p over the complex numbers, and conjugate-linear in q. If $p(x) = \sum_{|\alpha| \le N} a_\alpha \, x^\alpha$ for some nonnegative integer N and complex coefficients a_α, then

$$\langle p, p \rangle = \sum_{|\alpha| \le N} |a_\alpha|^2 \, \alpha!. \tag{2.12.6}$$

In particular, this is strictly positive, except when $p = 0$. It follows that (2.12.1) defines an inner product on $\mathcal{P}(\mathbf{R}^n, \mathbf{C})$, as a vector space over the complex numbers, as in the proof of Proposition 2.47 in Section G of Chapter 2 of [32], and on p176 of [76], p69 of [119], and p139 of [125].

Note that the Laplacian maps $\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$ into $\mathcal{P}_{k-2}(\mathbf{R}^n, \mathbf{C})$ for every integer $k \ge 2$. Let us use this inner product to show that

$$\Delta(\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})) = \mathcal{P}_{k-2}(\mathbf{R}^n, \mathbf{C}) \tag{2.12.7}$$

when $k \ge 2$, as in the proof of Proposition 2.47 in Section G of Chapter 2 of [32], and of Theorem 2.1 on p139 of [125]. Suppose that $q \in \mathcal{P}_{k-2}(\mathbf{R}^n, \mathbf{C})$ is orthogonal to every element of $\Delta(\mathcal{P}_k(\mathbf{R}^n, \mathbf{C}))$ with respect to this inner product, so that

$$\langle q, \Delta(p) \rangle = 0 \tag{2.12.8}$$

for every $p \in \mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$. This means that

$$q(\partial)(\Delta(\overline{p})) = 0, \tag{2.12.9}$$

because $\overline{\Delta(p)} = \Delta(\overline{p})$. This is the same as saying that

$$\Delta(q(\partial)(\overline{p})) = 0. \tag{2.12.10}$$

If we take $p(x) = |x|^2 \, q(x)$, then we get that

$$\langle p, p \rangle = p(\partial)(\overline{p}) = \Delta(q(\partial)(\overline{p})) = 0. \tag{2.12.11}$$

This implies that $p = 0$, as before. This means that $q = 0$, because of the way that we chose p. It follows that (2.12.7) holds, by standard arguments in linear algebra. This also uses the fact that $\mathcal{P}_{k-2}(\mathbf{R}^n, \mathbf{C})$ has finite dimension, as a vector space over $\mathbf{C}$.

2.13 An orthogonality argument

Let us continue with the same notation as in the previous section. If k is any nonnegative integer, then let

$$\mathcal{A}_k = \{p \in \mathcal{P}_k(\mathbf{R}^n, \mathbf{C}) : \Delta(p) = 0\} \tag{2.13.1}$$

be the space of homogeneous polynomials on $\mathbf{R}^n$ of degree k with complex coefficients that are harmonic, which is a linear subspace of $\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$. Of course, this is the same as $\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$ when $k = 0$ or 1. If $k \geq 2$, then put

$$\mathcal{B}_k = \{|x|^2\, q(x) : q \in \mathcal{P}_{k-2}(\mathbf{R}^n, \mathbf{C})\}, \tag{2.13.2}$$

which is also a linear subspace of $\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$.

Let $k \geq 2$ and $p \in \mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$ be given, and put

$$r_q(x) = |x|^2\, q(x) \tag{2.13.3}$$

for every $q \in \mathcal{P}_{k-2}(\mathbf{R}^n, \mathbf{C})$. Thus $r_q \in \mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$, and

$$\langle r_q, p\rangle = r_q(\partial)(\overline{p}) = q(\partial)(\Delta(\overline{p})) = \langle q, \Delta(p)\rangle. \tag{2.13.4}$$

Observe that

$$\langle q, \Delta(p)\rangle = 0 \tag{2.13.5}$$

for every $q \in \mathcal{P}_{k-2}(\mathbf{R}^n, \mathbf{C})$ if and only if $\Delta(p) = 0$. It follows that

$$\langle r_q, p\rangle = 0 \tag{2.13.6}$$

for every $q \in \mathcal{P}_{k-2}(\mathbf{R}^n, \mathbf{C})$ if and only if $\Delta(p) = 0$. This means that $\mathcal{A}_k$ is the orthogonal complement of $\mathcal{B}_k$ in $\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$ with respect to this inner product, as in the proof of Proposition 2.47 in Section G of Chapter 2 of [32], on p69 of [119], and p140 of [125].

This implies that every element of $\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$ can be expressed in a unique way as a sum of elements of $\mathcal{A}_k$ and $\mathcal{B}_k$, by standard arguments in linear algebra. More precisely, this uses the fact that $\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$ is a finite-dimensional vector space over $\mathbf{C}$. This also corresponds to Proposition 5.5 on p76 of [10].

We can repeat the process, to get that every element of $\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$ can be expressed as

$$\sum_{j=0}^{l} |x|^{2\,j}\, p_j(x), \tag{2.13.7}$$

where $2\,l \leq k$, and $p_j \in \mathcal{P}_{k-2\,j}(\mathbf{R}^n, \mathbf{C})$ is harmonic for each $j = 1, \ldots, l$. This corresponds to Theorem 5.7 on p77 of [10], Corollary 2.48 in Section G of Chapter 2 of [32], Proposition 4.1.1 on p176 of [76], some remarks on p70 of [119], and Theorem 2.1 on p139 of [125].

One can use this to get that every polynomial on $\mathbf{R}^n$ agrees with a harmonic polynomial on the unit sphere, as in some remarks on p77 of [10], Corollary 2.50 in Section G of Chapter 2 of [32], Corollary 4.1.2 on p177 of [76], mentioned on p70 of [119], and Corollary 2.2 on p140 of [125]. This corresponds to the Dirichlet problem on the open unit ball in $\mathbf{R}^n$, for the restriction to the unit sphere of a polynomial on $\mathbf{R}^n$.

2.14 The binomial theorem

If m is a positive integer and x, y are real or complex numbers, then the *binomial theorem* states that

$$(x+y)^m = \sum_{j=0}^{m} \binom{m}{j} x^j \, y^{m-j}, \tag{2.14.1}$$

where

$$\binom{m}{j} = \frac{m!}{j! \, (m-j)!} \tag{2.14.2}$$

is the usual *binomial coefficient* for each $j = 0, 1, \ldots, n$. If we take $y = 1$ in (2.14.1), then we get that

$$(x+1)^m = \sum_{j=0}^{m} \binom{m}{j} x^j. \tag{2.14.3}$$

Conversely, (2.14.1) can be obtained from (2.14.3) by replacing x with x/y when $y \neq 0$.

Of course, it is easy to see that $(x + 1)^m$ can be expressed as a sum of positive integer multiples of x^j, $0 \le j \le m$. To get that the multiples are given by binomial coefficients as before, one can look at the jth derivative of $(x+1)^m$ at 0 for each $j = 0, 1, \ldots, m$. In particular, this shows that the binomial coefficients are positive integers. Alternatively, one can verify (2.14.3) more directly, using induction on m.

One can expand $(x+1)^m$ into a sum of 2^m terms, each of which is a product of m factors, where every factor is equal to x or to 1. The coefficient of x^j in $(x + 1)^m$ is the same as the number of these terms with exactly j factors of x, and $m - j$ factors of 1. This is also the same as the number of subsets of $\{1, \ldots, m\}$ with exactly j elements.

Let k and n be positive integers. It is well known that the number of multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n)$ with order $|\alpha| = k$ is equal to

$$\binom{n+k-1}{k} = \binom{n+k-1}{n-1}. \tag{2.14.4}$$

This corresponds to Problem 2 in Section 1.5 of [29]. Equivalently, this is the dimension of the spaces $\mathcal{P}_k(\mathbf{R}^n, \mathbf{R})$, $\mathcal{P}_k(\mathbf{R}^n, \mathbf{C})$ of homogeneous polynomials of degree k on $\mathbf{R}^n$ with real or complex coefficients, as vector spaces over $\mathbf{R}$ or $\mathbf{C}$, as appropriate. This is mentioned on p78 of [10], in Proposition 2.52 of Section G of Chapter 2 of [32], on p174f of [76], and on p139 of [125].

If $x \in \mathbf{R}^n$ or $\mathbf{C}^n$, then the *multinomial theorem* states that

$$(x_1 + \cdots + x_n)^k = \sum_{|\alpha|=k} \binom{|\alpha|}{\alpha} x^\alpha, \tag{2.14.5}$$

as in Problem 3 in Section 1.5 of [29]. More precisely, the sum is taken over all multi-indices α with order $|\alpha| = k$, and we put

$$\binom{|\alpha|}{\alpha} = \frac{|\alpha|!}{\alpha!}, \tag{2.14.6}$$

where $\alpha!$ is as in Section 1.10. Note that (2.14.5) is trivial when $n = 1$, and that the $n = 2$ case is the same as the binomial theorem.

Let α be a multi-index, and let $x, y \in \mathbf{R}^n$ or $\mathbf{C}^n$ be given. One can check that

$$(x + y)^\alpha = \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\,\gamma!}\, x^\beta\, y^\gamma, \tag{2.14.7}$$

where the sum is taken over all multi-indices β, γ with $\beta + \gamma = \alpha$. This is the same as the binomial theorem when $n = 1$.

If $p(x)$ is a polynomial in $x_1, \ldots, x_n$ with real or complex coefficients and $b \in \mathbf{R}^n$ or $\mathbf{C}^n$, then $p(x + b)$ can be expressed as a polynomial in x with real or complex coefficients, as appropriate, as in Section 2.5. One can use (2.14.7) to get a more precise version of this.

2.15 Leibniz' formula

Let n be a positive integer, and let α, β be multi-indices. If

$$\beta_j \le \alpha_j \tag{2.15.1}$$

for each $j = 1, \ldots, n$, then put

$$\beta \le \alpha, \tag{2.15.2}$$

as in Problem 4 in Section 1.5 of [29]. Equivalently, this means that $\alpha - \beta$ is a multi-index too. In this case, we put

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta!\,(\alpha - \beta)!}, \tag{2.15.3}$$

as in [29].

Let u, v be smooth real-valued functions on $\mathbf{R}^n$. *Leibniz' formula* states that

$$\partial^\alpha (u\, v) = \sum_{\beta \le \alpha} \binom{\alpha}{\beta} (\partial^\beta u)\, (\partial^{\alpha-\beta} v), \tag{2.15.4}$$

as in Problem 4 in Section 1.5 of [29]. More precisely, the sum is taken over all multi-indices β with $\beta \le \alpha$. Of course, this also works when u, v are $|\alpha|$-times continuously-differentiable real or complex-valued functions on a nonempty open subset of $\mathbf{R}^n$. If $|\alpha| = 1$, then this reduces to the usual product rule for partial derivatives.

Let U be a nonempty open subset of $\mathbf{R}^n$, and let N, $\widetilde{N}$ be nonnegative integers. Suppose that for each multi-index α with order $|\alpha| \le N$, a_α is a real or complex-valued function on U. This permits us to define the corresponding differential operator

$$L = \sum_{|\alpha| \le N} a_\alpha \, \partial^\alpha, \tag{2.15.5}$$

as in Section 2.4. Similarly, suppose that b_β is a real or complex-valued function on U for each multi-index β with $|\beta| \le \widetilde{N}$. This permits us to define the differential operator

$$L = \sum_{|\beta| \le \widetilde{N}} b_\beta \, \partial^\beta, \tag{2.15.6}$$

as before.

Suppose that b_β is N-times continuously differentiable on U for each multi-index β with $|\beta| \le \widetilde{N}$. If n is an $(N + \widetilde{N})$-times continuously-differentiable real or complex-valued function on U, then $\widetilde{L}(u)$ is N-times continuously differentiable on U, so that $L(\widetilde{L}(u))$ is defined as a real or complex-valued function on U, as appropriate, as in Section 2.4. In fact, this can be expressed as $\widehat{L}(u)$, where

$$\widehat{L} = \sum_{|\gamma| \le N + \widetilde{N}} c_\gamma \, \partial^\gamma, \tag{2.15.7}$$

and c_γ is a real or complex-valued function on U for every multi-index γ with $|\gamma| \le N + \widetilde{N}$, as before. Remember that the c_γ's can be expressed in terms of sums of products of the a_α's with the b_β's and their derivatives of order less than or equal to N. This can be described more precisely using (2.15.4).

Chapter 3

Some integrals and other matters

3.1 Eigenfunctions of differential operators

Let n be a positive integer, let U be a nonempty open set in $\mathbf{R}^n$. Also let N be a positive integer, and let a_α be a complex-valued function on U for each multi-index α with order $|\alpha| \le N$. If f is an N-times continuously-differentiable complex-valued function on U, then put

$$L(f) = \sum_{|\alpha| \le N} a_\alpha \, \partial^\alpha f \tag{3.1.1}$$

on U.

We say that f is an *eigenfunction* for L with *eigenvalue* $\lambda \in \mathbf{C}$ if

$$L(f) = \lambda \, f \tag{3.1.2}$$

on U. One may wish to ask that f satisfy additional boundary conditions or other restrictions, depending on the circumstances.

Let us identify $\mathbf{R}^n \times \mathbf{R}$ with $\mathbf{R}^{n+1}$, so that

$$V = U \times \mathbf{R} \tag{3.1.3}$$

may be considered as an open subset of $\mathbf{R}^{n+1}$. Put

$$u(x,t) = \exp(\lambda \, t) \, f(x) \tag{3.1.4}$$

for every $x \in U$ and $t \in \mathbf{R}$, which defines an N-times continuously-differentiable complex-valued function on V. It is easy to see that

$$\frac{\partial u}{\partial t} = L(u) \tag{3.1.5}$$

on V. We also have that

$$u(x,0) = f(x) \tag{3.1.6}$$

for every $x \in U$.

Suppose that $\mu \in \mathbf{C}$ satisfies

$$\mu^2 = \lambda, \tag{3.1.7}$$

and put

$$v(x,t) = \exp(\mu\, t)\, f(x) \tag{3.1.8}$$

for every $x \in U$ and $t \in \mathbf{R}$. Observe that v and all of its partial derivatives in t are N-times continuously-differentiable on V, and that

$$\frac{\partial^2 v}{\partial t^2} = L(v) \tag{3.1.9}$$

on V. In addition,

$$v(x,0) = f(x) \tag{3.1.10}$$

and

$$\frac{\partial v}{\partial t}(x,0) = \mu\, f(x) \tag{3.1.11}$$

for every $x \in U$.

Similarly,

$$w(x,t) = \exp(-\mu\, t)\, f(x) \tag{3.1.12}$$

and all of its partial derivaitves in t are N-times continuously-differentiable on V. As before,

$$\frac{\partial^2 w}{\partial t^2} = L(w) \tag{3.1.13}$$

on V, because $(-\mu)^2 = \lambda$ too. In this case,

$$w(x,0) = f(x) \tag{3.1.14}$$

and

$$\frac{\partial w}{\partial t}(x,0) = -\mu\, f(x) \tag{3.1.15}$$

for every $x \in U$.

Of course, one may consider multiple eigenfunctions of L on U, with possibly different eigenvalues, to get more solutions of partial differential equations like these on V. One may also consider infinite sums, under suitable conditions.

Mark Kac' famous question of whether one can hear the shape of a drum involves eigenvalues for the Laplacian. See [22, 43, 44, 45, 68] for more information.

3.2 The spherical Laplacian

Let n be a positive integer, and let

$$S^{n-1} = \{x \in \mathbf{R}^n : |x| = 1\} \tag{3.2.1}$$

be the unit sphere in $\mathbf{R}^n$, with respect to the standard Euclidean norm. If u is a complex-valued function on $\mathbf{R}^n \setminus \{0\}$ that is homogeneous of degree $b \in \mathbf{C}$, then

$$u(x) = |x|^b \, u(|x|^{-1} \, x) \tag{3.2.2}$$

for every $x \in \mathbf{R}^n \setminus \{0\}$. In particular, this means that u is uniquely determined by its restriction to the unit sphere. Similarly, any real or complex-valued function on the unit sphere can be extended to a function on $\mathbf{R}^n \setminus \{0\}$ that is homogeneous of any given degree in $\mathbf{C}$.

Suppose that $n \geq 2$, and let u be a twice continuously-differentiable real or complex-valued function on $\mathbf{R}^n \setminus \{0\}$ that is homogeneous of degree 0. The restriction of u to the unit sphere may be considered as a twice continuously-differentiable function on $\mathbf{S}^{n-1}$. Smoothness of functions on S^{n-1} can be defined in terms of suitable local coordinates, but it is more convenient for us to look at it here in terms of smoothness of homogeneous extensions to $\mathbf{R}^n \setminus \{0\}$.

The *spherical Laplacian* of u is the function $\Delta_S u$ defined on S^{n-1} by

$$\Delta_S u = \Delta u \text{ on } S^{n-1}. \tag{3.2.3}$$

Note that Δu is homogeneous of degree -2 on $\mathbf{R}^n \setminus \{0\}$, as in Section 2.8. Thus

$$|x|^2 \, (\Delta u)(x) \tag{3.2.4}$$

is homogeneous of degree 0 on $\mathbf{R}^n \setminus \{0\}$. Of course, this is the same as (3.2.3) on S^{n-1}.

Now let v be a twice continuously-differentiable complex-valued function on $\mathbf{R}^n \setminus \{0\}$ that is homogeneous of degree $b \in \mathbf{C}$. Observe that

$$|x|^{-b} \, v(x) \tag{3.2.5}$$

is a twice continuously-differentiable function on $\mathbf{R}^n \setminus \{0\}$ that is homogeneous of degree 0 and equal to v on S^{n-1}. The spherical Laplacian of the restriction of v to S^{n-1} is

$$(\Delta_S v)(x) = \Delta(|x|^{-b} \, v(x)) \text{ on } S^{n-1}. \tag{3.2.6}$$

Suppose that p is a homogeneous polynomial of degree $k \geq 0$ on $\mathbf{R}^n$, and that p is harmonic on $\mathbf{R}^n$. It is well known that the spherical Laplacian of the restriction of p to S^{n-1} satisfies

$$\Delta_S p = -k \, (k + n - 2) \, p \text{ on } S^{n-1}, \tag{3.2.7}$$

as in Lemma 2.61 in Section G of Chapter 2 of [32], and on p70 of [119].

3.3 Connected components

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. It is well known that U can be expressed in a unique way as a union of a family of pairwise-disjoint nonempty connected open subsets of $\mathbf{R}^n$. These nonempty connected open sets are called the *connected components* of U. If U is connected, then U is the only connected component of itself.

In fact, one can define the notion of connected components for any subset E of $\mathbf{R}^n$, as well as subsets of arbitrary metric spaces or topological spaces. One can show that the connected components of E are relatively closed in E, but they are not necessarily relatively open in E.

The connected component of E that contains a point $x \in E$ can be obtained by taking the union of all of the connected subsets of E that contain x. One can verify that this is a connected set too. By construction, this is the largest possible connected subset of E that contains x.

Connected components of open subsets of $\mathbf{R}^n$ are open sets, basically because $\mathbf{R}^n$ is locally connected. This follows from the connectedness of open balls in $\mathbf{R}^n$.

More precisely, $\mathbf{R}^n$ is locally path connected, because every point in an open ball in $\mathbf{R}^n$ can be connected to the center of the ball by a line segment, which is contained in the ball. Because of this, the connected components of U are the same as the *path connected components*. These can be defined by saying that $x, y \in U$ are in the same path connected component of U when there is a continuous path in U connecting x and y.

If V is a connected component of U, then it is easy to see that

$$\overline{V} \subseteq \overline{U}, \tag{3.3.1}$$

because $V \subseteq U$. In particular, this implies that

$$\partial V \subseteq \overline{U}. \tag{3.3.2}$$

One can check that

$$\partial V \subseteq \partial U \tag{3.3.3}$$

under these conditions. Otherwise, if there is an element of ∂V in U, then one can show that that point should be in V, to get a contradiction.

Suppose that $U \neq \mathbf{R}^n$, so that $V \neq \mathbf{R}^n$. This implies that $\partial V \neq \emptyset$, and thus that ∂V has an element in ∂U.

3.4 Smoothness near the boundary

Let n be a positive integer, and let V be a nonempty open subset of $\mathbf{R}^n$. We shall sometimes be concerned with smoothness properties of functions on $\overline{V}$, including on the boundary. Suppose that U is an open subset of $\mathbf{R}^n$ with

$$\overline{V} \subseteq U. \tag{3.4.1}$$

If u is a function on U with some smoothness property, then the restriction of u to $\overline{V}$ may be considered as having that property on $\overline{V}$.

However, we may also be concerned with functions that are defined only on $\overline{V}$. If k is a positive integer, then we let $C^k(\overline{V}, \mathbf{R})$, $C^k(\overline{V}, \mathbf{C})$ be the spaces of k-times continuously-differentiable real or complex-valued functions u on V, respectively, such that u and all of its derivatives $\partial^\alpha u$ with $|\alpha| \leq k$ can be extended continuously to $\overline{V}$, as in Section A of Chapter 0 of [32]. A continuous extension of a function on V to $\overline{V}$ is unique when it exists, by standard arguments, and so we may consider u and its derivatives of order less than or equal to k as being defined on $\overline{V}$ in this case.

This is initially defined a bit differently in Appendix A.3 of [29], where one considers k-times continuously-differentiable functions u on V such that u and its derivatives $\partial^\alpha u$ with $|\alpha| \leq k$ are uniformly continuous on bounded subsets of V. It is well known that continuous functions on compact sets are uniformly continuous, which implies that continuous functions of $\overline{V}$ are uniformly continuous on bounded subsets of $\overline{V}$. Conversely, if a function on V is uniformly continuous on bounded subsets of V, then it is well known and not too difficult to show that the function has a continuous extension to $\overline{V}$.

If m is a positive integer, then we may also be concerned with continuity or smoothness properties of functions with values in $\mathbf{R}^m$ or $\mathbf{C}^m$. Such a function may be considered as an m-tuple of real or complex-valued functions, and the continuity or smoothness properties of the function are equivalent to the analogous properties holding for each of the corresponding m components.

We may be concerned with smoothness properties of the boundary of V as well. Properties like these are discussed in Appendix C.1 of [29], and Section B of Chapter 0 of [32].

3.5 The divergence theorem

Let $n \geq 2$ be an integer, although one could include $n = 1$, with suitable interpretations. Also let V be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary. Thus we may consider n-dimensional integrals over V, and surface integrals over ∂V, of suitable functions on V and ∂V, respectively.

Let w be a continuously-differentiable function on $\overline{V}$ with values in $\mathbf{R}^n$ or $\mathbf{C}^n$. The *divergence theorem* states that

$$\int_V \operatorname{div} w(x)\, dx = \int_{\partial V} w(y') \cdot \nu(y')\, dy', \tag{3.5.1}$$

where dy' is the element of surface area on ∂V, and $\nu(y')$ is the outward-pointing unit normal to ∂V at $y' \in \partial V$.

Let u be a twice continuously-differentiable real or complex-valued function on $\overline{V}$. If we take

$$w_j = \frac{\partial u}{\partial x_j} \tag{3.5.2}$$

for each $j = 1, \ldots, n$, then w defines a continuously-differentiable function on $\overline{V}$ with values in $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate. In this case, the divergence theorem implies that

$$\int_V (\Delta u)(x)\, dx = \int_{\partial V} (D_{v(y')} u)(y')\, dy', \tag{3.5.3}$$

where $D_{\nu(y')}$ denotes the directional derivative in the direction of $\nu(y')$. In particular, if u is also harmonic on V, then

$$\int_{\partial V} (D_{\nu(y')} u)(y')\, dy' = 0. \tag{3.5.4}$$

Suppose that v is a continuously-differentiable real or complex-valued function on $\overline{V}$, as appropriate. Under these conditions,

$$\begin{aligned} \int_V (\Delta u)(x)\, v(x)\, dx + \int_V \sum_{j=1}^n \frac{\partial u}{\partial x_j}(x)\, \frac{\partial v}{\partial x_j}(x)\, dx \\ = \int_{\partial V} (D_{\nu(y')} u)(y')\, v(y')\, dy'. \end{aligned} \tag{3.5.5}$$

This follows from the divergence theorem, with

$$w_j(x) = v(x)\, \frac{\partial u}{\partial x_j}(x) \tag{3.5.6}$$

for each $j = 1, \ldots, n$. In particular, we can take $v = \overline{u}$, to get that

$$\int_V (\Delta u)(x)\, \overline{u(x)}\, dx + \int_V \sum_{j=1}^n \left|\frac{\partial u}{\partial x_j}(x)\right|^2 dx = \int_{\partial V} (D_{\nu(y')} u)(y')\, \overline{u(y')}\, dy'. \tag{3.5.7}$$

If u is any continuously-differentiable real or complex-valued function on $\overline{V}$, then

$$\int_V \sum_{j=1}^n \left|\frac{\partial u}{\partial x_j}(x)\right|^2 dx \tag{3.5.8}$$

is called the *Dirichlet integral* of u on V, as in Section E of Chapter 2 of [32], and Section 4 of Chapter 5 of [123]. This is equal to 0 exactly when all of the first partial derivatives of u are equal to 0 on V. This happens if and only if u is constant on each of the connected components of V.

3.6 Some consequences

Let n be a positive integer, and let V be a nonempty proper open subset of $\mathbf{R}^n$. Suppose that u is a continuous real or complex-valued function on $\overline{V}$ that satisfies Dirichlet boundary conditions, so that $u = 0$ on ∂V. If u is constant on any connected component of V, then it is easy to see that $u = 0$ on that component. If u is constant on every connected component of V, then it follows that $u = 0$ on V.

Suppose now that V is bounded, with reasonably smooth boundary, and that u is twice continuously differentiable on $\overline{V}$. If u satisfies Dirichlet boundary conditions on $\overline{V}$, then (3.5.7) reduces to

$$\int_V (\Delta u)(x)\,\overline{u(x)}\,dx + \int_V \sum_{j=1}^n \left|\frac{\partial u}{\partial x_j}(x)\right|^2 dx = 0. \tag{3.6.1}$$

If u is harmonic on V, then it follows that $u = 0$ on V, as in the preceding paragraph. This corresponds to Theorem 16 in Section 2.2.5 a of [29]. The same conclusion could also be obtained using the maximum principle, as in Section 6.7.

Let $\nu(y')$ be the outward-pointing unit normal to ∂V at $y' \in \partial V$, as in the previous section. If

$$(D_{\nu(y')} u)(y') = 0 \tag{3.6.2}$$

for every $y' \in \partial V$, then u is said to satisfy *Neumann boundary conditions* on $\overline{V}$. Note that (3.5.7) reduces to (3.6.1) in this case too. If u is harmonic on V, then this implies that u is constant on every connected component of V. This corresponds to part (a) of Problem 10 in Section 6.6 of [29] and Proposition 3.3 in Section A of Chapter 3 of [32], and is related to part (a) of Exercise 18 on p108 of [10].

Part (b) of Problem 10 in Section 6.6 of [29] asks one to show the same statement using the maximum principle, under suitable smoothness conditions on V. This is related to Exercise 27 on p29 of [10].

Suppose that u is an eigenfunction for the Laplacian on V with eigenvalue $\lambda \in \mathbf{C}$, so that

$$\Delta u = \lambda\, u \tag{3.6.3}$$

on V. If u satisfies Dirichlet or Neumann boundary conditions on $\overline{V}$, then we get that

$$\lambda \int_V |u(x)|^2\,dx + \int_V \sum_{j=1}^n \left|\frac{\partial u}{\partial x_j}(x)\right|^2 dx = 0, \tag{3.6.4}$$

by (3.6.1). If $u \neq 0$ somewhere on V, then

$$\int_V |u(x)|^2\,dx > 0. \tag{3.6.5}$$

Under these conditions, we obtain that $\lambda \in \mathbf{R}$, and that $\lambda \le 0$. More precisely, if u satisfies Dirichlet boundary conditions on $\overline{V}$, then we get that $\lambda < 0$.

3.7 Another consequence

Let $n \ge 2$ be an integer, and let V be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary again. Also let u be a twice continuously-differentiable real or complex-valued function on $\overline{V}$, and let v be a continuously-differentiable real or complex-valued function on $\overline{V}$, as appropriate. Suppose that v satisfies Dirichlet boundary conditions on $\overline{V}$, so that $v(y') = 0$ for every $y' \in \partial V$. In this case, (3.5.5) reduces to

$$\int_V (\Delta u)(x)\, v(x)\, dx + \int_V \sum_{j=1}^n \frac{\partial u}{\partial x_j}(x)\, \frac{\partial v}{\partial x_j}(x)\, dx = 0. \tag{3.7.1}$$

This means that

$$\int_V \sum_{j=1}^n \frac{\partial u}{\partial x_j}(x)\, \frac{\partial v}{\partial x_j}(x)\, dx = 0 \tag{3.7.2}$$

if and only if

$$\int_V (\Delta u)(x)\, v(x)\, dx = 0. \tag{3.7.3}$$

Of course, (3.7.3) holds when u is harmonic on V.

Conversely, if (3.7.3) holds for every smooth function v on $\mathbf{R}^n$ with compact support contained in V, then one can check that u is harmonic on V. It follows that u is harmonic on V when (3.7.2) holds for every such v.

Suppose now that u is a twice continuously-differentiable real or complex-valued function on V, and that v is a continuously-differentiable real or complex-valued function on $\mathbf{R}^n$ with compact support contained in V. Under these conditions, for each $j = 1, \ldots, n$, we can define w_j as a continuously-differentiable real or complex-valued function on $\mathbf{R}^n$ with compact support contained in V, using (3.5.6) on V, and putting $w_j(x) = 0$ when $x \in \mathbf{R}^n \setminus V$. Similarly,

$$v(x)\, (\Delta u)(x) \tag{3.7.4}$$

and

$$\sum_{j=1}^n \frac{\partial u}{\partial x_j}(x)\, \frac{\partial v}{\partial x_j}(x) \tag{3.7.5}$$

can be extended to continuous real or complex-valued functions on $\mathbf{R}^n$ with compact support contained in V. One can use the divergence theorem to get that (3.7.1) holds in this case as well.

3.8 The Dirichlet principle

Let $n \geq 2$ be an integer, and let V be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary. Suppose that u and v are continuously-differentiable complex-valued functions on $\overline{V}$. Put

$$D(u, v) = \int_V \sum_{j=1}^n \frac{\partial u}{\partial x_j}(x)\, \overline{\frac{\partial v}{\partial x_j}(x)}\, dx, \tag{3.8.1}$$

as in Section E of Chapter 2 of [32]. It is easy to see that this is Hermitian symmetric, in the sense that

$$D(u, v) = \overline{D(v, u)}. \tag{3.8.2}$$

Of course, if u and v are real-valued, then $D(u, v)$ is a real number, and symmetric in u and v.

If $u = v$, then (3.8.1) is the same as the Dirichlet integral (3.5.8), which is a nonnegative real number. One might be interested in trying to minimize this quantity, under suitable conditions.

We can express v as $u + (v - u)$, to get that

$$\begin{aligned} D(v, v) &= D(u, u) + D(u, v - u) + D(v - u, u) + D(v - u, v - u) \\ &= D(u, u) + 2\,\mathrm{Re}\, D(u, v - u) + D(v - u, v - u). \end{aligned} \tag{3.8.3}$$

Suppose now that u is twice continuously differentiable on $\overline{V}$, and that

$$u = v \text{ on } \partial V. \tag{3.8.4}$$

This means that $v - u = 0$ on ∂U, so that

$$D(u, v - u) = -\int_V (\Delta u)(x)\,\overline{(v(x) - u(x))}\,dx, \tag{3.8.5}$$

as in the previous section.

If u is harmonic on U, then it follows that

$$D(u, v - u) = 0. \tag{3.8.6}$$

This implies that

$$D(v, v) = D(u, u) + D(v - u, v - u), \tag{3.8.7}$$

by (3.8.3). In particular, we get that

$$D(u, u) \le D(v, v), \tag{3.8.8}$$

which is part of the *Dirichlet principle*. More precisely, equality holds in (3.8.8) if and only if

$$D(v - u, v - u) = 0. \tag{3.8.9}$$

This condition holds if and only if $u = v$ on $\overline{V}$, because of (3.8.4).

Conversely, suppose that (3.8.8) holds whenever (3.8.4) holds. If $t \in \mathbf{C}$, then

$$w = u + t\,(v - u) \tag{3.8.10}$$

is another continuously-differentiable complex-valued function on $\overline{V}$, and $u = w$ on ∂U, by construction. This means that

$$D(u, u) \le D(w, w), \tag{3.8.11}$$

by hypothesis. Note that

$$D(w, w) = D(u, u) + 2\,\mathrm{Re}\,\overline{t}\,D(u, v - u) + |t|^2\,D(v - u, v - u), \tag{3.8.12}$$

as in (3.8.3). One can use this and (3.8.11) to get that (3.8.6) holds, because $t \in \mathbf{C}$ is arbitrary. This implies that u is harmonic on U, because of (3.8.5), as in the previous section. This is another part of the Dirichlet principle. See also Section 4 of Chapter 5 of [123].

3.9 Another helpful fact about integrals

Let $n \ge 2$ be an integer, and let V be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary. One could also include $n = 1$, with suitable interpretations, as before. If u, v are twice continuously-differentiable real or complex-valued functions on $\overline{V}$, then

$$\int_V (u(x)\,(\Delta v)(x) - v(x)\,(\Delta u)(x))\,dx = \int_{\partial V} (u(y')\,(D_{\nu(y')} v)(y') - v(y')\,(D_{\nu(y')} u)(y'))\,dy'. \tag{3.9.1}$$

Here $\nu(y')$ is the outward-pointing unit normal to ∂V at $y' \in \partial V$, and $D_{\nu(y')}$ denotes the directional derivative in the direction of $\nu(y')$, as usual. This can be obtained from the divergence theorem, with

$$w_j = u\,\frac{\partial v}{\partial x_j} - v\,\frac{\partial u}{\partial x_j} \tag{3.9.2}$$

for each $j = 1, \ldots, n$.

Suppose for the moment that u and v both satisfy Dirichlet boundary conditions on $\overline{V}$, so that

$$u = v = 0 \text{ on } \partial V. \tag{3.9.3}$$

In this case, (3.9.1) reduces to

$$\int_V (u(x)\,(\Delta v)(x) - v(x)\,(\Delta u)(x))\,dx = 0. \tag{3.9.4}$$

Similarly, suppose that u and v both satisfy Neumann boundary conditions on $\overline{V}$, which is to say that

$$(D_{\nu(y')} u)(y') = (D_{\nu(y')} v)(y') = 0 \tag{3.9.5}$$

for every $y' \in \partial V$. Clearly (3.9.1) reduces to (3.9.4) in this case as well.

Suppose now that u and v are eigenfunctions for the Laplacian on V with eigenvalues λ and μ, respectively. This means that

$$\Delta u = \lambda\, u \tag{3.9.6}$$

and

$$\Delta v = \mu\, v \tag{3.9.7}$$

on V. Suppose also that either u and v both satisfy Dirichlet boundary conditions on $\overline{V}$, or that they both satisfy Neumann boundary conditions on $\overline{V}$, so that (3.9.4) holds. This means that

$$(\mu - \lambda) \int_V u(x)\,v(x)\,dx = 0. \tag{3.9.8}$$

If $\lambda \ne \mu$, then it follows that

$$\int_V u(x)\,v(x)\,dx = 0. \tag{3.9.9}$$

3.10 Some remarks about zero sets

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let ϕ be a continuous real-valued function on U, and consider the corresponding zero set of ϕ in U,

$$\{x \in U : \phi(x) = 0\}. \tag{3.10.1}$$

This is a relatively closed set in U.

Suppose now that ϕ is continuously-differentiable on U. Let w be an element of U such that

$$\phi(w) = 0 \tag{3.10.2}$$

and

$$\frac{\partial \phi}{\partial x_l}(w) \neq 0 \tag{3.10.3}$$

for some $l \in \{1, \ldots, n\}$. Under these conditions, the *implicit function theorem* implies that near w, the zero set (3.10.1) can be represented as the graph of a continuously-differentiable real-valued function of the other variables x_j, $j \neq l$. Note that the implicit function theorem for real-valued functions can be shown more directly than for $\mathbf{R}^m$-valued functions with $m \geq 2$, as in Theorem 3.2.1 on p36 of [84].

One can also look at this in terms of the *inverse function theorem*. Consider the mapping Φ from U into $\mathbf{R}^n$ defined by

$$\Phi(x) = (x_1, \ldots, x_{l-1}, \phi(x), x_{l+1}, \ldots, x_n) \tag{3.10.4}$$

for each $x \in U$. Equivalently, the jth coordinate of $\Phi(x)$ is defined to be x_j when $j \neq l$, and to be $\phi(x)$ when $j = l$. This mapping is continuously differentiable on U, because ϕ is continuously differentiable on U, by hypothesis.

The differential of Φ at a point $x \in U$ is the linear mapping from $\mathbf{R}^n$ into itself that sends $v \in \mathbf{R}^n$ to the directional derivative $(D_v \Phi)(x)$ of Φ in the direction v at x. This linear mapping corresponds to the matrix of partial derivatives of the components of Φ. One can check that the differential of Φ at w is invertible as a linear mapping on $\mathbf{R}^n$, because of (3.10.3).

Under these conditions, the inverse function theorem implies that the restriction of Φ to a small neighborhood of w is invertible, where the inverse mapping is continuously differentiable too. Of course, the zero set (3.10.1) is the same as the inverse image of the $x_l = 0$ hyperplane under Φ.

3.11 The Neumann problem

Let n be a positive integer, and let U be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary. If $y' \in \partial U$, then we let $\nu(y')$ be the outward-pointing unit normal to ∂U at y', and we let $D_{\nu(y')}$ denote the directional derivative in the direction $\nu(y')$, as usual.

Let f be a real or complex-valued function on U, and let g be a real or complex-valued function on ∂U. A version of the *Neumann problem* asks one to find a real or complex-valued function u on $\overline{U}$, as appropriate, such that

$$\Delta u = f \tag{3.11.1}$$

on U, and

$$(D_{\nu(y')} u)(y') = g(y') \tag{3.11.2}$$

for every $y' \in \partial U$. This is discussed in Section C of Chapter 2 of [32].

Of course, one could add a constant to u without affecting (3.11.1) or (3.11.2). More precisely, one could add a different constant to u on each connected component of U, without affecting these conditions.

If u is twice-continuously differentiable on $\overline{U}$, harmonic on U, and satisfies Neumann boundary conditions on $\overline{U}$, then u is constant on every connected component of U, as in Section 3.6. This implies an appropriate uniqueness result for the Neumann problem.

Suppose that u is a twice continuously-differentiable real or complex-valued function on $\overline{U}$ that satisfies (3.11.1) and (3.11.2). If V is a connected component of U, then the restriction of u to $\overline{V}$ satisfies the analogous conditions there. It follows that

$$\int_V f(x)\,dx = \int_{\partial V} g(y')\,dy', \tag{3.11.3}$$

as in Section 3.5.

One may be particularly concerned with the Neumann problem with $f = 0$ on U, which may be described as the Neumann problem for harmonic functions. Of course, (3.11.3) reduces to

$$\int_{\partial V} g(y')\,dy' = 0 \tag{3.11.4}$$

in this case. Alternatively, one may be particularly concerned with the case where $g = 0$ on ∂U, so that u satisfies Neumann boundary conditions on $\overline{U}$. In this case, (3.11.3) reduces to

$$\int_V f(x)\,dx = 0. \tag{3.11.5}$$

As with the Dirichlet problem, these two cases of the Neumann problem are related to each other. If one can solve the Poisson equation (3.11.1) without (3.11.2), then a solution to a Neumann problem for harmonic functions could be used to obtain (3.11.2). If one can solve the Poisson equation with Neumann boundary conditions, then one can use that to try to solve the Neumann problem for harmonic functions.

The Dirichlet and Neumann problems for harmonic functions are also discussed in Chapter 3 of [32]. Another approach is discussed in Chapter 7 of [32]. See also Problem 4 in Section 6.6 of [29].

3.12 The unit ball in $\mathbf{R}^n$

Let n be a positive integer, and let us consider the case where $U = B(0,1)$ is the open unit ball in $\mathbf{R}^n$ in the previous section. If y' is an element of $\partial U = \partial B(0,1)$, which is the unit sphere in $\mathbf{R}^n$, then

$$\nu(y') = y' \tag{3.12.1}$$

is the outward-pointing unit normal to $\partial B(0,1)$ at y'. If u is a continuously-differentiable complex-valued function on $\overline{U} = \overline{B}(0,1)$, then

$$(D_{\nu(y')}u)(y') \tag{3.12.2}$$

is the same as the Euler operator applied to u at y', as in Section 2.8.

Suppose that p is a polynomial on $\mathbf{R}^n$ with complex coefficients that is homogeneous of degree k for some nonnegative integer k. If $y' \in \partial B(0,1)$, then

$$(D_{\nu(y')}p)(y') = k\,p(y'), \tag{3.12.3}$$

as in Section 2.8. If $k \geq 1$, and

$$q = k^{-1}\,p, \tag{3.12.4}$$

then

$$(D_{\nu(y')}q)(y') = p(y'). \tag{3.12.5}$$

This may be considered as an instance of the Neumann problem for harmonic functions on $\overline{B}(0,1)$ when p is harmonic, so that q is harmonic as well.

If p is a harmonic polynomial on $\mathbf{R}^n$ that is homogeneous of degree $k \geq 1$, then

$$\int_{\partial B(0,1)} p(y')\,dy' = 0. \tag{3.12.6}$$

This follows from (3.5.4) and (3.12.3).

If g is any polynomial on $\mathbf{R}^n$ with complex coefficients, then g agrees with a harmonic polynomial on $\partial B(0,1)$, as in Section 2.13. More precisely, g is equal to a sum of homogeneous harmonic polynomials on $\partial B(0,1)$, as before. If these homogeneous harmonic polynomials are all homogeneous of positive degree, then one can get a polynomial solution to the corresponding Neumann problem for harmonic functions on $\overline{B}(0,1)$, using (3.12.5). This condition on g is equivalent to asking that

$$\int_{\partial B(0,1)} g(y')\,dy' = 0, \tag{3.12.7}$$

because of (3.12.6). This is related to Exercise 18 on p108 of [10].

Let p_1, p_2 be harmonic polynomials on $\mathbf{R}^n$ with complex coefficients that are homogeneous of degrees $k_1, k_2 \geq 0$, respectively. Observe that

$$(k_1 - k_2) \int_{\partial B(0,1)} p_1(y')\,p_2(y')\,dy' = 0, \tag{3.12.8}$$

because of (3.9.1) and (3.12.3). If $k_1 \neq k_2$, then we get that

$$\int_{\partial B(0,1)} p_1(y')\, p_2(y')\, dy' = 0. \tag{3.12.9}$$

Note that this includes (3.12.6) as a particular case. This corresponds to Proposition 5.9 on p79 of [10], part of Theorem 2.51 in Section G of Chapter 2 of [32], Proposition 4.1.5 on p179 of [76], 3.1.1 on p69 of [119], and Corollary 2.4 on p141 of [125].

3.13 Some integrals over spheres

Let n be a positive integer, and let p be a polynomial on $\mathbf{R}^n$ with complex coefficients. If p is harmonic on $\mathbf{R}^n$ and homogeneous of degree $k \geq 1$, then

$$\int_{\partial B(0,r)} p(y')\, dy' = 0 \tag{3.13.1}$$

for every $r > 0$. This is the same as (3.12.6) when $r = 1$. Otherwise, one can reduce to that case using a change of variables, or use an analogous argument for any $r > 0$.

If p is any polynomial on $\mathbf{R}^n$ of degree less than or equal to N for some nonnegative integer N, then p can be expressed in a unique way as a sum of homogeneous polynomials of degrees from 0 to N, as in Section 2.9. If p is a harmonic polynomial, then one can use this to get that p can be expressed as a sum of harmonic homogeneous polynomials of degrees from 0 to N. This uses the fact that the Laplacian of a homogeneous polynomial of degree l is a homogeneous polynomial of degree $l-2$ when $l \geq 2$. Of course, the Laplacian of a homogeneous polynomial of degree l is 0 when $l = 0$ or 1.

If p is a harmonic polynomial on $\mathbf{R}^n$, then

$$\frac{1}{|\partial B(0,r)|} \int_{\partial B(0,r)} p(y')\, dy' = p(0) \tag{3.13.2}$$

for every $r > 0$. Here $|\partial B(0,r)|$ denotes the $(n-1)$-dimensional surface area of $\partial B(0,r)$. More precisely, (3.13.2) follows from (3.13.1) when p is homogeneous of degree $k \geq 1$. If p is homogeneous of degree 0, and thus a constant, then (3.13.2) is clear. One can reduce to the case where p is homogeneous of some degree $k \geq 0$, using the remarks in the preceding paragraph.

It follows that

$$\frac{1}{|\partial B(a,r)|} \int_{\partial B(a,r)} p(y')\, dy' = p(a) \tag{3.13.3}$$

for every $a \in \mathbf{R}^n$ and $r > 0$ under these conditions. This uses the fact that $p(x+a)$ is also a harmonic polynomial in x on $\mathbf{R}^n$. If one replaces $p(x)$ with $p(x+a)$ in (3.13.2), then the result is the same as (3.13.3), using a translation by a to go from an integral over $\partial B(0,r)$ to an integral over $\partial B(a,r)$. Of course, the surface area $|\partial B(a,r)|$ of $\partial B(a,r)$ is the same as the surface area of $\partial B(0,r)$.

This is known as the *mean-value property* of p, which will be discuseed further in Chapter 6. Any harmonic function on an open subset of $\mathbf{R}^n$ has a suitable version of this property, as in Section 6.2. A twice continuously-differentiable function with the mean-value property is harmonic, as in Section 6.3. One can also use the mean-value property to get smoothness, as in Section 6.4.

3.14 Some remarks about compositions

Let W be a nonempty open subset of $\mathbf{R}^2$, and suppose that f is a continuously-differentiable complex-valued function on W. If $v \in \mathbf{R}^2$, then the directional derivative of f in the direction v is equal to

$$D_v f = v_1\, \partial_1 f + v_2\, \partial_2 f \tag{3.14.1}$$

on W. If we identify v with the complex number $v_1 + i\, v_2$, then it is easy to see that

$$D_v f = v\, \frac{\partial f}{\partial z} + \overline{v}\, \frac{\partial f}{\partial \overline{z}}, \tag{3.14.2}$$

where $\partial f/\partial z$ and $\partial f/\partial \overline{z}$ are as in Section 2.2.

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a continuously-differentiable complex-valued function on U. Suppose that

$$u(U) \subseteq W, \tag{3.14.3}$$

where W is considered as a subset of $\mathbf{C}$, and that f is holomorphic on W. Under these conditions, we get that

$$\frac{\partial}{\partial x_j}(f(u(x))) = f'(u(x))\, \frac{\partial u}{\partial x_j}(x) \tag{3.14.4}$$

on U for each $j = 1, \ldots, n$, where $f' = \partial f/\partial z$ is the usual complex derivative of f.

Suppose for the moment that $n = 2$, and that u is holomorphic on U. In this case, one can check that $f \circ u$ is holomorphic on U, using (3.14.4).

It is well known that the complex exponential function is holomorphic on $\mathbf{C}$, with complex derivative equal to itself. If u is any continuously-differentiable complex-valued function on U, then it follows that

$$\frac{\partial}{\partial x_j}(\exp u(x)) = (\exp u(x))\, \frac{\partial u}{\partial x_j}(x) \tag{3.14.5}$$

on U for each $j = 1, \ldots, n$.

Suppose now that u is a continuously-differentiable real-valued function on U, and that W is an open subset of $\mathbf{R}$ that satisfies (3.14.3). If f is any continuously-differentiable complex-valued function on W, then (3.14.4) holds on U, by the usual chain rule.

3.15 More on first-order operators

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let $a_1, \ldots, a_n$ be n complex-valued functions on U, so that $a = (a_1, \ldots, a_n)$ may be considered as a mapping from U into $\mathbf{C}^n$. If u is a continuously-differentiable complex-valued function on U, then

$$L_a(u) = \sum_{j=1}^{n} a_j \, \frac{\partial u}{\partial x_j} \tag{3.15.1}$$

defines a complex-valued function on U, as in Section 2.3.

Let b be another complex-valued function on U, and put

$$L(u) = L_a(u) + b\, u. \tag{3.15.2}$$

This defines a differential operator on U, as in Section 2.4, with $N = 1$.

If c is a continuously-differentiable complex-valued function on U, then

$$L_a(c\, u) = L_a(c)\, u + c\, L_a(u) \tag{3.15.3}$$

on U, as in Section 2.3. If $c(x) \neq 0$ for every $x \in U$, then we get that

$$c^{-1}\, L_a(c\, u) = L_a(u) + c^{-1}\, L_a(c)\, u \tag{3.15.4}$$

on U. This is the same as (3.15.2) when

$$b = c^{-1}\, L_a(c). \tag{3.15.5}$$

If γ is a continuously-differentiable complex-valued function on U, then

$$c(x) = \exp \gamma(x) \tag{3.15.6}$$

is a continuously-differentiable complex-valued function on U with $c(x) \neq 0$ for every $x \in U$. We also have that

$$c^{-1}\, L_a(c) = L_a(\gamma) \tag{3.15.7}$$

on U, as in (3.14.5).

Suppose that u, v are continuously-differentiable complex-valued functions on U that are eigenfunctions for L_a, with eigenvalues $\lambda, \mu \in \mathbf{C}$, respectively. Observe that

$$L_a(u\, v) = L_a(u)\, v + u\, L_a(v) = (\lambda + \mu)\, u\, v \tag{3.15.8}$$

on U, so that $u\, v$ is an eigenfunction for L_a with eigenvalue $\lambda + \mu$.

Chapter 4

First-order equations

4.1 Some real first-order operators

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let $a_1, \ldots, a_n$ be n real-valued functions on U. Alternatively,

$$a(x) = (a_1(x), \ldots, a_n(x)) \tag{4.1.1}$$

defines a mapping from U into $\mathbf{R}^n$.

If u is a continuously-differentiable real-valued function on U, then put

$$L_a(u) = \sum_{j=1}^{n} a_j \, \frac{\partial u}{\partial x_j}. \tag{4.1.2}$$

This defines a real-valued function on U. The value of this function at $x \in U$ is the directional derivative of u at x in the direction $a(x)$.

Let I be a nonempty open interval in the real line, or an open half-line, or the whole real line, and let $w(t)$ be a continuously-differentiable function of $t \in I$ with values in $\mathbf{R}^n$. Equivalently, this means that the jth component $w_j(t)$ of $w(t)$ is a continuously-differentiable real-valued function on I for each $j = 1, \ldots, n$. One could also allow I to contain one or both endpoints, with suitable interpretations using one-sided derivatives at the endpoints. Suppose that

$$w(t) \in U \text{ for every } t \in I, \tag{4.1.3}$$

so that

$$z(t) = u(w(t)) \tag{4.1.4}$$

defines a real-valued function on I. It is well known that $z(t)$ is continuously differentiable on I under these conditions, with

$$z'(t) = \sum_{j=1}^{n} w_j'(t) \, (\partial_j u)(w(t)) \tag{4.1.5}$$

for every $t \in I$.

Suppose that

$$w_j'(t) = a_j(w(t)) \tag{4.1.6}$$

for every $t \in I$. This is the same as saying that

$$w'(t) = a(w(t)) \tag{4.1.7}$$

for every $t \in I$, as elements of $\mathbf{R}^n$. In this case, we get that

$$z'(t) = \sum_{j=1}^{n} a_j(w(t))\,(\partial_j u)(w(t)) = (L_a(u))(w(t)) \tag{4.1.8}$$

for every $t \in I$.

Suppose for the moment that we also have that

$$L_a(u) = 0 \text{ on } U. \tag{4.1.9}$$

This implies that

$$z'(t) = 0 \tag{4.1.10}$$

for every $t \in I$, so that $z(t)$ is constant on I.

Suppose now that u satisfies the semilinear equation

$$(L_a(u))(x) + b(u(x), x) = 0 \tag{4.1.11}$$

for some real-valued function b on $\mathbf{R} \times U$. In this case, we get that

$$z'(t) + b(z(t), w(t)) = 0 \tag{4.1.12}$$

for every $t \in I$.

