PROOF Assume there is a set E with $m(E) = 0$ but $\mu(E) > 0$. Put

(1) $$E_n = \{x \,\varepsilon\, E : (\bar{D}\mu)(x) < n\} \qquad (n = 1, 2, 3, \ldots).$$

Since $E = \bigcup E_n$, we have $\mu(E_n) > 0$ for some n. Fix this n. With the notation of Definition 8.3, put

(2) $$A_j = \{x \,\varepsilon\, E_n : \bar{\Delta}_{1/j}(x) < n\} \qquad (j = 1, 2, 3, \ldots).$$

Since $E_n = \bigcup A_j$, we have $\mu(A_j) > 0$ for some j. The regularity of μ (Theorem 2.18) now shows that there is a compact set $K \subset A_j$ with $\mu(K) > 0$.

Our construction shows that K has the following property: If $x \,\varepsilon\, K$, $x \,\varepsilon\, I$, $I \,\varepsilon\, \Omega$, and diam $I < 1/j$, then $\mu(I) < n \cdot m(I)$.

Let $\epsilon > 0$ be given. Since $K \subset E$, $m(K) = 0$, so there is an open set $V \supset K$ with $m(V) < \epsilon$.

Partition R^k into disjoint cubical boxes B, as in Sec. 2.19, whose diameter is less than $1/j$ and is so small that any box which intersects K lies in V. Keep those B's which intersect K, and enlarge each of them so as to obtain open cubes $I_i \supset B_i$ with $m(I_i) < 2m(B_i)$, diam $I_i < 1/j$. Then

$$\mu(K) \le \Sigma\mu(B_i) \le \Sigma\mu(I_i) < n \cdot \Sigma m(I_i)$$
$$< 2n \cdot \Sigma m(B_i) \le 2n \cdot m(V) < 2n\epsilon.$$

Since ϵ was arbitrary, we have $\mu(K) = 0$, a contradiction.

Functions of Bounded Variation

8.12 Definitions We associate with each complex function f on R^1 its *total variation function* T_f defined by

(1) $$T_f(x) = \sup \sum_{j=1}^{N} |f(x_j) - f(x_{j-1})| \qquad (-\infty < x < \infty),$$

where the supremum is taken over all N and over all choices of $\{x_j\}$ such that

(2) $$-\infty < x_0 < x_1 < \cdots < x_N = x.$$

In general,

(3) $$0 \le T_f(x) \le T_f(y) \le \infty \qquad (x < y).$$

If T_f is a bounded function, then (3) implies that

(4) $$V(f) = \lim_{x \to +\infty} T_f(x)$$

exists and is finite. In that case we say that f is of *bounded variation*, and we call $V(f)$ the *total variation of f;* the class of all such f will be denoted by BV.

If $-\infty < x \le \infty$, f is said to have a *left-hand limit* at x, written $f(x-)$, if there corresponds to each $\epsilon > 0$ a real number $\alpha < x$ such that

(5) $$\alpha < t < x \quad \text{implies} \quad |f(t) - f(x-)| < \epsilon.$$

If $f(x-) = f(x)$, f is said to be *left-continuous* at x.

Right-hand limits and continuity from the right are defined similarly on $[-\infty, \infty)$.

We call a function $f \,\varepsilon\, BV$ *normalized* if $f(x) \to 0$ as $x \to -\infty$ and f is left-continuous at every point of R^1. The class of these functions will be denoted by NBV.

Instead of considering only functions defined on all of R^1 we could equally well consider functions defined on any segment or interval of R^1. Neither the preceding definitions nor the theorems which follow would be affected in any significant way.

8.13 Theorem

(a) *If $f \,\varepsilon\, BV$ and $x < y$, then*

$$|f(y) - f(x)| \le T_f(y) - T_f(x).$$

(b) *If $f \,\varepsilon\, BV$, then $f(x-)$ exists at every point of $(-\infty, \infty]$, $f(x+)$ exists at every point of $[-\infty, \infty)$, the set of points at which f is discontinuous is at most countable, and there is a unique constant c and a unique function $g \,\varepsilon\, NBV$ such that*

$$f(x) = c + g(x)$$

at all points of continuity of f. Also, $V(g) \le V(f)$.

