Monday, January 14, 2013

1 Real Numbers

There are two ways to introduce the real numbers. The first is to give them in an axiomatic way, the second is to construct them starting from the natural numbers. We will use the first method.

The *real numbers* are a set $\mathbb{R}$ with two binary operations, *addition*

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$
$$(x, y) \mapsto x + y$$

and *multiplication*

$$\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$
$$(x, y) \mapsto x \cdot y$$

and a relation $\leq$ such that $(\mathbb{R}, +, ., \leq)$ is an *ordered field*, that is,

(A) $(\mathbb{R}, +)$ is an *commutative group*, that is,

- (A_1) for every $a, b \in \mathbb{R}$, $a + b = b + a$,
- (A_2) for every $a, b, c \in \mathbb{R}$, $(a + b) + c = a + (b + c)$,
- (A_3) there exists a unique element in $\mathbb{R}$, called *zero* and denoted 0, such that $0 + a = a + 0 = a$ for every $a \in \mathbb{R}$,
- (A_4) for every $a \in \mathbb{R}$ there exists a unique element in $\mathbb{R}$, called the *opposite* of a and denoted $-a$, such that $(-a) + a = a + (-a) = 0$,

(M)

- (M_1) for every $a, b \in \mathbb{R}$, $a \cdot b = b \cdot a$,
- (M_2) for every $a, b, c \in \mathbb{R}$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$,
- (M_3) there exists a unique element in $\mathbb{R}$, called *one* and denoted 1, such that $1 \neq 0$ and $1 \cdot a = a \cdot 1 = a$ for every $a \in \mathbb{R}$ with $a \neq 0$,
- (M_4) for every $a \in \mathbb{R}$ with $a \neq 0$ there exists a unique element in $\mathbb{R}$, called the *inverse* of a and denoted a^{-1}, such that $a^{-1} \cdot a = a \cdot a^{-1} = 1$,

(O) $\leq$ is a *total order relation*, that is,

- (O_1) for every $a, b \in \mathbb{R}$ either $a \leq b$ or $b \leq a$,
- (O_2) for every $a, b, c \in \mathbb{R}$ if $a \leq b$ and $b \leq c$, then $a \leq c$,
- (O_3) for every $a, b \in \mathbb{R}$ if $a \leq b$ and $b \leq a$, then $a = b$,
- (O_4) for every $a \in \mathbb{R}$ we have $a \leq a$,

(AM) for every $a, b, c \in \mathbb{R}$, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$,

(AO) for every $a, b, c \in \mathbb{R}$ if $a \leq b$, $a + c \leq b + c$,

(MO) for every $a, b \in \mathbb{R}$ if $0 \leq a$ and $0 \leq b$, then $0 \leq a \cdot b$,

(S) **(supremum property)**

Remark 1 *Properties* (A), (M), (O), (AM), (AO), (MO), *and* (S) *completely characterize the real numbers in the sense that if* $(\mathbb{R}', \oplus, \odot, \preccurlyeq)$ *satisfies the same properties, then there exists a bijection* $T : \mathbb{R} \to \mathbb{R}'$ *such that* T *is an isomorphism between the two fields, that is,*

$$T(a + b) = T(a) \oplus T(b), \qquad T(a \cdot b) = T(a) \odot T(b)$$

for all $a, b \in \mathbb{R}$, *and* $a \leq b$ *if and only if* $T(a) \preccurlyeq T(b)$. *Hence, for all practical purposes, we cannot distinguish* $\mathbb{R}$ *from* $\mathbb{R}'$.

If $a \leq b$ and $a \neq b$, we write $a < b$.

Exercise 2 *Using **only** the axioms* (A), (M), (O), (AO), (AM) *and* (MO) *of* $\mathbb{R}$, *prove the following properties of* $\mathbb{R}$:

(i) *if* $a \cdot b = 0$ *then either* $a = 0$ *or* $b = 0$,

(ii) *if* $a \geq 0$ *then* $-a \leq 0$,

(iii) *if* $a \leq b$ *and* $c < 0$ *then* $ac \geq bc$,

(iv) *for every* $a \in \mathbb{R}$ *we have* $a^2 \geq 0$,

(v) $1 > 0$.

Definition 3 *Let* $E \subseteq \mathbb{R}$ *be a nonempty set.*

(i) *An element* $L \in \mathbb{R}$ *is called an* upper bound *of* E *if* $x \leq L$ *for all* $x \in E$;

(ii) E *is said to be* bounded from above *if it has at least an upper bound;*

(iii) *if* E *is bounded from above, the least of all its upper bounds, if it exists, is called the* supremum *of* E *and is denoted* $\sup E$.

(iv) E *has a* maximum *if there exists* $L \in E$ *such that* $x \leq L$ *for all* $x \in E$. *We write* $L = \max E$.

We are now ready to state the supremum property.

(S) **(supremum property)** every nonempty set $E \subseteq \mathbb{R}$ bounded from above has a supremum in $\mathbb{R}$.

The supremum property says that in $\mathbb{R}$ the supremum of a nonempty set bounded from above always exists in $\mathbb{R}$.

Remark 4 *(i) Note that if a set has a maximum L, then L is also the supremum of the set.*

(ii) If $E \subseteq \mathbb{R}$ is a set bounded from below, to prove that a number $L \in \mathbb{R}$ is the supremum of E, we need to show that L is an upper bound of E, that is, that $x \leq L$ for every $x \in E$, and that any number $s < L$ cannot be an upper bound of E, that is, that there exists $x \in E$ such that $s < x$.

Definition 5 *Let $E \subseteq \mathbb{R}$ be a nonempty set.*

(i) An element $\ell \in \mathbb{R}$ is called a lower bound *of E if $\ell \leq x$ for all $x \in E$;*

(ii) E is said to be bounded from below *if it has at least an lower bound;*

(iii) if E is bounded from above, the greatest of all its lower bounds, if it exists, is called the infimum *of E and is denoted* $\inf E$*;*

(iv) E has a minimum *if there exists $\ell \in E$ such that $\ell \leq x$ for all $x \in E$. We write $\ell = \min E$.*

Remark 6 *(i) Note that if a set has a minimum ℓ, then ℓ is also the infimum of the set.*

(ii) If $E \subseteq \mathbb{R}$ is a set bounded from above, to prove that a number $\ell \in \mathbb{R}$ is the infimum of E, we need to show that ℓ is a lower bound of E, that is, that $\ell \leq x$ for every $x \in E$, and that any number $\ell < s$ cannot be a lower bound of E, that is, that there exists $x \in E$ such that $x < s$.

Example 7 *Consider the set*

$$E = \left\{ y \in \mathbb{R} : \, y = \arctan\left(\left|x^2 - 1\right| + 1\right), \, x \in \mathbb{R} \right\}.$$

Since $-\frac{\pi}{2} \leq \arctan t \leq \frac{\pi}{2}$ for all $t \in \mathbb{R}$, the set E is bounded. Let's prove that the minimum of E is

$$\inf E = \arctan 1 = \frac{1}{4}\pi.$$

Since

$$\frac{d}{dt}(\arctan t) = \frac{1}{t^2 + 1} > 0,$$

the function $\arctan t$ *is strictly increasing. If $x \neq \pm 1$, we have that $\left|x^2 - 1\right| + 1 > 1$, and so* $\arctan\left(\left|x^2 - 1\right| + 1\right) > \arctan 1$*, while if $x = \pm 1$, $y = \arctan(0 + 1) = \frac{\pi}{4}$. Hence, $\min E = \frac{1}{4}\pi$. To find the supremum note that since $\arctan t \leq \frac{\pi}{2}$ for all $t \in \mathbb{R}$, we have that $\frac{\pi}{2}$ is an upper bound of the set E. To prove that it is the supremum, we need to show that any $s < \frac{\pi}{2}$ is not an upper bound of E. Since*

$$\lim_{x \to \infty} \arctan\left(\left|x^2 - 1\right| + 1\right) = \frac{\pi}{2},$$

for every $\varepsilon > 0$ we can find $M > 0$ such that

$$\left|\arctan\left(\left|x^2 - 1\right| + 1\right) - \frac{\pi}{2}\right| < \varepsilon$$

for all $x > M$, that is,

$$\frac{\pi}{2} - \varepsilon < \arctan\left(\left|x^2 - 1\right| + 1\right) < \frac{\pi}{2} + \varepsilon$$

for all $x > M$. Taking $0 < \varepsilon < \frac{\pi}{2} - s$, we have that $\frac{\pi}{2} - \varepsilon > s$, and so

$$s < \frac{\pi}{2} - \varepsilon < \arctan\left(\left|x^2 - 1\right| + 1\right)$$

for all $x > M$, which shows that s is not an upper bound of E. Hence, $\frac{\pi}{2}$ is the least upper bound of E. This implies that

$$\sup E = \frac{\pi}{2}.$$

Wednesday, January 16, 2013

Example 8 *Consider the set*

$$E = \left\{ t \in \mathbb{R} : \, t = \frac{xy}{x^2 + y^2}, \, x, y \in \mathbb{R}, \, x < y \right\}.$$

The set E is bounded since $-\frac{1}{2} \leq \frac{xy}{x^2+y^2} \leq \frac{1}{2}$ for all $x, y \in \mathbb{R}$, with $x < y$. Moreover, by taking $x = -1$ and $y = 1$, we get $t = \frac{-1}{2}$, so

$$\inf E = \min E = -\frac{1}{2}.$$

Let's prove that

$$\sup E = \frac{1}{2}.$$

We need to show that any $s < \frac{1}{2}$ is not an upper bound for the set E. If $s \leq -\frac{1}{2}$, then we can take $x = -1$ and $y = 0$, so that $s < t = 0$. Thus assume that $-\frac{1}{2} < s < \frac{1}{2}$. Take the $y = 1$. Since

$$\lim_{x \to 1^-} \frac{x}{x^2 + 1} = \frac{1}{2},$$

for every $\varepsilon > 0$ we can find $\delta > 0$ such that

$$\left| \frac{x}{x^2 + 1} - \frac{1}{2} \right| < \varepsilon$$

for all $1 - \delta < x < 1$, that is,

$$\frac{1}{2} - \varepsilon < \frac{x}{x^2 + 1} < \frac{1}{2} + \varepsilon$$

for all $1 - \delta < x < 1$. Taking $0 < \varepsilon < \frac{1}{2} - s$, we have that $\frac{1}{2} - \varepsilon > s$, and so

$$s < \frac{1}{2} - \varepsilon < \frac{x}{x^2 + 1}$$

for all $1 - \delta < x < 1$, which shows that s is not an upper bound of E. Thus, $\frac{1}{2}$ is the supremum of the set. Note that $t = \frac{xy}{x^2+y^2} = \frac{1}{2}$ only if $x = y$, which is not allowed, so the set does not have a maximum.

2 The Euclidean Space

Definition 9 *A* vector space*, or* linear space*, over* $\mathbb{R}$ *is a nonempty set* X*, whose elements are called* vectors*, together with two operations,* addition *and* multiplication by scalars*,*

$$\begin{array}{ccc} X \times X \to X & & \mathbb{R} \times X \to X \\ (x,y) \mapsto x+y & \text{and} & (t,x) \mapsto tx \end{array}$$

with the properties that

(i) $(X,+)$ *is a commutative group, that is,*

(a) $x+y=y+x$ *for all* $x,y \in X$ *(commutative property),*

(b) $x+(y+z)=(x+y)+z$ *for all* $x,y,z \in X$ *(associative property),*

(c) there is a vector $0 \in X$*, called* zero*, such that* $x+0=0+x$ *for all* $x \in X$*,*

(d) for every $x \in X$ *there exists a vector in* X*, called the* opposite *of* x *and denoted* $-x$*, such that* $x+(-x)=0$*,*

(ii) for all $x,y \in X$ *and* $s,t \in \mathbb{R}$*,*

(a) $s\,(tx)=(st)\,x$*,*

(b) $1x=x$*,*

(c) $s\,(x+y)=(sx)+(sy)$*,*

(d) $(s+t)x=(sx)+(tx)$*.*

Remark 10 *Instead of using real numbers, one can use* $\mathbb{C}$ *or a field* F*. For most or our purposes the real numbers will suffice.* From now on, whenever we don't specify, it is understood that a vector space is over $\mathbb{R}$.

Example 11 *Some important examples of vector spaces over* $\mathbb{R}$ *are the following.*

(i) The Euclidean space $\mathbb{R}^N$ *is the space of all* N*-tuples* $\mathbf{x}=(x_1,\ldots,x_N)$ *of real numbers. The elements of* $\mathbb{R}^N$ *are called* vectors *or* points*. The Euclidean space is a vector space with the following operations*

$$\mathbf{x}+\mathbf{y} := (x_1+y_1,\ldots,x_N+y_N)\,, \quad t\mathbf{x} := (tx_1,\ldots,tx_N)$$

for every $t \in \mathbb{R}$ *and* $\mathbf{x}=(x_1,\ldots,x_N)$ *and* $\mathbf{y}=(y_1,\ldots,y_N)$ *in* $\mathbb{R}^N$*.*

(ii) The collection of all polynomials $p:\mathbb{R}\to\mathbb{R}$*.*

(iii) The space of continuous functions $f:[a,b]\to\mathbb{R}$*.*

Definition 12 *An* inner product*, or* scalar product*, on a vector space X is a function*

$$(\cdot,\cdot) : X \times X \to \mathbb{R}$$

such that

(i) $(x,x) \geq 0$ *for every* $x \in X$*,* $(x,x) = 0$ *if and only if* $x = 0$ *(positivity);*

(ii) $(x,y) = (y,x)$ *for all* $x, y \in X$ *(symmetry);*

(iii) $(sx + ty, z) = s\,(x,z) + t\,(y,z)$ *for all* $x, y, z \in X$ *and* $s, t \in \mathbb{R}$ *(bilinearity).*

Remark 13 *If X is a vector space over $\mathbb{C}$, then an inner product is a function* $(\cdot,\cdot) : X \times X \to \mathbb{C}$ *satisfying properties (i),*

(ii)′ $(x,y) = \overline{(y,x)}$ *for all* $x, y \in X$ *(skew-symmetry), where given* $z \in \mathbb{C}$*, the number* $\overline{z}$ *is the complex conjugate;*

(iii) $(sx + ty, z) = s\,(x,z) + t\,(y,z)$ *for all* $x, y, z \in X$ *and* $s, t \in \mathbb{C}$ *(bilinearity).*

Example 14 *Some important examples of inner products are the following.*

(i) Consider the Euclidean space $\mathbb{R}^N$*, then*

$$\mathbf{x} \cdot \mathbf{y} := x_1 y_1 + \cdots + x_N y_N,$$

where $\mathbf{x} = (x_1, \ldots, x_N)$ *and* $\mathbf{y} = (y_1, \ldots, y_N)$*, is an inner product.*

(ii) Consider the space of continuous functions $f : [a,b] \to \mathbb{R}$*. Then*

$$(f,g)_{L^2([a,b])} := \int_a^b f(x)\, g(x)\, dx$$

is an inner product.

Friday, January 18, 2013

Solutions of the test.

Monday, January 21, 2013

Martin Luther King's Day. No class

Wednesday, January 23, 2013

Definition 15 *A* norm *on a vector space X is a map*

$$\|\cdot\| : X \to [0,\infty)$$

such that

(i) $\|x\| = 0$ *implies* $x = 0$*;*

(ii) $\|tx\| = |t|\,\|x\|$ *for all* $x \in X$ *and* $t \in \mathbb{R}$*;*

(iii) $\|x + y\| \leq \|x\| + \|y\|$ *for all* $x, y \in X$*.*

A *normed space* $(X, \|\cdot\|)$ is a vector space X endowed with a norm $\|\cdot\|$. For simplicity, we often say that X is a normed space.

Given an inner product $(\cdot,\cdot) : X \times X \to \mathbb{R}$ on a vector space X, it turns out that the function
$$\|x\| := \sqrt{(x,x)}, \quad x \in X, \tag{1}$$
is a norm. This follows from the following result.

Proposition 16 (Cauchy–Schwarz's inequality) *Given an inner product* $(\cdot,\cdot) : X \times X \to \mathbb{R}$ *on a vector space* X*,*

$$|(x,y)| \leq \|x\| \, \|y\|$$

for all $x, y \in X$*.*

Proof. If $y = 0$, then both sides of the previous inequality are zeros, and so there is nothing to prove. Thus, assume that $y \neq 0$ and let $t \in \mathbb{R}$. By properties (i)-(iii),
$$0 \leq (x + ty, x + ty) = \|x\|^2 + t^2 \|y\|^2 + 2t\,(x,y)\,. \tag{2}$$
Taking
$$t := -\frac{(x,y)}{\|y\|^2}$$
in the previous inequality gives

$$0 \leq \|x\|^2 + \frac{(x,y)^2}{\|y\|^4} \|y\|^2 - 2\frac{(x,y)}{\|y\|^2},$$

or, equivalently,
$$(x,y)^2 \leq \|x\|^2 \|y\|^2 \,.$$
It now suffices to take the square root on both sides. ■

Remark 17 *It follows from the proof that equality holds in the Cauchy–Schwarz inequality if and only you have equality in (2), that is, if* $x + ty = 0$ *for some* $t \in \mathbb{R}$ *or* $y = 0$*.*

Corollary 18 *Given a scalar product* $(\cdot,\cdot) : X \times X \to \mathbb{R}$ *on a vector space* X*, the function*
$$\|x\| := \sqrt{(x,x)}, \quad x \in X,$$
is a norm.

Proof. By property (i), $\|\cdot\|$ is well-defined and $\|x\| = 0$ if and only if $x = 0$. Taking $t = 1$ in (2) and using the Cauchy–Schwarz inequality gives

$$\begin{aligned} 0 \leq \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2\,(x,y) \\ &\leq \|x\|^2 + \|y\|^2 + 2\,\|x\|\,\|y\| = (\|x\| + \|y\|)^2\,, \end{aligned}$$

which is the triangle inequality for the norm. Moreover, by properties (ii) and (iii) for every $t \in \mathbb{R}$,

$$\|tx\| = \sqrt{(tx, tx)} = \sqrt{t\,(x, tx)} = \sqrt{t\,(tx, x)} = \sqrt{t^2\,(x, x)} = |t|\,\|x\|\,.$$

Thus $\|\cdot\|$ is a norm. ■

Example 19 *Other important norms that one can put in $\mathbb{R}^N$ are*

$$\begin{aligned}\|\mathbf{x}\|_{\ell^\infty} &:= \max\{|x_1|, \ldots, |x_N|\}\,,\\ \|\mathbf{x}\|_{\ell^1} &:= |x_1| + \cdots + |x_N|\,,\\ \|\mathbf{x}\|_{\ell^p} &:= (|x_1|^p + \cdots + |x_N|^p)^{1/p}\,,\end{aligned}$$

for $\mathbf{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ and where $1 \le p < \infty$.

The proof of the following theorem is left as an exercise.

Theorem 20 *Let $(X, \|\cdot\|)$ be a normed space. Then there exists an inner product $(\cdot,\cdot) : X \times X \to \mathbb{R}$ such that $\|x\| = \sqrt{(x, x)}$ for all $x \in X$ if and only if $\|\cdot\|$ satisfies the parallelogram law*

$$\|x + y\|^2 + \|x - y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2$$

for all $x, y \in X$.

Example 21 *Using the previous theorem, we can prove that some normed spaces are not inner product spaces. The space $\mathbb{R}^N$ with the norm $\|\cdot\|_{\ell^p}$ for $p \ne 2$ is not an inner product . Take $\mathbf{x} = (1, 1, 0, \ldots)$, $\mathbf{y} = (1, -1, 0, \ldots)$. Then $\mathbf{x} + \mathbf{y} = (2, 0, \ldots)$, $\mathbf{x} - \mathbf{y} = (0, 2, 0, \ldots)$. Hence,*

$$\begin{aligned}\|\mathbf{x} + \mathbf{y}\|_{\ell^p}^2 + \|\mathbf{x} - \mathbf{y}\|_{\ell^p}^2 &= (2^p)^{\frac{2}{p}} + (2^p)^{\frac{2}{p}} = 8\\ &\ne 2\,\|\mathbf{x}\|_{\ell^p}^2 + 2\,\|\mathbf{y}\|_{\ell^p}^2\\ &= 2\,(1^p + 1^p)^{\frac{2}{p}} + 2\,(1^p + 1^p)^{\frac{2}{p}} = 2^{2+\frac{2}{p}}.\end{aligned}$$

Definition 22 *Given a set E and a function $f : E \to \mathbb{R}$, we say that f is* bounded from above *if the set*

$$f(E) := \{y \in \mathbb{R} : \, y = f(x)\,,\, x \in E\}$$

is bounded from above. We say that f is bounded from below *if the set $f(E)$ is bounded from below. Finally, we say that f is* bounded *if the set $f(E)$ is bounded.*

Given a set E and a function $f : E \to \mathbb{R}$, we write

$$\sup_E f := \sup f(E)\,, \quad \inf_E f := \inf f(E)\,.$$

Exercise 23 *Given a set E, consider the vector space $X := \{f : E \to \mathbb{R}$ bounded$\}$. For $f \in X$, define*

$$\|f\| := \sup_E |f| .$$

Prove that $\|\cdot\|$ is a norm.

Definition 24 *Given a set E and functions $f_n : E \to \mathbb{R}$, $n \in \mathbb{N}$, we say that the sequence of functions $\{f_n\}$* converges pointwise *in a set $F \subseteq E$ to some function $f : F \to \mathbb{R}$ if*

$$\lim_{n\to\infty} f_n(x) = f(x)$$

for all $x \in F$. We say that the sequence of functions $\{f_n\}$ converges uniformly *in a set $F \subseteq E$ to some function $f : F \to \mathbb{R}$ if*

$$\lim_{n\to\infty} \sup_F |f_n - f| = 0.$$

Example 25 *Let $E = [0,1]$ and let $f_n(x) = x^n$, $x \in [0,1]$. Then*

$$\lim_{n\to\infty} x^n = \begin{cases} 0 & \text{if } 0 \le x < 1, \\ 1 & \text{if } x = 1, \end{cases}$$

and so the sequence $\{f_n\}$ converges pointwise in $[0,1]$ to the function f defined by

$$f(x) = \begin{cases} 0 & \text{if } 0 \le x < 1, \\ 1 & \text{if } x = 1. \end{cases}$$

On the other hand,

$$|f_n(x) - f(x)| = \begin{cases} x^n & \text{if } 0 \le x < 1, \\ 0 & \text{if } x = 1. \end{cases}$$

Fix $n \in \mathbb{N}$. We claim that

$$\sup_{[0,1]} |f_n - f| = 1.$$

To see this, note that

$$|f_n - f|([0,1]) = [0,1) .$$

Indeed, if $0 \le y < 1$, let $x := \sqrt[n]{y}$. Then $x^n = y$.

Let's prove that

$$\sup [0,1) = 1.$$

Note that 1 is an upper bound for the set $[0,1)$, and for any $s < 1$, if $s < 0$ then s is not an upper bound of $[0,1)$, while if $0 \le s < 1$, then the middle point $\frac{1+s}{2}$ belongs to $[0,1)$ and is greater than s. Thus, s is not an upper bound of $|f_n - f|([0,1]) = [0,1)$, which proves that

$$\sup_{[0,1]} |f_n - f| = 1.$$

In turn,

$$\lim_{n\to\infty} \sup_{[0,1]} |f_n - f| = \lim_{n\to\infty} 1 = 1,$$

which shows that the sequence does not converges uniformly in $[0,1]$.

Friday, January 24, 2013

Example 26 (Continued) *Now let's try to find a smaller set of* $[0,1]$ *where there is uniform convergence. Since the problem is at* 1*, let's remove a small set near* 1*. Consider* $F = [0,a]$*, where* $0 < a < 1$*. Since* $f_n - f = f_n - 0$ *is increasing in* $[0,a]$*,*

$$\begin{aligned}\sup_{[0,a]} |f_n - f| &= \max_{[0,a]} |f_n| \\ &= f_n(a) = a^n \to 0.\end{aligned}$$

Hence, there is uniform convergence in F*.*

Example 27 *Consider the sequence of functions*

$$f_n(x) = \frac{nx}{1+n^4x^4}, \quad x \in \mathbb{R}.$$

Let $x \in \mathbb{R}$*. We have that*

$$f_n(0) = 0 \to 0.$$

while if $x \neq 0$*, we have that*

$$\lim_{n\to\infty} \frac{nx}{1+n^4x^4} = \lim_{n\to\infty} \frac{n}{n^4} \frac{x}{\frac{1}{n^4}+x^4} = \lim_{n\to\infty} \frac{1}{n^3} \frac{x}{\frac{1}{n^4}+x^4} = 0 \cdot \frac{x}{0+x^4} = 0.$$

Thus, there is pointwise convergence to the function $f = 0$ *in* $\mathbb{R}$*.*

To study uniform convergence, for each fixed n *we sketch the graph of the function* $f_n - f$ *in* E*. Since* $f = 0$*, we need to sketch the graph of* f_n*. Note that*

$$f_n(-x) = -f_n(x),$$

which means that the function f_n *is odd. So it is enough to study the function for* $x \geq 0$*. We have that* $f_n(x) = \frac{nx}{1+n^4x^4} \geq 0$ *for* $x \geq 0$*. Moreover,*

$$\lim_{x\to\infty} \frac{nx}{1+n^4x^4} = \lim_{x\to\infty} \frac{x}{x^4} \frac{n}{\frac{1}{x^4}+n^4} = \lim_{x\to\infty} \frac{1}{x^3} \frac{n}{\frac{1}{x^4}+n^4} = 0 \cdot \frac{n}{0+n^4} = 0.$$

Hence the supremum of $|f_n|$ *will be in the interior. To find it, let's calculate*

$$\begin{aligned}f_n'(x) &= \frac{n1\left(1+n^4x^4\right) - nx\left(0+4n^4x^3\right)}{\left(1+n^4x^4\right)^2} \\ &= \frac{n+n^5x^4-4n^5x^4}{\left(1+n^4x^4\right)^2} = \frac{n-3n^5x^4}{\left(1+n^4x^4\right)^2} \geq 0\end{aligned}$$

for $n - 3n^5x^4 \geq 0$*. Thus* $f_n'(x) \geq 0$ *for* $x^4 \leq \frac{1}{3n^4}$*, that is, for* $0 \leq x \leq \frac{1}{n\sqrt[4]{3}}$*. Hence,* f_n *is increasing in* $0 \leq x \leq \frac{1}{n\sqrt[4]{3}}$ *and decreasing for* $x > \frac{1}{n\sqrt[4]{3}}$*. It follows* *that* f_n *has a maximum at the point* $x = \frac{1}{n\sqrt[4]{3}}$. *It follows that*

$$\begin{aligned}\sup_{x\in\mathbb{R}} |f_n(x) - f(x)| &= \sup_{x\in\mathbb{R}} \left| \frac{nx}{1+n^4x^4} - 0 \right| = \max_{x\in\mathbb{R}} \left| \frac{nx}{1+n^4x^4} \right| \\ &= f_n\left(\frac{1}{n\sqrt[4]{3}}\right) = \frac{n\frac{1}{n\sqrt[4]{3}}}{1+n^4\left(\frac{1}{n\sqrt[4]{3}}\right)^4} \\ &= \frac{\frac{1}{\sqrt[4]{3}}}{1+\frac{1}{3}} \nrightarrow 0.\end{aligned}$$

Hence, there is no uniform convergence in $\mathbb{R}$.

Since $x = \frac{1}{n\sqrt[4]{3}} \to 0$, *the problem is near zero. However if we remove only* 0, *we have that*

$$\begin{aligned}\sup_{x\in\mathbb{R}\setminus\{0\}} |f_n(x) - f(x)| &= \max_{x\in\mathbb{R}\setminus\{0\}} \left| \frac{nx}{1+n^4x^4} \right| \\ &= f_n\left(\frac{1}{n\sqrt[4]{3}}\right) \\ &= \frac{\frac{1}{\sqrt[4]{3}}}{1+\frac{1}{3}} \nrightarrow 0.\end{aligned}$$

So in $\mathbb{R}\setminus\{0\}$, *we still do not have uniform convergence. Let's remove a small set near* 0. *Consider* $F = \mathbb{R}\setminus(-a,a)$. *Since* $a > 0$ *and* $\frac{1}{n\sqrt[4]{3}} \to 0$, *there exists* n_1 *such that* $0 < \frac{1}{n\sqrt[4]{3}} < a$ *for all* $n \geq n_1$. *It follows that* f_n *is decreasing for* $x \geq a$. *In turn,*

$$\begin{aligned}\sup_{x\in F} |f_n(x) - f(x)| &= \max_{x\in F} \left| \frac{nx}{1+n^4x^4} \right| \\ &= f_n(a) \\ &= \frac{na}{1+n^4a^4} \\ &= \frac{1}{n^4}\frac{a}{\frac{1}{n^4}+a^4} \to 0.\end{aligned}$$

This shows that there is uniform convergence in F.

Definition 28 *A* metric *on a set* X *is a map* $d : X \times X \to [0,\infty)$ *such that*

(i) $d(x,y) = 0$ *if and only if* $x = y$,

(ii) $d(x,y) = d(y,x)$ *for all* $x, y \in X$ *(symmetry),*

(iii) $d(x,y) \leq d(x,z) + d(z,y)$ *for all* $x, y, z \in X$ *(triangle inequality).*

A metric space (X,d) *is a set* X *endowed with a metric* d. *When there is no possibility of confusion, we abbreviate by saying that* X *is a metric space.*

Proposition 29 *Let* $(X, \|\cdot\|)$ *be a normed space. Then*

$$d(x, y) := \|x - y\|$$

is a metric.

Proof. By property (i) in Definition 15, we have that $0 = d(x, y) = \|x - y\|$ if and only if $x - y = 0$, that is, $x = y$.

By property (ii) in Definition 15, we obtain that

$$d(y, x) = \|y - x\| = \|-1(x - y)\| = |-1| \cdot \|x - y\| = \|x - y\| = d(x, y).$$

Finally, by property (ii) in Definition 15,

$$d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y).$$

■

Exercise 30 *Prove that in* $\mathbb{R}$ *the function*

$$d_1(x, y) := \left| \frac{x}{1 + |x|} - \frac{y}{1 + |y|} \right| \tag{3}$$

is a metric.

Monday, January 28, 2013

3 Topological Properties of the Euclidean Space

Definition 31 *Given a point* $\mathbf{x}_0 \in \mathbb{R}^N$ *and* $r > 0$*, the* ball *centered at* $\mathbf{x}_0$ *and of radius* r *is the set*

$$B(\mathbf{x}_0, r) := \left\{ \mathbf{x} \in \mathbb{R}^N : \|\mathbf{x} - \mathbf{x}_0\| < r \right\}.$$

Remark 32 *We can give a similar definition in a metric space* (X, d)*, precisely, given a point* $x_0 \in X$ *and* $r > 0$*, the* ball *centered at* x_0 *and of radius* r *is the set*

$$B(x_0, r) := \{x \in X : d(x, x_0) < r\}.$$

Definition 33 *Given a set* $E \subseteq \mathbb{R}^N$*, a point* $\mathbf{x} \in E$ *is called an* interior point *of* E *if there exists* $r > 0$ *such that* $B(\mathbf{x}, r) \subseteq E$*. The* interior E° *of a set* $E \subseteq \mathbb{R}^N$ *is the union of all its interior points. A subset* $U \subseteq \mathbb{R}^N$ *is* open *if every* $\mathbf{x} \in U$ *is an interior point of* U*.*

Example 34 *The ball* $B(\mathbf{x}_0, r)$ *is open. To see this, let* $\mathbf{x} \in B(\mathbf{x}_0, r)$*. Then* $B(\mathbf{x}, r - \|\mathbf{x} - \mathbf{x}_0\|)$ *is contained in* $B(\mathbf{x}_0, r)$*. Indeed, if* $\mathbf{y} \in B(\mathbf{x}, r - \|\mathbf{x} - \mathbf{x}_0\|)$*, then*

$$\|\mathbf{y} - \mathbf{x}_0\| \leq \|\mathbf{y} - \mathbf{x}\| + \|\mathbf{x} - \mathbf{x}_0\| < r - \|\mathbf{x} - \mathbf{x}_0\| + \|\mathbf{x} - \mathbf{x}_0\| = r,$$

and so $\mathbf{y} \in B(\mathbf{x}_0, r)$*.*

Example 35 *Some simple examples of sets that are open and of some that are not.*

(i) The set $(a, \infty) = \{x \in \mathbb{R} : x > a\}$ *is open. Indeed, if* $x > a$*, take* $r := x - a > 0$*. Then* $B(x, r) \subset (a, \infty)$*. Similarly, the set* $(-\infty, a)$ *is open.*

(ii) The set $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ *is open. Indeed, given* $a < x < b$*, take* $r := \min\{b - x, x - a\} > 0$*. Then* $B(x, r) \subseteq (a, b)$*.*

(iii) The set $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$ *is not open, since* b *belongs to the set but there is no ball* $B(b, r)$ *contained in* $(a, b]$*.*

Example 36 *Consider the set*

$$U = \mathbb{R} \setminus \left(\{0\} \cup \left\{ \frac{1}{n} : n \in \mathbb{N} \right\} \right).$$

Let's prove that U *is open. If* $x < 0$*, take* $r = -x > 0$*, then* $B(x, r) = (-2x, 0) \subseteq U$*. If* $x > 1$*, take* $r = x - 1$*, then* $B(x, r) = (1, 2x - 1) \subseteq U$*. If* $\frac{1}{n+1} < x < \frac{1}{n}$*, take* $r = \min\left\{\frac{1}{n} - x, x - \frac{1}{n+1}\right\} = \frac{1}{n+1}$*, then* $B(x, r) \subseteq U$*. Hence,* U *is open.*

Example 37 *Consider the set*

$$E = \mathbb{R} \setminus \left(\left\{ \frac{1}{n} : n \in \mathbb{N} \right\} \right).$$

Let's prove that E *is not open. The point* $x = 0$ *belongs to* E*, but for every* $r > 0$*, by the Archimedean principle we can find* $n \in \mathbb{N}$ *such that* $n > \frac{1}{r}$*, and so* $0 < \frac{1}{n} < r$*, which shows that* $\frac{1}{n} \in (-r, r)$*. Since* $\frac{1}{n}$ *does not belong to* E*, the ball* $(-r, r)$ *is not contained in* E *for* any $r > 0$*. Hence,* E *is not open.*

The main properties of open sets are given in the next proposition.

In what follows by an arbitrary family of sets of $\mathbb{R}^N$ we mean that there exists a set I and a function

$$f : I \to \mathcal{P}(\mathbb{R}^N)$$
$$\alpha \in I \mapsto f(\alpha) = U_\alpha$$

We write $\{U_\alpha\}$ or $\{U_\alpha\}_I$ or $\{U_\alpha\}_{\alpha \in I}$ to denote the set $\{f(\alpha) : \alpha \in I\}$.

Proposition 38 *The following properties hold:*

(i) $\emptyset$ *and* $\mathbb{R}^N$ *are open.*

(ii) If $U_i \subseteq \mathbb{R}^N$*,* $i = 1, \ldots, n$*, is a finite family of open sets of* $\mathbb{R}^N$*, then* $U_1 \cap \cdots \cap U_n$ *is open.*

(iii) If $\{U_\alpha\}_\alpha$ *is an arbitrary collection of open sets of* $\mathbb{R}^N$*, then* $\bigcup_\alpha U_\alpha$ *is open.*

Proof. To prove (ii), let $\mathbf{x} \in U_1 \cap \cdots \cap U_M$. Then $\mathbf{x} \in U_i$ for every $i = 1, \ldots, n$, and since U_i is open, there exists $r_i > 0$ such that $B(\mathbf{x}, r_i) \subseteq U_i$. Take $r := \min\{r_1, \ldots, r_n\} > 0$. Then

$$B(\mathbf{x}, r) \subseteq U_1 \cap \cdots \cap U_n,$$

which shows that $U_1 \cap \cdots \cap U_n$ is open.

To prove (iii), let $\mathbf{x} \in U := \bigcup_\alpha U_\alpha$. Then there is α such that $\mathbf{x} \in U_\alpha$ and since U_α is open, there exists $r > 0$ such that $B(\mathbf{x}, r) \subseteq U_\alpha \subseteq U$. This shows that U is open. ■

Remark 39 *The same proof continues to hold for a metric space.*

Properties (i)–(iii) are used to define topological spaces.

Definition 40 *Let X be a nonempty set and let τ be a family of sets of X. The pair (X, τ) is called a* topological space *if the following hold.*

(i) $\emptyset, X \in \tau$.

(ii) If $U_i \in \tau$ for $i = 1, \ldots, M$, then $U_1 \cap \ldots \cap U_M \in \tau$.

(iii) If $\{U_\alpha\}_\alpha$ is an arbitrary collection of elements of τ, then $\bigcup_\alpha U_\alpha \in \tau$.

The elements of the family τ are called open sets.

Remark 41 *The intersection of infinitely many open sets is not open in general. Take $U_n := \left(-\frac{1}{n}, \frac{1}{n}\right)$ for $n \in \mathbb{N}$. Then*

$$\bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\},$$

but $\{0\}$ is not open. Indeed, for every $r > 0$, the ball $(-r, r)$ is not contained in $\{0\}$.

Remark 42 *Proposition 38 shows that the family of open sets in $\mathbb{R}^N$ defined in Definition 33 is a a topology, called the* Euclidean topology. *Unless specified, in $\mathbb{R}^N$ we will always consider the Euclidean topology.*

Example 43 *Given a nonempty set X, there are always at least two topologies on X, namely,*

$$\tau_1 = \{\emptyset, X\}$$

(so according to τ_1, the only open sets are the empty set and X) and

$$\tau_2 = \{\text{all subsets of } X\}$$

(so according to τ_2 every set $E \subseteq X$ is open).

Exercise 44 *Prove that in* $\mathbb{R}^N$ *the norms*

$$\begin{aligned} \|\mathbf{x}\|_{\ell^\infty} &:= \max\{|x_1|, \dots, |x_N|\}, \\ \|\mathbf{x}\|_{\ell^1} &:= |x_1| + \cdots + |x_N|, \\ \|\mathbf{x}\|_{\ell^p} &:= \left(|x_1|^p + \cdots + |x_N|^p\right)^{1/p}, \end{aligned}$$

generate the same topology.

Wednesday, January 30, 2013

Remark 45 *For a topological space* (X, τ)*, given a point* $x \in X$*, a* neighborhood *of* x *is an open set containing* x*. Neighborhoods play the role of balls in metric spaces. Thus, given a set* $E \subseteq X$*, a point* $x \in E$ *is called an* interior point *of* E *if there exists a neighborhood* U *of* x *such that* $U \subseteq E$*.*

A subset $C \subseteq \mathbb{R}^N$ is *closed* if its complement $\mathbb{R}^N \setminus C$.
The main properties of closed sets are given in the next proposition.

Proposition 46 *The following properties hold:*

(i) $\emptyset$ *and* $\mathbb{R}^N$ *are closed.*

(ii) If $C_i \subseteq \mathbb{R}^N$*,* $i = 1, \dots, n$*, is a finite family of closed sets of* $\mathbb{R}^N$*, then* $C_1 \cup \cdots \cup C_n$ *is closed.*

(iii) If $\{C_\alpha\}_\alpha$ *is an arbitrary collection of closed sets of* $\mathbb{R}^N$*, then* $\bigcap_\alpha C_\alpha$ *is closed.*

The proof follows from Proposition 38 and De Morgan's laws. If $\{E_\alpha\}_\alpha$ is an arbitrary collection of subsets of a set $\mathbb{R}^N$, then *De Morgan's laws* are

$$\begin{aligned} \mathbb{R}^N \setminus \left(\bigcup_\alpha E_\alpha\right) &= \bigcap_\alpha \left(\mathbb{R}^N \setminus E_\alpha\right), \\ \mathbb{R}^N \setminus \left(\bigcap_\alpha E_\alpha\right) &= \bigcup_\alpha \left(\mathbb{R}^N \setminus E_\alpha\right). \end{aligned}$$

Definition 47 *Given a set* $E \subseteq \mathbb{R}^N$*, a point* $\mathbf{x} \in \mathbb{R}^N$ *is an* accumulation point*, or* cluster point *of* E *if for every* $r > 0$ *the ball* $B(\mathbf{x}, r)$ *contains at least one point of* E *different from* $\mathbf{x}$*.*

Note that $\mathbf{x}$ does not necessarily belong to the set E.