The equations (4.1.7) and (4.1.12) are called the *characteristic equations* for (4.1.11). This is related to some remarks in Section 3.2.2 a of [29], and Section B of Chapter 1 of [32].

It is interesting to consider the case where a is a nonzero constant, as in Section 2.1 in [29].

4.2 Quasilinear first-order equations

Let n be a positive integer again, and let U be a nonempty open subset of $\mathbf{R}^n$. In this section, we let $a_1, \ldots, a_n$ and b be real-valued functions on $\mathbf{R} \times U$. Consider the quasi-linear first-order partial differential equation

$$\sum_{j=1}^{n} a_j(u(x), x)\,\frac{\partial u}{\partial x_j}(x) + b(u(x), x) = 0, \tag{4.2.1}$$

where u is a continuously-differentiable real-valued function on U.

Let I be an interval in the real line with nonempty interior, which may be unbounded, as in the previous section. Also let $w(t)$ be a continuously-differentiable function of $t \in I$ with values in $\mathbf{R}^n$, and in fact in U, as before. If u is a continuously-differentiable real-valued function on U, then

$$z(t) = u(w(t)) \tag{4.2.2}$$

is a continuously-differentiable real-valued function of $t \in I$, with derivative as in (4.1.5).

Suppose that

$$w_j'(t) = a_j(u(w(t)), w(t)) \tag{4.2.3}$$

for every $t \in I$. If we consider $a = (a_1, \ldots, a_n)$ as an $\mathbf{R}^n$-valued function on $\mathbf{R} \times U$, then this is the same as saying that

$$w'(t) = a(u(w(t)), w(t)) \tag{4.2.4}$$

for every $t \in I$, as elements of $\mathbf{R}^n$. Under these conditions, (4.2.1) implies that

$$z'(t) + b(z(t), w(t)) = 0 \tag{4.2.5}$$

for every $t \in I$.

Observe that (4.2.3) is the same as saying that

$$w_j'(t) = a_j(z(t), w(t)) \tag{4.2.6}$$

for every $t \in I$. Equivalently, this means that

$$w'(t) = a(z(t), w(t)) \tag{4.2.7}$$

for every $t \in I$, as elements of $\mathbf{R}^n$. This together with (4.2.5) form a coupled system of ordinary differential equations for $w(t)$ and $z(t)$ that does not depend on u. These are the *characteristic equations* for (4.2.1). This is related to some remarks in Section 3.2.2 b of [29], and in Section B of Chapter 1 of [32].

There is an important difference between this case and the one discussed in the previous section. It is well known that solutions of the initial value problem for systems of ordinary differential equations are unique under suitable conditions. This implies that different curves corresponding to solutions of (4.1.7) cannot cross each other, under suitable conditions. Although one also has uniqueness for the initial value problem for the system (4.2.5), (4.2.7) under suitable conditions, it is possible for the curves corresponding to the $w(t)$'s to cross each other. This corresponds to some remarks in Sections 3.2.5 a, b of [29].

4.3 Fully nonlinear first-order equations

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n$, and let $F(q, y, x)$ be a real-valued function on

$$\mathbf{R}^n \times \mathbf{R} \times U. \tag{4.3.1}$$

Consider the fully nonlinear first-order partial differential equation

$$F(Du(x), u(x), x) = 0, \tag{4.3.2}$$

where u is a continuously-differentiable real-valued function on U.

As in the previous sections, we would like to find some systems of ordinary differential equations that are related to (4.3.2). Let I be an interval in the real line with nonempty interior, and which may be unbounded, and let $w(t)$ be a continuously-differentiable function of $t \in I$ with values in U again. Suppose that u is a continuously-differentiable real-valued function on U, so that

$$z(t) = u(w(t)) \tag{4.3.3}$$

is a continuously-differentiable real-valued function of $t \in I$, as before.

If $t \in I$, then let $p(t) \in \mathbf{R}^n$ be defined by

$$p(t) = Du(w(t)), \tag{4.3.4}$$

so that

$$p_j(t) = (\partial_j u)(w(t)) \tag{4.3.5}$$

for each $j = 1, \ldots, n$. We would like to find a nice system of ordinary differential equations for $w(t)$, $z(t)$, and $p(t)$. To do this, we suppose that u is twice continuously-differentiable on U, so that $p(t)$ is continuously differentiable on I. More precisely,

$$p_j'(t) = \sum_{l=1}^{n} w_l'(t)\,(\partial_j \partial_l u)(w(t)) \tag{4.3.6}$$

for every $j = 1, \ldots, n$ and $t \in I$.

Suppose that u satisfies (4.3.2) on U, and that F is continuously differentiable on (4.3.1). If we differentiate the left side of (4.3.2) with respect to x_j, then we get that

$$\begin{aligned}\sum_{l=1}^{n} \frac{\partial F}{\partial q_l}(Du(x), u(x), x)\,\frac{\partial^2 u}{\partial x_j\,\partial x_l}(x) + \frac{\partial F}{\partial y}(Du(x), u(x), x)\,\frac{\partial u}{\partial x_j}(x)& \\ + \frac{\partial F}{\partial x_j}(Du(x), u(x), x) = 0.&\end{aligned} \tag{4.3.7}$$

If $t \in I$, then we can take $x = w(t)$, to get that

$$\begin{aligned}\sum_{l=1}^{n} \frac{\partial F}{\partial q_l}(p(t), z(t), w(t))\,(\partial_j \partial_l u)(w(t)) + \frac{\partial F}{\partial y}(p(t), z(t), w(t))\,p_j(t)& \\ + \frac{\partial F}{\partial x_j}(p(t), z(t), w(t)) = 0&\end{aligned} \tag{4.3.8}$$

for each $j = 1, \ldots, n$.

Suppose that

$$w_l'(t) = \frac{\partial F}{\partial q_l}(p(t), z(t), w(t)) \tag{4.3.9}$$

for each $l = 1, \ldots, n$ and $t \in I$. Using this and (4.3.8), we obtain that

$$p_j'(t) = -\frac{\partial F}{\partial y}(p(t), z(t), w(t))\, p_j(t) - \frac{\partial F}{\partial x_j}(p(t), z(t), w(t)) \tag{4.3.10}$$

for each $j = 1, \ldots, n$ and $t \in I$.

Remember that $z'(t)$ can be expressed as in (4.1.5). Using (4.3.9), we get that

$$z'(t) = \sum_{j=1}^{n} \frac{\partial F}{\partial q_j}(p(t), z(t), w(t))\, p_j(t) \tag{4.3.11}$$

for every $t \in I$.

Thus (4.3.9), (4.3.10), and (4.3.11) form a coupled system of ordinary differential equations for $w(t)$, $z(t)$, and $p(t)$ that does not depend on u. These are the *characteristic equations* for (4.3.2). This corresponds to some remarks in Section 3.2.1 of [29], and Section B of Chapter 1 of [32].

4.4 More on fully nonlinear equations

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n$, and let $F(q, y, x)$ be a continuously-differentiable real-valued function on $\mathbf{R}^n \times \mathbf{R} \times U$ again. If $w(t)$, $z(t)$, and $p(t)$ are any continuously-differentiable functions on I with values in $\mathbf{R}^n$, $\mathbf{R}$, and U, respectively, then

$$\begin{aligned} \frac{d}{dt} F(p(t), q(t), w(t)) &= \sum_{j=1}^{n} \frac{\partial F}{\partial q_j}(p(t), z(t), w(t))\, p_j'(t) \\ &\quad + \frac{\partial F}{\partial y}(p(t), z(t), w(t))\, z'(t) \\ &\quad + \sum_{j=1}^{n} \frac{\partial F}{\partial x_j}(p(t), z(t), w(t))\, w_j'(t) \end{aligned} \tag{4.4.1}$$

on I.

If $w(t)$, $z(t)$, and $p(t)$ satisfy the characteristic equations (4.3.9), (4.3.10), and (4.3.11) on I, then it is easy to see that

$$\frac{d}{dt} F(p(t), q(t), w(t)) = 0 \tag{4.4.2}$$

on I. Of course, this means that $F(p(t), z(t), w(t))$ is constant on I.

If

$$F(p(t), z(t), w(t)) = 0 \tag{4.4.3}$$

for some $t \in I$, then it follows that this holds for all $t \in I$ under these conditions. This corresponds to step 2 in the proof of Theorem 2 in Section 3.2.4 of [29], and a remark near the end of Section B of Chapter 1 of [32].

Of course, if $w(t)$, $p(t)$, and $z(t)$ are associated to a solution u of (4.3.2) as in the previous section, then (4.4.3) holds by construction. Alternatively, one may consider initial value problems for the characteristic equations (4.3.7), (4.3.10), and (4.3.11) that satisfy (4.4.3) for some t, and thus for all t. Solutions to initial value problems of this type can be used to try to find solutions of (4.3.2), as in the next section.

In the quasilinear case, we have that

$$F(q, y, x) = \sum_{j=1}^{n} a_j(y, x)\, q_j + b(y, x) \tag{4.4.4}$$

for some real-valued functions $a_j(y, x)$, $1 \leq j \leq n$, and $b(y, x)$ on $\mathbf{R} \times U$. Note that (4.3.9) is the same as (4.2.6) in this case. If (4.4.3) holds, then (4.3.11) is the same as (4.2.5). This corresponds to a remark in Section 3.2.2 b of [29], and just after (1.14) in Section B of Chapter 1 of [32].

If c is a real number, then

$$\widehat{F}(q, y, x) = F(q, y, x) - c \tag{4.4.5}$$

is another continuously-differentiable real-valued function on $\mathbf{R}^n \times \mathbf{R} \times U$. If u is a continuously-differentiable real-valued function on U, then the first-order partial differential equation

$$\widehat{F}(Du(x), u(x), x) = 0 \tag{4.4.6}$$

on U is the same as saying that

$$F(Du(x), u(x), x) = c \tag{4.4.7}$$

on U. Note that the characteristic equations associated to $\widehat{F}$ as in the previous section are the same as for F, because they only involve the derivatives of F, $\widehat{F}$.

If F is as in (4.4.4), then

$$\widehat{F}(q, y, x) = \sum_{j=1}^{n} a_j(y, x)\, q_j + \widehat{b}(y, x), \tag{4.4.8}$$

with $\widehat{b}(y, x) = b(y, x) - c$. However, the characteristic equations associated to the quasilinear equations corresponding to F and $\widehat{F}$ as in Section 4.2 are not the same when $c \neq 0$. The equations for $w_j'(t)$ are the same, but the analogue of the equation (4.2.5) for $z'(t)$ with $\widehat{b}(y, x)$ instead of $b(y, x)$ is a bit different. The conditions under which this equation is supposed to be the same as in the previous section are also a bit different.

4.5 Non-characteristic conditions

Let n be a positive integer, and let U be a nonempty open set in $\mathbf{R}^n$. In Sections 4.1 – 4.3, we started with a solution u of a first-order partial differential equation on U, and found systems of ordinary differential equations that described the behavior of u along certain curves. These systems of ordinary differential equations do not depend on u, and may be used to try to find solutions of the partial differential equation, at least locally, as in Section 3.2.4 of [29], and Section B of Chapter 1 of [32].

More precisely, this normally involves additional regularity conditions on the functions used to define the original partial differential equation, in order to use appropriate results about systems of ordinary differential equations. One might suppose that u is given along a nice $(n-1)$-dimensional submanifold Σ of $\mathbf{R}^n$, with suitable regularity on Σ, as in [29, 32]. One would like to a find a solution to the partial differential equation near Σ, with the given values on Σ, perhaps at least near a given point on Σ. This may be considered as an *initial value problem* or *Cauchy problem* for the partial differential equation.

Remember that we considered systems of ordinary differential equations for $w(t)$, $z(t)$, and possibly $p(t)$ defined on an interval I in the rel line. To deal with the initial value problem for the partial differential equation along Σ, we want to consider suitable initial value problems for these systems of ordinary differential equations associated to points in Σ. Let $\sigma \in \Sigma$ and $t_0 \in \mathbf{R}$ be given, although one might normally simply take $t_0 = 0$. The initial conditions for $w(t)$ and $z(t)$ at $t = t_0$ are

$$w(t_0) = \sigma \tag{4.5.1}$$

and

$$z(t_0) = u(\sigma). \tag{4.5.2}$$

In the fully nonlinear case, we would also need to specify $p(t_0)$, and we shall return to that later.

We would like to define u near Σ in such a way that

$$u(w(t)) = z(t). \tag{4.5.3}$$

In particular, we would like to be able to reach points in U near Σ by such a path $w(t)$. In order to do this, there is an additional *non-characteristic condition*, as in Section 3.2.3 c of [29], and Section B of Chapter 1 of [32]. The non-characteristic condition at σ asks that $w'(t_0)$ not be tangent to Σ at $w(t_0)$. If $\nu(\sigma)$ is a nonzero element of $\mathbf{R}^n$ that is normal to Σ at σ, then this means that

$$w'(t_0) \cdot \nu(w(t_0)) = w'(t_0) \cdot \nu(\sigma) \neq 0. \tag{4.5.4}$$

In Section 4.1, the ordinary differential equations for w are given in terms of the functions $a_j(x)$, $1 \leq j \leq n$, and the non-characteristic condition at $\sigma \in \Sigma$ can be expressed as

$$a(\sigma) \cdot \nu(\sigma) = \sum_{j=1}^{n} a_j(\sigma)\, \nu_j(\sigma) \neq 0. \tag{4.5.5}$$

In Section 4.2, the ordinary differential equations for w are coupled with those for z, and the non-characteristic condition at σ can be expressed as

$$a(u(\sigma), \sigma) \cdot \nu(\sigma) = \sum_{j=1}^{n} a_j(u(\sigma), \sigma)\, \nu_j(\sigma) \neq 0. \tag{4.5.6}$$

In particular, this depends on the value of u at σ.

In Section 4.3, the ordinary differential equations for w are coupled with those for z and p, and the non-characteristic condition at σ can be expressed as

$$\begin{aligned} &\sum_{j=1}^{n} \frac{\partial F}{\partial q_j}(p(t_0), z(t_0), w(t_0))\, \nu_j(w(t_0)) \\ &\quad = \sum_{j=1}^{n} \frac{\partial F}{\partial q_j}(p(t_0), u(\sigma), \sigma)\, \nu_j(\sigma) \neq 0. \end{aligned} \tag{4.5.7}$$

This depends on the value of u at σ, and $p(t_0)$, which is supposed to represent the values of the first partial derivatives of u at σ.

The directional derivative of u at σ in a direction that is tangent to Σ at σ is determined by the restriction of u to Σ. Another condition on the first partial derivatives of u at σ is given by the partial differential equation. One basically needs to be able to choose $p(t_0)$ in a way that is compatible with these conditions, and the non-characteristic condition (4.5.7) depends on the choice of $p(t_0)$.

In particular, $p(t_0)$ should satisfy

$$F(p(t_0), z(t_0), w(t_0)) = F(p(t_0), u(\sigma), \sigma) = 0. \tag{4.5.8}$$

This implies that (4.4.3) holds along the curve, as before.

In Section 3.2.3 c of [29], one starts with a suitable choice of $p(t_0)$ for a point $\sigma \in \Sigma$. If the non-characteristic condition (4.5.7) holds at σ, then one can use the implicit function theorem to get suitable initial conditions for p corresponding to other points in Σ that are close to σ.

In Section B of Chapter 1 of [32], one simply asks to have suitable initial conditions for p corresponding to points along Σ. An important case where this is easy to get will be discussed in Section 4.10.

4.6 More on the Euler operator

Let n be a positive integer, and put $a_j(x) = x_j$ on $\mathbf{R}^n$ for each $j = 1, \ldots, n$. Thus $a = (a_1, \ldots, a_n)$ is the identity mapping on $\mathbf{R}^n$, and

$$L_a = \sum_{j=1}^{n} x_j \, \frac{\partial}{\partial x_j} \tag{4.6.1}$$

is the Euler operator, as in Section 2.8. In this case, (4.1.6) reduces to

$$w_j'(t) = w_j(t). \tag{4.6.2}$$

This is solved on the real line by

$$w_j(t) = c_j \exp t, \tag{4.6.3}$$

with $c_j \in \mathbf{R}$ for $j = 1, \ldots, n$. Equivalently, (4.1.7) reduces to

$$w'(t) = w(t), \tag{4.6.4}$$

which is solved on the real line by

$$w(t) = (\exp t)\, c, \tag{4.6.5}$$

where $c = (c_1, \ldots, c_n) \in \mathbf{R}^n$.

Let u be a continuously-differentiable real or complex-valued function on $\mathbf{R}^n \setminus \{0\}$. Suppose that $c \neq 0$, so that (4.6.5) is nonzero for each $t \in \mathbf{R}$. Observe that

$$\frac{d}{dt}(u((\exp t)\, c)) = \sum_{j=1}^{n} (\exp t)\, c_j\, (\partial_j u)((\exp t)\, c) = (L_a(u))((\exp t)\, c) \tag{4.6.6}$$

for every $t \in \mathbf{R}$. This is the same as (4.1.8) in this case. Of course, this is analogous to considering $u(\tau\, x)$ for $x \in \mathbf{R}^n \setminus \{0\}$ and $\tau > 0$, and differentiating in τ, as in Section 2.8.

If

$$L_a(u) = b\, u \tag{4.6.7}$$

on $\mathbf{R}^n \setminus \{0\}$ for some $b \in \mathbf{C}$, then (4.6.6) implies that

$$\frac{d}{dt}(u((\exp t)\, c)) = b\, u((\exp t)\, c) \tag{4.6.8}$$

for every $t \in \mathbf{R}$. This means that

$$u((\exp t)\, c) = u(c)\, \exp(b\, t) \tag{4.6.9}$$

for every $t \in \mathbf{R}$, which holds automatically when $t = 0$. One can use this to get that u is homogeneous of degree b on $\mathbf{R}^n \setminus \{0\}$, as in Section 2.8. Conversely, if u is homogeneous of degree b on $\mathbf{R}^n \setminus \{0\}$, then (4.6.9) holds, which implies that (4.6.8) holds, and thus (4.6.7) holds.

4.7 Angular derivatives in the plane

Let $a_1(x)$, $a_2(x)$ be the real-valued functions on $\mathbf{R}^2$ defined by

$$a_1(x) = -x_2, \;\; a_2(x) = x_1. \tag{4.7.1}$$

Thus

$$a(x) = (a_1(x), a_2(x)) = (-x_2, x_1) \tag{4.7.2}$$

defines a mapping from $\mathbf{R}^2$ onto itself. If we identify $x = (x_1, x_2) \in \mathbf{R}^2$ with $x_1 + x_2\, i \in \mathbf{C}$, then

$$a(x) = -x_2 + x_1\, i = i\, x. \tag{4.7.3}$$

The corresponding system of ordinary differential equations (4.1.6) reduces to

$$w_1'(t) = -w_2(t), \ w_2'(t) = w_1(t) \tag{4.7.4}$$

in this case. If we identify $w(t) = (w_1(t), w_2(t))$ with $w_1(t) + w_2(t)\, i$, as before, then this is the same as saying that

$$w'(t) = i\, w(t), \tag{4.7.5}$$

as in (4.1.7). This is solved on the real line by

$$w(t) = (\exp(i\, t))\, c, \tag{4.7.6}$$

where $c = (c_1, c_2) \in \mathbf{R}^2$ is identified with $c_1 + c_2\, i \in \mathbf{C}$, as usual. Note that $w(0) = c$.

Let U be a nonempty open subset of $\mathbf{R}^2$, and let u be a continuously-differentiable real or complex-valued function on U. Let $L_a(u)$ be the continuous real or complex-valued function on U, as appropriate, defined by

$$(L_a(u))(x) = -x_2\, \frac{\partial u}{\partial x_1}(x) + x_1\, \frac{\partial u}{\partial x_2}(x) \tag{4.7.7}$$

for every $x \in U$, as in Section 4.1. This is the same as the directional derivative of u at x, in the direction corresponding to $i\, x$, because of (4.7.3).

Let $x \in \mathbf{R}^2$ be given, and consider

$$\{t \in \mathbf{R} : \exp(i\, t)\, x \in U\}, \tag{4.7.8}$$

where $\mathbf{R}^2$ is identified with $\mathbf{C}$ as before. This is an open subset of $\mathbf{R}$, and

$$u(\exp(i\, t)\, x) \tag{4.7.9}$$

may be considered as a continuously-differentiable real or complex-valued function of t in (4.7.8). Observe that

$$\frac{d}{dt}(u(\exp(i\, t)\, x)) = (L_a(u))(\exp(i\, t)\, x) \tag{4.7.10}$$

for every t in (4.7.8).

A nice example related to this case is discussed in Section 3.2.2 a of [29].

4.8 Another example on $\mathbf{R}^2$

Now let $a_1(x)$, $a_2(x)$ be the real-valued functions on $\mathbf{R}^2$ defined by

$$a_1(x) = x_1,\ a_2(x) = -x_2, \tag{4.8.1}$$

and put

$$a(x) = (a_1(x), a_2(x)) = (x_1, -x_2) \tag{4.8.2}$$

for every $x \in \mathbf{R}^2$. This leads to the system of ordinary differential equations

$$w_1'(t) = w_1(t),\ w_2'(t) = -w_2(t), \tag{4.8.3}$$

as in (4.1.6) again. These equations are solved on the real line by

$$w_1(t) = c_1 \exp t,\ w_2(t) = c_2 \exp(-t), \tag{4.8.4}$$

with $c_1, c_2 \in \mathbf{R}$. If we put $w(t) = (w_1(t), w_2(t))$ and $c = (c_1, c_2)$, then we get that $w(0) = c$. It follows from (4.8.4) that

$$w_1(t)\, w_2(t) = c_1\, c_2 \tag{4.8.5}$$

for every $t \in \mathbf{R}$.

Let U be a nonempty subset of $\mathbf{R}^2$ again, and let u be a continuously-differentiable real or complex-valued function on U. Also let $L_a(u)$ be the continuous real or complex-valued function defined on U by

$$(L_a(u))(x) = x_1 \frac{\partial u}{\partial x_1}(x) - x_2 \frac{\partial u}{\partial x_2}(x), \tag{4.8.6}$$

as in Section 4.1.

Let $x \in \mathbf{R}^2$ be given, and note that

$$\{r \in \mathbf{R}_+ : (r\, x_1, r^{-1}\, x_2) \in U\} \tag{4.8.7}$$

is an open subset of $\mathbf{R}$. We may consider

$$u(r\, x_1, r^{-1}\, x_2) \tag{4.8.8}$$

as a continuously-differentiable real or complex-valued function of r on (4.8.7). If r is in (4.8.7), then

$$\begin{aligned} \frac{d}{dr}(u(r\, x_1, r^{-1}\, x_2)) &= x_1\, (\partial_1 u)(r\, x_1, r^{-1}\, x_2) - r^{-2}\, x_2\, (\partial_2 u)(r\, x_1, r^{-1}\, x_2) \\ &= r^{-1}\, (L_a(u))(r\, x_1, r^{-1}\, x_2). \end{aligned} \tag{4.8.9}$$

Alternatively,

$$\{t \in \mathbf{R} : ((\exp t)\, x_1, (\exp(-t))\, x_2) \in U\} \tag{4.8.10}$$

is an open subset of $\mathbf{R}$, and

$$u((\exp t)\, x_1, (\exp(-t))\, x_2) \tag{4.8.11}$$

may be considered as a continuously-differentiable real or complex-valued function of t on (4.8.10). If t is in (4.8.10), then

$$\frac{d}{dt}(u((\exp t)\, x_1, (\exp(-t))\, x_2)) = (L_a(u))((\exp t)\, x_1, (\exp(-t))\, x_2). \tag{4.8.12}$$

This is related to Exercise (3) in Section B of Chapter 1 of [32]. In particular, if f is a continuously-differentiable real or complex-valued function on an open subset V of the real line, then

$$u(x_1, x_2) = f(x_1\, x_2) \tag{4.8.13}$$

is a continuously-differentiable function on the open set

$$\{(x_1, x_2) \in \mathbf{R}^n : x_1\, x_2 \in V\} \tag{4.8.14}$$

in the plane, and

$$L_a(u) = 0 \tag{4.8.15}$$

on (4.8.14).

4.9 Some simpler quasilinear equations

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let $a_1, \ldots, a_n$ be real-valued functions on $\mathbf{R} \times U$, and let b be a real-valued function on $\mathbf{R}$. Consider the quasilinear first-order partial differential equation

$$\sum_{j=1}^{n} a_j(u(x), x)\, \frac{\partial u}{\partial x_j}(x) + b(u(x)) = 0, \tag{4.9.1}$$

where u is a continuously-differentiable real-valued function on U. This is the same as in Section 4.2, with b not depending on $x \in U$.

Let I be an interval in the real line with nonempty interior, and which may be unbounded, and let $w(t)$ be a continuously-differentiable function of $t \in I$ with values in U, as before. We previously considered a system of ordinary differential equations for $w(t)$ and a continuously-differentiable real-valued function $z(t)$ of $t \in I$. The equation for $w(t)$ is

$$w'(t) = a(z(t), w(t)) \tag{4.9.2}$$

for every $t \in I$, as before. In this case, the equation for $z(t)$ is

$$z'(t) + b(z(t)) = 0 \tag{4.9.3}$$

for every $t \in I$. This does not depend on $w(t)$, and so a solution to (4.9.3) can be used to get that (4.9.2) is a system of ordinary differential equations for $w(t)$ on I.

If $b \equiv 0$ on $\mathbf{R}$, then (4.9.1) reduces to

$$\sum_{j=1}^{n} a_j(u(x), x) \frac{\partial u}{\partial x_j}(x) = 0, \tag{4.9.4}$$

on U. Similarly, (4.9.3) reduces to

$$z'(t) = 0 \tag{4.9.5}$$

for every $t \in I$. Of course, this means that $z(t)$ is constant on I.

Suppose now that $a_1, \ldots, a_n$ are real-valued functions on $\mathbf{R}$, which is to say that they do not depend on $x \in U$. This means that (4.9.1) reduces to

$$\sum_{j=1}^{n} a_j(u(x)) \frac{\partial u}{\partial x_j}(x) + b(u(x)) = 0 \tag{4.9.6}$$

on U. Similarly, (4.9.2) reduces to

$$w'(t) = a(z(t)) \tag{4.9.7}$$

on I.

If we also ask that $b \equiv 0$ on $\mathbf{R}$ again, then (4.9.4) reduces to

$$\sum_{j=1}^{n} a_j(u(x)) \frac{\partial u}{\partial x_j}(x) = 0 \tag{4.9.8}$$

on U. The right side of (4.9.7) is constant on I under these conditions, as before. This means that the curve corresponding to $w(t)$ follows a straight line, at constant speed.

However, it is possible for curves like these to cross each other, as mentioned in Section 3.2.5 b of [29]. This can lead to limitations on continuously-differentiable solutions of (4.9.8), as in [29].

Some equations like these will be mentioned in Section 4.12.

4.10 A simplification with x_n

Let n be an integer greater than or equal to 2, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let $F(q, y, x) = F(q_1, \ldots, q_n, y, x)$ be a real-valued function on $\mathbf{R}^n \times \mathbf{R} \times U$, as in Section 4.3. If u is a continuously-differentiable real-valued function on U, then the first-order partial differential equation corresponding to $F(q, y, x)$ can be expressed as

$$F\Big(\frac{\partial u}{\partial x_1}(x), \ldots, \frac{\partial u}{\partial x_n}(x), u(x), x\Big) = 0. \tag{4.10.1}$$

Suppose that $F(q, y, x)$ can be expressed as

$$F(q_1, \ldots, q_{n-1}, q_n, y, x) = q_n + \widetilde{F}(q_1, \ldots, q_{n-1}, y, x) \tag{4.10.2}$$

for some real-valued function $\widetilde{F}(q_1, \ldots, q_{n-1}, y, x)$ on $\mathbf{R}^{n-1} \times \mathbf{R} \times U$. In this case, (4.10.1) is the same as saying that

$$\frac{\partial u}{\partial x_n}(x) + \widetilde{F}\Big(\frac{\partial u}{\partial x_1}(x), \ldots, \frac{\partial u}{\partial x_{n-1}}(x), u(x), x\Big) = 0. \tag{4.10.3}$$

Suppose that $\widetilde{F}$ is continuously differentiable on $\mathbf{R}^{n-1} \times \mathbf{R} \times U$, so that F is continuously differentiable on $\mathbf{R}^n \times \mathbf{R} \times U$. This leads to a coupled system of ordinary differential equations for $w(t)$, $z(t)$, and $p(t)$ as in Section 4.3. The differential equation for the nth component $w_n(t)$ of $w(t)$ reduces to

$$w_n'(t) = 1 \tag{4.10.4}$$

for every t in the interval I.

In the quasilinear case, as in Section 4.2, the analogous condition is that

$$a_n \equiv 1 \tag{4.10.5}$$

on $\mathbf{R} \times U$. In this case, we have a coupled system of ordinary differential equations for $w(t)$ and $z(t)$, as before. The differential equation for $w_n(t)$ reduces to (4.10.4) again. Similarly, one may consider the condition (4.10.5) in Section 4.1, where a_n is a real-valued function on U. The system of ordinary differential equations for $w(t)$ depends only on a, and the differential equation for $w_n(t)$ reduces to (4.10.4).

Suppose that the hypersurface Σ mentioned in Section 4.5 is contained in a hyperplane where x_n is constant. Note that the non-characteristic condition holds when the differential equation for $w_n(t)$ is as in (4.10.4).

The directional derivatives of u at a point in Σ in directions tangent to Σ are determined by the restriction of u to Σ, as before. In this case, this means that the derivative of u with respect to x_j on Σ is determined by the restriction of u to Σ for $j = 1, \ldots, n-1$. If u satisfies a partial differential equation as in (4.10.3), then it follows that the derivative of u with respect to x_n on Σ is determined by the restriction of u to Σ as well.

If u is given on Σ, then this makes it easy to get initial conditions for p at points in Σ, as in Section 4.5. More precisely, the initial condition for p_j at a point in Σ is given by the derivative of u with respect to x_j at the point when $j = 1, \ldots, n-1$, and is determined by (4.10.3) when $j = n$.

4.11 Some simpler fully nonlinear equations

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let $F(q, x)$ be a real-valued function on $\mathbf{R}^n \times U$, and consider the fully nonlinear first-order partial differential equation

$$F(Du(x), x) = 0, \tag{4.11.1}$$

where u is a continuously-differentiable real-valued function on U. This is the same as in Section 4.3, where the function $F(q, y, x)$ on $\mathbf{R}^n \times \mathbf{R} \times U$ does not depend on $y \in \mathbf{R}$.

Let I be an interval in the real line with nonempty interior, and which may be unbounded, and let $w(t)$, $z(t)$, and $p(t)$ be continuously-differentiable functions of $t \in I$ with values in U, $\mathbf{R}$, and $\mathbf{R}^n$, respectively. If $F(q, x)$ is continuously differentiable on $\mathbf{R}^n \times U$, then we get a system of ordinary differential equations for $w(t)$, $z(t)$, and $p(t)$, as in Section 4.3, which has some simplifications in this case. The equations for $w'(t)$ are now

$$w_l'(t) = \frac{\partial F}{\partial q_l}(p(t), w(t)) \tag{4.11.2}$$

for each $l = 1, \ldots, n$ and $t \in I$. The equations for $p'(t)$ reduce to

$$p_j'(t) = -\frac{\partial F}{\partial x_j}(p(t), w(t)) \tag{4.11.3}$$

for each $j = 1, \ldots, n$ and $t \in I$. The equation for $z'(t)$ is

$$z'(t) = \sum_{j=1}^{n} \frac{\partial F}{\partial q_j}(p(t), w(t))\, p_j(t) \tag{4.11.4}$$

on I.

The right sides of these equations do not involve $z(t)$. Thus (4.11.2) and (4.11.3) form a system of ordinary differential equations for $w(t)$ and $p(t)$. If one has solutions for these equations, then (4.11.4) can be solved directly.

Suppose now that $n \geq 2$, and that $F(q, x)$ can be expressed as

$$F(q_1, \ldots, q_{n-1}, q_n, x) = q_n + \widetilde{F}(q_1, \ldots, q_{n-1}, x) \tag{4.11.5}$$

for some real-valued function $\widetilde{F}(q_1, \ldots, q_{n-1}, x)$ on $\mathbf{R}^{n-1} \times U$, as in the previous section. This means that (4.11.1) is the same as saying that

$$\frac{\partial u}{\partial x_n}(x) + \widetilde{F}\Big(\frac{\partial u}{\partial x_1}(x), \ldots, \frac{\partial u}{\partial x_{n-1}}(x), x\Big) = 0, \tag{4.11.6}$$

on U, as before. Of course, (4.11.2) reduces to (4.10.4) when $l = n$. Similarly, (4.11.4) can be reexpressed as

$$z'(t) = \sum_{j=1}^{n-1} \frac{\partial \widetilde{F}}{\partial q_j}(p_1(t), \ldots, p_{n-1}(t), w(t))\, p_j(t) + p_n(t) \tag{4.11.7}$$

in this case.

If $\widetilde{F}(q_1, \ldots, q_{n-1}, x) = \widetilde{F}(q_1, \ldots, q_{n-1}, x_1, \ldots, x_{n-1}, x_n)$ does not depend on x_n, then (4.11.6) is the same as the *Hamilton–Jacobi equation*. Under these conditions, (4.11.3) says that

$$p_n'(t) = 0 \tag{4.11.8}$$

for every $t \in I$ when $j = n$, so that p_n is constant on I. This type of equation is discussed in [29], starting in Section 3.2.5 c. These equations are normally expressed a bit differently, as in the next section.

4.12 Other notation in $n + 1$ variables

Let n be a positive integer, and let us identify $\mathbf{R}^n \times \mathbf{R}$ with $\mathbf{R}^{n+1}$ in the usual way. An element of $\mathbf{R}^n \times \mathbf{R}$ may be expressed as (x, τ), where $x \in \mathbf{R}^n$ and $\tau \in \mathbf{R}$.

Let U be a nonempty open subset of $\mathbf{R}^n \times \mathbf{R}$, and let

$$F(q_1, \ldots, q_n, q_{n+1}, y, x, \tau) \tag{4.12.1}$$

be a real-valued function on $\mathbf{R}^{n+1} \times \mathbf{R} \times U$. This means that (4.12.1) is defined for $q_1, \ldots, q_n, q_{n+1}, y \in \mathbf{R}$ and $(x, \tau) \in U$. If $u(x, \tau)$ is a continuously-differentiable real-valued function on U, then the first-order partial differential equation corresponding to (4.12.1) can be expressed as

$$F\Big(\frac{\partial u}{\partial x_1}(x, \tau), \ldots, \frac{\partial u}{\partial x_n}(x, \tau), \frac{\partial u}{\partial \tau}(x, \tau), u(x, \tau), x, \tau\Big) = 0. \tag{4.12.2}$$

Suppose that (4.12.1) can be expressed as

$$q_{n+1} + \widetilde{F}(q_1, \ldots, q_n, y, x, \tau) \tag{4.12.3}$$

for some real-valued function $\widetilde{F}(q_1, \ldots, q_n, y, x, \tau)$ on $\mathbf{R}^n \times \mathbf{R} \times U$. Under these coditions, (4.12.2) is the same as saying that

$$\frac{\partial u}{\partial \tau}(x, \tau) + \widetilde{F}\Big(\frac{\partial u}{\partial x_1}(x, \tau), \ldots, \frac{\partial u}{\partial x_n}(x, \tau), u(x, \tau), x, \tau\Big) = 0. \tag{4.12.4}$$

This corresponds to (4.10.3), in this notation.

If $\widetilde{F}$ is continuously differentiable on $\mathbf{R}^n \times \mathbf{R} \times U$, then we can consider the associated system of ordinary differential equations, as before. The analogue of (4.10.4) with n replaced by $n + 1$ permits us to identify t with τ, perhaps with a suitable translation. Of course, an equation of the form (4.12.4) is often expressed with t in place of τ.

If $\widetilde{F}(q_1, \ldots, q_n, y, x, \tau)$ does not depend on y or τ, then (4.12.4) is the same as the Hamilton–Jacobi equation, as in the previous section, with slightly different notation.

Let Φ be a continuously-differentiable function on the real line with values in $\mathbf{R}^n$. The partial differential equation

$$\frac{\partial u}{\partial \tau} + \operatorname{div} \Phi(u) = 0 \tag{4.12.5}$$

is called a *scalar conservation law*, as in Example 5 in Section 3.2.5 b of [29]. More precisely, the divergence is taken in the x variables here. Equivalently, this can be expressed as

$$\frac{\partial u}{\partial \tau} + \sum_{j=1}^{n} \Phi_j'(u)\, \frac{\partial u}{\partial x_j} = 0, \tag{4.12.6}$$

where Φ_j is the jth component of Φ for each $j = 1, \ldots, n$. This may be considered as one of the types of equations mentioned in Section 4.9.

If $n = 1$ and $b \in \mathbf{R}$, then

$$\frac{\partial u}{\partial \tau}(x, \tau) + u(x, \tau)\,\frac{\partial u}{\partial x}(x, \tau) = b \tag{4.12.7}$$

is a quasilinear first-order equation with some additional properties as in Section 4.9. This is the inviscid form of *Burger's equation* when $b = 0$, which is discussed in Section 3.4.1 of [29]. Note that Burger's equation is an example of a scalar conservation law. This equation with $b = 1$ is mentioned in Problem 5 (c) in Section 3.5 of [29], as well as Example 2 and exercise (4) in Section B of Chapter 1 of [32].

4.13 Some other fully nonlinear equations

Let n be a positive integer, and let $F(q, y)$ be a real-valued function on $\mathbf{R}^n \times \mathbf{R}$. Also let U be a nonempty open subset of $\mathbf{R}^n$, and consider the fully nonlinear first-order partial differential equation

$$F(Du(x), u(x)) = 0, \tag{4.13.1}$$

where u is a continuously-differentiable real-valued function on U. This is the same as in Section 4.3 again, where the function $F(q, y, x)$ on $\mathbf{R}^n \times \mathbf{R} \times U$ does not depend on $x \in U$.

Let I be an interval in the real line with nonempty interior, and which may be unbounded, and let $w(t)$, $z(t)$, and $p(t)$ be continuously-differentiable functions of $t \in I$ with values in U, $\mathbf{R}$, and $\mathbf{R}^n$, respectively, as before. If $F(q, y)$ is continuously differentiable on $\mathbf{R}^n \times \mathbf{R}$, then the system of ordinary differential equations for $w(t)$, $z(t)$, and $p(t)$ discussed in Section 4.3 can be simplified in this case too. The equations for $w'(t)$ are

$$w_l'(t) = \frac{\partial F}{\partial q_l}(p(t), z(t)) \tag{4.13.2}$$

for each $l = 1, \ldots, n$ and $t \in I$. The equations for $p'(t)$ now reduce to

$$p_j'(t) = -\frac{\partial F}{\partial y}(p(t), z(t))\, p_j(t) \tag{4.13.3}$$

for each $j = 1, \ldots, n$ and $t \in I$. The equation for $z'(t)$ reduces to

$$z'(t) = \sum_{j=1}^{n} \frac{\partial F}{\partial q_j}(p(t), z(t))\, p_j(t) \tag{4.13.4}$$

for every $t \in I$.

The right sides of these equations do not involve $w(t)$, so that (4.13.3) and (4.13.4) form a system of ordinary differential equations for $p(t)$ and $z(t)$. If one has solutions to these equations, then (4.13.2) can be solved directly, as before.

Suppose for the moment that the derivative of $F(q, y)$ in y does not depend on y. This means that

$$F(q, y) = F(q, 0) + \frac{\partial F}{\partial y}(q, 0)\, y \tag{4.13.5}$$

for every $q \in \mathbf{R}^n$ and $y \in \mathbf{R}$. Note that $F(q, 0)$ and $(\partial F/\partial y)(q, 0)$ can be arbitrary continuously-differentiable real-valued functions of $q \in \mathbf{R}^n$ here. It follows that the right side of (4.13.3) does not involve $z(t)$, so that one gets a system of ordinary differential equations for $p(t)$. If one has a solution for this system, then (4.13.4) gives an ordinary differential equation for $z(t)$.

Suppose now that the derivative of $F(q, y)$ in y is a constant $c \in \mathbf{R}$, so that

$$F(q, y) = F(q, 0) + c\, y \tag{4.13.6}$$

for every $q \in \mathbf{R}^n$ and $y \in \mathbf{R}$. In this case, (4.13.3) is the same as saying that

$$p_j'(t) = -c\, p_j(t) \tag{4.13.7}$$

on I for each $j = 1, \ldots, n$. This is solved by taking

$$p_j(t) = a_j \, \exp(-c\, t) \tag{4.13.8}$$

for some real numbers $a_1, \ldots, a_n$. Note that the right side of (4.13.4) does not involve $z(t)$ under these conditions. Example 3 in Section 3.2.2 c of [29] is a nice example of this type.

4.14 A simpler case

Let n be a positive integer again, and let $F(q)$ be a real-valued function on $\mathbf{R}^n$. Consider the fully nonlinear first-order partial differential equation

$$F(Du(x)) = 0, \tag{4.14.1}$$

where u is a continuously-differentiable real-valued function on a nonempty open subset U of $\mathbf{R}^n$. This is the same as in Section 4.3, where the function $F(q, y, x)$ on $\mathbf{R}^n \times \mathbf{R} \times U$ does not depend on either $y \in \mathbf{R}$ or $x \in U$. This may also be considered as a particular case of the classes of fully nonlinear equations discussed in Sections 4.11 and 4.13. If $a \in \mathbf{R}^n$ and $b \in \mathbf{R}$, then

$$u(x) = a \cdot x + b \tag{4.14.2}$$

satisfies (4.14.1) on $\mathbf{R}^n$ if and only if

$$F(a) = 0. \tag{4.14.3}$$

Let I be an interval n the real line with nonempty interior, and which may be unbounded, and let $w(t)$, $z(t)$, and $p(t)$ be continuously-differentiable functions on I with values in U, $\mathbf{R}$, and $\mathbf{R}^n$, respectively, as usual. If $F(q)$ is continuously differentiable on $\mathbf{R}^n$, then the system of ordinary differential equations for $w(t)$, $z(t)$, and $p(t)$ discussed in Section 4.3 can be simplified further, as follows. The equations for $w'(t)$ are

$$w_l'(t) = \frac{\partial F}{\partial q_l}(p(t)) \tag{4.14.4}$$

for each $l = 1, \ldots, n$ and $t \in I$. The equations for $p'(t)$ are simply

$$p_j'(t) = 0 \tag{4.14.5}$$

for each $j = 1, \ldots, n$ and $t \in I$. The equation for $z'(t)$ is

$$z'(t) = \sum_{j=1}^n \frac{\partial F}{\partial q_j}(p(t))\, p_j(t) \tag{4.14.6}$$

for every $t \in I$.

Of course, (4.14.5) implies that $p(t)$ is constant on I. This means that the right sides of (4.14.4) and (4.14.6) are constant on I as well.

The *eikonal equation*

$$|\nabla u(x)| = 1 \tag{4.14.7}$$

is a partial differential equation of this type. More precisely, this is equivalent to saying that

$$|\nabla u(x)|^2 = 1 \tag{4.14.8}$$

on U. This corresponds to taking

$$F(q) = |q|^2 - 1 = \sum_{j=1}^n q_j^2 - 1, \tag{4.14.9}$$

which is a smooth function on $\mathbf{R}^n$.

Suppose that $n \ge 2$, and that $F(q)$ can be expressed as

$$F(q_1, \ldots, q_n) = q_n + \widetilde{F}(q_1, \ldots, q_{n-1}) \tag{4.14.10}$$

for some real-valued function $\widetilde{F}(q_1, \ldots, q_{n-1})$ on $\mathbf{R}^{n-1}$, as in Section 4.10. In this case, (4.14.1) is the same as saying that

$$\frac{\partial u}{\partial x_n}(x) + \widetilde{F}\Big(\frac{\partial u}{\partial x_1}(x), \ldots, \frac{\partial u}{\partial x_{n-1}}(x)\Big) = 0 \tag{4.14.11}$$

on U, as before. Remember that (4.14.4) reduces to (4.10.4) when $l = n$. Similarly, (4.14.6) reduces to

$$z'(t) = \sum_{j=1}^{n-1} \frac{\partial \widetilde{F}}{\partial q_j}(p_1(t), \ldots, p_{n-1}(t))\, p_j(t) + p_n(t) \tag{4.14.12}$$

under these conditions. Of course, (4.14.11) is a type of Hamilton–Jacobi equation, as in Section 4.11. This may normally be expressed a bit differently, as in Section 4.12. This type of Hamilton–Jacobi equation is discussed in Section 3.3 of [29].

4.15 Quasilinearity and derivatives

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n$, and let $F(q, y, y)$ be a continuously-differentiable real-valued function on $\mathbf{R}^n \times \mathbf{R} \times U$. Also let u be a twice continuously-differentiable real-valued function on U, and suppose that

$$F(Du(x), u(x), x) \text{ is constant on } U. \tag{4.15.1}$$

This implies that

$$\frac{\partial}{\partial x_j}(F(Du(x), u(x), x)) = 0 \tag{4.15.2}$$

on U for each $j = 1, \ldots, n$. This can be expanded using the chain rule to get differential equations that are linear in the second derivatives of u, as in Section 4.3.

Suppose that there is a real number c such that

$$F(q, y, x) = F(q, 0, x) + c\, y \tag{4.15.3}$$

on $\mathbf{R}^n \times \mathbf{R} \times U$. This implies that the equations (4.15.2) only involve the first and second derivatives of u, and not u itself.

Suppose now that $n \geq 2$, and that $F(q, y, x)$ can be expressed as

$$F(q, y, x) = \widehat{F}(q_1, x) + \sum_{l=2}^{n} a_l\, q_l + c\, y \tag{4.15.4}$$

on $\mathbf{R}^n \times \mathbf{R} \times U$. Here $\widehat{F}(q_1, x)$ is a continuously-differentiable real-valued function on $\mathbf{R} \times U$, and $a_1, \ldots, a_n$ and c are real numbers. In this case, the equation (4.15.2) with $j = 1$ reduces to

$$\begin{aligned} \frac{\partial \widehat{F}}{\partial q_1}\Big(\frac{\partial u}{\partial x_1}(x), x\Big)\, \frac{\partial^2 u}{\partial x_1^2}(x) + \sum_{l=2}^{n} a_l\, \frac{\partial^2 u}{\partial x_1\, \partial x_l}(x) \\ + c\, \frac{\partial u}{\partial x_1}(x) + \frac{\partial \widehat{F}}{\partial x_1}\Big(\frac{\partial u}{\partial x_1}(x), x\Big) = 0. \end{aligned} \tag{4.15.5}$$

This may be considered as a first-order quasilinear partial differential equation in $\partial u / \partial x_1$.

Suppose that $n = 2$, and that $F(q, y, x)$ can be expressed as

$$F(q, y, x) = \widetilde{F}(q_1) + q_2, \tag{4.15.6}$$

where $\widetilde{F}$ is a continuously-differentiable real-valued function on $\mathbf{R}$. This means that (4.15.5) reduces to

$$\widetilde{F}'\Big(\frac{\partial u}{\partial x_1}(x)\Big)\, \frac{\partial^2 u}{\partial x_1^2}(x) + \frac{\partial^2 u}{\partial x_1\, \partial x_2}(x) = 0. \tag{4.15.7}$$

This may be considered as a scalar conservation law in $\partial u / \partial x_1$, as in Section 4.12. This corresponds to a remark about the initial value problem (26) in Section 3.4.2 in [29].

Chapter 5

Some flows and exponentials

5.1 Some flows on $\mathbf{R}^n$

Let n be a positive integer, and let us identify $\mathbf{R}^n \times \mathbf{R}$ with $\mathbf{R}^{n+1}$, as in the previous section. An element of $\mathbf{R}^n \times \mathbf{R}$ may be expressed as (x, τ), with $x \in \mathbf{R}^n$ and $\tau \in \mathbf{R}$, as before. Let I be an interval in $\mathbf{R}$ with nonempty interior, which may be unbounded, and let W be a nonempty open subset of $\mathbf{R}^n$.

Suppose that for each $t \in I$, ϕ_t is a mapping from W into itself. Typically we might have that $0 \in I$, and that ϕ_0 is the identity mapping on W. If $\xi \in W$, then we ask that $\phi_t(\xi)$ be differentiable as a function of $t \in I$ with values in $\mathbf{R}^n$. This should be interpreted in terms of one-sided derivatives at any endpoints of I that are contained in I.

Let $u(x, \tau)$ be a continuously-differentiable real or complex-valued function on $W \times I$. If $x \in W$ and τ is an endpoint of I that is contained in I, then the partial derivative of u at (x, τ) in x_j can be defined in the usual way for $j = 1, \ldots, n$, and the partial derivative in τ can be defined as a one-sided derivative.

If $\xi \in W$, then $u(\phi_t(\xi), t)$ is differentiable as a real or complex-valued function of $t \in I$, with

$$\frac{d}{dt}(u(\phi_t(\xi), t)) = \sum_{j=1}^{n} \frac{d\phi_{t,j}(\xi)}{dt} \, \frac{\partial u}{\partial x_j}(\phi_t(\xi), t) + \frac{\partial u}{\partial \tau}(\phi_t(\xi), t). \tag{5.1.1}$$

Here $\phi_{t,j}(\xi)$ is the jth coordinate of $\phi_t(\xi)$ for each $j = 1, \ldots, n$.

Note that

$$\Phi(\xi, t) = (\phi_t(\xi), t) \tag{5.1.2}$$

defines a mapping from $W \times I$ into itself. Suppose now that for each $t \in I$, ϕ_t is a one-to-one mapping from W onto itself. Equivalently, this means that Φ is a one-to-one mapping from $W \times I$ onto itself.

If $1 \le j \le n$, then let a_j be the real-valued function on $W \times I$ such that

$$a_j(\phi_t(\xi), t) = \frac{d\phi_{t,j}(\xi)}{dt} \tag{5.1.3}$$

for every $\xi \in W$ and $t \in I$. Also put

$$a_{n+1} \equiv 1 \tag{5.1.4}$$

on $W \times I$, so that

$$a = (a_1, \ldots, a_n, a_{n+1}) \tag{5.1.5}$$

defines a mapping from $W \times I$ into $\mathbf{R}^{n+1}$.

Put

$$L_a(u) = \sum_{j=1}^{n} a_j \frac{\partial u}{\partial x_j} + \frac{\partial u}{\partial \tau} \tag{5.1.6}$$

on $W \times I$, as in Section 4.1. By construction,

$$(L_a(u))(\phi_t(\xi), t) = \frac{d}{dt}(u(\phi_t(\xi), t)) \tag{5.1.7}$$

for every $\xi \in W$ and $t \in I$. Similarly, if $\xi \in W$, then

$$(\phi_t(\xi), t) \tag{5.1.8}$$

satisfies the system of ordinary differential equations associated to a as a function of $t \in I$ as for $w(t)$ in Section 4.1.

Suppose for the moment that $I = \mathbf{R}$, and that

$$\phi_{r+t}(\xi) = \phi_r(\phi_t(\xi)) \tag{5.1.9}$$

for every $\xi \in W$ and $r, t \in \mathbf{R}$. This implies that the derivative of $\phi_t(\xi)$ in t at t is the same as the derivative of $\phi_r(\phi_t(\xi))$ in r at $r = 0$. This means that

$$a(\phi_t(\xi), t) = a(\phi_t(\xi), 0), \tag{5.1.10}$$

so that $a(x, \tau)$ does not depend on τ. Note that (5.1.9) implies that ϕ_0 is the identity mapping on W, because ϕ_0 is supposed to map W onto itself.

5.2 A more local version

Let n be a positive integer, and let us identify $\mathbf{R}^n \times \mathbf{R}$ with $\mathbf{R}^{n+1}$ again. Let U be an open subset of $\mathbf{R}^n \times \mathbf{R}$, and put

$$U_t = \{x \in \mathbf{R}^n : (x, t) \in U\} \tag{5.2.1}$$

for each $t \in \mathbf{R}$, which is an open set in $\mathbf{R}^n$. Let V be another open subset of $\mathbf{R}^n \times \mathbf{R}$, and let V_t be as in (5.2.1) for each $t \in \mathbf{R}$. If $\xi \in \mathbf{R}^n$, then

$$\{t \in \mathbf{R} : (\xi, t) \in V\} \tag{5.2.2}$$

is an open subset of $\mathbf{R}$.