(c) *If $f \,\varepsilon\, NBV$, then $T_f \,\varepsilon\, NBV$.*

PROOF

(a) If $x < y$ and $\epsilon > 0$, there are points $x_0 < x_1 < \cdots < x_n = x$ so that

(1) $$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| > T_f(x) - \epsilon.$$

Hence

$$T_f(y) \ge |f(y) - f(x)| + \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| > |f(y) - f(x)| + T_f(x) - \epsilon.$$

This proves (a).

(b) It follows from (a) that if $\{x_i\}$ is a sequence for which $\{T_f(x_i)\}$ is a Cauchy sequence, then $\{f(x_i)\}$ is also a Cauchy sequence. Since monotone functions (and T_f in particular) have right- and left-hand limits at all points and since they have at most countably many discontinuities, the same therefore holds for f. Hence we can define

$$(2) \qquad c = \lim_{t\to-\infty} f(t), \qquad g(x) = f(x-) - c \qquad (x \in R^1).$$

It is clear that $g(x) \to 0$ as $x \to -\infty$. If $x \in R^1$ and $\epsilon > 0$, there exists an $\alpha < x$ such that $|f(t) - f(x-)| < \epsilon$ for all $t \in (\alpha,x)$. Since $f(t-)$ is a limit point of the set of all numbers $f(s)$, for $\alpha < s < t$, it follows that $|f(t-) - f(x-)| \le \epsilon$ if $\alpha < t < x$. Thus g is left-continuous.

If $x_0 < x_1 < \cdots < x_n$ and $\delta > 0$, then

$$(3) \qquad \sum_{i=1}^{n} |g(x_i) - g(x_{i-1})| = \lim_{\delta\to 0} \sum_{i=1}^{n} |f(x_i - \delta) - f(x_{i-1} - \delta)|,$$

and since none of the sums on the right of (3) exceeds $V(f)$, we have $V(g) \le V(f)$. In particular, $g \in BV$.

This proves (b), except for the uniqueness. But if two left-continuous functions coincide on a dense subset of R^1, then they are identical. The uniqueness of g now follows easily.

(c) If $f \in NBV$, fix $x \in R^1$, $\epsilon > 0$, and choose points

$$x_0 < x_1 < \cdots < x_n = x$$

so that (1) holds. If $t_0 < \cdots < t_N = x_0$, then

$$(4) \qquad T_f(x) \ge \sum_{j=1}^{N} |f(t_j) - f(t_{j-1})| + \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|.$$

By (1), the first sum in (4) is less than ϵ. Hence $T_f(x_0) \le \epsilon$, and this says that $T_f(t) \to 0$ as $t \to -\infty$.

Finally, choose t, so that $x_{n-1} < t < x_n$. Then

$$(5) \qquad \sum_{i=1}^{n-1} |f(x_i) - f(x_{i-1})| + |f(t) - f(x_{n-1})| \le T_f(t) \le T_f(x-) \le T_f(x).$$

If we let $t \to x_n = x$, the left side of (5) tends to the left side of (1), since $f(x) = f(x-)$, and this gives $T_f(x) - \epsilon < T_f(x-)$. Comparison with (5) now shows that $T_f(x-) = T_f(x)$, and the proof is complete.

The next theorem explains the importance of the class NBV. Observe how the correspondence between f and μ associates the total variation of f with that of μ, and how the existence of Lebesgue measure is used to construct μ in part (b).

8.14 Theorem

(a) If μ is a complex Borel measure on R^1 and if

$$(1) \qquad f(x) = \mu((-\infty,x)) \qquad (x \in R^1),$$

then $f \in NBV$.

(b) Conversely, to every $f \in NBV$ there corresponds a unique complex Borel measure μ such that (1) holds; for this μ,

$$(2) \qquad T_f(x) = |\mu|((-\infty,x)) \qquad (x \in R^1).$$

(c) If (1) holds, then f is continuous precisely at those points x at which $\mu(\{x\}) = 0$.

PROOF If f is defined by (1) and if $x_1 < x_2 < \cdots$, $x_n \to x$, then $f(x_n) \to f(x)$, since

$$(3) \qquad (-\infty,x) = \bigcup_{n=1}^{\infty} (-\infty,x_n).$$

Thus f is left-continuous. If $x_1 > x_2 > \cdots$, $x_n \to -\infty$, then $\bigcap(-\infty,x_n) = \varnothing$, so $f(x) \to 0$ as $x \to -\infty$, by Theorem 1.19(e). If $x_0 < x_1 < \cdots < x_n = x$, then

$$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| = \sum_{i=1}^{n} |\mu([x_{i-1},x_i))| \le |\mu|((-\infty,x))$$

so that

$$(4) \qquad T_f(x) \le |\mu|((-\infty,x)).$$

This proves (a).