Remark 48 *An interior point of* $E \subseteq \mathbb{R}^N$ *is an accumulation point of* E*.*

Example 49 *Consider the set*

$$E := \left\{\frac{1}{n}\right\}_{n\in\mathbb{N}} \cup \left\{1+\frac{1}{n}\right\}_{n\in\mathbb{N}}.$$

We want to prove that 0 *and* 1 *are accumulation points of* E. *Note that* $0 \notin E$, *while* $1 \in E$ *(so accumulation points may or may not be in the set* E*). For* $r > 0$, *by taking a natural number* $n > \frac{1}{r}$, *we have that* $0 < \frac{1}{n} < r$, *and so* $\frac{1}{n} \in B(0,r) \cap E$ *(of course* $\frac{1}{n} \neq 0$*). This shows that* 0 *is an accumulation points of* E.

Similarly, for $r > 0$, *by taking a natural number* $n > \frac{1}{r}$, *we have that* $0 < \frac{1}{n} < r$, *and so* $1 < 1 + \frac{1}{n} < 1 + r$, *which shows that* $1 + \frac{1}{n} \in B(1,r) \cap E$ *(of course* $1 + \frac{1}{n} \neq 1$*). This shows that* 1 *is an accumulation points of* E. *Next we show that there are no other accumulation points of* E.

Indeed, if $x < 0$, *take* $r = -x > 0$, *then* $B(x,r) = (-2x, 0)$ *does not intersect* E. *If* $x > 2$, *take* $r = x - 1$, *then* $B(x,r) = (1, 2x-1)$ *does not intersect* E.

If $\frac{1}{n+1} < x < \frac{1}{n}$, *take* $r = \min\left\{\frac{1}{n} - x, x - \frac{1}{n+1}\right\} = \frac{1}{n+1}$, *then* $B(x,r)$ *does not intersect* E. *If* $x = \frac{1}{n}$, *with* $n > 1$, *take* $r = \min\left\{\frac{1}{n} - \frac{1}{n+1}, \frac{1}{n-1} - \frac{1}{n}\right\}$. *Then* $B(x,r)$ *intersects* E *only in* $\frac{1}{n}$. *Hence,* U *is open.*

If $1 + \frac{1}{n+1} < x < 1 + \frac{1}{n}$, *take* $r = \min\left\{1 + \frac{1}{n} - x, x - \left(1 + \frac{1}{n+1}\right)\right\} = \frac{1}{n+1}$, *then* $B(x,r)$ *does not intersect* E. *If* $x = 1 + \frac{1}{n}$, *take* $r = \min\left\{\frac{1}{n} - \frac{1}{n+1}, \frac{1}{n-1} - \frac{1}{n}\right\}$. *Then* $B(x,r)$ *intersects* E *only in* $1 + \frac{1}{n}$.

The set of all accumulation points of E is denoted $\operatorname{acc} E$.

Remark 50 *Note take if* $\mathbf{x} \in \mathbb{R}^N$ *is an accumulation point of* E, *then by taking* $r = \frac{1}{n}$, $n \in \mathbb{N}$, *there exists a sequence* $\{\mathbf{x}_n\} \subseteq E$ *with* $\mathbf{x}_n \neq \mathbf{x}$ *for all* $n \in \mathbb{N}$ *such that* $\|\mathbf{x}_n - \mathbf{x}\| < \frac{1}{n} \to 0$. *Thus* $\{x_n\}$ *converges to* x. *Conversely, if there exists* $\{\mathbf{x}_n\} \subseteq E$ *with* $\mathbf{x}_n \neq \mathbf{x}$ *for all* $n \in \mathbb{N}$ *such that* $\|\mathbf{x}_n - \mathbf{x}\| \to 0$, *then* $\mathbf{x}$ *is an accumulation point of* E.

Exercise 51 *Prove that a set* $E \subseteq \mathbb{R}^N$ *is closed if and only if it contains all its accumulation points.*

Definition 52 *Given a set* $E \subseteq \mathbb{R}^N$, *a point* $\mathbf{x} \in \mathbb{R}^N$ *is a* boundary point *of* E *if for every* $r > 0$ *the ball* $B(\mathbf{x}, r)$ *contains at least one point of* E *and one point of* $\mathbb{R}^N \setminus E$. *The set of boundary points of* E *is denoted* ∂E.

Exercise 53 *Prove that a set* $E \subseteq \mathbb{R}^N$ *is closed if and only if it contains all its boundary points.*

Friday, February 1, 2013

4 Functions

Consider a function $\mathbf{f} : E \to \mathbb{R}^M$, where $E \subseteq \mathbb{R}^N$. The set E is called the *domain* of $\mathbf{f}$. If E is not specified, then E should be taken to be the largest set of $\mathbf{x}$ for which $\mathbf{f}(\mathbf{x})$ makes sense. This means that:

If there are even roots, their arguments should be nonnegative. If there are logarithms, their arguments should be strictly positive. Denominators should be different from zero. If a function is raised to an irrational number, then the function should be nonnegative.

Given a set $F \subseteq E$, the set $\mathbf{f}(F) = \left\{\mathbf{y} \in \mathbb{R}^M : \mathbf{y} = \mathbf{f}(\mathbf{x}) \text{ for some } \mathbf{x} \in F\right\}$ is called the *image* of F through $\mathbf{f}$. The function $\mathbf{f}$ is said to be bounded from above in F, bounded from below in F, bounded in F if the set $\mathbf{f}(F)$ is bounded from above, bounded from below, bounded, respectively.

Given a set $G \subseteq \mathbb{R}$, the set $\mathbf{f}^{-1}(G) = \{\mathbf{x} \in E : \mathbf{f}(\mathbf{x}) \in G\}$ is called the *inverse image* of F through $\mathbf{f}$. It has NOTHING to do with the inverse function. It is just one of those unfortunate cases in which we use the same symbol for two different objects.

The *graph* of a function is the set of $\mathbb{R}^N \times \mathbb{R}^M$ defined by

$$\operatorname{gr} \mathbf{f} = \{(\mathbf{x}, \mathbf{f}(\mathbf{x})) : \mathbf{x} \in E\}.$$

A function $\mathbf{f}$ is said to be

- *one-to-one* or *injective* if $\mathbf{f}(\mathbf{x}) \neq \mathbf{f}(\mathbf{z})$ for all $\mathbf{x}, \mathbf{z} \in E$ with $\mathbf{x} \neq \mathbf{z}$.

If $\mathbf{f} : E \to F$, where $E, F \subseteq \mathbb{R}$, then $\mathbf{f}$ is said to be

- *onto* or *surjective* if $\mathbf{f}(E) = F$,
- *bijective* or *invertible* if it is one-to-one and onto. The function $\mathbf{f}^{-1} : F \to E$, which assigns to each $\mathbf{y} \in F = \mathbf{f}(E)$ the unique $\mathbf{x} \in E$ such that $\mathbf{f}(\mathbf{x}) = \mathbf{y}$, is called the *inverse* function of $\mathbf{f}$.

If $N = M = 1$ a function $f : E \to \mathbb{R}$ is said to be

- *increasing* if $f(x) \leq f(y)$ for all $x, y \in E$ with $x < y$,
- *strictly increasing* if $f(x) < f(y)$ for all $x, y \in E$ with $x < y$,
- *decreasing* if $f(x) \geq f(y)$ for all $x, y \in E$ with $x < y$,
- *strictly decreasing* if $f(x) > f(y)$ for all $x, y \in E$ with $x < y$,
- *monotone* if one of the four property above holds.

5 Limits of Functions

Definition 54 *Let* $E \subseteq \mathbb{R}^N$*, let* $\mathbf{x}_0 \in \mathbb{R}^N$ *be an accumulation point of* E*, and let* $\mathbf{f} : E \to \mathbb{R}^N$*. We say that a number* $\boldsymbol{\ell} \in \mathbb{R}^M$ *is the* limit *of* $\mathbf{f}(\mathbf{x})$ *as* $\mathbf{x}$ *approaches* $\mathbf{x}_0$ *if for every* $\varepsilon > 0$ *there exists a real number* $\delta = \delta(\varepsilon, \mathbf{x}_0) > 0$ *with the property that*

$$\|\mathbf{f}(\mathbf{x}) - \boldsymbol{\ell}\| < \varepsilon$$

for all $\mathbf{x} \in E$ *with* $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$*. We write*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \boldsymbol{\ell} \quad \textit{or} \quad \mathbf{f}(\mathbf{x}) \to \boldsymbol{\ell} \textit{ as } \mathbf{x} \to \mathbf{x}_0.$$

Note that even when $\mathbf{x}_0 \in E$, we cannot take $\mathbf{x} = \mathbf{x}_0$ since in the definition we require $0 < \|\mathbf{x} - \mathbf{x}_0\|$.

Remark 55 *If* (X, d_X) *and* (Y, d_Y) *are two metric spaces,* $E \subseteq X$*,* $x_0 \in E$ *is an accumulation point of* E *and* $f : E \to Y$*, we say that* $\ell \in Y$ *is the* limit *of* $f(x)$ *as* x *approaches* x_0 *if for every* $\varepsilon > 0$ *there exists a real number* $\delta = \delta(\varepsilon, x_0) > 0$ *with the property that*

$$d_Y(f(x), \ell) < \varepsilon$$

for all $x \in E$ *with* $0 < d_X(x, x_0) < \delta$*. We write*

$$\lim_{x \to x_0} f(x) = \ell.$$

If (X, τ_X) *and* (Y, τ_Y) *are two topological spaces,* $E \subseteq X$*,* $x_0 \in E$ *is an accumulation point of* E *and* $f : E \to Y$*, we say that* $\ell \in Y$ *is the* limit *of* $f(x)$ *as* x *approaches* x_0 *if for every neighborhood* V *of* ℓ *there exists a neighborhood* U *of* x_0 *with the property that*

$$f(x) \in V$$

for all $x \in E$ *with* $x \in U \setminus \{x_0\}$*. We write*

$$\lim_{x \to x_0} f(x) = \ell.$$

Note that unless the space Y *is Hausdorff, the limit may not be unique.*

Remark 56 *Let* $E \subseteq \mathbb{R}^N$*, let* $\mathbf{x}_0 \in \mathbb{R}^N$ *be an accumulation point of* E*, and let* $\mathbf{f} : E \to \mathbb{R}^N$*. Assume that there exists*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \boldsymbol{\ell}.$$

Then for every $\varepsilon > 0$ *there exists a real number* $\delta = \delta(\varepsilon, \mathbf{x}_0) > 0$ *with the property that*

$$\|\mathbf{f}(\mathbf{x}) - \boldsymbol{\ell}\| < \varepsilon \tag{4}$$

for all $\mathbf{x} \in E$ *with* $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$*. if* $F \subseteq E$ *is a subset such that* $\mathbf{x}_0 \in \mathbb{R}^N$ *be an accumulation point of* E *and we consider the resctriction of* $\mathbf{f}$ *to* F*, denoted*

$\mathbf{f}|_F$, *for every* $\varepsilon > 0$ *let* $\delta > 0$ *be the number given in (4). Then by restricting (4) we have that*

$$\|\mathbf{f}(\mathbf{x}) - \boldsymbol{\ell}\| < \varepsilon$$

for all $\mathbf{x} \in F$ *with* $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$. *Hence, there exists*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}|_F(\mathbf{x}) = \boldsymbol{\ell}.$$

It follows that if we can find two sets $F \subseteq E$ *and* $G \subseteq E$ *such that* $\mathbf{x}_0 \in \operatorname{acc} F$ *and* $\mathbf{x}_0 \in \operatorname{acc} G$

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}|_F(\mathbf{x}) = \boldsymbol{\ell}_1 \neq \boldsymbol{\ell}_2 = \lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}|_G(\mathbf{x}),$$

then the limit over E *cannot exist.*

Example 57 *Let's study the limit*

$$\lim_{(x,y) \to (0,0)} \frac{x^m y}{x^2 + y^2},$$

where $m \in \mathbb{N}$. *In this case* $f(x,y) = \frac{x^m y}{x^2+y^2}$ *and the domain is* $\mathbb{R}^2 \setminus \{(0,0)\}$.

For $m \geq 2$, *we have that the limit is* 0. *Indeed, using the fact that* $|x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2}$, *we have*

$$\left| \frac{x^m y}{x^2 + y^2} - 0 \right| = \frac{|x|^m |y|}{x^2 + y^2} \leq \frac{\left(x^2 + y^2\right)^{m/2} \left(x^2 + y^2\right)^{1/2}}{x^2 + y^2} = \left(x^2 + y^2\right)^{(m-1)/2} \to 0$$

as $(x,y) \to (0,0)$. *On the other hand, if* $m = 1$, *taking* $y = x$, *we have that*

$$\lim_{x \to 0} f(x,x) = \lim_{x \to 0} \frac{x^2}{x^2 + x^2} = \frac{1}{2},$$

while taking $y = 0$, *we have that*

$$f(x,0) = \frac{0}{x^2 + 0} = 0 \to 0,$$

and so the limit does not exist.

Remark 58 *Note that the degree of the numerator is* $m + 1$ *and the degree of the numerator is* 2, *so that in this particular example the limit exists if the degree of the numerator is higher than the degree of the numerator, that is, if* $m + 1 > 2$.

The next example shows that checking the limit on every line passing through $\mathbf{x}_0$ is not enough to gaurantee the existence of the limit.

Example 59 *Let*

$$f(x,y) := \begin{cases} 1 & \text{if } y = x^2,\ x \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Given the line $y = mx$, the line intersects the parabola $y = x^2$ only in $\mathbf{0}$ and in at most one point. Hence, if x is very small,

$$f(x, mx) = 0 \to 0$$

as $x \to 0$. However, since $f\left(x, x^2\right) = 1 \to 1$ as $x \to 0$, the limit does not exists.

Example 60 *Study the limit*

$$\lim_{(x,y)\to(0,0)} \frac{x^m y}{x^2 - y^2},$$

where $m \in \mathbb{N}$. In this case $f(x, y) = \frac{x^m y}{x^2 - y^2}$ and the domain is $\mathbb{R}^2 \backslash \{(x, x) : x \in \mathbb{R}\}$. Taking $y = 0$, we have that

$$f(x, 0) = \frac{0}{x^2 - 0} = 0 \to 0.$$

Let us take $y = x + x^a$, where $a > 1$. Then

$$\begin{aligned} f(x, x + x^a) &= \frac{x^m (x + x^a)}{x^2 - (x + x^a)^2} = \frac{x^{m+1} + x^{m+a}}{x^2 - x^2 - 2x^{a+1} - x^{2a}} \\ &= -\frac{x^{m+1} + x^{m+a}}{2x^{a+1} + x^{2a}}. \end{aligned}$$

Take $a = m$. Then

$$\begin{aligned} f(x, x + x^m) &= -\frac{x^{m+1} + x^{2m}}{2x^{m+1} + x^{2m}} = -\frac{x^{m+1}\left(1 + x^{m-1}\right)}{x^{m+1}\left(2 + x^{m-1}\right)} \\ &= -\frac{1 + x^{m-1}}{2 + x^{m-1}} \nrightarrow 0. \end{aligned}$$

Hence the limit does not exist.

Remark 61 *Note that the degree of the numerator is $m + 1$ and the degree of the numerator is 2, but in this case the limit never exists no matter how high is the degree of the numerator.*

Monday, February 4, 2013

Definition 62 *Let $E \subseteq \mathbb{R}^N$, let $\mathbf{x}_0 \in \mathbb{R}^N$ be an accumulation point of E, and let $f : E \to \mathbb{R}$. We say that*

- *∞ is the* limit *of $f(\mathbf{x})$ as $\mathbf{x}$ approaches $\mathbf{x}_0$ if for every $L > 0$ there exists a real number $\delta = \delta(L, \mathbf{x}_0) > 0$ with the property that*

$$f(\mathbf{x}) > L$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$. We write

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \infty \quad \text{or} \quad f(\mathbf{x}) \to \infty \text{ as } \mathbf{x} \to \mathbf{x}_0,$$

- $-\infty$ *is the* limit *of* $f(\mathbf{x})$ *as* $\mathbf{x}$ *approaches* $\mathbf{x}_0$ *if for every* $L > 0$ *there exists a real number* $\delta = \delta(L, \mathbf{x}_0) > 0$ *with the property that*

$$f(\mathbf{x}) < -L$$

for all $\mathbf{x} \in E$ *with* $0 < \|\mathbf{x} - \mathbf{x}_0\| \leq \delta$. *We write*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = -\infty \quad \text{or} \quad f(\mathbf{x}) \to -\infty \text{ as } \mathbf{x} \to \mathbf{x}_0.$$

Theorem 63 *Let* $E \subseteq \mathbb{R}^N$ *and let* $\mathbf{x}_0 \in \mathbb{R}^N$ *be an accumulation point of* E. *Given a function* $\mathbf{f} : E \to \mathbb{R}^M$, *if the limit*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(x)$$

exists, it is unique.

Proof. Assume by contradiction that there exist

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \boldsymbol{\ell} \quad \text{and} \quad \lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \boldsymbol{L}$$

with $\boldsymbol{\ell} \neq \boldsymbol{L}$. Then $\|\boldsymbol{\ell} - \boldsymbol{L}\| > 0$. Fix $0 < \varepsilon < \frac{1}{2}\|\boldsymbol{\ell} - \boldsymbol{L}\|$. Since $\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \boldsymbol{\ell}$, there exists $\delta_1 > 0$ with the property that

$$\|\mathbf{f}(\mathbf{x}) - \boldsymbol{\ell}\| < \varepsilon$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_1$, while, since $\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \boldsymbol{L}$, there exists $\delta_2 > 0$ with the property that

$$\|\mathbf{f}(\mathbf{x}) - \boldsymbol{L}\| < \varepsilon$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_2$.

Take $\delta = \min\{\delta_1, \delta_2\} > 0$ and take $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$. Note that such $\mathbf{x}$ exists because $\mathbf{x}_0$ is an accumulation point of E. Then by the properties of the norm,

$$\begin{aligned} \|\boldsymbol{\ell} - \boldsymbol{L}\| &= \|\boldsymbol{\ell} - \mathbf{f}(\mathbf{x}) + \mathbf{f}(\mathbf{x}) - \boldsymbol{L}\| \leq \|\boldsymbol{\ell} - \mathbf{f}(\mathbf{x})\| + \|\mathbf{f}(\mathbf{x}) - \boldsymbol{L}\| \\ &< \varepsilon + \varepsilon < \|\boldsymbol{\ell} - \boldsymbol{L}\|, \end{aligned}$$

which implies that $\|\boldsymbol{\ell} - \boldsymbol{L}\| < \|\boldsymbol{\ell} - \boldsymbol{L}\|$. This contradiction proves the theorem. ■

Remark 64 *This proof continues to hold in metric spaces, since we only used the fact that the balls* $B(\boldsymbol{\ell}, \varepsilon)$ *and* $B(\mathbf{L}, \varepsilon)$ *are disjoint whenever* $0 < \varepsilon < \frac{1}{2}\|\boldsymbol{\ell} - \mathbf{L}\|$. *For topological spaces in general the limit is not unique. Given* (X, τ_X) *and* (Y, τ_Y) *are two topological spaces,* $E \subseteq X$, $x_0 \in E$ *is an accumulation point of* E *and* $f : E \to Y$, *it can be shown that the limit is unique if the space* Y *is Hausdorff. A topological space* Y *is a* Hausdorff space, *if for every* x *and* $y \in Y$, *with* $x \neq y$, *there exist disjoint neighborhoods of* x *and* y.

We now list some important operations for limits.

Theorem 65 *Let $E \subseteq \mathbb{R}^N$ and let $\mathbf{x}_0 \in \mathbb{R}^N$ be an accumulation point of E. Given two functions $\boldsymbol{f}, \boldsymbol{g} : E \to \mathbb{R}^M$, assume that there exist*

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \boldsymbol{f}(\mathbf{x}) = \boldsymbol{\ell}_1 \in \mathbb{R}^M, \qquad \lim_{\mathbf{x}\to\mathbf{x}_0} \boldsymbol{g}(\mathbf{x}) = \boldsymbol{\ell}_2 \in \mathbb{R}^M.$$

Then

(i) there exists $\lim_{\mathbf{x}\to\mathbf{x}_0} (\boldsymbol{f} + \boldsymbol{g})(\mathbf{x}) = \boldsymbol{\ell}_1 + \boldsymbol{\ell}_2$,

(ii) there exists $\lim_{\mathbf{x}\to\mathbf{x}_0} (\boldsymbol{f} \cdot \boldsymbol{g})(\mathbf{x}) = \boldsymbol{\ell}_1 \boldsymbol{\ell}_2$,

(iii) if $M = 1$, $\ell_2 \neq 0$ and $F := \{\mathbf{x} \in E : g(\mathbf{x}) \neq 0\}$, then $\mathbf{x}_0$ is an accumulation point of F and there exists $\lim_{\mathbf{x}\to\mathbf{x}_0} \left(\left. \frac{f}{g} \right|_F \right)(\mathbf{x}) = \frac{\ell_1}{\ell_2}$.

Proof. Parts (i) and (ii) are left as an exercise. We prove part (iii). The fact that $\mathbf{x}_0$ is an accumulation point of F is left as an exercise. Write

$$\begin{aligned}
\left| \frac{f(\mathbf{x})}{g(\mathbf{x})} - \frac{\ell_1}{\ell_2} \right| &= \left| \frac{f(\mathbf{x})\,\ell_2 - g(\mathbf{x})\,\ell_1}{g(\mathbf{x})\,\ell_2} \right| = \left| \frac{f(\mathbf{x})\,\ell_2 \pm \ell_1\ell_2 - g(\mathbf{x})\,\ell_1}{g(\mathbf{x})\,\ell_2} \right| \\
&= \frac{1}{|g(\mathbf{x})|\,|\ell_2|} \left| \ell_2 (f(\mathbf{x}) - \ell_1) + \ell_1 (g(\mathbf{x}) - \ell_2) \right| \\
&\leq \frac{1}{|g(\mathbf{x})|} |f(\mathbf{x}) - \ell_1| + \frac{1}{|g(\mathbf{x})|} \frac{|\ell_1|}{|\ell_2|} |g(\mathbf{x}) - \ell_2| .
\end{aligned}$$

Thus we need to bound $\frac{1}{|g(\mathbf{x})|}$ from above, or, equivalently, we need $|g(\mathbf{x})|$ to stay away from zero. Since $\ell_2 \neq 0$, taking $\varepsilon = \frac{|\ell_2|}{2} > 0$, there exist $\delta_1 > 0$ such that

$$|g(\mathbf{x}) - \ell_2| \leq \frac{|\ell_2|}{2}$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_1$. Hence,

$$|g(\mathbf{x})| = |g(\mathbf{x}) \pm \ell_2| \geq |\ell_2| - |g(\mathbf{x}) - \ell_2| \geq |\ell_2| - \frac{|\ell_2|}{2} = \frac{|\ell_2|}{2}$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_1$. It follows that

$$\begin{aligned}
\left| \frac{f(\mathbf{x})}{g(\mathbf{x})} - \frac{\ell_1}{\ell_2} \right| &\leq \frac{1}{|g(\mathbf{x})|} |f(\mathbf{x}) - \ell_1| + \frac{1}{|g(\mathbf{x})|} \frac{|\ell_1|}{|\ell_2|} |g(\mathbf{x}) - \ell_2| \\
&\leq \frac{2}{|\ell_2|} |f(\mathbf{x}) - \ell_1| + \frac{2}{|\ell_2|} \frac{|\ell_1|}{|\ell_2|} |g(\mathbf{x}) - \ell_2| .
\end{aligned}$$

Given $\varepsilon > 0$ there exist $\delta_2 > 0$ such that

$$|f(\mathbf{x}) - \ell_1| \leq \frac{\varepsilon\,|\ell_2|}{4}$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_2$ and $\delta_3 > 0$ such that

$$|g(\mathbf{x}) - \ell_2| \leq \frac{\varepsilon |\ell_2|^2}{4(1 + |\ell_1|)}$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_3$. Then for for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta = \min\{\delta_1, \delta_2, \delta_3\}$, we have that

$$\begin{aligned} \left| \frac{f(\mathbf{x})}{g(\mathbf{x})} - \frac{\ell_1}{\ell_2} \right| &\leq \frac{2}{|\ell_2|} |f(\mathbf{x}) - \ell_1| + \frac{2}{|\ell_2|} \frac{|\ell_1|}{|\ell_2|} |g(\mathbf{x}) - \ell_2| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \frac{|\ell_1|}{1 + |\ell_1|} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} 1 = \varepsilon. \end{aligned}$$

This completes the proof. ■

Remark 66 *The previous theorem continues to hold if $\ell_1, \ell_1 \in [-\infty, \infty]$, provided we avoid the cases $\infty - \infty$, 0∞, $\frac{0}{0}$, $\frac{\infty}{\infty}$.*

Theorem 67 (Squeeze Theorem) *Let $E \subseteq \mathbb{R}^N$ and let $\mathbf{x}_0 \in \mathbb{R}^N$ be an accumulation point of E. Given three functions $f, g, h : E \to \mathbb{R}$, assume that there exist*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \lim_{\mathbf{x} \to \mathbf{x}_0} g(\mathbf{x}) = \ell.$$

and that $f(\mathbf{x}) \leq h(\mathbf{x}) \leq g(\mathbf{x})$ for every $\mathbf{x} \in E$. Then there exists $\lim_{\mathbf{x} \to \mathbf{x}_0} h(\mathbf{x}) = \ell$.

Proof. We prove the case in which $\ell \in \mathbb{R}$ and leave the cases $\ell = \infty$ and $\ell = -\infty$ as an exercise. Given $\varepsilon > 0$ there exist $\delta_1 > 0$ such that

$$|f(\mathbf{x}) - \ell| \leq \varepsilon$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_1$ and $\delta_2 > 0$ such that

$$|g(\mathbf{x}) - \ell| \leq \varepsilon$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_2$. Then for for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta = \min\{\delta_1, \delta_2\}$, we have that

$$\ell - \varepsilon \leq f(\mathbf{x}) \leq g(\mathbf{x}) \leq h(\mathbf{x}) \leq \ell + \varepsilon.$$

Hence,

$$|h(\mathbf{x}) - \ell| \leq \varepsilon$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$, which shows that $\lim_{\mathbf{x} \to \mathbf{x}_0} h(\mathbf{x}) = \ell$. ■

Theorem 68 *Let $E \subseteq \mathbb{R}^N$ and let $\mathbf{x}_0 \in \mathbb{R}$ be an accumulation point of E. Given two functions $\boldsymbol{f}, \boldsymbol{g} : E \to \mathbb{R}^M$, assume that there exists*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \boldsymbol{f}(\mathbf{x}) = \mathbf{0},$$

and that $\boldsymbol{g}$ is bounded, that is, $\|\boldsymbol{g}(\mathbf{x})\| \leq L$ for all $\mathbf{x} \in E$ and for some $L > 0$. Then there exists $\lim_{\mathbf{x} \to \mathbf{x}_0} (\boldsymbol{f} \cdot \boldsymbol{g})(\mathbf{x}) = 0$.

Proof. Given $\varepsilon > 0$ there existd $\delta > 0$ such that

$$\|f(\mathbf{x}) - \mathbf{0}\| < \frac{\varepsilon}{1+L}$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$. Hence, by the Cauchy–Schwarz's inequality

$$|(\boldsymbol{f} \cdot \boldsymbol{g})(\mathbf{x}) - 0| = |(\boldsymbol{f} \cdot \boldsymbol{g})(\mathbf{x})| \leq \|f(\mathbf{x})\| \|g(\mathbf{x})\| < \frac{\varepsilon}{1+L} L < \varepsilon$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$. ■

Example 69 *The previous theorem can be used for example to show that for* $a > 0$

$$\lim_{x \to 0} x^a \sin \frac{1}{x} = 0.$$

Wednesday, February 6, 2013

We next study the limit of composite functions.

Theorem 70 *Let* $E \subseteq \mathbb{R}^N$, $F \subseteq \mathbb{R}^M$ *and let* $\mathbf{x}_0 \in \mathbb{R}$ *be an accumulation point of* E. *Given two functions* $\mathbf{f} : E \to F$ *and* $g : F \to \mathbb{R}^P$, *assume that there exist*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \boldsymbol{\ell} \in \mathbb{R}^M,$$

that $\boldsymbol{\ell}$ *is an accumulation point of* F, *and that there exists*

$$\lim_{\mathbf{y} \to \boldsymbol{\ell}} \mathbf{g}(\mathbf{y}) = \mathbf{L} \in \mathbb{R}^P.$$

Assume that either there exists $\delta_1 > 0$ *such that* $\mathbf{f}(\mathbf{x}) \neq \boldsymbol{\ell}$ *for all* $\mathbf{x} \in E$ *with* $0 < |\mathbf{x} - \mathbf{x}_0| \leq \delta_1$, *or that* $\boldsymbol{\ell} \in F$ *and* $\mathbf{g}(\boldsymbol{\ell}) = \mathbf{L}$. *Then there exists* $\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{L}$.

Proof. Fix $\varepsilon > 0$ and find $\eta = \eta(\varepsilon, \boldsymbol{\ell}) > 0$ such that

$$\|\mathbf{g}(\mathbf{y}) - \mathbf{L}\| < \varepsilon \tag{5}$$

for all $\mathbf{y} \in F$ with $0 < \|\mathbf{y} - \boldsymbol{\ell}\| < \eta$.

Since $\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \boldsymbol{\ell}$, there exists $\delta_2 = \delta_2(\mathbf{x}_0, \eta) > 0$ such that

$$\|\mathbf{f}(\mathbf{x}) - \boldsymbol{\ell}\| < \eta$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_1$.

We now distinguish two cases.

Case 1: Assume that $\mathbf{f}(\mathbf{x}) \neq \boldsymbol{\ell}$ for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_1$. Then taking $\delta = \min\{\delta_1, \delta_2\}$, we have that for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$,

$$0 < \|\mathbf{f}(\mathbf{x}) - \boldsymbol{\ell}\| < \eta.$$

Hence, taking $\mathbf{y} = \mathbf{f}(\mathbf{x})$, by (5), it follows that

$$\|\mathbf{g}(\mathbf{f}(\mathbf{x})) - \mathbf{L}\| < \varepsilon$$

for all $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$. This shows that there exists $\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{L}$.

Case 2: Assume that $\boldsymbol{\ell} \in F$ and $\mathbf{g}(\boldsymbol{\ell}) = \mathbf{L}$. Let $\mathbf{x} \in E$ with $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta_1$. If $\mathbf{f}(\mathbf{x}) = \boldsymbol{\ell}$, then $\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{L}$, and so

$$\|\mathbf{g}(\mathbf{f}(\mathbf{x})) - \mathbf{L}\| = 0 < \varepsilon,$$

while if $\mathbf{f}(\mathbf{x}) \neq \boldsymbol{\ell}$, then taking $\mathbf{y} = \mathbf{f}(\mathbf{x})$, by (5), it follows that

$$\|\mathbf{g}(\mathbf{f}(\mathbf{x})) - \mathbf{L}\| < \varepsilon.$$

■

Example 71 *Let's prove that the previous theorem fails without the hypotheses that either* $\mathbf{f}(\mathbf{x}) \neq \boldsymbol{\ell}$ *for all* $\mathbf{x} \in E$ *near* $\mathbf{x}_0$ *or* $\boldsymbol{\ell} \in F$, $L \in \mathbb{R}$ *and* $\mathbf{g}(\boldsymbol{\ell}) = L$. *Consider the function*

$$g(y) := \begin{cases} 1 & \text{if } y \neq 0, \\ 2 & \text{if } y = 0. \end{cases}$$

Then there exists

$$\lim_{y \to 0} g(y) = 1.$$

So $L = 1$. *Consider the function* $f(x) := 0$ *for all* $x \in \mathbb{R}$. *Then for every* $x_0 \in \mathbb{R}$, *we have that*

$$\lim_{x \to x_0} f(x) = 0.$$

So $\ell = 0$. *However,* $g(f(x)) = g(0) = 2$ *for all* $x \in \mathbb{R}$. *Hence,*

$$\lim_{x \to x_0} g(f(x)) = \lim_{x \to x_0} 2 = 2 \neq 1,$$

which shows that the conclusion of the theorem is violated.

Example 72 *We list below some important limits.*

$$\lim_{x \to 0} \frac{\sin x}{x} = 1, \qquad \lim_{x \to 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}, \qquad \lim_{x \to 0} \frac{\log(1 + x)}{x} = 1,$$
$$\lim_{x \to 0} \frac{(1 + x)^a - 1}{x} = a \quad \text{for } a \in \mathbb{R}, \qquad \lim_{x \to 0} \frac{e^x - 1}{x} = 1.$$

Note that the previous theorem can be used to change variables in limits.

Example 73 *Let's try to calculate*

$$\lim_{x \to 0} \frac{\log(1 + \sin x)}{x}.$$

For $\sin x \neq 0$, *we have*

$$\frac{\log(1 + \sin x)}{x} = \frac{\log(1 + \sin x)}{x} \frac{\sin x}{\sin x} = \frac{\log(1 + \sin x)}{\sin x} \frac{\sin x}{x}.$$

Since $\lim_{x \to 0} \frac{\sin x}{x} = 1$, *it remains to study*

$$\lim_{x \to 0} \frac{\log(1 + \sin x)}{\sin x}.$$

Consider the function $g(y) = \frac{\log(1+y)}{y}$ *and the function* $f(x) = \sin x$. *As* $x \to 0$, *we have that* $\sin x \to 0 = \ell$, *while*

$$\lim_{y \to 0} \frac{\log(1 + y)}{y} = 1.$$

Moreover $\sin x \neq 0$ *for all* $x \in E := \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \setminus \{0\}$. *Hence, we can apply the previous theorem to conclude that*

$$\lim_{x \to 0} \frac{\log(1 + \sin x)}{\sin x} = 1.$$

In turn,

$$\lim_{x \to 0} \frac{\log(1 + \sin x)}{x} = 1.$$

6 Continuity

Definition 74 *Let* $E \subseteq \mathbb{R}^N$. *A point* $\mathbf{x}_0 \in E$ *is called an* isolated point *of* E *if there exists* $\delta > 0$ *such that*

$$B(\mathbf{x}_0, \delta) \cap E = \{\mathbf{x}_0\}.$$

It is clear that if a point of the set E is not an isolated point of E then it is an accumulation point of E.

Definition 75 *Let* $E \subseteq \mathbb{R}^N$ *and let* $\mathbf{x}_0 \in E$. *Given a function* $\mathbf{f} : E \to \mathbb{R}^M$ *we say that* $\mathbf{f}$ *is* continuous *at* $\mathbf{x}_0$ *if for every* $\varepsilon > 0$ *there exists a real number* $\delta = \delta(\varepsilon, \mathbf{x}_0) > 0$ *such that for all* $\mathbf{x} \in E$ *with* $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ *we have*

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| < \varepsilon.$$

If $\mathbf{f}$ *is continuous at every point of* E *we say that* $\mathbf{f}$ *is* continuous *on* E *and we write* $\mathbf{f} \in C(E)$ *or* $\mathbf{f} \in C^0(E)$.

Remark 76 *If* (X, d_X) *and* (Y, d_Y) *are two metric spaces,* $E \subseteq X$, $x_0 \in E$, *and* $f : E \to Y$, *we say that* f *is* continuous *at* x_0 *if for every* $\varepsilon > 0$ *there exists a real number* $\delta = \delta(\varepsilon, x_0) > 0$ *with the property that*

$$d_Y(f(x), f(x_0)) < \varepsilon$$

for all $x \in E$ *with* $d_X(x, x_0) < \delta$.

If (X, τ_X) *and* (Y, τ_Y) *are two topological spaces,* $E \subseteq X$, $x_0 \in E$, *and* $f : E \to Y$, *we say that* f *is* continuous *at* x_0 *if for every neighborhood* V *of* $f(x_0)$ *there exists a neighborhood* U *of* x_0 *with the property that*

$$f(x) \in V$$

for all $x \in E$ *with* $x \in U$.

Theorem 77 *Let* $E \subseteq \mathbb{R}^N$ *and let* $\mathbf{x}_0 \in E$*. Given a function* $\mathbf{f} : E \to \mathbb{R}^M$*,*

(i) if $\mathbf{x}_0$ *is an isolated point of* E *then* $\mathbf{f}$ *is continuous at* $\mathbf{x}_0$*;*

(ii) if $\mathbf{x}_0$ *is an accumulation point of* E *then* $\mathbf{f}$ *is continuous at* $\mathbf{x}_0$ *if and only if there exists* $\lim_{\mathbf{x} \to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0)$*.*

Proof of part (i). If $\mathbf{x}_0$ is an isolated point of E then there exists $\delta_0 > 0$ such that

$$B(\mathbf{x}_0, \delta_0) \cap E = \{\mathbf{x}_0\}.$$

Fix $\varepsilon > 0$ and take $\delta := \delta_0$ in the definition of continuity. Clearly if $\mathbf{x} \in E$ and $\|\mathbf{x} - \mathbf{x}_0\| < \delta$ then necessarily $\mathbf{x} = \mathbf{x}_0$ so that we have $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0)\| = 0$. ■

Exercise 78 *Prove that the functions* $\sin x$*,* $\cos x$*,* x^n*, where* $n \in \mathbb{N}$*, are continuous.*

The following theorems follows from the analogous results for limits.

Theorem 79 *Let* $E \subseteq \mathbb{R}^N$ *and let* $\mathbf{x}_0 \in E$*. Given two functions* $f, g : E \to \mathbb{R}$ *assume that* f *and* g *are continuous at* $\mathbf{x}_0$*. Then*

(i) $f + g$ *and* fg *are continuous at* $\mathbf{x}_0$*;*

(ii) if $g(\mathbf{x}_0) \neq 0$ *then* $\frac{f}{g}$ *restricted to the set* $F := \{\mathbf{x} \in E : g(\mathbf{x}) \neq 0\}$ *is continuous at* $\mathbf{x}_0$*.*

Example 80 *In view of Exercise 78 and the previous theorem, the functions* $\tan x = \frac{\sin x}{\cos x}$ *and* $\cot x = \frac{\cos x}{\sin x}$ *are continuous in their domain of definition.*

Theorem 81 *Let* $E \subseteq \mathbb{R}^N$*,* $F \subseteq \mathbb{R}^M$ *and let* $\mathbf{x}_0 \in E$*. Given two functions* $\boldsymbol{f} : E \to F$ *and* $\boldsymbol{g} : F \to \mathbb{R}^P$ *assume that* $\boldsymbol{f}$ *is continuous at* $\mathbf{x}_0$ *and that* $\boldsymbol{g}$ *is continuous at* $\boldsymbol{f}(\mathbf{x}_0)$*. Then* $\boldsymbol{g} \circ \boldsymbol{f} : E \to \mathbb{R}^P$ *is continuous at* $\mathbf{x}_0$*.*

We now discuss the continuity of inverse functions and of composite functions. If a continuous function $\mathbf{f}$ is invertible its inverse function $\mathbf{f}^{-1}$ may not be continuous.