Suppose that for each $t \in \mathbf{R}$, ϕ_t is a mapping from V_t into U_t. This means that

$$\Phi(\xi, t) = (\phi_t(\xi), t) \tag{5.2.3}$$

defines a mapping from V into U. If $\xi \in \mathbf{R}^n$, then we ask that $\phi_t(\xi)$ be differentiable as a function of t in (5.2.2) with values in $\mathbf{R}^n$.

Let u be a continuously-differentiable real or complex-valued function on U, and let $\xi \in \mathbf{R}^n$ be given. If t is an element of (5.2.2), then $(\xi, t) \in V$, $\xi \in V_t$, $\phi_t(\xi) \in U_t$, and thus

$$(\phi_t(\xi), t) \in U. \tag{5.2.4}$$

This means that

$$u(\phi_t(\xi), t) \tag{5.2.5}$$

is defined as a real or complex-valued function on (5.2.2). In fact, (5.2.5) is differentiable as a real or complex-valued function of t in (5.2.2), with derivative in t as in (5.1.1).

Suppose now that for each $t \in \mathbf{R}$, ϕ_t is a one-to-one mapping from V_t onto U_t. Equivalently, this means that the mapping Φ in (5.2.3) is a one-to-one mapping from V onto U. If $1 \le j \le n$, then let a_j be the real-valued function on U such that

$$a_j(\phi_t(\xi), t) = \frac{d\phi_{t,j}(\xi)}{dt} \tag{5.2.6}$$

for every $(\xi, t) \in V$. Also put $a_{n+1} \equiv 1$ on U, so that $a = (a_1, \ldots, a_n, a_{n+1})$ defines a mapping from U into $\mathbf{R}^{n+1}$.

Let $L_a(u)$ be defined on U as in (5.1.6). If $(\xi, t) \in V$, then

$$(L_a(u))(\phi_t(\xi), t) = \frac{d}{dt}(u(\phi_t(\xi), t)), \tag{5.2.7}$$

as before. Similarly, if $\xi \in \mathbf{R}^n$, then

$$(\phi_t(\xi), t) \tag{5.2.8}$$

satisfies the system of ordinary differential equations associated to a as a function of t in (5.2.2) as for $w(t)$ in Section 4.1.

Let $(\xi, t) \in V$ be given, and suppose that

$$(\phi_t(\xi), 0) \in V. \tag{5.2.9}$$

This implies that

$$(\phi_t(\xi), r) \in V \tag{5.2.10}$$

for every $r \in \mathbf{R}$ with $|r|$ sufficiently small. Of course, we also have that

$$(\xi, t + r) \in V \tag{5.2.11}$$

when $|r|$ is sufficiently small.

Suppose that

$$\phi_{r+t}(\xi) = \phi_r(\phi_t(\xi)) \tag{5.2.12}$$

when r is sufficiently small. This implies that the derivative of $\phi_t(\xi)$ in t at t is equal to the derivative of $\phi_r(\phi_t(\xi))$ in r at $r = 0$, as in the previous section. This means that

$$a(\phi_t(\xi), t) = a(\phi_t(\xi), 0), \tag{5.2.13}$$

as before.

5.3 Some basic first-order operators

Let n be a positive integer, and suppose that $a_j(x)$ is a real-valued linear function on $\mathbf{R}^n$ for each $j = 1, \dots, n$. This can be expressed as

$$a_j(x) = \sum_{l=1}^{n} a_{j,l}\, x_l \tag{5.3.1}$$

for $x \in \mathbf{R}^n$ and $j = 1, \dots, n$, where $(a_{j,l}) = (a_{j,l})_{j,l=1}^n$ is an $n \times n$ matrix of real numbers. Equivalently,

$$a(x) = (a_1(x), \dots, a_n(x)) \tag{5.3.2}$$

is a linear mapping from $\mathbf{R}^n$ into itself, which corresponds to this matrix in the usual way.

Let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a continuously-differentiable real or complex-valued function on U. Thus

$$(L_a(u))(x) = \sum_{j=1}^{n} a_j(x)\, \frac{\partial u}{\partial x_j}(x) \tag{5.3.3}$$

defines a continuous real or complex-valued function on U, as appropriate. Note that the examples mentioned in Sections 2.8, 4.7, and 4.8 are of this form.

Suppose for the moment that $U = \mathbf{R}^n \setminus \{0\}$, and that u is homogeneous of degree $b \in \mathbf{C}$. It is easy to see that $L_a(u)$ is homogeneous of degree b as well, because the partial derivatives of u are homogeneous of degree $b-1$, as in Section 2.8. Similarly, if p is a polynomial on $\mathbf{R}^n$ with real or complex coefficients that is homogeneous of degree k for some nonnegative integer k, then $L_a(p)$ is a homogeneous polynomial of degree k on $\mathbf{R}^n$ too.

Let $b_1(x), \dots, b_n(x)$ be n more real-valued linear functions on $\mathbf{R}^n$, and let b and L_b be as before. Observe that

$$c_j = L_a(b_j) - L_b(a_j) \tag{5.3.4}$$

is a real-valued linear function on $\mathbf{R}^n$ for each $j = 1, \dots, n$, as in the preceding paragraph. If c and L_c are as before, then L_c corresponds to the commutator of L_a and L_b, as in Section 2.3.

Let $(b_{j,l})$ and $(c_{j,l})$ be the matrices corresponding to b and c, respectively, as before. Clearly

$$L_a(b_j) = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{k,l}\, x_l\, \frac{\partial b_j}{\partial x_k} = \sum_{k=1}^{n} \sum_{l=1}^{n} a_{k,l} b_{j,k}\, x_l \tag{5.3.5}$$

for each $j = 1, \ldots, n$. Similarly,

$$L_b(a_j) = \sum_{k=1}^{n} \sum_{l=1}^{n} b_{k,l}\, a_{j,k}\, x_l \tag{5.3.6}$$

for each $j = 1, \ldots, n$. It follows that

$$c_{j,l} = \sum_{k=1}^{n} b_{j,k}\, a_{k,l} - \sum_{k=1}^{n} a_{j,k}\, b_{k,l} \tag{5.3.7}$$

for each $j, l = 1, \ldots, n$.

5.4 Exponentiating real matrices

Let n be a positive integer, and let A be a linear mapping from $\mathbf{R}^n$ into itself. This corresponds to an $n \times n$ matrix of real numbers in a standard way, as in the previous section. Of course, the composition of two linear mappings on $\mathbf{R}^n$ is another linear mapping on $\mathbf{R}^n$. It is well known and not difficult to see that this corresponds to matrix multiplication of the corresponding matrices.

If j is a positive integer, then A^j denotes the composition of A with itself a total of $j - 1$ times, so that there are j factors of A. This is interpreted as being the identity mapping I on $\mathbf{R}^n$ when $j = 0$. One would like to define the exponential of A by

$$\exp A = \sum_{j=0}^{\infty} (1/j!)\, A^j, \tag{5.4.1}$$

as another linear mapping on $\mathbf{R}^n$.

More precisely, it is well known and not difficult to show that there is a nonnegative real number C such that

$$|A(v)| \le C\, |v| \tag{5.4.2}$$

for every $v \in \mathbf{R}^n$. The smallest such C is known as the *operator norm* of A with respect to the standard Euclidean norm on $\mathbf{R}^n$. It follows that

$$|A^j(v)| \le C^j\, |v| \tag{5.4.3}$$

for every $j \ge 1$ and $v \in \mathbf{R}^n$. This also works with $j = 0$, and C^j interpreted as being equal to 1, as usual.

If $v \in \mathbf{R}^n$, then

$$\sum_{j=0}^{\infty} (1/j!)\, C^j\, |v| \tag{5.4.4}$$

is a convergent series of nonnegative real numbers, with sum equal to

$$(\exp C)\, |v|, \tag{5.4.5}$$

because of the usual series expansion for $\exp C$. It follows that

$$\sum_{j=0}^{\infty} (1/j!)\, |A^j(v)| \tag{5.4.6}$$

is a convergent series of nonnegative real numbers, with sum less than or equal to (5.4.5), because of (5.4.3) and the comparison test. Let $(A^j(v))_l$ be the lth coordinate of $A^j(v) \in \mathbf{R}^n$ for every $l = 1, \ldots, n$, so that

$$|(A^j(v))_l| \le |A^j(v)| \tag{5.4.7}$$

for each $j \ge 0$ and $l = 1, \ldots, n$. Thus

$$\sum_{j=0}^{\infty} (1/j!)\, |(A^j(v))_l| \tag{5.4.8}$$

is a convergent series of nonnegative real numbers for every $l = 1, \ldots, n$. This means that

$$\sum_{j=0}^{\infty} (1/j!)\, (A^j(v))_l \tag{5.4.9}$$

is an absolutely convergent series of real numbers for every $l = 1, \ldots, n$.

We would like to put

$$(\exp A)(v) = \sum_{j=0}^{\infty} (1/j!)\, A^j(v), \tag{5.4.10}$$

as an element of $\mathbf{R}^n$. The lth coordinate of the right side is equal to (5.4.9) for every $l = 1, \ldots, n$. It is easy to see that this defines a linear mapping from $\mathbf{R}^n$ into itself. One could also look at this in terms of matrices, where the entries of the matrix corresponding to $\exp A$ can be expressed as absolutely convergent series of real numbers.

Suppose that v is an eigenvector of A with eigenvalue $\lambda \in \mathbf{R}$, so that

$$A(v) = \lambda\, v. \tag{5.4.11}$$

This implies that

$$A^j(v) = \lambda^j\, v \tag{5.4.12}$$

for every $j \geq 0$. It follows that

$$(\exp A)(v) = (\exp \lambda)\, v. \tag{5.4.13}$$

Let T be a one-to-one linear mapping from $\mathbf{R}^n$ onto itself, so that the inverse mapping T^{-1} is linear on $\mathbf{R}^n$ too. It is easy to see that

$$T \circ A^j \circ T^{-1} = (T \circ A \circ T^{-1})^j \tag{5.4.14}$$

for every $j \geq 0$. This means that

$$T \circ (\exp A) \circ T^{-1} = \exp(T \circ A \circ T^{-1}). \tag{5.4.15}$$

5.5 Exponentials of sums

Let n be a positive integer, and let A, B be linear mappings from $\mathbf{R}^n$ into itself. Suppose that A and B commute on $\mathbf{R}^n$, so that

$$A \circ B = B \circ A. \tag{5.5.1}$$

If l is a positive integer, then one can check that

$$(A + B)^l = \sum_{j=0}^{l} \binom{l}{j} A^j \circ B^{l-j}, \tag{5.5.2}$$

as in the binomial theorem.

This implies that

$$\begin{aligned} \exp(A + B) &= \sum_{l=0}^{\infty} (1/l!)\, (A + B)^l \\ &= \sum_{l=0}^{\infty} \Big(\sum_{j=0}^{l} (1/j!)\, (1/(l-j)!)\, A^j \circ B^{l-j} \Big). \end{aligned} \tag{5.5.3}$$

The right side corresponds to the Cauchy product of the series used to define $\exp A$ and $\exp B$. In particular, this means that the same terms are being summed, but in different ways. One can use this to show that

$$\exp(A + B) = (\exp A) \circ (\exp B) \tag{5.5.4}$$

under these conditions. More precisely, this also uses absolute convergence of the sums, to ensure that the different ways of arranging the sums lead to the same results.

Note that $\exp A$ automatically commutes with A. Similarly, if A commutes with B, then $\exp A$ commutes with B. Of course, if $A = 0$, then $\exp A = I$. If A is any linear mapping on $\mathbf{R}^n$, then

$$(\exp A) \circ (\exp(-A)) = (\exp(-A)) \circ (\exp A) = I, \tag{5.5.5}$$

by (5.5.4). This implies that $\exp A$ is invertible on $\mathbf{R}^n$, with inverse equal to $\exp(-A)$.

Let A, B be any two linear mappings on $\mathbf{R}^n$, and let A', B' be the linear mappings corresponding to them as in Section 1.15. If $v, w \in \mathbf{R}^n$, then

$$\begin{aligned} (A \circ B)(v) \cdot w = A(B(v)) \cdot w &= B(v) \cdot A'(w) \\ &= v \cdot B'(A'(w)) = v \cdot (B' \circ A')(w). \end{aligned} \tag{5.5.6}$$

This means that

$$(A \circ B)' = B' \circ A'. \tag{5.5.7}$$

In particular,

$$(A^j)' = (A')^j \tag{5.5.8}$$

for each $j \geq 0$. It follows that

$$(\exp A)' = \exp(A'). \tag{5.5.9}$$

If

$$A' = -A, \tag{5.5.10}$$

then we get that

$$(\exp A)' = \exp(A') = \exp(-A) = (\exp A)^{-1}. \tag{5.5.11}$$

This means that $\exp A$ is an orthogonal transformation on $\mathbf{R}^n$, as in Section 1.15.

5.6 The exponential of $t\,A$

Let n be a positive integer, let A be a linear mapping from $\mathbf{R}^n$ into itself, and let t be a real number. Of course, $t\,A$ may be considered as a linear mapping on $\mathbf{R}^n$, with $(t\,A)(v) = t\,A(v)$ for every $v \in \mathbf{R}^n$. Thus the exponential of $t\,A$ may be defined as before, so that

$$\exp(t\,A) = \sum_{j=0}^{\infty} (1/j!)\, t^j\, A^j. \tag{5.6.1}$$

This may be considered as a power series in t, whose coefficients are linear mappings on $\mathbf{R}^n$. If $v \in \mathbf{R}^n$, then

$$(\exp(t\,A))(v) = \sum_{j=0}^{\infty} (1/j!)\, t^j\, A^j(v) \tag{5.6.2}$$

may be considered as a power series in t, with coefficients in $\mathbf{R}^n$.

More precisely, for each $l = 1, \ldots, n$, the lth coordinate of $(\exp(t\,A))(v)$ is

$$((\exp(t\,A))(v))_l = \sum_{j=0}^{\infty} (1/j!)\, t^j\, (A^j(v))_l. \tag{5.6.3}$$

This is an absolutely convergent power series in t with coefficients in $\mathbf{R}$. Similarly, the entries of the matrix associated to $\exp(t\,A)$ may be expressed as absolutely convergent power series in t with real coefficients.

In particular, these are smooth functions of t on $\mathbf{R}$, by standard results about power series. We can differentiate these series termwise, to get that

$$\frac{d}{dt}((\exp(t\,A))(v)) = A((\exp(t\,A))(v)) \tag{5.6.4}$$

for every $v \in \mathbf{R}^n$. This can be expressed by

$$\frac{d}{dt}(\exp(t\,A)) = A \circ (\exp(t\,A)). \tag{5.6.5}$$

Let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a continuously-differentiable real or complex-valued function on U. Put

$$(L_A(u))(x) = \sum_{l=1}^{n} (A(x))_l \, \frac{\partial u}{\partial x_l}(x) \tag{5.6.6}$$

for each $x \in U$. This is the same as in Section 5.3, with different notation. This is related to the system of ordinary differential equations

$$w'(t) = A(w(t)), \tag{5.6.7}$$

as in Section 4.1, where $w(t)$ is a continuously-differentiable function on an interval in the real line with nonempty interior, and with values in $\mathbf{R}^n$. If $v \in \mathbf{R}^n$, then

$$w(t) = (\exp(t\,A))(v) \tag{5.6.8}$$

satisfies (5.6.7), as in (5.6.4).

Let I be an open interval in the real line, which may be unbounded, with

$$(\exp(t\,A))(v) \in U \tag{5.6.9}$$

for each $t \in I$. Under these conditions,

$$\frac{d}{dt} u((\exp(t\,A))(v)) = (L_A(u))((\exp(t\,A))(v)) \tag{5.6.10}$$

on I, as in Section 4.1. This can be used to analyze first-order semilinear equations on U involving L_A, as before.

5.7 Traces and determinants

Let n be a positive integer, and let $(a_{j,l})$ be an $n \times n$ matrix of real numbers. The *trace* of this matrix is defined as usual as

$$\sum_{j=1}^{n} a_{j,j}. \tag{5.7.1}$$

The *determinant* of $(a_{j,l})$ is defined in a standard way, that we shall not repeat here.

If A is a linear mapping from $\mathbf{R}^n$ into itself, then A corresponds to an $n \times n$ matrix $(a_{j,l})$ of real numbers in a standard way. The trace $\operatorname{tr} A$ and determinant $\det A$ of A are defined as the trace and determinant of $(a_{j,l})$, respectively.

Let B be another linear mapping from $\mathbf{R}^n$ into itself. It is well known and not difficult to verify that

$$\operatorname{tr}(A \circ B) = \operatorname{tr}(B \circ A). \tag{5.7.2}$$

It is also well known that

$$\det(A \circ B) = (\det A)\,(\det B). \tag{5.7.3}$$

If t is a real number, then $I + t\,A$ is another linear mapping from $\mathbf{R}^n$. It is clear from the definition of the determinant that

$$\det(I + t\,A) \tag{5.7.4}$$

is a polynomial in t of degree at most n. One can check that this polynomial is of the form

$$1 + (\operatorname{tr} A)\,t + \cdots, \tag{5.7.5}$$

where the additional terms are multiples of t^j, $2 \le j \le n$. This means that the derivative of (5.7.4) in t at $t = 0$ is equal to $\operatorname{tr} A$.

It is well known that

$$\det(\exp A) = \exp(\operatorname{tr} A). \tag{5.7.6}$$

One way to see this is to use calculus to show that

$$\det(\exp(t\,A)) = \exp(t\ \operatorname{tr} A) \tag{5.7.7}$$

for every $t \in \mathbf{R}$. Note that both sides of this equation are equal to 1 at $t = 0$.

The right side of (5.7.7) satisfies the differential equation

$$f'(t) = (\operatorname{tr} A)\,f(t) \tag{5.7.8}$$

on $\mathbf{R}$. We would like to check that the left side of (5.7.7) satisfies the same differential equation. If we can do that, then (5.7.7) follows, by standard arguments.

One can verify directly that the left side of (5.7.7) satisfies (5.7.8) at $t = 0$. Let $t_0 \in \mathbf{R}$ be given, and observe that

$$\exp(t\,A) = (\exp((t - t_0)\,A)) \circ (\exp(t_0\,A)) \tag{5.7.9}$$

for every $t \in \mathbf{R}$, as in Section 5.5. One can use this to obtain that the left side of (5.7.7) satisifes (5.7.8) at t_0 from the analogous statement at 0.

5.8 Exponentiating complex matrices

Let m be a positive integer, and let A be a linear mapping from $\mathbf{C}^m$ into itself, as a vector space over the complex numbers. This corresponds to an $m \times m$ matrix of complex numbers in the usual way. The composition of two linear mappings on $\mathbf{C}^m$ corresponds to matrix multiplication of the corresponding matrices of complex numbers.

If j is a positive integer, then A^j denotes the composition of A with itself a total of $j - 1$ times, so that there are j factors of A, and which is interpreted as being the identity mapping I on $\mathbf{C}^m$ when $j = 0$. As in the real case, it is well known and not difficult to show that there is a nonnegative real number C such that

$$|A(v)| \le C\,|v| \tag{5.8.1}$$

for every $v \in \mathbf{C}^m$, and the smallest such C is the *operator norm* of A with respect to the standard Euclidean norm on $\mathbf{C}^m$. This implies that

$$|A^j(v)| \le C^j\,|v| \tag{5.8.2}$$

for every $j \ge 0$ and $v \in \mathbf{C}^m$.

One would like to define the exponential of A as another linear mapping on $\mathbf{C}^m$ by

$$\exp A = \sum_{j=0}^{\infty} (1/j!)\, A^j, \tag{5.8.3}$$

as in Section 5.4. More precisely, if $v \in \mathbf{C}^m$, then we would like to put

$$(\exp A)(v) = \sum_{j=0}^{\infty} (1/j!)\, A^j(v), \tag{5.8.4}$$

as an element of $\mathbf{C}^m$, as before. This means that for each $l = 1, \ldots, m$, the lth coordinate of $(\exp A)(v)$ is equal to

$$((\exp A)(v))_l = \sum_{j=0}^{\infty} (1/j!)\, (A^j(v))_l. \tag{5.8.5}$$

The right side is an absolutely convergent series of complex numbers, by the comparison test. This defines a linear mapping on $\mathbf{C}^m$, and the entries of the corresponding matrix can be expressed as absolutely convergent series of complex numbers in an analogous way.

Note that a linear mapping from $\mathbf{R}^m$ into itself, as a vector space over the real numbers, has a unique extension to a linear mapping from $\mathbf{C}^m$ into itself, as a vector space over the complex numbers. Both linear mappings correspond to the same $m \times m$ matrix of real numbers, which may be considered as an $m \times m$ matrix of complex numbers too. The exponential of the linear mapping on $\mathbf{C}^m$ is the same as the extension of the exponential of the linear mapping on $\mathbf{R}^m$ to a linear mapping on $\mathbf{C}^m$.

Suppose that $v \in \mathbf{C}^m$ is an eigenvector of A with eigenvalue $\lambda \in \mathbf{C}$, so that

$$A(v) = \lambda\, v. \tag{5.8.6}$$

It is easy to see that

$$(\exp A)(v) = (\exp \lambda)\, v, \tag{5.8.7}$$

as before. If T is a one-to-one linear mapping from $\mathbf{C}^m$ onto itself, then

$$T \circ (\exp A) \circ T^{-1} = \exp(T \circ A \circ T^{-1}), \tag{5.8.8}$$

as before.

Let B be another linear mapping from $\mathbf{C}^m$ into itself, and suppose that A and B commute on $\mathbf{C}^m$, so that

$$A \circ B = B \circ A. \tag{5.8.9}$$

Under these conditions,

$$\exp(A + B) = (\exp A) \circ (\exp B), \tag{5.8.10}$$

as in Section 5.5. We also have that $\exp A$ commutes with B in this case, as before. If we take $B = -A$, then we get that $\exp A$ is invertible on $\mathbf{C}^m$, with inverse equal to $\exp(-A)$, as before.

The trace and determinant of an $m \times m$ matrix of complex numbers can be defined in the same way as for real numbers. Similarly, the trace and determinant of A are defined to be the trace and determinant of the matrix corresponding to A, repectively. These satisfy the same basic properties as in the real case. In particular, it is well known that

$$\det(\exp A) = \exp(\operatorname{tr} A), \tag{5.8.11}$$

which can be shown using an argument like the one in Section 5.7. Alternatively, one can use results from linear algebra to reduce to the case where A corresponds to an upper triangular matrix, for which (5.8.11) can be verified more directly.

5.9 More on $\mathbf{C}^m$

Let m be a positive integer, and let $\langle v, w\rangle = \langle v, w\rangle_{\mathbf{C}^m}$ be the standard inner product on $\mathbf{C}^m$, as in Section 2.6. If $v, w \in \mathbf{C}^m$, then

$$\begin{aligned} |v + w|^2 = \langle v + w, v + w\rangle &= \langle v, v\rangle + \langle v, w\rangle + \langle w, v\rangle + \langle w, w\rangle \\ &= |v|^2 + 2\ \mathrm{Re}\langle v, w\rangle + |w|^2. \end{aligned} \tag{5.9.1}$$

If we replace w with $i\, w$, then we get that

$$|v + iw|^2 = |v|^2 + 2\,\mathrm{Re}(-i\,\langle v, w\rangle) + |w|^2 = |v|^2 + 2\ \mathrm{Im}\langle v, w\rangle + |w|^2. \tag{5.9.2}$$

It follows that

$$(5.9.3)\ \langle v, w\rangle = (1/2)\,(|v+w|^2 - |v|^2 - |w|^2) + (i/2)\,(|v+i\,w|^2 - |v|^2 - |w|^2).$$

This is another *polarization identity*.

Let T be a linear mapping from $\mathbf{C}^m$ into itself, as a vector space over the complex numbers. As in the real case,

$$\ker T = \{v \in \mathbf{C}^m : T(v) = 0\} \tag{5.9.4}$$

is a linear subspace of $\mathbf{C}^m$, called the *kernel* of T. This is equal to $\{0\}$ if and only if T is one-to-one, as before. It is is well known that T is one-to-one on $\mathbf{C}^m$ if and only if T maps $\mathbf{C}^m$ onto itself, in which case the inverse mapping T^{-1} is linear on $\mathbf{C}^m$ too.

A one-to-one linear mapping T from $\mathbf{C}^m$ onto itself is said to be *unitary* if

$$\langle T(v), T(w)\rangle = \langle v, w\rangle \tag{5.9.5}$$

for every $v, w \in \mathbf{C}^m$. Note that this implies that T^{-1} is unitary as well. In this case, we can take $v = w$ in (5.9.5), to get that

$$|T(v)| = |v|. \tag{5.9.6}$$

Conversely, if (5.9.5) holds for every $v \in \mathbf{C}^m$, then (5.9.5) holds for every v. w in $\mathbf{C}^m$, because of the polarization identity (5.9.3). Of course, (5.9.6) implies that $\ker T = \{0\}$.

If T is any linear mapping from $\mathbf{C}^m$ into itself, then it is well known that there is a unique linear mapping T^* from $\mathbf{C}^m$ into itself such that

$$\langle T(v), w\rangle = \langle v, T^*(w)\rangle \tag{5.9.7}$$

for every $v, w \in \mathbf{C}^m$. This is called the *adjoint* of T. As in the real case, every linear mapping from $\mathbf{C}^m$ into itself corresponds to an $m \times m$ matrix of complex numbers in a standard way. The matrix associated to T^* is obtained by taking the complex conjugates of the entries of the transpose of the matrix associated to T.

If T is a unitary transformation on $\mathbf{C}^m$, then one can verify that T^* is the same as the inverse of T. Conversely, if T is an invertible linear mapping on $\mathbf{C}^m$, with inverse equal to T^*, then T is a unitary transformation on $\mathbf{C}^m$.

Let A, B be linear mappings from $\mathbf{C}^m$ into itself, and let t be a complex number. Under these conditions, $A + B$ and $t\,A$ are linear mappings on $\mathbf{C}^m$, and one can check that

$$(A + B)^* = A^* + B^* \tag{5.9.8}$$

and

$$(t\,A)^* = \bar{t}\,A^*. \tag{5.9.9}$$

One can also verify that

$$(A \circ B)^* = B^* \circ A^*. \tag{5.9.10}$$

This implies that

$$(A^j)^* = (A^*)^j \tag{5.9.11}$$

for each nonnegative integer j, so that

$$(\exp A)^* = \exp(A^*). \tag{5.9.12}$$

If

$$A^* = -A, \tag{5.9.13}$$

then it follows that

$$(\exp A)^* = \exp(A^*) = \exp(-A) = (\exp A)^{-1}, \tag{5.9.14}$$

so that $\exp A$ is a unitary transformation on $\mathbf{C}^m$.

A linear mapping A on $\mathbf{C}^m$ is said to be *self-adjoint* with respect to the standard inner product on $\mathbf{C}^m$ if

$$A^* = A. \tag{5.9.15}$$

If T is any linear mapping on $\mathbf{C}^m$, then it is easy to see that

$$(T^*)^* = T. \tag{5.9.16}$$

One can use this to check that

$$A = (1/2)\,(T + T^*) \tag{5.9.17}$$

and

$$B = (-i/2)\,(T - T^*) \tag{5.9.18}$$

are self-adjoint. Note that

$$T = A + i\,B. \tag{5.9.19}$$

5.10 The exponential of $z\,A$

Let m be a positive integer, let A be a linear mapping from $\mathbf{C}^m$ into itself, and let z be a complex number. Thus $z\,A$ is another linear mapping from $\mathbf{C}^m$ into itself, whose exponential

$$\exp(z\,A) = \sum_{j=0}^{\infty} (1/j!)\, z^j\, A^j \tag{5.10.1}$$

may be considered as a power series in z, with coefficients that are linear mappings on $\mathbf{C}^m$. If $v \in \mathbf{C}^m$, then

$$(\exp(z\,A))(v) = \sum_{j=0}^{\infty} (1/j!)\, z^j\, A^j(v) \tag{5.10.2}$$

may be considered as a power series in z, with coefficients in $\mathbf{C}^m$.

As in Section 5.6, the lth coordinate of $(\exp(z\,A))(v)$ is

$$((\exp(z\,A))(v))_l = \sum_{j=0}^{\infty} (1/j!)\, z^j\, (A^j(v))_l \tag{5.10.3}$$

for each $l = 1, \ldots, m$, which is an absolutely convergent power series in z with complex coefficients. Similarly, the entries of the matrix associated to $\exp(z\,A)$ may be expressed as absolutely convergent power series in z with complex coefficients. One can differentiate these series termwise, to get that they are holomorphic functions of z, with

$$\frac{\partial}{\partial z}((\exp(z\,A))(v)) = A((\exp(z\,A))(v)) \tag{5.10.4}$$

for every $v \in \mathbf{C}^m$. This can be expressed by

$$\frac{\partial}{\partial z}(\exp(z\,A)) = A \circ (\exp(z\,A)), \tag{5.10.5}$$

as before.

Let r be a nonnegative integer, and suppose that

$$A^{r+1} = 0 \tag{5.10.6}$$

on $\mathbf{C}^m$. In this case, A is said to be *nilpotent* on $\mathbf{C}^m$. It is well known that if A is nilpotent on $\mathbf{C}^m$, then one can take $r \le m - 1$. Of course, if (5.10.6) holds, then $A^j = 0$ when $j \ge r + 1$. This means that

$$\exp(z\,A) = \sum_{j=0}^{r} (1/j!)\, z^j\, A^j \tag{5.10.7}$$

is a polynomial in z, with coefficients that are linear mappings on $\mathbf{C}^m$.

Note that

$$\exp(c\,z\,I) = (\exp(c\,z))\, I \tag{5.10.8}$$

for every $c, z \in \mathbf{C}$, where I is the identity maping on $\mathbf{C}^m$. If A is any linear mapping on $\mathbf{C}^m$, then A commutes with $c\,I$ on $\mathbf{C}^m$. This implies that

$$\exp(z\,(c\,I + A)) = (\exp(c\,z\,I)) \circ (\exp(z\,A)) = (\exp(c\,z))\, \exp(z\,A). \tag{5.10.9}$$

5.11 Polynomials and differential operators

Let n be a positive integer, and remember that $\mathcal{P}(\mathbf{R}^n, \mathbf{R})$ and $\mathcal{P}(\mathbf{R}^n, \mathbf{C})$ are the spaces of polynomials on $\mathbf{R}^n$ with real and complex coefficients, respectively, as in Section 2.9. If k is a nonnegative integer, then let $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ and $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$ be the spaces of polynomials on $\mathbf{R}^n$ with real and complex coefficients and degree less than or equal to k, respectively. These are linear subspaces of $\mathcal{P}(\mathbf{R}^n, \mathbf{R})$ and $\mathcal{P}(\mathbf{R}^n, \mathbf{C})$, as vector spaces over $\mathbf{R}$ and $\mathbf{C}$, respectively.

Consider the collection of monomials x^β, where β is a multi-index with order $|\beta| \le k$. This collection is a basis for $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ and $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$, as vector spaces over $\mathbf{R}$ and $\mathbf{C}$, respectively. In particular, $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ and $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$ have the same finite dimension, as vector spaces over $\mathbf{R}$ and $\mathbf{C}$, respectively.

Let N be a nonnegative integer, and suppose that a_α is a polynomial on $\mathbf{R}^n$ with real or complex coefficients for each multi-index α with $|\alpha| \le N$, so that

$$L = \sum_{|\alpha| \le N} a_\alpha \, \partial^\alpha \tag{5.11.1}$$

defines a differential operator on $\mathbf{R}^n$ with polynomial coefficients, as in Section 2.9. Remember that L maps $\mathcal{P}(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}(\mathbf{R}^n, \mathbf{C})$ into itself, as appropriate. Suppose that

$$\deg a_\alpha \le |\alpha| \tag{5.11.2}$$

for each α, $|\alpha| \le N$. If p is a polynomial on $\mathbf{R}^n$ with real or complex coefficients, as appropriate, then

$$\deg L(p) \le \deg p. \tag{5.11.3}$$

This means that L maps $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$ into itself for each $k \ge 0$, as appropriate.

Similarly, let c be a nonnegative integer, and suppose that

$$\deg a_\alpha \le |\alpha| - c \tag{5.11.4}$$

for each α, $|\alpha| \le N$. This is interpreted to mean that

$$a_\alpha = 0 \text{ when } |\alpha| < c. \tag{5.11.5}$$

If p is a polynomial on $\mathbf{R}^n$ with real or complex coefficients, as appropriate, then

$$\deg L(p) \le \deg p - c. \tag{5.11.6}$$

As before, this means that

$$L(p) = 0 \text{ when } \deg p < c. \tag{5.11.7}$$

If j is a positive integer, then we get that

$$\deg L^j(p) \le \deg p - c\, j. \tag{5.11.8}$$

This means that

$$L^j(p) = 0 \text{ when } \deg p < c\, j, \tag{5.11.9}$$

as usual. Suppose that $c \ge 1$, and let k be a nonnegative integer. If

$$k < c\, j, \tag{5.11.10}$$

then it follows that

$$L^j = 0 \text{ on } \mathcal{P}^k(\mathbf{R}^n, \mathbf{R}) \text{ or } \mathcal{P}^k(\mathbf{R}^n, \mathbf{C}), \tag{5.11.11}$$

as appropriate. Thus the restriction of L to $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$, as appropriate, is nilpotent under these conditions.

5.12 Some related differential equations

Let n be a positive integer, let N be a nonnegative integer, and let L be a differential operator of order less than or equal to N on $\mathbf{R}^n$ with polynomial coefficients, as in the previous section. Suppose that the coefficients satisfy (5.11.2) for each α, $|\alpha| \le N$, and let k be a nonnegative integer. Thus L maps $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$ into itself, as before. Let L_k be the restriction of L to $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$, as appropriate.

Let $m = m(k)$ be the number of multi-indices β with order $|\beta| \le k$. We can identify $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ and $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$ with $\mathbf{R}^m$ and $\mathbf{C}^m$, respectively, by listing the coefficients of a polynomial on $\mathbf{R}^n$ with degree less than or equal to k in any reasonable way. This means that we can identify L_k with a linear mapping from $\mathbf{R}^m$ or $\mathbf{C}^m$ into itself, as appropriate.

If $t \in \mathbf{R}$, then we can define the exponential of $t\,L_k$ as a linear mapping on $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$, as appropriate, as before. Let p be a polynomial on $\mathbf{R}^n$ with real or complex coefficients, as appropriate, and of degree less than or equal to k. Thus

$$(\exp(t\,L_k))(p) \tag{5.12.1}$$

is another polynomial on $\mathbf{R}^n$ with real or complex coefficients, as appropriate, and degree less than or equal to k. Of course, the coefficients of (5.12.1), as a polynomial on $\mathbf{R}^n$, depend on t, and in fact they are smooth functions of t. It follows that

$$u(x,t) = ((\exp(t\,L_k))(p))(x) \tag{5.12.2}$$

is smooth as a function of (x,t) on $\mathbf{R}^n \times \mathbf{R}$, which we can identify with $\mathbf{R}^{n+1}$.

Suppose for the moment that the coefficients of L satisfy (5.11.4) for some $c \ge 1$. This implies that L_k is nilpotent on $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$, as appropriate, as in the previous section. It follows that $\exp(t\,L_k)$ is a polynomial in t with coefficients that are linear mappings on $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$, as appropriate, as in Section 5.10. This means that (5.12.2) is a polynomial in x and t in this case.

Note that

$$u(x,0) = p(x) \tag{5.12.3}$$

for every $x \in \mathbf{R}^n$. We also have that

$$\frac{\partial}{\partial t}((\exp(t\,L_k))(p)) = L_k((\exp(t\,L_k))(p)), \tag{5.12.4}$$

as before. This means that

$$\frac{\partial u}{\partial t} = L(u) \tag{5.12.5}$$

on $\mathbf{R}^n \times \mathbf{R}$.

5.13 Some additional related equations

Let us continue with the same notation and hypotheses as at the beginning of the previous section. Suppose now that we are interested in the partial differential equation

$$\frac{\partial^2 u}{\partial t^2} = L(u) \tag{5.13.1}$$

on $\mathbf{R}^n \times \mathbf{R}$. If we put

$$v = \frac{\partial u}{\partial t}, \tag{5.13.2}$$

then (5.13.1) is the same as saying that

$$\frac{\partial v}{\partial t} = L(u). \tag{5.13.3}$$

Let us consider (5.13.2) and (5.13.3) as a system of partial differential equations in u and v on $\mathbf{R}^n \times \mathbf{R}$.

Of course, we can identify $\mathbf{R}^m \times \mathbf{R}^m$ and $\mathbf{C}^m \times \mathbf{C}^m$ with $\mathbf{R}^{2m}$ and $\mathbf{C}^{2m}$, respectively. Similarly, we can identify

$$\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}) \times \mathcal{P}^k(\mathbf{R}^n, \mathbf{R}) \tag{5.13.4}$$

and

$$\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}) \times \mathcal{P}^k(\mathbf{R}^n, \mathbf{C}) \tag{5.13.5}$$

with $\mathbf{R}^{2m}$ and $\mathbf{C}^{2m}$, respectively, using the analogous identifications mentioned in the previous section. Let T_k be the mapping from (5.13.4) or (5.13.5) into itself, as appropriate, defined by

$$T_k(p, q) = (q, L_k(p)) \tag{5.13.6}$$

for every $p, q \in \mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$, as appropriate. Observe that

$$T_k^2(p, q) = T_k(T_k(p, q)) = T_k(q, L_k(p)) = (L_k(p), L_k(q)) \tag{5.13.7}$$

for all such p, q. We can identify T_k with a linear mapping from $\mathbf{R}^{2m}$ or $\mathbf{C}^{2m}$ into itself, as before.

If $t \in \mathbf{R}$, then we can define the exponential of $t\,T_k$ as a linear mapping on (5.13.4) or (5.13.5), as appropriate, in the usual way. Let p, q be elements of $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$, as appropriate, so that

$$(\exp(t\,T_k))(p, q) \tag{5.13.8}$$

is an element of (5.13.4) or (5.13.5), as appropriate. Let $u(\cdot, t)$, $v(\cdot, t)$ be the elements of $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$, as appropriate, such that

$$(\exp(t\,T_k))(p, q) = (u(\cdot, t), v(\cdot, t)). \tag{5.13.9}$$

The coefficients of $u(x, t)$ and $v(x, t)$, as polynomials in x on $\mathbf{R}^n$, are smooth functions of t, as before. This implies that $u(x, t)$ and $v(x, t)$ are smooth as functions of (x, t) on $\mathbf{R}^n \times \mathbf{R}$, which we can identify with $\mathbf{R}^{n+1}$, as usual.

Note that

$$\frac{\partial}{\partial t}((\exp(t\,T_k))(p, q)) = T_k((\exp(t\,T_k))(p, q)), \tag{5.13.10}$$

as before. This means that

$$\frac{\partial}{\partial t}(u(\cdot,t),v(\cdot,t)) = T_k(u(\cdot,t),v(\cdot,t)) = (v(\cdot,t),L(u(\cdot,t))), \tag{5.13.11}$$

which is the same as saying that u and v satisfy (5.13.2) and (5.13.3). We also have that

$$u(\cdot,0) = p,\ v(\cdot,0) = q. \tag{5.13.12}$$

Suppose that the coefficients of L satisfy (5.11.4) for some $c \geq 1$, so that L_k is nilpotent on $\mathcal{P}^k(\mathbf{R}^n,\mathbf{R})$ or $\mathcal{P}^k(\mathbf{R}^n,\mathbf{C})$, as appropriate, as before. This implies that T_k is nilpotent on (5.13.4) or (5.13.5), as appropriate, because of (5.13.7). This means that $\exp(t\,T_k)$ is a polynomial in t with coefficients that are linear mappings on (5.13.4) or (5.13.5), as appropriate, as in Section 5.10. It follows that $u(x,t)$ and $v(x,t)$ are polynomials in x and t under these conditions.

5.14 Some products with $\exp(b \cdot x)$

Let n be a positive integer, and let $b \in \mathbf{R}^n$ or $\mathbf{C}^n$ be given. Also let N be a nonnegative integer, and let p be a polynomial on $\mathbf{R}^n$ with real or complex coefficients of degree less than or equal to N. Thus

$$p_b(x) = p(x+b) \tag{5.14.1}$$

can be expressed as a polynomial of degree less than or equal to N with real or complex coefficients, as appropriate, as in Section 2.5.

Let $p(\partial)$ and $p_b(\partial)$ be the differential operators corresponding to p and p_b as in Section 1.7, respectively. If f is a continuously-differentiable real or complex-valued function on $\mathbf{R}^n$, then

$$\frac{\partial}{\partial x_j}(f(x)\,\exp(b\cdot x)) = \Big(\frac{\partial f}{\partial x_j}(x) + b_j\,f(x)\Big)\,\exp(b\cdot x). \tag{5.14.2}$$

If f is N-times continuously differentiable on $\mathbf{R}^n$, then we get that

$$p(\partial)(f(x)\,\exp(b\cdot x)) = (p_b(\partial)(f))(x)\,\exp(b\cdot x). \tag{5.14.3}$$

If $b \in \mathbf{R}^n$, then let

$$\mathcal{P}(\mathbf{R}^n,\mathbf{R})\,\exp(b\cdot x) \tag{5.14.4}$$

be the space of functions on $\mathbf{R}^n$ of the form

$$q(x)\,\exp(b\cdot x), \tag{5.14.5}$$

where $q \in \mathcal{P}(\mathbf{R}^n,\mathbf{R})$. This is a linear subspace of $C^\infty(\mathbf{R}^n,\mathbf{R})$, as a vector space over the real numbers. If p is a polynomial with real coefficients, then p_b is a polynomial with real coefficients as well. In this case, $p(\partial)$ maps (5.14.4) into itself, because of (5.14.3).

Similarly, if $b \in \mathbf{C}^n$, then let

$$\mathcal{P}(\mathbf{R}^n, \mathbf{C}) \exp(b \cdot x) \tag{5.14.6}$$

be the space of functions on $\mathbf{R}^n$ of the form (5.14.5), with $q \in \mathcal{P}(\mathbf{R}^n, \mathbf{C})$. This is a linear subspace of $C^\infty(\mathbf{R}^n, \mathbf{C})$, as a vector space over the complex numbers. We also have that $p(\partial)$ maps (5.14.6) into itself, because of (5.14.3), as before.

Let k be a nonnegative integer, and if $b \in \mathbf{R}^n$, then let

$$\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}) \exp(b \cdot x) \tag{5.14.7}$$

be the space of functions on $\mathbf{R}^n$ of the form (5.14.5), with $q \in \mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$. This is a linear subspace of (5.14.4), as a vector space over the real numbers. If p is a polynomial with real coefficients, then $p(\partial)$ maps (5.14.7) into itself, because of (5.14.3) again.

If $b \in \mathbf{C}^n$, then let

$$\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}) \exp(b \cdot x) \tag{5.14.8}$$

be the space of functions on $\mathbf{R}^n$ of the form (5.14.5), with $q \in \mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$. This is a linear subspace of (5.14.6), as a vector space over the complex numbers. As usual, $p(\partial)$ maps (5.14.8) into itself, because of (5.14.3).

Suppose that

$$p_b(0) = p(b) = 0. \tag{5.14.9}$$

If q is a polynomial on $\mathbf{R}^n$ with real or complex coefficients, then

$$\deg(p_b(\partial))(q) \le \deg q - 1. \tag{5.14.10}$$

This implies that the restriction of $p_b(\partial)$ to $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$ is nilpotent, as in Section 5.11. It follows that the restriction of $p(\partial)$ to (5.14.8) is nilpotent, because of (5.14.3). If $b \in \mathbf{R}^n$, and p is a polynomial with real coefficients, then the restriction of $p(\partial)$ to (5.14.7) is nilpotent, for the same reasons.

5.15 Some remarks about derivatives

Let m be a positive integer, and let I be an interval in the real line, which may be unbounded, and which has nonempty interior. One can define continuity of a mapping from I into $\mathbf{C}^m$ in the usual way, using the restriction of the standard Euclidean metric on $\mathbf{R}$ to I, and the standard Euclidean metric on $\mathbf{C}^m$. It is well known and not difficult to see that this is equivalent to the continuity of the corresponding m component functions, as complex-valued functions on I. Similarly, a complex-valued function on I is continuous if and only if its real and imaginary parts are continuous.

Suppose that for each $t \in I$, $A(t)$ is a linear mapping from $\mathbf{C}^m$ into itself, as a vector space over the complex numbers. The continuity of $A(t)$ as a function on I with values in the space $\mathcal{L}(\mathbf{C}^m)$ of linear mappings from $\mathbf{C}^m$ into itself can also be defined in the usual way, using the restrction of the standard Euclidean metric on $\mathbf{R}$, and a suitable version of the standard Euclidean metric on $\mathcal{L}(\mathbf{C}^m)$. More prcisely, we can use the standard correspondence between elements of $\mathcal{L}(\mathbf{C}^m)$ and $m \times m$ matrices of complex numbers to identify $\mathcal{L}(\mathbf{C}^m)$ with $\mathbf{C}^{m^2}$, and use the standard Euclidean metric there. The continuity of $A(t)$ on I is equivalent to the continuity of the m^2 complex-valued functions on I corresponding to the matrix entries of $A(t)$. This is equivalent to the continuity of

$$(A(t))(v) \tag{5.15.1}$$

for each $v \in \mathbf{C}^m$, as a function of $t \in I$ with values in $\mathbf{C}^m$.

One can define differentiability of a mapping from I into $\mathbf{C}^m$ directly, using one-sided derivatives at any endpoints of I. This is equivalent to the differentiability of the m component functions, as complex-valued functions on I. The differentiability of a complex-valued function on I is equivalent to the differentiability of its real and imaginary parts.

Differentiability of $A(t)$ on I can be defined directly, and is equivalent to the differentiability of the m^2 complex-valued functions on I corresponding to the matrix entries of $A(t)$. This is equivalent as well to the differentiability of (5.15.1) for each $v \in \mathbf{C}^m$, as a function of $t \in I$ with values in $\mathbf{C}^m$.

Let $v(t)$ be a function on I with values in $\mathbf{C}^m$, so that

$$(A(t))(v(t)) \tag{5.15.2}$$

is an element of $\mathbf{C}^m$ for each $t \in I$. Of course, the components of (5.15.2) can be expressed as a sum or products of matrix entries of $A(t)$ and components of $v(t)$ in the usual way. If $v(t)$ is continuous at a point $t_0 \in I$, and if $A(t)$ is continuous at t_0, then (5.15.2) is continuous at t_0 too, as a function of $t \in I$ with values in $\mathbf{C}^m$. If $v(t)$ is differentiable at t_0, and $A(t)$ is differentiable at t_0, then (5.15.2) is differentiable at t_0, with derivative equal to

$$(A'(t_0))(v(t_0)) + (A(t_0))(v'(t_0)). \tag{5.15.3}$$

This is basically another version of the product rule.

Let B be a linear mapping from $\mathbf{C}^m$ into itself, and consider

$$A(t) = \exp(-t\,B). \tag{5.15.4}$$

This is a differentiable function of $t \in \mathbf{R}$ with values in $\mathcal{L}(\mathbf{C}^m)$, with derivative

$$A'(t) = -B \circ A(t) = -A(t) \circ B. \tag{5.15.5}$$

Suppose that $v(t)$ is differentiable on I, and put

$$w(t) = (A(t))(v(t)) = (\exp(-t\,B))(v(t)) \tag{5.15.6}$$

for each $t \in I$. Thus $w(t)$ is differentiable on I, with

$$w'(t) = -B(w(t)) + (A(t))(v'(t)), \tag{5.15.7}$$

by (5.15.5). Note that

$$v(t) = (\exp(t\,B))(w(t)) \tag{5.15.8}$$

for each $t \in I$.

Suppose for the moment that

$$v'(t) = B(v(t)) \tag{5.15.9}$$

on I. In this case,

$$w'(t) = 0 \tag{5.15.10}$$

on I, by (5.15.7). This means that $w(t)$ is constant on I, because of the analogous statement for real-valued functions.

Similarly, consider the differential equation

$$v'(t) = B(v(t)) + z(t), \tag{5.15.11}$$

where $z(t)$ is a function of $t \in I$ with values in $\mathbf{C}^m$. This corresponds to the differential equation

$$w'(t) = (\exp(-t\,B))(z(t)) \tag{5.15.12}$$

on I. Note that the right side is continuous on I when $z(t)$ is continuous on I.

Chapter 6

More on harmonic functions

Some nice references concerning harmonic functions include [10, 29, 32, 125], and some additional information may be found in [119]. See also [114, 123], for instance.

6.1 Some particular harmonic functions

It is well known and not difficult to verify that

$$|x|^{2-n} \tag{6.1.1}$$

is harmonic on $\mathbf{R}^n \setminus \{0\}$ when $n \geq 3$. This implies that

$$|x - a|^{2-n} \tag{6.1.2}$$

is harmonic on $\mathbf{R}^n \setminus \{a\}$ for every $a \in \mathbf{R}^n$ when $n \geq 3$.

Similarly, one can check that

$$\log |x| = (1/2) \log |x|^2 \tag{6.1.3}$$

is harmonic on $\mathbf{R}^2 \setminus \{0\}$. This means that

$$\log |x - a| \tag{6.1.4}$$

is harmonic on $\mathbf{R}^2 \setminus \{a\}$ for every $a \in \mathbf{R}^2$, as before.

If we put $z = x_1 + i\, x_2$, then we can express (6.1.3) as

$$(1/2) \log |z|^2. \tag{6.1.5}$$

Let $\partial/\partial z$ and $\partial/\partial \overline{z}$ be the differential operators defined in Section 2.2. Observe that

$$\frac{\partial}{\partial z}((1/2) \log |z|^2) = \frac{1}{2\,|z|^2} \frac{\partial}{\partial z}(|z|^2) = \frac{1}{2\,|z|^2} \frac{\partial}{\partial z}(z\,\overline{z}) = \frac{1}{2\,|z|^2}\,\overline{z} = \frac{1}{2\,z} \tag{6.1.6}$$

when $z \neq 0$.

It is well known and not difficult to check that

$$\frac{\partial}{\partial \overline{z}}\Big(\frac{1}{z}\Big) = 0 \tag{6.1.7}$$

for $z \neq 0$, which is to say that $1/z$ is holomorphic for $z \neq 0$. It follows that (6.1.3) is harmonic on $\mathbf{R}^2 \setminus \{0\}$, as in Section 2.2.

If $n \geq 3$ and $1 \leq j \leq n$, then

$$\begin{aligned} \frac{\partial}{\partial x_j}(|x|^{2-n}) &= \frac{\partial}{\partial x_j}((|x|^2)^{(2-n)/2}) \\ &= ((2-n)/2)\,(|x|^2)^{((2-n)/2)-1}\,(2\,x_j) = (2-n)\,\frac{x_j}{|x|^n} \end{aligned} \tag{6.1.8}$$

on $\mathbf{R}^n \setminus \{0\}$. Similarly,

$$\frac{\partial}{\partial x_j}(\log|x|) = \frac{\partial}{\partial x_j}((1/2)\,\log|x|^2) = (1/2)\,|x|^{-2}\,(2\,x_j) = \frac{x_j}{|x|^2} \tag{6.1.9}$$

on $\mathbf{R}^2 \setminus \{0\}$ for $j = 1, 2$, which is basically the same as (6.1.6). Note that these are harmonic functions too.

6.2 The mean-value property

Let $n \geq 2$ be an integer, and let V be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary. It is convenient to use $|V|$ for the n-dimensional volume of V, and $|\partial V|$ for the $(n-1)$-dimensional surface area of ∂V.

