➢ 28th September 1999

Introduction

Differentiation has a **geometric** aspect (measuring the *gradient* of an arc and then letting Δt tend to zero) and an **algebraic** aspect (linear approximation: $f(x) = f(a) + {}^{dt}/_{dx}(a)(x-a)$ + a bit). **Recalling**, $\mathbf{R}^m$ is an *m-dimensional* Euclidean space: $\underline{x} = (x_1, ..., x_m)$, $\underline{x}.\underline{y} = x_1y_1+x_2y_2+...+x_my_m = \Sigma^m{}_{i=i}\, x_iy_i$. The **norm** is $|\underline{x}| = (\underline{x}.\underline{x})^{1/2} = (\Sigma_{i=1}{}^m x_i{}^2)^{1/2}$. The *distance* $|\underline{x}-\underline{y}|$ is the distance between x and y. T: $\mathbf{R}^m \rightarrow \mathbf{R}^n$ is a *linear* function, with $T(r\underline{x}+s\underline{y}) = rT(\underline{x})+sT(\underline{y})$.

If we are given **bases** for $\mathbf{R}^m$ and $\mathbf{R}^n$, we can work out T in terms of a *matrix*. A plane through O in $\mathbf{R}^3$ has equation $\underline{a}.\underline{x} = 0$. $a = (a_1, a_2, a_3)$ is **normal** to the plane. For a plane going through a *vector* $\underline{b}$, the equation is $\underline{a}.(\underline{x}-\underline{b}) = 0$. Again, $\underline{a}$ is in the **normal** direction. This generalises to $\mathbf{R}^n$ for *all* n.

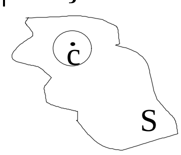
Section 1: Multivariable Differential Calculus

Let S be a subset of $\mathbf{R}^n$; and let f be a *function* from S to $\mathbf{R}^n$, f: $S \rightarrow \mathbf{R}^n$. If $\underline{c} \in S$ and $u \in \mathbf{R}^n$, and we have $\underline{c}+\underline{u} \in S$, $(\underline{u} \neq 0)$, *a problem* could be to study the change in f(x) as x **goes** from $\underline{c}$ to $\underline{c}+\underline{u}$ along the line *segment* $\underline{x} = \underline{c}+t\underline{u}$, where $t \in [0,1]$.

1.1. Directional Derivative

Study f as it *changes* in the direction of u. We need a fairly general S, since the domains of functions can be complicated. For example, with $f(x,y) = {}^1/_{x^2+y^2-1}$, if $x, y \in S$, then $S = \{(x,y) \mid x^2+y^2 \neq 1\}$. Similarly, for $f(x,y) = 1/(x^2+y^2-1)^{1/2}$, we need a *suitable* S. We usually take $\underline{u}$ to be small. In the above, S is *everything* but the circle $x^2+y^2 = 1$. If we take *u to be* "big", then we may not have $c+tu \in S$ for all $t \in [0,1]$.

Usually, we want S to be "**open**". Suppose that $r > 0$. Let $B(\underline{c};r) = \{\underline{x} \mid \|\underline{x}-\underline{c}\| < r\}$ be the open ball with centre $\underline{c}$, radius r. *Definition*: $S \subseteq \mathbf{R}^m$ is open if for all $\underline{c} \in S$, there is *some* $r > 0$ with $B(\underline{c};r) \subseteq S$. If S is *open* and t is small enough, then $\underline{c}+t\underline{u}$ will be in $B(\underline{c};r)$, because $\|(\underline{c}+t\underline{u})-\underline{c}\| = t\|\underline{u}\|$, which can be as *small* as it needs to be (it will be $< r$ if $t < {}^r/_{\|u\|}$).



Definition (for n = 1): The *directional derivative* of f at $\underline{c}$, in the direction of $\underline{u}$, denoted by $f'(\underline{c};\underline{u})$, is defined by $f'(\underline{c};\underline{u}) = \lim_{t\to 0} (f(\underline{c}+t\underline{u})-f(\underline{c})/t)$, wherever the *limit* exists. Pick a **co-ordinate** system. The partial derivatives are (directional derivatives in the *direction* of the axes) $\frac{\delta f}{\delta x_k}\big|_c = f'(\underline{c};e_k)$, where e_k is the **unit** vector (0,...,0,1,0,...,0), with 1 in the k^{th} position.