In the proof of (b), let us first assume that $f \in NBV$ and f is non-decreasing, $f \not\equiv 0$. Associate with each point $x \in R^1$ a set $\Phi[x]$, as follows: If f is continuous at x, $\Phi[x]$ is the point $f(x)$; if $f(x+) > f(x)$, then $\Phi[x]$ is the interval $[f(x),f(x+)]$. If $E \subset R^1$, let $\Phi[E]$ be the union of all sets $\Phi[x]$, for $x \in E$. We claim that the definition

$$(5) \qquad \mu(E) = m(\Phi[E])$$

gives us a measure which satisfies (1); here m is Lebesgue measure on R^1.

Put $J = \Phi[R^1]$. Then J is a 1-cell (i.e., a bounded interval with or without its end points). There are at most countably many points $y_i \in J$ such that $f^{-1}(y)$ consists of more than one point; for these y_i, $f^{-1}(y_i)$ is a 1-cell.

Let Σ be the class of all $E \subset R^1$ such that $\Phi[E]$ is a Borel set. If E is a 1-cell, so is $\Phi[E]$, hence $E \in \Sigma$. For any $E \subset R^1$, $\Phi[E^c]$ is the union of $J - \Phi[E]$ plus an at most countable set (a subset of $\{y_i\}$); thus $E \in \Sigma$ implies $E^c \in \Sigma$. Next,

$$\Phi[E_1 \cup E_2 \cup \cdots] = \Phi[E_1] \cup \Phi[E_2] \cup \cdots .$$

This proves that Σ is a σ-algebra which contains all segments, hence all Borel sets, so $m(\Phi[E])$ is defined for all Borel sets E. Moreover, μ is countably additive, for if $\{E_i\}$ is a disjoint collection of Borel sets in R^1, then $\{\Phi[E_i]\}$ is disjoint, except for our at most countable set $\{y_i\}$, and this does not affect the countable additivity of μ since $m(E) = 0$ for every countable set E.

Thus (5) defines a Borel measure. Since $\Phi[(-\infty,x)]$ is a 1-cell whose end points are 0 and $f(x)$, (5) shows that (1) holds.

We now turn to the general case of (*b*). If $f \in NBV$, then $f = u + iv$, u and v real, $u \in NBV$, and $T_u \in NBV$ by Theorem 8.13(*c*). Put

(6) $$u_1 = \tfrac{1}{2}(T_u + u), \qquad u_2 = \tfrac{1}{2}(T_u - u).$$

Then u_1 and $u_2 \in NBV$, and they are nondecreasing; this follows easily from Theorem 8.13(*a*). The preceding construction associates measures μ_1 and μ_2 with u_1 and u_2, and $\mu_1 - \mu_2$ will be associated with $u = u_1 - u_2$. If we deal similarly with v and combine the results, we obtain a measure μ which corresponds to f in the sense that (1) holds.

If two regular measures (note Theorem 2.18) coincide on all segments of the form $(-\infty,x)$, they coincide on all 1-cells of the form $[\alpha,\beta)$, hence on all open sets, hence on all Borel sets. This proves the uniqueness assertion of (*b*).

Finally, let λ be the measure associated with T_f in the same way. If $\alpha < \beta$, then

(7) $$\mu([\alpha,\beta)) = f(\beta) - f(\alpha), \qquad \lambda([\alpha,\beta)) = T_f(\beta) - T_f(\alpha).$$

The inequality

(8) $$|\mu(E)| \le \lambda(E)$$

therefore holds if $E = [\alpha,\beta)$. Since every open set in R^1 is a countable disjoint union of such 1-cells, (8) holds for every open set, hence for every Borel set. The definition of the total variation $|\mu|$ of μ now implies that $|\mu| \le \lambda$. In particular,

(9) $$|\mu|((-\infty,x)) \le \lambda((-\infty,x)) = T_f(x).$$

Now (2) follows from (4) and (9).

The proof of (*c*) is left as an exercise.