In particular, if $a \in \mathbf{R}^n$ and $r > 0$, then $|B(a,r)|$ denotes the volume of $B(a,r)$, and $|\partial B(a,r)|$ denotes the surface area of $\partial B(a,r)$. Note that

$$|B(a,r)| = r^n\,|B(0,1)| \tag{6.2.1}$$

and

$$|\partial B(a,r)| = r^{n-1}\,|\partial B(0,1)|. \tag{6.2.2}$$

Let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a twice continuously-differentiable real or complex-valued function on U that is harmonic on U. Also let $a \in U$ and $r > 0$ be given, with

$$\overline{B}(a,r) \subseteq U. \tag{6.2.3}$$

Under these conditions, it is well known that

$$u(a) = \frac{1}{|\partial B(a,r)|}\int_{\partial B(a,r)} u(y')\,dy'. \tag{6.2.4}$$

To see this, it suffices to show that

$$\frac{1}{|\partial B(a,t)|}\int_{\partial B(a,t)} u(y')\,dy' = \frac{1}{|\partial B(a,r)|}\int_{\partial B(a,r)} u(y')\,dy' \tag{6.2.5}$$

when $0 < t < r$. Indeed, one can check that

$$\lim_{t\to 0+}\frac{1}{|\partial B(a,t)|}\int_{\partial B(a,t)} u(y')\,dy' = u(a), \tag{6.2.6}$$

because u is continuous at a. This permits one to obtain (6.2.4) from (6.2.5).

Note that

$$\int_{\partial B(a,r)} (D_{\nu(y')}u)(y')\,dy' = 0, \tag{6.2.7}$$

where $\nu(y')$ is the outward-pointing unit normal to $\partial B(a,r)$ at a point y' in $\partial B(a,r)$. This follows from (3.5.4), with $V = B(a,r)$. Similarly,

$$\int_{\partial B(a,t)} (D_{\nu(y')}u)(y')\,dy' = 0, \tag{6.2.8}$$

where $\nu(y')$ is the outward-pointing unit normal to $\partial B(a,t)$ at a point y' in $\partial B(a,t)$.

To get (6.2.5), consider

$$V = B(a,r)\setminus \overline{B}(a,t), \tag{6.2.9}$$

which is a nonempty bounded open subset of $\mathbf{R}^n$. Observe that

$$\partial V = (\partial B(a,r)) \cup (\partial B(a,t)). \tag{6.2.10}$$

The outward-pointing unit normal to ∂V is the same as the outward-pointing unit normal to $\partial B(a,r)$ at points in $\partial B(a,r)$, and it is -1 times the outward-pointing unit normal to $\partial B(a,t)$ at points in $\partial B(a,t)$.

Put

$$v(x) = |x-a|^{2-n} \tag{6.2.11}$$

on $\mathbf{R}^n \setminus \{a\}$ when $n \ge 3$, and

$$v(x) = \log|x-a| \tag{6.2.12}$$

on $\mathbf{R}^2 \setminus \{a\}$ when $n = 2$. In both cases, $v(x)$ is harmonic on $\mathbf{R}^n \setminus \{a\}$, as in the previous section.

We would like to use (3.9.1) in this case. The left side of that equation is equal to 0, because u and v are harmonic on V. One can check that the part of the right side of the equation involving the normal derivative of u is equal to 0, because of (6.2.7) and (6.2.8). This also uses the fact that v is constant on $\partial B(a,r)$ and $\partial B(a,t)$.

It follows that the part of the right side of the equation involving the normal derivative of v is equal to 0. One can use this to get (6.2.5), as desired.

Alternatively,

$$\frac{1}{|\partial B(a,t)|}\int_{\partial B(a,t)} u(y')\,dy' = \frac{1}{|\partial B(0,1)|}\int_{\partial B(0,1)} u(a+t\,z')\,dz' \tag{6.2.13}$$

when $0 < t \le r$. The derivative of the right side in t is equal to

$$\begin{aligned} &\frac{1}{|\partial B(0,1)|}\int_{\partial B(0,1)} \sum_{j=1}^{n} \frac{\partial u}{\partial x_j}(a+t\,z')\,z_j'\,dz' \\ &= \frac{1}{|\partial B(a,t)|}\int_{\partial B(a,t)} \sum_{j=1}^{n} \frac{\partial u}{\partial x_j}(y')\,t^{-1}\,(y_j'-a_j)\,dy' \\ &= \frac{1}{|\partial B(a,t)|}\int_{\partial B(a,t)} (D_{\nu(y')}u)(y')\,dy', \end{aligned} \tag{6.2.14}$$

where $\nu(y')$ is the outward-pointing unit normal to $\partial B(a,t)$ at $y' \in \partial B(a,t)$ again. More precisely, one can verify that differentiation under the integral sign is permitted here, using the continuous differentiability of u. If u is harmonic on U, then the right side of (6.2.14) is equal to 0, as in (6.2.8). This implies that the right side of (6.2.13) is constant for $0 < t \le r$, so that (6.2.5) holds.

6.3 More on mean values

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a continuous real or complex-valued function on U. Let us say that u has the *mean-value property* on U if for every $a \in U$ and $r > 0$ such that $\overline{B}(a,r) \subseteq U$, we have that (6.2.4) holds. Equivalently, this means that

$$\int_{\partial B(a,r)} u(y')\,dy' = |\partial B(a,r)|\,u(a) = r^{n-1}\,|\partial B(0,1)|\,u(a). \tag{6.3.1}$$

In this case, we get that

$$\int_{B(a,r)} u(x)\,dx = |B(a,r)|\,u(a) = r^n\,|B(0,1)|\,u(a), \tag{6.3.2}$$

by integrating in r. Of course, this is the same as saying that

$$u(a) = \frac{1}{|B(a,r)|}\int_{B(a,r)} u(x)\,dx. \tag{6.3.3}$$

Conversely, one can get (6.3.1) from (6.3.2), by differentiating in r.

One can check that

$$\int_{\partial B(a,r)} (y_j'-a_j)\,dy' = \int_{B(a,t)} (x_j-a_j)\,dx = 0 \tag{6.3.4}$$

for every $a \in \mathbf{R}^n$, $r > 0$, and $j = 1, \ldots, n$. Similarly,

$$\int_{\partial B(a,t)} (y'_j - a_j)\,(y'_l - a_l)\,dy' = \int_{B(a,t)} (x_j - a_j)\,(x_l - a_l)\,dx = 0 \tag{6.3.5}$$

when $j \neq l$. We also have that

$$\int_{\partial B(a,r)} (y'_j - a_j)^2\,dy' = \int_{\partial B(a,r)} (y'_l - a_l)^2\,dy' \tag{6.3.6}$$

and

$$\int_{B(a,r)} (x_j - a_j)^2\,dx = \int_{B(a,r)} (x_l - a_l)^2\,dx \tag{6.3.7}$$

for every $j, l = 1, \ldots, n$. One can use these remarks to show directly that a polynomial on $\mathbf{R}^n$ of degree less than or equal to 2 satisfies the mean value property if and only if it is harmonic.

If u is twice continuously differentiable on U, and u has the mean-value property on U, then u is harmonic on U. This can be seen using the Taylor approximation to u at a point $a \in U$ of degree 2, to estimate the difference between the average of u on balls or spheres centered at a with small radius and $u(a)$.

Alternatively, the mean-value property implies that the right side of (6.2.14) is 0 when $\overline{B}(a,t) \subseteq U$. This means that

$$\int_{B(a,t)} (\Delta u)(x)\,dx = 0, \tag{6.3.8}$$

because of (3.5.3). One can use this to get that $(\Delta u)(a) = 0$.

6.4 Mean values and smoothness

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let u be a continuous real or complex-valued function on U with the mean-value property. Let $r > 0$ be given, and let ϕ be a continuous real-valued function on $\mathbf{R}^n$ supported in $\overline{B}(0,r)$. Suppose too that ϕ is a radial function on $\mathbf{R}^n$, so that $\phi(x)$ depends only on $|x|$.

Let $a \in U$ be given, and suppose that $\overline{B}(a,r) \subseteq U$. If $0 < t \leq r$, then

$$\int_{\partial B(a,t)} u(y')\,\phi(y'-a)\,dy' = \Big(\int_{\partial B(a,t)} \phi(y'-a)\,dy'\Big)\,u(a). \tag{6.4.1}$$

This uses the mean-value property of u, and the fact that $\phi(y'-a)$ is constant as a function of y' on $\partial B(a,t)$, because ϕ is radial on $\mathbf{R}^n$. It follows that

$$\int_{\partial B(a,t)} u(y')\,\phi(y'-a)\,dy' = \Big(\int_{\partial B(0,t)} \phi(z')\,dz'\Big)\,u(a). \tag{6.4.2}$$

We can integrate over t to get that

$$\int_{B(a,t)} u(x)\,\phi(x-a)\,dx = \Big(\int_{B(0,r)} \phi(w)\,dw\Big)\,u(a). \tag{6.4.3}$$

If

$$\int_{B(0,r)} \phi(w)\,dw = 1, \tag{6.4.4}$$

then we get that

$$\int_{B(a,r)} u(x)\,\phi(x-a)\,dx = u(a). \tag{6.4.5}$$

Of course, we can get (6.4.4) by dividing ϕ by its integral over $B(0,r)$, as long as the integral is not zero. It is easy to see that the integral is positive when ϕ is nonnegative and not equal to 0 at every point in $B(0,r)$.

Remember that $\overline{B}(a,r) \subseteq U$ implies that

$$\overline{B}(a, r+\epsilon) \subseteq U \tag{6.4.6}$$

for some $\epsilon > 0$, as in Section 1.13. If $b \in \mathbf{R}^n$ and $|a-b| \le \epsilon$, then it follows that

$$\overline{B}(b,r) \subseteq \overline{B}(a, r+\epsilon) \subseteq U, \tag{6.4.7}$$

using the triangle inequality in the first step. This means that

$$u(b) = \int_{B(b,r)} u(x)\,\phi(x-b)\,dx, \tag{6.4.8}$$

as before. This can also be expressed as

$$u(b) = \int_{B(a,r+\epsilon)} u(x)\,\phi(x-b)\,dx, \tag{6.4.9}$$

because ϕ is supported in $\overline{B}(0,r)$.

Suppose that ϕ is a smooth function on $\mathbf{R}^n$ too, which can be arranged by taking a suitable smooth function of $|x|^2$. Under these conditions, one can differentiate under the integral sign in (6.4.9), to get that u is smooth near a.

One can use this type of argument at every point in U, to get that u is smooth on U. It follows that u is harmonic on U, as in the previous section.

If u is twice continuously differentiable and harmonic on U, then u has the mean value property, as in Section 6.2. This implies that u is smooth on U, as in the preceding paragraph.

6.5 Uniform convergence

Let E be a nonempty set, let $\{f_j\}_{j=1}^\infty$ be a sequence of real or complex-valued functions on E, and let f be another real or complex-valued function on E. We say that $\{f_j\}_{j=1}^\infty$ converges to f *pointwise* on E if for every $x \in E$, $\{f_j(x)\}_{j=1}^\infty$

converges to $f(x)$ in the usual sense, as a sequence of real or complex numbers. We say that $\{f_j\}_{j=1}^{\infty}$ converges *uniformly* to f on E if for every $\epsilon > 0$ there is a positive integer L such that

$$|f_j(x) - f(x)| < \epsilon \tag{6.5.1}$$

for every $x \in E$ and $j \geq L$. Uniform convergence on E clearly implies pointwise convergence on E.

Let n be a positive integer, and suppose now that E is a nonempty subset of $\mathbf{R}^n$. If $\{f_j\}_{j=1}^{\infty}$ is a sequence of continuous real or complex-valued functions on E that converges uniformly to a real or complex-valued function f on E, as appropriate, then it is well known that f is continuous on E too.

Let U be a nonempty open subset of $\mathbf{R}^n$, let $\{f_j\}_{j=1}^{\infty}$ be a sequence of real or complex-valued functions on U, and let f be a real or complex-valued function on U. We say that $\{f_j\}_{j=1}^{\infty}$ converges to f *uniformly on compact subsets of* U if for every compact subset E of $\mathbf{R}^n$ such that $E \subseteq U$, $\{f_j\}_{j=1}^{\infty}$ converges to f uniformly on E. Uniform convergence on U implies uniform convergence on compact subsets of U, and uniform convergence on compact subsets of U implies pointwise convergence on U.

If $\{f_j\}_{j=1}^{\infty}$ is a sequence of continuous real or complex-valued functions on U that converges to f uniformly on compact subsets of U, then f is continuous on U as well.

Let $\{u_j\}_{j=1}^{\infty}$ be a sequence of harmonic functions on U that converges to a function u on U, uniformly on compact sets contained in U. This implies that u is continuous on U, as before. One can use the mean-value property for u_j for each j to get that u has the mean-value property on U too, because of standard results about uniform convergence and integration. This means that u is harmonic on U, as in the previous section. One can also show that derivatives of the u_j's converge to the corresponding derivatives of u, uniformly on compact subsets of U, by expressing the derivatives in terms of integrals of the functions, as in the previous section.

6.6 Liouville's theorem

Let n be a positive integer, and let u be a bounded harmonic function on $\mathbf{R}^n$. Under these conditions, *Liouville's theorem* states that u is a constant function on $\mathbf{R}^n$. To see this, let $x, y \in \mathbf{R}^n$ and $r > 0$ be given, so that

$$\begin{aligned} u(x) - u(y) &= \frac{1}{|B(x,r)|} \int_{B(x,r)} u(w)\,dw - \frac{1}{|B(y,r)|} \int_{B(y,r)} u(w) \\ &= \frac{1}{|B(0,1)|\,r^n} \int_{B(x,r)\setminus \overline{B}(y,r)} u(w)\,dw \\ &\quad - \frac{1}{|B(0,1)|\,r^n} \int_{B(y,r)\setminus \overline{B}(x,r)} u(w)\,dw \end{aligned} \tag{6.6.1}$$

If $r > |x - y|$, then one can check that

$$B(x,r) \setminus \overline{B}(y,r) \subseteq B(x,r) \setminus \overline{B}(x, r - |x-y|), \tag{6.6.2}$$

and similarly with the roles of x and y interchanged. The n-dimensional volume of the right side is equal to

$$\begin{aligned} |B(x,r)| - |B(x, r - |x-y|)| &= |B(0,1)|\,(r^n - (r - |x-y|)^n) \\ &= |B(0,1)| \sum_{j=0}^{n-1} (-1)^{n-j+1}\, r^j\, |x-y|^{n-j}, \end{aligned} \tag{6.6.3}$$

and similarly with the roles of x and y interchanged.

If u is bounded on $\mathbf{R}^n$, then one can use this to check that right side of (6.6.1) tends to 0 as $r \to \infty$. This implies that $u(x) = u(y)$, as desired.

Alternatively, we can use arguments like those in Section 6.4 to estimate first derivatives of harmonic functions. These estimates will show that bounded harmonic functions on $\mathbf{R}^n$ have all of their first derivatives equal to 0.

Let ϕ be a smooth real-valued radial function on $\mathbf{R}^n$ supported on the closed unit ball $\overline{B}(0,1)$, and with

$$\int_{B(0,1)} \phi(w)\, dw = 1. \tag{6.6.4}$$

Put

$$\phi_r(w) = r^{-n}\, \phi(r^{-1}\, w) \tag{6.6.5}$$

for every $w \in \mathbf{R}^n$ and $r > 0$. It is easy to see that ϕ_r is a smooth real-valued radial function on $\mathbf{R}^n$ that is supported on $\overline{B}(0,r)$ and satisfies

$$\int_{B(0,r)} \phi_r(w)\, dw = 1. \tag{6.6.6}$$

Let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a real or complex-valued harmonic function on U. If $a \in U$, $r > 0$, and $\overline{B}(a,r) \subseteq U$, then

$$u(a) = \int_{B(a,r)} u(x)\, \phi_r(x - a), \tag{6.6.7}$$

as in (6.4.5). If $\epsilon > 0$ is as in (6.4.6), $b \in \mathbf{R}^n$, and $|a - b| \le \epsilon$, then we get that

$$u(b) = \int_{B(a,r+\epsilon)} u(x)\, \phi_r(x - b)\, dx, \tag{6.6.8}$$

as in (6.4.9).

Observe that

$$\frac{\partial}{\partial w_j}(\phi_r(w)) = r^{-n-1} (\partial_j \phi)(r^{-1} w) \tag{6.6.9}$$

for each $j = 1, \ldots, n$. We can differentiate under the integral sign in (6.6.8) to get that

$$(\partial_j u)(a) = -r^{-n-1} \int_{B(a,r)} u(x)\,(\partial_j \phi)(r^{-1}\,(x-a))\,dx \tag{6.6.10}$$

for each $j = 1, \ldots, n$. This also uses the fact that $\partial_j \phi$ is supported in $\overline{B}(0,1)$.

If $U = \mathbf{R}^n$ and u is bounded on $\mathbf{R}^n$, then one can check that the right side of (6.6.10) tends to 0 as $r \to \infty$. This implies that $\partial_j u = 0$ on $\mathbf{R}^n$ for each $j = 1, \ldots, n$, so that u is constant on $\mathbf{R}^n$.

6.7 The maximum principle

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a continuous real-valued function on U. Suppose that for every $a \in U$ there is an $r > 0$ such that $\overline{B}(a,r) \subseteq U$ and the average of u on $B(a,r)$ is equal to $u(a)$, as in (6.3.3). In particular, this happens when u is harmonic on U, as in Section 6.2.

Let A be a real number such that

$$u(x) \le A \tag{6.7.1}$$

for every $x \in U$. Note that

$$\{x \in U : u(x) = A\} \tag{6.7.2}$$

is a relatively closed set in U, because u is continuous on U. Equivalently,

$$\{x \in U : u(x) < A\} \tag{6.7.3}$$

is an open set.

Suppose that

$$u(a) = A \tag{6.7.4}$$

for some $a \in U$. If $\overline{B}(a,r) \subseteq U$ and (6.3.3) holds, then one can verify that

$$u(x) = A \tag{6.7.5}$$

for every $x \in B(a,r)$.

This shows that (6.7.2) is an open set under these conditions. If (6.7.2) is nonempty, and U is connected, then it follows that (6.7.2) is equal to U. This is often called the *strong maximum principle*.

Suppose now that U is also bounded, and let u be a continuous real-valued function on $\overline{U}$. As before, we ask that for each $a \in U$ there be an $r > 0$ such that $\overline{B}(a,r) \subseteq U$ and (6.3.3) holds. Note that $\overline{U}$ is a nonempty compact subset of $\mathbf{R}^n$, so that u attains its maximum on $\overline{U}$, by the extreme value theorem.

Suppose that u attains its maximum on $\overline{U}$ at a point $a \in U$. If V is the connected component of U that contains a, then it follows that u is constant on V, as before. More precisely, u is constant on $\overline{V}$, by continuity. This implies that u attains its maximum on $\overline{U}$ at a point in ∂V, which is contained in ∂U, as in Section 3.3.

Otherwise, if u does not attain its maximum on $\overline{U}$ at a point in U, then u attains its maximum on $\overline{U}$ at a point in ∂U. This means that u attains its maximum on $\overline{U}$ at a point in ∂U in either case, which is another version of the *maximum principle*.

In particular, if

$$u(x) = 0 \text{ for every } x \in \partial U, \tag{6.7.6}$$

then we get that $u(x) \leq 0$ for every $x \in \overline{U}$. The same argument can be used for $-u$, to obtain that

$$u(x) = 0 \text{ for every } x \in \overline{U}. \tag{6.7.7}$$

6.8 A helpful integral formula

Let $n \geq 2$ be an integer, and let $N(x)$ be the real-valued function defined on $\mathbf{R}^n \setminus \{0\}$ by

$$\begin{aligned} N(x) &= \frac{|x|^{2-n}}{(2-n)\,|\partial B(0,1)|} \quad \text{when } n \geq 3 \\ &= \frac{1}{2\pi} \log |x| \qquad\qquad \text{when } n = 2. \end{aligned} \tag{6.8.1}$$

Thus $N(x)$ is harmonic on $\mathbf{R}^n \setminus \{0\}$, as in Section 6.1.

Let V be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary, and let u be a twice continuously-differentiable real or complex-valued function on $\overline{V}$, as in Section 3.4. Also let $a \in V$ be given, and suppose that

$$\overline{B}(a,r) \subseteq V \tag{6.8.2}$$

for some $r > 0$. Put

$$V_r = V \setminus \overline{B}(a,r), \tag{6.8.3}$$

which is an open subset of $\mathbf{R}^n$. Note that

$$\overline{V_r} = \overline{V} \setminus B(a,r) \tag{6.8.4}$$

and that

$$\partial V_r = (\partial V) \cup (\partial B(a,r)). \tag{6.8.5}$$

We would like to use (3.9.1), with V replaced with V_r, and $v(x) = N(x-a)$. This implies that

$$\begin{aligned} &-\int_{V_r} N(x-a)\,(\Delta u)(x)\,dx \\ &= \int_{\partial V_r} (u(y')\,(D_{\nu_r(y')}N)(y'-a) - N(y'-a)\,(D_{\nu_r(y')}u)(y'))\,dy', \end{aligned} \tag{6.8.6}$$

where $D_{\nu_r(y')}$ denoted the directional derivative in the direction $\nu_r(y')$ of the outward-pointing unit normal to ∂V_r at $y' \in \partial V_r$. It follows that

$$\begin{aligned}
&-\int_{V_r} N(x-a)\,(\Delta u)(x)\,dx \\
&\quad= \int_{\partial V} (u(y')\,(D_{\nu(y')}N)(y'-a) - N(y'-a)\,(D_{\nu(y')}u)(y'))\,dy' \\
&\qquad - \int_{\partial B(a,r)} (u(y')\,(D_{\mu(y')}N)(y'-a) - N(y'-a)\,(D_{\mu(y')}u)(y'))\,dy', \qquad (6.8.7)
\end{aligned}$$

where $\nu(y')$ is the outward-pointing unit normal to ∂V at $y' \in \partial V$, and $\mu(y')$ is the outward-pointing unit normal to $\partial B(a,r)$ at $y' \in \partial B(a,r)$. Of course,

$$\begin{aligned}
\nu_r(y') &= \nu(y') && \text{when } y' \in \partial V \\
&= -\mu(y') && \text{when } y' \in \partial B(a,r).
\end{aligned} \qquad (6.8.8)$$

One can check that

$$\int_{\partial B(a,r)} u(y')\,(D_{\mu(y')}N)(y'-a)\,dy' = \frac{1}{|\partial B(a,r)|}\int_{\partial B(a,r)} u(y')\,dy', \qquad (6.8.9)$$

by expressing the partial derivatives of N as in Section 6.1. This tends to $u(a)$ as $r \to 0+$, because u is continuous at a.

One can also verify that

$$\lim_{r\to 0+} \int_{\partial B(a,r)} N(y'-a)\,(D_{\mu(y')}u)(y')\,dy' = 0. \qquad (6.8.10)$$

This uses the continuous differentiability of u to get that the first derivatives of u are bounded near a.

This implies that

$$\begin{aligned}
&\lim_{r\to 0+} \int_{V_r} N(x-a)\,(\Delta u)(x)\,dx \qquad (6.8.11)\\
&\quad= \int_{\partial V} (N(y'-a)\,(D_{\nu(y')}u)(y') - u(y')\,(D_{\nu(y')}N)(y'-a))\,dy' + u(a).
\end{aligned}$$

The left side may be considered as

$$\int_V N(x-a)\,(\Delta u)(x)\,dx, \qquad (6.8.12)$$

defined as an improper integral, because $N(x-a)$ is unbounded near a. One can check that

$$\int_{V_r} |N(x-a)|\,|(\Delta u)(x)|\,dx \qquad (6.8.13)$$

stays bounded as $r \to 0+$, using polar coordinates near a, and the fact that $|(\Delta u)(x)|$ is bounded near a, because Δu is continuous, by hypothesis. In particular, (6.8.12) can be defined as a Lebesgue integral.

If u is harmonic on V, then we get that

$$u(a) = \int_{\partial V} (u(y')\,(D_{\nu(y')}N)(y'-a) - N(y'-a)\,(D_{\nu(y')}u)(y'))\,dy'. \qquad (6.8.14)$$

6.9 Poisson's equation on $\mathbf{R}^n$

Let $n \geq 2$ be an integer, and let $N(x)$ be as in (6.8.1). If f is a real or complex-valued function on $\mathbf{R}^n$, then one might like to define u as a real or complex-valued function on $\mathbf{R}^n$, as appropriate, by

$$u(x) = \int_{\mathbf{R}^n} N(x-y)\, f(y)\, dy = \int_{\mathbf{R}^n} N(y)\, f(x-y)\, dy. \tag{6.9.1}$$

If f is a continuous function on $\mathbf{R}^n$ with compact support, then the integral on the right may be considered as an improper integral over a bounded region for each $x \in \mathbf{R}^n$. One can use the Lebesgue integral to define u as a locally integrable function on $\mathbf{R}^n$ under suitable integrability conditions on f.

One would like to have

$$\Delta u = f \tag{6.9.2}$$

on $\mathbf{R}^n$, under suitable conditions, or interpreted in a suitable way. Suppose that f is twice continuously differentiable on $\mathbf{R}^n$, with compact support. In this case, one can show that u is twice continuously differentiable on $\mathbf{R}^n$, by differentiating the second integral in (6.9.1) under the integral sign. One can also use this to get (6.9.2), by taking V large enough in the previous section so that the support of f is contained in V. This corresponds to some remarks on p193 of [10], and to Theorem 1 in Section 2.2.1 b in [29].

Let v be a twice continuously-differentiable real or complex-valued function on $\mathbf{R}^n$ with compact support. Thus

$$\int_{\mathbf{R}^n} N(x-y)\, (\Delta v)(x)\, dx = v(y) \tag{6.9.3}$$

for every $y \in \mathbf{R}^n$, as in the previous section, with V taken large enough to contain the support of v. Under suitable integrability conditions on f, we have that

$$\begin{aligned} \int_{\mathbf{R}^n} u(x)\, (\Delta v)(x)\, dx &= \int_{\mathbf{R}^n} \Big(\int_{\mathbf{R}^n} N(x-y)\, f(y)\, dy\Big)\, (\Delta v)(x)\, dx \\ &= \int_{\mathbf{R}^n} \Big(\int_{\mathbf{R}^n} N(x-y)\, (\Delta v)(x)\, dx\Big)\, f(y)\, dy \\ &= \int_{\mathbf{R}^n} f(y)\, v(y)\, dy. \end{aligned} \tag{6.9.4}$$

This means that u satisfies (6.9.2) in the sense of distributions, as in Theorem 2.16 in Section B of Chapter 2 of [32].

If f is continuous on $\mathbf{R}^n$, and u is twice continuously differentiable on $\mathbf{R}^n$, then one can use (6.9.4) to get that (6.9.2) holds on $\mathbf{R}^n$. If f has a bit more regularity, then one can get that u is twice continuously differentiable under suitable conditions, as in Theorem 2.17 in Section B of Chapter 2 of [32].

Some topics related to integrals like those in (6.9.1) are discussed in Chapter 5 of [119].

6.10 The Poisson kernel

Let $n \ge 2$ be an integer, and put

$$p(w', x) = \frac{1}{|\partial B(0,1)|} \, \frac{(1 - |x|^2)}{|x - w'|^n} \tag{6.10.1}$$

for every $w', x \in \mathbf{R}^n$ with $|w'| = 1$ and $x \ne w'$. This is the *Poisson kernel* associated to the unit ball in $\mathbf{R}^n$.

Let $w' \in \mathbf{R}^n$ with $|w'| = 1$ be given, and let us check that $p(w', x)$ is harmonic as a function of x for $x \ne w'$. Observe that

$$\begin{aligned} |x|^2 &= |(x - w') + w'|^2 = |x - w'|^2 + 2\,(x - w') \cdot w' + |w'|^2 \\ &= |x - w'|^2 + 2\,(x - w') \cdot w' + 1. \end{aligned} \tag{6.10.2}$$

Thus

$$p(w', x) = \frac{1}{|\partial B(0,1)|} \Big(\frac{-1}{|x - w'|^{n-2}} - 2\,\frac{(x - w') \cdot w'}{|x - w'|^n} \Big). \tag{6.10.3}$$

The first term on the right is harmonic in x for $x \ne w'$, as mentioned in Section 6.1 when $n \ge 3$, and trivially when $n = 2$. The second term on the right can be expressed in terms of derivatives of harmonic functions in x for $x \ne w'$, as before, which are harmonic too.

Suppose that $w', x' \in \mathbf{R}^n$, $|w'| = |x'| = 1$, $r \in \mathbf{R}$, and $0 \le r < 1$. It is easy to see that

$$|r\,x' - w'| = |x' - r\,w'|, \tag{6.10.4}$$

by squaring both sides and expanding using the dot product on $\mathbf{R}^n$. This means that

$$p(w', r\,x') = p(x', r\,w'). \tag{6.10.5}$$

The mean-value property for harmonic functions implies that

$$\frac{1}{|\partial B(0,1)|} \int_{\partial B(0,1)} p(x', r\,w')\,dw' = p(x', 0) = \frac{1}{|\partial B(0,1)|}. \tag{6.10.6}$$

This implies that

$$\int_{\partial B(0,1)} p(w', r\,x')\,dw' = 1, \tag{6.10.7}$$

because of (6.10.5).

Note that

$$p(w', x) > 0 \tag{6.10.8}$$

for every $w', x \in \mathbf{R}^n$ with $|w'| = 1$ and $|x| < 1$. If $x' \in \mathbf{R}^n$, $|x'| = 1$, and η is a positive real number, then one can check that

$$\int_{(\partial B(0,1)) \setminus B(x', \eta)} p(w', r\,x')\,dw' \to 0 \tag{6.10.9}$$

as $r \to 1-$, uniformly in x'. More precisely,

$$1 = |w'| \le |w' - r\,x'| + |r\,x'| = |w' - r\,x'| + r \tag{6.10.10}$$

for every $r \ge 0$, so that

$$1 - r \le |w' - r\,x'|. \tag{6.10.11}$$

This implies that

$$|w' - x'| \le |w' - r\,x'| + |r\,x' - x'| = |w - r\,x'| + 1 - r \le 2\,|w' - r\,x'| \tag{6.10.12}$$

when $0 \le r \le 1$.

6.11 The Poisson integral

Let us continue with the same notation and hypotheses as in the previous section. Let f be a continuous complex-valued function on the unit sphere $\partial B(0,1)$. Consider the complex-valued function u defined on the closed unit ball $\overline{B}(0,1)$ by

$$\begin{aligned} u(x) &= \int_{\partial B(0,1)} f(w')\,p(w',x)\,dw' \quad \text{when } |x| < 1 \\ &= f(x) \qquad\qquad\qquad\qquad\qquad \text{when } |x| = 1. \end{aligned} \tag{6.11.1}$$

It is not too difficult to show that u is harmonic on $B(0,1)$, because $p(w',x)$ is harmonic in x on $B(0,1)$ for every $w' \in \partial B(0,1)$, as in the previous section. One way to do this is to use standard results about differentiation under the integral sign. Another way to do this is to check that u is continuous and satisfies the mean-value property on $B(0,1)$. This uses the mean-value property for $p(w',x)$ in x for each w', and well known results about interchanging the order of integration.

One can also show that u is continuous on $\overline{B}(0,1)$. The continuity of u on $B(0,1)$ is reasonably straightforward, as in the preceding paragraph. If $y' \in \partial B(0,1)$, then one would like to verify that u is continuous at y', as a function on $\overline{B}(0,1)$. Equivalently, this means that

$$u(x) \to u(y') = f(y') \tag{6.11.2}$$

as $x \in \overline{B}(0,1)$ tends to y'. More precisely, it suffices to consider only $x \in B(0,1)$ here, because f is continuous on $\partial B(0,1)$, by hypothesis.

Note that

$$\int_{\partial B(0,1)} p(w',x)\,dw' = 1 \tag{6.11.3}$$

for every $x \in B(0,1)$, by (6.10.7). This implies that

$$u(x) - f(y') = \int_{\partial B(0,1)} p(w',x)\,(f(w') - f(y'))\,dw' \tag{6.11.4}$$

for every $x \in B(0,1)$. It follows that

$$|u(x) - f(y')| \le \int_{\partial B(0,1)} p(w',x)\,|f(w') - f(y')|\,dw' \tag{6.11.5}$$

for every $x \in B(0,1)$, because of (6.10.8).

We would like to get that the right side of (6.11.5) is as small as we like when x is sufficiently close to y'. If $\eta > 0$, then the right side of (6.11.5) can be expressed as the sum of

$$\int_{(\partial B(0,1))\cap B(y',\eta)} p(w',x)\,|f(w') - f(y')|\,dw' \tag{6.11.6}$$

and

$$\int_{(\partial B(0,1))\setminus B(y',\eta)} p(w',x)\,|f(w') - f(y')|\,dw'. \tag{6.11.7}$$

If η is sufficiently small, then

$$|f(w') - f(y')| \tag{6.11.8}$$

is as small as we like when $|w' - y'| < \eta$, because f is continuous at y', by hypothesis. We can use this to get that (6.11.6) is as small as we like, because of (6.11.3). Let us now fix $\eta > 0$ in this way.

With η fixed, we can get that (6.11.7) is as small as we like when x is sufficiently close to y'. Note that f is bounded on $\partial B(0,1)$, because f is continuous on $\partial B(0,1)$, and $\partial B(0,1)$ is compact. If x is sufficiently close to y', then $|x|$ is as close as we like to 1, and $x/|x|$ is as close as we like to y'. We can use this to get that (6.11.7) is as small as we like, as in (6.10.9).

If v is any continuous complex-valued function on $\overline{B}(0,1)$ that is harmonic on $B(0,1)$ and equal to f on $\partial B(0,1)$, then $v = u$ on $\overline{B}(0,1)$, as in Section 6.7.

6.12 Some more integral formulas

Let n be a positive integer, and let $N(x)$ be the real-valued function defined on $\mathbf{R}^n \setminus \{0\}$ as in Section 6.8. Put

$$\begin{aligned} c_r &= \frac{r^{2-n}}{(2-n)\,|\partial B(0,1)|} && \text{when } n \ge 3 \\ &= \frac{1}{2\pi}\log r && \text{when } n = 2 \end{aligned} \tag{6.12.1}$$

for each $r > 0$, so that $N(x) = c_r$ when $|x| = r$. Let $a \in \mathbf{R}^n$ and $r > 0$ be given, and suppose that u is a twice continuously-differentiable real or complex-valued function on $\overline{B}(a,r)$, as in Section 3.4.

Let $0 < t < r$ be given, and put

$$V = B(a,r) \setminus \overline{B}(a,t). \tag{6.12.2}$$

If $y' \in \partial V = (\partial B)(a,r)) \cup (\partial B)(a,t))$, then let $\nu(y')$ be the outward pointing unit normal to ∂V at y', as usual. We would like to use (3.9.1), with

$$v(x) = N(x-a) - c_r. \tag{6.12.3}$$

This implies that

$$\begin{aligned} -\int_V v(x)\,(\Delta u)(x)\,dx & \\ = \int_{\partial V} (u(y')\,(D_{\nu(y')} v)(y') - v(y')\,(D_{\nu(y')} u)(y'))\,dy'. \end{aligned} \tag{6.12.4}$$

If $\rho > 0$ and $y' \in \partial B(a,\rho)$, then put

$$\mu_\rho(y') = \rho^{-1}\,(y'-a), \tag{6.12.5}$$

which is the outward-pointing unit normal to $\partial B(a,\rho)$ at y'. Thus

$$\begin{aligned} \nu(y') &= \mu_r(y') \quad \text{when } y' \in \partial B(a,r) \\ &= -\mu_t(y') \quad \text{when } y' \in \partial B(a,t). \end{aligned} \tag{6.12.6}$$

Using this and (6.12.4), we get that

$$\begin{aligned} &-\int_V v(x)\,(\Delta u)(x)\,dx \\ &= \int_{\partial B(a,r)} u(y')\,(D_{\mu_r(y')} v)(y') dy' \\ &\quad - \int_{\partial B(a,t)} (u(y')\,(D_{\mu_t(y')} v)(y') - v(y')\,(D_{\mu_t(y')} u)(y'))\,dy', \end{aligned} \tag{6.12.7}$$

because $v = 0$ on $\partial B(a,r)$, by construction.

It follows from this and (6.12.3) that

$$\begin{aligned} &\int_V (c_r - N(x-a))\,(\Delta u)(x)\,dx \\ &= \frac{1}{|\partial B(a,r)|} \int_{\partial B(a,r)} u(y')\,dy' - \frac{1}{|\partial B(a,t)|} \int_{\partial B(a,t)} u(y')\,dy' \\ &\quad + (c_t - c_r) \int_{\partial B(a,t)} (D_{\mu_t(y')} u)(y')\,dy', \end{aligned} \tag{6.12.8}$$

using also (6.8.9) and its analogue for $\partial B(a,t)$. Remember that

$$\int_{\partial B(a,t)} (D_{\mu_t(y')} u)(y')\,dy' = \int_{B(a,t)} (\Delta u)(x)\,dx, \tag{6.12.9}$$

as in Section 3.5. Using this, we can reexpress (6.12.8) as

$$\begin{aligned} &\int_{B(a,r)} \min(c_r - N(x-a), c_r - c_t)\,(\Delta u)(x)\,dx \\ &= \frac{1}{|\partial B(a,r)|} \int_{\partial B(a,r)} u(y')\,dy' - \frac{1}{|\partial B(a,t)|} \int_{\partial B(a,t)} u(y')\,dy'. \end{aligned} \tag{6.12.10}$$

We can take the limit as $t \to 0+$ on both sides of (6.12.8) or (6.12.10) to get that

$$\int_{B(a,r)} (c_r - N(x-a))\,(\Delta u)(x)\,dx = \frac{1}{|\partial B(a,r)|} \int_{\partial B(a,r)} u(y')\,dy' - u(a), \tag{6.12.11}$$

as in Section 6.8. More precisely, the left side of the equation should be considered as an improper integral, or a Lebesgue integral, as before.

Alternatively, if $0 < \rho \le r$, then

$$\frac{d}{d\rho}\Big(\frac{1}{|\partial B(a,\rho)|} \int_{\partial B(a,\rho)} u(y')\,dy'\Big) = \frac{1}{|\partial B(a,\rho)|} \int_{\partial B(a,\rho)} (D_{\mu_\rho(y')} u)(y')\,dy', \tag{6.12.12}$$

as in Section 6.2. This means that

$$\frac{d}{d\rho}\Big(\frac{1}{|\partial B(a,\rho)|} \int_{\partial B(a,\rho)} u(y')\,dy'\Big) = \frac{1}{|\partial B(a,\rho)|} \int_{B(a,\rho)} (\Delta u)(x)\,dx, \tag{6.12.13}$$

by (6.12.9). One can get (6.12.10) by integrating both sides of (6.12.13) in ρ from t to r. This also involves interchanging the order of integration on the right side. In some cases, we may be particularly interested simply in the nonnegativity of some of these integrals of Δu when $\Delta u \ge 0$, as in the next section.

6.13 Subharmonic functions

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. A twice continuously-differentiable real-valued function u on U is said to be *subharmonic* if

$$\Delta u \ge 0 \tag{6.13.1}$$

on U. Equivalently, u may be considered as a *subsolution* of the Laplace equation in this case. If $n = 1$, then this corresponds to convexity of u.

Let $a \in U$ and $r > 0$ be given, with $\overline{B}(a,r) \subseteq U$. If u is subharmonic on U, then it is well known that

$$u(a) \le \frac{1}{|\partial B(a,r)|} \int_{\partial B(a,r)} u(y')\,dy'. \tag{6.13.2}$$

This can be obtained from suitable integral formulas, as in the previous section. One can use this to get that

$$u(a) \le \frac{1}{|B(a,r)|} \int_{B(a,r)} u(x)\,dx, \tag{6.13.3}$$

as before. Conditions like these may be used to extend the notion of subharmonicity to functions with less regularity.

Suppose that u is a continuous real-valued function on U, and that there is a real number A such that

$$u(x) \le A \tag{6.13.4}$$

for every $x \in U$. Suppose also for the moment that

$$u(a) = A \tag{6.13.5}$$

for some $a \in U$, and that (6.13.3) holds for some $r > 0$ such that $\overline{B}(a,r) \subseteq U$. Under these conditions, one can check that

$$u(x) = A \tag{6.13.6}$$

for every $x \in B(a,r)$. More precisely, $A - u(x) \ge 0$ for every $x \in U$, and

$$\int_{B(a,r)} (A - u(x))\,dx \le 0, \tag{6.13.7}$$

because of (6.13.3) and (6.13.5).

Suppose now that for every $a \in U$ there is an $r > 0$ such that $\overline{B}(a,r) \subseteq U$ and (6.13.3) holds. This implies that the set of $x \in U$ such that (6.13.6) holds is an open set, as in the preceding paragraph. This set is relatively closed in U as well, because u is continuous on U. If this set is nonempty, and U is connected, then this set is equal to U, so that

$$u \equiv A \text{ on } U. \tag{6.13.8}$$

This is another version of the strong maximum principle.

Suppose that U is bounded, and that u is a continuous real-valued function on $\overline{U}$ such that for every $a \in U$ there is an $r > 0$ with $\overline{B}(a,r) \subseteq U$ and for which (6.13.3) holds. The extreme value theorem implies that u attains its maximum on $\overline{U}$. In fact, the maximum of u on $\overline{U}$ is attained as a point in ∂U, as in Section 6.7. More precisely, this uses the remarks in the previous paragraph too. This is another version of the maximum priciple.

6.14 Another approach to local maxima

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a twice continuously-differentiable real-valued function on U. If u has a local maximum at a point $a \in U$, then a is a critical point of U, and the second derivatives of u at a in any direction are less than or equal to 0, by the second-derivative test. In particular, this means that

$$(\Delta u)(a) \le 0. \tag{6.14.1}$$

If

$$(\Delta u)(x) > 0 \tag{6.14.2}$$

for every $x \in U$, then it follows that u has no local maxima in U.

Suppose from now on in this section that U is bounded, and that u is a continuous real-valued function on $\overline{U}$ that is twice continuously differentiable on U. The extreme value theorem implies that u attains its maximum on $\overline{U}$. If (6.14.2) holds at every point in U, then the maximum of u on $\overline{U}$ cannot be attained at a point in U, as in the preceding paragraph. This implies that the maximum of u on $\overline{U}$ is attained at a point in ∂U.

Suppose that u is subharmonic on U, so that $\Delta u \geq 0$ on U. Let $\epsilon > 0$ be given, and put

$$v_\epsilon(x) = u(x) + \epsilon\,|x|^2 \tag{6.14.3}$$

for every $x \in \overline{U}$. Note that v_ϵ is continuous on $\overline{U}$, twice continuously differentiable on U, and that

$$(\Delta v_\epsilon)(x) \geq 2\,n\,\epsilon > 0 \tag{6.14.4}$$

for every $x \in U$. It follows that the maximum of v_ϵ on $\overline{U}$ is attained at a point in ∂U, as in the previous paragraph.

Of course, $\overline{U}$ is bounded in $\mathbf{R}^n$, because U is bounded, so that there is a nonnegative real number R such that

$$|x| \leq R \tag{6.14.5}$$

for every $x \in \overline{U}$. This means that

$$v_\epsilon(x) \leq u(x) + \epsilon\,R^2 \tag{6.14.6}$$

for every $x \in \overline{U}$. It follows that

$$\max_{x \in \overline{U}} v_\epsilon(x) = \max_{x \in \partial U} v_\epsilon(x) \leq \max_{x \in \partial U} u(x) + \epsilon\,R^2, \tag{6.14.7}$$

using the remarks in the previous paragraph in the first step. This implies that

$$\max_{x \in \overline{U}} u(x) \leq \max_{x \in \partial U} u(x) + \epsilon\,R^2, \tag{6.14.8}$$

because $u \leq v_\epsilon$ on $\overline{U}$, by construction. Thus

$$\max_{x \in \overline{U}} u(x) \leq \max_{x \in \partial U} u(x), \tag{6.14.9}$$

because $\epsilon > 0$ is arbitrary.

This is the same as saying that the maximum of u on $\overline{U}$ is attained at a point in ∂U. This is another approach to the maximum principle under these conditions.

6.15 Positive harmonic functions

Let n be a positive integer, and suppose for the moment that u is a positive real-valued harmonic function on $\mathbf{R}^n$. Another version of *Liouville's theorem* states that u has to be constant on $\mathbf{R}^n$. This can be shown in a way that is somewhat analogous to the first proof in Section 6.6, with some adjustments. This is Theorem 3.1 on p45 of [10].

Now let u be a positive harmonic function on a nonempty open subset U of $\mathbf{R}^n$. Suppose that $x, y \in U$ and $r > 0$ satisfy

$$|x - y| \le r \tag{6.15.1}$$

and

$$\overline{B}(x, 2\,r) \subseteq U. \tag{6.15.2}$$

It is easy to see that

$$\overline{B}(y, r) \subseteq \overline{B}(x, 2\,r), \tag{6.15.3}$$

using (6.15.1) and the triangle inequality. It follows that

$$\begin{aligned} u(y) = \frac{1}{|B(y,r)|} \int_{B(y,r)} u(z)\,dz \quad &\le \quad \frac{1}{|B(y,r)|} \int_{B(x,2\,r)} u(z)\,dz \\ &= \quad \frac{2^n}{|B(x,2\,r)|} \int_{B(x,2\,r)} u(z)\,dz = 2^n\,u(x). \end{aligned} \tag{6.15.4}$$

Similarly, if (6.15.1) holds and

$$\overline{B}(y, 2\,r) \subseteq U, \tag{6.15.5}$$

then

$$u(x) \le 2^n\,u(y). \tag{6.15.6}$$

Note that

$$\overline{B}(y, 2\,r) \subseteq \overline{B}(x, 3\,r), \tag{6.15.7}$$

by (6.15.1) and the triangle inequality again. If

$$\overline{B}(x, 3\,r) \subseteq U, \tag{6.15.8}$$

then (6.15.7) implies (6.15.5).

Suppose that U is connected, and that K is compact subset of $\mathbf{R}^n$ that is contained in U. In this case, it is well known that there is a real number $C \ge 1$ such that

$$C^{-1}\,u(x) \le u(y) \le C\,u(x) \tag{6.15.9}$$

for every $x, y \in K$. More precisely, this constant C does not depend on u. This is *Harnack's inequality*, as in Theorem 3.6 on p48 of [10], and Theorem 11 in Section 2.2.3 f of [29].

One can get more precise estimates on balls using the Poisson integral formula, as in 3.4, 3.5 on p47f of [10].

Chapter 7

The heat equation

7.1 Some basic solutions

Let n be a positive integer, and let us identify $\mathbf{R}^n \times \mathbf{R}$ with $\mathbf{R}^{n+1}$, as usual. Let U be a nonempty open subset of $\mathbf{R}^n \times \mathbf{R}$, and let u be a twice continuously-differentiable real or complex-valued function on U. We shall use $\Delta u = \Delta_x u$ to refer to the Laplacian of $u(x,t)$ as a function of x, with t fixed.

We say that $u(x,t)$ satisfies the *heat equation* on U if

$$\frac{\partial u}{\partial t} = \Delta u = \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_j^2} \tag{7.1.1}$$

on U. One may also consider continuously-differentiable functions $u(x,t)$ on U whose second derivatives in x exist and are continuous on U.

Let V be a nonempty open subset of $\mathbf{R}^n$, and let v be a twice continuously-differentiable real or complex-valued function on V. Thus $W = V \times \mathbf{R}$ is an open set in $\mathbf{R}^n \times \mathbf{R}$, and

$$w(x,t) = v(x) \tag{7.1.2}$$

is twice continuously-differentiable on W. Clearly w satisfies the heat equation on W if and only if v is harmonic on V.

Let $a \in \mathbf{C}$ and $b \in \mathbf{C}^n$ be given, and put

$$u(x,t) = \exp(a\,t + b \cdot x) \tag{7.1.3}$$

for every $x \in \mathbf{R}^n$ and $t \in \mathbf{R}$. This satisfies the heat equation on $\mathbf{R}^n \times \mathbf{R}$ if and only if

$$a = b \cdot b. \tag{7.1.4}$$

If $b \in \mathbf{R}^n$, then it follows that $a \geq 0$. If $b = i\,c$ for some $c \in \mathbf{R}^n$, then (7.1.4) implies that

$$a = -c \cdot c \leq 0. \tag{7.1.5}$$

Put

$$K(x,t) = (4\pi\, t)^{-n/2} \exp(-|x|^2/(4\, t)) \tag{7.1.6}$$

for $x \in \mathbf{R}^n$ and $t > 0$. One can check directly that this satisfies the heat equation on $\mathbf{R}^n \times \mathbf{R}_+$. This is known as the *Gauss–Weierstrass* or *heat kernel*, as in Section A of Chapter 4 of [32] and p8 of [125]. This is also discussed in Section 2.3.1 a of [29].

With this normalization, we have that

$$\int_{\mathbf{R}^n} K(x,t)\, dx = 1 \tag{7.1.7}$$

for every $t > 0$. The integral on the left may be considered as an improper integral or as a Lebesgue integral, and this will be discussed in the next two sections.

Put

$$K(x,t) = 0 \tag{7.1.8}$$

when $t = 0$ and $x \neq 0$, and for every $x \in \mathbf{R}^n$ when $t < 0$. This together with (7.1.6) defines $K(x,t)$ on $(\mathbf{R}^n \times \mathbf{R}) \setminus \{(0,0)\}$. One can verify that $K(x,t)$ is smooth on this set, and satisfies the heat equation there, as in Section 2.3.1 b of [29], and Section A of Chapter 4 of [32].

Of course, the heat equation is invariant under translations. In particular, if $y \in \mathbf{R}^n$ and $r \in \mathbf{R}$, then

$$K(x-y, t-r) \tag{7.1.9}$$

is smooth as a function of (x,t) on $(\mathbf{R}^n \times \mathbf{R}) \setminus \{(y,r)\}$, and satisfies the heat equation there.

Note that

$$(-t)^{-n/2} \exp(-|x|^2/(4t)) \tag{7.1.10}$$

satisfies the heat equation on $\mathbf{R}^n \times (-\infty, 0)$, for the same reasons as before. If $y \in \mathbf{R}^n$ and $r \in \mathbf{R}$, then it follows that

$$(r-t)^{-n/2} \exp(|x-y|^2/(4(r-t))) \tag{7.1.11}$$

satisfies the heat equation as a function of (x,t) on $\mathbf{R}^n \times (-\infty, r)$.

7.2 Integrable continuous functions

Let f be a nonnegative real-valued continuous function on the real line. If a, b are real numbers with $a \le b$, then

$$\int_a^b f(x)\, dx \tag{7.2.1}$$

is defined as a Riemann integral, and is a nonnegative real number. Let us say that f is *integrable* on $\mathbf{R}$ if the integrals (7.2.1) are bounded. In this case,

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{\mathbf{R}} f(x)\, dx \tag{7.2.2}$$

may be defined as the supremum of the integrals in (7.2.1) over all $a, b \in \mathbf{R}$ with $a \le b$. This could also be considered as an improper integral, which is obtained by taking a suitable limit of (7.2.1) as $a \to -\infty$ and $b \to \infty$.

One could also use the Lebesgue integral to define (7.2.2) as a nonnegative extended real number for any nonnegative continuous function on $\mathbf{R}$. Integrability of f in the sense considered in the preceding paragraph is the same as the finiteness of (7.2.2) as a Lebesgue integral, which implies that f is Lebesgue integrable on $\mathbf{R}$.

If f is a real-valued continuous function on $\mathbf{R}$, then

$$f_+ = \max(f, 0),\ f_- = \max(-f, 0) \tag{7.2.3}$$

are nonnegative continuous functions on $\mathbf{R}$ such that

$$f = f_+ - f_-,\ |f| = f_+ + f_-. \tag{7.2.4}$$

Let us say that f is *integrable* on $\mathbf{R}$ if $|f|$ is integrable as a nonnegative continuous function on $\mathbf{R}$, which happens if and only if f_+ and f_- are integrable as nonnegative continuous functions on $\mathbf{R}$. This permits us to define the integral (7.2.2) as the difference of the integrals of f_+ and f_- on $\mathbf{R}$. This could also be considered as an improper integral, as before. This is the same as the Lebesgue integral of f on $\mathbf{R}$ as well.

Similarly, a complex-valued continuous function f on $\mathbf{R}$ is said to be *integrable* on $\mathbf{R}$ if $|f|$ is integrable as a nonnegative real-vaued continuous function on $\mathbf{R}$. This happens if and only if the real and imaginary parts of f are integrable as real-valued continuous functions on $\mathbf{R}$, and the real and imaginary parts of the integral (7.2.2) are defined as the integrals of the real and imaginary parts of f on $\mathbf{R}$. This could be considered as an improper integral too, and it is the same as the Lebesgue integral of f on $\mathbf{R}$.