Example 82 *Let*

$$f(x) := \begin{cases} x & \text{if } 0 \leq x \leq 1, \\ x - 1 & \text{if } 2 < x \leq 3. \end{cases}$$

Then $f^{-1} : [0, 2] \to R$ *is given by*

$$f^{-1}(x) := \begin{cases} x & \text{if } 0 \leq x \leq 1, \\ x + 1 & \text{if } 1 < x \leq 2, \end{cases}$$

which is not continuous at $x = 1$*.*

We will see that this cannot happen if E is an interval or a compact set.

Theorem 83 *Let $K \subset \mathbb{R}^N$ be a closed and bounded set and let $\mathbf{f} : K \to \mathbb{R}^N$ be one-to-one and continuous. Then the inverse function $\mathbf{f}^{-1} : f(K) \to \mathbb{R}^N$ is continuous.*

Theorem 84 *Let $I \subseteq \mathbb{R}$ be an interval and let $f : I \to \mathbb{R}$ be one-to-one and continuous. Then the inverse function $f^{-1} : f(I) \to \mathbb{R}$ is continuous.*

Friday, February 8, 2013

7 Directional Derivatives and Differentiability

Let $E \subseteq \mathbb{R}^N$, let $f : E \to \mathbb{R}$ and let $\mathbf{x}_0 \in E$. Given a direction $\mathbf{v} \in \mathbb{R}^N$, let L be the line through $\mathbf{x}_0$ in the direction $\mathbf{v}$, that is,

$$L := \left\{ \mathbf{x} \in \mathbb{R}^N : \mathbf{x} = \mathbf{x}_0 + t\mathbf{v},\, t \in \mathbb{R} \right\},$$

and assume that $\mathbf{x}_0$ is an accumulation point of the set $E \cap L$. The *directional derivative* of f at $\mathbf{x}_0$ in the direction $\mathbf{v}$ is defined as

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) := \lim_{t \to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t},$$

provided the limit exists in $\mathbb{R}$. In the special case in which $\mathbf{v} = \mathbf{e}_i$, the directional derivative $\frac{\partial f}{\partial \mathbf{e}_i}(\mathbf{x}_0)$, if it exists, is called the *partial derivative* of f with respect to x_i and is denoted $\frac{\partial f}{\partial x_i}(\mathbf{x}_0)$ or $f_{x_i}(\mathbf{x}_0)$ or $D_i f(\mathbf{x}_0)$.

Remark 85 *The previous definition continues to hold if in place of $\mathbb{R}^N$ one takes a normed space V, so that $f : E \to \mathbb{R}$ where $E \subseteq V$.*

Remark 86 *If $N = 1$, taking the direction $v = 1$, we have that*

$$\frac{\partial f}{\partial 1}(x_0) = \lim_{t \to 0} \frac{f(x_0 + t) - f(x_0)}{t},$$

which is the definition of derivative $\frac{df}{dx}(x_0)$ or $f'(x_0)$.

Next we show that even if the directional derivatives at $\mathbf{x}_0$ exist and are finite in every direction, then f does not have to be continuous at $\mathbf{x}_0$.

Example 87 *Let*

$$f(x, y) := \begin{cases} \frac{x^2 y}{x^4 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

Let's find the directional derivatives of f at $\mathbf{0}$. Given a direction $\mathbf{v} = (v_1, v_2)$, with $v_1^2 + v_2^2 = 1$, we have

$$f(0 + tv_1, 0 + tv_2) = 0.$$

It follows that

$$\frac{f(0+tv_1,0+tv_2)-f(0,0)}{t} = \frac{\frac{(tv_1)^2tv_2}{(tv_1)^4+(tv_2)^2}-0}{t}$$
$$= \frac{t^3v_1^2v_2}{t^5v_1^4+t^3v_2^2}.$$

If $v_2 = 0$ *then*

$$\frac{f(0+tv_1,0+tv_2)-f(0,0)}{t} = \frac{0}{t^5v_1^4+0} = 0 \to 0$$

as $t \to 0$*, so* $\frac{\partial f}{\partial x}(0,0) = 0$*. If* $v_2 \neq 0$*, then,*

$$\frac{f(0+tv_1,0+tv_2)-f(0,0)}{t} = \frac{v_1^2v_2}{t^2v_1^4+v_2^2} \to \frac{v_1^2v_2}{0+v_2^2} = \frac{v_1^2}{v_2},$$

so

$$\frac{\partial f}{\partial \mathbf{v}}(0,0) = \frac{v_1^2}{v_2}.$$

In particular, $\frac{\partial f}{\partial y}(0,0) = \frac{0}{1} = 0$*. Now let's prove that* f *is not continuous at* $\mathbf{0}$*. We have*

$$f(x,0) = \frac{0}{0+y^2} = 0 \to 0$$

as $x \to 0$*, while*

$$f\left(x,x^2\right) = \frac{x^2x^2}{x^4+x^4} = \frac{1}{2} \to \frac{1}{2}$$

as $x \to 0$*. Hence, the limit* $\lim_{(x,y)\to(0,0)} f(x,y)$ *does not exists and so* f *is not continuous at* $(0,0)$*. Note that* f *is continuous at all other points* $(x,y) \neq (0,0)$ *by Theorem 79, since* $h(x,y) = x$ *and* $g(x,y) = y$ *are continuous functions in* $\mathbb{R}^2$*.*

The previous examples show that in dimension $N \geq 2$ partial derivatives do not give the same kind of results as in the case $N = 1$. To solve this problem, we introduce a stronger notion of derivative, namely, the notion of differentiability.

We recall that a function $T : \mathbb{R}^N \to \mathbb{R}$ is *linear* if

$$T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ and

$$T(s\mathbf{x}) = sT(\mathbf{x})$$

for all $s \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^N$. Write $\mathbf{x} = \sum_{i=1}^N x_i\mathbf{e}_i$. Then by the linearity of T,

$$T(\mathbf{x}) = T\left(\sum_{i=1}^N x_i\mathbf{e}_i\right) = \sum_{i=1}^N x_iT(\mathbf{e}_i).$$

Define $\mathbf{b} := (T(\mathbf{e}_1), \ldots, T(\mathbf{e}_N)) \in \mathbb{R}^N$. Then the previous calculation shows that

$$T(\mathbf{x}) = \mathbf{b}\cdot\mathbf{x} \qquad \text{for all } \mathbf{x} \in \mathbb{R}^N.$$

Definition 88 *Let $E \subseteq \mathbb{R}^N$, let $f : E \to \mathbb{R}$, and let $\mathbf{x}_0 \in E$ be an accumulation point of E. The function f is* differentiable at $\mathbf{x}_0$ *if there exists a linear function $T : \mathbb{R}^N \to \mathbb{R}$ (depending on f and $\mathbf{x}_0$) such that*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - T(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

provided the limit exists. The function T, if it exists, is called the differential *of f at $\mathbf{x}_0$ and is denoted $df(\mathbf{x}_0)$ or $df_{\mathbf{x}_0}$.*

Exercise 89 *Prove that if $N = 1$, then f is differentiable at x_0 if and only there exists the derivative $f'(x_0) \in \mathbb{R}$.*

Remark 90 *The previous definition continues to hold if in place of $\mathbb{R}^N$ one takes a normed space V , so that $f : E \to \mathbb{R}$ where $E \subseteq V$. In this case, however, we require $T : V \to \mathbb{R}$ to be linear and continuous. Note that in $\mathbb{R}^N$ every linear function is continuous.*

The next theorem shows that differentiability in dimension $N \geq 2$ plays the same role of the derivative in dimension $N = 1$.

Theorem 91 *Let $E \subseteq \mathbb{R}^N$, let $f : E \to \mathbb{R}$, and let $\mathbf{x}_0 \in E$ be an accumulation point of E. If f is differentiable at $\mathbf{x}_0$, then f is continuous at $\mathbf{x}_0$.*

Proof. Let T be the differential of f at $\mathbf{x}_0$. Write $T(\mathbf{x}) = \mathbf{b} \cdot \mathbf{x}$. We have

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{x}_0) &= f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{b} \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{b} \cdot (\mathbf{x} - \mathbf{x}_0) \\ &= \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{b} \cdot (\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} \|\mathbf{x} - \mathbf{x}_0\| + \mathbf{b} \cdot (\mathbf{x} - \mathbf{x}_0). \end{aligned}$$

Hence, by Cauchy's inequality for $\mathbf{x} \in E$, $\mathbf{x} \neq \mathbf{x}_0$,

$$\begin{aligned} 0 &\leq |f(\mathbf{x}) - f(\mathbf{x}_0)| \leq \left| \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{b} \cdot (\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} \right| \|\mathbf{x} - \mathbf{x}_0\| + |\mathbf{b} \cdot (\mathbf{x} - \mathbf{x}_0)| \\ &\leq \left| \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{b} \cdot (\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} \right| \|\mathbf{x} - \mathbf{x}_0\| + \|\mathbf{b}\| \|\mathbf{x} - \mathbf{x}_0\| \\ &\to |0| \cdot 0 + \|\mathbf{b}\| \cdot 0 = 0 \end{aligned}$$

as $\mathbf{x} \to \mathbf{x}_0$. It follows by the squeeze theorem that f is continuous at $\mathbf{x}_0$. ■

Monday, February 11, 2013

Next we study the relation between directional derivatives and differentiability. The next theorem gives a formula for the vector $\mathbf{b}$ used in the previous proof and hence determines T. Here we need $\mathbf{x}_0$ to be an interior point of E.

Theorem 92 *Let $E \subseteq \mathbb{R}^N$, let $f : E \to \mathbb{R}$ be differentiable at some point $\mathbf{x}_0 \in E^\circ$ and T be the differential of f at $\mathbf{x}_0$. Then* *(i) all the directional derivatives of f at $\mathbf{x}_0$ exist and*

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = T(\mathbf{v}),$$

(ii) for every direction $\mathbf{v}$,

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\, v_i. \tag{6}$$

Proof. Since $\mathbf{x}_0$ is an interior point, there exists $B(\mathbf{x}_0, r) \subseteq E$. Let $\mathbf{v} \in \mathbb{R}^N$ be a direction and $\mathbf{x} = \mathbf{x}_0 + t\mathbf{v}$. Note that for $|t| < r$, we have that

$$\|\mathbf{x} - \mathbf{x}_0\| = \|\mathbf{x}_0 + t\mathbf{v} - \mathbf{x}_0\| = \|t\mathbf{v}\| = |t|\,\|\mathbf{v}\| = |t|\,1 < r$$

and so $\mathbf{x}_0 + t\mathbf{v} \in B(\mathbf{x}_0, r) \subseteq E$. Moreover, $\mathbf{x} \to \mathbf{x}_0$ as $t \to 0$ and so, since f is differentiable at $\mathbf{x}_0$,

$$0 = \lim_{\mathbf{x}\to\mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - T(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} = \lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - T(\mathbf{x}_0 + t\mathbf{v} - \mathbf{x}_0)}{\|\mathbf{x}_0 + t\mathbf{v} - \mathbf{x}_0\|}$$
$$= \lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - tT(\mathbf{v})}{|t|}.$$

Since $\frac{|t|}{t}$ is bounded by one, it follows that

$$\lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - tT(\mathbf{v})}{t} = \lim_{t\to 0} \frac{|t|}{t} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - tT(\mathbf{v})}{|t|} = 0.$$

But then

$$0 = \lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) - tT(\mathbf{v})}{t} = \lim_{t\to 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} - T(\mathbf{v}),$$

which shows that there exists $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = T(\mathbf{v})$.

Part (ii) follows from the linearity of T. Indeed, writing $\mathbf{v} = \sum_{i=1}^{N} v_i \mathbf{e}_i$, by the linearity of T,

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = T(\mathbf{v}) = T\left(\sum_{i=1}^{N} v_i \mathbf{e}_i\right) = \sum_{i=1}^{N} v_i T(\mathbf{e}_i) = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\, v_i.$$

■

Remark 93 *If in the previous theorem $\mathbf{x}_0$ is not an interior point but for some direction $\mathbf{v} \in \mathbb{R}^N$, the point $\mathbf{x}_0$ is an accumulation point of the set $E \cap L$, where L is the line through $\mathbf{x}_0$ in the direction $\mathbf{v}$, then as in the first part of the proof we can show that there exists the directional derivative $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$ and*

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = T(\mathbf{v}).$$

If all the partial derivatives of f at $\mathbf{x}_0$ exist, the vector

$$\left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \ldots, \frac{\partial f}{\partial x_N}(\mathbf{x}_0)\right) \in \mathbb{R}^N$$

is called the *gradient* of f at $\mathbf{x}_0$ and is denoted by $\nabla f(\mathbf{x}_0)$ or $\operatorname{grad} f(\mathbf{x}_0)$ or $Df(\mathbf{x}_0)$. Note that part (ii) of the previous theorem shows that

$$df_{\mathbf{x}_0}(\mathbf{v}) = T(\mathbf{v}) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v} = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\mathbf{x}_0) v_i. \tag{7}$$

for all directions $\mathbf{v}$. Hence, only at *interior points* of E, to check differentiability it is enough to prove that

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} = 0. \tag{8}$$

The next theorem gives an important sufficient condition for differentiability at a point $\mathbf{x}_0$.

Theorem 94 *Let $E \subseteq \mathbb{R}^N$, let $f : E \to \mathbb{R}$, let $\mathbf{x}_0 \in E^\circ$. Assume that there exists $r > 0$ such that $B(\mathbf{x}_0, r) \subseteq E$ and the partial derivatives $\frac{\partial f}{\partial x_j}$, $j = 1, \ldots, N$, exist for every $\mathbf{x} \in B(\mathbf{x}_0, r)$ and are continuous at $\mathbf{x}_0$. Then f is differentiable at $\mathbf{x}_0$.*

Wednesday, February 13, 2013

Example 95 *Let*

$$f(x,y) := \begin{cases} \frac{x^2|y|}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Let's study continuity, partial derivatives and differentiability. For $(x,y) \neq (0,0)$, we have that f is continuous by Theorem 79, while for $(x,y) = (0,0)$, we need to check that

$$\lim_{(x,y) \to (0,0)} f(x,y) = f(0,0).$$

We have

$$0 \leq |f(x,y) - f(0,0)| = \left| \frac{x^2|y|}{x^2+y^2} - 0 \right| = \frac{x^2|y|}{x^2+y^2} \leq \frac{(x^2+y^2)|y|}{x^2+y^2} = |y| \to 0$$

as $(x,y) \to (0,0)$. Hence, f is continuous at $(0,0)$.

Next, let's study partial derivatives. For $(x,y) \neq (0,0)$, by the quotient rule, we have

$$\frac{\partial f}{\partial x}(x,y) = \frac{2x|y|(x^2+y^2) - x^2|y|(2x+0)}{(x^2+y^2)^2}, \tag{9}$$

while for $(x,y)=(0,0)$,

$$\frac{\partial f}{\partial x}(0,0)=\lim_{t\to0}\frac{f(0+t1,0+t0)-f(0,0)}{t}=\lim_{t\to0}\frac{\frac{t^2|0|}{t^2+0}-0}{t}=\lim_{t\to0}\frac{0}{t^3}=0.$$

For $y\neq0$, *by the quotient rule, we have*

$$\frac{\partial f}{\partial y}(x,y)=\frac{x^2\frac{y}{|y|}\left(x^2+y^2\right)-x^2|y|\left(0+y^2\right)}{\left(x^2+y^2\right)^2}, \tag{10}$$

while at a point $(x_0,0)$,

$$\frac{\partial f}{\partial y}(x_0,0)=\lim_{t\to0}\frac{f(x_0+t0,0+t1)-f(x_0,0)}{t}=\lim_{t\to0}\frac{\frac{x_0^2|t|}{x_0^2+t^2}-0}{t}=\lim_{t\to0}\frac{|t|}{t}\frac{x_0^2}{x_0^2+t^2}.$$

If $x_0=0$, *then* $\frac{|t|}{t}\frac{x_0^2}{x_0^2+t^2}=\frac{|t|}{t}\frac{0}{0+t^2}=0\to0$ *as* $t\to0$, *so* $\frac{\partial f}{\partial y}(0,0)=0$, *while if* $x_0\neq0$, *we have*

$$\lim_{t\to0^+}\frac{|t|}{t}\frac{x_0^2}{x_0^2+t^2}=\lim_{t\to0^+}\frac{t}{t}\frac{x_0^2}{x_0^2+t^2}=\lim_{t\to0^+}\frac{x_0^2}{x_0^2+t^2}=\frac{x_0^2}{x_0^2+0}=1,$$
$$\lim_{t\to0^-}\frac{|t|}{t}\frac{x_0^2}{x_0^2+t^2}=\lim_{t\to0^+}\frac{-t}{t}\frac{x_0^2}{x_0^2+t^2}=-\lim_{t\to0^+}\frac{x_0^2}{x_0^2+t^2}=-\frac{x_0^2}{x_0^2+0}=-1.$$

Hence, $\frac{\partial f}{\partial y}(x_0,0)$ *does not exist at* $(x_0,0)$ *for* $x_0\neq0$, *and so by Theorem 92,* f *is not differentiable at* $(x_0,0)$ *for* $x_0\neq0$.

On the other hand, at points (x,y) *with* $y\neq0$, *we have that* $\frac{\partial f}{\partial x}$ *and* $\frac{\partial f}{\partial y}$ *exist in a small ball centered at* (x,y) *(see (9) and (10)) and they are continuous by Theorem 79. Hence, we can apply Theorem 94 to conclude that* f *is differentiable at all points* (x,y) *with* $y\neq0$.

It remains to study differentiability at $(0,0)$. *By 8, we need to prove that*

$$\lim_{(x,y)\to(0,0)}\frac{f(x,y)-f(0,0)-\nabla f(0,0)\cdot((x,y)-(0,0))}{\|(x,y)-(0,0)\|}=0.$$

We have

$$\frac{f(x,y)-f(0,0)-\nabla f(0,0)\cdot((x,y)-(0,0))}{\|(x,y)-(0,0)\|}=\frac{\frac{x^2|y|}{x^2+y^2}-0-(0,0)\cdot((x,y)-(0,0))}{\sqrt{x^2+y^2}}$$
$$=\frac{x^2|y|}{\left(x^2+y^2\right)^{3/2}}.$$

Taking $y=x$, *with* $x>0$, *we get*

$$\frac{x^2|x|}{\left(x^2+x^2\right)^{3/2}}=\frac{x^2x^3}{\left(x^2+x^2\right)^{3/2}}=\frac{1}{(2)^{3/2}}\nrightarrow0.$$

Hence, f *is not differentiable at* $(0,0)$.

Friday, February 15, 2013

The next exercise shows that the conditions in Theorem 94 are sufficient but not necessary for differentiability.

Example 96 *Let*

$$f(x,y) := \begin{cases} \left(x^2+y^2\right)\sin\frac{1}{x+y} & \text{if } (x,y)\neq(0,0) \text{ or } x+y\neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Let's study continuity, partial derivatives and differentiability. For $x+y \neq 0$, we have that f is continuous by Theorem 79 and Theorem 81. For $(x,y)=(0,0)$, we need to check that

$$\lim_{(x,y)\to(0,0)} f(x,y) = f(0,0).$$

We have

$$0 \leq |f(x,y) - f(0,0)| = \left|\left(x^2+y^2\right)\sin\frac{1}{x+y} - 0\right| \leq \left(x^2+y^2\right)1 \to 0$$

as $(x,y)\to(0,0)$. Hence, f is continuous at $(0,0)$. At a point $(x_0,-x_0)$ with $x_0 \neq 0$, taking $x = x_0$ and $y = -x_0 + t$, where $t \to 0$. We have

$$\begin{aligned} f(x_0,-x_0+t) &= \left(x_0^2 + (-x_0+t)^2\right)\sin\frac{1}{x_0 - x_0 + t} \\ &= \left(x_0^2 + (-x_0+t)^2\right)\sin\frac{1}{t}. \end{aligned}$$

Take $t = \frac{1}{\frac{\pi}{2}+2n\pi}$. Then

$$\begin{aligned} f\left(x_0, -x_0 + \frac{1}{\frac{\pi}{2}+2n\pi}\right) &= \left(x_0^2 + \left(-x_0 + \frac{1}{\frac{\pi}{2}+2n\pi}\right)^2\right)\sin\left(\frac{\pi}{2}+2n\pi\right) \\ &= \left(x_0^2 + \left(-x_0 + \frac{1}{\frac{\pi}{2}+2n\pi}\right)^2\right)1 \to \left(x_0^2 + (-x_0+0)^2\right) = 2x_0^2 \neq 0 \end{aligned}$$

as $n\to\infty$. Hence f is not continuous at $(x_0,-x_0)$ with $x_0 \neq 0$, and so by Theorem 91, f is not differentiable at $(x_0,-x_0)$ with $x_0 \neq 0$. Next, let's study partial derivatives. If $x+y \neq 0$,

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= (2x+0)\sin\frac{1}{x+y} + \left(x^2+y^2\right)\left(\cos\frac{1}{x+y}\right)\left(-\frac{1}{(x+y)^2}\right), \\ \frac{\partial f}{\partial y}(x,y) &= (0+2y)\sin\frac{1}{x+y} + \left(x^2+y^2\right)\left(\cos\frac{1}{x+y}\right)\left(-\frac{1}{(x+y)^2}\right). \end{aligned}$$

Since for every (x,y), with $x+y\neq 0$, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist in a small ball centered at (x,y) and they are continuous by Theorem 79 and Theorem 81, we can apply Theorem 94 to conclude that f is differentiable at all points (x,y) with $x+y\neq 0$.

Next, let"s study the partial derivatives at $(0,0)$*. We have*

$$\frac{f(x,0)-f(0,0)}{x-0}=\frac{\left(x^2+0\right)\sin\frac{1}{x+0}-0}{x-0}=x\sin\frac{1}{x}\to 0$$

as $x\to 0$*, since* $\sin\frac{1}{x}$ *is bounded. Hence,* $\frac{\partial f}{\partial x}(0,0)=0$*. Similarly,* $\frac{\partial f}{\partial y}(0,0)=0$*.*

For $(x_0,-x_0)$*, with* $x_0\neq 0$*,*

$$\begin{aligned}\frac{\partial f}{\partial x}(x_0,-x_0)&=\lim_{t\to 0}\frac{f(x_0+t0,-x_0+t1)-f(x_0,-x_0)}{t}=\lim_{t\to 0}\frac{\left(x_0^2+(-x_0+t)^2\right)\sin\frac{1}{x_0-x_0+t}-0}{t}\\&=\lim_{t\to 0}\frac{\left(x_0^2+(-x_0+t)^2\right)\sin\frac{1}{t}-0}{t}.\end{aligned}$$

Taking $t=\frac{1}{\frac{\pi}{2}+2n\pi}$ *and* $t=\frac{1}{2n\pi}$ *shows that this limit does not exist. Indeed,*

$$\frac{\left(x_0^2+\left(-x_0+\frac{1}{\frac{\pi}{2}+2n\pi}\right)^2\right)\sin\frac{1}{\frac{1}{\frac{\pi}{2}+2n\pi}}}{\frac{1}{\frac{\pi}{2}+2n\pi}}=\frac{\left(x_0^2+\left(-x_0+\frac{1}{\frac{\pi}{2}+2n\pi}\right)^2\right)1}{\frac{1}{\frac{\pi}{2}+2n\pi}}\to\frac{2x_0^2}{0^+}=\infty$$

as $n\to\infty$*, while*

$$\frac{\left(x_0^2+\left(-x_0+\frac{1}{2n\pi}\right)^2\right)\sin\frac{1}{\frac{1}{2n\pi}}}{\frac{1}{2n\pi}}=\frac{\left(x_0^2+\left(-x_0+\frac{1}{2n\pi}\right)^2\right)0}{\frac{1}{2n\pi}}=\frac{0}{\frac{1}{2n\pi}}=0\to 0$$

as $n\to\infty$*. Similarly,* $\frac{\partial f}{\partial y}(x_0,-x_0)$ *does not exist.*

Let's prove that f *is differentiable at* $(0,0)$ *at* $(0,0)$*. We have,*

$$\begin{aligned}\frac{f(x,0)-f(0,0)-\frac{\partial f}{\partial x}(0,0)(x-0)-\frac{\partial f}{\partial y}(0,0)(y-0)}{\sqrt{(x-0)^2+(y-0)^2}}&=\begin{cases}\frac{\left(x^2+y^2\right)\sin\frac{1}{x+y}}{\sqrt{x^2+y^2}} & \text{if } x+y\neq 0,\\ 0 & \text{otherwise.}\end{cases}\\&=\begin{cases}\sqrt{x^2+y^2}\sin\frac{1}{x+y} & \text{if } x+y\neq 0,\\ 0 & \text{otherwise.}\end{cases}\\&\to 0\end{aligned}$$

as $(x,y)\to(0,0)$ *since* $\sin\frac{1}{x+y}$ *is bounded and* $\sqrt{x^2+y^2}\to 0$*. Hence,* f *is differentiable at* $(0,0)$*, even if nearby* $\frac{\partial f}{\partial x}$ *and* $\frac{\partial f}{\partial y}$ *do not always exist.*

8 Higher Order Derivatives

Let $E\subseteq\mathbb{R}^N$, let $f:E\to\mathbb{R}$ and let $\mathbf{x}_0\in E$. Let $i\in\{1,\dots,N\}$ and assume that there exists the partial derivatives $\frac{\partial f}{\partial x_i}(\mathbf{x})$ for all $\mathbf{x}\in E$. If $j\in\{1,\dots,N\}$ and $\mathbf{x}_0$ is an accumulation point of $E\cap L$, where L is the line through $\mathbf{x}_0$ in the direction $\mathbf{e}_j$, then we can consider the partial derivative of the function $\frac{\partial f}{\partial x_i}$ with respect to x_j, that is,

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) =: \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Note that in general the order in which we take derivatives is important.

Example 97 *Let*

$$f(x,y) := \begin{cases} y^2 \arctan \frac{x}{y} & \text{if } y \neq 0, \\ 0 & \text{if } y = 0. \end{cases}$$

If $y \neq 0$, then

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= \frac{\partial}{\partial x}\left(y^2 \arctan \frac{x}{y}\right) = y^2 \frac{1}{1+\left(\frac{x}{y}\right)^2} \frac{\partial}{\partial x}\left(\frac{x}{y}\right) \\ &= \frac{y^3}{x^2+y^2},\end{aligned}$$

and

$$\begin{aligned}\frac{\partial f}{\partial y}(x,y) &= \frac{\partial}{\partial y}\left(y^2 \arctan \frac{x}{y}\right) = 2y \arctan \frac{x}{y} + y^2 \frac{1}{1+\left(\frac{x}{y}\right)^2} \frac{\partial}{\partial y}\left(\frac{x}{y}\right) \\ &= 2y \arctan \frac{x}{y} - \frac{xy^2}{x^2+y^2},\end{aligned}$$

while at points $(x_0, 0)$ we have:

$$\begin{aligned}\frac{\partial f}{\partial x}(x_0,0) &= \lim_{t\to 0} \frac{f(x_0+t,0)-f(x_0,0)}{t} = \lim_{t\to 0} \frac{0-0}{t} = 0, \\ \frac{\partial f}{\partial y}(x_0,0) &= \lim_{t\to 0} \frac{f(x_0,0+t)-f(x_0,0)}{t} = \lim_{t\to 0} \frac{t^2 \arctan \frac{x_0}{t} - 0}{t} \\ &= \lim_{t\to 0} t \arctan \frac{x_0}{t} - = 0,\end{aligned}$$

where we have used the fact that $\arctan \frac{x_0}{t}$ *is bounded and $t \to 0$. Thus,*

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{y^3}{x^2+y^2} & \text{if } y \neq 0, \\ 0 & \text{if } y = 0, \end{cases} \qquad \frac{\partial f}{\partial y}(x,y) = \begin{cases} 2y \arctan \frac{x}{y} - \frac{xy^2}{x^2+y^2} & \text{if } y \neq 0, \\ 0 & \text{if } y = 0. \end{cases}$$

To find $\frac{\partial^2 f}{\partial y \partial x}(0,0)$, we calculate

$$\begin{aligned}\frac{\partial^2 f}{\partial y \partial x}(0,0) &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0) = \lim_{t\to 0} \frac{\frac{\partial f}{\partial x}(0,0+t) - \frac{\partial f}{\partial x}(0,0)}{t} \\ &= \lim_{t\to 0} \frac{\frac{t^3}{0+t^2} - 0}{t} = \lim_{t\to 0} 1 = 1,\end{aligned}$$

while

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0) = \lim_{t\to 0} \frac{\frac{\partial f}{\partial y}(0+t,0) - \frac{\partial f}{\partial y}(0,0)}{t}$$
$$= \lim_{t\to 0} \frac{0-0}{t} = \lim_{t\to 0} 0 = 0.$$

Hence, $\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$.

Exercise 98 *Let*

$$f(x,y) := \begin{cases} \frac{x^3 y - x y^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Prove that $\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$.

The next important theorem shows that at "good points" the order in which we take derivatives does not matter.

Theorem 99 (Schwartz) *Let* $E \subseteq \mathbb{R}^N$*, let* $f : E \to \mathbb{R}$*, let* $\mathbf{x}_0 \in E^\circ$*, and let* $i, j \in \{1, \ldots, N\}$*. Assume that there exists* $r > 0$ *such that* $B(\mathbf{x}_0, r) \subseteq E$ *and for all* $\mathbf{x} \in B(\mathbf{x}_0, r)$*, the partial derivatives* $\frac{\partial f}{\partial x_i}(\mathbf{x})$*,* $\frac{\partial f}{\partial x_j}(\mathbf{x})$*, and* $\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x})$ *exist. Assume also that* $\frac{\partial^2 f}{\partial x_j \partial x_i}$ *is continuous at* $\mathbf{x}_0$*. Then there exists* $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)$ *and*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}_0).$$

Monday, February 18, 2013

Next we prove Taylor's formula in higher dimensions. We set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. A *multi-index* $\boldsymbol{\alpha}$ is a vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N) \in (\mathbb{N}_0)^N$. The *length* of a multi-index is defined as

$$|\boldsymbol{\alpha}| := \alpha_1 + \cdots + \alpha_N.$$

Given a multi-index $\boldsymbol{\alpha}$, the partial derivative $\frac{\partial^{\boldsymbol{\alpha}}}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}$ is defined as

$$\frac{\partial^{\boldsymbol{\alpha}}}{\partial \mathbf{x}^{\boldsymbol{\alpha}}} := \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial x_1^{\alpha_1} \cdots \partial x_N^{\alpha_N}},$$

where $\mathbf{x} = (x_1, \ldots, x_N)$. If $\boldsymbol{\alpha} = \mathbf{0}$, we set $\frac{\partial^{\mathbf{0}} f}{\partial \mathbf{x}^{\mathbf{0}}} := f$.

Example 100 *If* $N = 3$ *and* $\boldsymbol{\alpha} = (2,1,0)$*, then*

$$\frac{\partial^{(2,1,0)}}{\partial (x,y,z)^{(2,1,0)}} = \frac{\partial^3}{\partial x^2 \partial y}.$$

Given a multi-index $\boldsymbol{\alpha}$ and $\mathbf{x} \in \mathbb{R}^N$, we set

$$\boldsymbol{\alpha}! := \alpha_1! \cdots \alpha_N!, \quad \mathbf{x}^{\boldsymbol{\alpha}} := x_1^{\alpha_1} \cdots x_N^{\alpha_N}.$$

If $\boldsymbol{\alpha} = \mathbf{0}$, we set $\mathbf{x}^{\mathbf{0}} := 1$.

Using this notation, we can extend the binomial theorem.

Theorem 101 (Multinomial Theorem) *Let $\mathbf{x} = (x_1, \ldots, x_N) \in \mathbb{R}^N$ and let $n \in \mathbb{N}$. Then*

$$(x_1 + \cdots + x_N)^n = \sum_{\boldsymbol{\alpha} \text{ multi-index}, \ |\boldsymbol{\alpha}| = n} \frac{n!}{\boldsymbol{\alpha}!} \mathbf{x}^{\boldsymbol{\alpha}}.$$

Proof. Exercise. ■

Given an open set $U \subseteq \mathbb{R}^N$, for every nonnegative integer $m \in \mathbb{N}_0$, we denote by $C^m(U)$ the space of all functions that are continuous together with their partial derivatives up to order m. We set $C^\infty(U) := \bigcap_{m=0}^{\infty} C^m(U)$.

Theorem 102 (Taylor's Formula) *Let $U \subseteq \mathbb{R}^N$ be an open set, let $f \in C^m(U)$, $m \in \mathbb{N}$, and let $\mathbf{x}_0 \in U$. Then for every $\mathbf{x} \in U$,*

$$f(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \text{ multi-index}, \ 0 \le |\boldsymbol{\alpha}| \le m} \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{\boldsymbol{\alpha}} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)^{\boldsymbol{\alpha}} + R_m(\mathbf{x}),$$

where

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{R_m(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^m} = 0.$$

We write $R_m(\mathbf{x}) = o(\|\mathbf{x} - \mathbf{x}_0\|^m)$ and we say that R_m is a little o *of $\|\mathbf{x} - \mathbf{x}_0\|^m$ as $\mathbf{x} \to \mathbf{x}_0$.*

Example 103 *Let's calculate the limit*

$$\lim_{(x,y) \to (0,0)} \frac{(1+x)^y - 1}{\sqrt{x^2 + y^2}}.$$

By substituting we get $\frac{0}{0}$. Consider the function

$$f(x,y) = (1+x)^y - 1 = e^{\log(1+x)^y} - 1 = e^{y \log(1+x)} - 1,$$

which is defined in the set $U := \{(x,y) \in \mathbb{R}^2 : 1 + x > 0\}$. The function f is of class C^∞. Let's use Taylor's formula of order $m = 1$ at $(0,0)$,

$$f(x,y) = f(0,0) + \frac{\partial f}{\partial x}(0,0)(x - 0) + \frac{\partial f}{\partial y}(0,0)(y - 0) + o\left(\sqrt{x^2 + y^2}\right).$$

We have

$$\frac{\partial f}{\partial x}(x,y) = \frac{\partial}{\partial x}\left(e^{y\log(1+x)}-1\right) = e^{y\log(1+x)}y\frac{1}{1+x},$$
$$\frac{\partial f}{\partial y}(x,y) = \frac{\partial}{\partial y}\left(e^{y\log(1+x)}-1\right) = e^{y\log(1+x)}y\log(1+x),$$

and so

$$f(x,y) = 0 + 0(x-0) + 0(y-0) + o\left(\sqrt{x^2+y^2}\right),$$

which means that

$$\lim_{(x,y)\to(0,0)} \frac{(1+x)^y-1}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{o\left(\sqrt{x^2+y^2}\right)}{\sqrt{x^2+y^2}} = 0.$$

Note that if we had to calculate the limit

$$\lim_{(x,y)\to(0,0)} \frac{(1+x)^y-1}{x^2+y^2},$$

then we would need Taylor's formula of order $m=2$ at $(0,0)$,

$$\begin{aligned} f(x,y) = f(0,0) &+ \frac{\partial f}{\partial x}(0,0)(x-0) + \frac{\partial f}{\partial y}(0,0)(y-0) \\ &+ \frac{1}{(2,0)!}\frac{\partial^2 f}{\partial x^2}(0,0)(x-0)^2 + \frac{1}{(1,1)!}\frac{\partial^2 f}{\partial x \partial y}(0,0)(x-0)(y-0) \\ &+ \frac{1}{(0,2)!}\frac{\partial^2 f}{\partial y^2}(0,0)(y-0)^2 + o\left(x^2+y^2\right). \end{aligned}$$

Another simpler method would be to use the Taylor's formulas for e^t and $\log(1+s)$.

Wednesday, February 20, 2013

First midterm

Friday, February 22, 2013

Solutions First midterm.

Example 104 (Example 103, continued) ***Second method:*** *If either $x = 0$ or $y = 0$, we get*

$$\frac{(1+x)^y-1}{\sqrt{x^2+y^2}} = \frac{0}{\sqrt{x^2+y^2}} = 0.$$

If $x \neq 0$ and $y \neq 0$, then

$$\frac{(1+x)^y-1}{\sqrt{x^2+y^2}} = \frac{e^{y\log(1+x)}-1}{y\log(1+x)}\frac{\log(1+x)}{x}\frac{xy}{\sqrt{x^2+y^2}}.$$

Now, using the limits $\lim_{t\to 0}\frac{e^t-1}{t} = 1$ and $\lim_{t\to 0}\frac{\log(1+t)}{t} = 1$, we have

$$\lim_{(x,y)\to(0,0)} \frac{e^{y\log(1+x)}-1}{y\log(1+x)} = 1, \quad \frac{\log(1+x)}{x},$$

while

$$0 \leq \left|\frac{xy}{\sqrt{x^2+y^2}}\right| = \frac{\sqrt{x^2}\,|y|}{\sqrt{x^2+y^2}} \leq \frac{\sqrt{x^2+y^2}\,|y|}{\sqrt{x^2+y^2}} = |y| \to 0$$

as $(x,y) \to (0,0)$.

Monday, February 25, 2013

Example 105 *Let's calculate*

$$\lim_{(x,y)\to(0,0)} \frac{\log\left(1+\sin^2(xy)\right) - x^2y^2}{\left(x^2+y^2\right)^2}$$

Taylor's formula of $\sin t$ *of order one is given by*

$$\sin t = t + o\left(t^2\right)$$

and so

$$\begin{aligned}\sin^2 t &= \left(t+o\left(t^2\right)\right)^2 = t^2 + \left(o\left(t^2\right)\right)^2 + 2to\left(t^2\right) \\ &= t^2 + o\left(t^3\right)\end{aligned}$$

where we have used the properties of the little o. Hence

$$\log\left(1+\sin^2 t\right) = \log\left(1+t^2+o\left(t^3\right)\right),$$

Let's use now Taylor's formula

$$\log(1+s) = s + o(s),$$

where for us $s = \sin^2 t = t^2 + o\left(t^3\right)$. *We get*

$$\begin{aligned}&\log\left(1+\sin^2 t\right) = \log\left(1+t^2+o\left(t^3\right)\right) \\ &= \left(t^2+o\left(t^3\right)\right) + o\left(t^2+o\left(t^3\right)\right) = t^2 + o\left(t^2\right).\end{aligned}$$

Hence,

$$\begin{aligned}\frac{\log\left(1+\sin^2(xy)\right) - x^2y^2}{\left(x^2+y^2\right)^2} &= \frac{x^2y^2 + o\left(x^2y^2\right) - x^2y^2}{\left(x^2+y^2\right)^2} \\ &= \frac{o\left(x^2y^2\right)}{\left(x^2+y^2\right)^2} = \frac{x^2y^2}{\left(x^2+y^2\right)^2}\frac{o\left(x^2y^2\right)}{x^2y^2}\end{aligned}$$

if $x \neq 0$ *and* $y \neq 0$. *Now*

$$0 \leq \frac{x^2y^2}{\left(x^2+y^2\right)^2} \leq \frac{1}{2},$$

and so

$$\frac{x^2y^2}{\left(x^2+y^2\right)^2}\frac{o\left(x^2y^2\right)}{x^2y^2} \to 0$$

by Theorem while if either $x = 0$ or $y = 0$, we get

$$\frac{\log\left(1+\sin^2(xy)\right)-x^2y^2}{\left(x^2+y^2\right)^2}=\frac{0}{\left(x^2+y^2\right)^2}=0.$$

Hence,

$$\lim_{(x,y)\to(0,0)}\frac{\log\left(1+\sin^2(xy)\right)-x^2y^2}{\left(x^2+y^2\right)^2}=0$$

Exercise 106 *Calculate the limit*

$$\lim_{(x,y)\to(0,0)}\frac{\log\left(1+\sin^2(xy)\right)-x^2y^2}{\left(x^2+y^2\right)^4}.$$

Wednesday, February 27, 2013

9 Local Minima and Maxima

Definition 107 *Let $f : E \to \mathbb{R}$, where $E \subseteq \mathbb{R}^N$, and let $\mathbf{x}_0 \in E$. We say that*

(i) f attains a local minimum *at $\mathbf{x}_0$ if there exists $r > 0$ such that $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ for all $\mathbf{x} \in E \cap B(\mathbf{x}_0, r)$,*

(ii) f attains a global minimum *at $\mathbf{x}_0$ if $f(\mathbf{x}) \geq f(\mathbf{x}_0)$ for all $\mathbf{x} \in E$,*

(iii) f attains a local maximum *at $\mathbf{x}_0$ if there exists $r > 0$ such that $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ for all $\mathbf{x} \in E \cap B(\mathbf{x}_0, r)$,*

(iv) f attains a global maximum *at $\mathbf{x}_0$ if $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ for all $\mathbf{x} \in E$.*

Theorem 108 *Let $f : E \to \mathbb{R}$, where $E \subseteq \mathbb{R}^N$. Assume that f attains a local minimum (or maximum) at some point $\mathbf{x}_0 \in E$. If there exists a direction $\mathbf{v}$ and $\delta > 0$ such that the set*

$$\{\mathbf{x}_0 + t\mathbf{v} : t \in (-\delta, \delta)\} \subseteq E \tag{11}$$

and if there exists $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0)$, then necessarily, $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}_0) = 0$. In particular, if $\mathbf{x}_0$ is an interior point of E and f is differentiable at $\mathbf{x}_0$, then all partial derivatives and directional derivatives of f at $\mathbf{x}_0$ are zero.