Example: m = 2, n = 1, $f(x,y,z) = xy+yz^2$ with $S = \mathbf{R}^3$. ${}^{\delta f}/_{\delta x} = y$, ${}^{\delta f}/_{\delta y} = x+z^2$ and ${}^{\delta f}/_{\delta z} = 2yz$. (Taken at a *general* point c = (x,y,z)). **Example**: $\underline{a} = (a_1,...,a_m) \in \mathbf{R}^m$, $f(\underline{x}) = \underline{a}.\underline{x}$. Find $\delta f/\delta x_k$ from *first* principles. Now $f(\underline{x}+t\underline{e}_k) = \underline{a}.(\underline{x}+t\underline{e}_k) = \underline{a}.\underline{x} + t\underline{a}.\underline{e}_k = \underline{a}.\underline{x}+ta_k$. Now $\lim_{t\to 0} f(\underline{x}+t\underline{e}_k)-f(x)/t = \lim_{t\to 0} \underline{a}.\underline{x} + ta_k-\underline{a}.\underline{x}/t = a_k$. In **general,** for $f'(\underline{x};\underline{u})$, $f(\underline{x}+t\underline{u}) = \underline{a}.\underline{x}+t\underline{a}.\underline{u} = \underline{a}.\underline{x}+t\Sigma a_iu_i$. $f'(\underline{x};\underline{u}) = \Sigma a_iu_i = \underline{a}.\underline{u} = \nabla f.\underline{u}$ (where $\nabla f = (\delta f/\delta x_1,...,\delta f/\delta x_m)$).

Theorem: Suppose f: S→**R**, and assume all *partial* derivatives exist **everywhere** in S. Set $\nabla f = (\delta f/\delta x_1,...,\delta f/\delta x_m) = \Sigma_{i=1}^{m}\ \delta f/\delta x_i\ \underline{e}_i$, the *gradient* of f. Then for all c and u, f'(c;u) *exists*, and f'(c;u) = ∇f.u.

Tutorial

Remember in **partial** differentiation, only differentiate in the *order* indicated, and only w.r.t. the variable(s) *involved*. Rules for stationary points (Revision): If z = f(x, y), then to test for **stationary** points, solve $^{\delta z}/_{\delta x} = {}^{\delta z}/_{\delta y} = 0$. (**Important**). Then let $A = {}^{\delta^2 z}/_{\delta x^2}$, $B = {}^{\delta^2 z}/_{\delta x \delta y}$ and $C = {}^{\delta^2 z}/_{\delta y^2}$. Form the **Hessian** matrix of f, $H = ({}^{A}{}_{B}\ {}^{B}{}_{C})$. If H is *positive* definite at a stationary point P, then it is a local min. If it is positive *negative,* P is a local max. If it's **indefinite,** then P is a saddle point. If det(H) is zero, then the *test* fails.

So if $AC > B^2$ and $A < 0$, we have a local max; if $AC > B^2$ and $A > 0$, we have a local min; if $AC < B^2$, it is a saddle point; and if $AC = B^2$, the test fails. The **tangent** plane at the point $P(x_0,y_0,z_0)$ on the surface f(x,y,z) = c is the **plane** $f_x(P)(x-x_0)+f_y(P)(y-y_0)+f_z(P)(z-z_0) = 0$. The **normal** line of the surface at P is the line $x = x_0+f_x(P)t$, $y = y_0+f_y(P)t$, $z = z_0+f_z(P)t$.

Convex functions. A function $f(x_1,...,x_n)$ is a convex function on a *convex* set S (in $\mathbf{R}^n$) if for any x' ∈ S and x'' ∈ S, $f(cx'+(1-c)x^n) \le cf(x')+(1-c)f(x^n)$ holds for $0 \le c \le 1$. Define the **Hessian**, H, of $f(x_1,...,x_n)$, as being the n×n matrix whose ij^{th} entry is $\delta^2 f/\delta x_i \delta x_j$. Winston quotes the **following** theorem: Suppose that $f(x_1,...,x_n)$ has continuous second-order partial *derivatives* for each point $x = (x_1,...,x_n) \in S$. Then $f(x_1,...,x_n)$ is a convex function on S if and only if for each x ∈ S, all **principal minors** of H are non-negative.

Example: Find the equation of the *tangent* plane and the normal line to the quadric surface $3x^2+2y^2+z^2+4x+5y+6z = 3$ at P(-1,-1,1). A: Let f(x,y,z) be $3x^2+2y^2+z^2+4x+5y+6z-3 = 0$. Let P be the **point** (-1,-1,1). $F_x(P) = 6x+4 = 6(-1)+4 = -2$. $F_y(P) = 4y+5 = 4(-1)+5 = 1$. $F_z(P) = 2z+6 = 2(1)+6 = 8$. The *tangent* plane is $f_x(P)(x-x_0)+f_y(P)(y-y_0)+f_z(P)(z-z_0) = 0$; -2(x--1)+1(y--1)+8(z-1) = 0; -2x-2+y+1+8z-8 = 0; -2x+y+8z = 9. The line **normal** to the *surface* at P is given by $x = x_0+f_x(P)t$; x = -1-2t. $y = y_0+f_y(P)t$; y = -1+t. $z = z_0+f_z(P)t$; z = 1+8t.