There are analogous notions on $\mathbf{R}^n$ for any positive integer n. If f is a nonnegative real-valued continuous function on $\mathbf{R}^n$, then the *integrability* of f on $\mathbf{R}^n$ can be defined in terms of the boundedness of the integrals of f over any reasonable family of balls, cubes, or other regions that exhaust $\mathbf{R}^n$, and the integral of f on $\mathbf{R}^n$ can be defined as the supremum of these integrals. If f is a real or complex-valued continuous function on $\mathbf{R}^n$, then the integrability of f is defined to mean that $|f|$ is integrable, and this can be used to define the integral of f on $\mathbf{R}^n$ as before. This implies that f is Lebesgue integrable on $\mathbf{R}^n$, and the integral of f on $\mathbf{R}^n$ is the same as the Lebesgue integral.

7.3 Some examples of integrable functions

Let n be a positive integer, and let a be a positive real number. Note that

$$\min(1, |x|^{-a}) \tag{7.3.1}$$

is continuous on $\mathbf{R}^n$, which is interpreted as being equal to 1 at $x = 0$. One can check that this function is integrable on $\mathbf{R}^n$ exactly when $a > n$.

It is easy to see that $\exp(-|x|^2)$ is integrable on $\mathbf{R}^n$. It is well known that

$$\int_{\mathbf{R}^n} \exp(-|x|^2)\, dx = \pi^{n/2}. \tag{7.3.2}$$

More precisely, the $n = 2$ case can be obtained using polar coordinates. The 2-dimensional integral is the same as the square of the one-dimensional integral, which can be used to get the $n = 1$ case. Similarly, the n-dimensional integral is equal to the nth power of the one-dimensional integral.

If a is a positive real number again, then $\exp(-a\,|x|^2)$ is integrable on $\mathbf{R}^n$. One can check that

$$\int_{\mathbf{R}^n} \exp(-a\,|x|^2)\, dx = (\pi/a)^{n/2}, \tag{7.3.3}$$

using a change of variables.

If $b \in \mathbf{R}^n$, then it is easy to see that

$$\exp(-a\,|x|^2 + b \cdot x) \tag{7.3.4}$$

is integrable on $\mathbf{R}^n$. Observe that

$$\exp(-a\,|x|^2 + b \cdot x) = \exp(-a\,|x - (2\,a)^{-1} b|^2 + (4\,a)^{-1} |b|^2) \tag{7.3.5}$$

for every $x \in \mathbf{R}^n$. It follows that

$$\int_{\mathbf{R}^n} \exp(-a\,|x|^2 + b \cdot x)\, dx = (\pi/a)^{n/2} \exp((4\,a)^{-1}\,|b|^2), \tag{7.3.6}$$

using (7.3.3) and a change of variables.

In fact, (7.3.4) is integrable on $\mathbf{R}^n$ when $b \in \mathbf{C}^n$. It is well known that

$$\int_{\mathbf{R}^n} \exp(-a\,|x|^2 + b \cdot x)\, dx = (\pi/a)^{n/2} \exp((4\,a)^{-1}\, b \cdot b) \tag{7.3.7}$$

for every $b \in \mathbf{C}^n$, which is the same as (7.3.6) when $b \in \mathbf{R}^n$. One can first reduce to the case where $n = 1$, because the left side is the same as the product of n analogous integrals over $\mathbf{R}$. If $n = 1$, then both sides of (7.3.7) are holomorphic functions of $b \in \mathbf{C}$. This permits one to reduce to the case where $b \in \mathbf{R}$, using standard results in complex analysis.

Alternatively, one can use the fact that

$$\exp(-a\,z^2 + b\,z) = \exp(-a\,(z - (b/(2\,a)))^2 + b^2/(4\,a)) \tag{7.3.8}$$

is a holomorphic function of $z \in \mathbf{C}$. One can reduce to the case where $b \in \mathbf{R}$ again, using Cauchy's theorem to make a suitable change of contour.

As another approach, one can reduce to the case where b is purely imaginary, using a change of variables in x, as before. This corresponds to a Fourier transform, as on p105f of [76], and Theorem 1.4 on p138 of [121].

7.4 Some integral solutions

Let n be a positive integer, and let f be a continuous real or complex-valued function on $\mathbf{R}^n$. If $x \in \mathbf{R}^n$ and $t > 0$, then we would like to put

$$u(x,t) = \int_{\mathbf{R}^n} K(x-y,t)\, f(y)\, dy = \int_{\mathbf{R}^n} (4\pi\, t)^{-n/2} \exp(-|x-y|^2/(4\, t))\, f(y)\, dy. \tag{7.4.1}$$

This is called the *Gauss–Weierstrass integral* of f. The integral on the right is defined as long as

$$\exp(-|x-y|^2/(4\, t))\, f(y) \tag{7.4.2}$$

is integrable as a function of y on $\mathbf{R}^n$, as in Section 7.2. Equivalently, this means that

$$\exp((2x \cdot y - |y|^2)/(4\, t))\, |f(y)| \tag{7.4.3}$$

is integrable as a function of y on $\mathbf{R}^n$.

Let τ be a positive real number, and suppose that there is a nonnegative real number $C(\tau)$ such that

$$|f(y)| \le C(\tau) \exp(|y|^2/(4\, \tau)) \tag{7.4.4}$$

for every $y \in \mathbf{R}^n$. If $0 < t < \tau$ and $x \in \mathbf{R}^n$, then it follows that (7.4.3) is integrable as a function of y on $\mathbf{R}^n$. Thus $u(x,t)$ can be defined as in (7.4.1) in this case.

One can differentiate under the integral sign, to show that $u(x,t)$ is smooth on $\mathbf{R}^n \times (0,\tau)$ under these conditions. In particular, one can check that any number of derivatives of (7.4.2) in x and t is integrable as a function of y on $\mathbf{R}^n$ when $0 < t < \tau$, because of (7.4.4). We also get that $u(x,t)$ satisfies the heat equation on $\mathbf{R}^n \times (0,\tau)$, because $K(x-y,t)$ satisfies the heat equation as a function of (x,t) for every $y \in \mathbf{R}^n$.

If $x, z \in \mathbf{R}^n$ and $0 < t < \tau$, then

$$u(x,t) - f(z) = \int_{\mathbf{R}^n} K(x-y,t)\, (f(y) - f(z))\, dy, \tag{7.4.5}$$

because of (7.1.7). This implies that

$$|u(x,t) - f(z)| \le \int_{\mathbf{R}^n} K(x-y,t)\, |f(y) - f(z)|\, dy, \tag{7.4.6}$$

because $K(x-y,t) \ge 0$. One can use this to show that

$$u(x,t) \to f(z) \tag{7.4.7}$$

as $(x,t) \to (z,0)$ in $\mathbf{R}^n \times \mathbf{R}$. This means that $u(x,t)$ extends continuously to $t \ge 0$, by taking it to be equal to $f(x)$ when $t = 0$. Properties like these are mentioned in Theorem 1 in Section 2.3.1 b of [29], and Theorem 4.3 in Section A of Chapter 4 of [32].

A related convergence property is that

$$u(x,t) \to f(x) \tag{7.4.8}$$

uniformly on bounded subsets of $\mathbf{R}^n$ as $t \to 0$. This can be obtained using the uniform continuity of f on compact subsets of $\mathbf{R}^n$. The previous convergence property can be obtained from this one and the continuity of f on $\mathbf{R}^n$. This convergence property could also be obtained from the continuous extension of $u(x,t)$ to $t \geq 0$, and the uniform continuity of this extension on compact sets.

Suppose now that for every $\tau > 0$ there is a nonnegative real number $C(\tau)$ such that (7.4.4) holds. This implies that $u(x,t)$ can be defined as in (7.4.1) for every $x \in \mathbf{R}^n$ and $t > 0$.

Of course, this condition holds when f is bounded on $\mathbf{R}^n$. This condition also holds when f is the exponential of a linear function on $\mathbf{R}^n$.

7.5 Some related integrability conditions

Let n be a positive integer, and let f be a continuous real or complex-valued function on $\mathbf{R}^n$. Also let τ_1 be a positive real number, and suppose that

$$\exp(-|y|^2/(4\,\tau_1))\,|f(y)| \tag{7.5.1}$$

is integrable on $\mathbf{R}^n$, as in Section 7.2. This implies that (7.4.3) is integrable as a function of y on $\mathbf{R}^n$ when $0 < t < \tau_1$ and $x \in \mathbf{R}^n$. This means that $u(x,t)$ can be defined as in (7.4.1) under these conditions.

If f satisfies (7.4.4) for some $\tau > 0$, then it is easy to see that (7.5.1) is integrable on $\mathbf{R}^n$ for $0 < \tau_1 < \tau$. If (7.5.1) is integrable on $\mathbf{R}^n$ for some $\tau_1 > 0$, then $u(x,t)$ satisfies the same properties on $\mathbf{R}^n \times (0,\tau_1)$ as mentioned in the previous section when (7.4.4) holds.

If (7.5.1) is integrable on $\mathbf{R}^n$ for every $\tau_1 > 0$, then $u(x,t)$ can be defined as in (7.4.1) for every $x \in \mathbf{R}^n$ and $t > 0$. In particular, this holds when for every $\tau > 0$ there is a $C(\tau) \geq 0$ such that (7.4.4) holds. Of course, if f is integrable on $\mathbf{R}^n$, then (7.5.1) is integrable on $\mathbf{R}^n$ for every $\tau_1 > 0$.

If one is familiar with Lebesgue integrals, then one may consider real or complex-valued Lebesgue measurable functions f on $\mathbf{R}^n$. The integral on the right side of (7.4.1) can be defined as a Lebesgue integral when (7.4.2) is Lebesgue integrable as a function of y on $\mathbf{R}^n$. This is equivalent to the Lebesgue integrability of (7.4.3) as a function of y on $\mathbf{R}^n$, as before. Note that this implies that f is locally integrable with respect to Lebesgue measure on $\mathbf{R}^n$.

If (7.5.1) is Lebesgue integrable on $\mathbf{R}^n$ for some $\tau_1 > 0$, then (7.4.3) is Lebesgue integrable as a function of y on $\mathbf{R}^n$ when $0 < t < \tau_1$ and $x \in \mathbf{R}^n$, as before. This implies that $u(x,t)$ can be defined as in (7.4.1) on $\mathbf{R}^n \times (0,\tau_1)$. One can differentiate under the integral sign under these conditions too, to get that $u(x,t)$ is smooth on $\mathbf{R}^n \times (0,\tau_1)$. Note that any number of derivatives of (7.4.2) in x and t is Lebesgue integrable as a function of y on $\mathbf{R}^n$ when $0 < t < \tau_1$, because of the Lebesgue integrability of (7.5.1) on $\mathbf{R}^n$. We also have that $u(x,t)$ satisfies the heat equation on $\mathbf{R}^n \times (0, \tau_1)$, as before.

However, the convergence of $u(x,t)$ to $f(x)$ as $t \to 0+$ is more complicated in this case. Some results along these lines are mentioned in Theorem 4.3 in Section A of Chapter 4 of [32], and Theorems 1.18 and 1.25 on p10, 13 of [125].

There are continuity and convergence results like those mentioned in the previous section at points where f is continuous. One can also use Riemann integrals on suitable regions in $\mathbf{R}^n$, and corresponding improper integrals on $\mathbf{R}^n$, to deal with some types of functions that may not be continuous, instead of Lebesgue integrals.

7.6 Translations and integrability

Let n be a positive integer, and let f be a continuous real or complex-valued function on $\mathbf{R}^n$. If $a \in \mathbf{R}^n$, then

$$f_a(x) = f(x-a) \tag{7.6.1}$$

is a continuous function on $\mathbf{R}^n$ as well. Note that f is integrable on $\mathbf{R}^n$ if and only if f_a is integrable on $\mathbf{R}^n$, in which case

$$\int_{\mathbf{R}^n} f_a(x)\,dx = \int_{\mathbf{R}^n} f(x)\,dx. \tag{7.6.2}$$

Of course, this also holds with $|f|$ in place of f, so that

$$\int_{\mathbf{R}^n} |f_a(x)|\,dx = \int_{\mathbf{R}^n} |f(x)|\,dx. \tag{7.6.3}$$

Let $x \in \mathbf{R}^n$ and $t > 0$ be given. Observe that

$$\begin{aligned} &\exp(-|x-y|^2/(4\,t))\,f(y-a) \\ &= \exp(-|(x-a)-(y-a)|^2/(4\,t))\,f(y-a) \end{aligned} \tag{7.6.4}$$

is the same as

$$\exp(-|(x-a)-y|^2/(4\,t))\,f(y) \tag{7.6.5}$$

with y replaced by $y-a$. Thus (7.6.4) is integrable on $\mathbf{R}^n$ if and only if (7.6.5) is integrable on $\mathbf{R}^n$, as in the preceding paragraph. In this case, we get that

$$u(x-a,t) = \int_{\mathbf{R}^n} (4\pi\,t)^{-n/2}\,\exp(-|x-y|^2/(4\,t))\,f(y-a)\,dy, \tag{7.6.6}$$

where the left side is defined as in (7.4.1).

Note that

$$\exp(-|y|^2/(4\,t))\,|f(y-a)| \tag{7.6.7}$$

is integrable on $\mathbf{R}^n$ if and only if

$$\exp(-|y + a|^2/(4\,t))\,|f(y)| \tag{7.6.8}$$

is integrable on $\mathbf{R}^n$. This is equivalent to the integrability of

$$\exp((-2\,a \cdot y - |y|^2)/(4\,t))\,|f(y)| \tag{7.6.9}$$

If (7.5.1) is integrable on $\mathbf{R}^n$ for some $\tau_1 > 0$, then (7.6.9) is integrable on $\mathbf{R}^n$ when $0 < t < \tau_1$, as in the previous section. This means that (7.6.7) is integrable on $\mathbf{R}^n$ when $0 < t < \tau_1$. Of course, there are analogous statements for Lebesgue measurable functions f using Lebesgue integrability.

Similarly, (7.6.7) is bounded on $\mathbf{R}^n$ if and only if (7.6.8) is bounded on $\mathbf{R}^n$, which is equivalent to the boundedness of (7.6.9) on $\mathbf{R}^n$. Suppose that

$$\exp(-|y|^2/(4\,\tau))\,|f(y)| \tag{7.6.10}$$

is bounded on $\mathbf{R}^n$ for some $\tau > 0$, which is the same as saying that (7.4.4) holds for some $C(\tau) \geq 0$. If $0 < t < \tau$, then it follows that (7.6.9) is bounded on $\mathbf{R}^n$, so that (7.6.7) is bounded on $\mathbf{R}^n$.

Let $0 < \tau_0 \leq +\infty$ be given. Consider the condition that (7.5.1) be integrable on $\mathbf{R}^n$ for every positive real number $\tau_1 < \tau_0$. This implies that f_a satisfies the analogous condition, by the earlier remarks. Similarly, consider the condition that (7.6.10) be bounded on $\mathbf{R}^n$ for every $0 < \tau < \tau_0$. This implies that f_a satisfies the analogous condition, as in the preceding paragraph.

7.7 Some properties of these solutions

Let n be a positive integer, and let f be a continuous real or complex-valued function on $\mathbf{R}^n$. Also let $x \in \mathbf{R}^n$ and $t > 0$ be given, and suppose for the moment that (7.4.2) or equivalently (7.4.3) is integrable as a function of y on $\mathbf{R}^n$. Thus $u(x,t)$ may be defined as in (7.4.1), and we have that

$$\begin{aligned} |u(x,t)| &\leq \int_{\mathbf{R}^n} K(x-y,t)\,|f(y)|\,dy \\ &= \int_{\mathbf{R}^n} (4\pi\,t)^{-n/2}\,\exp(-|x-y|^2/(4\,t))\,|f(y)|\,dy. \end{aligned} \tag{7.7.1}$$

This also works when f is Lebesgue measurable on $\mathbf{R}^n$, and (7.4.2) or equivalently (7.4.3) is Lebesgue integrable as a function of y on $\mathbf{R}^n$.

Of course, if f is real-valued on $\mathbf{R}^n$, then $u(x,t) \in \mathbf{R}$. If f is also nonnegative on $\mathbf{R}^n$, then

$$u(x,t) \geq 0. \tag{7.7.2}$$

Similarly, if $f(y) \geq a$ for some $a \in \mathbf{R}$ and every $y \in \mathbf{R}^n$, then

$$u(x,t) \geq a, \tag{7.7.3}$$

because of (7.1.7). If $f(y) \le b$ for some $b \in \mathbf{R}$ and every $y \in \mathbf{R}^n$ then

$$u(x,t) \le b, \tag{7.7.4}$$

for basically the same reasons.

If f is a bounded continuous complex-valued function on $\mathbf{R}^n$, then $u(x,t)$ is defined for every $x \in \mathbf{R}^n$ and $t > 0$, as in Section 7.4. More precisely, suppose that

$$|f(y)| \le C \tag{7.7.5}$$

for some $C \ge 0$ and every $y \in \mathbf{R}^n$. This implies that

$$|u(x,t)| \le C \tag{7.7.6}$$

for every $x \in \mathbf{R}^n$ and $t > 0$, because of (7.1.7) and (7.7.1). This works when f is a bounded Lebesgue measurable function on $\mathbf{R}^n$ as well. If f is a constant on $\mathbf{R}^n$, then $u(x,t)$ is equal to the same constant for every $x \in \mathbf{R}^n$ and $t > 0$.

Suppose now that f is a real or complex-valued function on $\mathbf{R}^n$ that is continuous and integrable, or simply Lebesgue integrable. This implies that $u(x,t)$ may be defined as in (7.4.1) for every $x \in \mathbf{R}^n$ and $t > 0$. In this case, $u(x,t)$ is integrable as a function of x on $\mathbf{R}^n$ for every $t > 0$, with

$$\int_{\mathbf{R}^n} |u(x,t)|\,dx \le \int_{\mathbf{R}^n} |f(y)|\,dy. \tag{7.7.7}$$

This can be obtained from (7.7.1) by interchanging the order of integration and using (7.1.7). Similarly,

$$\int_{\mathbf{R}^n} u(x,t)\,dx = \int_{\mathbf{R}^n} f(y)\,dy \tag{7.7.8}$$

for every $t > 0$.

One can also show that

$$\lim_{t \to 0+} \int_{\mathbf{R}^n} |u(x,t) - f(x)|\,dx = 0. \tag{7.7.9}$$

This corresponds to taking $p = 1$ in Theorem 4.3 in Section A of Chapter 4 of [32], and Theorem 1.18 on p10 of [125]. This is simpler when f is a continuous function on $\mathbf{R}^n$ with compact support. Otherwise, one can approximate f by such functions.

7.8 Parabolic boundaries and maxima

Let n be a positive integer, let V be a nonempty bounded open subset of $\mathbf{R}^n$, and let T be a positive real number. Thus

$$U = V \times (0,T) \tag{7.8.1}$$

is a bounded open subset of $\mathbf{R}^n \times \mathbf{R}$, which we identify with $\mathbf{R}^{n+1}$, as usual. The closure $\overline{U}$ of U in $\mathbf{R}^n \times \mathbf{R}$ is given by

$$\overline{U} = \overline{V} \times [0, T], \tag{7.8.2}$$

where $\overline{V}$ is the closure of V in $\mathbf{R}^n$. The boundary ∂U of U in $\mathbf{R}^n \times \mathbf{R}$ is given by

$$\partial U = (\overline{V} \times \{0\}) \cup ((\partial V) \times [0, T]) \cup (\overline{V} \times \{T\}), \tag{7.8.3}$$

where ∂V is the boundary of V in $\mathbf{R}^n$.

Note that

$$(\overline{V} \times \{0\}) \cup ((\partial V) \times [0, T]) \tag{7.8.4}$$

is a closed set in $\mathbf{R}^n \times \mathbf{R}$ that is contained in ∂U. This may be called the *parabolic boundary* of U, as in Section 2.3.2 of [29], although the term is used slightly differently there. This is the same as

$$(V \times \{0\}) \cup ((\partial V) \times [0, T]), \tag{7.8.5}$$

because $(\partial V) \times \{0\}$ is contained in the second part of the union.

Let u be a continuous real-valued function on $\overline{U}$, and suppose that u is continuously differentiable on U, and that the second derivatives of $u(x, t)$ in x exist and are continuous on U. Remember that u attains its maximum on $\overline{U}$, by the extreme value theorem. If u satisfies the heat equation on U, then it is well known that the maximum of u on $\overline{U}$ is attained on the parabolic boundary of U. This is the *maximum principle* for the heat equation, as in Theorem 4.10 in Section B of Chapter 4 of [32].

This corresponds to part (i) of Theorem 4 in Section 2.3.3 a of [29]. Part (ii) of that theorem is a version of the strong maximum principle for the heat equation. The proof uses a mean-value property for the heat equation in Theorem 3 of Section 2.3.2 of [29]. A more direct approach to the first part is indicated in Problem 16 in Section 2.5 of [29], which is similar to the argument in [32], that we shall follow here. This version of the maximum principle also works for subsolutions of the heat equation, which will be discussed in the next section.

If $u = 0$ on the parabolic boundary of U, then the maximum principle implies that $u \le 0$ on $\overline{U}$. The same argument could be applied to $-u$, to get that $u \equiv 0$ on $\overline{U}$. This corresponds to Theorem 5 in Section 2.3.3 a of [29], and to Corollary 4.11 in Section B of Chapter 4 of [32].

Of course, the parabolic boundary (7.8.4) of U is closed and bounded in $\mathbf{R}^n \times \mathbf{R}$, and thus compact. If u is any continuous real-valued function on $\overline{U}$, then the maximum of u on the parabolic boundary of U is attained, by the extreme value theorem. In order to show that the maximum of u on $\overline{U}$ is attained on the parabolic boundary of U, it suffices to show that for each $(x, t) \in \overline{U}$, $u(x, t)$ is less than or equal to the maximum of u on the parabolic boundary of U.

Suppose now that u is a continuous complex-valued function on $\overline{U}$ that is continuously differentiable on U, and that the second derivatives of $u(x, t)$ in x exist and are continuous on U. Thus the previous statements for real-valued functions can be applied to the real and imaginary parts of u. Similarly, if $\alpha \in \mathbf{C}$, then the previous statements can be applied to

$$\mathrm{Re}(\alpha\, u(x,t)). \tag{7.8.6}$$

If $w \in \mathbf{C}$, then it is easy to see that

$$|w| = \max\{\mathrm{Re}(\alpha\, w) : \alpha \in \mathbf{C},\ |\alpha| = 1\}. \tag{7.8.7}$$

One can use this to show that the maximum of $|u|$ on $\overline{U}$ is attained on the parabolic boundary of U, because of the analogous statement for (7.8.6).

7.9 Subsolutions of the heat equation

Let n be a positive integer, let U be a nonempty open subset of $\mathbf{R}^n \times \mathbf{R}$, and let u be a real-valued function on U. Suppose that u is continuously differentiable on U, and that the second derivatives of $u(x,t)$ exist and are continuous on U. If

$$\frac{\partial u}{\partial t} \le \Delta u \tag{7.9.1}$$

on U, then u is said to be a *subsolution* of the heat equation, as in Problem 17 in Section 2.5 of [29]. Let us say that u is a *strict subsolution* of the heat equation if

$$\frac{\partial u}{\partial t} < \Delta u \tag{7.9.2}$$

on U.

Suppose that u has a local maximum at $(\xi,\tau) \in U$. This implies that (ξ,τ) is a critical point of u, and that the second derivative of u at (ξ,τ) in x_j is less than or equal to 0 for each $j = 1,\ldots,n$. It follows that u is not a strict subsolution of the heat equation on U.

Now let V, T, and U be as in the previous section, and let u be a continuous real-valued function on $\overline{U}$. Suppose that u is continuously differentiable on U again, and that the second derivatives of $u(x,t)$ in x exist and are continuous on U. Suppose for the moment that u is a strict subsolution of the heat equation on U.

Let R be a positive real number with $R < T$, and note that

$$\overline{V} \times [0,R] \tag{7.9.3}$$

is closed and bounded in $\mathbf{R}^n \times \mathbf{R}$, and thus compact. This means that the maximum of u on (7.9.3) is attained, by the extreme value theorem. The maximum of u on (7.9.3) cannot be attained at a point in

$$V \times (0,R), \tag{7.9.4}$$

as before.

Suppose for the sake of a contradiction that u has a local maximum at (ξ, R) for some $\xi \in V$, as a function on (7.9.3). In particular, (ξ, R) is a local maximum for u as a function on

$$V \times \{R\}, \tag{7.9.5}$$

so that ξ is a critical point for $u(x, R)$ as a function of x, and the second derivative of u at (ξ, R) in x_j is less than or equal to 0 for each $j = 1, \ldots, n$. We also get that the derivative of u in t at (ξ, R) is greater than or equal to 0, because (ξ, R) is a local maximum for u on (7.9.3). This is not possible, because u is supposed to be a strict subsolution of the heat equation on U.

It follows that the maximum of u on (7.9.3) can only be attained at a point in

$$(\overline{V} \times \{0\}) \cup ((\partial V) \times [0, R]). \tag{7.9.6}$$

This is the parabolic boundary of (7.9.4), as in the previous section.

Remember that the maximum of u on the parabolic boundary (7.8.4) of U is attained, by the extreme value theorem. Of course, the maximum of u on (7.9.6) is less than or equal to the maximum of u on (7.8.4), because (7.9.6) is contained in (7.8.4). This implies that the maximum of u on (7.9.3) is less than or equal to the maximum of u on (7.8.4), by the statement in the preceding paragraph. One can use this to get that the maximum of u on $\overline{U}$ is attained on (7.8.4), because the previous statement holds for all $R \in (0, T)$.

In [29], one typically asks that the regularity properties of u extend to the "parabolic cylinder", which includes

$$V \times \{T\}. \tag{7.9.7}$$

In this case, one can get directly that the maximum of u on $\overline{U}$ can only be attained on the parabolic boundary of U, as before.

Suppose now that u is a non-strict subsolution of the heat equation on U, and let $\epsilon > 0$. It is easy to see that

$$u_\epsilon(x, t) = u(x, t) - \epsilon\, t \tag{7.9.8}$$

and

$$v_\epsilon(x, t) = u(x, t) + \epsilon\, |x|^2 \tag{7.9.9}$$

are strict subsolutions of the heat equation on U. Thus the maxima of u_ϵ and v_ϵ on $\overline{U}$ are attained on the parabolic boundary of U, as before. One can use either of these to get that the maximum of u on $\overline{U}$ is attained on the parabolic boundary of U. This is a version of the maximum principle for subsolutions of the heat equation, as mentioned in the previous section.

7.10 Another approach to uniqueness

Let n be a positive integer, let V be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary, and let T be a positive real number. Put $U = V \times (0, T)$, which is a bounded open subset of $\mathbf{R}^n \times \mathbf{R}$, as before. Let u

be a continuously-differentiable real or complex-valued function on $\overline{U}$, which is twice continuously differentiable in x. More precisely, this means that $u(x,t)$ is twice continuously differentiable as a function of x on $\overline{V}$ for each $t \in [0,T]$, and that all of the second derivatives of $u(x,t)$ in x are continuous on $\overline{U}$.

If $0 \le t \le T$, then put

$$e(t) = \int_V |u(x,t)|^2 \, dx. \tag{7.10.1}$$

Observe that

$$\begin{aligned} \frac{\partial}{\partial t}(|u(x,t)|^2) &= \frac{\partial}{\partial t}(u(x,t)\,\overline{u(x,t)}) = \frac{\partial u}{\partial t}(x,t)\,\overline{u(x,t)} + u(x,t)\,\overline{\frac{\partial u}{\partial t}(x,t)} \\ &= 2\ \mathrm{Re}\Big(\overline{u(x,t)}\,\frac{\partial u}{\partial t}(x,t)\Big). \end{aligned} \tag{7.10.2}$$

We can differentiate under the integral sign under these conditions, to get that

$$\frac{de}{dt}(t) = 2\ \mathrm{Re} \int_V \overline{u(x,t)}\,\frac{\partial u}{\partial t}(x,t)\,dx. \tag{7.10.3}$$

If u satisfies the heat equation, then this implies that

$$\frac{de}{dt}(t) = 2\ \mathrm{Re} \int_V \overline{u(x,t)}\,(\Delta u)(x,t)\,dx. \tag{7.10.4}$$

Suppose that either

$$u(y',t) = 0 \text{ on } (\partial V) \times [0,T] \tag{7.10.5}$$

or

$$(D_{\nu(y')}u)(y',t) = 0 \text{ on } (\partial V) \times [0,T], \tag{7.10.6}$$

where $\nu(y')$ is the outward-pointing unit normal to ∂V in $\mathbf{R}^n$ at $y' \in \partial V$, as usual, and $D_{\nu(y')}$ indicates the directional derivative in the direction $\nu(y')$. In both cases, we can use the divergence theorem, as in Section 3.5, to get that

$$\frac{de}{dt}(t) = -2 \int_V |\nabla u(x,t)|^2 \, dx. \tag{7.10.7}$$

More precisely, $\nabla u(x,t) = \nabla_x u(x,t)$ refers to the gradient of $u(x,t)$ in x. In particular, the right side of (7.10.7) is less than or equal to 0, so that $e(t)$ decreases monotonically on $[0,T]$.

If we also have that

$$u(x,0) = 0 \text{ on } V, \tag{7.10.8}$$

then we get that $e(0) = 0$. This implies that $e(t) = 0$ for every $t \in [0,T]$, because $e(t)$ decreases monotonically on $[0,T]$. This means that

$$u(x,t) = 0 \text{ on } \overline{U} = \overline{V} \times [0,T]. \tag{7.10.9}$$

This corresponds to Theorem 10 in Section 2.3.4 a of [29] in the case of the Dirichlet boundary conditions (7.10.5), and Problem 1 in Section 7.5 of [29] for the Neumann boundary conditions (7.10.6).

7.11 Some integrals of products

Let n be a positive integer, let V be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary, and let a, b be real numbers with $a < b$. Thus $U = V \times (a, b)$ is a bounded open subset of $\mathbf{R}^n \times \mathbf{R}$, with closure $\overline{U} = \overline{V} \times [a, b]$. Let u, v be continuously-differentiable real or complex-valued functions on $\overline{U}$ that are twice continuously differentiable in x. Suppose that u satisfies the heat equation, and that v satisfies the "backwards" heat equation

$$\frac{\partial v}{\partial t} = -\Delta v. \tag{7.11.1}$$

Equivalently, this means that $v(x, -t)$ satisfies the heat equation.

Observe that

$$\frac{\partial}{\partial t}(u\, v) = \frac{\partial u}{\partial t}\, v + u\, \frac{\partial v}{\partial t} = (\Delta u)\, v - u\, (\Delta v). \tag{7.11.2}$$

If $a \le t \le b$, then one can differentiate under the integral sign to get that

$$\frac{d}{dt} \int_V u(x, t)\, v(x, t)\, dx = \int_V ((\Delta u)(x, t)\, v(x, t) - u(x, t)\, (\Delta v)(x, t))\, dx. \tag{7.11.3}$$

This implies that

$$\begin{aligned} &\frac{d}{dt} \int_V u(x, t)\, v(x, t)\, dx \\ &\quad = \int_{\partial V} ((D_{\nu(y')} u)(y', t)\, v(y', t) - u(y', t)\, (D_{\nu(y')} v)(y', t))\, dy', \end{aligned} \tag{7.11.4}$$

as in Section 3.9. Here $\nu(y')$ is the outward-pointing unit normal to ∂V in $\mathbf{R}^n$ at $y' \in \partial V$, and $D_{\nu(y')}$ indicates the directional derivative in the direction $\nu(y')$, as before.

If we integrate in t over $[a, b]$, then we get that

$$\begin{aligned} &\int_V u(x, b)\, v(x, b)\, dx - \int_V u(x, a)\, v(x, a)\, dx \\ &\quad = \int_a^b \int_{\partial V} ((D_{\nu(y')} u)(y', t)\, v(y', t) - u(y', t)\, (D_{\nu(y')} v)(y', t)))\, dy'\, dt. \end{aligned} \tag{7.11.5}$$

This corresponds to Problem 3 in Section 7.5 of [29]. This could also be obtained from the divergence theorem on U, as in the proof of Theorem 4.4 in Section A of Chapter 4 of [32].

Let $K(x, t)$ be the heat kernel as defined on $(\mathbf{R}^n \times \mathbf{R}) \setminus \{(0, 0)\}$ in Section 7.1, so that $K(x, t)$ is smooth and satisfies the heat equation on this set. Let $z \in \mathbf{R}^n$ and $t_1 \in \mathbf{R}$ with $t_1 > b$ be given, and consider

$$v(x, t) = K(x - z, t_1 - t), \tag{7.11.6}$$

which is a smooth function on $(\mathbf{R}^n \times \mathbf{R}) \setminus \{(z, t_1)\}$ that satisfies the backward heat equation. In particular, (7.11.6) is smooth and satisfies the backward heat equation on $\overline{U}$, so that (7.11.4) and (7.11.5) hold in this case.

If $z \in V$, then

$$\int_V u(x,b)\, K(x-z, t_1 - b)\, dx \to u(z,b) \tag{7.11.7}$$

as $t_1 \to b+$, as in Section 7.4. This implies that

$$\begin{aligned} u(z,b) \quad = \quad & \int_V u(x,a)\, K(x-z, b-a)\, dx \\ & + \int_a^b \int_{\partial V} (D_{\nu(y')} u)(y')\, K(y'-z, b-t)\, dy'\, dt \\ & - \int_a^b \int_{\partial V} u(y',t)\, (D_{\nu(y')} K)(y'-z, b-t)\, dy'\, dt, \end{aligned} \tag{7.11.8}$$

by taking the limit as $t_1 \to b+$ in the other terms in (7.11.5).

Suppose now that u is a continuously-differentiable real or complex-valued function on $\mathbf{R}^n \times [a,b]$ that is twice continuously differentiable in x and satisfies the heat equation. If $z \in \mathbf{R}^n$, then we would like to use (7.11.8) to get that

$$\begin{aligned} u(z,b) \quad &= \quad \int_{\mathbf{R}^n} u(x,a)\, K(x-z, b-a)\, dx \\ &= \quad \int_{\mathbf{R}^n} u(x,a)\, (4\pi\,(b-a))^{-n/2}\, \exp(-|x-z|^2/(4\,(b-a)))\, dx \end{aligned} \tag{7.11.9}$$

under suitable conditions, as in the proof of Theorem 4.4 in Section A of Chapter 4 of [32].

Suppose that there are real numbers $b_1 > b$ and $C, C' \ge 0$ such that

$$|u(x,t)| \le C\, \exp(|x|^2/(4\,(b_1 - t))) \tag{7.11.10}$$

and

$$|\nabla u(x,t)| \le C'\, \exp(|x|^2/(4\,(b_1 - t))) \tag{7.11.11}$$

for every $x \in \mathbf{R}^n$ and $a \le t \le b$. Here $\nabla u(x,t) = \nabla_x u(x,t)$ refers to the gradient of $u(x,t)$ in x, as before. In particular, if we take $t = a$ in (7.11.10), then we get that the integrand on the right side of (7.11.9) is integrable on $\mathbf{R}^n$.

If r is a positive real number with $|z| < r$, then we can take $V = B(0,r)$ in (7.11.8). The second and third terms on the right side of (7.11.8) tend to 0 as $r \to \infty$, because of (7.11.10) and (7.11.11). The first term on the right side of (7.11.8) tends to the right side of (7.11.9) as $r \to \infty$, because of (7.11.10) with $t = a$. Thus (7.11.9) holds, as desired.

Suppose that $0 < T \le \infty$, and let $u(x,t)$ is a continuously-differentiable real or complex-valued function on $\mathbf{R}^n \times (0,T)$ that is twice continuously differentiable in x and satisfies the heat equation. If $0 < a < b < T$, then (7.11.9) holds for every $z \in \mathbf{R}^n$ under the conditions mentioned earlier. If $u(x,t)$ has boundary values as $t \to 0+$ in an appropriate sense, then one can use (7.11.9) to express u as the Gauss–Weierstrass integral of its boundary values, under suitable conditions. A version of this is given by Theorem 4.4 in Section A of [32] and its proof.

7.12 Upper bounds and $t = 0$

Let n be a positive integer, and let T be a positive real number. Also let u be a continuous real-valued function on $\mathbf{R}^n \times [0,T]$. Suppose that on $\mathbf{R}^n \times (0,T)$, $u(x,t)$ is continuously differentiable, twice continuously differentiable in x, and satisfies the heat equation. If

$$u(x,0) \le 0 \text{ for every } x \in \mathbf{R}^n, \tag{7.12.1}$$

then we would like to be able to say that

$$u(x,t) \le 0 \text{ for every } (x,t) \in \mathbf{R}^n \times [0,T], \tag{7.12.2}$$

at least under suitable conditions.

We shall do this here when

$$|x|^{-2} \max(u(x,t),0) \to 0 \text{ as } |x| \to \infty, \tag{7.12.3}$$

uniformly over $t \in [0,T]$. An analogous statement with a much weaker condition on $u(x,t)$ is given in Theorem 6 in Section 2.3.3 a of [29], which will be discussed in the next section. More precisely, it suffices to ask that u be a subsolution of the heat equation on $\mathbf{R}^n \times (0,T)$, instead of satisfying the heat equation there.

Let $y \in \mathbf{R}^n$ be given, and observe that

$$|x|^2 + 2\,n\,t \tag{7.12.4}$$

satisfies the heat equation on $\mathbf{R}^n \times \mathbf{R}$. Thus, for each $\epsilon > 0$,

$$v(x,t) = u(x,t) - \epsilon\,(|x|^2 + 2\,n\,t) \tag{7.12.5}$$

is a subsolution of the heat equation on $\mathbf{R}^n \times (0,T)$. Note that

$$v(x,0) \le u(x,0) \le 0 \tag{7.12.6}$$

for every $x \in \mathbf{R}^n$. We also have that

$$v(x,t) \le 0 \tag{7.12.7}$$

for every $t \in [0,T]$ when $|x|$ is sufficiently large, by hypothesis.

It follows that (7.12.7) holds for every $(x,t) \in \mathbf{R}^n \times [0,T]$, by the maximum principle. This implies (7.12.2), because $\epsilon > 0$ is arbitrary.

Suppose now that u satisfies the heat equation on $\mathbf{R}^n \times (0, T)$, and that

$$u(x, 0) = 0 \text{ for every } x \in \mathbf{R}^n. \tag{7.12.8}$$

If

$$|x|^{-2}\, u(x, t) \to 0 \text{ as } |x| \to \infty, \tag{7.12.9}$$

then

$$u(x, t) = 0 \text{ for every } (x, t) \in \mathbf{R}^n \times [0, T], \tag{7.12.10}$$

by the previous argument for u and $-u$. This corresponds to Theorem 7 in Section 2.3.3 a of [29], which has a much weaker condition on the size of $u(x, t)$, as before.

7.13 A weaker condition on $u(x, t)$

Let n, T, and u be as at the beginning of the previous section, and suppose that (7.12.1) holds. Suppose also that there are nonnegative real numbers a, A such that

$$u(x, t) \le A \exp(a\, |x|^2) \tag{7.13.1}$$

for every $(x, t) \in \mathbf{R}^n \times [0, T]$. Under these conditions, we have that (7.12.2) holds, as in Theorem 6 in Section 2.3.3 a of [29]. As in the previous section, it suffices to ask that u be a subsolution of the heat equation on $\mathbf{R}^n \times (0, T)$, instead of satisfying the heat equation there.

As in [29], we suppose first that

$$4\, a\, T < 1. \tag{7.13.2}$$

This implies that

$$4\, a\, (T + \eta) < 1 \tag{7.13.3}$$

for some $\eta > 0$. Note that

$$(T + \eta - t)^{-n/2} \exp(|x|^2 / (4\, (T + \eta - t))) \tag{7.13.4}$$

satisfies the heat equation on $\mathbf{R}^n \times (-\infty, T + \eta)$, as in Section 7.1. It follows that for each $\mu > 0$,

$$w(x, t) = u(x, t) - \mu\, (T + \eta - t)^{-n/2} \exp(|x|^2 / (4\, (T + \eta - t))) \tag{7.13.5}$$

is a subsolution of the heat equation on $\mathbf{R}^n \times (0, T)$.

Clearly

$$w(x, 0) \le u(x, 0) \le 0 \tag{7.13.6}$$

for every $x \in \mathbf{R}^n$. One can check that

$$w(x, t) \le 0 \tag{7.13.7}$$

for every $t \in [0,T]$ when $|x|$ is sufficiently large, using (7.13.1) and (7.13.3). This implies that (7.13.7) holds for every $(x,t) \in \mathbf{R}^n \times [0,T]$, by the maximum principle. It follows that (7.12.2) holds, because $\mu > 0$ is arbitrary.

If (7.13.2) does not hold, then we can use the same argument on smaller intervals that satisfy this condition. One can use this repeatedly to get the same conclusion as before, as in [29].

If u satisfies the heat equation on $\mathbf{R}^n \times (0,T)$, (7.12.8), and

$$|u(x,t)| \le A\, \exp(a\,|x|^2) \tag{7.13.8}$$

for some $a, A \ge 0$ and all $(x,t) \in \mathbf{R}^n \times [0,T]$, then (7.12.10) holds, as in the previous section. This corresponds to Theorem 7 in Section 2.3.3 a of [29], as before.

7.14 Some more integrals of products

Let n be a positive integer, let V be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary, and let a, b be real numbers with $a < b$, as in Section 7.11. Put $U = V \times (a,b)$, and let u, v be continuously-differentiable real or complex-valued functions on $\overline{U}$ that are twice continuously differentiable in x.

If $a \le t \le b$, then

$$\frac{d}{dt}\int_V u(x,t)\,v(x,t)\,dx = \int_V \Big(\frac{\partial u}{\partial t}(x,t)\,v(x\,t) + u(x,t)\,\frac{\partial v}{\partial t}(x,t)\Big)\,dx. \tag{7.14.1}$$

We can combine this with an identity from Section 3.9 to get that

$$\begin{aligned} \frac{d}{dt}&\int_V u(x,t)\,v(x,t)\,dx \\ &= \int_V \Big(\frac{\partial u}{\partial t}(x,t) - (\Delta u)(x,t)\Big)\,v(x,t)\,dx \\ &\quad + \int_V u(x,t)\,\Big(\frac{\partial v}{\partial t}(x,t) + (\Delta v)(x,t)\Big)\,dx \\ &\quad + \int_{\partial V} (u(y',t)\,(D_{\nu(y')}v)(y',t) - v(y',t)(D_{\nu(y')}u)(y',t))\,dy'. \end{aligned} \tag{7.14.2}$$

Here $\nu(y')$ is the outward-pointing unit normal to ∂V in $\mathbf{R}^n$ at $y' \in \partial V$, and $D_{\nu(y')}$ indicates the directional derivative in the direction $\nu(y')$, as usual.

Let us integrate in t over $[a,b]$, to get that

$$\begin{aligned} \int_V &u(x,b)\,v(x,b)\,dx - \int_V u(x,a)\,v(x,a)\,dx \\ &= \int_a^b \int_V \Big(\frac{\partial u}{\partial t}(x,t) - (\Delta u)(x,t)\Big)\,v(x,t)\,dx\,dt \\ &\quad + \int_a^b \int_V u(x,t)\,\Big(\frac{\partial v}{\partial t}(x,t) + (\Delta v)(x,t)\Big)\,dx\,dt \end{aligned} \tag{7.14.3}$$

$$+\int_a^b \int_{\partial V} (u(y',t)\,(D_{\nu(y')}v)(y',t) - v(y',t)\,(D_{\nu(y')}u)(y',t))\,dy'\,dt.$$

Of course, this can be simplified when u satisfies the heat equation, or v satisfies the backward heat equation.

Let $K(x,t)$ be the heat kernel as defined on $(\mathbf{R}^n \times \mathbf{R}) \setminus \{(0,0)\}$ in Section 7.1, which is smooth and satisfies the heat equation on this set. If $z \in \mathbf{R}^n$ and $t_1 \in \mathbf{R}$, then

$$v(x,t) = K(x-z, t_1 - t) \tag{7.14.4}$$

is a smooth function on $(\mathbf{R}^n \times \mathbf{R}) \setminus \{(z,t_1)\}$ that satisfies the backward heat equation, as in Section 7.11. If $t_1 > b$, then (7.14.4) is smooth and satisfies the backward heat equation on $\overline{U}$, as before. In this case, we get that

$$\begin{aligned}
\int_V u(x,b)\,K(x-z,t_1-b)\,dx &- \int_V u(x,a)\,K(x-z,t_1-a)\,dx \\
= &\int_a^b \int_V \Big(\frac{\partial u}{\partial t}(x,t) - (\Delta u)(x,t)\Big)\,K(x-z,t_1-t)\,dx\,dt \\
&+\int_a^b \int_{\partial V} u(y',t)\,(D_{\nu(y')}K)(y'-z,t_1-t)\,dy'\,dt \\
&-\int_a^b \int_{\partial V} K(y'-z,t_1-t)\,(D_{\nu(y')}u)(y',t)\,dy'\,dt.
\end{aligned} \tag{7.14.5}$$

Suppose that $z \in V$, and consider the limit as $t_1 \to b+$ of both sides of the equation, as in Section 7.11. We would like to say that

$$\begin{aligned}
u(z,b) &- \int_V u(x,a)\,K(x-z,b-a)\,dx \\
= &\int_a^b \int_V \Big(\frac{\partial u}{\partial t}(x,t) - (\Delta u)(x,t)\Big)\,K(x-z,b-t)\,dx\,dt \\
&+\int_a^b \int_{\partial V} u(y',t)\,(D_{\nu(y')}K)(y'-z,b-t)\,dy'\,dt \\
&-\int_a^b \int_{\partial V} K(y'-z,b-t)\,(D_{\nu(y')}u)(y',t)\,dy'\,dt.
\end{aligned} \tag{7.14.6}$$

More precisely, the first term on the right side should be handled a bit carefully, as in the next section.

Similarly, if $z \in \mathbf{R}^n$ and $t_0 \in \mathbf{R}$, then

$$u(x,t) = K(x-z, t-t_0) \tag{7.14.7}$$

is a smooth function on $(\mathbf{R}^n \times \mathbf{R}) \setminus \{(z,t_0)\}$ that satisfies the heat equation. If $t_0 < a$, then (7.14.7) is smooth and satisfies the heat equation on $\overline{U}$. If v is as at the beginning of the section again, then we obtain that

$$\int_V K(x-z,b-t_0)\,v(x,b)\,dx - \int_V K(x-z,a-t_0)\,v(x,a)\,dx$$

$$
\begin{aligned}
&= \int_a^b \int_V K(x-z, t-t_0) \Big(\frac{\partial v}{\partial t}(x,t) + (\Delta v)(x,t)\Big)\, dx\, dt \\
(7.14.8) \qquad &+ \int_a^b \int_{\partial V} K(y'-z, t-t_0)\, (D_{\nu(y')} v)(y', t)\, dy'\, dt \\
&- \int_a^b \int_{\partial V} v(y', t)\, (D_{\nu(y')} K)(y'-z, t-t_0)\, dy'\, dt.
\end{aligned}
$$

Suppose that $z \in V$ again, and consider the limit as $t_0 \to a-$ of both sides of the equation. Of course, this is basically the same as the previous version, and we get that

$$
\begin{aligned}
&\int_V K(x-z, b-a)\, v(x,b)\, dx - v(z,a) \\
&\quad = \int_a^b \int_V K(x-z, t-a) \Big(\frac{\partial v}{\partial t}(x,t) + (\Delta v)(x,t)\Big)\, dx\, dt \\
(7.14.9) \qquad &\quad + \int_a^b \int_{\partial V} K(y'-z, t-a)\, (D_{\nu(y')} v)(y', t)\, dy'\, dt \\
&\quad - \int_a^b \int_{\partial V} v(y', t)\, (D_{\nu(y')} K)(y'-z, t-a)\, dy'\, dt.
\end{aligned}
$$

7.15 Some integrals with $K(x,t)$

Let $K(x,t)$ be the heat kernel as defined on $(\mathbf{R}^n \times \mathbf{R}) \setminus \{(0,0)\}$, as in Section 7.1. Remember that

$$\int_{\mathbf{R}^n} K(x,t)\, dx = 1 \tag{7.15.1}$$

for every $t > 0$. This implies that

$$\int_{r_1}^{r_2} \int_{\mathbf{R}^n} K(x,t)\, dx\, dt = r_2 - r_1 \tag{7.15.2}$$

when r_1, r_2 are positive real numbers with $r_1 \le r_2$. One could also allow $r_1 = 0$ here, by considering the integral over t as an improper integral, or defining the integrand at $t = 0$, or using Lebesgue integrals.

One may consider the left side of (7.15.2) as an $(n+1)$-dimensional integral over $\mathbf{R}^n \times [r_1, r_2]$, even when $r_1 = 0$, using suitable improper integrals, or Lebesgue integrals. In particular, $K(x,t)$ is locally integrable on $\mathbf{R}^n \times \mathbf{R}$, with respect to $(n+1)$-dimensional Lebesgue measure.

Let W be a nonempty bounded open subset of $\mathbf{R}^n$, let T be a positive real number, and let f be a continuous real or complex-valued function on $\overline{W} \times [0,T]$. Note that f is bounded on $\overline{W} \times [0,T]$, so that

$$|f(x,t)| \le C \tag{7.15.3}$$

for some $C \ge 0$ and every $x \in \overline{W}$, $t \in [0,T]$. If one is using Riemann integrals, then one should ask for a bit more regularity of the boundary of W, or that f is equal to 0 on $(\partial W) \times [0, T]$. If one is using Lebesgue integrals, then one might simply ask that f be bounded and measurable on $W \times [0, T]$.

If $0 < t \le T$, then

$$\int_W K(x,t)\,|f(x,t)|\,dx \le C, \tag{7.15.4}$$

by (7.15.1) and (7.15.3). This implies that

$$\int_{r_1}^{r_2} \int_W K(x,t)\,|f(x,t)|\,dx\,dt \le C\,(r_2 - r_1) \tag{7.15.5}$$

when $0 < r_1 \le r_2 \le T$. This also works with $r_1 = 0$, with suitable interpretations when $0 \in \overline{W}$, or using Lebesgue integrals, as before. In particular, this means that $K(x,t)\,|f(x,t)|$ is integrable with respect to $(n+1)$-dimensional Lebesgue measure on $W \times [0, T]$.

Similarly, if η is a positive real number, then

$$\int_W K(x, t+\eta)\,|f(x,t)|\,dx \le C \tag{7.15.6}$$

for every $t \in [0, T]$. It follows that

$$\int_{r_1}^{r_2} \int_W K(x,t+\eta)\,|f(x,t)|\,dx\,dt \le C\,(r_2 - r_1) \tag{7.15.7}$$

when $0 \le r_1 \le r_2 \le T$.

Of course,

$$\left| \int_W K(x,t)\, f(x,t)\,dx \right| \le C \tag{7.15.8}$$

when $0 < t \le T$, by (7.15.4). If $r \in [0, T]$, then

$$\int_0^r \int_W K(x,t)\, f(x,t)\,dx\,dt \tag{7.15.9}$$

may be defined directly unless $0 \in \overline{W}$, in which case the integral over t may be considered as an improper integral, or defined a bit carefully as a Riemann integral, or using Lebesgue integrals, as before. Note that

$$\left| \int_0^r \int_W K(x,t)\, f(x,t)\,dx\,dt \right| \le C\,r. \tag{7.15.10}$$

If $0 < r_1 \le r_2 \le T$, then

$$\begin{aligned} \lim_{\eta \to 0+} \int_{r_1}^{r_2} \int_W K(x,t+\eta)\, f(x,t)\,dx\,dt \\ = \int_{r_1}^{r_2} \int_W K(x,t)\, f(x,t)\,dx\,dt, \end{aligned} \tag{7.15.11}$$

by standard arguments. It is not too difficult to show that this works with $r_1 = 0$ as well, using the previous remarks.