Proof. Exercise. ■

Remark 109 *In view of the previous theorem, when looking for local minima and maxima, we have to search among the following:*

- *Interior points at which f is differentiable and $\nabla f(\mathbf{x}) = \mathbf{0}$, these are called* critical points*. Note that if $\nabla f(\mathbf{x}_0) = \mathbf{0}$, the function f may not attain a local minimum or maximum at $\mathbf{x}_0$. Indeed, consider the function $f(x) = x^3$. Then $f'(0) = 0$, but f is strictly increasing, and so f does not attain a local minimum or maximum at 0.*

- *Interior points at which f is not differentiable. The function $f(x) = |x|$ attains a global minimum at $x = 0$, but f is not differentiable at $x = 0$.*
- *Boundary points.*

To find sufficient conditions for a critical point to be a point of local minimum or local maximu, we study the second order derivatives of f.

Definition 110 *Let $f : E \to \mathbb{R}$, where $E \subseteq \mathbb{R}^N$, and let $\mathbf{x}_0 \in E$. The Hessian matrix of f at $\mathbf{x}_0$ is the $N \times N$ matrix*

$$
\begin{aligned}
H_f(\mathbf{x}_0) :&= \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_N \partial x_1}(\mathbf{x}_0) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_N}(\mathbf{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_N^2}(\mathbf{x}_0) \end{pmatrix} \\
&= \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) \right)_{i,j=1}^N ,
\end{aligned}
$$

whenever it is defined.

Remark 111 *If the hypotheses of Schwartz's theorem are satisfied for all $i, j = 1, \ldots, N$, then*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}_0),$$

which means that the Hessian matrix $H_f(\mathbf{x}_0)$ is symmetric.

Given an $N \times N$ matrix H, the *characteristic polynomial* of H is the polynomial

$$p(t) := \det(tI_N - H), \quad t \in \mathbb{R}.$$

Theorem 112 *Let H be an $N \times N$ matrix. If H is symmetric, then all roots of the characteristic polynomial are real.*

Theorem 113 *Given a polynomial of the form*

$$p(t) = t^N + a_{N-1}t^{N-1} + a_{N-2}t^{N-2} + \cdots + a_1 t + a_0, \quad t \in \mathbb{R},$$

where the coefficients a_i are real for every $i = 0, \ldots, N-1$, assume that all roots of p are real. Then

(i) all roots of p are positive if and only if the coefficients alternate sign, that is, $a_{N-1} < 0$, $a_{N-2} > 0$, $a_{N-3} < 0$, etc.

(ii) all roots of p are negative if and only $a_i > 0$ for every $i = 0, \ldots, N-1$.

The next theorem gives necessary and sufficient conditions for a point to be of local minimum or maximum.

Theorem 114 *Let $U \subseteq \mathbb{R}^N$ be open, let $f : U \to \mathbb{R}$ be of class $C^2(U)$ and let $\mathbf{x}_0 \in U$ be a critical point of f.*

(i) If $H_f(\mathbf{x}_0)$ is positive definite, then f attains a local minimum at $\mathbf{x}_0$,

(ii) if f attains a local minimum at $\mathbf{x}_0$, then $H_f(\mathbf{x}_0)$ is positive semidefinite,

(iii) if $H_f(\mathbf{x}_0)$ is negative definite, then f attains a local maximum at $\mathbf{x}_0$,

(iv) if f attains a local maximum at $\mathbf{x}_0$, then $H_f(\mathbf{x}_0)$ is negative semidefinite.

Remark 115 *Note that in view of the previous theorem, if at a critical point $\mathbf{x}_0$ the characteristic polynomial of $H_f(\mathbf{x}_0)$ has one positive root and one negative root, then f does not admit a local minimum or a local maximum at $\mathbf{x}_0$.*

Theorem 114 shows that if $H_f(\mathbf{x}_0)$ is positive definite, then f admits a local minimum at $\mathbf{x}_0$. The following exercise shows that we cannot weakened this hypothesis to $H_f(\mathbf{x}_0)$ positive semidefinite.

Example 116 *Let $f(x,y) := x^2 - y^4$. Let's find the critical points of f. We have*

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= \frac{\partial}{\partial x}\left(x^2 - y^4\right) = 2x - 0 = 0, \\ \frac{\partial f}{\partial y}(x,y) &= \frac{\partial}{\partial y}\left(x^2 - y^4\right) = 0 - 4y^3 = 0. \end{aligned}$$

Hence, $(0,0)$ is the only critical point. Let's find the Hessian matrix at these points. Note that the function is of class C^∞, so we can apply Schwartz's theorem. We have

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x,y) &= \frac{\partial}{\partial x}(2x-) = 2, \\ \frac{\partial^2 f}{\partial y^2}(x,y) &= \frac{\partial}{\partial y}\left(-4y^3\right) = -12y^2, \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) &= \frac{\partial}{\partial y}(2x) = 0. \end{aligned}$$

Hence,

$$H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix},$$

so that

$$\begin{aligned} 0 &= \det\left(\lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - H_f(0,0)\right) \\ &= \det\begin{pmatrix} \lambda - 2 & 0 \\ 0 & \lambda - 0 \end{pmatrix} = (\lambda - 2)(\lambda - 0) - 0 \\ &= (\lambda - 2)\lambda, \end{aligned}$$

and so the roots are $\lambda = 2$ or $\lambda = 0$. Hence, at $(0,0)$ we cannot have a local maximum. But it could be a local minimum. However, taking

$$f(0,y) = -y^4,$$

which has a strict maximum at $y = 0$. This shows that f does not admit a local minimum or a local maximum at $(0,0)$.Theorem 114 shows that if $H_f(\mathbf{x}_0)$ is positive definite, then f admits a local minimum at $\mathbf{x}_0$. The following exercise shows that we cannot weakened this hypothesis to $H_f(\mathbf{x}_0)$ positive semidefinite.

A critical point at which f does not admit a local minimum or a local maximum is called a *saddle point.*

Friday, March 1, 2013

Example 117 *Consider the function*

$$f(x,y) := (x-y)\, e^{-x^2-y^2}$$

defined in the set

$$E := \overline{B((0,0),1)}.$$

Let's find the critical points of f. We have

$$\frac{\partial f}{\partial x}(x,y) = \frac{\partial}{\partial x}\left((x-y)\, e^{-x^2-y^2}\right) = e^{-x^2-y^2}\left(-2x^2+2yx+1\right) = 0,$$
$$\frac{\partial f}{\partial y}(x,y) = \frac{\partial}{\partial y}\left((x-y)\, e^{-x^2-y^2}\right) = -e^{-x^2-y^2}\left(-2y^2+2xy+1\right) = 0.$$

By subtracting the two equations, we get

$$-x^2+y^2 = 0,$$
$$-2y^2+2xy+1 = 0.$$

If $x = y$, then $-2x^2+2x^2+1 = 0$, which has no solutions, while if $x = -y$, then $-4y^2+1 = 0$. Hence, the critical points are $\left(\frac{1}{2}, -\frac{1}{2}\right)$ and $\left(-\frac{1}{2}, \frac{1}{2}\right)$. Let's find the Hessian matrix at these points. Note that the function is of class C^∞, so we can apply Schwartz's theorem. We have

$$\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial}{\partial x}\left(e^{-x^2-y^2}\left(-2x^2+2yx+1\right)\right) = 2e^{-x^2-y^2}\left(2x^3-2yx^2-3x+y\right),$$
$$\frac{\partial^2 f}{\partial y^2}(x,y) = \frac{\partial}{\partial y}\left(-e^{-x^2-y^2}\left(-2y^2+2xy+1\right)\right) = -2e^{-x^2-y^2}\left(2y^3-2xy^2-3y+x\right),$$
$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial}{\partial y}\left(e^{-x^2-y^2}\left(-2x^2+2yx+1\right)\right) = 2e^{-x^2-y^2}\left(2x^2y-2xy^2+x-y\right).$$

Hence,

$$H_f\left(\frac{1}{2}, -\frac{1}{2}\right) = \begin{pmatrix} -3e^{-\frac{1}{2}} & e^{-\frac{1}{2}} \\ e^{-\frac{1}{2}} & -3e^{-\frac{1}{2}} \end{pmatrix},$$

so that

$$
\begin{aligned}
0 &= \det\left(\lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - H_f\left(\frac{1}{2}, -\frac{1}{2}\right)\right) \\
&= \det\begin{pmatrix} \lambda + 3e^{-\frac{1}{2}} & -e^{-\frac{1}{2}} \\ -e^{-\frac{1}{2}} & \lambda + 3e^{-\frac{1}{2}} \end{pmatrix} = \left(\lambda + 3e^{-\frac{1}{2}}\right)^2 - e^{-1} \\
&= \lambda^2 + 6e^{-\frac{1}{2}}\lambda + 8e^{-1},
\end{aligned}
$$

and so the eigenvalues are both negative. Hence, at $\left(\frac{1}{2}, -\frac{1}{2}\right)$ *we have a local maximum. On the other hand*

$$
H_f\left(-\frac{1}{2}, \frac{1}{2}\right) = \begin{pmatrix} 3e^{-\frac{1}{2}} & -e^{-\frac{1}{2}} \\ -e^{-\frac{1}{2}} & 3e^{-\frac{1}{2}} \end{pmatrix},
$$

so that

$$
\begin{aligned}
0 &= \det\left(\lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - H_f\left(-\frac{1}{2}, +\frac{1}{2}\right)\right) \\
&= \det\begin{pmatrix} \lambda - 3e^{-\frac{1}{2}} & -e^{-\frac{1}{2}} \\ -e^{-\frac{1}{2}} & \lambda - 3e^{-\frac{1}{2}} \end{pmatrix} = \left(\lambda - 3e^{-\frac{1}{2}}\right)^2 - e^{-1} \\
&= \lambda^2 - 6e^{-\frac{1}{2}}\lambda + 8e^{-1},
\end{aligned}
$$

and so the eigenvalues are all positive by Theorem 113. Hence, at $\left(\frac{1}{2}, -\frac{1}{2}\right)$ *we have a local minimum.*

It remains to study the boundary. Using polar coordinates, we have $x = \cos\theta$, $y = \sin\theta$, *so that*

$$
g(\theta) := f(\cos\theta, \sin\theta) = (\cos\theta - \sin\theta)\, e^{-1}, \quad \theta \in [0, 2\pi].
$$

We have

$$
g'(\theta) = (-\sin\theta - \cos\theta)\, e^{-1} \geq 0
$$

for $\frac{3\pi}{4} \leq \theta \leq \frac{7\pi}{4}$. *Thus the point* $\left(\cos\frac{3\pi}{4}, \sin\frac{3\pi}{4}\right) = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ *could be a point of local minimum and the point* $\left(\cos\frac{7\pi}{4}, \sin\frac{7\pi}{4}\right) = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ *could be a point of local maximum. Are they?*

Monday, March 4, 2013

Example 118 ***First method:*** *To see if they are consider the restriction* $y = -x$. *Let*

$$
h(x) = f(x, -x) = 2xe^{-2x^2}, \quad x \in \left[-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right]
$$

Then

$$
h'(x) = 2e^{-2x^2} + 2xe^{-2x^2}(-4x) = 2e^{-2x^2}\left(1 - 4x^2\right) > 0
$$

for $-\frac{1}{2} < x < \frac{1}{2}$. *Hence,* h *decreases for* $x < -\frac{1}{2}$, *increases for* $-\frac{1}{2} < x < \frac{1}{2}$, *and decreases for* $x > \frac{1}{2}$. *Since* $h'\left(-\frac{\sqrt{2}}{2}\right) < 0$, *we have that* $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ *is not a point of local minimum, while, since* $h'\left(\frac{\sqrt{2}}{2}\right) < 0$, *we have that* $\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ *is not a point of local maximum.*

Second method: *Another way to look at this was to study the partial and directional derivatives. We have*

$$\frac{\partial f}{\partial x}\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = e^{-1}\left(-1-1+1\right) = -e^{-1},$$
$$\frac{\partial f}{\partial y}\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = -e^{-1}\left(-1-1+1\right) = e^{-1}.$$

Since $\frac{\partial f}{\partial x}\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) < 0$, *for points* $\left(x, \frac{\sqrt{2}}{2}\right)$ *with* $x > -\frac{\sqrt{2}}{2}$ *close to* $-\frac{\sqrt{2}}{2}$ *we have that* $f\left(x, \frac{\sqrt{2}}{2}, y\right) < f\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$. *To see this, note that*

$$-e^{-1} = \frac{\partial f}{\partial x}\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = \lim_{t\to 0}\frac{f\left(-\frac{\sqrt{2}}{2}+t, \frac{\sqrt{2}}{2}\right) - f\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)}{t},$$

and so taking $0 < \varepsilon < e^{-1}$, *we can find* $\delta > 0$ *such that*

$$-e^{-1} - \varepsilon < \frac{f\left(-\frac{\sqrt{2}}{2}+t, \frac{\sqrt{2}}{2}\right) - f\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)}{t} < -e^{-1} + \varepsilon < 0$$

for all $0 < |t-0| < \delta$. *In particular, if* $0 < t < \delta$, *then* $f\left(-\frac{\sqrt{2}}{2}+t, \frac{\sqrt{2}}{2}\right) < f\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$, *which shows that* $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$ *is not a point of local minimum.*

Similarly,

$$\frac{\partial f}{\partial x}\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = e^{-1}\left(-1-1+1\right) = -e^{-1} < 0,$$

and so reasoning as before we can find $\delta_1 > 0$ *such that*

$$-e^{-1} - \varepsilon < \frac{f\left(\frac{\sqrt{2}}{2}+t, -\frac{\sqrt{2}}{2}\right) - f\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)}{t} < -e^{-1} + \varepsilon < 0$$

for all $0 < |t-0| < \delta_1$. *In particular, if* $-\delta_1 < t < 0$, *then* $f\left(\frac{\sqrt{2}}{2}+t, -\frac{\sqrt{2}}{2}\right) > f\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$, *which shows that* $\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ *is not a point of local maximum.*

Example 119 *Let* $f(x,y,z) := x^2 + y^4 + y^2 + z^3 - 2xz$. *We have*

$$\left\{\begin{array}{l}\frac{\partial f}{\partial x} = 2x - 2z = 0\\ \frac{\partial f}{\partial y} = 4y^3 + 2y = 0\\ \frac{\partial f}{\partial z} = 3z^2 - 2x = 0\end{array}\right. \Longleftrightarrow \left\{\begin{array}{c}x - z = 0\\ y\left(2y^2+1\right) = 0\\ 3z^2 - 2x = 0\end{array}\right. \Longleftrightarrow \left\{\begin{array}{c}x - z = 0\\ y = 0\\ 3z^2 - 2z = 0\end{array}\right.$$

$$\Longleftrightarrow \left\{\begin{array}{c}x - z = 0\\ y = 0\\ z(3z-2) = 0\end{array}\right.$$

and so the critical points are $(0,0,0)$ *and* $\left(\frac{2}{3}, 0, \frac{2}{3}\right)$. *Note that* $(0,0,0)$ *is not a point of local minimum or maximum, since* $f(0,0,z) = z^3$ *which changes sign near* 0. *Let's study the point* $\left(\frac{2}{3}, 0, \frac{2}{3}\right)$. *We have*

$$H_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial z \partial x}\\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial z \partial y}\\ \frac{\partial^2 f}{\partial x \partial z} & \frac{\partial^2 f}{\partial y \partial z} & \frac{\partial^2 f}{\partial z^2}\end{pmatrix} = \begin{pmatrix} 2 & 0 & -2\\ 0 & 12y^2+2 & 0\\ -2 & 0 & 6z\end{pmatrix}$$

and so

$$H_f\left(\frac{2}{3}, 0, \frac{2}{3}\right) = \begin{pmatrix} 2 & 0 & -2\\ 0 & 2 & 0\\ -2 & 0 & 6\end{pmatrix}.$$

We have

$$\begin{aligned} 0 &= \det\left(t\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} - \begin{pmatrix}2 & 0 & -2\\ 0 & 2 & 0\\ -2 & 0 & 4\end{pmatrix}\right)\\ &= \det\begin{pmatrix} t-2 & 0 & 2\\ 0 & t-2 & 0\\ 2 & 0 & t-4\end{pmatrix} = t^3 - 8t^2 + 16t - 8.\end{aligned}$$

The eigenvalues are all positive by Theorem 113. Hence, at $\left(\frac{2}{3}, 0, \frac{2}{3}\right)$ *we have a local minimum.*

Theorem 120 (Weierstrass) *Let* $K \subset \mathbb{R}^N$ *be closed and bounded and let* $f : K \to \mathbb{R}$ *be a continuous function. Then there exists* $\mathbf{x}_0, \mathbf{x}_1 \in K$ *such that*

$$f(\mathbf{x}_0) = \min_{\mathbf{x}\in K} f(\mathbf{x}), \quad f(\mathbf{x}_1) = \max_{\mathbf{x}\in K} f(\mathbf{x}).$$

Definition 121 *Given a set* $E \subseteq \mathbb{R}^N$ *and a function* $\mathbf{f} : E \to \mathbb{R}^M$, *the* Jacobian matrix *of* $\mathbf{f} = (f_1, \ldots, f_M)$ *at some point* $\mathbf{x}_0 \in E$, *whenever it exists, is the* $M \times N$ *matrix*

$$J_{\mathbf{f}}(\mathbf{x}_0) := \begin{pmatrix} \nabla f_1(\mathbf{x}_0)\\ \vdots\\ \nabla f_M(\mathbf{x}_0)\end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_1}{\partial x_N}(\mathbf{x}_0)\\ \vdots & & \vdots\\ \frac{\partial f_M}{\partial x_1}(\mathbf{x}_0) & \cdots & \frac{\partial f_M}{\partial x_N}(\mathbf{x}_0)\end{pmatrix}.$$

It is also denoted

$$\frac{\partial\left(f_{1},\ldots,f_{M}\right)}{\partial\left(x_{1},\ldots,x_{N}\right)}\left(\mathbf{x}_{0}\right).$$

When $M=N$*,* $J_{\mathbf{f}}\left(\mathbf{x}_{0}\right)$ *is an* $N\times N$ *square matrix and its determinant is called the* Jacobian determinant *of* $\mathbf{f}$ *at* $\mathbf{x}_{0}$*. Thus,*

$$\det J_{\mathbf{f}}\left(\mathbf{x}_{0}\right)=\det\left(\frac{\partial f_{j}}{\partial x_{i}}\left(\mathbf{x}_{0}\right)\right)_{i,j=1,\ldots,N}.$$

10 Lagrange Multipliers

In Section 9 (see Theorem 114) we have seen how to find points of local minima and maxima of a function $f:E\rightarrow\mathbb{R}$ in the interior E° of E. Now we are ready to find points of local minima and maxima of a function $f:E\rightarrow\mathbb{R}$ on the boundary ∂E of E. We assume that the boundary of E has a special form, that is, it is given by a set of the form

$$\left\{\mathbf{x}\in\mathbb{R}^{N}:\,\mathbf{g}\left(\mathbf{x}\right)=\mathbf{0}\right\}.$$

Definition 122 *Let* $f:E\rightarrow\mathbb{R}$*, where* $E\subseteq\mathbb{R}^{N}$*, let* $F\subseteq E$ *and let* $\mathbf{x}_{0}\in F$*. We say that*

(i) f *attains a* constrained local minimum *at* $\mathbf{x}_{0}$ *if there exists* $r>0$ *such that* $f\left(\mathbf{x}\right)\geq f\left(\mathbf{x}_{0}\right)$ *for all* $\mathbf{x}\in F\cap B\left(\mathbf{x}_{0},r\right)$*,*

(ii) f *attains a* constrained local maximum *at* $\mathbf{x}_{0}$ *if there exists* $r>0$ *such that* $f\left(\mathbf{x}\right)\leq f\left(\mathbf{x}_{0}\right)$ *for all* $\mathbf{x}\in F\cap B\left(\mathbf{x}_{0},r\right)$*.*

The set F is called the *constraint.*

Wednesday, March 6, 2013

Theorem 123 (Lagrange Multipliers) *Let* $U\subseteq\mathbb{R}^{N}$ *be an open set, let* $f:U\rightarrow\mathbb{R}$ *be a function of class* C^{1} *and let* $\mathbf{g}:U\rightarrow\mathbb{R}^{M}$ *be a class of function* C^{1}*, where* $M<N$*, and let*

$$F:=\left\{\mathbf{x}\in U:\,\mathbf{g}\left(\mathbf{x}\right)=\mathbf{0}\right\}.$$

Let $\mathbf{x}_{0}\in F$ *and assume that* f *attains a constrained local minimum (or maximum) at* $\mathbf{x}_{0}$*. If* $J_{\mathbf{g}}\left(\mathbf{x}_{0}\right)$ *has maximum rank* M*, then there exist* $\lambda_{1},\ldots,\lambda_{M}\in\mathbb{R}$ *such that*

$$\nabla f\left(\mathbf{x}_{0}\right)=\lambda_{1}\nabla g_{1}\left(\mathbf{x}_{0}\right)+\cdots+\lambda_{M}\nabla g_{M}\left(\mathbf{x}_{0}\right).$$

Example 124 *We want to find points of local minima and maxima of the function* $f\left(x,y,z\right):=x-y+2z$ *over the set*

$$E=\left\{\left(x,y,z\right)\in\mathbb{R}^{3}:\,x^{2}+y^{2}+2z^{2}\leq2\right\}.$$

Let's find critical points of f in the interior. Since $\frac{\partial f}{\partial x}(x,y,z) = 1 \neq 0$, there are no critical points in the interior. Thus, points of local minima and maxima, if they exist, must be found on the boundary of E, that is,

$$\partial E = \left\{(x,y,z) \in \mathbb{R}^3 : \, x^2 + y^2 + 2z^2 = 2\right\}.$$

In other words, we are looking at a constrained problem. The constraint is given by $x^2 + y^2 + 2z^2 = 2$. Define $g(x,y,z) := x^2 + y^2 + 2z^2 - 2$. In this case the Jacobian matrix of g coincides with the gradient, that is,

$$J_g(x,y,z) = \nabla g(x,y,z) = (2x, 2y, 4z).$$

In this case the maximum rank is one, so it is enough to check that $(2x, 2y, 4z) \neq (0,0,0)$ at points in the constraint ∂E. But $(2x, 2y, 4z) = (0,0,0)$ only at the origin and this point does not belong to the constraint.

By the theorem on Lagrange multipliers, we need to find (x,y,z) and λ such that

$$\begin{cases} 0 = \frac{\partial f}{\partial x}(x,y,z) - \lambda \frac{\partial g}{\partial x}(x,y,z) = \frac{\partial}{\partial x}\left(x - y + 2z - \lambda\left(x^2 + y^2 + 2z^2 - 2\right)\right) = 1 - 2x\lambda, \\ 0 = \frac{\partial f}{\partial y}(x,y,z) - \lambda \frac{\partial g}{\partial y}(x,y,z) = \frac{\partial}{\partial y}\left(x - y + 2z - \lambda\left(x^2 + y^2 + 2z^2 - 2\right)\right) = -1 - 2y\lambda, \\ 0 = \frac{\partial f}{\partial z}(x,y,z) - \lambda \frac{\partial g}{\partial z}(x,y,z) = \frac{\partial}{\partial z}\left(x - y + 2z - \lambda\left(x^2 + y^2 + 2z^2 - 2\right)\right) = 2 - 4z\lambda, \\ 0 = g(x,y,z) = x^2 + y^2 + 2z^2 - 2. \end{cases}$$

If $\lambda = 0$, we have no solution, while if $\lambda \neq 0$, we get

$$\begin{cases} x = \frac{1}{2\lambda}, \\ y = -\frac{1}{2\lambda}, \\ z = \frac{1}{2\lambda}, \\ x^2 + y^2 + 2z^2 - 2 = 0. \end{cases}$$

Substituting gives $\left(\frac{1}{2\lambda}\right)^2 + \left(-\frac{1}{2\lambda}\right)^2 + 2\left(\frac{1}{2\lambda}\right)^2 - 2 = 0$, that is, $2\lambda^2 - 1 = 0$, and so $\lambda = \pm\frac{1}{\sqrt{2}}$. Thus, the only two possible points are $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ and $\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$. Now, since E is closed and bounded (why?), it is compact. Since f is continuous, it follows by the Weierstrass theorem that there exist the minimum and the maximum of f over E. Hence, $f\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = -\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2} - 2\frac{\sqrt{2}}{2}$ is the minimum and $f\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right) = \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} + 2\frac{\sqrt{2}}{2}$ is the maximum.

Example 125 *Given the set*

$$E = \left\{(x,y,z) \in \mathbb{R}^3 : \, x^2 - xy + y^2 - z^2 = 1, \, x^2 + y^2 = 1\right\},$$

we want to find the points of E having minimal distance from the origin. So we need to find

$$\min\left\{\sqrt{(x-0)^2 + (y-0)^2 + (z-0)^2} : \, (x,y,z) \in E\right\}.$$

To avoid dealing with square roots, we consider the function $f(x,y,z) = x^2 + y^2 + z^2$. *We want to find the minimum of* f *over the st* E.

First method. *Let* $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^3$ *be the function*

$$\mathbf{g}(x,y,z) = \left(x^2 - xy + y^2 - z^2 - 1, x^2 + y^2 - 1\right).$$

By the theorem on Lagrange multipliers we need to find (x,y,z) *and* λ_1, λ_2 *such that*

$$\begin{cases} 0 = \frac{\partial f}{\partial x} - \lambda_1 \frac{\partial g_1}{\partial x} - \lambda_2 \frac{\partial g_1}{\partial x} = \frac{\partial}{\partial x}\left(x^2 + y^2 + z^2 - \lambda_1\left(x^2 - xy + y^2 - z^2 - 1\right) - \lambda_2\left(x^2 + y^2 - 1\right)\right), \\ 0 = \frac{\partial f}{\partial y} - \lambda_1 \frac{\partial g_1}{\partial y} - \lambda_2 \frac{\partial g_1}{\partial y} = \frac{\partial}{\partial y}\left(x^2 + y^2 + z^2 - \lambda_1\left(x^2 - xy + y^2 - z^2 - 1\right) - \lambda_2\left(x^2 + y^2 - 1\right)\right), \\ 0 = \frac{\partial f}{\partial z} - \lambda_1 \frac{\partial g_1}{\partial z} - \lambda_2 \frac{\partial g_1}{\partial z} = \frac{\partial}{\partial z}\left(x^2 + y^2 + z^2 - \lambda_1\left(x^2 - xy + y^2 - z^2 - 1\right) - \lambda_2\left(x^2 + y^2 - 1\right)\right), \\ 0 = g_1(x,y,z) = x^2 - xy + y^2 - z^2 - 1 = 0, \\ 0 = g_2(x,y,z) = x^2 + y^2 - 1 = 0. \end{cases}$$

We have

$$\begin{cases} 0 = 2x - \lambda_1(2x - y) - \lambda_2 2x, \\ 0 = 2y - \lambda_1(-x + 2y) - \lambda_2 2y, \\ 0 = 2z - \lambda_1(-2z), \\ 0 = x^2 - xy + y^2 - z^2 - 1, \\ 0 = x^2 + y^2 - 1. \end{cases}$$

Adding the first two equations and substituting the last in the fourth one, we get

$$\begin{cases} 0 = (x + y)(2 - \lambda_1 - 2\lambda_2), \\ 0 = 2y - \lambda_1(-x + 2y) - \lambda_2 2y, \\ 0 = z(1 + \lambda_1), \\ 0 = -xy - z^2, \\ 0 = x^2 + y^2 - 1. \end{cases}$$

From the third equation $0 = z(1 + \lambda_1)$ *we have that either* $z = 0$ *or* $\lambda_1 = -1$. *If* $z = 0$ *we get*

$$\begin{cases} 0 = (x + y)(2 - \lambda_1 - 2\lambda_2), \\ 0 = 2y - \lambda_1(-x + 2y) - \lambda_2 2y, \\ 0 = z, \\ 0 = -xy, \\ 0 = x^2 + y^2 - 1. \end{cases}$$

Hence, either $x = 0$ *or* $y = 0$. *If* $x = 0$, *then from the last last equation* $y = \pm 1$, *so that*

$$\begin{cases} 0 = 2 - \lambda_1 - 2\lambda_2, \\ 0 = 1 - \lambda_1 - \lambda_2, \\ 0 = z, \\ 0 = x, \\ \pm 1 = y. \end{cases} \iff \begin{cases} 0 = 1 - \lambda_2, \\ 0 = 1 - \lambda_1 - \lambda_2, \\ 0 = z, \\ 0 = x, \\ \pm 1 = y. \end{cases} \iff \begin{cases} 1 = \lambda_2, \\ 0 = \lambda_1, \\ 0 = z, \\ 0 = x, \\ \pm 1 = y. \end{cases}$$

If $y = 0$ then from the last last equation $x = \pm 1$, so that

$$\begin{cases} 0 = (2 - \lambda_1 - 2\lambda_2), \\ 0 = \lambda_1, \\ 0 = z, \\ 0 = y, \\ \pm 1 = x. \end{cases} \iff \begin{cases} -1 = \lambda_2, \\ 0 = \lambda_1, \\ 0 = z, \\ 0 = y, \\ \pm 1 = x. \end{cases}$$

If $\lambda_1 = -1$, then

$$\begin{cases} 0 = (x + y)(3 - 2\lambda_2), \\ 0 = 4y - x - \lambda_2 2y, \\ -1 = \lambda_1, \\ 0 = -xy - z^2, \\ 0 = x^2 + y^2 - 1. \end{cases}$$

In the first equation either $x + y = 0$ or $3 - 2\lambda_2 = 0$. In the first case, we get

$$\begin{cases} x = -y, \\ 0 = x(-5 + \lambda_2 2), \\ -1 = \lambda_1, \\ 0 = x^2 - z^2, \\ 1 = 2x^2, \end{cases} \iff \begin{cases} x = -y, \\ \frac{5}{2} = \lambda_2, \\ -1 = \lambda_1, \\ z = \pm x, \\ \pm \frac{1}{\sqrt{2}} = x, \end{cases}$$

while in the second, $\frac{3}{2} = \lambda_2$, and so

$$\begin{cases} \frac{3}{2} = \lambda_2, \\ 0 = 4y - x - 3y, \\ -1 = \lambda_1, \\ 0 = -xy - z^2, \\ 0 = x^2 + y^2 - 1. \end{cases} \iff \begin{cases} \frac{3}{2} = \lambda_2, \\ x = y, \\ -1 = \lambda_1, \\ 0 = -x^2 - z^2, \\ 0 = 2x^2 - 1. \end{cases}$$

The fourth and fifth equation are incompatible, so there are no solutions in this case.

In conclusion, we have found the points $(0, \pm 1, 0)$ with $\lambda_1 = 0$ and $\lambda_2 = 1$, $(\pm 1, 0, 0)$ with $\lambda_1 = 0$ and $\lambda_2 = -1$, $\left(\pm \frac{1}{\sqrt{2}}, \mp \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right)$ and $\left(\pm \frac{1}{\sqrt{2}}, \mp \frac{1}{\sqrt{2}}, \mp \frac{1}{\sqrt{2}}\right)$ with $\lambda_1 = -1$ and $\lambda_2 = \frac{5}{2}$.

Let's prove that E is bounded. We have that $x^2 + y^2 = 1$, so $|x| \leq 1$ and $|y| \leq 1$, while

$$z^2 = -1 - x^2 + xy - y^2 \leq xy \leq 1.$$

Hence, E is bounded and closed. It follows by the Weierstrass theorem that f has a mimimum over E. Since

$$f\left(\pm \frac{1}{\sqrt{2}}, \mp \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right) = f\left(\pm \frac{1}{\sqrt{2}}, \mp \frac{1}{\sqrt{2}}, \mp \frac{1}{\sqrt{2}}\right) = \frac{3}{2} > 1 = f(\pm 1, 0, 0) = f(0, \pm 1, 0),$$

it follows that $(\pm 1, 0, 0)$ and $(0, \pm 1, 0)$ are the two points in E of minimal distance from the origin.

Second method. *Exercise. Find a simpler way to solve this problem.*

Friday, March 8, 2013

Midsemester break, no classes

Monday, March 11, 2013

Spring break, no classes

Wednesday, March 13, 2013

Spring break, no classes

Friday, March 15, 2013

Spring break, no classes

Monday, March 18, 2013

11 Chain Rule

Definition 126 *Let $E \subseteq \mathbb{R}^N$, let $\mathbf{f} : E \to \mathbb{R}^M$, and let $\mathbf{x}_0 \in E$ be an accumulation point of E. The function $\mathbf{f}$ is* differentiable at $\mathbf{x}_0$ *if there exists a linear function $\mathbf{T} : \mathbb{R}^N \to \mathbb{R}^M$ (depending on $\mathbf{f}$ and $\mathbf{x}_0$) such that*

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) - \mathbf{T}(\mathbf{x}-\mathbf{x}_0)}{\|\mathbf{x}-\mathbf{x}_0\|_N} = \mathbf{0}. \tag{12}$$

provided the limit exists. The function T, if it exists, is called the differential *of $\mathbf{f}$ at $\mathbf{x}_0$ and is denoted $d\mathbf{f}(\mathbf{x}_0)$ or $d\mathbf{f}_{\mathbf{x}_0}$.*

Theorem 127 *Let $E \subseteq \mathbb{R}^N$, let $\mathbf{f} : E \to \mathbb{R}^M$, and let $\mathbf{x}_0 \in E$ be an accumulation point of E. Then $\mathbf{f} = (f_1, \ldots, f_M)$ is differentiable at $\mathbf{x}_0$ if and only if all its components f_j, $j = 1, \ldots, M$, are differentiable at $\mathbf{x}_0$. Moreover, $d\mathbf{f}_{\mathbf{x}_0} = \left(d(f_1)_{\mathbf{x}_0}, \ldots, d(f_M)_{\mathbf{x}_0}\right)$.*

Proof. Exercise. ■

We study the differentiability of composite functions.

Theorem 128 (Chain Rule) *Let $F \subseteq \mathbb{R}^M$, $E \subseteq \mathbb{R}^N$, let $\mathbf{g} : F \to E$, $\mathbf{g} = (\mathbf{g}_1, \ldots, \mathbf{g}_N)$, and let $f : E \to \mathbb{R}$. Assume that at some point $\mathbf{y}_0 \in F^\circ$ there exist the directional derivatives $\frac{\partial g_1}{\partial \mathbf{v}}(\mathbf{y}_0)$, ..., $\frac{\partial g_N}{\partial \mathbf{v}}(\mathbf{y}_0)$ for some direction $\mathbf{v}$ and f is differentiable at the point $\mathbf{g}(\mathbf{y}_0)$. Then the composite function $f \circ \mathbf{g}$ admits a directional derivative at $\mathbf{y}_0$ in the direction $\mathbf{v}$ and*

$$\begin{aligned}\frac{\partial (f\circ\mathbf{g})}{\partial \mathbf{v}}(\mathbf{y}_0) &= \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(\mathbf{g}(\mathbf{y}_0)) \frac{\partial g_i}{\partial \mathbf{v}}(\mathbf{y}_0) \\ &= \nabla f(\mathbf{g}(\mathbf{y}_0)) \cdot \frac{\partial \mathbf{g}}{\partial \mathbf{v}}(\mathbf{y}_0).\end{aligned}$$

Moreover, if $\mathbf{g}$ is differentiable at $\mathbf{y}_0$ and f is differentiable at $\mathbf{g}(\mathbf{y}_0)$, then $f \circ \mathbf{g}$ is differentiable at $\mathbf{y}_0$.