Differentiation of Point Functions

8.15 Absolutely Continuous Functions A complex function f on R^1 is said to be *absolutely continuous* if to every $\epsilon > 0$ there corresponds a $\delta > 0$ such that

(1) $$\sum_{i=1}^{N} (\beta_i - \alpha_i) < \delta \quad \text{implies} \quad \sum_{i=1}^{N} |f(\beta_i) - f(\alpha_i)| < \epsilon,$$

whenever $(\alpha_1,\beta_1), \ldots, (\alpha_N,\beta_N)$ are disjoint segments.

Observe that every absolutely continuous function is uniformly continuous (take $N = 1$) and that the restriction of any absolutely continuous function to a bounded interval is of bounded variation. However, if $f(x) = \sin x$, or if $f(x) = x + |x|$, then f is absolutely continuous, but $f \notin BV$.

The two meanings of the term "absolutely continuous" are related as follows:

8.16 Theorem *Suppose $f \in NBV$ and μ is associated with f as in Theorem 8.14. Then $\mu \ll m$ if and only if f is absolutely continuous.*

(Here m denotes Lebesgue measure on R^1.)

PROOF Suppose f is absolutely continuous. Let E be a Borel set such that $m(E) = 0$, choose $\epsilon > 0$, and choose $\delta > 0$ in accordance with Sec. 8.15. The regularity of μ shows that there are open sets $W_1 \supset W_2 \supset \cdots \supset E$ such that $m(W_1) < \delta$ and $\mu(W_n) \to \mu(E)$ as $n \to \infty$. Since W_n is a disjoint union of segments $I_j = (\alpha_j,\beta_j)$, and $\Sigma(\beta_j - \alpha_j) < \delta$, it follows that

$$|\mu(W_n)| \le \sum_j |\mu(I_j)| = \sum_j |f(\beta_j) - f(\alpha_j)| \le \epsilon.$$

Consequently, $\mu(E) = 0$. This proves that $\mu \ll m$.

The converse follows from Theorem 6.11.

We are now in a position to translate our earlier results on differentiation of set functions into point-function language. Theorems 8.17 to 8.19 are classical results, due to Lebesgue. Theorems 8.18 and 8.21 generalize the fundamental theorem of calculus.

8.17 Theorem *If $g \in L^1(R^1)$, and if*

(1) $$f(x) = \int_{-\infty}^{x} g(t)\,dt \qquad (-\infty < x < \infty),$$

then $f \in NBV$, f is absolutely continuous, and

(2) $$f'(x) = g(x) \text{ a.e. } [m].$$

PROOF Define

$$(3) \qquad \mu(E) = \int_E g(t)\,dt$$

for every Borel set E. Then $f(x) = \mu((-\infty,x))$ and $\mu \ll m$. By Theorem 8.14, $f \in NBV$; by Theorem 8.16, f is absolutely continuous; Theorems 8.1 and 8.6 imply that

$$(4) \qquad f'(x) = (D\mu)(x) = g(x) \text{ a.e. } [m]$$

if $D\mu$ is computed relative to the family of all open segments in R^1.

8.18 Theorem *If $f \in NBV$, then f is differentiable a.e. $[m]$, $f' \in L^1(R^1)$, and there is a function $f_s \in NBV$ with $f_s'(x) = 0$ a.e. $[m]$ such that*

$$(1) \qquad f(x) = f_s(x) + \int_{-\infty}^{x} f'(t)\,dt \qquad (-\infty < x < \infty);$$

$f_s = 0$ if and only if f is absolutely continuous; if f is nondecreasing, so is f_s.

We call f_s the *singular part* of f. It is a perhaps unexpected fact that there exist continuous singular functions which are not constant. Examples are given in Sec. 8.20(b). The word "singular" as applied to measures has its origin in this phenomenon.

PROOF By Theorem 8.14 there is a complex measure μ on R^1 such that $\mu((-\infty,x)) = f(x)$. By Theorem 8.6,

$$(2) \qquad \mu(E) = \mu_s(E) + \int_E (D\mu)(t)\,dt,$$

where $D\mu$ is computed relative to the open segments in R^1. Put

$$(3) \qquad f_s(x) = \mu_s((-\infty,x)) \qquad (-\infty < x < \infty).$$

Theorems 8.6 and 8.1 show that $f_s'(x) = 0$ a.e. $[m]$ and that

$$f'(x) = (D\mu)(x) \text{ a.e. } [m].$$

Hence (1) follows from (2) if we take $E = (-\infty,x)$.