Chapter 8

Some more equations and solutions

8.1 Another uniqueness argument

Let n be a positive integer, let V be a nonempty bounded open subset of $\mathbf{R}^n$ with reasonably smooth boundary, and let T be a positive real number. Put $U = V \times (0, T)$, which is a bounded open subset of $\mathbf{R}^n \times \mathbf{R}$, with closure $\overline{U} = \overline{V} \times [0, T]$. Let $u(x, t)$ be a twice continuously-differentiable real or complex-valued function on $\overline{U}$. If $0 \le t \le T$, then put

$$E(t) = \frac{1}{2} \int_V \left(\left| \frac{\partial u}{\partial t}(x, t) \right|^2 + \sum_{j=1}^n \left| \frac{\partial u}{\partial x_j}(x, t) \right|^2 \right) dx. \tag{8.1.1}$$

We can differentiate under the integral sign in t, to get that

$$\frac{d}{dt} E(t) = \mathrm{Re} \int_V \Big(\overline{\frac{\partial u}{\partial t}(x, t)} \, \frac{\partial^2 u}{\partial t^2}(x, t) + \sum_{j=1}^n \frac{\partial u}{\partial x_j}(x, t) \, \overline{\frac{\partial^2 u}{\partial x_j \, \partial t}(x, t)} \Big) \, dx. \tag{8.1.2}$$

Suppose that

$$u(y', t) = 0 \text{ on } (\partial V) \times [0, T], \tag{8.1.3}$$

so that

$$\frac{\partial u}{\partial t}(y', t) = 0 \text{ on } (\partial V) \times [0, T]. \tag{8.1.4}$$

Under these conditions,

$$\int_V (\Delta u)(x, t) \, \overline{\frac{\partial u}{\partial t}(x, t)} \, dx + \int_V \sum_{j=1}^n \frac{\partial u}{\partial x_j}(x, t) \, \overline{\frac{\partial^2 u}{\partial x_j \, \partial t}(x, t)} \, dx = 0, \tag{8.1.5}$$

as in Section 3.5, where $(\Delta u)(x, t)$ refers to the Laplacian of $u(x, t)$ in x. This also works when

$$(D_{\nu(y')} u)(y', t) = 0 \text{ on } (\partial V) \times [0, T], \tag{8.1.6}$$

where $\nu(y')$ is the outward-pointing unit normal to ∂V in $\mathbf{R}^n$ at $y' \in \partial V$, and $D_{\nu(y')}$ indicates the directional derivative in the direction $\nu(y')$. Combining (8.1.2) and (8.1.5), we obtain that

$$\frac{d}{dt}E(t) = \mathrm{Re} \int_V \overline{\frac{\partial u}{\partial t}(x,t)} \Big(\frac{\partial^2 u}{\partial t^2}(x,t) - (\Delta u)(x,t)\Big)\, dx. \tag{8.1.7}$$

Suppose now that u satisfies the *wave equation*

$$\frac{\partial^2 u}{\partial t^2} = \Delta u \tag{8.1.8}$$

on U. In this case, (8.1.7) reduces to

$$\frac{d}{dt}E(t) = 0. \tag{8.1.9}$$

If

$$E(0) = 0, \tag{8.1.10}$$

then it follows that

$$E(t) = 0 \tag{8.1.11}$$

for every $t \in [0,T]$. This means that the first derivatives of $u(x,t)$ in x and t are equal to 0 on U.

Of course, (8.1.10) holds when

$$u(x,0) = \frac{\partial u}{\partial t}(x,0) = 0 \tag{8.1.12}$$

for every $x \in V$. If

$$\frac{\partial u}{\partial t}(x,t) = 0 \tag{8.1.13}$$

on U, then (8.1.12) implies that

$$u(x,t) = 0 \tag{8.1.14}$$

on U. Thus (8.1.14) holds on U when (8.1.12) holds, and (8.1.11) holds for each $t \in [0,T]$. This means that (8.1.14) holds on U when u satisfies (8.1.3), (8.1.8), and (8.1.12). This corresponds to Theorem 5 in Section 2.4.3 a in [29].

A more localized version of this will be discussed in the next section.

8.2 A more localized version

Let n be a positive integer, let T be a positive real number, and let $u(x,t)$ be a twice continuously-differentiable real or complex-valued function on $\mathbf{R}^n \times [0,T]$. Also let $\xi \in \mathbf{R}^n$ and a positive real number $t_0 \le T$ be given, and if $0 \le t \le t_0$, then put

$$e(t) = \frac{1}{2} \int_{B(\xi, t_0 - t)} \bigg(\Big|\frac{\partial u}{\partial t}(x,t)\Big|^2 + \sum_{j=1}^n \Big|\frac{\partial u}{\partial x_j}(x,t)\Big|^2\bigg)\, dx. \tag{8.2.1}$$

Here $B(\xi, t_0 - t)$ is the open ball in $\mathbf{R}^n$ centered at ξ with radius $t_0 - t$, which may be interpreted as the empty set when $t = t_0$. Observe that

$$\begin{aligned}
\frac{d}{dt} e(t) &= \operatorname{Re} \int_{B(\xi, t_0 - t)} \Big(\overline{\frac{\partial u}{\partial t}(x,t)} \, \frac{\partial^2 u}{\partial t^2}(x,t) + \sum_{j=1}^{n} \frac{\partial u}{\partial x_j}(x,t) \, \overline{\frac{\partial^2 u}{\partial x_j \, \partial t}(x,t)} \Big) \, dx \\
&\quad - \frac{1}{2} \int_{\partial B(\xi, t_0 - t)} \left(\left| \frac{\partial u}{\partial t}(y', t) \right|^2 + \sum_{j=1}^{n} \left| \frac{\partial u}{\partial x_j}(y', t) \right|^2 \right) dy'.
\end{aligned} \tag{8.2.2}$$

As in Section 3.5, we have that

$$\begin{aligned}
&\int_{B(\xi, t_0 - t)} (\Delta u)(x,t) \, \overline{\frac{\partial u}{\partial t}(x,t)} \, dx \\
&\qquad + \int_{B(\xi, t_0 - t)} \sum_{j=1}^{n} \frac{\partial u}{\partial x_j}(x,t) \, \overline{\frac{\partial^2 u}{\partial x_j \, \partial t}(x,t)} \, dx \\
&\qquad = \int_{\partial B(\xi, t_0 - t)} (D_{\nu_t(y')} u)(y', t) \, \overline{\frac{\partial u}{\partial t}(y', t)} \, dy'.
\end{aligned} \tag{8.2.3}$$

More precisely, if $y' \in \partial B(\xi, t_0 - t)$, then $\nu_t(y')$ denotes the outward-pointing unit normal to $\partial B(\xi, t_0 - t)$ at y', and $D_{\nu_t(y')}$ indicates the directional derivative in the direction $\nu_t(y')$, as usual. If $u(x,t)$ satisfies the wave equation (8.1.8), then the integral in the first term on the right side of (8.2.2) is the same as the left side of (8.2.3). This means that

$$\begin{aligned}
\frac{d}{dt} e(t) &= \operatorname{Re} \int_{\partial B(\xi, t_0 - t)} (D_{\nu_t(y')} u)(y', t) \, \overline{\frac{\partial u}{\partial t}(y', t)} \, dy' \\
&\quad - \frac{1}{2} \int_{\partial B(\xi, t_0 - t)} \left(\left| \frac{\partial u}{\partial t}(y', t) \right|^2 + \sum_{j=1}^{n} \left| \frac{\partial u}{\partial x_j}(y', t) \right|^2 \right) dy'.
\end{aligned} \tag{8.2.4}$$

We would like to use this to get that

$$\frac{d}{dt} e(t) \le 0 \tag{8.2.5}$$

To do this, note that

$$\begin{aligned}
\operatorname{Re} \left((D_{\nu_t(y')} u)(y', t) \, \overline{\frac{\partial u}{\partial t}(y', t)} \right) &\le |(D_{\nu_t(y')} u)(y', t)| \left| \frac{\partial u}{\partial t}(y', t) \right| \\
&\le |(\nabla u)(y', t)| \left| \frac{\partial u}{\partial t}(y', t) \right|
\end{aligned} \tag{8.2.6}$$

for every $y' \in \partial B(\xi, t_0 - t)$, where $(\nabla u)(x,t)$ is the gradient of $u(x,t)$ in x, as before. The right side of this inequality is less than or equal to

$$\frac{1}{2} \left(|(\nabla u)(y', t)|^2 + \left| \frac{\partial u}{\partial t}(y', t) \right|^2 \right), \tag{8.2.7}$$

because of the well-known fact that $2\,a\,b \le a^2 + b^2$ for all $a, b \in \mathbf{R}$. One can use this to obtain (8.2.5) from (8.2.4).

This shows that $e(t)$ decreases monotonically on $[0, t_0]$. If

$$e(0) = 0, \tag{8.2.8}$$

then we get that

$$e(t) = 0 \tag{8.2.9}$$

when $0 \le t \le t_0$. Suppose now that

$$u(x, 0) = \frac{\partial u}{\partial t}(x, 0) = 0 \tag{8.2.10}$$

for every $x \in B(\xi, t_0)$, which implies that (8.2.8) holds. If

$$0 \le t \le t_0 \text{ and } x \in \overline{B}(\xi, t_0 - t), \tag{8.2.11}$$

then it follows that

$$(\nabla u)(x, t) = \frac{\partial u}{\partial t}(x, t) = 0, \tag{8.2.12}$$

by (8.2.9).

One can use (8.2.10) and (8.2.12) to get that

$$u(x, t) = 0 \tag{8.2.13}$$

when (8.2.11) holds. More precisely, in this argument, we only need that $u(x, t)$ is twice continuously differentiable and satisfies the wave equation on the set where (8.2.11) holds. This corresponds to Theorem 6 in Section 2.4.3 b of [29], and Theorem 5.3 in Section A of Chapter 5 of [32].

8.3 Some differential equations on $\mathbf{R}^2$

Let $w(y_1, y_2)$ be a twice continuously-differentiable real or complex-valued function on $\mathbf{R}^2$, and consider the partial differential equation

$$\frac{\partial^2 w}{\partial y_1\, \partial y_2} = 0. \tag{8.3.1}$$

This equation obviously holds when $w(y_1, y_2)$ depends only on y_1 or y_2. Conversely, it is well known and not too difficult to show that if $w(y_1, y_2)$ satisfies (8.3.1) on $\mathbf{R}^2$, then $w(y_1, y_2)$ can be expressed as the sum of a function of y_1 and a function of y_2. More precisely, (8.3.1) implies that $\partial w/\partial y_1$ does not depend on y_2, and one can use this to get the desired representation of w. Alternatively, one could use (8.3.1) to get that $\partial w/\partial y_2$ does not depend on y_1, and use this to get the same type of representation of w.

Let $u(x, t)$ be a twice continuously-differentiable real or complex-valued function on $\mathbf{R}^2$, and consider the partial differential equation

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0. \tag{8.3.2}$$

This is the same as the wave equation with $n = 1$, and it can also be expressed as

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) u = 0. \tag{8.3.3}$$

It is easy to see that (8.3.1) corresponds to (8.3.3) under the change of variables

$$y_1 = x + t,\ y_2 = x - t. \tag{8.3.4}$$

Clearly any function of $x + t$ or of $x - t$ satisfies (8.3.2). Conversely, if $u(x,t)$ satisfies (8.3.2) on $\mathbf{R}^2$, then $u(x,t)$ can be expressed as a sum of a function of $x + t$ and a function of $x - t$, as before.

Alternatively, put

$$v(x,t) = \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) u(x,t), \tag{8.3.5}$$

which is a continuously differentiable function on $\mathbf{R}^2$. Thus (8.3.3) is the same as saying that

$$\frac{\partial v}{\partial t} + \frac{\partial v}{\partial x} = 0. \tag{8.3.6}$$

This is a linear first-order partial differential equation in v, as in Section 4.1, whose solutions are given by functions of $x - t$ on $\mathbf{R}^2$. Given such a solution, (8.3.5) may be considered as a linear first-order partial differential equation in u. This corresponds to some remarks in Section 2.4.1 a in [29]. Of course, we could consider

$$\widetilde{v}(x,t) = \left(\frac{\partial}{\partial t} + \frac{\partial}{\partial x}\right) u(x,t) \tag{8.3.7}$$

instead, which is another continuously differentiable function on $\mathbf{R}^2$. Using this, (8.3.3) is the same as saying that

$$\frac{\partial \widetilde{v}}{\partial t} - \frac{\partial \widetilde{v}}{\partial x} = 0. \tag{8.3.8}$$

Suppose that

$$u(x,t) = \phi(x+t) + \psi(x-t) \tag{8.3.9}$$

for some continuously-differentiable real or complex-valued functions ϕ, ψ on the real line. This implies that

$$u(x,0) = \phi(x) + \psi(x) \tag{8.3.10}$$

and

$$\frac{\partial u}{\partial t}(x,0) = \phi'(x) - \psi'(x) \tag{8.3.11}$$

for every $x \in \mathbf{R}$. Of course,

$$\frac{\partial u}{\partial x}(x,0) = \phi'(x) + \psi'(x) \tag{8.3.12}$$

for every $x \in \mathbf{R}$, by (8.3.10). It follows that ϕ' and ψ' are uniquely determined by $(\partial u/\partial t)(x,0)$ and $(\partial u/\partial x)(x,0)$ on $\mathbf{R}$. This means that ϕ and ψ are uniquely determined on $\mathbf{R}$, up to adding a constant to ϕ and subtracting the same constant from ψ, by $u(x,0)$ and $(\partial u/\partial t)(x,0)$ on $\mathbf{R}$. Note that the right side of (8.3.9) is not affected by adding a constant to ϕ, and subtracting the same constant from ψ. Thus we get that $u(x,t)$ is uniquely determined on $\mathbf{R}^2$ by $u(x,0)$ and $(\partial u/\partial t)(x,0)$ on $\mathbf{R}$. This corresponds to some more remarks in Section 2.4.1 a of [29], and some remarks in Section B of Chapter 5 of [32].

Observe that ϕ' and ψ' may be arbitrary continuous real or complex-valued functions on $\mathbf{R}$, so that the right sides of (8.3.11) and (8.3.12) may be arbitrary continuous functions on $\mathbf{R}$. Similarly, the right side of (8.3.10) may be any continuously-differentiable function on $\mathbf{R}$, which can be chosen at the same time as the right side of (8.3.11), as an arbitrary continuous function on $\mathbf{R}$. If we take ϕ and ψ to be twice continuously-differentiable functions on $\mathbf{R}$, then the right side of (8.3.10) can be any twice continuously-differentiable function on $\mathbf{R}$, which can be chosen at the same time as the right side of (8.3.11), as an arbitrary continuously-differentiable function on $\mathbf{R}$. In this case, (8.3.9) is a twice continuously-differentiable function on $\mathbf{R}^2$ that satisfies the wave equation (8.3.2). This corresponds to Theorem 1 in Section 2.4.1 a of [29], and Theorem 5.6 in Section B of Chapter 5 of [32].

8.4 Some remarks about the Laplacian

Let n be a positive integer, and let a, b be real numbers with

$$0 \le a < b, \tag{8.4.1}$$

although one could also permit $b = +\infty$ here. Also let f be a twice continuously-differential real or complex-valued on (a,b). Note that

$$\{x \in \mathbf{R}^n : a < |x| < b\} \tag{8.4.2}$$

is an open set in $\mathbf{R}^n$. Put

$$F(x) = f(|x|) \tag{8.4.3}$$

on (8.4.2), which is a twice continuously-differentiable function on this set. One can check that

$$\Delta F(x) = f''(|x|) + (n-1)\,|x|^{-1}\,f'(|x|) \tag{8.4.4}$$

on this set, as in Lemma 2.60 in Section G of Chapter 2 of [32].

Let p be a homogeneous polynomial of degree $k \ge 0$ on $\mathbf{R}^n$, and suppose that p is harmonic on $\mathbf{R}^n$. Thus

$$q(x) = |x|^{-k}\,p(x) \tag{8.4.5}$$

is a twice continuously-differentiable function on $\mathbf{R}^n \setminus \{0\}$ that is homogeneous of degree 0 and equal to p on the unit sphere. One can check that

$$\Delta q(x) = -k\,(k+n-2)\,|x|^{-2}\,q(x) \tag{8.4.6}$$

on $\mathbf{R}^n \setminus \{0\}$, as in Lemma 2.61 in Section G of Chapter 2 of [32]. This was mentioned in Section 3.2 when $|x| = 1$.

Under these conditions, one can also verify that

$$\begin{aligned} &\Delta(F\,q)(x) \\ &= \big(f''(|x|) + (n-1)\,|x|^{-1}\,f'(|x|) - k\,(k+n-2)\,|x|^{-2}\,f(|x|)\big)\,q(x) \end{aligned} \tag{8.4.7}$$

on (8.4.2) as in Lemma 2.62 in Section G of Chapter 2 of [32].

Suppose that

$$f''(r) + (n-1)\,r^{-1}\,f'(r) - k\,(k+n-2)\,r^{-2}\,f(r) = \mu\,f(r) \tag{8.4.8}$$

on (a, b) for some real or complex number μ. Combining this with (8.4.7), we get that

$$\Delta(F\,q) = \mu\,F\,q \tag{8.4.9}$$

on (8.4.2). This is discussed further in Section G of Chapter 2 of [32].

8.5 Some spherical means

Let n be a positive integer, and let ϕ be a continuous real or complex-valued function on $\mathbf{R}^n$. If $x \in \mathbf{R}^n$ and $r \in \mathbf{R}$, then the corresponding *spherical mean* of ϕ may be defined by

$$M_\phi(x; r) = \frac{1}{|\partial B(0,1)|} \int_{\partial B(0,1)} \phi(x + r\,y')\,dy', \tag{8.5.1}$$

as in Section B of Chapter 5 of [32]. Note that

$$M_\phi(x; -r) = M_\phi(x; r), \tag{8.5.2}$$

and that

$$M_\phi(x; 0) = \phi(x), \tag{8.5.3}$$

as in [32]. If $r > 0$, then

$$M_\phi(x; r) = \frac{1}{|\partial B(x,r)|} \int_{\partial B(x,r)} \phi(z')\,dz'. \tag{8.5.4}$$

If $x \in \mathbf{R}^n$ is fixed, then (8.5.1) is continuous as a function of $r \in \mathbf{R}$. Similarly, if ϕ is k-times continuously differentiable on $\mathbf{R}^n$ for some positive integer k, then (8.5.1) is k-times continuously differentiable as a function of $r \in \mathbf{R}$, as in [32].

Let a and b be real numbers with $a < b$, although we could also allow $a = -\infty$ or $b = +\infty$ here. Suppose now that $\phi(x, t)$ is a continuous real or complex-valued function on $\mathbf{R}^n \times (a, b)$. If $x \in \mathbf{R}^n$ and $t \in (a, b)$, then put

$$M_\phi(x; r, t) = \frac{1}{|\partial B(0,1)|} \int_{\partial B(0,1)} \phi(x + r\,y', t)\,dy' \tag{8.5.5}$$

for each $r \in \mathbf{R}$. This is the same as (8.5.1) for $\phi(x,t)$ as a function of x on $\mathbf{R}^n$ for each $t \in (a,b)$. In particular,

$$M_\phi(x; -r, t) = M_\phi(x; r, t) \tag{8.5.6}$$

for each $r \in \mathbf{R}$,

$$M_\phi(x; 0, t) = \phi(x,t), \tag{8.5.7}$$

and

$$M_\phi(x; r, t) = \frac{1}{|\partial B(x,r)|} \int_{\partial B(x,r)} \phi(z', t)\, dz' \tag{8.5.8}$$

when $r > 0$, as before.

If $x \in \mathbf{R}^n$ is fixed, then (8.5.5) is continuous as a function of $(r,t) \in \mathbf{R} \times (a,b)$, as before. Similarly, if $\phi(w,t)$ is k-times continuously differentiable on $\mathbf{R}^n \times (a,b)$, then (8.5.5) if k-times continuously-differentiable as a function of $(r,t) \in \mathbf{R} \times (a,b)$.

Let us now take $a = 0$ and $b = +\infty$, and let $u(x,t)$ be a twice continuously-differentiable real or complex-valued function on $\mathbf{R}^n \times \mathbf{R}_+$ that satisfies the wave equation. If $x \in \mathbf{R}^n$, then $M_u(x; r, t)$ satisfies

$$\frac{\partial^2 M_u}{\partial t^2}(x; r, t) = \frac{\partial^2 M_u}{\partial r^2}(x; r, t) + \frac{n-1}{r} \frac{\partial M_u}{\partial r}(x; r, t) \tag{8.5.9}$$

for $r, t > 0$, as in Proposition 5.8 in Section B of Chapter 5 of [32]. This corresponds to part of Lemma 1 in Section 2.4.1 b of [29]. This is known as the *Euler–Poisson–Darboux equation*, as in [29].

Chapter 9

Some distribution theory

9.1 Fundamental solutions

Let n be a positive integer, and let u, v be complex-valued functions on $\mathbf{R}^n$, at least one of which has compact support in $\mathbf{R}^n$. If α is a multi-index, and if u, v are $|\alpha|$-times continuously differentiable on $\mathbf{R}^n$, then

$$\int_{\mathbf{R}^n} (\partial^\alpha u)(x)\, v(x)\, dx = (-1)^{|\alpha|} \int_{\mathbf{R}^n} u(x)\, (\partial^\alpha v)(x)\, dx, \tag{9.1.1}$$

by integration by parts.

Let N be a nonnegative integer, and let

$$p(w) = \sum_{|\alpha| \le N} a_\alpha\, w^\alpha \tag{9.1.2}$$

be a polynomial in the n variables $w_1, \ldots, w_n$ of degree less than or equal to N, and with complex coefficients a_α. Put

$$\widetilde{p}(w) = p(-w) = \sum_{|\alpha| \le N} (-1)^{|\alpha|}\, a_\alpha\, w^\alpha, \tag{9.1.3}$$

which is another polynomial in $w_1, \ldots, w_n$ of degree less than or equal to n with complex coefficients, as appropriate. Using these polynomials, we get corresponding differential operators $p(\partial)$ and $\widetilde{p}(\partial)$, as in Section 1.7. If u, v are N-times continuously differentiable on $\mathbf{R}^n$, then

$$\int_{\mathbf{R}^n} (p(\partial)u)(x)\, v(x)\, dx = \int_{\mathbf{R}^n} u(x)\, (\widetilde{p}(\partial)v)(x)\, dx, \tag{9.1.4}$$

because of (9.1.1).

A complex-valued function E on $\mathbf{R}^n$ is said to be a *fundamental solution* of $p(\partial)$ if

$$\int_{\mathbf{R}^n} E(x)\, (\widetilde{p}(\partial)v)(x)\, dx = v(0) \tag{9.1.5}$$

for every smooth function v on $\mathbf{R}^n$ with compact support. This is interpreted as meaning that $(p(\partial))(E)$ is the Dirac delta function δ_0 associated to 0, in the sense of distributions. More precisely, this can be extended to distributions E on $\mathbf{R}^n$. If $p \neq 0$, then a famous theorem of Ehrenpreis and Malgrange states that $p(\partial)$ has a fundamental solution on $\mathbf{R}^n$, which may be a distribution. See [27, 32, 37, 101, 111, 115, 138] for more information.

A fundamental solution for the Laplacian is given by the function N defined in Section 6.8. A fundamental solution for the heat operator

$$\frac{\partial}{\partial t} - \Delta \tag{9.1.6}$$

is given by the heat kernel. Once one has a fundamental solution E for $p(\partial)$, one can solve

$$(p(\partial))(u) = f \tag{9.1.7}$$

by convolving E with f under suitable conditions, in the sense of distributions. If f is a smooth function with compact support on $\mathbf{R}^n$, then this gives a smooth solution u of (9.1.7).

Some basic aspects of distribution theory will be discussed in the next sections, and some additional references about this include [14, 34, 35, 36, 47, 65, 70, 88, 98, 124, 128, 133, 156].

9.2 Spaces of test functions

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Remember that a real or complex-valued function f on U is said to have compact support in U if there is a compact set $E \subseteq \mathbf{R}^n$ such that $E \subseteq U$ and $f = 0$ on $U \setminus E$, as in Section 1.9. Note that the union of two compact subsets of $\mathbf{R}^n$ is compact too. Using this, it is easy to see that the sum of two real or complex-valued functions on U with compact support in U has compact support in U as well.

Let $C_{com}(U, \mathbf{R})$, $C_{com}(U, \mathbf{C})$ be the spaces of continuous real and complex-valued functions on U with compact support, respectively. These are linear subspaces of the spaces $C(U, \mathbf{R})$, $C(U, \mathbf{C})$ of all continuous real or complex-valued functions on U, as vector spaces over the real and complex numbers, respectively. Similarly, if k is a positive integer, then let $C^k_{com}(U, \mathbf{R})$, $C^k_{com}(U, \mathbf{C})$ be the spaces of k-times continuously-differentiable real and complex-valued functions on U with compact support. It is sometimes convenient to use the same notation with $k = 0$ for the analogous spaces of continuous functions, as before. The spaces of smooth real or complex-valued functions on U with compact support are denoted $C^\infty_{com}(U, \mathbf{R})$, $C^\infty_{com}(U, \mathbf{C})$.

Equivalently,

$$C^k_{com}(U, \mathbf{R}) = C^k(U, \mathbf{R}) \cap C_{com}(U, \mathbf{R}), \tag{9.2.1}$$

$$C^k_{com}(U, \mathbf{C}) = C^k(U, \mathbf{C}) \cap C_{com}(U, \mathbf{C}) \tag{9.2.2}$$

for every $k \geq 1$. These are linear subspaces of $C^k(U, \mathbf{R})$ and $C^k(U, \mathbf{C})$, respectively. Similarly,

$$C^\infty_{com}(U, \mathbf{R}) = C^\infty(U, \mathbf{R}) \cap C_{com}(U, \mathbf{R}), \tag{9.2.3}$$

$$C^\infty_{com}(U, \mathbf{C}) = C^\infty(U, \mathbf{C}) \cap C_{com}(U, \mathbf{C}), \tag{9.2.4}$$

which are linear subspaces of $C^\infty(U, \mathbf{R})$ and $C^\infty(U, \mathbf{C})$, respectively.

If f is a real or complex-valued function on U, then we can extend f to a function on $\mathbf{R}^n$, simply by putting $f = 0$ on $\mathbf{R}^n \backslash U$. If f is a continuous function on U with compact support in U, then this extension of f to $\mathbf{R}^n$ is continuous too. Similarly, if f is k-times continuously differentiable on U for some $k \geq 1$, or if f is smooth on U, and f has compact support in U, then this extension has the analogous property on $\mathbf{R}^n$.

If f is a continuously-differentiable real or complex-valued function on U with compact support in U, then it is easy to see that the partial derivatives of f have compact support in U as well. If α is a multi-index and $|\alpha| \leq k$, then ∂^α defines linear mappings from $C^k_{com}(U, \mathbf{R})$, $C^k_{com}(U, \mathbf{C})$ into $C^{k-|\alpha|}_{com}(U, \mathbf{R})$, $C^{k-|\alpha|}_{com}(U, \mathbf{C})$, respectively. Similarly, ∂^α defines linear mappings from $C^\infty_{com}(U, \mathbf{R})$, $C^\infty_{com}(U, \mathbf{C})$ into themselves.

If f, g are continuous real or complex-valued functions on U, then it is well known that their product $f\,g$ is continuous on U. If f and g are both k-times continuously differentiable on U, or both smooth on U, then $f\,g$ has the same property. If either f or g has compact support in U, then $f\,g$ has compact support in U.

Smooth functions on U with compact support in U are also known as *test functions* on U.

9.3 Distributions

A *linear functional* on a vector space V over the real or complex numbers is a linear mapping from V into $\mathbf{R}$ or $\mathbf{C}$, as appropriate. Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. A *distribution* on U is a linear functional on the space $C^\infty_{com}(U, \mathbf{C})$ of complex-valued test functions on U that is continuous in a certain sense. Before describing the continuity condition, let us mention some basic examples.

Let f be a continuous complex-valued function on U. If ϕ is a test function on U, then put

$$\lambda_f(\phi) = \int_U f(x)\,\phi(x)\,dx. \tag{9.3.1}$$

The right side may be interpreted as a Riemann integral over any suitable region that contains the support of ϕ. This defines a linear functional on $C^\infty_{com}(U, \mathbf{C})$. One can check that $\lambda_f(\phi) = 0$ for every test function ϕ on U only when $f \equiv 0$ on U. This implies that f is uniquely determined by λ_f on $C^\infty_{com}(U, \mathbf{C})$.

Similarly, if f is a complex-valued function on U that is locally integrable on U with respect to Lebesgue measure, then the right side of (9.3.1) may be interpreted as a Lebesgue integral. In this case, one can show that f is determined almost everywhere on U with respect to Lebesgue measure by λ_f on $C^\infty_{com}(U, \mathbf{C})$.

If $x \in U$, then

$$\delta_x(\phi) = \phi(x) \tag{9.3.2}$$

defines a linear functional on $C^\infty_{com}(U, \mathbf{C})$. This is the *Dirac distribution* on U associated to x.

The continuity condition used to define distributions can be described in terms of a suitable notion of convergent sequences of test functions. Let $\{\phi_j\}_{j=1}^\infty$ be a sequence of test functions on U, and let ϕ be another test function on U. We say that

$$\{\phi_j\}_{j=1}^\infty \textit{converges} \text{ to } \phi \text{ in } C^\infty_{com}(U, \mathbf{C}) \tag{9.3.3}$$

if the following two conditions hold. First, there is a compact set $E \subseteq \mathbf{R}^n$ such that $E \subseteq U$ and

$$\phi_j = 0 \text{ on } U \setminus E \tag{9.3.4}$$

for every $j \ge 1$. Second, for every multi-index α, we have that

$$\{\partial^\alpha \phi_j\}_{j=1}^\infty \text{ converges to } \partial^\alpha \phi \text{ uniformly on } U. \tag{9.3.5}$$

In particular, we can take $\alpha = 0$, to get that $\{\phi_j\}_{j=1}^\infty$ converges to ϕ uniformly on U. It follows that

$$\phi = 0 \text{ on } U \setminus E, \tag{9.3.6}$$

because of (9.3.4).

A linear functional λ on $C^\infty_{com}(U, \mathbf{C})$ is said to be a distribution on U if for every sequence $\{\phi_j\}_{j=1}^\infty$ of test functions on U that converges to a test function ϕ on U, in the sense described in the preceding paragraph, we have that

$$\lim_{j \to \infty} \lambda(\phi_j) = \lambda(\phi). \tag{9.3.7}$$

Alternatively, there is a standard topology defined on $C^\infty_{com}(U, \mathbf{C})$, and it is well known that a linear functional on $C^\infty_{com}(U, \mathbf{C})$ is continuous with respect to this topology if and only if it satisfies this continuity condition in terms of convergent sequences. More precisely, one can show that convergence of sequences in $C^\infty_{com}(U, \mathbf{C})$ with respect to this topology is equivalent to the notion of convergence mentioned in the preceding paragraph. In particular, this implies that continuity with respect to this topology on $C^\infty_{com}(\mathbf{R}^n, \mathbf{C})$ automatically implies the sequential continuity condition (9.3.7). It is well known that the converse holds for linear functionals, but this is more complicated in this case than for metric spaces, for instance.

The space of distributions on U may be denoted

$$C^\infty_{com}(U, \mathbf{C})'. \tag{9.3.8}$$

This is a vector space over the complex numbers, with respect to pointwise addition and scalar multiplication of linear functionals on $C^\infty_{com}(U, \mathbf{C})$.

If f is a continuous complex-valued function on U, or a locally integrable function on U with respect to Lebesgue measure, then it is easy to see that (9.3.1) defines a distribution on U. In this case, it is enough to take $\alpha = 0$ in (9.3.5). We also have that

$$f \mapsto \lambda_f \tag{9.3.9}$$

is a linear mapping from $C(U, \mathbf{C})$ into $C^\infty_{com}(U, \mathbf{C})'$, or from the space of locally integrable functions f on U into $C^\infty_{com}(U, \mathbf{C})'$. It is very easy to see that the Dirac distribution associated to $x \in U$ is indeed a distribution on U.

9.4 Some basic properties of distributions

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Suppose that $\{\phi_j\}_{j=1}^\infty$ is a sequence of test functions on U that converges to a test function ϕ on U, in the sense described in the previous section. This implies that

$$\{\partial_l \phi_j\}_{j=1}^\infty \text{ converges to } \partial_l \phi \tag{9.4.1}$$

in the same sense for each $l = 1, \ldots, n$. Similarly, if a is a smooth complex-valued function on U, then one can check that

$$\{a\, \phi_j\}_{j=1}^\infty \text{ converges to } a\, \phi \tag{9.4.2}$$

in this sense as well. This uses the fact that a and its derivatives of any order are bounded on any compact subset of $\mathbf{R}^n$ that is contained in U.

Let λ be a distribution on U, and for each $l = 1, \ldots, n$, put

$$(\partial_l \lambda)(\phi) = -\lambda(\partial_l \phi) \tag{9.4.3}$$

for every test function ϕ on U. It is easy to see that this defines a distribution on U, which is considered as the partial derivative of λ in the lth variable. If f is a continuously-differentiable complex-valued function on U, and λ_f is the distribution associated to f as in (9.3.1), then

$$\partial_l \lambda_f = \lambda_{\partial_l f} \tag{9.4.4}$$

is the distribution associated to $\partial_l f$ on U. This basically corresponds to integration by parts. More precisely, this works when f is a continuous function on U such that the partial derivative $\partial f / \partial x_l$ in the lth variable exists at every point in U, and is continuous on U.

It is easy to see that

$$\partial_j(\partial_l \lambda) = \partial_l(\partial_j \lambda) \tag{9.4.5}$$

for each $j, l = 1, \ldots, n$, using the analogous statement for smooth functions. If α is any multi-index, then one can differentiate λ repeatedly, to get that

$$(\partial^\alpha \lambda)(\phi) = (-1)^{|\alpha|}\, \lambda(\partial^\alpha \phi) \tag{9.4.6}$$

for all test functions ϕ on U. If f is an $|\alpha|$-times continuously-differentiable complex-valued function on U, and λ_f is the distribution on U associated to f as before, then

$$\partial^\alpha \lambda_f = \lambda_{\partial^\alpha f} \tag{9.4.7}$$

is the distribution associated to $\partial^\alpha f$ on U.

If a is a smooth complex-valued function on U, then put

$$(a\,\lambda)(\phi) = \lambda(a\,\phi) \tag{9.4.8}$$

for every test function ϕ on U. This defines a distribution on U, which is considered as the product of a and λ. If f is a continuous or simply locally-integrable complex-valued function on U, then

$$a\,\lambda_f = \lambda_{a\,f} \tag{9.4.9}$$

is the distribution on U associated to the usual product $a\,f$ of a and f on U. One can check that

$$\partial_l(a\,\lambda) = (\partial_l a)\,\lambda + a\,(\partial_l \lambda), \tag{9.4.10}$$

as distributions on U, using the usual product rule for partial derivatives of smooth functions on U.

One may consider λ to be real-valued as a distribution on U if

$$\lambda(\phi) \in \mathbf{R} \tag{9.4.11}$$

for every real-valued test function ϕ on U. In this case, one may say that λ is nonnegative as a distribution on U, or

$$\lambda \ge 0, \tag{9.4.12}$$

if

$$\lambda(\phi) \ge 0 \tag{9.4.13}$$

for every nonnegative real-valued test function ϕ on U. If λ is the distribution associated to a continuous function f on U, then these conditions correspond to their usual versions for f. If f is locally integrable with respect to Lebesgue measure, and not necessarily continuous, then the analogous conditions on f should be interpreted as holding almost everywhere with respect to Lebesgue measure, as usual.

9.5 Using a fixed compact set

Let n be a positive integer, and let K be a nonempty compact subset of $\mathbf{R}^n$. Consider the space $C_K^\infty(\mathbf{R}^n, \mathbf{C})$ of smooth complex-valued functions ϕ on $\mathbf{R}^n$ such that

$$\operatorname{supp}\phi \subseteq K. \tag{9.5.1}$$

Equivalently, this means that $\phi = 0$ on $\mathbf{R}^n \backslash K$. Note that $C_K^\infty(\mathbf{R}^n, \mathbf{C})$ is a linear subspace of the space $C_{com}^\infty(\mathbf{R}^n, \mathbf{C})$ of all smooth complex-valued functions on $\mathbf{R}^n$ with compact support, as a vector space over the complex numbers.

Let $\{\phi_j\}_{j=1}^\infty$ be a sequence of elements of $C_K^\infty(\mathbf{R}^n, \mathbf{C})$, and let ϕ be another element of $C_K^\infty(\mathbf{R}^n, \mathbf{C})$. Let us say that $\{\phi_j\}_{j=1}^\infty$ *converges* to ϕ in $C_K^\infty(\mathbf{R}^n, \mathbf{C})$ if for every multi-index α,

$$\{\partial^\alpha \phi_j\}_{j=1}^\infty \text{ converges to } \partial^\alpha \phi \text{ uniformly on } K. \tag{9.5.2}$$

Let λ be a linear functional on $C_K^\infty(\mathbf{R}^n, \mathbf{C})$. We can use the notion of convergent sequences in $C_K^\infty(\mathbf{R}^n, \mathbf{C})$ described in the preceding paragraph to define a natural continuity condition for λ. This condition asks that

$$\lim_{j \to \infty} \lambda(\phi_j) = \lambda(\phi) \tag{9.5.3}$$

for every sequence $\{\phi_j\}_{j=1}^\infty$ of elements of $C_K^\infty(\mathbf{R}^n, \mathbf{C})$ that converges to an element ϕ of $C_K^\infty(\mathbf{R}^n, \mathbf{C})$ in this sense.

Alternatively, there is a standard topology defined on $C_K^\infty(\mathbf{R}^n, \mathbf{C})$, for which convergence of sequences is equivalent to the notion of convergence mentioned before. In this case, one can get the equivalence between continuity and sequential continuity more directly. In particular, the continuity condition for a linear functional mentioned in the preceding paragraph is equivalent to continuity with respect to this topology.

One can show that λ is continuous on $C_K^\infty(\mathbf{R}^n, \mathbf{C})$ with respect to this topology if and only if there are a nonnegative real number C and a nonnegative integer N such that

$$|\lambda(\phi)| \le C \sum_{|\alpha| \le N} \Big(\max_{x \in K} |(\partial^\alpha \phi)(x)| \Big) \tag{9.5.4}$$

for every $\phi \in C_K^\infty(\mathbf{R}^n, \mathbf{C})$. The sum on the right is taken over all multi-indices α with $|\alpha| \le N$, as usual. This condition implies that (9.5.3) holds whenever (9.5.2) holds for all such multi-indices.

9.6 Compact sets in open sets

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Every element of $C_{com}^\infty(U, \mathbf{C})$ can be extended to an element of $C^\infty(\mathbf{R}^n, \mathbf{C})$, by putting it equal to 0 on $\mathbf{R}^n \backslash U$, as in Section 9.2. Using this, we can identify $C_{com}^\infty(U, \mathbf{C})$ with a linear subspace of $C_{com}^\infty(\mathbf{R}^n, \mathbf{C})$. With this identification, $C_{com}^\infty(U, \mathbf{C})$ corresponds to the union of $C_K^\infty(\mathbf{R}^n, \mathbf{C})$ over all nonempty compact subsets K of $\mathbf{R}^n$ such that $K \subseteq U$.

If K is a nonempty compact subset of $\mathbf{R}^n$ that is contained in U, then a convergent sequence in $C_K^\infty(\mathbf{R}^n, \mathbf{C})$, in the sense described in the previous section, may be considered as a convergent sequence in $C_{com}^\infty(U, \mathbf{C})$, in the sense of Section 9.3. Conversely, any convergent sequence in $C_{com}^\infty(U, \mathbf{C})$, in the sense of Section 9.3, corresponds to a convergent sequence in $C^\infty_K(\mathbf{R}^n, \mathbf{C})$ for some nonempty compact subset K of $\mathbf{R}^n$ that is contained in U.

Let λ be a linear functional on $C^\infty_{com}(U, \mathbf{C})$. If K is a nonempty compact subset of $\mathbf{R}^n$ that is contained in U, then the restriction of λ to $C^\infty_K(\mathbf{R}^n, \mathbf{C})$ defines a linear functional on that vector space. Observe that λ satisfies the continuity condition on $C^\infty_{com}(U, \mathbf{C})$ described in Section 9.3 if and only if the restriction of λ to $C^\infty_K(\mathbf{R}^n, \mathbf{C})$ satisfies the continuity condition described in the previous section for every nonempty compact subset K of $\mathbf{R}^n$ that is contained in U. This follows from the remarks about convergent sequences in the preceding paragraph.

It is not too difficult to show that there are sequences $\{K_j\}_{j=1}^\infty$ of nonempty compact subsets of $\mathbf{R}^n$ such that

$$\bigcup_{j=1}^\infty K_j = U \tag{9.6.1}$$

and K_j is contained in the interior of K_{j+1} for each j. If K is any compact subset of $\mathbf{R}^n$ that is contained in U, then it follows that

$$K \subseteq K_j \tag{9.6.2}$$

for some j. This implies that $C^\infty_{com}(U, \mathbf{C})$ corresponds to

$$\bigcup_{j=1}^\infty C^\infty_{K_j}(\mathbf{R}^n, \mathbf{C}), \tag{9.6.3}$$

as a linear subspace of $C^\infty_{com}(\mathbf{R}^n, \mathbf{C})$.

9.7 The Schwartz class

Let n be a positive integer. The *Schwartz class* $\mathcal{S}(\mathbf{R}^n)$ is the space of smooth complex-valued functions f on $\mathbf{R}^n$ such that

$$x^\alpha\, (\partial^\beta f)(x) \tag{9.7.1}$$

is bounded on $\mathbf{R}^n$ for all multi-indices α, β. Equivalently, this means that

$$(1 + |x|^2)^k\, |(\partial^\beta f)(x)| \tag{9.7.2}$$

is bounded on $\mathbf{R}^n$ for every nonnegative integer k and multi-index β. It is easy to see that $\mathcal{S}(\mathbf{R}^n)$ is a linear subspace of the space $C^\infty(\mathbf{R}^n, \mathbf{C})$ of all complex-valued smooth functions on $\mathbf{R}^n$, as a vector space over the complex numbers.

Clearly

$$C^\infty_{com}(\mathbf{R}^n, \mathbf{C}) \subseteq \mathcal{S}(\mathbf{R}^n). \tag{9.7.3}$$

If a is a positive real number and $b \in \mathbf{C}^n$, then one can check that

$$\exp(-a\,|x|^2 + b \cdot x) \in \mathcal{S}(\mathbf{R}^n). \tag{9.7.4}$$

If $f \in \mathcal{S}(\mathbf{R}^n)$ and $c \in \mathbf{R}^n$, then one can verify that

$$f(x+c) \in \mathcal{S}(\mathbf{R}^n). \tag{9.7.5}$$

It is easy to see that

$$\partial^\gamma f \in \mathcal{S}(\mathbf{R}^n) \tag{9.7.6}$$

for every multi-index γ in this case too.

If $f \in \mathcal{S}(\mathbf{R}^n)$ and p is a polynomial on $\mathbf{R}^n$ with complex coefficients, then one can check that

$$p\, f \in \mathcal{S}(\mathbf{R}^n). \tag{9.7.7}$$

More precisely, this holds for smooth complex-valued functions p on $\mathbf{R}^n$ with the following property. If γ is any multi-index, then we ask that $\partial^\gamma p$ has at most polynomial growth at infinity on $\mathbf{R}^n$. This means that for every such γ, there are a nonnegative real number $C(\gamma)$ and a nonnegative integer $N(\gamma)$ such that

$$|(\partial^\gamma p)(x)| \le C(\gamma)\,(1+|x|^2)^{N(\gamma)} \tag{9.7.8}$$

for every $x \in \mathbf{R}^n$. Of course, polynomials on $\mathbf{R}^n$ satisfy these conditions.

Let $\{f_j\}_{j=1}^\infty$ be a sequence of elements of $\mathcal{S}(\mathbf{R}^n)$, and let f be another element of $\mathcal{S}(\mathbf{R}^n)$. We say that $\{f_j\}_{j=1}^\infty$ *converges* to f in $\mathcal{S}(\mathbf{R}^n)$ if for every pair of multi-indices α, β,

$$x^\alpha\,(\partial^\beta f_j)(x) \to x^\alpha\,(\partial^\beta f)(x) \text{ as } j \to \infty, \tag{9.7.9}$$

uniformly on $\mathbf{R}^n$. This is the same as saying that for every nonnegative integer N and multi-index β,

$$(1+|x|^2)^N\,|(\partial^\beta f_j)(x) - (\partial^\beta f)(x)| \to 0 \text{ as } j \to \infty, \tag{9.7.10}$$

uniformly on $\mathbf{R}^n$. This is also equivalent to the convergence of $\{f_j\}_{j=1}^\infty$ to f with respect to a standard topology on $\mathcal{S}(\mathbf{R}^n)$. Note that a convergent sequence in $C^\infty_{com}(\mathbf{R}^n, \mathbf{C})$, in the sense described in Section 9.3, converges as a sequence in $\mathcal{S}(\mathbf{R}^n)$.

If $\{f_j\}_{j=1}^\infty$ converges to f in $\mathcal{S}(\mathbf{R}^n)$, then it is easy to see that $\{\partial^\gamma f_j\}_{j=1}^\infty$ converges to $\partial^\gamma f$ in $\mathcal{S}(\mathbf{R}^n)$ for every multi-index γ. If p is a smooth complex-valued function on $\mathbf{R}^n$ whose derivatives of all orders grow at most polynomially on $\mathbf{R}^n$, as before, then one can check that $\{p\, f_j\}_{j=1}^\infty$ converges to $p\, f$ in $\mathcal{S}(\mathbf{R}^n)$. In particular, this holds when p is a polynomial on $\mathbf{R}^n$.

9.8 Tempered distributions

Let n be a positive integer, and let λ be a linear functional on $\mathcal{S}(\mathbf{R}^n)$. Let us say that λ is continuous on $\mathcal{S}(\mathbf{R}^n)$ if for every sequence $\{\phi_j\}_{j=1}^\infty$ of elements of $\mathcal{S}(\mathbf{R}^n)$ that converges to an element ϕ of $\mathcal{S}(\mathbf{R}^n)$, in the sense described in the previous section, we have that

$$\lim_{j\to\infty} \lambda(\phi_j) = \lambda(\phi). \tag{9.8.1}$$

This is equivalent to the continuity of ϕ with respect to the standard topology on $\mathcal{S}(\mathbf{R}^n)$, which was mentioned in the previous section. Under these conditions, λ is said to be a *tempered distribution* on $\mathbf{R}^n$. The space

$$\mathcal{S}(\mathbf{R}^n)' \tag{9.8.2}$$

of tempered distributions on $\mathbf{R}^n$ is a vector space over the complex numbers, with respect to pointwise addition and scalar multiplication of linear functionals on $\mathcal{S}(\mathbf{R}^n)$.

Let f be a continuous complex-valued function on $\mathbf{R}^n$, and suppose that f grows at most polynomially on $\mathbf{R}^n$, so that

$$|f(x)| \le C\,(1 + |x|^2)^k \tag{9.8.3}$$

for some nonnegative real number C, nonnegative integer k, and every $x \in \mathbf{R}^n$. If $\phi \in \mathcal{S}(\mathbf{R}^n)$, then put

$$\lambda_f(\phi) = \int_{\mathbf{R}^n} f(x)\,\phi(x)\,dx, \tag{9.8.4}$$

where the right side may be defined as in Section 7.2. One can check that this defines a tempered distribution on $\mathbf{R}^n$. More precisely, this works when

$$f(x)\,(1 + |x|^2)^{-l} \tag{9.8.5}$$

is integrable on $\mathbf{R}^n$ for some nonnegative integer l. This also works when f is a locally integrable function on $\mathbf{R}^n$ with respect to Lebesgue measure such that (9.8.5) is integrable, in which case (9.8.4) should be interpreted as a Lebesgue integral.

One can define derivatives of tempered distributions in the same way as in Section 9.4. Let p be a smooth complex-valued function on $\mathbf{R}^n$ whose derivatives of all orders grow at most polynomially, as in the previous section. If λ is any tempered distribution on $\mathbf{R}^n$, then

$$(p\,\lambda)(\phi) = \lambda(p\,\phi) \tag{9.8.6}$$

defines another tempered distribution on $\mathbf{R}^n$. If f is as in the preceding paragraph, then $p\,f$ satisfies an analogous condition, so that $\lambda_{p\,f}$ is defined as a tempered distribution on $\mathbf{R}^n$ too. Of course,

$$p\,\lambda_f = \lambda_{p\,f}, \tag{9.8.7}$$

as tempered distributions on $\mathbf{R}^n$.

If λ is a tempered distribution on $\mathbf{R}^n$, then it is easy to see that

(9.8.8) the restriction of λ to $C^\infty_{com}(\mathbf{R}^n, \mathbf{C})$ defines a distribution on $\mathbf{R}^n$.

It is well known that

(9.8.9) λ is uniquely determined by its restriction to $C^\infty_{com}(\mathbf{R}^n, \mathbf{C})$.

More precisely,

$$C^\infty_{com}(\mathbf{R}^n, \mathbf{C}) \text{ is dense in } \mathcal{S}(\mathbf{R}^n), \tag{9.8.10}$$

with respect to the standard topology on $\mathbf{S}(\mathbf{R}^n)$. Equivalently, this means that if $\phi \in \mathcal{S}(\mathbf{R}^n)$, then there is a sequence $\{\phi_j\}_{j=1}^\infty$ of elements of $C^\infty_{com}(\mathbf{R}^n)$ that converges to ϕ in $\mathcal{S}(\mathbf{R}^n)$. The ϕ_j's can be obtained by multiplying ϕ by suitable smooth functions on $\mathbf{R}^n$ with compact support, which are equal to 1 on large bounded subsets of $\mathbf{R}^n$.

Of course, if $x \in \mathbf{R}^n$, then $\delta_x(\phi) = \phi(x)$ defines a tempered distribution on $\mathbf{R}^n$, which is another version of the Dirac distribution associated to x. See [85, 121, 125] for more information about the Schwartz class and tempered distributions, in addition the references about distributions mentioned in Section 9.1.

9.9 More on $\mathcal{S}(\mathbf{R}^n)$, $\mathcal{S}(\mathbf{R}^n)'$

Let n be a positive integer, let f be an element of the Schwartz class $\mathcal{S}(\mathbf{R}^n)$, and let $c \in \mathbf{R}^n$ be given. One can check that $f(x + c) \in \mathcal{S}(\mathbf{R}^n)$, as a function of $x \in \mathbf{R}^n$, as mentioned in Section 9.7. More precisely, if α and β are multi-indices, then the boundedness of

$$x^\alpha\, (\partial^\beta f)(x + c) \tag{9.9.1}$$

on $\mathbf{R}^n$ is equivalent to the boundedness of

$$(x - c)^\alpha\, (\partial^\beta f)(x) \tag{9.9.2}$$

on $\mathbf{R}^n$. This can be obtained from the boundedness of

$$x^\gamma\, (\partial^\beta f)(x) \tag{9.9.3}$$

on $\mathbf{R}^n$, for multi-indices γ with $\gamma_j \le \alpha_j$ for each $j = 1, \ldots, n$. This argument also shows that one can get a bound for the absolute value of (9.9.1) on $\mathbf{R}^n$ that grows at most polynomially in $|c|$.

Let λ be a linear functional on $\mathcal{S}(\mathbf{R}^n)$. Suppose that there are a nonnegative real number C and nonnegative integers N_1, N_2 such that

$$|\lambda(\phi)| \le C \sum_{|\alpha| \le N_1} \sum_{|\beta| \le N_2} \Big(\sup_{x \in \mathbf{R}^n} |x^\alpha\, (\partial^\beta \phi)(x)| \Big) \tag{9.9.4}$$

for every $\phi \in \mathcal{S}(\mathbf{R}^n)$. Here the first sum is taken over all multi-indices α with $|\alpha| \le N_1$, and the second sum is taken over all multi-indices β with $|\beta| \le N_2$, as usual. Under these conditions, one can check that λ is a tempered distribution on $\mathbf{R}^n$.