Example 129 *Consider the function*

$$g(\mathbf{x}) := f(\|\mathbf{x}\|) = f\left(\sqrt{x_1^2 + x_2^2 + \cdots + x_N^2}\right),$$

where $f : [0, \infty) \to \mathbb{R}$ is differentiable. Then for all $\mathbf{x} \neq \mathbf{0}$,

$$\frac{\partial g}{\partial x_i}(\mathbf{x}) = f'\left(\sqrt{x_1^2 + x_2^2 + \cdots + x_N^2}\right) \frac{2x_i}{2\left(\sqrt{x_1^2 + x_2^2 + \cdots + x_N^2}\right)}.$$

Example 130 *Consider the function*

$$h(x, y) = \int_{f(x,y)}^{g(x,y)} e^{-t^2}\, dt,$$

where $f : \mathbb{R}^2 \to \mathbb{R}$ and $g : \mathbb{R}^2 \to \mathbb{R}$ are differentiable. To apply the chain rule, consider the function

$$k(z, w) = \int_w^z e^{-t^2}\, dt,$$

and observe that

$$h(x, y) = k(g(x, y), f(x, y)).$$

Hence, by the chain rule,

$$\frac{\partial h}{\partial x}(x, y) = \frac{\partial k}{\partial z}(g(x, y), f(x, y)) \frac{\partial g}{\partial x}(x, y) + \frac{\partial k}{\partial w}(g(x, y), f(x, y)) \frac{\partial f}{\partial x}(x, y),$$
$$\frac{\partial h}{\partial y}(x, y) = \frac{\partial k}{\partial z}(g(x, y), f(x, y)) \frac{\partial g}{\partial y}(x, y) + \frac{\partial k}{\partial w}(g(x, y), f(x, y)) \frac{\partial f}{\partial y}(x, y).$$

Now

$$\frac{\partial k}{\partial z}(z, w) = e^{-z^2},$$

while

$$\frac{\partial k}{\partial w}(z, w) = -e^{-z^2},$$

and so

$$\frac{\partial h}{\partial x}(x, y) = e^{-(g(x,y))^2} \frac{\partial g}{\partial x}(x, y) - e^{-(f(x,y))^2} \frac{\partial f}{\partial x}(x, y),$$
$$\frac{\partial h}{\partial y}(x, y) = e^{-(g(x,y))^2} \frac{\partial g}{\partial y}(x, y) - e^{-(f(x,y))^2} \frac{\partial f}{\partial y}(x, y).$$

Example 131 *Consider the functions $f : \mathbb{R}^2 \to \mathbb{R}$ and $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^2$ defined by*

$$f(x_1, x_2) = x_1 x_2 - 1, \quad \mathbf{g}(y_1, y_2) = (y_1 y_2 - e^{y_1}, y_1 \sin(y_1 y_2)).$$

Then

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 1x_2 - 0, \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = x_1 1 - 0,$$
$$\frac{\partial g_1}{\partial y_1}(y_1, y_2) = 1y_2 - e^{y_1}, \quad \frac{\partial g_1}{\partial y_2}(y_1, y_2) = y_1 1 - 0,$$
$$\frac{\partial g_2}{\partial y_1}(y_1, y_2) = 1 \sin(y_1 y_2) + y_1 \cos(y_1 y_2)(1y_2), \quad \frac{\partial g_2}{\partial y_2}(y_1, y_2) = y_1 \cos(y_1 y_2)(y_1 1).$$

First method: *Consider the composition*

$$(f \circ \mathbf{g})(y_1, y_2) = f(g_1(y_1, y_2), g_2(y_1, y_2)).$$

Then by the chain rule,

$$\begin{aligned}\frac{\partial(f \circ \mathbf{g})}{\partial y_1}(y_1, y_2) &= \frac{\partial f}{\partial x_1}(g_1(y_1, y_2), g_2(y_1, y_2))\frac{\partial g_1}{\partial y_1}(y_1, y_2) + \frac{\partial f}{\partial x_2}(g_1(y_1, y_2), g_2(y_1, y_2))\frac{\partial g_2}{\partial y_1}(y_1, y_2) \\ &= y_1 \sin(y_1 y_2)(1y_2 - e^{y_1}) + (y_1 y_2 - e^{y_1})(1\sin(y_1 y_2) + y_1 \cos(y_1 y_2)(1y_2)),\end{aligned}$$

while

$$\begin{aligned}\frac{\partial(f \circ \mathbf{g})}{\partial y_2}(y_1, y_2) &= \frac{\partial f}{\partial x_1}(g_1(y_1, y_2), g_2(y_1, y_2))\frac{\partial g_1}{\partial y_2}(y_1, y_2) + \frac{\partial f}{\partial x_2}(g_1(y_1, y_2), g_2(y_1, y_2))\frac{\partial g_2}{\partial y_2}(y_1, y_2) \\ &= y_1 \sin(y_1 y_2)(y_1 1 - 0) + (y_1 y_2 - e^{y_1})(y_1 \cos(y_1 y_2)(y_1 1)).\end{aligned}$$

Second method: *Write the explicit formula for* $(f \circ \mathbf{g})$*, that is,*

$$\begin{aligned}(f \circ \mathbf{g})(y_1, y_2) &= f(g_1(y_1, y_2), g_2(y_1, y_2)) \\ &= (y_1 y_2 - e^{y_1})(y_1 \sin(y_1 y_2)) - 1.\end{aligned}$$

Then

$$\begin{aligned}\frac{\partial(f \circ \mathbf{g})}{\partial y_1}(y_1, y_2) &= \frac{\partial}{\partial y_1}[(y_1 y_2 - e^{y_1})(y_1 \sin(y_1 y_2)) - 1] \\ &= (1y_2 - e^{y_1})(y_1 \sin(y_1 y_2)) + (y_1 y_2 - e^{y_1})(1\sin(y_1 y_2) + y_1 \cos(y_1 y_2)(1y_2)) - 0\end{aligned}$$

and

$$\begin{aligned}\frac{\partial(f \circ \mathbf{g})}{\partial y_2}(y_1, y_2) &= \frac{\partial}{\partial y_2}[(y_1 y_2 - e^{y_1})(y_1 \sin(y_1 y_2)) - 1] \\ &= (y_1 1 - 0)(y_1 \sin(y_1 y_2)) + (y_1 y_2 - e^{y_1})(y_1 \cos(y_1 y_2)(y_1 1)) - 0.\end{aligned}$$

As a corollary of Theorem 128, we have the following result.

Corollary 132 *Let* $F \subseteq \mathbb{R}^M$*,* $E \subseteq \mathbb{R}^N$*, let* $\mathbf{g} : F \to E$*,* $\mathbf{g} = (\mathbf{g}_1, \ldots, \mathbf{g}_N)$*, and let* $\mathbf{f} : E \to \mathbb{R}^P$*. Assume that* $\mathbf{g}$ *is differentiable at some point* $\mathbf{y}_0 \in F^\circ$ *and that* $\mathbf{f}$ *is differentiable at the point* $\mathbf{g}(\mathbf{y}_0)$ *and that* $\mathbf{g}(\mathbf{y}_0) \in E^\circ$*. Then the composite function* $\mathbf{f} \circ \mathbf{g}$ *is differentiable at* $\mathbf{y}_0$ *and*

$$J_{\mathbf{f} \circ \mathbf{g}}(\mathbf{y}_0) = J_{\mathbf{f}}(\mathbf{g}(\mathbf{y}_0)) J_{\mathbf{g}}(\mathbf{x}_0).$$

Wednesday, March 20, 2013

12 Implicit and Inverse Function

Given a function f of two variables $(x, y) \in \mathbb{R}^2$, consider the equation

$$f(x, y) = 0.$$

We want to solve for y, that is, we are interested in finding a function $y = g(x)$ such that

$$f(x, g(x)) = 0.$$

We will see under which conditions we can do this. The result is going to be local.

In what follows given $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y} \in \mathbb{R}^M$ and $\mathbf{f}(\mathbf{x}, \mathbf{y})$, we write

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{y}) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial f_1}{\partial x_N}(\mathbf{x}, \mathbf{y}) \\ \vdots & & \vdots \\ \frac{\partial f_M}{\partial x_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial f_M}{\partial x_N}(\mathbf{x}, \mathbf{y}) \end{pmatrix}$$

and

$$\frac{\partial \mathbf{f}}{\partial \mathbf{y}}(\mathbf{x}, \mathbf{y}) := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial f_1}{\partial y_M}(\mathbf{x}, \mathbf{y}) \\ \vdots & & \vdots \\ \frac{\partial f_M}{\partial y_1}(\mathbf{x}, \mathbf{y}) & \cdots & \frac{\partial f_M}{\partial y_M}(\mathbf{x}, \mathbf{y}) \end{pmatrix}.$$

Theorem 133 (Implicit Function) *Let $U \subseteq \mathbb{R}^N \times \mathbb{R}^M$ be open, let $\mathbf{f} : U \to \mathbb{R}^M$, and let $(\mathbf{a}, \mathbf{b}) \in U$. Assume that $\mathbf{f} \in C^m(U)$ for some $m \in \mathbb{N}$, that*

$$\mathbf{f}(\mathbf{a}, \mathbf{b}) = \mathbf{0} \quad \textit{and} \quad \det \frac{\partial \mathbf{f}}{\partial \mathbf{y}}(\mathbf{a}, \mathbf{b}) \neq 0.$$

Then there exist $B_N(\mathbf{a}, r_0) \subset \mathbb{R}^N$ and $B_M(\mathbf{b}, r_1) \subset \mathbb{R}^M$, with $B_N(\mathbf{a}, r_0) \times B_M(\mathbf{b}, r_1) \subseteq U$, and a unique function

$$\mathbf{g} : B_N(\mathbf{a}, r_0) \to B_M(\mathbf{b}, r_1)$$

of class C^m such that $\mathbf{f}(\mathbf{x}, \mathbf{g}(\mathbf{x})) = \mathbf{0}$ for all $\boldsymbol{x} \in B_N(\mathbf{a}, r_0)$ and $\mathbf{g}(\mathbf{a}) = \mathbf{b}$.

The next examples show that when $\det \frac{\partial \mathbf{f}}{\partial \mathbf{y}}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$, then anything can happen.

Example 134 *In all these examples $N = M = 1$ and $\frac{\partial f}{\partial y}(x_0, y_0) = 0$.*

(i) Consider the function

$$f(x, y) := (y - x)^2.$$

Then $f(0,0) = 0$, $\frac{\partial f}{\partial y}(0,0) = 0$ and $g(x) = x$ satisfies $f(x, g(x)) = 0$.

(ii) Consider the function

$$f(x, y) := x^2 + y^2.$$

Then $f(0,0) = 0$, $\frac{\partial f}{\partial y}(0,0) = 0$ but there is no function g defined near $x = 0$ such that $f(x, g(x)) = 0$.

(iii) Consider the function

$$f(x,y) := (xy-1)\left(x^2+y^2\right).$$

Then $f(0,0)=0$, $\frac{\partial f}{\partial y}(0,0)=0$ *but*

$$g(x) = \begin{cases} 0 & \text{if } x=0, \\ \frac{1}{x} & \text{if } x \neq 0, \end{cases}$$

which is discontinuous.

Next we give some examples on how to apply the implicit function theorem.

Example 135 *Consider the function*

$$f(x,y) := xe^y - y.$$

Then $f(0,0)=0$, $\frac{\partial f}{\partial y}(x,y) = xe^y - 1$ *so that* $\frac{\partial f}{\partial y}(0,0) = -1$. *By the implicit function theorem there exist* $r>0$ *and a function* $g \in C^\infty((-r,r))$ *such that* $g(0)=0$ *and* $f(x,g(x))=0$. *To find the behavior of* g *near* $x=0$, *we can use Taylor's formula. Let's find* $g'(0)$ *and* $g''(0)$. *We have*

$$xe^{g(x)} - g(x) = 0.$$

Hence, differentiating twice

$$1e^{g(x)} + xg'(x)e^{g(x)} - g'(x) = 0,$$
$$e^{g(x)}g'(x) + g'(x)e^{g(x)} + xg''(x)e^{g(x)} + x\left(g'(x)\right)^2 e^{g(x)} - g''(x) = 0.$$

Substituting $x=0$ *and using* $g(0)=0$ *we get*

$$1e^0 + 0 - g'(0) = 0,$$
$$e^0 g'(0) + g'(0)e^0 + 0 + 0 - g''(0) = 0.$$

So that $g'(0)=1$ *and* $g''(0)=2$. *Hence,*

$$\begin{aligned} g(x) &= g(0) + g'(0)(x-0) + \frac{1}{2}g''(0)(x-0)^2 + o\left((x-0)^2\right) \\ &= 0 + 1(x-0) + \frac{1}{2}2(x-0)^2 + o\left((x-0)^2\right). \end{aligned}$$

Friday, March 22, 2013

Correction homework.

Monday, March 25, 2013

Example 136 *Consider the function*

$$\mathbf{f}(x,y,z) = \left(y\cos(xz) - x^2 + 1, y\sin(xz) - x\right).$$

Let's prove that there exist $r > 0$ and $\mathbf{g} : (1-r, 1+r) \to \mathbb{R}^2$ of class C^∞ such that $\mathbf{g}(1) = \left(1, \frac{\pi}{2}\right)$ and $\mathbf{f}(x, \mathbf{g}(x)) = 0$. Note that $\mathbf{f}$ is of class C^∞. Here the point is $\left(1, 1, \frac{\pi}{2}\right)$ and

$$\mathbf{f}\left(1, 1, \frac{\pi}{2}\right) = \left(1\cos\left(1\frac{\pi}{2}\right) - 1 + 1, 1\sin\left(1\frac{\pi}{2}\right) - 1\right) = (0, 0).$$

Moreover,

$$\begin{aligned}\frac{\partial \mathbf{f}}{\partial (y,z)}(x,y,z) &= \begin{pmatrix} \frac{\partial f_1}{\partial y}(x,y,z) & \frac{\partial f_1}{\partial z}(x,y,z) \\ \frac{\partial f_2}{\partial y}(x,y,z) & \frac{\partial f_1}{\partial z}(x,y,z) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial}{\partial y}\left(y\cos(xz) - x^2 + 1\right) & \frac{\partial}{\partial z}\left(y\cos(xz) - x^2 + 1\right) \\ \frac{\partial}{\partial y}\left(y\sin(xz) - x\right) & \frac{\partial}{\partial z}\left(y\sin(xz) - x\right) \end{pmatrix} \\ &= \begin{pmatrix} 1\cos(xz) - 0 - 0 & -xy\sin(xz) - 0 - 0 \\ 1\sin(xz) - 0 & xy\cos(xz) - 0 \end{pmatrix}\end{aligned}$$

and so

$$\det \frac{\partial \mathbf{f}}{\partial (y,z)}\left(1, 1, \frac{\pi}{2}\right) = \det \begin{pmatrix} \cos\left(1\frac{\pi}{2}\right) & -1\sin\left(1\frac{\pi}{2}\right) \\ \sin\left(1\frac{\pi}{2}\right) & 1\cos\left(1\frac{\pi}{2}\right) \end{pmatrix} = 1 \neq 0.$$

Hence, by the implicit function theorem there exist $r > 0$ and $\mathbf{g} : (1-r, 1+r) \to \mathbb{R}^2$ of class C^∞ such that $\mathbf{g}(1) = \left(1, \frac{\pi}{2}\right)$ and $\mathbf{f}(x, \mathbf{g}(x)) = 0$ for all $x \in (1-r, 1+r)$, that is,

$$\begin{cases} g_1(x)\cos(xg_2(x)) - x^2 + 1 = 0, \\ g_1(x)\sin(xg_2(x)) - x = 0. \end{cases}$$

Reasoning as before, we can use Taylor's formula to find the behavior of g_1 and g_2 near $x = 1$, that is,

$$\begin{aligned} g_1(x) &= g_1(1) + g_1'(1)(x-1) + o((x-1)), \\ g_2(x) &= g_2(1) + g_2'(1)(x-1) + o((x-1)). \end{aligned}$$

Let's differentiate the two equations. We get

$$\begin{cases} g_1'(x)\cos(xg_2(x)) - g_1(x)(1g_2(x) + xg_2'(x))\sin(xg_2(x)) - 2x + 0 = 0, \\ g_1'(x)\sin(xg_2(x)) + g_1(x)(1g_2(x) + xg_2'(x))\cos(xg_2(x)) - 1 = 0. \end{cases}$$

Taking $x = 1$ and using the fact that $g_1(1) = 1$ and $g_2(1) = \frac{\pi}{2}$, we obtain

$$\begin{cases} g_1'(1)\cos\left(1\frac{\pi}{2}\right) - 1\left(\frac{\pi}{2} + 1g_2'(1)\right)\sin\left(1\frac{\pi}{2}\right) - 2 = 0, \\ g_1'(1)\sin\left(1\frac{\pi}{2}\right) + 1\left(\frac{\pi}{2} + 1g_2'(1)\right)\cos\left(1\frac{\pi}{2}\right) - 1 = 0, \end{cases}$$

that is,

$$\begin{cases} 0 - 1\left(\frac{\pi}{2} + g_2'(1)\right)1 - 2 = 0, \\ g_1'(1)\,1 + 0 - 1 = 0, \end{cases}$$

and so $g_1'(1) = 1$ *and* $g_2'(1) = -2 - \frac{\pi}{2}$. *Hence,*

$$g_1(x) = 1 + 1(x-1) + o((x-1)),$$
$$g_2(x) = \frac{\pi}{2} + \left(-2 - \frac{\pi}{2}\right)(x-1) + o((x-1)).$$

Next we give the inverse function theorem.

Theorem 137 (Inverse Function) *Let* $U \subseteq \mathbb{R}^N$ *be open, let* $\mathbf{f} : U \to \mathbb{R}^N$, *and let* $\mathbf{a} \in U$. *Assume that* $\mathbf{f} \in C^m(U)$ *for some* $m \in \mathbb{N}$ *and that*

$$\det J_{\mathbf{f}}(\mathbf{a}) \neq 0.$$

Then there exists $B(\mathbf{a}, r_0) \subseteq U$ *such that* $\mathbf{f}(B(\mathbf{a}, r_0))$ *is open, the function*

$$\mathbf{f} : B(\mathbf{a}, r_0) \to \mathbf{f}(B(\mathbf{a}, r_0))$$

is invertible and $\mathbf{f}^{-1} \in C^m(\mathbf{f}(B(\mathbf{a}, r_0)))$. *Moreover,*

$$J_{\mathbf{f}^{-1}}(\mathbf{y}) = \left(J_{\mathbf{f}}\left(\mathbf{f}^{-1}(\mathbf{y})\right)\right)^{-1}.$$

Exercise 138 *Consider the function* $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ *defined by*

$$f_1(x,y) = \begin{cases} 0 & \text{if } x = 0, \\ x + 2x^2 \sin\frac{1}{x} & \text{if } x \neq 0, \end{cases}$$
$$f_2(x,y) = y.$$

Prove that $\mathbf{f} = (f_1, f_2)$ *is differentiable in* $(0,0)$ *and* $J_{\mathbf{f}}(0,0) = 1$. *Prove that* $\mathbf{f}$ *is not one-to-one in any neighborhood of* $(0,0)$.

The next example shows that the existence of a local inverse at every point does not imply the existence of a global inverse.

Example 139 *Consider the function* $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ *defined by*

$$\mathbf{f}(x,y) = (e^x \cos y, e^x \sin y).$$

The function $\mathbf{f}$ *is not injective. Indeed,* $\mathbf{f}(x,y) = \mathbf{f}(x, y + 2\pi)$ *for all* $(x,y) \in \mathbb{R}^2$. *Hence,* $\mathbf{f}$ *does not admit a global inverse. However,*

$$J_{\mathbf{f}}(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix},$$

and so

$$\det J_{\mathbf{f}}(x,y) = \det \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} = \left(\cos^2 y\right) e^{2x} + e^{2x} \sin^2 y = e^{2x} \neq 0$$

for all $(x,y) \in \mathbb{R}^2$. *Hence, by the inverse function theorem* $\mathbf{f}$ *admits a local inverse in a neighborhood of every point* $(x,y) \in \mathbb{R}^2$.

Wednesday, March 27, 2013

13 Curves

Let $I \subseteq \mathbb{R}$ be an interval and let $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ be a function. As the *parameter* t traverses I, $\boldsymbol{\varphi}(t)$ traverses a curve in $\mathbb{R}^N$. Rather than calling $\boldsymbol{\varphi}$ a curve, it is better to regard any vector function $\mathbf{g}$ obtained from $\boldsymbol{\varphi}$ by a suitable change of parameter as representing the same curve as $\boldsymbol{\varphi}$. Thus, one should define a curve as an equivalence class of equivalent parametric representations.

Definition 140 *Given two intervals $I, J \subseteq \mathbb{R}$ and two functions $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ and $\boldsymbol{\psi} : J \to \mathbb{R}^N$, we say that they are* equivalent *if there exists a continuous, bijective function $h : I \to J$ such that*

$$\boldsymbol{\varphi}(t) = \boldsymbol{\psi}(h(t))$$

for all $t \in I$. We write $\boldsymbol{\varphi} \sim \boldsymbol{\psi}$ and we call $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ parametric representations *and the function h a* parameter change.

Note that in view of a theorem done in class, $h^{-1} : J \to I$ is also continuous.

Exercise 141 *Prove that $\sim$ is an equivalence relation.*

Definition 142 *A* curve *$\boldsymbol{\gamma}$ is an equivalence class of parametric representations.*

Given a curve $\boldsymbol{\gamma}$ with parametric representation $\boldsymbol{\varphi} : I \to \mathbb{R}^N$, where $I \subseteq \mathbb{R}$ is an interval, the *multiplicity* of a point $\mathbf{y} \in \mathbb{R}^N$ is the (possibly infinite) number of points $t \in I$ such that $\boldsymbol{\varphi}(t) = \mathbf{y}$. Since every parameter change $h : I \to J$ is bijective, the multiplicity of a point does not depend on the particular parametric representation. The *range* of $\boldsymbol{\gamma}$ is the set of points of $\mathbb{R}^N$ with positive multiplicity, that is, $\boldsymbol{\varphi}(I)$. A point in the range of $\boldsymbol{\gamma}$ with multiplicity one is called a *simple point.* If every point of the range is simple, then $\boldsymbol{\gamma}$ is called a *simple arc.*

If $I = [a, b]$ and $\boldsymbol{\varphi}(a) = \boldsymbol{\varphi}(b)$, then the curve $\boldsymbol{\gamma}$ is called a *closed curve.* A closed curve is called *simple* if every point of the range is simple, with the exception of $\boldsymbol{\varphi}(a)$, which has multiplicity two.

Example 143 *The two functions*

$$\begin{aligned} \boldsymbol{\varphi}(t) &:= (\cos t, \sin t), \quad t \in [0, 2\pi], \\ \boldsymbol{\psi}(s) &:= (\cos 2s, \sin 2s), \quad s \in [0, \pi], \end{aligned}$$

are equivalent, take $h(t) = \frac{t}{2}$, while the function

$$\mathbf{p}(r) := (\cos r, \sin r), \quad r \in [0, 4\pi],$$

is not equivalent to the previous two, since the first curve is simple and the second no. Note that the range is the same, the unit circle.

Example 144 *Consider the curve with parametric representations* $\boldsymbol{\varphi} : \mathbb{R} \to \mathbb{R}^2$*, given by*

$$\boldsymbol{\varphi}(t) := (t(t-1), t(t-1)(2t-1)), \quad t \in \mathbb{R}.$$

Let's try to sketch the range of the curve. Set $x(t) := t(t-1)$ *and* $y(t) := t(t-1)(2t-1)$*. We have* $x(t) \geq 0$ *for* $t(t-1) \geq 0$*, that is, when* $t \geq 1$ *or* $t \leq 0$*, while* $x'(t) = 1(t-1) + t(1-0) = 2t - 1 \geq 0$ *for* $t \geq \frac{1}{2}$*. Moreover,*

$$\lim_{t\to-\infty} x(t) = \lim_{t\to-\infty} t(t-1) = -\infty(-\infty-1) = \infty,$$
$$\lim_{t\to\infty} x(t) = \lim_{t\to\infty} t(t-1) = \infty(\infty-1) = \infty.$$

On the other hand, $y(t) \geq 0$ *for* $t(t-1)(2t-1) \geq 0$*, that is, when* $0 \leq t \leq \frac{1}{2}$ *or* $t \geq 1$*, while* $y'(t) = 6t^2 - 6t + 1 \geq 0$ *for* $t \geq \frac{3+\sqrt{3}}{6}$ *or* $t \leq \frac{3-\sqrt{3}}{6}$*. Moreover,*

$$\lim_{t\to-\infty} y(t) = \lim_{t\to-\infty} t(t-1)(2t-1) = -\infty,$$
$$\lim_{t\to\infty} y(t) = \lim_{t\to\infty} t(t-1)(2t-1) = \infty.$$

To draw the range of the curve in the xy plane, note that x increasing means that we are moving to the right, x decreasing to the left, y increasing means that we are moving upwards, y decreasing downwards.

Next we start discussing the regularity of a curve.

Definition 145 *The curve* $\boldsymbol{\gamma}$ *is said to be* continuous *if one (and so all) of its parametric representations is continuous.*

The next result shows that the class of continuous curves is somehow too large for our intuitive idea of curve. Indeed, we construct a continuous curve that fills the unit square in $\mathbb{R}^2$. The first example of this type was given by Peano in 1890.

Theorem 146 (Peano) *There exists a continuous function* $\boldsymbol{\varphi} : [0,1] \to \mathbb{R}^2$ *such that* $\boldsymbol{\varphi}([0,1]) = [0,1] \times [0,1]$*.*

Friday, March 29, 2013

In view of the previous result, to recover the intuitive idea of a curve, we restrict the class of continuous curves to those with finite length.

In what follows, given an interval $I \subseteq \mathbb{R}$, a *partition* of I is a finite set $P := \{t_0, \ldots, t_n\} \subset I$, where

$$t_0 < t_1 < \cdots < t_n.$$

Definition 147 *Let* $I \subseteq \mathbb{R}$ *be an interval and let* $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ *be a function. The* pointwise variation *of* $\boldsymbol{\varphi}$ *on the interval* I *is*

$$\mathrm{Var}_I\, \boldsymbol{\varphi} := \sup\left\{\sum_{i=1}^{n} \|\boldsymbol{\varphi}(t_i) - \boldsymbol{\varphi}(t_{i-1})\|\right\}, \tag{13}$$

where the supremum is taken over all partitions $P := \{t_0, \ldots, t_n\}$ *of* I*,* $n \in \mathbb{N}$*. A function* $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ *has* finite *or* bounded pointwise variation *if* $\mathrm{Var}_I\, \boldsymbol{\varphi} < \infty$*.*

Exercise 148 *Prove that if two parametric representations* $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ *and* $\boldsymbol{\psi} : J \to \mathbb{R}^N$ *are equivalent, then* $\operatorname{Var}_I \boldsymbol{\varphi} = \operatorname{Var}_J \boldsymbol{\psi}$*.*

We are now ready to define the length of a curve.

Definition 149 *Given a curve* $\boldsymbol{\gamma}$*, let* $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ *be a parametric representation of* $\boldsymbol{\gamma}$*, where* $I \subseteq \mathbb{R}$ *is an interval. We define the* length *of* $\boldsymbol{\gamma}$ *as*

$$L(\boldsymbol{\gamma}) := \operatorname{Var}_I \boldsymbol{\varphi}.$$

We say that the curve $\boldsymbol{\gamma}$ *is* rectifiable *if* $L(\boldsymbol{\gamma}) < \infty$*.*

In view of the previous exercise, the definition makes sense, that is, the length of the curve does not depend on the particular representation of the curve.

Peano's example shows that continuous curves in general are not rectifiable, that is, they may have infinite length. Next we show that C^1 or piecewise C^1 curves are rectifiable.

Definition 150 *Given a function* $\boldsymbol{\varphi} : [a, b] \to \mathbb{R}^N$*, we say that* f *is* piecewise C^1 *if there exists a partition* $P = \{t_0, \ldots, t_n\}$ *of* $[a, b] \in I$*, with* $a = t_0 < t_1 < \cdots < t_n = b$ *such that* $\boldsymbol{\varphi}$ *is of class* C^1 *in each interval* $[t_{i-1}, t_i]$*,* $i = 1, \ldots, n$*.*

Note that at each t_i, $i = 1, \ldots, n-1$, the function $\boldsymbol{\varphi}$ has a right and a left derivative, but they might be different.

Example 151 *Consider the curve with parametric representations* $\boldsymbol{\varphi} : [-1, 1] \to \mathbb{R}^2$*, given by*

$$\boldsymbol{\varphi}(t) := (t, |t|), \quad t \in [-1, 1],$$

is not a curve of class C^1 *since at* $t = 0$ *the second component of* $\boldsymbol{\varphi}$ *is not differentiable, but it is piecewise* C^1*, since*

$$\boldsymbol{\varphi}(t) = (t, -t), \quad t \in [-1, 0],$$
$$\boldsymbol{\varphi}(t) = (t, t), \quad t \in [0, 1],$$

are both C^1 *functions. Note that the left derivative at* 0 *is* $\boldsymbol{\varphi}'_-(0) = (1, -1)$ *while the right derivative is* $\boldsymbol{\varphi}'_+(0) = (1, 1)$*. Hence, the curve has a corner at the point* $\boldsymbol{\varphi}(0) = (0, 0)$*.*

By requiring parametric representations, parameter changes and their inverses to be differentiable, or Lipschitz, or of class C^n, $n \in \mathbb{N}_0$, etc., or piecewise C^1, we may define curves $\boldsymbol{\gamma}$ that are differentiable, or Lipschitz, or of class C^n, $n \in \mathbb{N}_0$, or piecewise C^1, respectively.

Theorem 152 *Let* $\boldsymbol{\gamma}$ *be a curve of class* C^1*, with parametric representation* $\boldsymbol{\varphi} : [a, b] \to \mathbb{R}^N$*. Then*

$$\int_a^b \|\boldsymbol{\varphi}'(t)\| \, dt = L(\boldsymbol{\gamma}).$$

Given a curve $\boldsymbol{\gamma}$, with parametric representation $\boldsymbol{\varphi} : I \to \mathbb{R}^N$, where $I \subseteq \mathbb{R}$ is an interval. If $\boldsymbol{\varphi}$ is differentiable at a point $t_0 \in I$ and $\boldsymbol{\varphi}'(t_0) \neq 0$, the straight line parametrized by

$$\boldsymbol{\psi}(t) := \boldsymbol{\varphi}(t_0) + \boldsymbol{\varphi}'(t_0)(t - t_0), \quad t \in \mathbb{R},$$

is called a *tangent line* to $\boldsymbol{\gamma}$ at the point $\boldsymbol{\varphi}(t_0)$, and the vector $\boldsymbol{\varphi}'(t_0)$ is called a *tangent vector* to $\boldsymbol{\gamma}$ at the point $\boldsymbol{\varphi}(t_0)$. Note that if $\boldsymbol{\gamma}$ is a simple arc, then the tangent line at a point $\mathbf{y}$ of the range Γ of $\boldsymbol{\gamma}$, if it exists, is unique, while if the curve is self-intersecting at some point $\mathbf{y} \in \Gamma$, then there could be more than one tangent line at $\mathbf{y}$.

Given a function $f : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval, the graph of f is the range of a curve parametrized by

$$\boldsymbol{\varphi}(t) = (t, f(t)).$$

If f is differentiable at some point $t_0 \in I$, then the tangent vector is given by

$$\boldsymbol{\varphi}'(t_0) = (1, f'(t_0)).$$

Definition 153 *Given a rectifiable curve $\boldsymbol{\gamma}$, we say that $\boldsymbol{\gamma}$ is parametrized by* arclength *if it admits a parametrization $\boldsymbol{\psi} : [0, L(\gamma)] \to \mathbb{R}^N$ such that $\boldsymbol{\psi}$ is differentiable for all but finitely many $\tau \in [0, L(\gamma)]$ with $\left\|\boldsymbol{\psi}'(\tau)\right\| = 1$ for all but finitely many τ.*

The name arclength comes from the fact that if $\boldsymbol{\gamma}$ is piecewise C^1 then for every $\tau_1 < \tau_2$, by Theorem 152 we have that the length of the portion of the curve parametrized by $\boldsymbol{\psi} : [\tau_1, \tau_2] \to \mathbb{R}^N$ is given by

$$\int_{\tau_1}^{\tau_2} \left\|\boldsymbol{\psi}'(\tau)\right\| d\tau = \int_{\tau_1}^{\tau_2} 1 \, d\tau = \tau_2 - \tau_1.$$

Definition 154 *A curve $\boldsymbol{\gamma}$ is* regular *if is piecewise C^1 and it admits a parametric representation $\boldsymbol{\varphi} : [a, b] \to \mathbb{R}^N$ for which the left and right derivative are different from zero for all $t \in [a, b]$.*

Theorem 155 *If $\boldsymbol{\gamma}$ is a regular curve, then $\boldsymbol{\gamma}$ can be parametrized by arclength.*

Proof. Let $\boldsymbol{\varphi} : [a, b] \to \mathbb{R}^N$ be a piecewise C^1 parametric representation of $\boldsymbol{\gamma}$ and define the *length function*

$$s(t) := \int_a^t \|\boldsymbol{\varphi}'(r)\| \, dr.$$

Note that by by Theorem 152, $s(b) = L(\boldsymbol{\gamma})$, while $s(a) = 0$. Hence, $s : [a, b] \to [0, L(\gamma)]$ is onto. We claim that s is invertible. Indeed, if $t_1 < t_2$, then since $\|\boldsymbol{\varphi}'(r)\| > 0$ for all but finitely many t in $[t_1, t_2]$, we have that

$$s(t_2) - s(t_1) = \int_{t_1}^{t_2} \|\boldsymbol{\varphi}'(r)\| \, dr > 0$$

(why?). Hence, $s : [a, b] \to [0, L(\gamma)]$ is invertible. By a theorem in Real Analysis, $s^{-1} : [0, L(\gamma)] \to [a, b]$ is continuous. By the fundamental theorem of calculus, $s'(t) = \|\boldsymbol{\varphi}'(t)\| > 0$ for all but finitely many t. Hence, s is piecewise C^1. In turn, by a theorem in the Real Analysis, for any such t,

$$\frac{d}{d\tau}\left(s^{-1}\right)(s(t)) = \frac{1}{s'(t)}.$$

Hence, there exists

$$\frac{d}{d\tau}\left(s^{-1}\right)(\tau) = \frac{1}{s'(s^{-1}(\tau))} = \frac{1}{\|\boldsymbol{\varphi}'(s^{-1}(\tau))\|} \tag{14}$$

for all but finitely many τ. Since $\boldsymbol{\varphi}$ is piecewise C^1 and the left and right derivatives of $\boldsymbol{\varphi}$ are always different from zero, it follows from (14) and the continuity of s^{-1} that s^{-1} is piecewise C^1. Thus, the function $\boldsymbol{\psi} : [0, L(\gamma)] \to \mathbb{R}^N$, defined by

$$\boldsymbol{\psi}(\tau) := \boldsymbol{\varphi}\left(s^{-1}(\tau)\right), \quad \tau \in [0, L(\gamma)], \tag{15}$$

is equivalent to $\boldsymbol{\varphi}$. Moreover, by the chain rule and (14), for all but finitely many τ,

$$\boldsymbol{\psi}'(\tau) = \boldsymbol{\varphi}'\left(s^{-1}(\tau)\right)\frac{d}{d\tau}\left(s^{-1}\right)(\tau) = \frac{\boldsymbol{\varphi}'\left(s^{-1}(\tau)\right)}{\|\boldsymbol{\varphi}'(s^{-1}(\tau))\|}$$

and so $\left\|\boldsymbol{\psi}'(\tau)\right\| = 1$. ■

Example 156 *Consider the curve with parametric representations* $\boldsymbol{\varphi} : [0, 2\pi] \to \mathbb{R}^3$*, given by*

$$\boldsymbol{\varphi}(t) := (R\cos t, R\sin t, 2t), \quad t \in [0, 2\pi].$$

Since

$$\boldsymbol{\varphi}'(t) = (-R\sin t, R\cot st, 2) \neq (0, 0, 0),$$

the curve is regular. Let's parametrize it using arclength. Set

$$\begin{aligned} s(t) &:= \int_0^t \|\boldsymbol{\varphi}'(r)\|\, dr = \int_0^t \sqrt{R^2\sin^2 t + R^2\cos^2 t + 4}\, dr \\ &= \sqrt{R^2 + 4}t. \end{aligned}$$

In particular, $s(2\pi) = 2\pi\sqrt{R^2 + 4}$ *is the length of the curve. The inverse of* s *is* $t(s) := \frac{s}{\sqrt{R^2+4}}$, $s \in \left[0, 2\pi\sqrt{R^2 + 4}\right]$*. Hence,*

$$\boldsymbol{\psi}(s) := \boldsymbol{\varphi}(t(s)) = \left(R\cos\frac{s}{\sqrt{R^2+4}}, R\sin\frac{s}{\sqrt{R^2+4}}, \frac{2s}{\sqrt{R^2+4}}\right),$$

where $s \in \left[0, 2\pi\sqrt{R^2 + 4}\right]$*, is the parametrization by arclength.*

14 Curve Integrals

In this section we define the integral of a function along a curve.

Definition 157 *Given a piecewise C^1 curve $\boldsymbol{\gamma}$ with parametric representation $\boldsymbol{\varphi} : [a,b] \to \mathbb{R}^N$ and a bounded function $f : E \to \mathbb{R}$, where $\boldsymbol{\varphi}([a,b]) \subseteq E$, we define the* curve *(or* line*)* integral *of f along the curve $\boldsymbol{\gamma}$ as the number*

$$\int_{\boldsymbol{\gamma}} f\,ds := \int_a^b f(\boldsymbol{\varphi}(t)) \|\boldsymbol{\varphi}'(t)\|\,dt,$$

whenever the function $t \in [a,b] \mapsto f(\boldsymbol{\varphi}(t)) \|\boldsymbol{\varphi}'(t)\|$ is Riemann integrable.

Note that $\int_{\boldsymbol{\gamma}} f\,ds$ is always defined if f is continuous.

Next we show that the curve integral does not depend on the particular parametric representation.

Proposition 158 *Let $\boldsymbol{\gamma}$ be a piecewise C^1 curve and let $\boldsymbol{\varphi} : [a,b] \to \mathbb{R}^N$ and $\boldsymbol{\psi} : [c,d] \to \mathbb{R}^N$ be two parametric representations. Given a continuous function $f : E \to \mathbb{R}$, where E contains the range of $\boldsymbol{\gamma}$,*

$$\int_a^b f(\boldsymbol{\varphi}(t)) \|\boldsymbol{\varphi}'(t)\|\,dt = \int_c^d f(\boldsymbol{\psi}(\tau)) \left\|\boldsymbol{\psi}'(\tau)\right\|\,d\tau.$$

Proof. Exercise ■

Remark 159 *In particular, if $\boldsymbol{\gamma}$ is a regular curve and $\boldsymbol{\psi} : [0, L(\boldsymbol{\gamma})] \to \mathbb{R}^N$ is the parametric representation obtained using arclength, then $\left\|\boldsymbol{\psi}'(\tau)\right\| = 1$ for all but finitely many τ. Hence,*

$$\int_{\boldsymbol{\gamma}} f\,ds = \int_0^{L(\boldsymbol{\gamma})} f(\boldsymbol{\psi}(\tau))\,d\tau.$$

Monday, April 1, 2013

Next we introduce the notion of an oriented curve.

Definition 160 *Given a curve $\boldsymbol{\gamma}$ with parametric representations $\boldsymbol{\varphi} : I \to \mathbb{R}^N$ and $\boldsymbol{\psi} : J \to \mathbb{R}^N$, we say that $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ have the* same orientation *if the parameter change $h : I \to J$ is increasing and* opposite orientation *if the parameter change $h : I \to J$ is decreasing. If $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ have the same orientation, we write $\boldsymbol{\varphi} \overset{*}{\sim} \boldsymbol{\psi}$.*

Exercise 161 *Prove that $\overset{*}{\sim}$ is an equivalence relation.*

Definition 162 *An* oriented curve *$\boldsymbol{\gamma}$ is an equivalence class of parametric representations with the same orientation.*

Note that any curve $\boldsymbol{\gamma}$ gives rise to two oriented curves. Indeed, it is enough to fix a parametric representation $\boldsymbol{\varphi}: I \to \mathbb{R}^N$ and considering the equivalence class $\boldsymbol{\gamma}^+$ of parametric representations with the same orientation of $\boldsymbol{\varphi}$ and the equivalence class $\boldsymbol{\gamma}^-$ of parametric representations with the opposite orientation of $\boldsymbol{\varphi}$.