By Theorem 8.16, f is absolutely continuous if and only if $\mu \ll m$, i.e., if and only if $\mu_s = 0$.

Finally, if f is nondecreasing, then $\mu \geq 0$, hence $\mu_s \geq 0$, hence f_s is nondecreasing.

8.19 Theorem *If $f \in BV$, then f is differentiable a.e. $[m]$, and $f' \in L^1(R^1)$.*

PROOF By Theorem 8.13, there exists a $g \in NBV$ such that

$$f(x) = g(x) + c$$

at all points of continuity of f. Theorem 8.18 applies to g. Hence the following lemma (with $h = f - g$) implies the theorem:

Lemma *If $h \in BV$ and $h(x) = 0$ except on an at most countable set, then $h'(x) = 0$ a.e.*

To prove the lemma, let $S = \{x_i\}$ be the at most countable set at which $h(x_i) = c_i \neq 0$. Since $h \in BV$, it is easily seen that $\Sigma|c_i| < \infty$. Fix k, and let A_k be the set of all $x \notin S$ at which

$$(1) \qquad \left|\frac{h(y) - h(x)}{y - x}\right| > \frac{1}{k}$$

for infinitely many y. Thus $x \in A_k$ if and only if $|x - x_i| < k|c_i|$ for infinitely many i. If J_i is the segment with center at x_i and length $2k|c_i|$, it follows that

$$(2) \qquad \sum m(J_i) = 2k \sum |c_i| < \infty,$$

and hence $m(A_k) = 0$ for $k = 1, 2, 3, \ldots$, by Theorem 1.41.

But if $x \notin S \cup A_1 \cup A_2 \cup A_3 \cup \cdots$, then $h'(x) = 0$. This completes the proof.

Exercises 5 and 6 are relevant to this lemma.

8.20 Examples The preceding theorems show that the equation

$$(1) \qquad f(x) - f(a) = \int_a^x f'(t)\,dt$$

(in which the right side is a Lebesgue integral) holds for all x in some interval $[a,b]$ if and only if f is absolutely continuous on $[a,b]$. One may ask whether the existence of f' implies the absolute continuity of f. Put this way, the question is not precise enough. We shall give two examples which show how (1) can fail, and then give a theorem in which (1) is deduced from another set of sufficient conditions.

(a) Put $f(x) = x^2 \sin(x^{-2})$ if $x \neq 0$, $f(0) = 0$. Then f is differentiable at every point, but

$$(2) \qquad \int_0^1 |f'(t)|\,dt = \infty,$$

so $f' \notin L^1$. Also, $f \notin BV$ on $[0,1]$.

If we interpret the integral in (1) (with $[0,1]$ in place of $[a,b]$) as the limit, as $\epsilon \to 0$, of the integrals over $[\epsilon,1]$, then (1) still holds for this f.

More complicated situations can arise where this kind of passage to the limit is of no use. There are integration processes, due to Denjoy and Perron (see [18], [28]), which are so designed that (1) holds whenever f is differentiable at every point. These fail to have the property that the integrability of f implies that of $|f|$, and therefore do not play such an important role in analysis.

(b) Suppose f is continuous on $[a,b]$, f is differentiable at almost every point of $[a,b]$, and $f' \in L^1$ on $[a,b]$. Do these assumptions imply that (1) holds?

Answer: No.

Choose $\{\delta_n\}$ so that $1 = \delta_0 > \delta_1 > \delta_2 > \cdots$, $\delta_n \to 0$. Put $E_0 = [0,1]$. Suppose $n \geq 0$ and E_n is constructed so that E_n is the union of 2^n disjoint closed intervals, each of length $2^{-n}\delta_n$. Delete a segment in the center of each of these 2^n intervals, so that each of the remaining 2^{n+1} intervals has length $2^{-n-1}\delta_{n+1}$ (this is possible, since $\delta_{n+1} < \delta_n$), and let E_{n+1} be the union of these 2^{n+1} intervals. Then $E_1 \supset E_2 \supset \cdots$, $m(E_n) = \delta_n$, and if

$$E = \bigcap_{n=1}^{\infty} E_n, \tag{3}$$

then E is compact and $m(E) = 0$. (In fact, E is perfect.) Put

$$g_n = \delta_n^{-1}\chi_{E_n} \quad \text{and} \quad f_n(x) = \int_0^x g_n(t)\,dt \qquad (n = 0, 1, 2, \ldots). \tag{4}$$