In fact, let $\{\phi_j\}_{j=1}^\infty$ be a sequence of elements of $\mathcal{S}(\mathbf{R}^n)$, and let ϕ be another element of $\mathcal{S}(\mathbf{R}^n)$. Suppose that

$$x^\alpha\, (\partial^\beta \phi)(x) \to x^\alpha\, (\partial^\beta \phi)(x) \text{ as } j \to \infty, \tag{9.9.5}$$

uniformly on $\mathbf{R}^n$, for all multi-indices α, β with $|\alpha| \le N_1$ and $|\beta| \le N_2$. If (9.9.4) holds, then it is easy to see that $\lambda(\phi_j) \to \phi$ as $j \to \infty$.

Conversely, if λ is a tempered distribution on $\mathbf{R}^n$, then it is well known that (9.9.4) holds for some $C, N_1, N_2 \ge 0$. This can be obtained from the continuity of λ at 0, with respect to the standard topology on $\mathcal{S}(\mathbf{R}^n)$.

Let f be a continuous complex-valued function on $\mathbf{R}^n$, and let $\{c_j\}_{j=1}^\infty$ be a sequence of elements of $\mathbf{R}^n$ that converges to 0. Put

$$f_j(x) = f(x + c_j) \tag{9.9.6}$$

for each $x \in \mathbf{R}^n$ and $j \ge 1$. Note that

$$f_j \to f \text{ as } j \to \infty \tag{9.9.7}$$

pointwise on $\mathbf{R}^n$, because f is continuous on $\mathbf{R}^n$. More precisely, one can check that (9.9.7) holds uniformly on compact subsets of $\mathbf{R}^n$, because continuous functions are uniformly continuous on compact sets. If f is uniformly continuous on $\mathbf{R}^n$, then (9.9.7) holds uniformly on $\mathbf{R}^n$.

If f is smooth on $\mathbf{R}^n$, then for each multi-index α,

$$\partial^\alpha f_j \to \partial^\alpha f \text{ as } j \to \infty, \tag{9.9.8}$$

uniformly on compact subsets of $\mathbf{R}^n$. If f has compact support in $\mathbf{R}^n$, then one can verify that there is a compact subset of $\mathbf{R}^n$ that contains the supports of f and f_j for each j.

If $f \in \mathcal{S}(\mathbf{R}^n)$, then $f_j \in \mathcal{S}(\mathbf{R}^n)$ for each j, as before. In this case, it is not too difficult to show that (9.9.7) holds in $\mathcal{S}(\mathbf{R}^n)$, in the sense defined in Section 9.7.

9.10 Some convolutions

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. If $a \in \mathbf{R}^n$ and $E \subseteq \mathbf{R}^n$, then put

$$a + E = \{a + x : x \in E\} \tag{9.10.1}$$

and

$$-E = \{-x : x \in E\}. \tag{9.10.2}$$

Similarly, we put $a - E = a + (-E)$.

Let K be a nonempty compact subset of $\mathbf{R}^n$, and put

$$V = \{a \in \mathbf{R}^n : a - K \subseteq U\}. \tag{9.10.3}$$

One can check that this is an open subset of $\mathbf{R}^n$, because U is an open set.

If ϕ is a complex-valued function on $\mathbf{R}^n$, then put

$$\widetilde{\phi}(y) = \phi(-y) \tag{9.10.4}$$

for every $y \in \mathbf{R}^n$. Also let $\tau_a(\phi)$ be the complex-valued function on $\mathbf{R}^n$ defined by

$$(\tau_a(\phi))(y) = \phi(y - a) \tag{9.10.5}$$

for every $y \in \mathbf{R}^n$. Thus

$$(\tau_a(\widetilde{\phi}))(y) = \phi(a - y) \tag{9.10.6}$$

for every $y \in \mathbf{R}^n$.

Suppose now that ϕ is smooth on $\mathbf{R}^n$, with

$$\operatorname{supp} \phi \subseteq K, \tag{9.10.7}$$

and let λ be a distribution on U. Observe that

$$\operatorname{supp} \tau_a(\widetilde{\phi}) = a - \operatorname{supp} \phi \tag{9.10.8}$$

for every $a \in \mathbf{R}^n$. In particular, if $a \in V$, then

$$\operatorname{supp} \tau_a(\widetilde{\phi}) \subseteq U. \tag{9.10.9}$$

Under these conditions, the *convolution* of λ and ϕ is the complex-valued function $\lambda * \phi$ defined on V by

$$(\lambda * \phi)(a) = \lambda(\tau_a(\widetilde{\phi})). \tag{9.10.10}$$

If $b \in \mathbf{R}^n$, then

$$(\delta_b * \phi)(a) = \phi(a - b) \tag{9.10.11}$$

for every $a \in \mathbf{R}^n$, by (9.10.6).

If $\{a_j\}_{j=1}^\infty$ is a sequence of elements of V that converges to $a \in V$, then one can check that

$$\lim_{j \to \infty} (\lambda * \phi)(a_j) = (\lambda * \phi)(a). \tag{9.10.12}$$

Equivalently, this means that

$$\lim_{j \to \infty} \lambda(\tau_{a_j}(\widetilde{\phi})) = \lambda(\tau_a(\widetilde{\phi})). \tag{9.10.13}$$

This implies that $\lambda * \phi$ is continuous on V.

It is well known and not too difficult to show that the first partial derivatives of $\lambda * \phi$ exist on V, with

$$\partial_l(\lambda * \phi) = \lambda * (\partial_l \phi) \tag{9.10.14}$$

for each $l = 1, \ldots, n$. One can use this repeatedly, to get that $\lambda * \phi$ is smooth on V, with

$$\partial^\alpha(\lambda * \phi) = \lambda * (\partial^\alpha \phi) \tag{9.10.15}$$

on V for each multi-index α.

One can verify that

$$\lambda * (\partial^\alpha \phi) = (\partial^\alpha \lambda) * \phi \tag{9.10.16}$$

on V for every multi-index α. It follows that

$$\partial^\alpha(\lambda * \phi) = (\partial^\alpha \lambda) * \phi \tag{9.10.17}$$

on V, by (9.10.15).

If $\phi \in \mathcal{S}(\mathbf{R}^n)$ and $\lambda \in \mathcal{S}(\mathbf{R}^n)'$, then $\lambda * \phi$ can be defined on $\mathbf{R}^n$ as in (9.10.10). It is well known that this satisfies the same type of properties as before.

In this case, one can also show that

(9.10.18) $\lambda * \phi$ grows at most polynomially on $\mathbf{R}^n$,

using the remarks in the previous section. More precisely,

(9.10.19) the derivatives of $\lambda * \phi$ of all orders grow at most polynomially

too.

9.11 Local solvability

Let n be a positive integer, and let $p(w)$ be a nonzero polynomial on $\mathbf{R}^n$ with complex coefficients. As in Section 9.1, a theorem of Ehrenpreis and Malgrange states that there is a distribution E on $\mathbf{R}^n$ that is a fundamental solution of $p(\partial)$, in the sense that

$$p(\partial)(E) = \delta_0. \tag{9.11.1}$$

Let $f \in C^\infty_{com}(\mathbf{R}^n, \mathbf{C})$ be given, and put

$$u = E * f, \tag{9.11.2}$$

which is a smooth complex-valued function on $\mathbf{R}^n$, as in the previous section. Under these conditions, we have that

$$(p(\partial))(u) = (p(\partial))(E) * f = \delta_0 * f = f \tag{9.11.3}$$

on $\mathbf{R}^n$, as mentioned earlier.

Let

$$L = \sum_{|\alpha| \le N} a_\alpha(x)\, \partial^\alpha \tag{9.11.4}$$

be a differential operator whose coefficients $a_\alpha(x)$ are smooth complex-valued functions on $\mathbf{R}^n$. Here N is a nonnegative integer, and the sum is taken over all multi-indices α with $|\alpha| \le N$, as usual. We say that L is *locally solvable* at a point $x_0 \in \mathbf{R}^n$ if for any smooth complex-valued function f on $\mathbf{R}^n$ there is a function (or distribution) u on a neighborhood of x_0 in $\mathbf{R}^n$ that satisfies

$$L(u) = f \tag{9.11.5}$$

on that neighborhood, as in Section F of Chapter 1 of [32]. We may as well take f to have compact support in $\mathbf{R}^n$, as in [32], since otherwise we can multiply

f by a smooth function on $\mathbf{R}^n$ with compact support that is equal to 1 on a neighborhood of x_0. Similarly, we could start with any smooth complex-valued function f_0 defined on a neighborhood of x_0 in $\mathbf{R}^n$, and get a smooth function on $\mathbf{R}^n$ with compact support that is equal to f_0 on neighborhood of x_0.

If the coefficients of L are constants, not all equal to 0, then the theorem of Ehrenpreis and Malgrange implies that L is locally solvable at every point in $\mathbf{R}^n$. If

$$f \text{ and the coefficients } a_\alpha \text{ are real-analytic near } x_0, \tag{9.11.6}$$

and if

$$a_\alpha(x_0) \neq 0 \tag{9.11.7}$$

for some multi-index α with $|\alpha| = N$, then one can get real-analytic solutions to (9.11.5) near x_0 using a famous theorem of Cauchy and Kovalevskaya, as mentioned near the beginning of Section E of Chapter 1 of [32]. This theorem is discussed in Section 4.6.3 of [29], Section D of Chapter 1 of [32], and Section 2.8 of [83].

There is a famous example of H. Lewy of a first-order differential operator on $\mathbf{R}^3$ whose coefficients are constants or linear functions, and for which local solvability does not hold. See Section E of Chapter 1 of [32] for more information.

9.12 Sequences of distributions

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. A sequence $\{\lambda_j\}_{j=1}^\infty$ of distributions on U is said to *converge* to a distribution λ on U if

$$\lim_{j \to \infty} \lambda_j(\phi) = \lambda(\phi) \tag{9.12.1}$$

for every $\phi \in C^\infty_{com}(U, \mathbf{C})$. More precisely, this is the same as convergence with respect to the "weak* topology" on $C^\infty_{com}(\mathbf{R}^n, \mathbf{C})'$.

In this case, it is easy to see that

$$\{\partial^\alpha \lambda_j\}_{j=1}^\infty \text{ converges to } \partial^\alpha \lambda \tag{9.12.2}$$

in the same sense for every multi-index α. Similarly, if a is a smooth complex-valued function on U, then

$$\{a\, \lambda_j\}_{j=1}^\infty \text{ converges to } a\, \lambda \tag{9.12.3}$$

in this sense.

Let $\{f_j\}_{j=1}^\infty$ be a sequence of continuous complex-valued functions on U that converges to a complex-valued function f uniformly on compact sets contained in U. Under these conditions, it is easy to see that the corresponding sequence of distributions $\{\lambda_{f_j}\}_{j=1}^\infty$, as in Section 9.3, converges to the distribution λ_f corresponding to f, in the sense considered here. This also works when $\{f_j\}_{j=1}^\infty$ is a sequence of locally-integrable functions on U that converges to a locally-integrable function f on U with respect to the L^1 metric on any compact set contained in U.

Let $\{\lambda_j\}_{j=1}^\infty$ be a sequence of distributions on U again, and suppose for the moment that for each $\phi \in C^\infty_{com}(U, \mathbf{C})$,

$$\{\lambda_j(\phi)\}_{j=1}^\infty \text{ is a bounded sequence in } \mathbf{C}. \tag{9.12.4}$$

Let K be a nonempty compact subset of $\mathbf{R}^n$ that is contained in U. A famous theorem of Banach and Steinhaus implies that there are a nonnegative real number $C(K)$ and a nonnegative integer $N(K)$ such that

$$|\lambda_j(\phi)| \le C(K) \sum_{|\alpha| \le N(K)} \Big(\max_{x \in K} |(\partial^\alpha \phi)(x)| \Big) \tag{9.12.5}$$

for each $j \ge 1$ and $\phi \in C^\infty_K(\mathbf{R}^n, \mathbf{C})$. This uses the well-known fact that $C^\infty_K(\mathbf{R}^n, \mathbf{C})$ is a "Fréchet space".

Suppose now that for each $\phi \in C^\infty_{com}(U, \mathbf{C})$,

$$\{\lambda_j(\phi)\}_{j=1}^\infty \text{ converges in } \mathbf{C}. \tag{9.12.6}$$

This implies (9.12.4), because convergent sequences are bounded. If λ is defined on $C^\infty_{com}(U, \mathbf{C})$ as in (9.12.1), then it is easy to see that λ is a linear functional on $C^\infty_{com}(U, \mathbf{C})$. We also have that

$$|\lambda(\phi)| \le C(K) \sum_{|\alpha| \le N(K)} \Big(\max_{x \in K} |(\partial^\alpha \phi)(x)| \Big) \tag{9.12.7}$$

for every nonempty compact set K contained in U and $\phi \in C^\infty_K(\mathbf{R}^n, \mathbf{C})$, where $C(K)$ and $N(K)$ are as in the preceding paragraph. This implies that

$$\lambda \text{ is a distribution on } U, \tag{9.12.8}$$

as in Section 9.6. Thus $\{\lambda_j\}_{j=1}^\infty$ converges to λ in the sense described at the beginning of the section. This corresponds to Theorem 6.17 on p146 of [115].

Similarly, a sequence $\{\lambda_j\}_{j=1}^\infty$ of tempered distributions on $\mathbf{R}^n$ is said to *converge* to a tempered distribution λ on $\mathbf{R}^n$ if (9.12.1) holds for every ϕ in $\mathcal{S}(\mathbf{R}^n)$. This is the same as convergence with respect to the weak* topology on $\mathcal{S}(\mathbf{R}^n)'$, as before.

If α is a multi-index, then it follows that $\{\partial^\alpha \lambda_j\}_{j=1}^n$ converges to $\partial^\alpha \lambda$ in the same sense. If a is a smooth complex-valued function on $\mathbf{R}^n$ such that a and all of its derivatives grow at most polynomially on $\mathbf{R}^n$, then $\{a\, \lambda_j\}_{j=1}^\infty$ converges to $a\, \lambda$ in this sense too.

Let $\{f_j\}_{j=1}^\infty$ be a sequence of complex-valued functions on $\mathbf{R}^n$, let f be another complex-valued function on $\mathbf{R}^n$, and let l be a nonnegative integer. Suppose that the f_j's and f are continuous on $\mathbf{R}^n$, or at least locally integrable, and that the products $f_j(x)\,(1 + |x|^2)^{-l}$ and $f(x)\,(1 + |x|^2)^{-l}$ are integrable on $\mathbf{R}^n$. Thus we get tempered distributions λ_{f_j}, λ_f on $\mathbf{R}^n$, as in Section 9.8. If

$$\lim_{j \to \infty} \int_{\mathbf{R}^n} |f_j(x) - f(x)|\,(1 + |x|^2)^{-l}\, dx = 0, \tag{9.12.9}$$

then it is easy to see that $\{\lambda_{f_j}\}_{j=1}^\infty$ converges to λ_f, as tempered distributions on $\mathbf{R}^n$.

Suppose that $\{\lambda_j\}_{j=1}^\infty$ is a sequence of tempered distributions on $\mathbf{R}^n$ that satisfies (9.12.4) for every $\phi \in \mathcal{S}(\mathbf{R}^n)$. One can use the Banach–Steinhaus theorem to get that there are a nonnegative real number C and nonnegative integers N_1, N_2 such that

$$|\lambda_j(\phi)| \le C \sum_{|\alpha| \le N_1} \sum_{|\beta| \le N_2} \Big(\sup_{x \in \mathbf{R}^n} |x^\alpha \, (\partial^\beta \phi)(x)| \Big) \tag{9.12.10}$$

for each $j \ge 1$ and $\phi \in \mathcal{S}(\mathbf{R}^n)$. This uses the fact that $\mathcal{S}(\mathbf{R}^n)$ is a Fréchet space too.

Suppose that (9.12.6) holds for every $\phi \in \mathcal{S}(\mathbf{R}^n)$, so that (9.12.4) holds in particular, as before. This permits us to define λ as a linear functional on $\mathcal{S}(\mathbf{R}^n)$ by (9.12.1). Note that

$$|\lambda(\phi)| \le C \sum_{|\alpha| \le N_1} \sum_{|\beta| \le N_2} \Big(\sup_{x \in \mathbf{R}^n} |x^\alpha \, (\partial^\beta \phi)(x)| \Big) \tag{9.12.11}$$

for every $\phi \in \mathcal{S}(\mathbf{R}^n)$, by (9.12.10). This implies that λ is a tempered distribution on $\mathbf{R}^n$, as in Section 9.9.

Chapter 10

Vector-valued functions and systems

10.1 Vector-valued functions

Let n and l be positive integers, and let U be a nonempty open subset of $\mathbf{R}^n$. If $f_1, \ldots, f_l$ are real or complex-valued functions on U, then

$$f(x) = (f_1(x), \ldots, f_l(x)) \tag{10.1.1}$$

defines a mapping from U into $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate. The continuity of f on U can be defined in the usual way, using the standard Euclidean metrics on $\mathbf{R}^n$ and on $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate. It is well known and not difficult to show that this is equivalent to the continuity of $f_1, \ldots, f_l$ as real or complex-valued functions on U, as appropriate.

The spaces of continuous functions on U with values in $\mathbf{R}^l$ and $\mathbf{C}^l$ may be denoted $C(U, \mathbf{R}^l)$ and $C(U, \mathbf{C}^l)$, respectively. These are vector spaces over the real and complex numbers, respectively, with respect to pointwise addition and scalar multiplication of functions. These spaces may be identified with the spaces $C(U, \mathbf{R})^l$ and $C(U, \mathbf{C})^l$ of l-tuples of elements of $C(U, \mathbf{R})$ and $C(U, \mathbf{C})$, respectively.

One can define partial derivatives of f, when they exist, in the usual way, using the standard Euclidean metric on $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate. This is equivalent to the existence of the corresponding partial derivative of f_j for each $j = 1, \ldots, l$, in which case the jth component of the partial derivative of f is equal to the corresponding partial derivative of f_j. Similarly, the continuous differentiability of f on U can be defined directly, and is equivalent to the continuous differentiability of f_j on U for each $j = 1, \ldots, l$. If k is any positive integer, then the k-times continuous differentiability of f on U can be defined directly as well, and is equivalent to the k-times continuous differentiability of f_j on U for each $j = 1, \ldots, l$. If f is k-times continuously differentiable on U

for every $k \geq 1$, then f is said to be infinitely differentiable or smooth on U, as before.

Let $C^k(U, \mathbf{R}^l)$ and $C^k(U, \mathbf{C}^l)$ be the spaces of k-times continuously differentiable functions on U with values in $\mathbf{R}^l$ and $\mathbf{C}^l$, respectively, for each $k \geq 1$. We may use the same notation with $k = 0$ for the corresponding spaces of continuous functions, as before. Note that $C^k(U, \mathbf{R}^l)$ and $C^k(U, \mathbf{C}^l)$ are linear subspaces of $C(U, \mathbf{R}^l)$ and $C(U, \mathbf{C}^l)$, as vector spaces over the real and complex numbers, for each k. We may identify $C^k(U, \mathbf{R}^l)$ and $C^k(U, \mathbf{C}^l)$ with the spaces $C^k(U, \mathbf{R})^l$ and $C^k(U, \mathbf{C})^l$ of l-tuples of elements of $C^k(U, \mathbf{R})$ and $C^k(U, \mathbf{C})$, respectively, as usual.

Similarly, $C^\infty(U, \mathbf{R}^l)$ and $C^\infty(U, \mathbf{C}^l)$ denote the spaces of smooth functions on U with values in $\mathbf{R}^l$ and $\mathbf{C}^l$, respectively. These are linear subspaces of $C^k(U, \mathbf{R}^l)$ and $C^k(U, \mathbf{C}^l)$, respectively, for each k. We may identify $C^\infty(U, \mathbf{R}^l)$ and $C^\infty(U, \mathbf{C}^l)$ with the spaces $C^\infty(U, \mathbf{R})^l$ and $C^\infty(U, \mathbf{C})^l$ of l-tuples of elements of $C^\infty(U, \mathbf{R})$ and $C^\infty(U, \mathbf{C})$, respectively, as before.

10.2 Matrix-valued functions

Let l_1, l_2 be positive integers, and let $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$, $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$ be the spaces of linear mappings from $\mathbf{R}^{l_1}$, $\mathbf{C}^{l_1}$ into $\mathbf{R}^{l_2}$, $\mathbf{C}^{l_2}$, respectively, as vector spaces over the real and complex numbers. Note that $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$ and $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$ are vector spaces over the real and complex numbers, respectively, with respect to pointwise addition and scalar multiplication of linear mappings. Of course, these linear mappings can be represented in terms of matrices of real or complex numbers, as appropriate, in the usual way. One can use this to identify $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$ and $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$ with $\mathbf{R}^{l_1 l_2}$ and $\mathbf{C}^{l_1 l_2}$, respectively.

Let n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Suppose that $a(x)$ is a function of $x \in U$ with values in $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$. This can be identified with a function on U with values in $\mathbf{R}^{l_1 l_2}$ or $\mathbf{C}^{l_1 l_2}$, as appropriate, as before. In particular, this can be used to define the usual continuity and differentiability properties of $a(x)$ on U, using the standard Euclidean metric on $\mathbf{R}^{l_1 l_2}$ or $\mathbf{C}^{l_1 l_2}$, as appropriate. This is equivalent to the analogous continuity or differentiability properties of the $l_1 \, l_2$ real or complex-valued functions on U corresponding to the matrix entries of $a(x)$.

A version of this was mentioned in Section 5.15, for functions defined on an interval in the real line. Similarly, if $v \in \mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate, then

$$(a(x))(v) \tag{10.2.1}$$

defines a function of $x \in U$ with values in $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate. Continuity or differentiability properties of $a(x)$ on U are also equivalent to the analogous properties of (10.2.1) holding for every $v \in \mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate, as a function of $x \in U$ with values in $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate.

Suppose that $v(x)$ is a function on U with values in $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate, so that

$$(a(x))(v(x)) \tag{10.2.2}$$

is a function on U with values in $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate. If $a(x)$and $v(x)$ satisfy suitable continuity or differentiability properties on U, then (10.2.2) satisfies the same property on U, as in Section 5.15. In particular,

$$\frac{\partial}{\partial x_j}((a(x))(v(x))) = \Big(\frac{\partial a}{\partial x_j}(x)\Big)(v(x)) + (a(x))\Big(\frac{\partial v}{\partial x_j}(x)\Big) \tag{10.2.3}$$

when the partial derivatives of $a(x)$ and $v(x)$ exist, as before.

Let l_0 be another positive integer, and let $b(x)$ be a function of $x \in U$ with values in $\mathcal{L}(\mathbf{R}^{l_0}, \mathbf{R}^{l_1})$ or $\mathcal{L}(\mathbf{C}^{l_0}, \mathbf{C}^{l_1})$, as appropriate. If $x \in U$, then let

$$a(x)\, b(x) \tag{10.2.4}$$

be the composition of $b(x)$ with $a(x)$ as linear mappings, which defines an element of $\mathcal{L}(\mathbf{R}^{l_0}, \mathbf{R}_{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_0}, \mathbf{C}^{l_2})$, as appropriate. Of course, this corresponds to multiplication of the matrices associated to $b(x)$ and $a(x)$. If $a(x)$ and $b(x)$ satisfy suitable continuity or differentiability properties on U, then (10.2.4) satisfies the same property on U, as before. In particular,

$$\frac{\partial}{\partial x_j}(a(x)\, b(x)) = \Big(\frac{\partial a}{\partial x_j}(x)\Big)\, b(x) + a(x)\, \Big(\frac{\partial b}{\partial x_j}(x)\Big), \tag{10.2.5}$$

when the partial derivatives of $a(x)$ and $b(x)$ in x_j exist.

10.3 Matrix-valued coefficients

Let l_1, l_2, and n be a positive integer, and let U be a nonempty open subset of $\mathbf{R}^n$. Also let N be a nonnegative integer, and for each multi-index α with order $|\alpha| \le N$, let a_α be a function on U with values in $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$. If u is an N-times continuously-differentiable function on U with values in $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate, then let $L(u)$ be the function on U with values in $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate, defined by

$$(L(u))(x) = \sum_{|\alpha| \le N} (a_\alpha(x))((\partial^\alpha u)(x)) \tag{10.3.1}$$

for every $x \in U$. More precisely, if $x \in U$ and α is a multi-index with $|\alpha| \le N$, then $(\partial^\alpha u)(x)$ is an element of $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, $a_\alpha(x)$ is an element of $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$, and $(a_\alpha(x))((\partial^\alpha u)(x))$ is an element of $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate.

Suppose that a_α is r-times continuously differentiable on U for some nonnegative integer r, and each multi-index α with $|\alpha| \le N$. If u is $(N+r)$-times continuously differentiable on on U, then $L(u)$ is r-times continuously differentiable on U, as in Section 2.4. Under these conditions, L defines a linear mapping from $C^{N+r}(U, \mathbf{R}^{l_1})$ into $C^r(U, \mathbf{R}^{l_2})$, or from $C^{N+r}(U, \mathbf{C}^{l_1})$ into $C^r(U, \mathbf{C}^{l_2})$, as apprropiate.

Similarly, if a_α is smooth on U for every multi-index α with $|\alpha| \le N$, and u is smooth on U, then $L(u)$ is smooth on U as well. In this case, L defines a linear mapping from $C^\infty(U, \mathbf{R}^{l_1})$ into $C^\infty(U, \mathbf{R}^{l_2})$, or from $C^\infty(U, \mathbf{C}^{l_1})$ into $C^\infty(U, \mathbf{C}^{l_2})$, as appropriate.

Polynomials on $\mathbf{R}^n$ with values in $\mathbf{R}^l$ or $\mathbf{C}^l$ for some positive integer l will be discussed in the next section. One can check that the a_α's are uniquely determined by $L(u)$ for polynomials u with values in $\mathbf{R}^{l_1}$ of degree less than or equal to N, as in Section 2.4.

Let l_0 be another positive integer, and let $\widetilde{N}$ be another nonnegative integer. Suppose that for each multi-index β with $|\beta| \le \widetilde{N}$, b_β is a function on U with values in $\mathcal{L}(\mathbf{R}^{l_0}, \mathbf{R}^{l_1})$ or $\mathcal{L}(\mathbf{C}^{l_0}, \mathbf{C}^{l_1})$. If u is an $\widetilde{N}$-times continuously-differentiable function on U with values in $\mathbf{R}^{l_0}$ or $\mathbf{C}^{l_0}$, as appropriate, then

$$(\widetilde{L}(u))(x) = \sum_{|\beta| \le \widetilde{N}} (b_\beta(x))((\partial^\beta u)(x)) \tag{10.3.2}$$

defines a function on U with values in $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate. If b_β is N-times continuously-differentiable on U for every multi-index β with $|\beta| \le \widetilde{N}$, and u is $(N + \widetilde{N})$-times continuously differentiable on U, then $\widetilde{L}(u)$ is N-times continuously differentiable on U. This implies that

$$L(\widetilde{L}(u)) \tag{10.3.3}$$

is defined as a function on U with values in $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate.

As in Section 2.4, (10.3.3) may be expressed as

$$(\widehat{L}(u))(x) = \sum_{|\gamma| \le N + \widetilde{N}} (c_\gamma(x))((\partial^\gamma u)(x)). \tag{10.3.4}$$

Here c_γ is a function on U with values in $\mathcal{L}(\mathbf{R}^{l_0}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_0}, \mathbf{C}^{l_2})$ for each multi-index γ with $|\gamma| \le N + \widetilde{N}$. These functions can be expressed as sums of products of the a_α's with the b_β's and their derivatives of order less than or equal to N, as before. More precisely, these products correspond to compositions of linear mappings from $\mathbf{R}^{l_0}$ or $\mathbf{C}^{l_0}$ into $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$ with linear mappings from $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$ into $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$ to get linear mappings from $\mathbf{R}^{l_0}$ or $\mathbf{C}^{l_0}$ into $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate.

If a_α is r-times continuously differentiable on U for some $r \ge 0$ and every α with $|\alpha| \le N$, and if b_β is $(N + r)$-times continuously differentiable on U for every β with $|\beta| \le \widetilde{N}$, then c_γ is r-times continuously differentiable on U for every γ with $|\gamma| \le N + \widetilde{N}$. If u is also $(N + \widetilde{N} + r)$-times continuously differentiable on U, then $\widetilde{L}(u)$ is $(N + r)$-times continuously differentiable on U, and $\widehat{L}(u)$ is r-times continuously differentiable on U, as before. In particular, if the a_α's and b_β's are smooth on U, then the c_γ's are smooth on U. In this case, if u is smooth on U, then $\widetilde{L}(u)$ and $\widehat{L}(u)$ are smooth on U as well.

10.4 Vector-valued polynomials

Let n and l be positive integers again, and let $\mathcal{P}(\mathbf{R}^n, \mathbf{R}^l)$ and $\mathcal{P}(\mathbf{R}^n, \mathbf{C}^l)$ be the spaces of polynomials on $\mathbf{R}^n$ with coefficients in $\mathbf{R}^l$ and $\mathbf{C}^l$, respectively. These spaces can be identified with the spaces $\mathcal{P}(\mathbf{R}^n, \mathbf{R})^l$ and $\mathcal{P}(\mathbf{R}^n, \mathbf{C})^l$ of l-tuples of polynomials on $\mathbf{R}^n$ with real or complex coefficients, as appropriate. These are also linear subspaces of $C^\infty(\mathbf{R}^n, \mathbf{R}^l)$ and $C^\infty(\mathbf{R}^n, \mathbf{C}^l)$, respectively, as vector spaces over the real or complex numbers, as appropriate.

If k is a nonnegative integer, then let $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ and $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$ be the spaces of polynomials on $\mathbf{R}^n$ with coefficients in $\mathbf{R}^l$ and $\mathbf{C}^l$, respectively, and degree less than or equal to k. These are linear subspaces of $\mathcal{P}(\mathbf{R}^n, \mathbf{R}^l)$ and $\mathcal{P}(\mathbf{R}^n, \mathbf{C}^l)$, as vector spaces over the real or complex numbers, as appropriate. We can identify $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ and $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$ with the spaces $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})^l$ and $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})^l$ of l-tuples of polynomials on $\mathbf{R}^n$ with real and complex coefficients, respectively, of degree less than or equal to k.

Note that $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ and $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$ have the same finite dimension, as vector spaces over $\mathbf{R}$ and $\mathbf{C}$, respectively. This is equal to l times the dimension of $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R})$ and $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C})$, which is the same as the number of multi-indices β with order $|\beta| \le k$, as in Section 5.11.

Let l_1, l_2 be positive integers, and let

$$\mathcal{P}(\mathbf{R}^n, \mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})), \quad \mathcal{P}(\mathbf{R}^n, \mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})) \tag{10.4.1}$$

be the spaces of polynomials on $\mathbf{R}^n$ with coefficients in $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$, $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$, respectively. These spaces may be identified with $\mathcal{P}(\mathbf{R}^n, \mathbf{R}^{l_1\,l_2})$, $\mathcal{P}(\mathbf{R}^n, \mathbf{C}^{l_1\,l_2})$, respectively, as in Section 10.2.

If k is a nonnegative integer, then let

$$\mathcal{P}^k(\mathbf{R}^n, \mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})), \quad \mathcal{P}^k(\mathbf{R}^n, \mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})) \tag{10.4.2}$$

be the spaces of polynomials on $\mathbf{R}^n$ with coefficients in $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$, $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$, respectively, and degree less than or equal to k. These may be identified with the spaces $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^{l_1\,l_2})$, $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^{l_1\,l_2})$, respectively, as before.

Suppose that $a(x)$ be a polynomial on $\mathbf{R}^n$ with coefficients in $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$, and let $v(x)$ be a polynomial on $\mathbf{R}^n$ with coefficients in $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate. Observe that $a(x)(v(x))$ is a polynomial on $\mathbf{R}^n$ with coefficients in $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate, and that

$$\deg(a(x)(v(x))) \le \deg a(x) + \deg v(x). \tag{10.4.3}$$

Let l_0 be another positive integer, and let $b(x)$ be a polynomial on $\mathbf{R}^n$ with coefficients in $\mathcal{L}(\mathbf{R}^{l_0}, \mathbf{R}^{l_1})$ or $\mathcal{L}(\mathbf{C}^{l_0}, \mathbf{C}^{l_1})$, as appropriate. The product $a(x)\,b(x)$ may be defined as in Section 10.2, and is a polynomial on $\mathbf{R}^n$ with coefficients in $\mathcal{L}(\mathbf{R}^{l_0}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_0}.\mathbf{C}^{l_2})$, as appropriate. We also have that

$$\deg(a(x)\,b(x)) \le \deg a(x) + \deg b(x). \tag{10.4.4}$$

Let $\mathcal{L}(\mathbf{R}^l) = \mathcal{L}(\mathbf{R}^l, \mathbf{R}^l)$ and $\mathcal{L}(\mathbf{C}^l) = \mathcal{L}(\mathbf{C}^l, \mathbf{C}^l)$ be the spaces of linear mappings from $\mathbf{R}^l$ and $\mathbf{C}^l$ into themselves, respectively, as in Section 5.15. The spaces of polynomials on $\mathbf{R}^n$ with coefficients in $\mathcal{L}(\mathbf{R}^l)$ and $\mathcal{L}(\mathbf{C}^l)$ may be denoted $\mathcal{P}(\mathbf{R}^n, \mathcal{L}(\mathbf{R}^l))$ and $\mathcal{P}(\mathbf{R}^n, \mathcal{L}(\mathbf{C}^l))$, respectively. Similarly, the spaces of polynomials on $\mathbf{R}^n$ with coefficients in $\mathcal{L}(\mathbf{R}^l)$ and $\mathcal{L}(\mathbf{C}^l)$ and degree less than or equal to k may be denoted $\mathcal{P}^k(\mathbf{R}^n, \mathcal{L}(\mathbf{R}^l))$ and $\mathcal{P}^k(\mathbf{R}^n, \mathcal{L}(\mathbf{C}^l))$, respectively.

10.5 Matrix-valued polynomials

Let n, l_1, and l_2 be positive integers, and let N be a nonnegative integer. Also let

$$p(w) = \sum_{|\alpha| \le N} a_\alpha \, w^\alpha \tag{10.5.1}$$

be a polynomial in the n variables $w_1, \ldots, w_n$ with coefficients in $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$ of degree less than or equal to N. Thus, for each multi-index α with order $|\alpha| \le N$, a_α is a linear mapping from $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$ into $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate.

Using p, we get a differential operator

$$p(\partial) = \sum_{|\alpha| \le N} a_\alpha \, \partial^\alpha, \tag{10.5.2}$$

as in Section 1.7. More precisely, this is a differential operator with coefficients in $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$, as appropriate, as in Section 10.3.

Let b be an element of $\mathbf{R}^n$ or $\mathbf{C}^n$, as appropriate, so that $\exp(b \cdot x)$ defines a smooth real or complex-valued function of $x \in \mathbf{R}^n$. If $v \in \mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate, then

$$(\exp(b \cdot x)) \, v \tag{10.5.3}$$

defines a smooth function of $x \in \mathbf{R}^n$ with values in $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate. It is easy to see that

$$(p(\partial))((\exp(b \cdot x)) \, v) = (\exp(b \cdot x)) \, (p(b))(v), \tag{10.5.4}$$

which is a function of $x \in \mathbf{R}^n$ with values in $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate. More precisely, $p(b)$ is defined as a linear mapping from $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$ into $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate, which sends v to an element of $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate. In particular,

$$(p(\partial))((\exp(b \cdot x)) \, v) = 0 \tag{10.5.5}$$

if and only if

$$(p(b))(v) = 0. \tag{10.5.6}$$

Let l_0 be another positive integer, let N_0 be another nonnegative integer, and let $p_0(w)$ be a polynomial in $w_1, \ldots, w_n$ with coefficients in $\mathcal{L}(\mathbf{R}^{l_0}, \mathbf{R}^{l_1})$ or $\mathcal{L}(\mathbf{C}^{l_0}, \mathbf{C}^{l_1})$, as appropriate. Thus $p_0(\partial)$ is a differential operator with coefficients in $\mathcal{L}(\mathbf{R}^{l_0}, \mathbf{R}^{l_1})$ or $\mathcal{L}(\mathbf{C}^{l_0}, \mathbf{C}^{l_1})$, as appropriate. The product $p(w) \, p_0(w)$ is a polynomial in $w_1, \ldots, w_n$ with coefficients in $\mathcal{L}(\mathbf{R}^{l_0}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_0}, \mathbf{C}^{l_2})$, as appropriate, of degree less than or equal to $N_0 + N$, as in the previous section. This leads to a differential operator $(p \, p_0)(\partial)$ with coefficients in $\mathcal{L}(\mathbf{R}^{l_0}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_0}, \mathbf{C}^{l_2})$, as appropriate. One can check that

$$(p \, p_0)(\partial) = p(\partial) \, p_0(\partial), \tag{10.5.7}$$

as in Section 1.7.

Let U be a nonempty open subset of $\mathbf{R}^n$, and let u be a function on U with values in $\mathbf{R}^{l_0}$ or $\mathbf{C}^{l_0}$, as appropriate, that is $(N_0 + N)$-times continuously differentiable on U. Thus $(p_0(\partial))(u)$ is a function on U with values in $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate, that is N-times continuously differentiable on U. Under these conditions, we have that

$$((p\,p_0)(\partial))(u) = (p(\partial))((p_0(\partial))(u)) \tag{10.5.8}$$

on U, as in (10.5.7).

Let l be a positive integer, and let us now take $l_1 = l_2 = l$. Let b be an element of $\mathbf{R}^n$ or $\mathbf{C}^n$ again, as appropriate, so that $p(b)$ is a linear mapping from $\mathbf{R}^l$ or $\mathbf{C}^l$ into itself, as appropriate. Suppose that v is an element of $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate, that is an eigenvector of $p(b)$ with eigenvalue λ in $\mathbf{R}$ or $\mathbf{C}$, as appropriate. This implies that

$$(p(\partial))((\exp(b \cdot x))\,v) = \lambda\,(\exp(b \cdot x))\,v, \tag{10.5.9}$$

as in (10.5.4).

Note that $\det p(w)$ is a polynomial in $w_1, \ldots, w_n$ of degree less than or equal to $N\,l$ with real or complex coefficients, as appropriate.

10.6 Polynomials, vectors, and operators

Let n, l_1, and l_2 be positive integers, and let N be a nonnegative integer. Suppose that for each multi-index α with $|\alpha| \le N$, a_α is a polynomial on $\mathbf{R}^n$ with coefficients in $\mathcal{L}(\mathbf{R}^{l_1}, \mathbf{R}^{l_2})$ or $\mathcal{L}(\mathbf{C}^{l_1}, \mathbf{C}^{l_2})$. This can be identified with a polynomial with coefficients in $\mathbf{R}^{l_1\,l_2}$ or $\mathbf{C}^{l_1\,l_2}$, as in Section 10.2.

Using the a_α's, we can define a differential operator L acting on N-times continuously differentiable functions u on $\mathbf{R}^n$ with values in $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate, as in (10.3.1). In this case, L maps polynomials on $\mathbf{R}^n$ with coefficients in $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$ to polynomials on $\mathbf{R}^n$ with coefficients in $\mathbf{R}^{l_2}$ or $\mathbf{C}^{l_2}$, as appropriate.

Let c be an integer, and suppose that

$$\deg a_\alpha \le |\alpha| - c \tag{10.6.1}$$

for each α, $|\alpha| \le N$, which is interpreted as meaning that $a_\alpha = 0$ when $|\alpha| < c$, as usual. If p is a polynomial on $\mathbf{R}^n$ with coefficients in $\mathbf{R}^{l_1}$ or $\mathbf{C}^{l_1}$, as appropriate, then

$$\deg L(p) \le \deg p - c, \tag{10.6.2}$$

which means that $L(p) = 0$ when $\deg p < c$, as before.

Suppose now that $l_1 = l_2 = l$, so that L maps $\mathcal{P}(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}(\mathbf{R}^n, \mathbf{C}^l)$ into itself, as appropriate. If p is a polynomial on $\mathbf{R}^n$ with coefficients in $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate, then

$$\deg L^j(p) \le \deg p - c\,j \tag{10.6.3}$$

for each $j \ge 1$, by (10.6.2). This means that $L^j(p) = 0$ when $\deg p < c\,j$, as before.

Suppose that $c \ge 0$, and let k be a nonnegative integer. Thus L maps $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$ into itself, as appropriate. Let L_k be the restriction of L to $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate. If $c \ge 1$ and

$$k < c\,j, \tag{10.6.4}$$

then

$$L_k^j = 0, \tag{10.6.5}$$

by (10.6.3). In particular, this means that L_k is nilpotent when $c \ge 1$.

Let $m(k)$ be the number of multi-indices β with order $|\beta| \le k$, so that $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ and $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$ have dimension $l\,m(k)$ as vector spaces over $\mathbf{R}$ and $\mathbf{C}$, respectively, as in Section 10.4. This permits us to identify L_k with a linear mapping from $\mathbf{R}^{l\,m(k)}$ or $\mathbf{C}^{l\,m(k)}$ into itself, as appropriate. If $t \in \mathbf{R}$, then we can define the exponential of $t\,L_k$ as a linear mapping on $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate, as in Sections 5.4 and 5.8.

Let q be a polynomial on $\mathbf{R}^n$ with coefficients in $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate, and of degree less than or equal to k. Note that

$$(\exp(t\,L_k))(q) \tag{10.6.6}$$

is another polynomial on $\mathbf{R}^n$ with coefficients in $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate, and degree less than or equal to k. The coefficients of this polynomial depend on t, and are smooth functions of t. This implies that

$$u(x,t) = ((\exp(t\,L_k))(q))(x) \tag{10.6.7}$$

is smooth as a function of $(x,t) \in \mathbf{R}^n \times \mathbf{R}$ with values in $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate.

If $c \ge 1$, then $\exp(t\,L_k)$ is a polynomial in t whose coefficients are linear mappings on $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate, as in Section 5.10. It follows that (10.6.7) is a polynomial in x and t with coefficients in $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate, in this case.

Of course,

$$u(x,0) = q(x) \tag{10.6.8}$$

for every $x \in \mathbf{R}^n$, and

$$\frac{\partial}{\partial t}((\exp(t\,L_k))(q)) = L_k((\exp(t\,L_k))(q)), \tag{10.6.9}$$

as in Sections 5.6 and 5.10. This implies that

$$\frac{\partial u}{\partial t} = L(u) \tag{10.6.10}$$

on $\mathbf{R}^n \times \mathbf{R}$, as in Section 5.12.

A standard approach to dealing with equations with higher-order derivatives in t is to reduce to the case of systems of equations with only first-order derivatives in t. A basic version of this was discussed in Section 5.13. That is much easier to do here, since we are already working with systems of equations.

10.7 Some more products with $\exp(b \cdot x)$

Let n and l be positive integers, also let N be a nonnegative integer. Also let $p(w)$ be a polynomial in the n variables $w_1, \ldots, w_n$ with coefficients in $\mathcal{L}(\mathbf{R}^l)$ or $\mathcal{L}(\mathbf{C}^l)$ of degree less than or equal to N, as in Section 10.5. This leads to a differential operator $p(\partial)$, as before.

Let $b \in \mathbf{R}^n$ or $\mathbf{C}^n$ be given, as appropriate. Observe that

$$p_b(w) = p(w + b) \tag{10.7.1}$$

can be expressed as a polynomial in $w_1, \ldots, w_n$ with coefficients in $\mathcal{L}(\mathbf{R}^l)$ or $\mathcal{L}(\mathbf{C}^l)$, as appropriate, of degree less than or equal to N, as in Section 2.5.

Let $p_b(\partial)$ be the differential operator associated to $p_b(w)$, and let f be an N-times continuously-differentiable function on $\mathbf{R}^n$ with values in $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate. Under these conditions,

$$p(\partial)((\exp(b \cdot x))\, f(x)) = (\exp(b \cdot x))\, (p_b(\partial)(f))(x), \tag{10.7.2}$$

as in Section 5.14.

If $b \in \mathbf{R}^n$, then let

$$(\exp(b \cdot x))\, \mathcal{P}(\mathbf{R}^n, \mathbf{R}^l) \tag{10.7.3}$$

be the space of functions on $\mathbf{R}^n$ with values in $\mathbf{R}^l$ of the form

$$(\exp(b \cdot x))\, q(x), \tag{10.7.4}$$

where $q \in \mathcal{P}(\mathbf{R}^n, \mathbf{R}^l)$. This is a linear subspace of $C^\infty(\mathbf{R}^n, \mathbf{R}^l)$, as a vector space over the real numbers. Similarly, if k is a nonnegative integer, then let

$$(\exp(b \cdot x))\, \mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l) \tag{10.7.5}$$

be the space of functions on $\mathbf{R}^n$ with values in $\mathbf{R}^l$ of the form (10.7.4), with $q \in \mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$. This is a linear subspace of (10.7.3), as a vector space over $\mathbf{R}$. If $p(w)$ has coefficients in $\mathcal{L}(\mathbf{R}^l)$, then $p(\partial)$ maps (10.7.3) and (10.7.5) into themselves, because of (10.7.2), as in Section 5.14.

If $b \in \mathbf{C}^n$, then let

$$(\exp(b \cdot x))\, \mathcal{P}(\mathbf{R}^n, \mathbf{C}^l) \tag{10.7.6}$$

be the space of functions on $\mathbf{R}^n$ with values in $\mathbf{C}^l$ of the form (10.7.4), with $q \in \mathcal{P}(\mathbf{R}^n, \mathbf{C}^l)$. Similarly, if k is a nonnegative integer, then let

$$(\exp(b \cdot x))\, \mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l) \tag{10.7.7}$$

be the space of functions on $\mathbf{R}^n$ with values in $\mathbf{C}^l$ of the form (10.7.4), with $q \in \mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$. These are linear subspaces of $C^\infty(\mathbf{R}^n, \mathbf{C}^l)$, as a vector space over the complex numbers. If $p(w)$ has coefficients in $\mathcal{L}(\mathbf{C}^l)$, then $p(\partial)$ maps (10.7.6) and (10.7.7) into themselves, because of (10.7.2), as before.

10.8 Some remarks about nilpotency

Let n, l, N, and $p(w)$ be as at the beginning of the previous section. Suppose for the moment that $p(0)$ is nilpotent, so that

$$p(0)^{r+1} = 0 \tag{10.8.1}$$

on $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate, for some nonnegative integer r. If k is a nonnegative integer, then it follows that $p(\partial)^{r+1}$ is nilpotent on $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate, as in Section 10.6. Of course this means that $p(\partial)$ is nilpotent on $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate.

Put

$$L = p(\partial), \tag{10.8.2}$$

and let L_k be the restriction of L to $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate, for each nonnegative integer k. If $t \in \mathbf{R}$, then we can define

$$\exp(t\, L_k) \tag{10.8.3}$$

as a linear mapping on $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate, as in Sections 5.4 and 5.8. If $p(0)$ is nilpotent, so that L_k is nilpotent, then (10.8.3) is a polynomial in t whose coefficients are linear mappings on $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate, as in Section 5.10.

Let q be an element of $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate. If $p(0)$ is nilpotent, then $(\exp(t\, L_k))(q)$ may be considered as a polynomial in t with coefficients in $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate. In particular,

$$((\exp(t\, L_k))(q))(x) \tag{10.8.4}$$

is a polynomial in x and t with coefficients in $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate.

Let $b \in \mathbf{R}^n$ or $\mathbf{C}^n$ be given, as appropriate, and let $p_b(w)$ be as in (10.7.1). Suppose now that

$$p_b(0) = p(b) \text{ is nilpotent} \tag{10.8.5}$$

on $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate. If k is a nonnegative integer, then the restriction of $p_b(\partial)$ to $\mathcal{P}^k(\mathbf{R}^n, \mathbf{R}^l)$ or $\mathcal{P}^k(\mathbf{R}^n, \mathbf{C}^l)$, as appropriate, is nilpotent, by the remarks at the beginning of the section. This implies that the restriction of $p(\partial)$ to (10.7.5) or (10.7.7), as appropriate, is nilpotent, because of (10.7.2).

10.9 The characteristic polynomial

Let l be a positive integer, and let A be a linear mapping from $\mathbf{R}^l$ or $\mathbf{C}^l$ into itself. If t is a real or complex number, then $A - t\, I$ is another linear mapping from $\mathbf{R}^l$ or $\mathbf{C}^l$ into itself, as appropriate, where I is the identity mapping on $\mathbf{R}^l$ or $\mathbf{C}^l$. Thus

$$\mathrm{ch}_A(t) = \det(A - t\, I) \tag{10.9.1}$$

defines a real or complex-valued function on $\mathbf{R}$ or $\mathbf{C}$, as appropriate. More precisely, $\mathrm{ch}_A(t)$ is a polynomial of degree l in t, with real or complex coefficients, as appropriate. This is known as the *characteristic polynomial* of A.

The characteristic polynomial may be expressed as

$$\mathrm{ch}_A(t) = \sum_{j=0}^{l} c_j\, t^j, \tag{10.9.2}$$

with $c_j \in \mathbf{R}$ or $\mathbf{C}$ for each j, as appropriate. Remember that the determinant of an $l \times l$ matrix is given by a homogeneous polynomial of degree l in the l^2 entries of the matrix. This means that c_j is given by a homogeneous polynomial of degree $l - j$ in the l^2 entries of the $l \times l$ matrix corresponding to A. In particular,

$$c_l = (-1)^l, \tag{10.9.3}$$

and $c_0 = \det A$. The zeros of $\mathrm{ch}_A(t)$ in $\mathbf{R}$ or $\mathbf{C}$ are the same as the eigenvalues of A as a linear mapping on $\mathbf{R}^l$ or $\mathbf{C}^l$, as appropriate, by standard arguments.

A polynomial of degree l in t with complex coefficients is equal to the product of the coefficient of t^l and l linear factors, corresponding to the l zeros of the polynomial in $\mathbf{C}$, with their appropriate multiplicities, by the fundamental theorem of algebra. It follows that $\mathrm{ch}_A(t)$ is uniquely determined by the eigenvalues of A in the complex case, because of (10.9.3).

If A is nilpotent, then it is easy to see that 0 is the only eigenvalue of A. This implies that

$$\mathrm{ch}_A(t) = (-1)^l\, t^l \tag{10.9.4}$$

in the complex case, by the remarks in the preceding paragraph. Equivalently, this means that

$$c_j = 0,\ 0 \le j \le l - 1. \tag{10.9.5}$$

One can get the same conclusion in the real case using the unique extension of a linear mapping from $\mathbf{R}^l$ into itself to a linear mapping from $\mathbf{C}^l$ into itself, as a vector space over the complex numbers. Another proof of this will be mentioned in the next section.

If A is any linear mapping from $\mathbf{R}^l$ or $\mathbf{C}^l$ into itself, then the Cayley–Hamilton theorem states that

$$\mathrm{ch}_A(A) = \sum_{j=0}^{l} c_j\, A^j = 0, \tag{10.9.6}$$

where A^0 is interpreted as being equal to I. If (10.9.4) holds, then it follows that

$$A^l = 0. \tag{10.9.7}$$

10.10 More on nilpotent linear mappings

Let l be a positive integer again, and let A be a linear mapping from $\mathbf{R}^l$ or $\mathbf{C}^l$ into itself. If τ is a real or complex number, then $I - \tau\, A$ defines a linear mapping from $\mathbf{R}^l$ or $\mathbf{C}^l$ into itself, as appropriate. Let r be a nonnegative integer, and observe that

$$(I - \tau\, A) \sum_{j=0}^{r} \tau^j\, A^j = \sum_{j=0}^{r} \tau^j\, A^j - \sum_{j=1}^{r+1} \tau^j\, A^j = I - \tau^{r+1}\, A^{r+1}. \tag{10.10.1}$$

Similarly,

$$\Big(\sum_{j=0}^{r} \tau^j\, A^j\Big)\,(I - \tau\, A) = I - \tau^{r+1}\, A^{r+1}. \tag{10.10.2}$$

If

$$A^{r+1} = 0, \tag{10.10.3}$$

then it follows that $I - \tau\, A$ is invertible, with

$$(I - \tau\, A)^{-1} = \sum_{j=0}^{r} \tau^j\, A^j. \tag{10.10.4}$$

Note that

$$\det(I - \tau\, A) \tag{10.10.5}$$

and

$$\det\Big(\sum_{j=0}^{r} \tau^j\, A^j\Big) \tag{10.10.6}$$

are polynomials in τ with real or complex coefficients, as appropriate. Both of these polynomials are equal to 1 at $\tau = 0$. If (10.10.3) holds, then

$$\det(I - \tau\, A)\, \det\Big(\sum_{j=0}^{r} \tau^j\, A^j\Big) = 1 \tag{10.10.7}$$

for all τ in $\mathbf{R}$ or $\mathbf{C}$, as appropriate. One can use this to get that

$$\det(I - \tau\, A) = 1 \tag{10.10.8}$$

and

$$\det\Big(\sum_{j=0}^{r} \tau^j\, A^j\Big) = 1 \tag{10.10.9}$$

for each τ.