Let $E \subseteq \mathbb{R}^N$ and let $\mathbf{g}: E \to \mathbb{R}^N$ be a continuous function. Given a piecewise C^1 oriented curve $\boldsymbol{\gamma}$ with parametric representation $\boldsymbol{\varphi}: [a,b] \to \mathbb{R}^N$ such that $\boldsymbol{\varphi}([a,b]) \subseteq E$, we define

$$\int_{\boldsymbol{\gamma}} \mathbf{g} := \int_a^b \sum_{i=1}^N g_i(\boldsymbol{\varphi}(t))\, \varphi_i'(t)\, dt.$$

Definition 163 *Let $U \subseteq \mathbb{R}^N$ be an open set and let $\mathbf{g}: U \to \mathbb{R}^N$. We say that $\mathbf{g}$ is* conservative vector field *if there exists a differentiable function $f: U \to \mathbb{R}$ such that*

$$\nabla f(\mathbf{x}) = \mathbf{g}(\mathbf{x})$$

for all $\mathbf{x} \in U$. The function f is called a scalar potential *for $\mathbf{g}$.*

Definition 164 *Given $E \subseteq \mathbb{R}^N$, we say that*

(i) E is disconnected *if there exist two nonempty open sets $U, V \subseteq \mathbb{R}^N$ such that*

$$E \subseteq U \cup V, \quad E \cap U \neq \emptyset, \quad E \cap V \neq \emptyset, \quad E \cap U \cap V = \emptyset.$$

(ii) E is connected *if it is not disconnected.*

Theorem 165 *Let $U \subseteq \mathbb{R}^N$ be an open connected set and let $\mathbf{g}: U \to \mathbb{R}^N$ be a continuous function. Then the following conditions are equivalent.*

(i) $\mathbf{g}$ is a conservative vector field,

(ii) for every $\mathbf{x}$, $\mathbf{y} \in U$ and for every two piecewise C^1 oriented curves $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ with parametric representations $\varphi_1: [a,b] \to \mathbb{R}^N$ and $\varphi_2: [c,d] \to \mathbb{R}^N$, respectively, such that $\varphi_1(b) = \varphi_2(d) = \mathbf{x}$, $\varphi_1(a) = \varphi_2(c) = \mathbf{y}$, and $\varphi_1([a,b]), \varphi_2([c,d]) \subset U$,

$$\int_{\boldsymbol{\gamma}_1} \mathbf{g} = \int_{\boldsymbol{\gamma}_2} \mathbf{g}.$$

(iii) for every piecewise C^1 closed oriented curve $\boldsymbol{\gamma}$ with range contained in U,

$$\int_{\boldsymbol{\gamma}} \mathbf{g} = \mathbf{0}.$$

Next we give a simple necessary condition for a field $\mathbf{g}$ to be conservative.

Definition 166 *Let $U \subseteq \mathbb{R}^N$ be an open set and let $\mathbf{g} : U \to \mathbb{R}^N$ be differentiable. We say that $\mathbf{g}$ is an* irrotational vector field *or* a curl-free vector field *if*

$$\frac{\partial g_i}{\partial x_j}(\mathbf{x}) = \frac{\partial g_j}{\partial x_i}(\mathbf{x})$$

for all $i, j = 1, \dots, N$ and all $\mathbf{x} \in U$.

Theorem 167 *Let $U \subseteq \mathbb{R}^N$ be an open set and let $\mathbf{g} : U \to \mathbb{R}^N$ be a conservative vector field of class C^1. Then $\mathbf{g}$ is irrotational.*

Proof. Since $\mathbf{g}$ is a conservative vector field, there exists a a scalar potential $f : U \to \mathbb{R}$ with $\nabla f = \mathbf{g}$ in U. But since $\mathbf{g}$ is of class C^1, we have that f is of class C^2. Hence, we are in a position to apply the Schwartz theorem to conclude that

$$\frac{\partial g_i}{\partial x_j}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = \frac{\partial g_j}{\partial x_i}(\mathbf{x})$$

for all $i, j = 1, \dots, N$ and all $\mathbf{x} \in U$. ■

The next example shows that there exist irrotational vector fields that are not conservative.

Example 168 *Let $U := \mathbb{R}^2 \setminus \{(0,0)\}$ and consider the function*

$$\mathbf{g}(x, y) := \left(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2}\right).$$

Note that

$$\frac{\partial}{\partial y}\left(\frac{-y}{x^2 + y^2}\right) = \frac{-x^2 + y^2}{(x^2 + y^2)^2},$$
$$\frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2}\right) = \frac{-x^2 + y^2}{(x^2 + y^2)^2},$$

and so $\mathbf{g}$ is irrotational. However, if we consider the closed curve $\boldsymbol{\gamma}$ of parametric representation $\varphi(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$, we have that

$$\int_{\boldsymbol{\gamma}} \mathbf{g} = \int_0^{2\pi} \left(-\frac{\sin t}{\cos^2 t + \sin^2 t}(\cos t)' + \frac{\cos t}{\cos^2 t + \sin^2 t}(\sin t)'\right) dt$$
$$= \int_0^{2\pi} \left(\sin^2 t + \cos^2 t\right) dt = 2\pi \neq 0,$$

and so $\mathbf{g}$ cannot be conservative.

The problem here is the fact that the domain has a hole.

Wednesday, April 3, 2013

Definition 169 *A set $E \subseteq \mathbb{R}^N$ is* starshaped *with respect to a point $\mathbf{x}_0 \in E$ if for every $\mathbf{x} \in E$, the segment joining $\mathbf{x}$ and $\mathbf{x}_0$ is contained in E.*

Theorem 170 (Poincaré's Lemma) *Let $U \subseteq \mathbb{R}^N$ be an open set starshaped with respect to a point $\mathbf{x}_0$ and let $\mathbf{g} : U \to \mathbb{R}^N$ be an irrotational vector field of class C^1. Then $\mathbf{g}$ is a conservative vector field.*

Example 171 *Consider the function $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^2$ defined by*

$$\mathbf{g}(x, y) := (ye^x, e^x - \cos y).$$

We have,

$$\frac{\partial}{\partial y}(ye^x) = 1e^x,$$
$$\frac{\partial}{\partial x}(e^x - \cos y) = e^x - 0,$$

and so $\mathbf{g}$ is irrotational. Since $\mathbb{R}^2$ is convex, it is starshaped with respect to every point. Hence, by the previous theorem, $\mathbf{g}$ is conservative. To find a scalar potential, note that

$$\frac{\partial f}{\partial x}(x, y) = ye^x, \quad \frac{\partial f}{\partial y}(x, y) = e^x - \cos y$$

and so

$$\begin{aligned} f(x, y) - f(0, y) &= \int_0^x \frac{\partial f}{\partial x}(s, y)\, ds = \int_0^x ye^s\, ds \\ &= y(e^x - 1). \end{aligned}$$

Hence,

$$f(x, y) = f(0, y) + y(e^x - 1).$$

Differentiating with respect to y, we get

$$e^x - \cos y = \frac{\partial f}{\partial y}(x, y) = \frac{\partial f}{\partial y}(0, y) + e^x - 1$$

and so

$$\frac{\partial f}{\partial y}(0, y) = -\cos y + 1.$$

Hence,

$$\begin{aligned} f(0, y) - f(0, 0) &= \int_0^y \frac{\partial f}{\partial y}(0, t)\, dt = \int_0^y (-\cos t + 1)\, dt \\ &= y - \sin y \end{aligned}$$

and so

$$\begin{aligned} f(x, y) &= f(0, y) + y(e^x - 1) = \\ &= f(0, 0) + y - \sin y + ye^x - y = f(0, 0) - \sin y + ye^x. \end{aligned}$$

The Poincaré lemma continues to hold in a very important class of sets, namely *simply connected sets.*

Definition 172 *Given a set* $E \subseteq \mathbb{R}^N$*, two continuous closed curves, with parametric representations* $\varphi : [a, b] \to \mathbb{R}^N$ *and* $\psi : [a, b] \to \mathbb{R}^N$ *such that* $\varphi([a, b]) \subseteq E$ *and* $\psi([a, b]) \subseteq E$*, are* homotopic *in* E *if there exists a continuous function* $\mathbf{h} : [a, b] \times [0, 1] \to \mathbb{R}^N$ *such that* $\mathbf{h}([a, b] \times [0, 1]) \subseteq E$,

$$\mathbf{h}(t, 0) = \varphi(t) \text{ for all } t \in [a, b], \quad \mathbf{h}(t, 1) = \psi(t) \text{ for all } t \in [a, b],$$
$$\mathbf{h}(a, s) = \mathbf{h}(b, s) \text{ for all } s \in [0, 1].$$

The function $\mathbf{h}$ *is called a* homotopy in E *between the two curves.*

Roughly speaking, two curves are homotopic in E if it is possible to deform the first continuously until it becomes the second *without leaving* the set E.

Definition 173 *A set* $E \subseteq \mathbb{R}^N$ *is* pathwise connected *if for every two points* $\mathbf{x}, \mathbf{y} \in E$ *there exists a continuous curve joining* $\mathbf{x}$ *and* $\mathbf{y}$ *and with range contained in* E.

Definition 174 *A set* $E \subseteq \mathbb{R}^N$ *is* simply connected *if it is pathwise connected and if every continuous closed curve with range in* E *is homotopic in* E *to a point in* E *(that is, to a curve with parametric representation a constant function).*

Example 175 *A star-shaped set is simply connected. Indeed, let* $E \subseteq \mathbb{R}^N$ *be star-shaped with respect to some point* $\mathbf{x}_0 \in E$ *and consider a continuous closed curve* $\boldsymbol{\gamma}$ *with parametric representation* $\boldsymbol{\varphi} : [a, b] \to \mathbb{R}^N$ *such that* $\boldsymbol{\varphi}([a, b]) \subseteq E$. *Then the function*

$$\mathbf{h}(t, s) := s\boldsymbol{\varphi}(t) + (1 - s)\mathbf{x}_0$$

is an homotopy between $\boldsymbol{\gamma}$ *and the point* $\mathbf{x}_0$.

Remark 176 *It can be shown that* $\mathbb{R}^2 \setminus$ *a point is not simply connected, while* $\mathbb{R}^3 \setminus$ *a point is. On the other hand,* $\mathbb{R}^3 \setminus$ *a line is not simply connected.*

Friday, April 5, 2013

Given an open set $U \subseteq \mathbb{R}^N$ and a differentiable function $\mathbf{g} : U \to \mathbb{R}^N$, a differentiable function $h : U \to \mathbb{R}$ is called an *integrating factor* for $\mathbf{g}$ if the function

$$h\mathbf{g} = (hg_1, \dots, hg_N)$$

is irrotational.

Exercise 177 *Assume that* $N = 2$ *and that* $\mathbf{g} : U \to \mathbb{R}^2$ *is of class* C^2. *Prove that if*

$$\frac{\partial g_1}{\partial y}(x, y) - \frac{\partial g_2}{\partial x}(x, y) = a(x) g_2(x, y) - b(y) g_1(x, y)$$

for some continuous functions $a : I \to \mathbb{R}$ *and* $b : J \to \mathbb{R}$*, where* I *and* J *are intervals, then an integrating factor is given by*

$$h(x, y) = e^{\int_{x_0}^{x} a(t)\, dt} e^{\int_{y_0}^{y} b(s)\, ds},$$

where $x_0 \in I$ *and* $y_0 \in J$ *are some fixed points.*

Remark 178 *Conservative fields are useful is solving differential equations. Consider the differential equation*

$$y'(x) = -\frac{g_1(x, y(x))}{g_2(x, y(x))},$$

where

$$\mathbf{g}(x, y) = (g_1(x, y), g_2(x, y))$$

is of class C^1. If $\mathbf{g}$ conservative, then there exists a function $f : U \to \mathbb{R}$ such that $\nabla f = \mathbf{g}$, so that

$$y'(x) = -\frac{\frac{\partial f}{\partial x}(x, y(x))}{\frac{\partial f}{\partial y}(x, y(x))},$$

which implies that

$$\frac{\partial f}{\partial x}(x, y(x)) + \frac{\partial f}{\partial y}(x, y(x))\, y'(x) = 0.$$

Consider the function

$$h(x) := f(x, y(x)).$$

By the chain rule, we have that

$$h'(x) = \frac{\partial f}{\partial x}(x, y(x)) + \frac{\partial f}{\partial y}(x, y(x))\, y'(x) = 0.$$

Hence, h is constant in each connected domain of its domain. Assume that h is defined in some interval I, then

$$f(x, y(x)) = \text{constant}$$

for all $x \in I$. Thus, we have solved implicitly our differential equation.

If $\mathbf{g}$ is not irrotational, but it admits an integrating factor $h \neq 0$, then

$$y'(x) = -\frac{g_1(x, y(x))}{g_2(x, y(x))} = -\frac{h(x, y(x))\, g_1(x, y(x))}{h(x, y(x))\, g_2(x, y(x))},$$

and if the function $h\mathbf{g}$ is conservative with $\nabla f = h\mathbf{g}$, then we can conclude as before that

$$f(x, y(x)) = \text{constant}$$

for all $x \in I$, so we have solved implicitly our differential equation.

Example 179 *We want to solve the differential equations*

$$y' = \frac{2\log(xy) + 1}{\frac{x}{y}}.$$

Consider the function

$$\mathbf{g}(x, y) = \left(2\log(xy) + 1, -\frac{x}{y}\right).$$

Let's see if $\mathbf{g}$ *is irrotational. We have*

$$\frac{\partial}{\partial y}\left(2\log\left(xy\right)+1\right)=\frac{2}{y}\neq-\frac{1}{y}=\frac{\partial}{\partial x}\left(-\frac{x}{y}\right).$$

: Hence, $\mathbf{g}$ *is not irrotational, and so it cannot be conservative. Let's try to find an integrating factor* h*. We have*

$$\begin{aligned}\frac{\partial g_1}{\partial y}\left(x,y\right)-\frac{\partial g_2}{\partial x}\left(x,y\right)=\frac{2}{y}+\frac{1}{y}&=a\left(x\right)g_2\left(x,y\right)-b\left(y\right)g_1\left(x,y\right)\\&=a\left(x\right)\left(-\frac{x}{y}\right)-b\left(y\right)\left(2\log\left(xy\right)+1\right).\end{aligned}$$

Take $a\left(x\right)=-\frac{3}{x}$ *and* $b\left(y\right)=0$*. Then*

$$h\left(x,y\right)=e^{\int_{x_0}^{x}-\frac{3}{x}\,dt}e^{\int_{y_0}^{y}0\,ds}=e^{-3\log|x|+3\log x_0}=ce^{\log|x|^{-3}}=\frac{c}{|x|^3}$$

is an integrating factor. Indeed,

$$\frac{\partial}{\partial y}\left(\frac{1}{x^3}\left(2\log\left(xy\right)+1\right)\right)=\frac{2}{x^3y}=\frac{\partial}{\partial x}\left(-\frac{1}{x^2y}\right).$$

Hence,

$$h\left(x,y\right)\mathbf{g}\left(x,y\right)=\left(\frac{1}{|x|^3}\left(2\log\left(xy\right)+1\right),-\frac{x}{|x|^3y}\right)$$

is irrotational. Its domain is the set $U=\left\{(x,y)\in\mathbb{R}^2:\,xy>0\right\}$*, which is given by the disjoint union of two convex sets. In each of these convex sets* $h\mathbf{g}$ *is conservative by the the Poincaré lemma. Let's find a potential* f*. We have*

$$\frac{\partial f}{\partial x}\left(x,y\right)=\frac{1}{|x|^3}\left(2\log\left(xy\right)+1\right),\quad\frac{\partial f}{\partial y}\left(x,y\right)=-\frac{x}{|x|^3y}$$

and so

$$\begin{aligned}f\left(x,y\right)&=\int\frac{\partial f}{\partial y}\left(x,y\right)\,dy=\int-\frac{x}{|x|^3y}\,dy\\&=-\frac{x}{|x|^3}\log|y|+c\left(x\right).\end{aligned}$$

Hence,

$$f\left(x,y\right)=-\frac{x}{|x|^3}\log|y|+c\left(x\right).$$

If $x>0$ *and* $y>0$ *differentiating with respect to* x*, we get*

$$\frac{1}{x^3}\left(2\log\left(xy\right)+1\right)=\frac{\partial f}{\partial x}\left(x,y\right)=\frac{\partial}{\partial x}\left(-\frac{1}{x^2}\log y+c\left(x\right)\right)=\frac{2}{x^3}\log y+c'\left(x\right)$$

and so

$$\frac{2}{x^3}\log x+\frac{2}{x^3}\log y+\frac{1}{x^3}=\frac{2}{x^3}\log y+c'(x).$$

It follows that

$$c'(x)=\frac{2}{x^3}\log x+\frac{1}{x^3}.$$

Hence,

$$c(x)=\int c'(x)\,dx=\int\left(\frac{2}{x^3}\log x+\frac{1}{x^3}\right)dx=-\frac{1}{x^2}(\log x+1)+c_1.$$

: If $x<0$ and $y<0$ differentiating with respect to x, we get

$$-\frac{1}{x^3}(2\log|xy|+1)=\frac{\partial f}{\partial x}(x,y)=\frac{\partial}{\partial x}\left(\frac{1}{x^2}\log|y|+c(x)\right)=-\frac{2}{x^3}\log|y|+c'(x)$$

and so

$$-\frac{2}{x^3}\log|x|-\frac{2}{x^3}\log|y|+\frac{1}{x^3}=-\frac{2}{x^3}\log|y|+c'(x).$$

It follows that

$$c'(x)=-\frac{2}{x^3}\log|x|+\frac{1}{x^3}.$$

Hence,

$$c(x)=\int c'(x)\,dx=\int\left(-\frac{2}{x^3}\log(-x)+\frac{1}{x^3}\right)dx=\frac{1}{x^2}\log(-x)+c_2.$$

In conclusion the potential is

$$f(x,y)=\begin{cases}-\frac{x}{|x|^3}\log|y|-\frac{1}{x^2}(\log x+1)+c_1 & \text{if } x>0,\\ -\frac{x}{|x|^3}\log|y|+\frac{1}{x^2}\log(-x)+c_2 & \text{if } x<0.\end{cases}$$

Monday, April 8, 2013

15 Integration

Given N bounded intervals $I_1,\dots,I_N\subset\mathbb{R}$, a *rectangle* in $\mathbb{R}^N$ is a set of the form

$$R:=I_1\times\cdots\times I_N.$$

The elementary measure of a rectangle is given by

$$\operatorname{meas}R:=\operatorname{length}I_1\cdot\cdots\cdot\operatorname{length}I_N,$$

where if I_n has endpoints $a_n\le b_n$, then we set $\operatorname{length}I_n:=b_n-a_n$. To highlight the dependence on N, we will use the notation meas_N.

Remark 180 *Note that the intersection of two rectangles is still a rectangle. We will use this fact a lot in what follows.*

Given a rectangle R, by a *partition* $\mathcal{P}$ of R we mean a finite set of rectangles $R_1, \ldots, R_n$ such that $R_i \cap R_j = \emptyset$ if $i \neq j$ and

$$R = \bigcup_{i=1}^{n} R_i.$$

Exercise 181 *Let $R \subset \mathbb{R}^N$ be a rectangle and let $\mathcal{P} = \{R_1, \ldots, R_n\}$ be a partition of R. Prove that*

$$\operatorname{meas} R = \sum_{i=1}^{n} \operatorname{meas} R_i.$$

Hint: Use induction on N.

Given a rectangle R, consider a bounded function $f : R \to \mathbb{R}$ and a partition $\mathcal{P}$ of R. We define the *lower* and *upper sums* of f for the partition $\mathcal{P}$ respectively by

$$L(f, \mathcal{P}) := \sum_{i=1}^{n} \operatorname{meas} R_i \inf_{\mathbf{x} \in R_i} f(\mathbf{x}),$$
$$U(f, \mathcal{P}) := \sum_{i=1}^{n} \operatorname{meas} R_i \sup_{\mathbf{x} \in R_i} f(\mathbf{x}).$$

Since f is bounded, using the previous exercise, we have that

$$\begin{aligned} \operatorname{meas} R \inf_{\mathbf{x} \in R} f(\mathbf{x}) &= \sum_{i=1}^{n} \operatorname{meas} R_i \inf_{\mathbf{x} \in R} f(\mathbf{x}) \leq L(f, \mathcal{P}) \qquad (16) \\ &\leq U(f, \mathcal{P}) \leq \sum_{i=1}^{n} \operatorname{meas} R_i \sup_{\mathbf{x} \in R} f(\mathbf{x}) \leq \operatorname{meas} R \sup_{\mathbf{x} \in R} f(\mathbf{x}). \end{aligned}$$

The *lower* and *upper integrals* of f over R are defined respectively by

$$\underline{\int_R} f(\mathbf{x})\, d\mathbf{x} := \sup\{L(f, \mathcal{P}) : \mathcal{P} \text{ partition of } R\},$$
$$\overline{\int_R} f(\mathbf{x})\, d\mathbf{x} := \inf\{U(f, \mathcal{P}) : \mathcal{P} \text{ partition of } R\}.$$

Given a rectangle R and a bounded function $f : R \to \mathbb{R}$, we say that f is *Riemann integrable* over R if the lower and upper integral coincide. We call the common value the *Riemann integral* of f over R and we denote it by $\int_R f(\mathbf{x})\, d\mathbf{x}$. Thus, for a Riemann integrable function,

$$\int_R f(\mathbf{x})\, d\mathbf{x} := \underline{\int_R} f(\mathbf{x})\, d\mathbf{x} = \overline{\int_R} f(\mathbf{x})\, d\mathbf{x}.$$

Remark 182 *Note that the function $f = 1$ is Riemann integrable over a rectangle R and*

$$\int_R 1\, dx = \operatorname{meas} R.$$

To determine when a function is Riemann integrable we need to introduce the notion of sets of Lebesgue measure zero.

Definition 183 *A set $E \subset \mathbb{R}^N$ has* Lebesgue measure zero *if for every $\varepsilon > 0$ there exists a countable family of rectangles R_n such that*

$$E \subset \bigcup_n R_n \quad \text{and} \quad \sum_n \operatorname{meas} R_n \leq \varepsilon.$$

Example 184 *Let's see some examples.*

(i) A singleton $E = \{\mathbf{c}\}$ has Lebesgue measure zero. Given $\varepsilon > 0$, take R to be the cube centered at $\mathbf{c}$ and of side-length $\varepsilon^{1/N}$.

(ii) If a set E contains an open rectangle R, then it cannot have Lebesgue measure zero. Indeed, for any countable family of rectangles R_n, we have

$$R \subset E \subset \bigcup_n R_n,$$

and so

$$\operatorname{meas} R \leq \sum_n \operatorname{meas} R_n.$$

Taking $\varepsilon < \operatorname{meas} R$, we obtain a contradiction.

(iii) A countable set $E = \{\mathbf{x}_n\}_n$ has Lebesgue measure zero. Given $\varepsilon > 0$, take R_n to be the cube centered at $\mathbf{x}_n$ and of side-length $\left(\frac{\varepsilon}{2^n}\right)^{1/N}$. Then

$$\sum_n \operatorname{meas} R_n = \varepsilon \sum_n \frac{1}{2^n} \leq \varepsilon.$$

In particular, $\mathbb{N}$, $\mathbb{Q}$, and $\mathbb{Z}$ all have Lebesgue measure zero.

(iv) If $\{E_k\}_k$ is a countable family of sets, each with Lebesgue measure zero, then their union

$$E := \bigcup_k E_k$$

has Lebesgue measure zero. Indeed, given $\varepsilon > 0$ fix k. Since E_k has Lebesgue measure zero, there exists a countable family of rectangles $R_n^{(k)}$ such that

$$E_k \subset \bigcup_n R_n^{(k)} \quad \text{and} \quad \sum_n \operatorname{meas} R_n^{(k)} \leq \frac{\varepsilon}{2^k}.$$

Consider the family $\left\{R_n^{(k)}\right\}_{n,k}$. *It is still countable,*

$$E = \bigcup_k E_k \subset \bigcup_k \bigcup_n R_n^{(k)}$$

and

$$\sum_{n,k} \operatorname{meas} R_n^{(k)} = \sum_k \sum_n \operatorname{meas} R_n^{(k)} \leq \sum_k \frac{\varepsilon}{2^k} \leq \varepsilon.$$

(v) *There are uncountable sets that have Lebesgue measure zero. One such example is given by the Cantor set.*

Exercise 185 *Prove that the boundary of a rectangle* $R \subset \mathbb{R}^N$ *has Lebesgue measure zero.*

Exercise 186 *Prove that in Definition 183, the rectangles* R_n *can be assumed to be open.*

The following theorem characterizes Riemann integrable functions.

Theorem 187 *Given a rectangle* R*, a bounded function* $f : R \to \mathbb{R}$ *is Riemann integrable if and only if the set of its discontinuity points has Lebesgue measure zero.*

Wednesday, April 10, 2013

Second midterm

Friday, April 5, 2013

The next theorem is very important in exercises. It allows to calculate triple, double, etc.. integrals by integrating one variable at a time.

Theorem 188 (Repeated Integration) *Let* $S \subset \mathbb{R}^N$ *and* $T \subset \mathbb{R}^M$ *be rectangles, let* $f : S \times T \to \mathbb{R}$ *be Riemann integrable and assume that for every* $\mathbf{x} \in S$*, the function* $\mathbf{y} \in T \mapsto f(\mathbf{x}, \mathbf{y})$ *is Riemann integrable. Then the function* $\mathbf{x} \in S \mapsto \int_T f(\mathbf{x}, \mathbf{y})\, d\mathbf{y}$ *is Riemann integrable and*

$$\int_{S\times T} f(\mathbf{x}, \mathbf{y})\, d(\mathbf{x}, \mathbf{y}) = \int_S \left(\int_T f(\mathbf{x}, \mathbf{y})\, d\mathbf{y} \right) d\mathbf{x}. \tag{17}$$

Similarly, if for every $\mathbf{y} \in T$*, the function* $\mathbf{x} \in S \mapsto f(\mathbf{x}, \mathbf{y})$ *is Riemann integrable, then the function* $\mathbf{y} \in T \mapsto \int_S f(\mathbf{x}, \mathbf{y})\, d\mathbf{x}$ *is Riemann integrable and*

$$\int_{S\times T} f(\mathbf{x}, \mathbf{y})\, d(\mathbf{x}, \mathbf{y}) = \int_T \left(\int_S f(\mathbf{x}, \mathbf{y})\, d\mathbf{x} \right) d\mathbf{y}.$$

Example 189 *Let's calculate* $\int_{[0,1]\times[0,1]} f(x, y)\, d(x, y)$*, where*

$$f(x, y) = x \sin(x + y).$$

Since f is continuous and bounded, we can apply the previous theorem. We have

$$\begin{aligned}\int_{[0,1]\times[0,1]} x\sin(x+y)\ d(x,y) &= \int_0^1 \left(\int_0^1 x\sin(x+y)\ dy\right)\ dx \\ &= \int_0^1 x\left(\int_0^1 \sin(x+y)\ dy\right)\ dx \\ &= \int_0^1 x\left[-\cos(x+y)\right]_{y=0}^{y=1}\ dx \\ &= \int_0^1 x\left[-\cos(x+1)+\cos(x+0)\right]\ dx.\end{aligned}$$

We use integration by parts with $h(x) = x$ so that $h'(x) = 1$ and $g'(x) = -\cos(x+1) + \cos x$, so that $g(x) = -\sin(x+1) + \sin x$. We have

$$\begin{aligned}&\int_0^1 \underset{h}{x}\underset{g'}{[-\cos(x+1)+\cos x]}\,dx \\ &= \left[\underset{h}{x}\underset{g}{[-\sin(x+1)+\sin x]}\right]_{x=0}^{x=1} - \int_0^1 \underset{h'}{1}\underset{g}{[-\sin(x+1)+\sin x]}\,dx \\ &= \sin 1 - \sin 2 - 0 + \int_0^1 \left[\sin(x+1) - \sin x\right]\ dx.\end{aligned}$$

We can also define the Riemann integral over bounded sets E. Given a bounded set $E \subset \mathbb{R}^N$, let R be a rectangle containing E. We say that function $f : E \to \mathbb{R}$ is *Riemann integrable over E* if the function

$$g(\mathbf{x}) := \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \in R \\ 0 & \text{if } \mathbf{x} \in R \setminus E \end{cases}$$

is Riemann integrable over R and we define the *Riemann integral of f over E* to be

$$\int_E f(\mathbf{x})\ d\mathbf{x} := \int_R g(\mathbf{x})\ d\mathbf{x}.$$

Exercise 190 *Prove the previous definition does not depend on the choice of the particular rectangle R containing E.*

Given a bounded set $E \subset \mathbb{R}^N$, let R be a rectangle containing E. We say that E is *Peano–Jordan measurable* if the function χ_E is Riemann integrable over R. In this case the *Peano–Jordan measure* of E is given by

$$\operatorname{meas} E := \int_R \chi_E(\mathbf{x})\ d\mathbf{x}.$$

In dimension $N = 2$, the Peano–Jordan of a set is called area of the set and in dimension $N = 3$ it is called volume of a set.

Exercise 191 *Prove that the previous definition does not depend on the choice of the rectangle R containing E.*

Remark 192 *Note that a rectangle is Peano–Jordan measurable and its Peano–Jordan measure coincides with its elementary measure (see Remark 182).*

Theorem 193 *A bounded set $E \subset \mathbb{R}^N$ is Peano–Jordan measurable if and only if its boundary is Peano–Jordan measurable and it has Peano–Jordan measure zero.*

Definition 194 *A set $E \subset \mathbb{R}^{N+1}$ is called a* normal domain *if it can be written as*

$$E = \{(\mathbf{x}, y) \in R \times \mathbb{R} : \, \alpha(\mathbf{x}) \leq y \leq \beta(\mathbf{x})\},$$

where $R \subset \mathbb{R}^N$ is a rectangle, and $\alpha : R \to \mathbb{R}$ and $\beta : R \to \mathbb{R}$ are two Riemann integrable functions, with $\alpha(\mathbf{x}) \leq \beta(\mathbf{x})$ for all $\mathbf{x} \in R$.

Theorem 195 *A normal domain $E \subset \mathbb{R}^{N+1}$ is Peano–Jordan measurable and*

$$\operatorname{meas}_{N+1} E = \int_R \left(\int_{\alpha(\mathbf{x})}^{\beta(\mathbf{x})} 1 \, dy \right) d\mathbf{x}.$$

Theorem 196 *Let $E \subset \mathbb{R}^{N+1}$ be a normal domain and let $f : E \to \mathbb{R}$ be a bounded continuous function. Then f is Riemann integrable over E and*

$$\int_E f(\mathbf{x}, y) \, d(\mathbf{x}, y) = \int_R \left(\int_{\alpha(\mathbf{x})}^{\beta(\mathbf{x})} f(\mathbf{x}, y) \, dy \right) d\mathbf{x}.$$

Example 197 *Calculate the volume of the set $E \subset \mathbb{R}^3$ bounded by $z = \sqrt{x^2 + y^2}$ and $x^2 + y^2 + z^2 - z = 0$. Note that the first is a cone, while the second can be written as $x^2 + y^2 + \left(z - \frac{1}{2}\right)^2 = \frac{1}{4}$ and so it is a sphere of center $\left(0, 0, \frac{1}{2}\right)$ and radius $\frac{1}{2}$. To see where these surfaces intersect, let's solve the system*

$$\begin{cases} z = \sqrt{x^2 + y^2} \\ x^2 + y^2 + z^2 - z = 0 \end{cases} \Leftrightarrow \begin{cases} z^2 = x^2 + y^2 \\ x^2 + y^2 + z^2 - z = 0 \end{cases} \Leftrightarrow \begin{cases} z^2 = x^2 + y^2 \\ z^2 + z^2 - z = 0 \end{cases}$$
$$\Leftrightarrow \begin{cases} z^2 = x^2 + y^2 \\ z(2z - 1) = 0 \end{cases} \Leftrightarrow \begin{cases} z^2 = x^2 + y^2 \\ z = 0 \text{ or } z = \frac{1}{2} \end{cases}$$

and so they intersect at the origin or at the plane $z = \frac{1}{2}$. Note that E is a normal domain, since we can write it as

$$E = \left\{ (x, y, z) \in \mathbb{R}^3 : \, \sqrt{x^2 + y^2} \leq z \leq \frac{1}{2} + \sqrt{\frac{1}{4} - (x^2 + y^2)}, \, x^2 + y^2 \leq \frac{1}{4} \right\}.$$

We have

$$\operatorname{meas}_3 E = \int_{[-1,1] \times [-1,1] \times [0,1]} \chi_E(\mathbf{x}) \, d\mathbf{x}.$$

Let's apply Theorem 188,

$$\text{meas}_3\, E = \int_{[0,1]} \left(\int_{[-1,1]\times[-1,1]} \chi_E\left((x,y,z)\right)\, d(x,y) \right) dz.$$

So we are fixing $z_0 \in [0,1]$ and we are cutting the set E with the plane $z = z_0$ and calculating in $\mathbb{R}^2$ the area of the section

$$\begin{aligned} E_{z_0} &= \left\{ (x,y) \in \mathbb{R}^2 : (x,y,z_0) \in E \right\} \\ &= \begin{cases} \left\{ (x,y) \in \mathbb{R}^2 : x^2 + y^2 \le z_0 - z_0^2 \right\} & \text{if } z_0 \in \left[\frac{1}{2}, 1\right], \\ \left\{ (x,y) \in \mathbb{R}^2 : x^2 + y^2 \le z_0^2 \right\} & \text{if } z_0 \in \left[0, \frac{1}{2}\right]. \end{cases} \end{aligned}$$

Hence,

$$\begin{aligned} \text{meas}_3\, E &= \int_{\left[0,\frac{1}{2}\right]} \left(\int_{[-1,1]\times[-1,1]} \chi_E(x,y,z)\, d(x,y) \right) dz \\ &\quad + \int_{\left[\frac{1}{2},1\right]} \left(\int_{[-1,1]\times[-1,1]} \chi_E(x,y,z)\, d(x,y) \right) dz \\ &= \int_{\left[0,\frac{1}{2}\right]} \text{meas}_2\, E_z\, dz + \int_{\left[\frac{1}{2},1\right]} \text{meas}_2\, E_z\, dz \\ &= \int_{\left[0,\frac{1}{2}\right]} \pi z^2\, dz + \int_{\left[\frac{1}{2},1\right]} \pi\left(z - z^2\right) dz. \end{aligned}$$

Another possibility would have been to use

$$\text{meas}_3\, E = \int_{[-1,1]\times[-1,1]} \left(\int_{[0,1]} \chi_E(x,y,z)\, dz \right) d(x,y).$$

Mondy, April 15, 2013

Exercise 198 *Calculate the integral*

$$\int_E \log(xy)\, dxdy,$$

where E is the set of $\mathbb{R}^2$ bounded by the curves $xy = 1$, $x = y^2$, and $x = 3$.

Exercise 199 *Calculate the integral*

$$\int_E x \log\left(1 + y^2\right) dxdy,$$

where $E = \left\{ (x,y) \in \mathbb{R}^2 : \frac{x^2}{4} + \frac{y^2}{2} \le 1,\ x^2 - y^2 = 1,\ x \ge 0 \right\}$.

Exercise 200 *Find the volume of the set E of $\mathbb{R}^3$ enclosed by $z = \frac{1}{3}\left(x^2 + y^2\right)$ and $z = \sqrt{4 - x^2 - y^2}$.*

Next we discuss some properties of Riemann integration.

Proposition 201 *Given a rectangle R, let $f, g : R \to \mathbb{R}$ be Riemann integrable.*

(i) If $\lambda \in \mathbb{R}$, then λf is Riemann integrable and

$$\int_R \lambda f(\mathbf{x})\, d\mathbf{x} = \lambda \int_R f(\mathbf{x})\, d\mathbf{x}. \tag{18}$$

(ii) The functions $f + g$ and fg are Riemann integrable and

$$\int_R (f(\mathbf{x}) + g(\mathbf{x}))\, d\mathbf{x} = \int_R f(\mathbf{x})\, d\mathbf{x} + \int_R g(\mathbf{x})\, d\mathbf{x}. \tag{19}$$

(iii) If $f \leq g$, then

$$\int_R f(\mathbf{x})\, d\mathbf{x} \leq \int_R g(\mathbf{x})\, d\mathbf{x}.$$

(iv) The function $|f|$ is Riemann integrable and

$$\left| \int_R f(\mathbf{x})\, d\mathbf{x} \right| \leq \int_R |f(\mathbf{x})|\, d\mathbf{x}.$$

Exercise 202 *Give an example of a bounded function $f : R \to \mathbb{R}$ such that $|f|$ is Riemann integrable over R, but f is not.*

Remark 203 *If $E, F \subset \mathbb{R}^N$ are two Peano–Jordan measurable sets, with $E \subseteq F$, then $\chi_E \leq \chi_F$, and so by Proposition 201,*

$$\operatorname{meas} E \leq \operatorname{meas} F.$$

Exercise 204 *Prove that if $E \subset \mathbb{R}^N$ and $F \subset \mathbb{R}^N$ are Peano–Jordan measurable and R is a rectangle containing E, then $E \cup F$, $E \cap F$, $R \setminus E$ are Peano–Jordan measurable.*

Exercise 205 *Prove that if $E_1, \ldots, E_n \subset \mathbb{R}^N$ are Peano–Jordan measurable and pairwise disjoint, then*

$$\operatorname{meas} \left(\bigcup_{i=1}^n E_i \right) = \sum_{i=1}^n \operatorname{meas} E_i.$$

Exercise 206 *Prove that if $E_1, \ldots, E_n \subset \mathbb{R}^N$ are Peano–Jordan measurable and $E \subseteq \bigcup_{i=1}^n E_i$ is Peano–Jordan measurable, then*

$$\operatorname{meas} E \leq \sum_{i=1}^n \operatorname{meas} E_i.$$

Theorem 207 *Given a bounded set $E \subset \mathbb{R}^N$ and a Riemann integrable function $f : E \to \mathbb{R}$, if $F \subset E$ is Peano–Jordan measurable, then f is integrable over F.*

Theorem 208 *If $E \subset \mathbb{R}^N$ is given by the disjoint union of a finite number of Peano–Jordan measurable,*

$$E = \bigcup_{n=1}^{N} E_n,$$

and if $f : E \to \mathbb{R}$ is f is Riemann integrable over E, then

$$\int_E f\, d\mathbf{x} = \sum_{n=1}^{N} \int_{E_n} f\, d\mathbf{x}.$$

Corollary 209 *If $E \subset \mathbb{R}^N$ is given by the disjoint union of a finite number of normal domains,*

$$E = \bigcup_{n=1}^{N} E_n,$$

then

$$\operatorname{meas}_N E = \sum_{n=1}^{N} \operatorname{meas}_N E_n$$

and if $f : E \to \mathbb{R}$ is a bounded continuous function, then f is Riemann integrable over E and

$$\int_E f\, d\mathbf{x} = \sum_{n=1}^{N} \int_{E_n} f\, d\mathbf{x}.$$

Theorem 210 *Let $E \subset \mathbb{R}^N$ be symmetric with respect to the hyperplane $x_i = 0$, that is,*

$$x = (x_1, \ldots, x_i, \ldots, x_N) \in E \Rightarrow x = (x_1, \ldots, -x_i, \ldots, x_N) \in E$$

and let $f : E \to \mathbb{R}$ is Riemann integrable over E. If f is odd with respect to x_i, that is,

$$f(x_1, \ldots, -x_i, \ldots, x_N) = -f(x_1, \ldots, x_i, \ldots, x_N),$$

then

$$\int_E f\, d\mathbf{x} = 0,$$

while if is even with respect to x_i, that is,

$$f(x_1, \ldots, -x_i, \ldots, x_N) = f(x_1, \ldots, x_i, \ldots, x_N),$$

then

$$\int_E f\, d\mathbf{x} = 2\int_{E_1} f\, d\mathbf{x},$$

where $E_1 = \{\mathbf{x} \in E : \, x_i \geq 0\}$.