Then $f_n(0) = 0$, $f_n(1) = 1$, and each f_n is a monotonic function which is constant on each segment in the complement of E_n. If I is one of the 2^n intervals whose union is E_n, then

$$\int_I g_n(t)\,dt = \int_I g_{n+1}(t)\,dt = 2^{-n}. \tag{5}$$

It follows from (5) that

$$f_{n+1}(x) = f_n(x) \qquad (x \notin E_n) \tag{6}$$

and that

$$|f_n(x) - f_{n+1}(x)| \leq \int_I |g_n - g_{n+1}| < 2^{-n+1} \qquad (x \in E_n). \tag{7}$$

Hence $\{f_n\}$ converges uniformly to a continuous monotonic function f, with $f(0) = 0$, $f(1) = 1$, and $f'(x) = 0$ for all $x \notin E$. Since $m(E) = 0$, we have $f' = 0$ a.e.

Thus (1) fails. Incidentally, we have now constructed examples of continuous singular functions as defined after the statement of Theorem 8.18.

If $\delta_n = (2/3)^n$, the set E is Cantor's "middle thirds" set.

8.21 Theorem *Suppose f is a real function on $[a,b]$ which is differentiable at every point of $[a,b]$, and assume that $f' \in L^1$ on $[a,b]$. Then*

$$f(x) - f(a) = \int_a^x f'(t)\,dt \qquad (a \leq x \leq b). \tag{1}$$

Note that differentiability is assumed to hold at *every* point of $[a,b]$.

PROOF It is clear that it is enough to prove this for $x = b$. Fix $\epsilon > 0$. Theorem 2.24 ensures the existence of a lower semicontinuous function g on $[a,b]$ such that $g > f'$ and

$$\int_a^b g(t)\,dt < \int_a^b f'(t)\,dt + \epsilon. \tag{2}$$

Actually, Theorem 2.24 only gives $g \geq f'$, but since $m([a,b]) < \infty$, we can add a small constant to g without affecting (2). For any $\eta > 0$, define

$$F_\eta(x) = \int_a^x g(t)\,dt - f(x) + f(a) + \eta(x-a) \qquad (a \leq x \leq b). \tag{3}$$

Keep η fixed for the moment. To each $x \in [a,b)$ there corresponds a $\delta_x > 0$ such that

$$g(t) > f'(x) \quad \text{and} \quad \frac{f(t) - f(x)}{t - x} < f'(x) + \eta \tag{4}$$

for all $t \in (x, x + \delta_x)$, since g is lower semicontinuous and $g(x) > f'(x)$. For any such t we therefore have

$$\begin{aligned} F_\eta(t) - F_\eta(x) &= \int_x^t g(s)\,ds - [f(t) - f(x)] + \eta(t - x) \\ &> (t-x)f'(x) - (t-x)[f'(x) + \eta] + \eta(t-x) = 0. \end{aligned}$$

Since $F_\eta(a) = 0$ and F_η is *continuous*, there is a last point $x \in [a,b]$ at which $F_\eta(x) = 0$. If $x < b$, the preceding computation implies that $F_\eta(t) > 0$ for $t \in (x,b]$. In any case, $F_\eta(b) \geq 0$. Since this holds for every $\eta > 0$, (2) and (3) now give

$$f(b) - f(a) \leq \int_a^b g(t)\,dt < \int_a^b f'(t)\,dt + \epsilon, \tag{5}$$

and since ϵ was arbitrary, we conclude that

$$f(b) - f(a) \leq \int_a^b f'(t)\,dt. \tag{6}$$

If f satisfies the hypotheses of the theorem, so does $-f$; therefore (6) holds with $-f$ in place of f, and these two inequalities together give (1).

Differentiable Transformations

8.22 Definitions With any $x = (\xi_1, \ldots, \xi_k) \in R^k$ we associate the norm

$$\|x\| = \max(|\xi_1|, \ldots, |\xi_k|). \tag{1}$$

This norm is better adapted for dealing with cubes than is the ordinary euclidean norm

$$\|x\|_2 = (\xi_1^2 + \cdots + \xi_k^2)^{\frac{1}{2}}. \tag{2}$$