If t is a nonzero real or complex number, as appropriate, then it is easy to see that (10.9.4) is equivalent to (10.10.8), with $\tau = 1/t$. Of course, if (10.9.4) holds for all $t \neq 0$, then it holds when $t = 0$ too. One could also obtain (10.9.4) with $t = 0$ more directly from (10.10.3). This is another way to obtain (10.9.4) from (10.10.3), as mentioned in the previous section.

If

$$f(t) = \sum_{j=0}^{m} b_j\, t^j \tag{10.10.10}$$

is any polynomial with real or complex coefficients, then

$$f(A) = \sum_{j=0}^{m} b_j\, A^j \tag{10.10.11}$$

defines a linear mapping from $\mathbf{R}^l$ or $\mathbf{C}^l$ into itself, as appropriate. If A is any linear mapping from $\mathbf{R}^l$ or $\mathbf{C}^l$ into itself, then it is well known that one can find a nonzero polynomial $f(t)$ of degree at most l^2 such that

$$f(A) = 0. \tag{10.10.12}$$

Of course, this follows from the Cayley–Hamilton theorem, and it can also be obtained more directly from the fact that $\mathcal{L}(\mathbf{R}^l)$, $\mathcal{L}(\mathbf{C}^l)$ have dimension l^2, as vector spaces over the real and complex numbers, respectively.

If A is a linear mapping from $\mathbf{C}^l$ into itself, and $A - t\,I$ is invertible for every $t \in \mathbf{C}$ with $t \neq 0$, then (10.10.12) implies that A is nilpotent, because f can be expressed as the product of a nonzero constant and finitely many linear factors. This invertibility condition holds when (10.9.2) holds for every $t \in \mathbf{C}$, or equivalently (10.10.8) holds for every $\tau \in \mathbf{C}$. If A is a linear mapping from $\mathbf{R}^l$ into itself, then A has a unique extension to a linear mapping from $\mathbf{C}^l$ into itself, as a vector space over the complex numbers, that we may denote by A as well. If (10.9.4) holds for every $t \in \mathbf{R}$, or equivalently (10.10.8) holds for every $\tau \in \mathbf{R}$, then these conditions hold for all $t, \tau \in \mathbf{C}$, because the left sides of these equations are polynomials in t, τ, respectively. The argument in the complex case implies that A is nilpotent on $\mathbf{C}^l$, and thus on $\mathbf{R}^l$.

Chapter 11

Power series in several variables

11.1 Sums over multi-indices

Let n be a positive integer, and let $(\mathbf{Z}_+ \cup \{0\})^n$ be the set of n tuples of elements of the set $\mathbf{Z}_+ \cup \{0\}$ of nonnegative integers. Equivalently, this is the set of all multi-indices. If f is a real or complex-valued function on $(\mathbf{Z}_+ \cup \{0\})^n$, then we may be interested in a sum of the form

$$\sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} f(\alpha). \tag{11.1.1}$$

Of course, this can be reduced to a finite sum when $f(\alpha) = 0$ for all but finitely many multi-indices α. If $n = 1$, then this may be interpreted as an infinite series.

One can try to define (11.1.1) for any n by reducing to an infinite series. One way to do this is to use the fact that $(\mathbf{Z}_+ \cup \{0\})^n$ is countably infinite, so that one can find a sequence $\{\alpha(l)\}_{l=0}^\infty$ of multi-indices in which every multi-index occurs exactly once. Thus one may try to interpret the sum (11.1.1) as being equal to the infinite series

$$\sum_{l=0}^\infty f(\alpha(l)). \tag{11.1.2}$$

Alternatively, let $E_0, E_1, E_2, E_3, \ldots$ be an infinite sequence of nonempty finite subsets of $(\mathbf{Z}_+ \cup \{0\})^n$ such that

$$E_N \subseteq E_{N+1} \tag{11.1.3}$$

for every nonnegative integer N, and

$$\bigcup_{N=0}^\infty E_N = (\mathbf{Z}_+ \cup \{0\})^n. \tag{11.1.4}$$

One may wish to interpret the sum (11.1.1) as being equal to

$$\lim_{N \to \infty} \sum_{\alpha \in E_N} f(\alpha), \tag{11.1.5}$$

if the limit exists.

Let $\{\alpha(l)\}_{l=0}^{\infty}$ be an enumeration of $(\mathbf{Z}_+ \cup \{0\})^n$, as before. If we put

$$E_N = \{\alpha(0), \alpha(1), \ldots, \alpha(N)\} \tag{11.1.6}$$

for each nonnegative integer N, then we get a sequence of nonempty finite subsets of $(\mathbf{Z}_+ \cup \{0\})^n$ that satisfies (11.1.3) and (11.1.4). In this case,

$$\sum_{\alpha \in E_N} f(\alpha) = \sum_{l=0}^{N} f(\alpha(l)) \tag{11.1.7}$$

for each $N \ge 0$, so that (11.1.2) is the same as (11.1.5).

As another basic example, one can take E_N to be

$$\{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n : |\alpha| \le N\} \tag{11.1.8}$$

for each $N \ge 0$, where $|\alpha|$ is the order of α, as usual. Another possibility is to take E_N to be

$$\{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n : \alpha_j \le N \text{ for each } j = 1, \ldots, N\} \tag{11.1.9}$$

for every $N \ge 0$. These are the same when $n = 1$, in which case (11.1.5) is the same as the usual interpretation of (11.1.1) as an infinite series.

Let f be a nonnegative real-valued function on $(\mathbf{Z}_+ \cup \{0\})^n$. If A is a nonempty finite subset of $(\mathbf{Z}_+ \cup \{0\})^n$, then

$$\sum_{\alpha \in A} f(\alpha) \tag{11.1.10}$$

is a nonnegative real number. Let us say that f is *summable* on $(\mathbf{Z}_+ \cup \{0\})^n$ if the collection of these finite sums has an upper bound in $\mathbf{R}$. Under these conditions, the sum (11.1.1) may be defined as the supremum of the set of these finite sums. Otherwise, it is sometimes convenient to interpret (11.1.1) as being equal to $+\infty$.

If $n = 1$, then the summability of f is equivalent to the convergence of the corresponding infinite series of nonnegative real numbers, with the same value of the sum. If $\{\alpha(l)\}_{l=0}^{\infty}$ is any enumeration of $(\mathbf{Z}_+ \cup \{0\})^n$, then the summability of f is equivalent to the convergence of (11.1.2), with the same value of the sum.

Let $E_0, E_1, E_2, E_3, \ldots$ be an infinite sequence of nonempty finite subsets of $(\mathbf{Z}_+ \cup \{0\})^n$ that satisfies (11.1.3) and (11.1.4) again. It is easy to see that f is summable on $(\mathbf{Z}_+ \cup \{0\})^n$ if and only if the sums

$$\sum_{\alpha \in E_N} f(\alpha) \tag{11.1.11}$$

bounded. In this case, the supremum of these sums is the same as the supremum of the set of sums of the form (11.1.10). We also get that the limit in (11.1.5) exists and is equal to this supremum, because the sums (11.1.11) are monotonically increasing in N.

Suppose that f can be expressed as

$$f(\alpha) = \prod_{j=1}^{n} f_j(\alpha_j), \tag{11.1.12}$$

where f_j is a nonnegative real-valued function on the set $\mathbf{Z}_+ \cup \{0\}$ of nonnegative integers for each $j = 1, \ldots, n$. If E_N is as in (11.1.9), then

$$\sum_{\alpha \in E_N} f(\alpha) = \prod_{j=1}^{n} \Big(\sum_{\alpha_j = 0}^{N} f_j(\alpha_j) \Big) \tag{11.1.13}$$

for each $N \geq 0$. If f_j is summable on $\mathbf{Z}_+ \cup \{0\}$ for each $j = 1, \ldots, n$, then it follows that f is summable on $(\mathbf{Z}_+ \cup \{0\})^n$, with

$$\sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} f(\alpha) = \prod_{j=1}^{n} \Big(\sum_{\alpha_j = 0}^{\infty} f_j(\alpha_j) \Big). \tag{11.1.14}$$

Conversely, if f is summable on $(\mathbf{Z}_+ \cup \{0\})^n$, and if none of the f_j's is identically zero on $\mathbf{Z}_+ \cup \{0\}$, then f_j is summable on $\mathbf{Z}_+ \cup \{0\}$ for each j.

As a basic family of examples, let r be an element of the set $(\mathbf{R}_+ \cup \{0\})^n$ of n-tuples of nonnegative real numbers, and put

$$f(\alpha) = r^\alpha \tag{11.1.15}$$

for each multi-index α. This is of the form (11.1.12), with

$$f_j(\alpha_j) = r^{\alpha_j} \tag{11.1.16}$$

for each $j = 1, \ldots, n$. It follows that f is summable on $(\mathbf{Z}_+ \cup \{0\})^n$ if and only if $r_j < 1$ for each $j = 1, \ldots, n$, in which case

$$\sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} r^\alpha = \prod_{j=1}^{n} (1 - r_j)^{-1}. \tag{11.1.17}$$

11.2 Real and complex-valued functions

Let n be a positive integer, and let f be a real or complex-valued function on $(\mathbf{Z}_+ \cup \{0\})^n$. Let us say that f is *summable* on $(\mathbf{Z}_+ \cup \{0\})^n$ if $|f(\alpha)|$ is summable as a nonnegative real-valued function on $(\mathbf{Z}_+ \cup \{0\})^n$. If f is real valued, then this is equivalent to the summability of $f_+(\alpha) = \max(f(\alpha), 0)$ and $f_-(\alpha) = \max(-f(\alpha), 0)$ on $(\mathbf{Z}_+ \cup \{0\})^n$. If f is complex valued, then summability of f is equivalent to the summability of the real and imaginary parts of f.

If f is a summable real or complex-valued function on $(\mathbf{Z}_+ \cup \{0\})^n$, then the sum (11.1.1) may be defined as a real or complex number, as appropriate, by reducing to the case of nonnegative real-valued summable functions. More precisely, if f is real valued, then the sum may be defined as the difference of the analogous sums for f_+ and f_-. If f is complex valued, then the real and imaginary parts of the sum may be defined as the corresponding sums of the real and imaginary parts of f. In both cases, the sum (11.1.1) may be described equivalently as in (11.1.2) or (11.1.5), because of the analogous statements for nonnegative real-valued summable functions.

We also have that

$$\left| \sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} f(\alpha) \right| \le \sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} |f(\alpha)| \tag{11.2.1}$$

in both cases. If f is real valued, then this follows directly from the definition of the sum (11.1.1) mentioned in the preceding paragraph. If f is complex valued, and one tries to consider the real and imaginary parts of the sum directly, then one gets an extra factor of 2 on the right side, or $\sqrt{2}$ with a bit more effort. Of course, if $f(\alpha) = 0$ for all but finitely many $\alpha \in (\mathbf{Z}_+ \cup \{0\})^n$, then (11.2.1) follows from the triangle inequality for the absolute value of a complex number. One can use this to get (11.2.1), by expressing the sum (11.1.1) as in (11.1.2) or (11.1.5).

It is easy to see that the spaces of real and complex-valued summable functions on $(\mathbf{Z}_+ \cup \{0\})^n$ are linear subspaces of the spaces of all real and complex-valued functions on $(\mathbf{Z}_+ \cup \{0\})^n$, as vector spaces over $\mathbf{R}$ and $\mathbf{C}$, respectively. The linearity of the sum (11.1.1) in f can be obtained from the descriptions of the sum as in (11.1.2) or (11.1.5).

If $n = 1$, then the summability of f is equivalent to the absolute convergence of the corresponding infinite series. Similarly, if $\{\alpha(l)\}_{l=0}^\infty$ is any enumeration of $(\mathbf{Z}_+ \cup \{0\})^n$, then the summablility of f is equivalent to the absolute convergence of (11.1.2).

Suppose that f is as in (11.1.12), where f_j is a real or complex-valued summable function on $\mathbf{Z}_+ \cup \{0\}$ for each $j = 1, \ldots, n$. This implies that f is summable on $(\mathbf{Z}_+ \cup \{0\})^n$, as in the previous section. One can check that (11.1.14) holds under these conditions, using the same type of argument as before, or by reducing to the previous case.

Let $z \in \mathbf{C}^n$ be given, and put

$$f(\alpha) = z^\alpha \tag{11.2.2}$$

for each multi-index α. If $|z_j| < 1$ for each $j = 1, \ldots, n$, then f is summable on $(\mathbf{Z}_+ \cup \{0\})^n$, as in the previous section. In this case,

$$\sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} z^\alpha = \prod_{j=1}^n (1 - z_j)^{-1}, \tag{11.2.3}$$

by (11.1.14).

11.3 Cauchy products

Let n be a positive integer, and let f, g be real or complex-valued functions on $(\mathbf{Z}_+ \cup \{0\})^n$. If γ is a multi-index, then put

$$h(\gamma) = \sum_{\alpha+\beta=\gamma} f(\alpha)\, g(\beta). \tag{11.3.1}$$

More precisely, the sum on the right is taken over all multi-indices α, β such that $\alpha + \beta = \gamma$. Note that there are only finitely many such multi-indices α, β.

Suppose for the moment that $f(\alpha) = 0$ for all but finitely many multi-indices α, and that $g(\beta) = 0$ for all but finitely many multi-indices β. This implies that $h(\gamma) = 0$ for all but finitely many multi-indices γ. Under these conditions, one can verify that

$$\sum_{\gamma \in (\mathbf{Z}_+ \cup \{0\})^n} h(\gamma) = \Big(\sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} f(\alpha)\Big) \Big(\sum_{\beta \in (\mathbf{Z}_+ \cup \{0\})^n} g(\beta)\Big). \tag{11.3.2}$$

In fact, both sides of the equation are the same as the sum of $f(\alpha)\, g(\beta)$ over all multi-indices α, β. The sum on the left may be desacribed as the *Cauchy product* of the two sums on the right.

Suppose now that f, g are nonnegative real-valued functions on $(\mathbf{Z}_+ \cup \{0\})^n$, so that h is nonnegative as well. If N is a nonnegative integer, then let E_N be the set of multi-indices α with order $|\alpha| \le N$, as in (11.1.8). Observe that

$$\sum_{\gamma \in E_N} h(\gamma) \le \Big(\sum_{\alpha \in E_N} f(\alpha)\Big) \Big(\sum_{\beta \in E_N} g(\beta)\Big) \tag{11.3.3}$$

and

$$\Big(\sum_{\alpha \in E_N} f(\alpha)\Big) \Big(\sum_{\beta \in E_N} g(\beta)\Big) \le \sum_{\gamma \in E_{2N}} h(\gamma) \tag{11.3.4}$$

for each $N \ge 0$. If f and g are summable on $(\mathbf{Z}_+ \cup \{0\})^n$, then it follows that h is summable too, and that (11.3.2) holds.

If f and g are any real or complex-valued functions on $(\mathbf{Z}_+ \cup \{0\})^n$, then

$$|h(\gamma)| \le \sum_{\alpha+\beta=\gamma} |f(\alpha)|\, |g(\beta)| \tag{11.3.5}$$

for every multi-index γ. Suppose that f and g are summable on $(\mathbf{Z}_+ \cup \{0\})^n$, which implies that the right side of (11.3.5) is summable as a function of γ, as in the preceding paragraph. It follows that h is summable as well, with

$$\sum_{\gamma \in (\mathbf{Z}_+ \cup \{0\})^n} |h(\gamma)| \le \Big(\sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} |f(\alpha)|\Big) \Big(\sum_{\beta \in (\mathbf{Z}_+ \cup \{0\})^n} |g(\beta)|\Big). \tag{11.3.6}$$

One can check that (11.3.2) holds too, by reducing to the case of nonnegative real-valued summable functions.

Let $z \in \mathbf{C}^n$ be given, and suppose that

$$f(\alpha) = a_\alpha \, z^\alpha, \ g(\beta) = b_\beta \, z^\beta \tag{11.3.7}$$

for all multi-indices α, β, where a_α, b_β are complex numbers. If we put

$$c_\gamma = \sum_{\alpha+\beta=\gamma} a_\alpha \, b_\beta \tag{11.3.8}$$

for each multi-index γ, then we get that

$$h(\gamma) = c_\gamma \, z^\gamma. \tag{11.3.9}$$

11.4 Power series on closed polydisks

Let n be a positive integer, and let $z_0 = (z_{0,1}, \ldots, z_{0,n}) \in \mathbf{C}^n$ be given. Also let a_α be a complex number for each multi-index α, and consider the *power series*

$$f(z) = \sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} a_\alpha \, (z - z_0)^\alpha \tag{11.4.1}$$

in $z_1, \ldots, z_n$, centered at z_0. More precisely, the sum on the right is defined as a complex number for each $z \in \mathbf{C}^n$ such that

$$a_\alpha \, (z - z_0)^\alpha \tag{11.4.2}$$

is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$.

Let $r \in (\mathbf{R}_+ \cup \{0\})^n$ be given, and suppose for the moment that

$$|a_\alpha| \, r^\alpha \tag{11.4.3}$$

is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$. This implies that (11.4.2) is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$ when

$$|z_j - z_{0,j}| \le r_j, \ 1 \le j \le n. \tag{11.4.4}$$

This means that (11.4.1) defines a complex-valued function on the *closed polydisk*

$$\{z \in \mathbf{C}^n : |z_j - z_{0,j}| \le r_j, \, 1 \le j \le n\}, \tag{11.4.5}$$

which is a closed set in $\mathbf{C}^n$, with respect to the standard Euclidean metric.

Let $\epsilon > 0$ be given, and let $A(\epsilon)$ be a nonempty finite subset of $(\mathbf{Z}_+ \cup \{0\})^n$ such that

$$\sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} |a_\alpha| \, r^\alpha < \Big(\sum_{\alpha \in A(\epsilon)} |a_\alpha| \, r^\alpha \Big) + \epsilon. \tag{11.4.6}$$

The existence of such a set follows from the definition of the sum on the left, as the supremum of the corresponding sums over nonempty finite subsets of $(\mathbf{Z}_+ \cup \{0\})^n$, as in Section 11.1. Using this, we get that

$$\sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n \setminus A(\epsilon)} |a_\alpha| \, r^\alpha < \epsilon, \tag{11.4.7}$$

because of the linearity of the sum.

Let A be a nonempty finite subset of $(\mathbf{Z}_+ \cup \{0\})^n$ such that

$$A(\epsilon) \subseteq A. \tag{11.4.8}$$

If $z \in \mathbf{C}^n$ satisfies (11.4.4), then

$$\begin{aligned} \Big| f(z) - \sum_{\alpha \in A} a_\alpha \, (z - z_0)^\alpha \Big| &= \Big| \sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n \setminus A} a_\alpha \, (z - z_0)^\alpha \Big| \\ &\le \sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n \setminus A} |a_\alpha| \, r^\alpha \\ &\le \sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n \setminus A(\epsilon)} |a_\alpha| \, r^\alpha < \epsilon. \end{aligned} \tag{11.4.9}$$

Let $E_0, E_1, E_2, E_3, \ldots$ be an infinite sequence of nonempty finite subsets of $(\mathbf{Z}_+ \cup \{0\})^n$ that satisfy (11.1.3) and (11.1.4). If $z \in \mathbf{C}^n$ satisfies (11.4.4), then

$$\lim_{N \to \infty} \sum_{\alpha \in E_N} a_\alpha \, (z - z_0)^\alpha = f(z), \tag{11.4.10}$$

as in Sections 11.1 and 11.2. In fact, the convergence is uniform over (11.4.5), as in (11.4.9). This corresponds to a classical criterion for uniform convergence of Weierstrass. It follows that f is continuous on (11.4.5), because polynomials are continuous on $\mathbf{C}^n$.

11.5 Power series on open polydisks

Let n be a positive integer, let $z_0 \in \mathbf{C}^n$ be given, and let a_α be a complex number for each multi-index α. Also let t be an element of the set $(\mathbf{R}_+ \cup \{+\infty\})^n$ of positive extended real numbers. Suppose that if $r \in (\mathbf{R}_+ \cup \{0\})^n$ satisfies

$$r_j < t_j, \ 1 \le j \le n, \tag{11.5.1}$$

then (11.4.3) is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$. If $z \in \mathbf{C}^n$ satisfies

$$|z_j - z_{0,j}| < t_j, \ 1 \le j \le n, \tag{11.5.2}$$

then it follows that (11.4.2) is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$.

This implies that (11.4.1) defines a complex-valued function on

$$\{z \in \mathbf{C}^n : |z_j - z_{0,j}| < t_j,\ 1 \le j \le n\}, \tag{11.5.3}$$

which is an open set in $\mathbf{C}^n$, with respect to the standard Euclidean metric. This set may be described as an *open polydisk* in $\mathbf{C}^n$, at least when $t_1, \ldots, t_n$ are finite. One can check that f is continuous on (11.5.3), because its restriction to any closed polydisk (11.4.5) is continuous when (11.5.1) holds, as in the previous section.

Suppose for the moment that $t_1, \ldots, t_n$ are finite, and that

$$|a_\alpha|\, t^\alpha \tag{11.5.4}$$

is bounded as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$. If $r \in (\mathbf{R}_+ \cup \{0\})^n$ satisfies (11.5.1), then

$$r^\alpha\, t^{-\alpha} \tag{11.5.5}$$

is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$, as in Section 11.1. This implies that (11.4.3) is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$.

Let β be a multi-index. If α is another multi-index, then α^β can be defined as a nonnegative integer in the usual way. If $r \in (\mathbf{R}_+ \cup \{0\})^n$ satisfies (11.5.1), then

$$\alpha^\beta\, |a_\alpha|\, r^\alpha \tag{11.5.6}$$

is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$. To see this, one can use an n-tuple $r_0 = (r_{0,1}, \ldots, r_{0,n})$ of positive real numbers such that

$$r_j < r_{0,j} < t_j,\ 1 \le j \le n. \tag{11.5.7}$$

Under these conditions,

$$|a_\alpha|\, r_0^\alpha \tag{11.5.8}$$

is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$, by hypothesis, and

$$\alpha^\beta\, r^\alpha\, r_0^{-\alpha} \tag{11.5.9}$$

is bounded as a function of α on $(\mathbf{Z}_+ \cup \{0\})^n$, by well-known results.

If one differentiates the right side of (11.4.1) term-by-term, then one gets a power series of the same type, with suitable coefficients. The remarks in the preceding paragraph imply that this power series has the same summability properties as those considered for $f(z)$ in this section. It is well known that $f(z)$ is smooth on (11.5.3), with derivatives given by differentiating the power series termwise.

More precisely, $f(z)$ is holomorphic on (11.5.3), because polynomials in $z_1, \ldots, z_n$ are holomorphic on $\mathbf{C}^n$. If β is any multi-index, then

$$\frac{\partial^{|\beta|} f}{\partial z^\beta}(z_0) = \beta!\, a_\beta. \tag{11.5.10}$$

Conversely, it is well known that any holomorphic function on (11.5.3) can be expressed as a power series with these summability properties.

11.6 Double sums

Let m and n be positive integers, and let us refer to multi-indices associated to n as *n-multi-indices*, so that we may also consider m-multi-indices and $(m+n)$-multi-indices. Let us identify the set $(\mathbf{Z}_+ \cup \{0\})^{m+n}$ of all $(m+n)$-multi-indices with the set

$$(\mathbf{Z}_+ \cup \{0\})^m \times (\mathbf{Z}_+ \cup \{0\})^n \tag{11.6.1}$$

of ordered pairs (α, β), where α is an m-multi-index, and β is an n-multi-index.

Let $f(\alpha, \beta)$ be a nonnegative real-valued function on $(\mathbf{Z}_+ \cup \{0\})^{m+n}$, identified with (11.6.1). If $f(\alpha, \beta)$ is summable on $(\mathbf{Z}_+ \cup \{0\})^{m+n}$, then it is easy to see that for each $\alpha \in (\mathbf{Z}_+ \cup \{0\})^m$, $f(\alpha, \beta)$ is summable as a function of β on $(\mathbf{Z}_+ \cup \{0\})^n$. If A is a nonempty finite subset of $(\mathbf{Z}_+ \cup \{0\})^m$, then one can check that

$$\sum_{\alpha \in A} \Big(\sum_{\beta \in (\mathbf{Z}_+ \cup \{0\})^n} f(\alpha, \beta) \Big) \le \sum_{(\alpha, \beta) \in (\mathbf{Z}_+ \cup \{0\})^{m+n}} f(\alpha, \beta). \tag{11.6.2}$$

This implies that

$$\sum_{\beta \in (\mathbf{Z}_+ \cup \{0\})^n} f(\alpha, \beta) \tag{11.6.3}$$

is summable as a nonnegative real-valued function of α on $(\mathbf{Z}_+ \cup \{0\})^m$, with

$$\sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^m} \Big(\sum_{\beta \in \mathbf{Z}_+ \cup \{0\})^n} f(\alpha, \beta) \Big) \le \sum_{(\alpha, \beta) \in (\mathbf{Z}_+ \cup \{0\})^{m+n}} f(\alpha, \beta). \tag{11.6.4}$$

Conversely, suppose that for each $\alpha \in (\mathbf{Z}_+ \cup \{0\})^m$, $f(\alpha, \beta)$ is summable as a function of β on $(\mathbf{Z}_+ \cup \{0\})^n$, and that (11.6.3) is summable as a nonnegative real-valued function of α on $(\mathbf{Z}_+ \cup \{0\})^m$. One can check that $f(\alpha, \beta)$ is summable on $(\mathbf{Z}_+ \cup \{0\})^{m+n}$ under these conditions, with

$$\sum_{(\alpha, \beta) \in (\mathbf{Z}_+ \cup \{0\})^{m+n}} f(\alpha, \beta) \le \sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^m} \Big(\sum_{\beta \in (\mathbf{Z}_+ \cup \{0\})^n} f(\alpha, \beta) \Big). \tag{11.6.5}$$

This means that

$$\sum_{(\alpha, \beta) \in (\mathbf{Z}_+ \cup \{0\})^{m+n}} f(\alpha, \beta) = \sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} \Big(\sum_{\beta \in (\mathbf{Z}_+ \cup \{0\})^n} f(\alpha, \beta) \Big) \tag{11.6.6}$$

in both cases. Of course, there are analogous statements for summing over α first.

Suppose now that $f(\alpha, \beta)$ is a summable real or complex-valued function on $(\mathbf{Z}_+ \cup \{0\})^{m+n}$. This implies that for each $\alpha \in (\mathbf{Z}_+ \cup \{0\})^m$, $f(\alpha, \beta)$ is summable as a function of β on $(\mathbf{Z}_+ \cup \{0\})^n$, as before. We also have that

$$\Big| \sum_{\beta \in (\mathbf{Z}_+ \cup \{0\})^n} f(\alpha, \beta) \Big| \le \sum_{\beta \in (\mathbf{Z}_+ \cup \{0\})^n} |f(\alpha, \beta)| \tag{11.6.7}$$

for every $\alpha \in (\mathbf{Z}_+ \cup \{0\})^m$, as in Section 11.2. The right side is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^m$, as before. It follows that (11.6.3) is summable as a function of α on $(\mathbf{Z}_+ \cup \{0\})^m$. One can check that (11.6.6) holds here too, by reducing to the case of summable nonnegative real-valued functions on $(\mathbf{Z}_+ \cup \{0\})^{m+n}$. There are analogous statements for summing over α first, as before.

Now let $f(\alpha)$, $g(\beta)$ be summable real or complex-valued functions of α, β on $(\mathbf{Z}_+ \cup \{0\})^m$, $(\mathbf{Z}_+ \cup \{0\})^n$, respectively. One can check that

$$\phi(\alpha, \beta) = f(\alpha)\, g(\beta) \tag{11.6.8}$$

is summable on $(\mathbf{Z}_+ \cup \{0\})^{m+n}$, by summing

$$|\phi(\alpha, \beta)| = |f(\alpha)|\, |g(\beta)| \tag{11.6.9}$$

one variable at a time. Similarly,

$$\sum_{(\alpha,\beta)\in(\mathbf{Z}_+\cup\{0\})^{m+n}} f(\alpha)\, g(\beta) = \Big(\sum_{\alpha\in\mathbf{Z}_+\cup\{0\})^m} f(\alpha)\Big)\Big(\sum_{\beta\in(\mathbf{Z}_+\cup\{0\})^n} g(\beta)\Big). \tag{11.6.10}$$

11.7 Some more rearrangements

Let m and n be positive integers, and let $A(\gamma)$ be a finite set of n-multi-indices for each m-multi-index γ. Suppose that the $A(\gamma)$'s are pairwise disjoint, so that

$$A(\gamma) \cap A(\gamma') = \emptyset \tag{11.7.1}$$

for all m-multi-indices γ, γ' with $\gamma \neq \gamma'$, and that

$$\bigcup_{\gamma\in(\mathbf{Z}_+\cup\{0\})^m} A(\gamma) = (\mathbf{Z}_+ \cup \{0\})^n. \tag{11.7.2}$$

Let ϕ be a real or complex-valued function on $(\mathbf{Z}_+ \cup \{0\})^n$, and put

$$\psi(\gamma) = \sum_{\alpha\in A(\gamma)} \phi(\alpha) \tag{11.7.3}$$

for every m-multi-index γ. This is interpreted as being equal to 0 when $A(\gamma) = \emptyset$. Note that

$$|\psi(\gamma)| \le \sum_{\alpha\in A(\gamma)} |\phi(\alpha)| \tag{11.7.4}$$

for every n-multi-index γ.

Suppose for the moment that ϕ is a nonnegative real-valued function on $(\mathbf{Z}_+ \cup \{0\})^n$, so that ψ is a nonnegative real-valued function on $(\mathbf{Z}_+ \cup \{0\})^m$.

If ϕ is summable on $(\mathbf{Z}_+ \cup \{0\})^n$ then one can check that ψ is summable on $(\mathbf{Z}_+ \cup \{0\})^m$, with

$$\sum_{\gamma \in (\mathbf{Z}_+ \cup \{0\})^m} \psi(\gamma) \le \sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} \phi(\alpha). \tag{11.7.5}$$

Similarly, if ψ is summable on $(\mathbf{Z}_+ \cup \{0\})^m$, then one can verify that ϕ is summable on $(\mathbf{Z}_+ \cup \{0\})^n$, with

$$\sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} \phi(\alpha) \le \sum_{\gamma \in (\mathbf{Z}_+ \cup \{0\})^m} \psi(\gamma). \tag{11.7.6}$$

It follows that

$$\sum_{\gamma \in (\mathbf{Z}_+ \cup \{0\})^m} \psi(\gamma) = \sum_{\alpha \in (\mathbf{Z}_+ \cup \{0\})^n} \phi(\alpha) \tag{11.7.7}$$

in both cases.

If ϕ is a summable real or complex-valued function on $(\mathbf{Z}_+ \cup \{0\})^n$, then one can use (11.7.4) and the remarks in the preceding paragraph to get that ψ is summable on $(\mathbf{Z}_+ \cup \{0\})^m$. One can also check that (11.7.7) holds under these conditions, by reducing to the case of summable nonnegative real-valued functions on $(\mathbf{Z}_+ \cup \{0\})^n$.

As a basic class of examples, let us take $n = 2\,m$, and identify the set $(\mathbf{Z}_+ \cup \{0\})^{2m}$ of all $(2\,m)$-multi-indices with the set

$$(\mathbf{Z}_+ \cup \{0\})^m \times (\mathbf{Z}_+ \cup \{0\})^m \tag{11.7.8}$$

of all ordered pairs (α, β) of m-multi-indices, as in the previous section. If γ is an m-multi-index, then put

$$A(\gamma) = \{(\alpha, \beta) \in (\mathbf{Z}_+ \cup \{0\})^{2m} : \alpha + \beta = \gamma\}. \tag{11.7.9}$$

These are pairwise-disjoint nonempty finite subsets of $(\mathbf{Z}_+ \cup \{0\})^{2m}$, whose union is all of $(\mathbf{Z}_+ \cup \{0\})^{2m}$.

Let f, g be real or complex-valued functions on $(\mathbf{Z}_+ \cup \{0\})^m$, and let $\phi(\alpha, \beta)$ be defined on $(\mathbf{Z}_+ \cup \{0\})^{2m}$ as in (11.6.8). In this case,

$$\psi(\gamma) = \sum_{\alpha + \beta = \gamma} f(\alpha)\, g(\beta) \tag{11.7.10}$$

is the same as $h(\gamma)$ in Section 11.3, and the earlier properties of $h(\gamma)$ could also be obtained from the remarks in this and the previous section.

Bibliography

[1] R. Agnew, *Differential Equations*, 2nd edition, McGraw-Hill, 1960.

[2] R. Agnew, *Views and approximations of differential equations*, American Mathematical Monthly **60** (1953), 1–6.

[3] R. Agnew, *Classroom notes: Transformations of the Laplacian*, American Mathematical Monthly **60** (1953), 323–325.

[4] C. Allendoerfer, *Classroom notes: Editorial — The proof of Euler's equation*, American Mathematical Monthly **55** (1948), 94–95.

[5] S. Antman, *The equations for large vibrations of strings*, American Mathematical Monthly **87** (1980), 359–370.

[6] T. Archibald, *Connectivity and smoke-rings: Green's second identity in its first fifty years*, Mathematics Magazine **62** (1989), 219–232.

[7] A. Aksoy and M. Martelli, *The wave equation, mixed partial derivatives, and Fubini's theorem*, American Mathematical Monthly **111** (2004), 340–347.

[8] S. Axler, *Harmonic functions from a complex analysis viewpoint*, American Mathemtical Monthly **93** (1986), 246–258.

[9] S. Axler, P. Bourdon, and W. Ramey, *Bôcher's theorem*, American Mathematical Monthly **99** (1992), 51–55.

[10] S. Axler, P. Bourdon, and W. Ramey, *Harmonic Function Theory*, 2nd edition, Springer-Verlag, 2001.

[11] J. Baker, *Integration over spheres and the divergence theorem for balls*, American Mathematical Monthly **104** (1997), 36–47.

[12] J. Baker, *The Dirichlet problem for ellipsoids*, American Mathematical Monthly **106** (1999), 829–834.

[13] R. Beals, *Advanced Mathematical Analysis*, Springer-Verlag, 1973.

[14] R. Beals, *Analysis: An Introduction*, Cambridge University Press, 2004.

[15] R. Beals, *The scope of the power series method*, Mathematics Magazine **86** (2013), 56–62.

[16] F. Brackx, R. Delanghe, and F. Sommen, *Clifford Analysis*, Pitman, 1982.

[17] J. Bramble and L. Payne, *Mean value theorems for polyharmonic functions*, American Mathematical Monthly **73** (1966), 124–127.

[18] R. Burckel, *A strong converse to Gauss's mean-value theorem*, American Mathematical Monthly **87** (1980), 819–820.

[19] R. Burckel, *Three secrets about harmonic functions*, American Mathematical Monthly **104** (1997), 52–56.

[20] D. Cannell, *George Green: an enigmatic mathematician*, American Mathematical Monthly **106** (1999), 136–151.

[21] D. Cannell, *George Green, Mathematician & Physicist 1793–1841*, 2nd edition, Society for Industrial and Applied Mathematics (SIAM), 2001.

[22] S. Chapman, *Drums that sound the same*, American Mathematical Monthly **102** (1995), 124–138.

[23] S. Chu, *On a mean value property for solutions of a wave equation*, American Mathematical Monthly **74** (1967), 711–713.

[24] E. Coddington and N. Levinson, *Theory of Ordinary Differential quations*, McGraw-Hill, 1955.

[25] J. Cortissoz, *A note on harmonic functions on surfaces*, American Mathematical Monthly **123** (2016), 884–893.

[26] J. Dodziuk, *Eigenvalues of the Laplacian and the heat equation*, American Mathematical Monthly **88** (1981), 686-695.

[27] J. Duistermaat and J. Kolk, *Distributions: Theory and Applications*, translated from the Dutch by J. van Braam Houckgeest, Birkhäuser, 2010.

[28] G. Evans, *A necessary and sufficient condition of Wiener*, American Mathematical Monthly **54** (1947), 151–155.

[29] L. Evans, *Partial Differential Equations*, 2nd edition, American Mathematical Society, 2010.

[30] F. Ficken, *A derivation of the equation for a vibrating string*, American Mathematical Monthly **64** (1957), 155–157.

[31] F. Ficken, *Some uses of linear spaces in analysis*, American Mathematical Monthly **66** (1959), 259–275.

[32] G. Folland, *Introduction to Partial Differential Equations*, Princeton University Press, 1976. A second edition was published in 1995.

[33] G. Folland, *On characterizations of analyticity*, American Mathematical Monthly **93** (1986), 640–641.

[34] G. Folland, *Fourier Analysis and Its Applications*, Wadsworth & Brooks / Cole, 1992.

[35] G. Folland, *Real Analysis*, 2nd edition, Wiley, 1999.

[36] G. Folland, *A Guide to Advanced Real Analysis*, Mathematical Association of America, 2009.

[37] F. Friedlander, *Introduction to the Theory of Distributions*, 2nd edition, with additional material by M. Joshi, Cambridge University Press, 1998.

[38] L. Gårding, *Some Points of Analysis and their History*, American Mathematical Society, 1997.

[39] F. Gehring, G. Martin, and B. Palka, *An Introduction to the Theory of Higher-Dimensional Quasiconformal Mappings*, American Mathematical Society, 2017.

[40] J. Gilbert and M. Murray, *Clifford Algebras and Dirac Operators in Harmonic Analysis*, Cambridge University Press, 1991.

[41] E. González-Velasco, *Connections in mathematical analysis: the case of Fourier series*, American Mathematical Monthly **99** (1992), 427–441.

[42] E. González-Velasco, *Fourier Analysis and Boundary Value Problems*, Academic Press, 1996.

[43] C. Gordon, *Sunada's isospectrality technique: two decades later*, in *Spectral Analysis in Geometry and Number Theory*, 45–58, Contemporary Mathematics **484**, American Mathematical Society, 2009.

[44] C. Gordon, D. Webb, and S. Wolpert, *One cannot hear the shape of a drum*, Bulletin of the American Mathematical Society (N.S.) **27** (1992), 134–138.

[45] C. Gordon, D. Webb, and S. Wolpert, *Isospectral plane domains and surfaces via Riemannian orbifolds*, Inventiones Mathematicae **110** (1992), 1-22.

[46] P. Gorkin and J. Smith, *Dirichlet: his life, his principle, and his problem*, Mathematics Magazine **78** (2005), 283–296.

[47] U. Graf, *Introduction to Hyperfunctions and Their Integral Transforms*, Birkhäuser, 2010.

[48] I. Grattan-Guiness, *Why did George Green write his essay of 1828 on electricity and magnetism?*, American Mathematical Monthly **102** (1995), 387–396.

[49] J. Gray, *Miscellanea: Clear theory*, American Mathematical Monthly **92** (1985), 158–159.

[50] J. Gray, *The Real and the Complex: A History of Analysis in the 19th Century*, Springer, 2015.

[51] J. Gray, *Change and Variations — A History of Differential Equations to 1900*, Springer, 2021.

[52] R. Greene and S. Krantz, *Function Theory of One Complex Variable*, 3rd edition, American Mathematical Society, 2006.

[53] R. Guenther, *Some elementary properties of the fundamental solution of parabolic equations*, Mathematics Magazine **39** (1966), 294–298.

[54] R. Hall and K. Jocić, *The mathematics of musical instruments*, American Mathematical Monthly **108** (2001), 347–354.

[55] W. Hansen, *A strong version of Liouville's theorem*, American Mathematical Monthly **115** (2008), 583–595.

[56] P. Hartman, *A differential equation with non-unique solutions*, American Mathematical Monthly **70** (1963), 255–259.

[57] P. Hartman, *Ordinary Differential Equations*, corrected reprint of the second edition, Society for Industrial and Applied Mathematics (SIAM), 2002.

[58] P. Hartman and A. Wintner, *Mean value theorems and linear operators*, American Mathematical Monthly **62** (1955), 217–222.

[59] R. Hersh, *How to classify differential polynomials*, American Mathematical Monthly **80** (1973), 641–654.

[60] G. Herzog, *Polynomials solving Dirichlet boundary value problems*, American Mathematical Monthly **107** (2000), 934–936.

[61] I. Hirschman, Jr. and D. Widder, *A miniature theory in illustration of the convolution transform*, American Mathematical Monthly **57** (1950), 667–674.

[62] I. Hirschman and D. Widder, *The Convolution Transform*, Princeton University Press, 1955.

[63] P. Hodge, Jr., *On the methods of characteristics*, American mathematical Monthly **57** (1950), 621–623.

[64] L. Hörmander, *Notions of Convexity*, Birkhäuser, 2007.

[65] J. Horvàth, *An introduction to distributions*, American Mathematical Monthly **77** (1970), 227–240.

[66] F. John, *Partial Differential Equations*, 4th edition, Springer-Verlag, 1991.

[67] F. Jones, *Lebesgue Integration on Euclidean Space*, Jones and Bartlett, 1993.

[68] M. Kac, *Can one hear the shape of a drum?*, American Mathematical Monthly **73** (1966), No. 4, Part II, 1–23.

[69] A. Knapp, *Basic Real Analysis*, Birkhäuser, 2005.

[70] A. Knapp, *Advanced Real Analysis*, Birkhäuser, 2005.

[71] T. Körner, *Exercises for* Fourier Analysis ([73]), Cambridge University Press, 1993.

[72] T. Körner, *A Companion to Analysis: A Second First and First Second Course in Analysis*, American Mathematical Society, 2004.

[73] T. Körner, *Fourier Analysis*, Cambridge University Press, 2022.

[74] J. Král and W. Pfeffer, *Poisson integrals of Riemann integrable functions*, American Mathematical Monthly **98** (1991), 929–931.

[75] S. Krantz, *What is several complex variables?*, American Mathematical Monthly **94** (1987), 236–256.

[76] S. Krantz, *A Panorama of Harmonic Analysis*, Mathematical Association of America, 1999.

[77] S. Krantz, *Function Theory of Several Complex Variables*, AMS Chelsea, 2001.

[78] S. Krantz, *Complex analysis as catalyst*, American Mathematical Monthly **115** (2008), 775–794.

[79] S. Krantz, *Foundations of Analysis*, CRC Press, 2015.

[80] S. Krantz, *Real Analysis and Foundations*, 4th edition, CRC Press, 2017.

[81] S. Krantz, *Elementary Introduction to the Lebesgue Integral*, CRC Press, 2018.

[82] S. Krantz, *A calculus approach to complex variables*, Mathematics Magazine **94** (2021), 194–200.

[83] S. Krantz and H. Parks, *A Primer of Real-Analytic Functions*, 2nd edition, Birkhäuser, 2002.

[84] S. Krantz and H. Parks, *The Implicit Function Theorem: History, Theory, and Applications*, Birkhäuser / Springer, 2013.

[85] S. Lang, *Undergraduate Analysis*, 2nd edition, Springer-Verlag, 1997.

[86] R. Langer, *Fourier's Series: The Genesis and Evolution of a Theory*, The first Herbert Ellsworth Slaught Memorial Paper, American Mathematical Monthly **54** (1947), no. 7, part II.

[87] S. Lefschetz, *Differential Equations: Geometric Theory*, 2nd edition, Dover, 1977.

[88] J. Lützen, *The Prehistory of the Theory of Distributions*, Springer-Verlag, 1982.

[89] J. Lützen, *Euler's vision of a general partial differential calculus for a generalized kind of functions*, Mathematics Magazine **56** (1983), 299–306.

[90] J. Lützen, *Joseph Liouville 1809–1882: Master of Pure and Applied Mathematics*, Springer-Verlag, 1990.

[91] S. Mandelbrojt, *Obituary: Emile Picard, 1856–1941*, American Mathematical Monthly **49** (1942), 277–278.

[92] S. Mandelbrojt, *The mathematical work of Jacques Hadamard*, American Mathematical Monthly **60** (1953), 599–604.

[93] W. Martin, *Functions of several complex variables*, American Mathematical Monthly **52** (1945), 17–27.

[94] W. McEwen, *Spectral theory and its applications to differential eigenvalue problems*, American Mathematical Monthly **60** (1953), 223-233.

[95] E. Miles, *Three dimensional harmonic functions generated by analytic functions of a hypervariable*, American Mathematical Monthly **61** (1954), 694–697.

[96] E. Miles, Jr. and E. Williams, *A basic set of polynomial solutions for the Euler–Poisson–Darboux and Beltrami equations*, American Mathematical Monthly **63** (1956), 401–404.

[97] D. Minda, *The Dirchlet problem for a disk*, American Mathematical Monthly **97** (1990), 220–223.

[98] D. Mitrea, *Distributions, Partial Differential Equations, and Harmonic Analysis*, 2nd edition, Springer, 2018.

[99] T. Needham, *The geometry of harmonic functions*, Mathematics Magazine **67** (1994), 92–108.

[100] T. Needham, *Visual Complex Analysis*, 25th anniversary edition, Oxford University Press, 2023.

[101] N. Ortner and P. Wagner, *Fundamental Solutions of Linear Partial Differential Operators*, Springer, 2015.

[102] J. Ovall, *The Laplacian and mean and extreme values*, American Mathematical Monthly **123** (2016), 287–291.

[103] M. Protter, *The characteristic initial value problem for the wave equation and Riemann's method*, American Mathematical Monthly **61** (1954), 702–705.

[104] M. Protter and C. Morrey, Jr., *A First Course in Real Analysis*, 2nd edition, Springer-Verlag, 1991.

[105] M. Protter and H. Weinberger, *Maximum Principles in Differential Equations*, corrected reprint of the 1967 original, Springer-Verlag, 1984.

[106] R. Range, *Complex analysis: a brief tour into higher dimensions*, American Mathematical Monthly **110** (2003), 89–108.

[107] W. P. Reid, *Classroom notes: On numerical solutions to the one-dimensional wave equation*, American Mathematical Monthly **61** (1954), 115.

[108] W. P. Reid, *On some contradictions in boundary value problems*, Mathematics Magazine **37** (1964), 172–175.

[109] W. T. Reid, *Ordinary Differential Equations*, Wiley, 1971.

[110] W. T. Reid, *Anatomy of the ordinary differential equation*, American Mathematical Monthly **82** (1975), 971–984; addendum, **83** (1976), 718.

[111] J.-P. Rosay, *A very elementary proof of the Malgrange–Ehrenpreis theorem*, American Mathematical Monthly **98** (1991), 518–523.

[112] P. Rosenbloom and D. Widder, *A temperature function which vanishes initially*, American Mathematical Monthly **65** (1958), 607–609.

[113] W. Rudin, *Principles of Mathematical Analysis*, 3rd edition, McGraw-Hill, 1976.

[114] W. Rudin, *Real and Complex Analysis*, 2nd edition, McGraw-Hill, 1974. A third edition was published in 1987.

[115] W. Rudin, *Functional Analysis*, McGraw-Hill, 1973. A second edition was published in 1991.

[116] R. Seeley, *Spherical harmonics*, American Mathematical Monthly **73** (1966), no. 4, part II, 115–121.

[117] R. Shakarchi, *Problems and Solutions for* Undergraduate Analysis ([85]), Springer-Verlag, 1998.

[118] J. Slade, Jr., *Questions, discussions, and notes: Note on the Laplacian of a vector point function*, American Mathematical Monthly **40** (1933), 483–484.

[119] E. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, 1970.

[120] E. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, with the assistance of T. Murphy, Princeton University Press, 1993.

[121] E. Stein and R. Shakarchi, *Fourier Analysis: An Introduction*, Princeton University Press, 2003.

[122] E. Stein and R. Shakarchi, *Complex Analysis*, Princeton University Press, 2003.

[123] E. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton University Press, 2005.

[124] E. Stein and R. Shakarchi, *Functional Analysis: Introduction to Further Topics in Analysis*, Princeton University Press, 2011.

[125] E. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.

[126] W. Strauss, *Partial Differential Equations: An Introduction*, 2nd edition, Wiley, 2008.

[127] R. Strichartz, *The Way of Analysis*, Revised edition, Jones and Bartlett, 2000.

[128] R. Strichartz, *A Guide to Distribution Theory and Fourier Transforms*, World Scientific, 2003.

[129] R. Strichartz, *Differential Equations on Fractals*, Princeton University Press, 2006.

[130] A. Taylor, *A note on the Poisson kernel*, American Mathematical Monthly **57** (1950), 478–479.

[131] H. Thurston, *Classroom notes: On Euler's equation*, American Mathematical Monthly **67** (1960), 678–681.

[132] A. Torchinsky, *The Fourier transform and the wave equation*, American Mathematical Monthly **118** (2011), 599–609.

[133] F. Trèves, *Applications of distributions to PDE theory*, American Mathematical Monthly **77** (1970), 241–248.

[134] F. Trèves, *On the local solvability of linear partial differential equations*, Bulletin of the American Mathematical Society **76** (1970), 552–571.

[135] F. Trèves, *A treasure trove of geometry and analysis: the hyperquadric*, Notices of the American Mathematical Society **47** (2000), 1246–1256.

[136] F. Trèves, *Basic Linear Partial Differential Equations*, Dover, 2006.

[137] R. Vanderbei, *The falling slinky*, American Mathematical Monthly **124** (2017), 24–36.

[138] P. Wagner, *A new constructive proof of the Malgrange–Ehrenpreis theorem*, American Mathematical Monthly **116** (2009), 457–462.

[139] J. Walsh, *Lemniscates and equipotential curves of Green's function*, American Mathematical Monthly **42** (1935), 1–17.

[140] J. Walsh, *On the shape of level curves of Green's function*, American Mathematical Monthly **44** (1937), 202–213.

[141] J. Walsh, *Note on the shape of level curves of Green's function*, American Mathematical Monthly **60** (1953), 671–674.

[142] J. Walsh, *The circles of curvature of steepest descent of Green's function*, American Mathematical Monthly **68** (1961), 323–329.

[143] W. Walter, *An elementary proof of the Cauchy–Kowalevsky theorem*, American Mathemtical Monthly **92** (1985), 115–126.

[144] W. Walter, *Ordinary Differential Equations*, translated from the sixth German edition by R. Thompson, Springer-Verlag, 1998.

[145] S. Warschawski, *On the Green function of a star-shaped three dimensional region*, American Mathematical Monthly **57** (1950), 471–473.

[146] H. Weinberger, *A First Course in Partial Differential Equations with Complex Variables and Transform Methods*, corrected reprint of the 1965 original, Dover, 1995.

[147] G. Weiss, *Complex methods in harmonic analysis*, American Mathematical Monthly **77** (1970), 465–474.

[148] D. Widder, *The Laplace Transform*, Princeton University Press, 1941.

[149] D. Widder, *What is the Laplace transform?*, American Mathematical Monthly **52** (1945), 419–425.

[150] D. Widder, *The Heat Equation*, Academic Press, 1975.

[151] D. Widder, *The Airy transform*, American Mathematical Monthly **86** (1979), 271–277.

[152] D. Widder, *Advanced Calculus*, 3rd edition, Dover, 1989.

[153] C. Wilcox, *Positive temperatures with prescribed initial heat distributions*, American Mathematical Monthly **87** (1980), 183–186.

[154] C. Wilcox, *The Cauchy problem for the wave equation with distribution data: an elementary approach*, American Mathematical Monthly **98** (1991), 401–410.

[155] B. Wirth, *Green's function for the Neumann–Poisson problem on n-dimensional balls*, American Mathematical Monthly **127** (2020), 737–743.

[156] A. Zemanian, *Distribution Theory and Transform Analysis*, Dover, 1987.

Index