Example 211 *Let's compute the area of a circle of radius r centered at $(0,0)$. We have*

$$E = \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le r^2\right\}$$

and so

$$\operatorname{meas}_2 E = \int_E 1 \, dxdy.$$

Since $f = 1$ is even in x and in y and since E is symmetric with respect to both axes, we have that

$$\operatorname{meas}_2 E = 4 \int_{E_1} 1 \, dxdy,$$

where

$$\begin{aligned} E_1 &= \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le r^2, \, x \ge 0, \, y \ge 0\right\} \\ &= \left\{(x,y) \in \mathbb{R}^2 : 0 \le y \le \sqrt{r^2 - x^2}, \, 0 \le x \le r\right\}. \end{aligned}$$

Since E_1 is normal,

$$\begin{aligned} \operatorname{meas}_2 E &= 4 \int_0^r \left(\int_0^{\sqrt{r^2-x^2}} 1 \, dy \right) dx = 4 \int_0^r [y]_{y=0}^{y=\sqrt{r^2-x^2}} \, dx \\ &= 4 \int_0^r \left[\sqrt{r^2 - x^2} - 0\right] dx. \end{aligned}$$

Using the change of variables $x = r\cos\theta$, where $0 \le \theta \le \frac{\pi}{2}$, we get $dx = -r\sin\theta \, d\theta$ and $\sqrt{r^2 - x^2} = \sqrt{r^2 - r^2\cos^2\theta} = \sqrt{r^2\sin^2\theta} = r\sin\theta$, so that

$$\begin{aligned} \operatorname{meas}_2 E &= 4 \int_0^r \left[\sqrt{r^2 - x^2} - 0\right] dx = -4 \int_0^{\pi/2} \sqrt{r^2 - r^2\cos^2\theta} \, (-r\sin\theta) \, d\theta \\ &= 4r^2 \int_0^{\pi/2} \sin^2\theta \, d\theta = 4r^2 \frac{1}{4}\pi. \end{aligned}$$

16 Change of Variables

Theorem 212 *Let $U \subseteq \mathbb{R}^N$ be an open set and let $\mathbf{g} : U \to \mathbb{R}^M$ be a function of class C^1 , with $N \le M$. Let $E \subset \mathbb{R}^N$ be a Peano–Jordan measurable set with $\overline{E} \subseteq U$. Moreover, if $N = M$, assume that E has measure zero. Then $\mathbf{g}(E)$ is Peano–Jordan measurable with measure zero.*

Theorem 213 (Change of variables for multiple integrals) *Let $U \subseteq \mathbb{R}^N$ be an open set and let $\mathbf{g} : U \to \mathbb{R}^N$ be a one-to-one function of class C^1 such that $\det J_{\mathbf{g}}(\mathbf{x}) \ne 0$ for all $\mathbf{x} \in U$. Let $E \subset \mathbb{R}^N$ be Peano–Jordan measurable with $\overline{E} \subseteq U$ and let $f : \mathbf{g}(E) \to \mathbb{R}$ be Riemann integrable. Then the function $\mathbf{x} \in E \mapsto f(\mathbf{g}(\mathbf{x})) |\det J_{\mathbf{g}}(\mathbf{x})|$ is Riemann integrable and*

$$\int_{\mathbf{g}(E)} f(\mathbf{y}) \, d\mathbf{y} = \int_E f(\mathbf{g}(\mathbf{x})) \left|\det J_{\mathbf{g}}(\mathbf{x})\right| \, d\mathbf{x}.$$

Example 214 *Let's calculate the integral*

$$\iint_F (x+y)\ dxdy,$$

where

$$F := \left\{(x,y) \in \mathbb{R}^2 :\ 0 < x < y < 2x,\ 1 < xy < 2\right\}.$$

Consider the change of variables $u = xy$ *and* $v = \frac{y}{x}$*. Let's solve for* x *and* y*. We have* $x = \sqrt{\frac{u}{v}}$ *and* $y = \sqrt{uv}$*. Define* $\mathbf{g} : U \to \mathbb{R}^2$ *as follows*

$$\mathbf{g}(u,v) := \left(\sqrt{\frac{u}{v}}, \sqrt{uv}\right),$$

where

$$U := \left\{(u,v) \in \mathbb{R}^2 :\ u > 0,\ v > 0\right\}.$$

Then $\mathbf{g}$ *is one-to-one and its inverse is given by* $\mathbf{g}^{-1}(x,y) := \left(xy, \frac{y}{x}\right)$*. (Note that in general it is much simpler to check that* $\mathbf{g}$ *is one-to-one than finding the inverse). Moreover,*

$$\begin{aligned}\det J_{\mathbf{g}}(u,v) &= \det\begin{pmatrix} \frac{\partial g_1}{\partial u}(u,v) & \frac{\partial g_1}{\partial v}(u,v) \\ \frac{\partial g_2}{\partial u}(u,v) & \frac{\partial g_2}{\partial v}(u,v)\end{pmatrix} \\ &= \det\begin{pmatrix} \frac{1}{2\sqrt{uv}} & -\frac{\sqrt{u}}{2v^{3/2}} \\ \frac{\sqrt{v}}{2\sqrt{u}} & \frac{\sqrt{u}}{2\sqrt{v}}\end{pmatrix} = \frac{1}{2v}.\end{aligned}$$

Set $E = (1,2) \times (1,2)$ *and note* $\mathbf{g}(E) = F$*. By the change of variables theorem,*

$$\begin{aligned}\iint_F (x+y)\ dxdy &= \iint_E (g_1(u,v) + g_2(u,v))\left|\det J_{\mathbf{g}}(u,v)\right|\ dudv \\ &= \iint_E \left(\sqrt{\frac{u}{v}} + \sqrt{uv}\right)\frac{1}{2v}\, dudv \\ &= \frac{1}{2}\iint_E \left(\frac{\sqrt{u}}{v^{3/2}} + \frac{\sqrt{u}}{v}\right)\ dudv.\end{aligned}$$

It follows that

$$\begin{aligned}\iint_F (x+y)\ dxdy &= \frac{1}{2}\int_1^2 \left(\int_1^2 \left(\frac{\sqrt{u}}{v^{3/2}} + \frac{\sqrt{u}}{v}\right)\ dv\right)\ du = \frac{1}{2}\int_1^2 \left(\left[-2\frac{\sqrt{u}}{v^{1/2}} + \sqrt{u}\log v\right]_{v=1}^{v=2}\right)\ du \\ &= \frac{1}{2}\int_1^2 \sqrt{u}\log 2\, du = \frac{1}{3}(\log 2)\left(2\sqrt{2} - 1\right).\end{aligned}$$

Wednesday, April 17, 2013

The previous theorem unfortunately does not work for polar coordinates. The problem is that polar coordinates are not one-to-one.

Corollary 215 *Let $E \subset \mathbb{R}^N$ be a Peano–Jordan measurable set, let $\mathbf{g} : E \to \mathbb{R}^N$ be such that there exist $\frac{\partial g_i}{\partial x_j}$ in E for all $i, j = 1, \ldots, N$, and that they are bounded and continuous. Assume that there exists a Peano–Jordan measurable set $E_1 \subset E$ such that $\operatorname{meas}(E_1) = \operatorname{meas}(\mathbf{g}(E_1)) = 0$, $E \setminus E_1$ is open, $\det J_{\mathbf{g}}(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in E \setminus E_1$ and $\mathbf{g} : E \setminus E_1 \to \mathbb{R}^N$ is one-to-one. Let $f : \mathbf{g}(E) \to \mathbb{R}$ be Riemann integrable. Then the function $\mathbf{x} \in E \mapsto f(\mathbf{g}(\mathbf{x})) |\det J_{\mathbf{g}}(\mathbf{x})|$ is Riemann integrable and*

$$\int_{\mathbf{g}(E)} f(\mathbf{y})\, d\mathbf{y} = \int_E f(\mathbf{g}(\mathbf{x})) |\det J_{\mathbf{g}}(\mathbf{x})|\, d\mathbf{x}.$$

The most important applications of the previous corollary is given by polar coordinates in $\mathbb{R}^2$. Let $P = (x, y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, let $O = (0,0)$, and let θ be the angle that the OP forms with the positive x axis. Then $0 \leq \theta \leq 2\pi$ and $x = r\cos\theta$ and $y = r\sin\theta$, where $r = \sqrt{x^2 + y^2}$. Hence,

$$\mathbf{g}(r, \theta) := (r\cos\theta, r\sin\theta), \quad (r, \theta) \in [0, \infty) \times [0, 2\pi).$$

Note that partial derivatives of $\mathbf{g}$ of any order exist in $\mathbb{R}^2$ and are continuous. The Jacobian of $\mathbf{g}$ is given by

$$\begin{aligned} \det J_{\mathbf{g}}(r, \theta) &= \det \begin{pmatrix} \frac{\partial g_1}{\partial r}(r, \theta) & \frac{\partial g_1}{\partial \theta}(r, \theta) \\ \frac{\partial g_2}{\partial r}(r, \theta) & \frac{\partial g_2}{\partial \theta}(r, \theta) \end{pmatrix} \\ &= \det \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r\cos^2\theta + r\sin^2\theta = r, \end{aligned}$$

which is zero at $r = 0$. We observe that $\mathbf{g}$ is onto but $\mathbf{g}$ is one-to-one only in the set where $(0, \infty) \times [0, 2\pi)$. However, this is not open. Consider instead, the open sets $U := (0, \infty) \times (0, 2\pi)$ and $V := \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}$ and let's prove that $\mathbf{g} : U \to V$ is invertible. Note that we cannot apply Theorem 213 to $\mathbf{g}$ but we can apply Corollary 215, since given a Peano–Jordan measurable set $E \subset [0, \infty) \times [0, 2\pi)$ the "bad" set E_1 is given by the intersection of E with the two lines $r = 0$ and $\theta = 0$, precisely,

$$E_1 := E \setminus (\{(r, 0) : r \geq 0\} \cup \{(0, \theta) : \theta \in [0, 2\pi)\}).$$

Moreover $\mathbf{g}(E_1)$ is a bounded subset of the x-axis. Hence, $\operatorname{meas}(E_1) = \operatorname{meas}(\mathbf{g}(E_1)) = 0$.

Example 216 *Let's calculate the integral*

$$\iint_F \frac{\sqrt{x^2 + y^2}}{x}\, dxdy,$$

where

$$F := \left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 > 1,\ x^2 + y^2 - 2x < 0\right\}.$$

Note that the domain is symmetric with respect to the x-axis and f is even in y. Hence,

$$\iint_F \frac{\sqrt{x^2+y^2}}{x}\, dxdy = 2\iint_{F_1} \frac{\sqrt{x^2+y^2}}{x}\, dxdy,$$

where

$$F_1 = \left\{(x,y) \in \mathbb{R}^2 : x^2+y^2 > 1,\ x^2+y^2-2x < 0,\ y \geq 0\right\}.$$

Using polar coordinates we have $r^2\cos^2\theta + r^2\sin^2\theta > 1$, that is, $r > 1$, and $r^2\cos^2\theta + r^2\sin^2\theta - 2r\cos\theta < 0$, that is, $r(r-2\cos\theta) < 0$, which gives $r < 2\cos\theta$. So $1 < r < 2\cos\theta$ and $0 \leq \theta \leq \frac{\pi}{2}$. However, for the first inequality to make sense we need $1 < 2\cos\theta$, that is, $\frac{1}{2} < \cos\theta$, so $0 \leq \theta < \frac{\pi}{3}$. It follows that F_1 is the image of the set

$$E_1 = \left\{(r,\theta) \in \mathbb{R}^2 : 0 \leq \theta < \frac{\pi}{3},\ 1 < r < 2\cos\theta\right\},$$

which is a normal domain. Using polar coordinates, we get

$$\begin{aligned}
2\iint_{F_1} \frac{\sqrt{x^2+y^2}}{x}\, dxdy &= 2\iint_{E_1} \frac{\sqrt{r^2\cos^2\theta + r^2\sin^2\theta}}{r\cos\theta}\, |r|\, d\theta dr \\
&= 2\int_0^{\pi/3}\left(\int_1^{2\cos\theta} \frac{r}{\cos\theta}\, dr\right) d\theta = \int_0^{\pi/3} \frac{1}{\cos\theta}\left[r^2\right]_{r=1}^{r=2\cos\theta} d\theta \\
&= \int_0^{\pi/3} \frac{1}{\cos\theta}\left[4\cos^2\theta - 1\right] d\theta = \int_0^{\pi/3}\left(4\cos\theta - \frac{1}{\cos\theta}\right) d\theta.
\end{aligned}$$

For the second integral

$$\int_0^{\pi/3} \frac{1}{\cos\theta}\, d\theta = \int_0^{\pi/3} \frac{\cos\theta}{\cos^2\theta}\, d\theta = \int_0^{\pi/3} \frac{\cos\theta}{1-\sin^2\theta}\, d\theta$$

and we can now make the change of variables $t = \sin\theta$, so that $dt = \cos\theta\, dt$. It follows that

$$\begin{aligned}
\int_0^{\pi/3} \frac{\cos\theta}{1-\sin^2\theta}\, d\theta &= \int_0^{\sqrt{3}/2} \frac{1}{1-t^2}\, dt \\
&= \frac{1}{2}\int_0^{\sqrt{3}/2} \frac{1}{1-t} + \frac{1}{1+t}\, dt \\
&= \frac{1}{2}\left[\log|1-t| + \log|1+t|\right]_{t=0}^{t=\sqrt{3}/2} \\
&= \frac{1}{2}\log\left(1-\frac{1}{2}\sqrt{3}\right) + \frac{1}{2}\log\left(\frac{1}{2}\sqrt{3}+1\right).
\end{aligned}$$

Now let's introduce spherical coordinates in $\mathbb{R}^3$. Let $P = (x,y,z) \in \mathbb{R}^3 \setminus \{(0,0,0)\}$, let $O = (0,0,0)$, and let θ be the angle that the OP forms with the positive z axis. Then $0 \le \theta \le 2\pi$ and $z = r\cos\theta$, where $r = \sqrt{x^2+y^2+z^2}$. Project P in the plane $z = 0$ to get a point Q and let φ be the angle that the OQ forms with the positive x-axis. Then $x = r\sin\theta\cos\varphi$ and $y = r\sin\theta\sin\varphi$, where $0 \le \varphi \le 2\pi$. Then

$$\mathbf{g}(r,\theta,\varphi) := (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta), \quad (r,\theta,\varphi) \in [0,\infty)\times[0,2\pi]\times[0,2\pi].$$

Note that partial derivatives of $\mathbf{g}$ of any order exist in $\mathbb{R}^3$ and are continuous. The Jacobian of $\mathbf{g}$ is given by

$$\begin{aligned}\det J_{\mathbf{g}}(r,\theta,\varphi) &= \det\begin{pmatrix} \frac{\partial g_1}{\partial r}(r,\theta,\varphi) & \frac{\partial g_1}{\partial \theta}(r,\theta,\varphi) & \frac{\partial g_1}{\partial \varphi}(r,\theta,\varphi) \\ \frac{\partial g_2}{\partial r}(r,\theta,\varphi) & \frac{\partial g_2}{\partial \theta}(r,\theta,\varphi) & \frac{\partial g_2}{\partial \varphi}(r,\theta,\varphi) \\ \frac{\partial g_3}{\partial r}(r,\theta,\varphi) & \frac{\partial g_3}{\partial \theta}(r,\theta,\varphi) & \frac{\partial g_3}{\partial \varphi}(r,\theta,\varphi) \end{pmatrix} \\ &= \det\begin{pmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix} = r^2\sin\theta > 0\end{aligned}$$

for $r > 0$ and, say, $0 < \theta < \pi$.

Example 217 *Let's calculate the integral*

$$\iiint_F \frac{1}{\sqrt{x^2+y^2+z^2}}\,dxdydz,$$

where

$$F := \left\{(x,y,z) \in \mathbb{R}^3 : 1 \le x^2+y^2+z^2 < 4,\ z \ge 0\right\}.$$

Since $z \ge 0$, we have $\theta \in \left(0, \frac{\pi}{2}\right)$, while $1 < r < 2$, $0 \le \varphi \le 2\pi$. Hence, F is the image of the set

$$E = \left\{(r,\theta,\varphi) \in \mathbb{R}^3 : 1 \le r < 2,\ \theta \in \left(0, \frac{\pi}{2}\right),\ 0 \le \varphi \le 2\pi\right\}$$

and so

$$\begin{aligned}\iiint_F \frac{1}{\sqrt{x^2+y^2+z^2}}\,dxdydz &= \iiint_E \frac{1}{r}\left|r^2\sin\theta\right|\,drd\theta d\varphi \\ &= \left(\int_1^2 r\,dr\right)\left(\int_0^{\pi/2}\sin\theta\,d\theta\right)\left(\int_0^{2\pi} 1\,d\varphi\right) \\ &= \left[\frac{r^2}{2}\right]_{r=1}^{r=2}\left[-\cos\theta\right]_{\theta=0}^{\theta=\pi/2} 2\pi = \left(1 - \frac{1}{2}\right)(0-1)\,2\pi.\end{aligned}$$

Other important coordinates are cylindrical coordinates in $\mathbb{R}^3$. We project $P = (x,y,z)$ into the plane $z = 0$ and in the plane xy we use polar coordinates. Then

$$\mathbf{g}(\rho,\varphi,z) := (\rho\cos\varphi, \rho\sin\varphi, z), \quad (\rho,\varphi,z) \in [0,\infty)\times[0,2\pi]\times\mathbb{R},$$

where $\rho = \sqrt{x^2+y^2}$. Note that the partial derivatives of $\mathbf{g}$ of any order exist in $\mathbb{R}^3$ and are continuous. The Jacobian of $\mathbf{g}$ is given by

$$\det J_{\mathbf{g}}(\rho,\varphi,z) = \det\begin{pmatrix} \frac{\partial g_1}{\partial \rho}(\rho,\varphi,z) & \frac{\partial g_1}{\partial \varphi}(\rho,\varphi,z) & \frac{\partial g_1}{\partial z}(\rho,\varphi,z) \\ \frac{\partial g_2}{\partial \rho}(\rho,\varphi,z) & \frac{\partial g_2}{\partial \varphi}(\rho,\varphi,z) & \frac{\partial g_2}{\partial z}(\rho,\varphi,z) \\ \frac{\partial g_3}{\partial \rho}(\rho,\varphi,z) & \frac{\partial g_3}{\partial \varphi}(\rho,\varphi,z) & \frac{\partial g_3}{\partial z}(\rho,\varphi,z) \end{pmatrix}$$
$$= \det\begin{pmatrix} \cos\varphi & -\rho\sin\varphi & 0 \\ \sin\varphi & \rho\cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \rho > 0$$

for $\rho > 0$.

Example 218 *Let's calculate the volume of the bounded set F included between the cone $z = \sqrt{x^2+y^2}$ and the paraboloid $z = x^2+y^2$. Hence,*

$$F := \left\{(x,y,z)\in\mathbb{R}^3 : x^2+y^2 \le z \le \sqrt{x^2+y^2},\, x^2+y^2 \le 1\right\}.$$

We have $0\le\varphi\le 2\pi$, $\rho^2 \le z \le \rho$, $0 \le \rho \le 1$. Hence, F is the image of the set

$$E = \left\{(\rho,\varphi,z)\in\mathbb{R}^3 : \rho^2 \le z \le \rho,\, 0\le\rho\le 1,\, 0\le\varphi\le 2\pi\right\}.$$

Note that E is a normal domain. Hence,

$$\begin{aligned}\iiint_F 1\, dxdydz &= \iiint_E 1\,|\rho|\, d\rho d\varphi dz \\ &= \left(\int_0^1 \left(\int_{\rho^2}^{\rho} \rho\, dz\right) d\rho\right)\left(\int_0^{2\pi} 1\, d\varphi\right) \\ &= 2\pi \int_0^1 \rho\left(\int_{\rho^2}^{\rho} dz\right) d\rho = 2\pi\int_0^1 \rho\left([z]_{z=\rho^2}^{z=\rho}\right) d\rho \\ &= 2\pi\int_0^1 \rho\left(\rho-\rho^2\right) d\rho = 2\pi\left[\frac{\rho^3}{3}-\frac{\rho^4}{4}\right]_{\rho=0}^{\rho=1} \\ &= \frac{1}{6}\pi.\end{aligned}$$

Friday, April 19, 2013

Carnival, no classes

Monday, April 22, 2013

17 Differential Surfaces

Definition 219 *Given $1 \le k < N$, a nonempty set $M \subseteq \mathbb{R}^N$ is called a k-*dimensional differential surface *or* manifold *if for every $\mathbf{x}_0 \in M$ there exist an open set U containing $\mathbf{x}_0$ and a differentiable function $\varphi : W \to \mathbb{R}^N$, where $W \subseteq \mathbb{R}^k$ is an open set such that* *(i)* $\varphi : W \to M \cap U$ *is a homeomorphism, that is, it is invertible and continuous together with its inverse* $\varphi^{-1} : M \cap U \to W$*,*

(ii) $J_\varphi(\mathbf{y})$ *has maximum rank* k *for all* $\mathbf{y} \in W$*.*

The function φ *is called a* local chart *or a* system of local coordinates *or a* local parametrization *around* $\mathbf{x}_0$*. We say that* M *is of class* C^m*,* $m \in \mathbb{N}$*, (respectively,* C^∞*) if all local charts are of class* C^m *(respectively,* C^∞*).*

Roughly speaking a set $M \subset \mathbb{R}^N$ is a k-dimensional differential surface if for every point $\mathbf{x}_0 \in M$ we can "cut" a piece of M around $\mathbf{x}_0$ and deform it/flatten it in a smooth way to get, say, a ball of $\mathbb{R}^k$. Another way to say this is that locally M looks like $\mathbb{R}^k$. Thus the range of a simple differentiable curve is a 1-dimensional surface since locally it looks like $\mathbb{R}$, while the range of a self-intersecting curve is not, because at self-intersection point the range of the curve does not look like $\mathbb{R}$. Similarly, a sphere in $\mathbb{R}^3$ is a 2-dimensional surface because locally it looks like $\mathbb{R}^2$, while a cone is not because near the tip it does not look like $\mathbb{R}^2$.

A simple way to construct k-dimensional differential surface is to start with a set of $\mathbb{R}^k$ and then deform it in a smooth way.

Remark 220 *The difference between curves and surfaces, it's that a surface is a set of* $\mathbb{R}^N$*, while a curve is an equivalence class of functions. Also for surfaces we do not allow self-intersections. Also, a curve can be described by only one parametrization, while a surface usually needs more than one.*

Remark 221 *Note that Theorem 212 implies that if a bounded* k*-dimensional differential surface is Peano–Jordan measurable, then its measure must be zero.*

Example 222 *Consider the hyperbola*

$$M := \left\{(x,y) \in \mathbb{R}^2 : x^2 - y^2 = 1\right\}.$$

Note that the set M *is not the range of a curve (why?). To cover* M *we need at least two local charts, precisely, we can take the open sets*

$$V := \left\{(x,y) \in \mathbb{R}^2 : x > 0\right\}, \quad W := \left\{(x,y) \in \mathbb{R}^2 : x < 0\right\},$$

and the functions $\varphi : \mathbb{R} \to M \cap V$ *and* $\psi : \mathbb{R} \to M \cap W$ *defined by*

$$\varphi(t) := \left(\sqrt{1+t^2}, t\right), \quad \psi(t) := \left(-\sqrt{1+t^2}, t\right), \quad t \in \mathbb{R}.$$

Note that both φ *and* ψ *are of class* C^∞ *(the argument inside the square roots is never zero). Moreover,* $\varphi'(t) = \left(\frac{t}{\sqrt{1+t^2}}, 1\right)$ *and* $\psi'(t) := \left(-\frac{t}{\sqrt{1+t^2}}, 1\right)$*, and so the rank of* $J_\varphi(t)$ *and of* $J_\psi(t)$ *is one. Finally,* $\varphi^{-1} : M \cap V \to \mathbb{R}$ *and* $\psi^{-1} : M \cap W \to \mathbb{R}$ *are given by*

$$\varphi^{-1}(x,y) = y, \quad \psi^{-1}(x,y) = y,$$

which are continuous. Thus, M *is a 1-dimensional surface of class* C^∞*.*

Example 223 *Given an open set $V \subseteq \mathbb{R}^k$ and a differential function $f : V \to \mathbb{R}$, consider the graph of f,*

$$\operatorname{Gr} h := \{(\mathbf{y}, t) \in V \times \mathbb{R} : \, t = f(\mathbf{y})\}.$$

A chart is given by the function $\varphi : V \to \mathbb{R}^{k+1}$ given by $\varphi(\mathbf{y}) := (\mathbf{y}, f(\mathbf{y}))$. Then,

$$J_\varphi(\mathbf{y}) = \begin{pmatrix} I_k \\ \nabla f(\mathbf{y}) \end{pmatrix},$$

which has rank N. Note that φ is one-to-one and that $\varphi(V) = \operatorname{Gr} h$. Hence, there exists $\varphi^{-1} : \operatorname{Gr} f \to V$. Moreover, φ^{-1} is continuous, since the projection

$$\begin{aligned} \Pi : \mathbb{R}^{k+1} &\to \mathbb{R}^k \\ (\mathbf{y}, t) &\mapsto \mathbf{y} \end{aligned}$$

is of class C^∞ and φ^{-1} is given by the restriction of Π to $\operatorname{Gr} f$. Hence, $\operatorname{Gr} f$ is an k-dimensional differential surface.

The next proposition gives an equivalent definition of surfaces, which is very useful for examples.

Proposition 224 *Given $1 \leq k < N$, a nonempty set $M \subseteq \mathbb{R}^N$,and $m \in \mathbb{N}$, then the following are equivalent:*

(i) M is a k-dimensional surface of class C^m.

(ii) For every $\mathbf{x}_0 \in M$ there exist an open set $U \subseteq \mathbb{R}^N$ containing $\mathbf{x}_0$ and a function $\mathbf{g} : U \to \mathbb{R}^{N-k}$ of class C^m, such that

$$M \cap U = \{\mathbf{x} \in U : \, \mathbf{g}(\mathbf{x}) = \mathbf{0}\} \tag{20}$$

and $J_\mathbf{g}(\mathbf{x})$ has maximum rank $N - k$ for all $\mathbf{x} \in M \cap U$.

Wednesday, April 24, 2013

Example 225 (Torus) *A* torus *is a 2-dimensional surface M obtained from a rectangle of $\mathbb{R}^2$ by identifying opposite sides. Consider the chart $\varphi : (0, 2\pi) \times (0, 2\pi) \to \mathbb{R}^2$ given by*

$$\varphi(u, v) := ((r \cos u + a) \cos v, (r \cos u + a) \sin v, r \sin u), \tag{21}$$

where $a > 0$ and $r > 0$. Define $M := \varphi((0, 2\pi) \times (0, 2\pi)) \subset \mathbb{R}^3$.
First method. *Let's prove that M is a 2-dimensional surface. Note that φ is of class C^1 and*

$$J_\varphi(u, v) = \begin{pmatrix} -r \sin u \cos v & -(r \cos u + a) \sin v \\ -r \sin u \sin v & (r \cos u + a) \cos v \\ r \cos u & 0 \end{pmatrix}.$$

Observe that if $r\cos u + a = 0$, then

$$J_\varphi(u,v) = \begin{pmatrix} -r\sin u\cos v & 0 \\ -r\sin u\sin v & 0 \\ r\cos u & 0 \end{pmatrix}$$

and so $J_\varphi(u,v)$ cannot have rank 2. Thus, for M to be a 2-dimensional surface we need to assume that $r\cos u + a \neq 0$ for all $u \in (0,2\pi)$. In what follows we assume that

$$a > r.$$

The submatrix $\begin{pmatrix} -r\sin u\cos v & -(r\cos u + a)\sin v \\ -r\sin u\sin v & (r\cos u + a)\cos v \end{pmatrix}$ has determinant $-r\sin u\,(r\cos u + a)$. Then $r\cos u + a > 0$. If $\sin u \neq 0$, then the determinant is different from zero and so so $J_\varphi(u,v)$ has rank 2. On the other hand, when $\sin u = 0$, that is, when $u = \pi$, then

$$J_\varphi(\pi,v) = \begin{pmatrix} 0 & -(-r+a)\sin v \\ 0 & (-r+a)\cos v \\ -r & 0 \end{pmatrix}.$$

For any $v \in (0,2\pi)$, either $\cos v \neq 0$ or $\sin v \neq 0$ (or both) and so either the submatrix $\begin{pmatrix} 0 & (-r+a)\cos v \\ -r & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & -(-r+a)\sin v \\ -r & 0 \end{pmatrix}$ have determinant different from zero. Thus, we have shown that when $a > r$, $J_\varphi(u,v)$ has rank 2 for all $(u,v) \in (0,2\pi)\times(0,2\pi)$.

To see that φ is one-to-one, consider $(x,y,z) \in M$. We want to find a unique pair (u,v) such that $\varphi(u,v) = (x,y,z)$. Note that to find u we can use $\sin u = \frac{z}{r}$. The problem is that there are two values of u, one in $\left(0,\frac{\pi}{2}\right)\cup\left(\frac{3\pi}{2},2\pi\right)$ and one in $\left[\frac{\pi}{2},\frac{3\pi}{2}\right]$. If $x^2+y^2 \leq a^2$, then

$$x^2+y^2 = (r\cos u + a)^2\cos^2 v + (r\cos u + a)^2\sin^2 v = (r\cos u + a)^2 \leq a^2,$$

that is $\cos u\,(r\cos u + 2a) \leq 0$, so that $\cos u \leq 0$, that is, $\frac{\pi}{2} \leq u \leq \frac{3\pi}{2}$. So in this case this determines uniquely u in terms of (x,y,z). On the other hand, if $x^2+y^2 > a^2$, then $\cos u > 0$, and so either $0 < u < \frac{\pi}{2}$ or $\frac{3\pi}{2} \leq u < 2\pi$. Again, this determines uniquely u in terms of (x,y,z). In turn,

$$\cos v = \frac{x}{(r\cos u + a)}, \quad \sin v = \frac{y}{(r\cos u + a)},$$

which determine uniquely v. Thus, φ is one-to-one. We leave as an exercise to check that φ^{-1} is continuous.

Second method: *Let's write down M explicitly. We have $x^2+y^2 = (r\cos u + a)^2$, and so $\sqrt{x^2+y^2} = r\cos u + a > 0$ (since $a > r$). In turn, $\left(\sqrt{x^2+y^2} - a\right)^2 = r^2\cos^2 u$. Adding $r^2\sin^2 u = z^2$ to both sides gives*

$$M = \left\{(x,y,z) \in \mathbb{R}^3 : z^2 = r^2 - \left(\sqrt{x^2+y^2} - a\right)^2\right\}.$$

Consider the function

$$g(x,y,z) := z^2 + \left(\sqrt{x^2+y^2} - a\right)^2 - r^2.$$

Note that g is not of class C^1 in $\mathbb{R}^3$, since at points $(0,0,z)$, we have problems. However, these points do not belong to M. Indeed,

$$g(0,0,z) = z^2 + a^2 - r^2 \geq a^2 - r^2 > 0.$$

Hence, we can take $U = \left\{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 > 0\right\}$. Then g is of class C^1 in U since

$$\begin{aligned} J_g(x,y,z) &= \nabla g(x,y,z) \\ &= \left(2\left(\sqrt{x^2+y^2}-a\right)\frac{2x}{2\sqrt{x^2+y^2}}, 2\left(\sqrt{x^2+y^2}-a\right)\frac{2y}{2\sqrt{x^2+y^2}}, 2z\right) \end{aligned}$$

is continuous in U. To prove that $J_g(x,y,z)$ has rank one in U, note that if $z \neq 0$, then $\frac{\partial g}{\partial z} = 2z \neq 0$, while if $z = 0$, then $x^2+y^2 > 0$, so x and y cannot be both zero, so $\frac{\partial g}{\partial x}(x,y,z)$ or $\frac{\partial g}{\partial y}(x,y,z)$ are different from zero. Thus, $J_g(x,y,z)$ has maximum rank 1. This shows that M is a 2-dimensional surface of class C^∞.

Note that M is obtained by taking the circle of radius r and center $(0,a,0)$, namely,

$$(y-a)^2 + z^2 = r^2$$

and revolving it about the z.

The *Klein bottle* is obtained by starting from a rectangle of $\mathbb{R}^2$, reflecting one of its sides across its center and then performing the identification.

Exercise 226 (Klein Bottle) *Consider the function $\varphi : (0,2\pi)\times(0,2\pi) \to \mathbb{R}^4$ given by*

$$\varphi(u,v) := \left((r\cos v + a)\cos u, (r\cos v + a)\sin u, r\sin v\cos\frac{u}{2}, r\sin v\sin\frac{u}{2}\right).$$

Then Klein bottle is the set $M := \varphi((0,2\pi)\times(0,2\pi)) \subset \mathbb{R}^4$. Prove that M is a 2-dimensional surface of class C^∞.

Example 227 *Consider the unit sphere in $\mathbb{R}^N$, $N \geq 2$,*

$$M := \left\{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{x}\| = 1\right\}.$$

Given $\mathbf{x}_0 \in M$ consider the function $g(\mathbf{x}) := \|\mathbf{x}\|^2 - 1$ (here $N-k=1$ and so $k = N-1$). Then $J_g(\mathbf{x}) = \nabla g(\mathbf{x}) = 2\mathbf{x}$. By taking a ball $B(\mathbf{x}_0, r)$ so small that it does not intersect the origin, we have that $2\mathbf{x} \neq \mathbf{0}$ for all $\mathbf{x} \in M \cap B(\mathbf{x}_0, r)$ and so M is a $(N-1)$-dimensional surface of class C^∞.

Example 228 *Given an open set $U \subseteq \mathbb{R}^N$ and a function $f : U \to \mathbb{R}$ of class C^1, for every $c \in \mathbb{R}$ consider the level set*

$$B_c := \{\mathbf{x} \in U : f(\mathbf{x}) = c\}.$$

Given $\mathbf{x}_0 \in B_c$ consider the function $g(\mathbf{x}) := f(\mathbf{x}) - c$ (here $N - k = 1$ and so $k = N - 1$). Then $J_g(\mathbf{x}) = \nabla f(\mathbf{x})$. Hence, we have a problem if B_c contains some critical points. Thus, let

$$M := \{\mathbf{x} \in U : f(\mathbf{x}) = c\} \setminus \{\mathbf{x} \in U : \nabla f(\mathbf{x}) = \mathbf{0}\}.$$

Given $\mathbf{x}_0 \in M$, since $\nabla f(\mathbf{x}_0) \neq \mathbf{0}$ and f is of class C^1, by taking a small ball $B(\mathbf{x}_0, r)$ we have that $\nabla f(\mathbf{x}) \neq \mathbf{0}$ for all $\mathbf{x} \in M \cap B(\mathbf{x}_0, r)$ and so M is an $(N-1)$-dimensional surface of class C^1.

Friday, April 26, 2013

Definition 229 *Given a k-dimensional surface M and a point $\boldsymbol{x}_0 \in M$, a vector $\boldsymbol{t}$ is a* tangent vector *to M at $\boldsymbol{x}_0$ if there exists a function $\boldsymbol{\psi} : (-r, r) \to M$ such that $\boldsymbol{\psi}(0) = \boldsymbol{x}_0$ and $\boldsymbol{\psi}$ is differentiable at $t = 0$ with $\boldsymbol{\psi}'(0) = \boldsymbol{t}$. The set of all tangent vectors to M at $\boldsymbol{x}_0$ is called the tangent space to M at $\boldsymbol{x}_0$ and is denoted $T(\boldsymbol{x}_0)$.*

We will see that $T(\boldsymbol{x}_0)$ is a vector space of dimension k.

Definition 230 *Given a k-dimensional surface M and a point $\boldsymbol{x}_0 \in M$, a vector $\boldsymbol{n}$ is a* normal vector *to M at $\boldsymbol{x}_0$ if $\boldsymbol{n} \cdot \boldsymbol{t} = \mathbf{0}$ for all $\boldsymbol{t} \in T(\boldsymbol{x}_0)$.*

Theorem 231 *Given a k-dimensional surface M of class C^1 and a point $\boldsymbol{x}_0 \in M$, if $\varphi : W \to \mathbb{R}^N$ is a local chart with $\varphi(\boldsymbol{y}_0) = \boldsymbol{x}_0$ then the vectors $\frac{\partial \varphi}{\partial y_1}(\mathbf{y}_0)$, $\ldots$, $\frac{\partial \varphi}{\partial y_k}(\mathbf{y}_0)$ form a basis for the tangent space $T(\boldsymbol{x}_0)$. Hence, a vector $\boldsymbol{n}$ is a normal vector to M at $\boldsymbol{x}_0$ if $\boldsymbol{n} \cdot \boldsymbol{\frac{\partial \varphi}{\partial y_i}}(\mathbf{y_0}) = \mathbf{0}$ for all $i = 1, \ldots, k$.*

Theorem 232 *Given a k-dimensional surface M of class C^1 and a point $\boldsymbol{x}_0 \in M$, if M is represented as in (20), then the tangent space $T(\boldsymbol{x}_0)$ is given by the kernel (or null space) of the matrix $J_{\mathbf{g}}(\mathbf{x}_0)$. The vectors $\nabla g_1(\mathbf{x}_0)$, $\ldots$, $\nabla g_{N-k}(\mathbf{x}_0)$ forms a basis for the normal space.*

18 Surface Integrals

To define the integral of a function over a surface, we use local charts. Given $1 \leq k < N$ we define

$$\Lambda_{N,k} := \left\{\boldsymbol{\alpha} \in \mathbb{N}^k : 1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_k \leq N\right\}.$$

Given a k-dimensional surface M of class C^m, $m \in \mathbb{N}$, consider a local chart $\varphi : V \to M$, let $E \subseteq \varphi(V)$ be such that $\varphi^{-1}(E)$ is Peano–Jordan measurable and let $f : E \to \mathbb{R}$ be a bounded function. The *surface integral* of f over E is defined as

$$\int_E f\, d\mathcal{H}^k := \int_{\varphi^{-1}(E)} f(\varphi(\mathbf{y})) \sqrt{\sum_{\boldsymbol{\alpha}\in\Lambda_{N,k}} \left[\det \frac{\partial(\varphi_{\alpha_1},\ldots,\varphi_{\alpha_k})}{\partial(y_1,\ldots,y_k)}(\mathbf{y})\right]^2}\, d\mathbf{y}, \quad (22)$$

provided the function $\mathbf{y} \in \varphi^{-1}(E) \mapsto f(\varphi(\mathbf{y})) \sqrt{\sum_{\boldsymbol{\alpha}\in\Lambda_{N,k}} \left[\det \frac{\partial(\varphi_{\alpha_1},\ldots,\varphi_{\alpha_k})}{\partial(y_1,\ldots,y_k)}(\mathbf{y})\right]^2}$ is Riemann integrable. Note that $\int_E f\, d\mathcal{H}^k$ exists if f is bounded and continuous. In particular, if $f = 1$, the number

$$\mathcal{H}^k(E) := \int_E 1\, d\sigma = \int_{\varphi^{-1}(E)} \sqrt{\sum_{\boldsymbol{\alpha}\in\Lambda_{N,k}} \left[\det \frac{\partial(\varphi_{\alpha_1},\ldots,\varphi_{\alpha_k})}{\partial(y_1,\ldots,y_k)}(\mathbf{y})\right]^2}\, d\mathbf{y}$$

is called the k*-dimensional surface measure* of E.

Remark 233 *To understand the expression under the square root, observe that the Jacobian matrix of* φ,

$$J_\varphi(\mathbf{y}) = \begin{pmatrix} \frac{\partial\varphi_1}{\partial y_1}(\mathbf{y}) & \cdots & \frac{\partial\varphi_1}{\partial y_k}(\mathbf{y}) \\ \vdots & \cdots & \vdots \\ \frac{\partial\varphi_N}{\partial y_1}(\mathbf{y}) & \cdots & \frac{\partial\varphi_N}{\partial y_k}(\mathbf{y}) \end{pmatrix}$$

is an $N \times k$ *matrix, where* $1 \le k < N$. *The expression under square root is obtained by considering the sum of the squares of the determinant of all* $k \times k$ *submatrices of* $J_\varphi(\mathbf{y})$.

Remark 234 *It can be shown that the number* $\int_E f\, d\mathcal{H}^k$ *does not depend on the particular local chart* φ.

Example 235 *Consider the set*

$$M := \left\{(x,y,z,w) \in \mathbb{R}^4 : x^2+y^2=1,\ z^2+w^2=1,\ x,z>0\right\}.$$

Let's prove that it is a 2*-dimensional surface. A (local) chart* $\varphi : \left(-\frac{\pi}{2},\frac{\pi}{2}\right) \times \left(-\frac{\pi}{2},\frac{\pi}{2}\right) \to M$ *is given by*

$$\varphi(u,v) := (\cos u, \sin u, \cos v, \sin v),$$

so that

$$J_\varphi(u,v) = \begin{pmatrix} -\sin u & 0 \\ \cos u & 0 \\ 0 & -\sin v \\ 0 & \cos v \end{pmatrix}.$$

Since $\cos u, \cos v > 0$*, the submatrix* $\begin{pmatrix} \cos u & 0 \\ 0 & -\cos v \end{pmatrix}$ *has determinant* $-\cos u \cos v < 0$*, and so* $J_\varphi (u, v)$ *has rank* 2*. Moreover* φ *is one-to-one with continuous inverse and of class* C^1*. Hence,*

$$\begin{aligned}\sum_{\boldsymbol{\alpha} \in \Lambda_{4,2}} \left(\det \frac{\partial (\varphi_{\alpha_1}, \varphi_{\alpha_2})}{\partial (u, v)} (u, v) \right)^2 &= \left(\det \begin{pmatrix} -\sin u & 0 \\ \cos u & 0 \end{pmatrix} \right)^2 + \left(\det \begin{pmatrix} -\sin u & 0 \\ 0 & -\sin v \end{pmatrix} \right)^2 \\ &\quad + \left(\det \begin{pmatrix} -\sin u & 0 \\ 0 & \cos v \end{pmatrix} \right)^2 + \left(\det \begin{pmatrix} \cos u & 0 \\ 0 & -\sin v \end{pmatrix} \right)^2 \\ &\quad + \left(\det \begin{pmatrix} \cos u & 0 \\ 0 & \cos v \end{pmatrix} \right)^2 + \left(\det \begin{pmatrix} 0 & -\sin v \\ 0 & \sin v \end{pmatrix} \right)^2 \\ &= 0 + \sin^2 u \sin^2 v + \sin^2 u \cos^2 v + \cos^2 u \sin^2 v + \cos^2 u \cos^2 v + 0 \\ &= 1. \end{aligned}$$

Let's now calculate the surface integral $\int_M \left(x^2 + z^2\right) d\mathcal{H}^2$*. We have*

$$\begin{aligned}\int_M \left(x^2 + z^2\right) d\mathcal{H}^2 &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\cos^2 u + \cos^2 v\right) \sqrt{1}\, du dv \\ &= \pi \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 u \, du + \pi \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 v \, dv = \pi^2. \end{aligned}$$

Monday, April 29, 2013

19 Divergence Theorem

Definition 236 *An open and bounded set* $U \subset \mathbb{R}^N$ *is* regular *if for every* $\mathbf{x}_0 \in \partial U$ *there exist an open set* $V \subseteq \mathbb{R}^N$ *containing* $\mathbf{x}_0$ *and a function* $g : V \to \mathbb{R}$ *, with* $\nabla g (\mathbf{x}) \neq \mathbf{0}$ *for all* $\mathbf{x} \in V$*, such that*

$$\begin{aligned} U \cap V &= \{\mathbf{x} \in V : \, g (\mathbf{x}) < 0\}, \\ V \cap \partial U &= \{\mathbf{x} \in V : \, g (\mathbf{x}) = 0\}. \end{aligned}$$

It follows that ∂U is an $(N-1)$-dimensional surface of class C^1.

Definition 237 *Given an open, bounded, regular set* $U \subset \mathbb{R}^N$*, a unit normal vector* $\boldsymbol{\nu}$ *to* ∂U *at* $\mathbf{x}_0$ *is called a* unit outward normal *to* U *at* $\mathbf{x}_0$ *if there exists* $\delta > 0$ *such that* $\mathbf{x}_0 - t\boldsymbol{\nu} \in U$ *and* $\mathbf{x}_0 + t\boldsymbol{\nu} \in \mathbb{R}^N \setminus \overline{U}$ *for all* $0 < t < \delta$*.*

Exercise 238 *Prove that if* $U \subset \mathbb{R}^N$ *is an open, bounded, regular set, and* g *is as in Definition 236, then for every* $\mathbf{x}_0 \in \partial U$*, the unit outward normal to* U *at* $\mathbf{x}_0$ *is the unit vector*

$$\boldsymbol{\nu} (\mathbf{x}_0) = \frac{\nabla g (\mathbf{x}_0)}{\|\nabla g (\mathbf{x}_0)\|}.$$

Hence, $\boldsymbol{\nu} : \partial U \to \mathbb{R}^N$ *is a continuous function.*

We are ready to state the divergence theorem.

Theorem 239 (Divergence Theorem) *Let $U \subset \mathbb{R}^N$ be an open, bounded, regular set and let $\mathbf{f} : \overline{U} \to \mathbb{R}^N$ be such that $\mathbf{f}$ is bounded and continuous in $\overline{U}$ and there exist the partial derivatives of $\mathbf{f}$ in $\mathbb{R}^N$ at all $\mathbf{x} \in U$ and they are continuous and bounded. Then*

$$\int_U \operatorname{div} \mathbf{f}(\mathbf{x})\, d\mathbf{x} = \int_{\partial U} \mathbf{f}(\mathbf{x}) \cdot \boldsymbol{\nu}(\mathbf{x})\, d\mathcal{H}^{N-1}(\mathbf{x}),$$

where

$$\operatorname{div} \mathbf{f} := \sum_{i=1}^{N} \frac{\partial f_i}{\partial x_i}.$$

Remark 240 *In physics $\int_{\partial U} \mathbf{f}(\mathbf{x}) \cdot \boldsymbol{\nu}(\mathbf{x})\, d\mathcal{H}^{N-1}(\mathbf{x})$ represents the* outward flux *of a vector field $\mathbf{f}$ across the boundary of a region U.*

We recall that given two vectors $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ of $\mathbb{R}^3$, their *cross-product* or *vector-product* is defined by

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &:= (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1) \\ &= \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}. \end{aligned}$$

Remark 241 *For a 2-dimensional manifold M in $\mathbb{R}^3$, given a local chart $\varphi : W \to \mathbb{R}^2$, where $W \subseteq \mathbb{R}^2$ is an open set, we have that a basis for the tangent space at a point $\varphi(u, v) \in M$ is given by $\frac{\partial \varphi}{\partial u}(u, v)$ and $\frac{\partial \varphi}{\partial v}(u, v)$ and the two unit normals are given by*

$$\pm \frac{\frac{\partial \varphi}{\partial u}(u, v) \times \frac{\partial \varphi}{\partial v}(u, v)}{\left\| \frac{\partial \varphi}{\partial u}(u, v) \times \frac{\partial \varphi}{\partial v}(u, v) \right\|}.$$

It follows by direct computation that $\left\| \frac{\partial \varphi}{\partial u}(u, v) \times \frac{\partial \varphi}{\partial v}(u, v) \right\|$ is given exactly by $\sqrt{\sum_{\boldsymbol{\alpha} \in \Lambda_{3,2}} \left(\det \frac{\partial(\varphi_{\alpha_1}, \varphi_{\alpha_2})}{\partial(u, v)}(u, v) \right)^2}$. Hence,

$$\begin{aligned} &\int_{\varphi(W)} \mathbf{f}(x, y, z) \cdot \boldsymbol{\nu}(x, y, z)\, d\mathcal{H}^2(x, y, z) \\ &= \int_W \mathbf{f}(\varphi(u, v)) \cdot \frac{\pm \left(\frac{\partial \varphi}{\partial u}(u, v) \times \frac{\partial \varphi}{\partial v}(u, v) \right)}{\left\| \frac{\partial \varphi}{\partial u}(u, v) \times \frac{\partial \varphi}{\partial v}(u, v) \right\|} \sqrt{\sum_{\boldsymbol{\alpha} \in \Lambda_{3,2}} \left(\det \frac{\partial(\varphi_{\alpha_1}, \varphi_{\alpha_2})}{\partial(u, v)}(u, v) \right)^2} dudv \\ &= \pm \int_W \mathbf{f}(\varphi(u, v)) \cdot \left(\frac{\partial \varphi}{\partial u}(u, v) \times \frac{\partial \varphi}{\partial v}(u, v) \right) dudv, \end{aligned}$$

where the sign has to be chosen appropriately.

If $E \subseteq \mathbb{R}^N$ and $\mathbf{f} : E \to \mathbb{R}^N$ is differentiable, then $\mathbf{f}$ is called a *divergence-free field* or *solenoidal field* if

$$\operatorname{div} \mathbf{f} = 0.$$

Thus for a smooth solenoidal field, the outward flux across the boundary of a regular set U is zero.

Corollary 242 *The theorem continues to hold if $U \subset \mathbb{R}^N$ is open, bounded, and its boundary consists of two sets E_1 and E_2, where E_1 is a closed set contained in the finite union of closed and bounded manifolds of class C^1 and dimension less than $N - 1$, while for every $\mathbf{x}_0 \in E_2$ there exist a ball $B(\mathbf{x}_0, r)$ and a function $g \in C^1(B(\mathbf{x}_0, r))$ such that with $\nabla g(\mathbf{x}) \neq \mathbf{0}$ for all $\mathbf{x} \in B(\mathbf{x}_0, r) \cap \partial U$, such that*

$$\begin{aligned} B(\mathbf{x}_0, r) \cap U &= \{\mathbf{x} \in B(\mathbf{x}_0, r) : \, g(\mathbf{x}) < 0\}, \\ B(\mathbf{x}_0, r) \setminus \overline{U} &= \{\mathbf{x} \in B(\mathbf{x}_0, r) : \, g(\mathbf{x}) > 0\}, \\ B(\mathbf{x}_0, r) \cap \partial U &= \{\mathbf{x} \in B(\mathbf{x}_0, r) : \, g(\mathbf{x}) = 0\}. \end{aligned}$$

Remark 243 *In particular, the previous corollary works if $U \subset \mathbb{R}^N$ is open, bounded, and ∂U is given by a finite union of manifolds of dimension $N-1$ and a finite union of manifolds of dimension less than $N - 1$.*

Tuesday, April 30, 2013

Recitation.

Example 244 *Let's calculate the outward flux of the function*

$$\mathbf{f}(x, y, z) := \left(x, y, z^2\right)$$

across the boundary of the region

$$U := \left\{(x, y, z) \in \mathbb{R}^3 : \, -1 < z < -x^2 - y^2\right\}.$$

First method. *Let's use the divergence theorem. Note that U is not a regular open set (why?) but it satisfies the hypotheses of the previous corollary (why?).*

We have

$$\operatorname{div} \mathbf{f}(x, y, z) = \frac{\partial}{\partial x}(x) + \frac{\partial}{\partial y}(y) + \frac{\partial}{\partial z}\left(z^2\right) = 1 + 1 + 2z,$$

and so by the divergence theorem

$$\int_{\partial U} \mathbf{f} \cdot \boldsymbol{\nu} \, d\mathcal{H}^2 = 2 \int_U (1 + z) \, dxdydz.$$

Since U is a normal domain, we have

$$2\int_U (1+z)\ dxdydz = 2\int_E \left(\int_{-1}^{-x^2-y^2} (1+z)\ dz\right) dxdy$$
$$= 2\int_E \left[z+\frac{z^2}{2}\right]_{z=-1}^{z=-x^2-y^2} dxdy$$
$$= 2\int_E \left(\frac{1}{2}\left(x^2+y^2\right)^2 - x^2 - y^2 + \frac{1}{2}\right) dxdy,$$

where $E = \left\{(x,y)\in\mathbb{R}^2 :\ x^2+y^2<1\right\}$. Now using polar coordinates,

$$\int_E \left(\frac{1}{2}\left(x^2+y^2\right)^2 - x^2 - y^2 + \frac{1}{2}\right) dxdy = \int_0^{2\pi}\left(\int_0^1 \left(\frac{1}{2}r^4 - r^2 + \frac{1}{2}\right) r\, dr\right) d\theta$$
$$= 2\pi\int_0^1 \left(\frac{1}{2}r^5 - r^3 + \frac{1}{2}r\right) dr$$
$$= 2\pi\left[\frac{1}{12}r^6 - \frac{r^4}{4} + \frac{r^2}{4}\right]_{r=0}^{r=1}$$
$$= \frac{1}{6}\pi$$

Second method. *Let's use the definition of outward flux. The boundary of U is given by the union of two surfaces and by a curve (which is a lower dimensional manifold). Precisely,*

$$\partial U = M_1 \cup M_2 \cup M_3,$$

where

$$M_1 := \left\{(x,y,z)\in\mathbb{R}^3 :\ z=-1,\ x^2+y^2<1\right\},$$
$$M_2 := \left\{(x,y,z)\in\mathbb{R}^3 :\ z=-x^2-y^2,\ x^2+y^2<1\right\},$$
$$M_3 := \left\{(x,y,z)\in\mathbb{R}^3 :\ z=-1,\ x^2+y^2=1\right\}.$$

Note that M_1 and M_2 are 2-dimensional surfaces, since M_1 is the graph of the function $g_1(x,y)=1$ over $E=\left\{(x,y)\in\mathbb{R}^2 :\ x^2+y^2<1\right\}$ and M_2 is the graph of the function $g_2(x,y)=-x^2-y^2$ over E. On the other hand, M_3 is a one-dimensional manifold. Hence, we need to compute $\int_{M_1}\mathbf{f}\cdot\boldsymbol{\nu}\,d\mathcal{H}^2$ and $\int_{M_2}\mathbf{f}\cdot\boldsymbol{\nu}\,d\mathcal{H}^2$.

A chart for M_1 is given by

$$\boldsymbol{\varphi}(x,y) = (x,y,-1), \quad (x,y)\in E.$$

We have

$$\frac{\partial\boldsymbol{\varphi}}{\partial x}(x,y) = (1,0,0), \quad \frac{\partial\boldsymbol{\varphi}}{\partial y}(x,y) = (0,1,0).$$

Hence, to find the normal, we need a vector orthogonal to $(1,0,0)$ *and* $(0,1,0)$, *namely,* $\pm(0,0,1)$. *The outward normal is* $(0,0,-1)$. *Hence, using Remark 241*

$$\begin{aligned}\int_{M_1} \mathbf{f}\cdot\boldsymbol{\nu}\, d\mathcal{H}^2 &= \int_E (x,y,1)\cdot(0,0,-1)\; dxdy \\ &= -\int_E 1\, dxdy = -\pi.\end{aligned}$$

A chart for M_2 *is given by*

$$\boldsymbol{\varphi}(x,y) = \left(x,y,-x^2-y^2\right), \quad (x,y)\in E.$$

We have

$$\frac{\partial\boldsymbol{\varphi}}{\partial x}(x,y) = (1,0,-2x), \quad \frac{\partial\boldsymbol{\varphi}}{\partial y}(x,y) = (0,1,-2y).$$

Hence, to find the normal, we need

$$\begin{aligned}\frac{\partial\boldsymbol{\varphi}}{\partial x}(x,y)\times\frac{\partial\boldsymbol{\varphi}}{\partial y}(x,y) &= \det\begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 0 & -2x \\ 0 & 1 & -2y \end{pmatrix} \\ &= (2x,2y,1).\end{aligned}$$

Hence, the normals vectors are $\pm(2x,2y,1)$. *The outward normal is* $(2x,2y,1)$. *Hence, using Remark 241*

$$\begin{aligned}\int_{M_1} \mathbf{f}\cdot\boldsymbol{\nu}\, d\mathcal{H}^2 &= \int_E \left(x,y,\left(-x^2-y^2\right)^2\right)\cdot(2x,2y,1)\; dxdy \\ &= \int_E \left(2x^2+2y^2+\left(x^2+y^2\right)^2\right)\, dxdy.\end{aligned}$$

Now using polar coordinates,

$$\begin{aligned}\int_E \left(2x^2+2y^2+\left(x^2+y^2\right)^2\right)\, dxdy &= \int_0^{2\pi}\left(\int_0^1 \left(2r^2+r^4\right) r\, dr\right) d\theta \\ &= 2\pi\int_0^1 \left(2r^3+r^5\right)\, dr \\ &= 2\pi\left[\frac{1}{2}r^4+\frac{r^6}{6}\right]_{r=0}^{r=1} \\ &= \frac{4}{3}\pi.\end{aligned}$$

Example 245 *Let's calculate the outward flux of the function*

$$\mathbf{f}(x,y,z) := (0,yz,x)$$

across the boundary of the region

$$U := \left\{(x,y,z)\in\mathbb{R}^3 :\, x^2+y^2<z^2,\, x^2+y^2+z^2<2y,\, z>0\right\}.$$

Note that U is not a regular open set (why?) but it satisfies the hypotheses of the previous corollary (why?).

We have

$$\operatorname{div}\mathbf{f}(x,y,z)=\frac{\partial}{\partial x}(0)+\frac{\partial}{\partial y}(yz)+\frac{\partial}{\partial z}(x)=0+1z+0,$$

and so by the divergence theorem

$$\int_{\partial U}\mathbf{f}\cdot\boldsymbol{\nu}\,d\mathcal{H}^2=\int_U z\,dxdydz.$$

Using cylindrical coordinates $x=r\cos\theta$, $y=r\sin\theta$, $z=z$, we get $r^2\cos^2\theta+r^2\sin^2\theta<z^2$, that is, $r^2<z^2$, $r^2\cos^2\theta+r^2\sin^2\theta+z^2<2r\sin\theta$, that is, $r^2+z^2<2r\sin\theta$, and $z>0$, hence, $r^2<z^2<2r\sin\theta-r^2$, which implies that $r^2<2r\sin\theta-r^2$, or equivalently, $r<\sin\theta$. In turn, $\sin\theta$ should be positive, and so $\theta\in(0,\pi)$.

$$W:=\left\{(r,\theta,z)\in\mathbb{R}^3:\,0<\theta<\pi,\,r<\sin\theta,\,r<z<\sqrt{2r\sin\theta-r^2}\right\}.$$

Hence, by changing variables

$$\begin{aligned}\int_U z\,dxdydz&=\int_0^\pi\left(\int_0^{\sin\theta}\left(\int_r^{\sqrt{2r\sin\theta-r^2}}z\,dz\right)r\,dr\right)d\theta\\&=\int_0^\pi\left(\int_0^{\sin\theta}\left[\frac{z^2}{2}\right]_{z=r}^{z=\sqrt{2r\sin\theta-r^2}}r\,dr\right)d\theta\\&=\int_0^\pi\left(\int_0^{\sin\theta}\left(r^2\sin\theta-r^3\right)\,dr\right)d\theta\\&=\int_0^\pi\left[\frac{r^3}{3}\sin\theta-\frac{r^4}{4}\right]_{r=0}^{r=\sin\theta}d\theta=\int_0^\pi\left[\frac{1}{3}\sin^4\theta-\frac{1}{4}\sin^4\theta\right]d\theta\\&=\int_0^\pi\frac{1}{12}\sin^4\theta\,d\theta=\frac{1}{32}\pi.\end{aligned}$$

Wednesday, May 1, 2013

Corollary 246 (Integration by Parts) *Let $U\subset\mathbb{R}^N$ be an open, bounded, regular set and let $f:\overline{U}\to\mathbb{R}$ and $g:\overline{U}\to\mathbb{R}$ be such that f and g are bounded and continuous in $\overline{U}$ and there exist the partial derivatives of f and g in $\mathbb{R}$ at all $\mathbf{x}\in U$ and they are continuous and bounded. Then for every $i=1,\dots,N$,*

$$\int_U f(\mathbf{x})\frac{\partial g}{\partial x_i}(\mathbf{x})\,d\mathbf{x}=-\int_U g(\mathbf{x})\frac{\partial f}{\partial x_i}(\mathbf{x})\,d\mathbf{x}+\int_{\partial U}f(\mathbf{x})\,g(\mathbf{x})\,\nu_i(\mathbf{x})\,d\mathcal{H}^{N-1}(\mathbf{x}).$$

Proof. Fix $i\in\{1,\dots,N\}$. We apply the divergence theorem to the function $\mathbf{f}:\overline{U}\to\mathbb{R}^N$ defined by

$$f_j(\mathbf{x}):=\begin{cases}f(\mathbf{x})\,g(\mathbf{x}) & \text{if } j=i,\\ 0 & \text{if } j\neq i.\end{cases}$$

Then

$$\operatorname{div}\mathbf{f} = \sum_{j=1}^{N} \frac{\partial f_j}{\partial x_j} = \frac{\partial (fg)}{\partial x_i} = f\frac{\partial g}{\partial x_i} + g\frac{\partial f}{\partial x_i},$$

and so

$$\int_U \left(f\frac{\partial g}{\partial x_i} + g\frac{\partial f}{\partial x_i} \right) d\mathbf{x} = \int_U \operatorname{div}\mathbf{f}\, d\mathbf{x} = \int_{\partial U} \mathbf{f}\cdot\boldsymbol{\nu}\, d\mathcal{H}^{N-1} = \int_{\partial U} fg\nu_i\, d\mathcal{H}^{N-1}.$$

■

Another important application is given by the area's formulas in $\mathbb{R}^2$. Let $U \subset \mathbb{R}^2$ be an open, bounded set and assume that its boundary ∂U is the range of a closed, simple, regular curve $\boldsymbol{\gamma}$ with parametric representation $\boldsymbol{\varphi} : [a,b] \to \mathbb{R}^2$. Then the hypotheses of Corollary 242 are satisfied (why?). Given $t \in [a,b]$, the vector $\boldsymbol{\varphi}'(t)$ is tangent to the curve at the point $\boldsymbol{\varphi}(t) \in \partial U$. Hence, if $\boldsymbol{\varphi}(t) = (x(t), y(t))$, the outer normal to ∂U at the point $\boldsymbol{\varphi}(t)$ is either

$$\frac{(-y'(t), x'(t))}{\sqrt{(x'(t))^2 + (y'(t))^2}} \quad \text{or} \quad -\frac{(-y'(t), x'(t))}{\sqrt{(x'(t))^2 + (y'(t))^2}},$$

depending on the orientation of the curve. We say that $\boldsymbol{\gamma}$ has a *positive orientation* for ∂U if as we traverse the curve starting from $t = a$ we find U on our left. In this case, the outward normal at the point $\boldsymbol{\varphi}(t)$ is is given by

$$\boldsymbol{\nu}(\boldsymbol{\varphi}(t)) = \frac{(y'(t), -x'(t))}{\sqrt{(x'(t))^2 + (y'(t))^2}}.$$

Hence, if $f, g : \overline{U} \to \mathbb{R}$ are bounded and continuous in $\overline{U}$ and there exist the partial derivatives of f and g at all $(x,y) \in U$ and they are continuous and bounded, then applying the divergence theorem to the function $\mathbf{f}(x,y) := (f(x,y), g(x,y))$

$$\begin{aligned}
\int_U \left(\frac{\partial f}{\partial x}(x,y) + \frac{\partial g}{\partial y}(x,y) \right) dxdy &= \int_{\partial U} (f,g)\cdot\boldsymbol{\nu}\, d\mathcal{H}^1 \\
&= \int_a^b f(x(t), y(t)) \frac{y'(t)}{\sqrt{(x'(t))^2 + (y'(t))^2}} \sqrt{(x'(t))^2 + (y'(t))^2}\, dt \\
&\quad + \int_a^b g(x(t), y(t)) \frac{-x'(t)}{\sqrt{(x'(t))^2 + (y'(t))^2}} \sqrt{(x'(t))^2 + (y'(t))^2}\, dt, \\
&= \int_a^b [f(x(t), y(t))\, y'(t) - g(x(t), y(t))\, x'(t)]\, dt \\
&=: \int_{\partial U} f\, dy - \int_{\partial U} g\, dx.
\end{aligned}$$

Taking $g = 0$ or $f = 0$ we get the *Gauss-Green formulas*

$$\int_U \frac{\partial f}{\partial x}(x,y)\ dxdy = \int_{\partial U} f\,dy. \tag{23}$$

$$\int_U \frac{\partial g}{\partial y}(x,y)\ dxdy = -\int_{\partial U} g\,dx. \tag{24}$$

Taking $f(x,y) = x$ in (23), we get the *second area's formula*

$$\operatorname{meas} U = \int_U 1\,dxdy = \int_a^b x(t)\,y'(t)\ dt = \int_{\partial U} x\,dy,$$

while taking $g(x,y) = y$ in (24) we get the *second area's formula*

$$\operatorname{meas} U = \int_U 1\,dxdy = -\int_a^b y(t)\,x'(t)\ dt = -\int_{\partial U} y\,dx.$$

By adding these two identities, we get the *third area's formula*

$$\operatorname{meas} U = \frac{1}{2}\int_{\partial U} \left[-y(t)\,x'(t) + x(t)\,y'(t)\right]\ dt = \frac{1}{2}\left[\int_{\partial U} x\,dy - \int_{\partial U} y\,dx\right].$$

Example 247 *Let's find the area of the region U enclosed by the curve of parametric representation*

$$\begin{cases} x(\theta) = (1-\cos\theta)\cos\theta, \\ y(\theta) = (1-\cos\theta)\sin\theta, \end{cases} \qquad \theta \in [0, 2\pi].$$

To see that it is simple, note that for $0 < \theta < 2\pi$, $(1-\cos\theta) > 0$. The curve starts from the origin and goes around clockwise only once. Note that

$$x'(\theta) = -\sin\theta + 2\sin\theta\cos\theta = \sin\theta\,(-1 + 2\cos\theta),$$
$$y'(\theta) = \cos\theta + \sin^2\theta - \cos^2\theta = -2\cos^2\theta + \cos\theta + 1.$$

Thus, $x'(\theta) = 0$ only for $\sin\theta = 0$, that is $\theta = 0, \pi, 2\pi$,and $\cos\theta = \frac{1}{2}$, that is, for $\theta = \frac{1}{3}\pi, \frac{5}{3}\pi$. At those values,

$$y'(0) = y'(2\pi) = -2 + 1 + 1 = 0, \quad y'(\pi) = -2 - 1 + 1,$$
$$y'\left(\frac{1}{3}\pi\right) = -2\cos^2\frac{1}{3}\pi + \cos\frac{1}{3}\pi + 1 = 1,$$
$$y'\left(\frac{5}{3}\pi\right) = -2\cos^2\frac{5}{3}\pi + \cos\frac{5}{3}\pi + 1 = 1.$$

Hence, the curve is not regular, but since these are only a finite number of bad points, we can apply Corollary 242. Using the third area formula, we get

$$\begin{aligned}\operatorname{meas} U &= -\frac{1}{2}\int_0^{2\pi} (1-\cos\theta)\sin\theta\left[\sin\theta\cos\theta - (1-\cos\theta)\sin\theta\right]\ d\theta \\ &+ \frac{1}{2}\int_0^{2\pi} (1-\cos\theta)\cos\theta\left[\sin\theta\sin\theta + (1-\cos\theta)\cos\theta\right]\ d\theta \\ &= \frac{1}{2}\int_0^{2\pi} (1-\cos\theta)^2\left[\cos^2\theta + \sin^2\theta\right]\ d\theta = \frac{1}{2}\int_0^{2\pi} (1-\cos\theta)^2\ d\theta = \frac{3}{2}\pi.\end{aligned}$$

Friday, May 3, 2013

20 Stokes' Theorem

To state Stokes' theorem, we need to introduce the notion of a manifold with boundary.

Definition 248 *Given a set $E \subseteq \mathbb{R}^k$, a set $F \subseteq E$ is relatively open in E if there exists an open set $V \subseteq \mathbb{R}^k$ such that $F = E \cap V$.*

Denote $\mathbb{H}^k := \left\{ \mathbf{y} \in \mathbb{R}^k : x_k \geq 0 \right\}$. Note that $\mathbb{H}^k$ is a closed upper half-space.

Definition 249 *Given $1 \leq k < N$, a nonempty set $M \subseteq \mathbb{R}^N$ is called a k-*dimensional differential manifold with boundary *if for every $\mathbf{x}_0 \in M$ there exist an open set U containing $\mathbf{x}_0$ and a differentiable function $\varphi : W \to \mathbb{R}^N$, where $W \subseteq \mathbb{R}^k$ is either an open set or a relatively open set of $\mathbb{H}^k$ such that*

- *(i) $\varphi : W \to M \cap U$ is a homeomorphism, that is, it is invertible and continuous together with its inverse $\varphi^{-1} : M \cap U \to W$,*
- *(ii) $J_\varphi (\mathbf{y})$ has maximum rank k for all $\mathbf{y} \in W$.*

The function φ is called a local chart *or a* system of local coordinates *or a* local parametrization *around $\mathbf{x}_0$. We say that M is of class C^m, $m \in \mathbb{N}$, (respectively, C^∞) if all local charts are of class C^m (respectively, C^∞).*

A point $\mathbf{x}_0 \in M$ is an interior point *of M if there is a local chart $\varphi : W \to \mathbb{R}^N$, where $W \subseteq \mathbb{R}^k$ is open set, and is a* boundary point *of M, otherwise. The set of boundary points is called the* boundary *of M and is denoted ∂M.*

Theorem 250 *Let $M \subseteq \mathbb{R}^N$ be a k-dimensional differentiable manifold with boundary. Then ∂M is a $(k-1)$-dimensional differentiable manifold.*

Next we introduce the notion of orientation for a manifold with boundary. Let $M \subseteq \mathbb{R}^N$ be a k-dimensional manifold with boundary of class C^1. Given $\mathbf{x}_0 \in M$, the tangent space $T(\mathbf{x}_0)$ is a vector space of dimension k. We say that two bases $\mathcal{B} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$ and $\mathcal{B}_1 = \{\boldsymbol{f}_1, \ldots, \boldsymbol{f}_k\}$ of $T(\mathbf{x}_0)$ are equivalent if the matrix that transforms one basis into the other has positive determinant and we write $\mathcal{B} \sim \mathcal{B}_1$. An *orientation* at $\mathbf{x}_0$ is an equivalence class $[\mathcal{B}]$. Note that there are only two possible orientations at $\mathbf{x}_0$. A (possibly discontinuous) orientation O on M is a mapping that assigns to each $\mathbf{x} \in M$ an orientation $O_{\mathbf{x}}$. We say that an orientation O is *continuous* if there exist continuous functions $\boldsymbol{e}_1 : M \to \mathbb{R}^N$, $\ldots$, $\boldsymbol{e}_k : M \to \mathbb{R}^N$ such that for every $\mathbf{x} \in M$, $\mathcal{B}_{\mathbf{x}} = \{\boldsymbol{e}_1(\mathbf{x}), \ldots, \boldsymbol{e}_k(\mathbf{x})\}$ is a basis for $T(\mathbf{x})$ and $O_{\mathbf{x}} = [\mathcal{B}_{\mathbf{x}}]$. We say that M is *orientable* if it admits a continuous orientation.

Remark 251 *If $M \subseteq \mathbb{R}^{N-1}$ is an $(N-1)$-dimensional manifold of class C^1, then for each $\mathbf{x} \in M$, the tangent space $T(\mathbf{x})$ is a vector space of dimension*

$N-1$. *In turn, there are only two unit normals at the point* $\mathbf{x}\in M$. *It can be shown that* M *is orientable if and only if for every* $\mathbf{x}\in M$ *we can choose a unit normal vector* $\boldsymbol{n}(\mathbf{x})$ *to* M *at* $\mathbf{x}$ *in such a way that the function* $\boldsymbol{n}: M\to\mathbb{R}^N$ *is continuous.*

Theorem 252 *Let* $M\subseteq\mathbb{R}^N$ *be a* k*-dimensional orientable manifold with boundary of class* C^1. *Then a continuous orientation on* M *induces a continuous orientation on* ∂M.

The Klein bottle or the Möbius strip are a typical examples of manifolds that not orientable. Indeed, the normal as a function of (x,y,z) is discontinuous (but it is continuous if you look at it as a function of (u,v)). If you think of the normal as a person walking on the manifold, after going around once, that person would find himself/herself upside down, which means that there is a discontinuity.

Example 253 (The Möbius strip) *Let* $\varphi:[0,2\pi]\times[-1,1]\to\mathbb{R}^3$ *be given by*

$$\varphi(u,v):=\left(\left(2-v\sin\frac{u}{2}\right)\sin u,\left(2-v\sin\frac{u}{2}\right)\cos u, v\cos\frac{u}{2}\right).$$

It can be shown that $M=\varphi([0,2\pi]\times[-1,1])$ *is a* 2*-dimensional manifold. We have*

$$\frac{\partial\boldsymbol{\varphi}}{\partial u}(u,v)=\left(-(\cos u)\left(v\sin\frac{1}{2}u-2\right)-\frac{1}{2}v\sin u\cos\frac{1}{2}u,(\sin u)\left(v\sin\frac{1}{2}u-2\right)-\frac{1}{2}v\cos u\cos\frac{1}{2}u,-\frac{1}{2}v\sin\frac{1}{2}u\right)$$
$$\frac{\partial\boldsymbol{\varphi}}{\partial v}(u,v)=\left(-\sin u\sin\frac{1}{2}u,-\cos u\sin\frac{1}{2}u,\cos\frac{1}{2}u\right).$$

Hence, to find the normal, we need

$$\begin{aligned}&\frac{\partial\boldsymbol{\varphi}}{\partial u}(u,v)\times\frac{\partial\boldsymbol{\varphi}}{\partial v}(u,v)\\ &=\begin{vmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\\ -(\cos u)\left(v\sin\frac{1}{2}u-2\right)-\frac{1}{2}v\sin u\cos\frac{1}{2}u & (\sin u)\left(v\sin\frac{1}{2}u-2\right)-\frac{1}{2}v\cos u\cos\frac{1}{2}u & -\frac{1}{2}v\sin\frac{1}{2}u\\ -\sin u\sin\frac{1}{2}u & -\cos u\sin\frac{1}{2}u & \cos\frac{1}{2}u\end{vmatrix}\\ &=\left(\left(v\sin\frac{1}{2}u-2\right)\sin u\cos\frac{u}{2}-\frac{1}{2}v\cos u,\left(v\sin\frac{1}{2}u-2\right)\cos u\cos\frac{u}{2}+\frac{1}{2}v\sin u,-\left(v\sin\frac{1}{2}u-2\right)\sin\frac{u}{2}\right).\end{aligned}$$

Hence,

$$\frac{\partial\boldsymbol{\varphi}}{\partial u}(0,0)\times\frac{\partial\boldsymbol{\varphi}}{\partial v}(0,0)=(0,-2,0),$$
$$\frac{\partial\boldsymbol{\varphi}}{\partial u}(2\pi,0)\times\frac{\partial\boldsymbol{\varphi}}{\partial v}(2\pi,0)=(0,2,0).$$

Let $U \subseteq \mathbb{R}^3$ be an open set and let $\mathbf{f} : U \to \mathbb{R}^3$ be a differentiable function. The *curl* of $\mathbf{f}$ is the function $\operatorname{curl} \mathbf{f} : U \to \mathbb{R}^3$ defined by

$$\begin{aligned}\operatorname{curl} \mathbf{f}(x, y, z) &:= \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right) \\ &= \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1 & f_2 & f_3 \end{pmatrix} = \nabla \times \mathbf{f}.\end{aligned}$$

Theorem 254 (Stokes' Theorem) *Let $U \subseteq \mathbb{R}^3$ be an open set and let $\mathbf{f} : U \to \mathbb{R}^3$ be a function of class C^1. Let $M \subseteq U$ be a closed bounded 2-dimensional orientable manifold with boundary of class C^2. Then*

$$\int_M \operatorname{curl} \mathbf{f} \cdot \boldsymbol{n} \, d\mathcal{H}^2 = \int_{\partial M} \mathbf{f} \cdot \boldsymbol{t} \, d\mathcal{H}^1,$$

where $\boldsymbol{n} : M \to \mathbb{R}^3$ is a continuous unit normal to M and $\boldsymbol{t} : \partial M \to \mathbb{R}^3$ is a continuous unit tangent to ∂M (with the orientation induced by the orientation of M).

Theorem 255 *Given a k-dimensional manifold $M \subseteq \mathbb{R}^N$ with boundary of class C^1, let $\varphi : W \to \mathbb{R}^N$, where $W \subseteq \mathbb{R}^k$ is an open set. Consider a bounded open set V with $\overline{V} \subset W$ and let $M_1 := \varphi(\overline{V})$. Then M_1 is a k-dimensional orientable manifold with boundary of class C^1. Moreover, $\partial M_1 = \varphi(\partial V)$.*

The previous theorem shows that every manifold is locally orientable. It will be useful in the exercises.

Example 256 *Given the surface M of parametric representation*

$$\varphi(u, v) := \left(u - v, u, u^2 + v^2\right), \quad (u, v) \in V,$$

where $V := \left\{(u, v) \in \mathbb{R}^2 : \, u^2 + v^2 < 1\right\}$, let's calculate

$$\int_M \operatorname{curl} \mathbf{f} \cdot \boldsymbol{\nu} \, d\mathcal{H}^2,$$

where

$$\mathbf{f}(x, y, z) := \left(x^2 z, y, yz\right), \quad (x, y, z) \in \mathbb{R}^3.$$

Note that φ is defined in $\mathbb{R}^2$ so we can take $W = \mathbb{R}^2$. To see that φ is one-to-one, we find the inverse, we have $x = u - v$, $y = u$, $z = u^2 + v^2$ and so, $u = z$, $v = u - x = z - x$. Thus,

$$\varphi^{-1}(x, y, z) = (z, z - x),$$

which is continuous. Hence, φ is a homeomorphism. Moreover,

$$J_\varphi(u, v) = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 2u & 2v \end{pmatrix}.$$

The submatrix $\begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$ *has determinant different from zero, and so* $J_{\varphi}(u,v)$ *has always maximal rank two. In view of the previous theorem,* M *is orientable. Thus, we are in a position to apply Stokes' theorem. The boundary of* V *is the unit circle. Thus a parametric representation of* $\boldsymbol{\gamma}$ *is* $\boldsymbol{\psi} : [0, 2\pi] \to \mathbb{R}^2$*, where* $\boldsymbol{\psi}(t) := (\cos t, \sin t)$*. In turn, a parametric representation for* ∂M *is given by* $\phi := \varphi \circ \boldsymbol{\psi} : [0, 2\pi] \to \mathbb{R}^3$*, where*

$$\phi(t) = \left(\cos t - \sin t, \cos t, \cos^2 t + \sin^2 t\right).$$

In turn,

$$\phi'(t) = (-\sin t - \cos t, -\sin t, 0).$$

Hence, by Stokes' theorem

$$\begin{aligned}
\int_M \operatorname{curl} \mathbf{f} \cdot \boldsymbol{\nu} \, d\mathcal{H}^2 &= \int_{\partial M} \mathbf{f} \cdot \boldsymbol{t} \, d\mathcal{H}^1 \\
&= \int_0^{2\pi} \left[(\cos t - \sin t)^2 \, 1 \, (-\sin t - \cos t) + (\cos t)(-\sin t) + (\cos t)(1) \, 0 \right] dt \\
&= -\int_0^{2\pi} \left[\left(\cos^2 t - \sin^2 t\right)(\cos t - \sin t) + \cos t \sin t \right] dt \\
&= -\int_0^{2\pi} \left[\left(1 - 2\sin^2 t\right) \cos t - \left(-1 + 2\cos^2 t\right) \sin t + \cos t \sin t \right] dt \\
&= -\left[\sin t - \frac{2}{3} \sin^3 t + \cos t + \frac{2}{3} \cos^3 t + \frac{1}{2} \sin^2 t \right]_{t=0}^{t=2\pi} = 0.
\end{aligned}$$

Exercise 257 *In the previous example calculate* $\int_M \operatorname{curl} \mathbf{f} \cdot \boldsymbol{\nu} \, d\mathcal{H}^2$ *directly, without using Stokes's theorem.*