Q: Verify that the **following** function is convex on $\mathbf{R}^3$: $f(x,y,z) = 3x^2+2y^2+z^2+4x+5y+6z$. A: First we need to *calculate* the Hessian Matrix with the entries shown in yellow. $^{\delta f}/_{\delta x} = 6x+4$. $^{\delta^2 f}/_{\delta x^2} = 6$. And $^{\delta^2 f}/_{\delta x \delta y} = {}^{\delta^2 f}/_{\delta x \delta z} = 0$. $^{\delta f}/_{\delta y} = 4y+5$. $^{\delta^2 f}/_{\delta y^2} = 4$. And $^{\delta^2 f}/_{\delta y \delta x} = {}^{\delta^2 f}/_{\delta y \delta z} = 0$. $^{\delta f}/_{\delta z} = 2z+6$. $^{\delta^2 f}/_{\delta z^2} = 2$. And $^{\delta^2 f}/_{\delta z \delta x} = {}^{\delta^2 f}/_{\delta z \delta y} = 0$. Therefore, H = the **Green** Matrix.

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$\begin{bmatrix} \delta^2 f/\delta x^2 & \delta^2 f/\delta x\delta y & \delta^2 f/\delta x\delta z \\ \delta^2 f/\delta y\delta x & \delta^2 f/\delta y^2 & \delta^2 f/\delta y\delta z \\ \delta^2 f/\delta z\delta x & \delta^2 f/\delta z\delta y & \delta^2 f/\delta z^2 \end{bmatrix}$$

We now must test to see if all the *principle* minors of H are non-negative. If they are, then f(x,y,z) is a convex function. By deleting rows/columns 1 and 2, we obtain the **1st-order** principle minor 2, which is > 0. Similarly by deleting rows/columns 1 and 3, we get 4 > 0, and by deleting rows/*columns* 2 and 3, we get 6 > 0. By deleting row/column 1, we find the 2nd order principle minor, $\det[{}^{4}{}_{0}{}^{0}{}_{2}] = 8 > 0$. Row/column 2 gives $\det[{}^{6}{}_{0}{}^{0}{}_{2}] = 12 > 0$. And row/column 3 gives $\det[{}^{6}{}_{0}{}^{0}{}_{4}] = 24 > 0$. The **3rd-order principle minor** is the determinant of H itself, worked out expanding by *column 1* to be 48 > 0. Since for all (x,y,z), all principle minors of the Hessian are non-negative, f(x,y,z) **is** a convex function on $\mathbf{R}^3$.

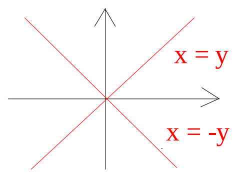
➢ 4th October 1999

When do Directional Derivatives Exist?

If f'(c;u) exists in *every* direction, then the partial derivatives exist, but things go **wrong** for the converse. Limits of continuity in *many* variables (see T&F as well). Consider f: $\mathbf{R}^2 \to \mathbf{R}$, where f(x,y) = x+y if x = 0 or y = 0, and 1 otherwise. Then $\left.{}^{\delta f}/_{\delta x}\right|_0 = 1$ and $\left.{}^{\delta f}/_{\delta y}\right|_0 = 1$. But $f(\underline{0}+t\underline{u})-t(0)/t$ for $u = (u_1,u_2)$ $(u_1, u_2 \neq 0) = {}^1/_t$, which does not have a *limit* as $t \to 0$.

Consider **S** to be open connected, $S \subseteq \mathbf{R}^2$. Consider f: $S \to \mathbf{R}^m$. For m = 1, $\lim_{x\to x_0} f(x) = L$ means "to get values of f(x) near to L, we need only take *values* of x near x_0", i.e. given $\varepsilon > 0$, there exists a $\delta > 0$ such that if $|x-x_0| < \delta$, then $|f(x)-L| < \varepsilon$. Definition (in general): $\lim_{x\to x_0} f(x) = L$ if given $\varepsilon > 0$, there exists a $\delta > 0$ such that if $\|x-x_0\| < \delta$ and $x \in S$, then $\|f(x)-L\| < \varepsilon$.

Example: consider $f(x,y) = {}^{x^2-y^2}/_{x^2+y^2}$ for $x \neq 0$, and 0 for $x = \underline{0}$. Does $\lim_{x\to 0} f(x)$ exist? (i) On the lines x = y, x = -y, f is **zero.** (ii) Along y = 0, f(x,0) = 1 if $x \neq 0$ and 0 if x = 0. (iii) Along x = 0, f(0,y) = -1 if $y \neq 0$, and 0 if y = 0. Clearly, if I $= \lim_{x\to x_0}$ exists, it **must** be at the same time 0, 1 and -1.



Try $f(x,y) = 2x^2y/(x^4+y^2)$, $(x,y) \neq (0,0)$ Does f(x) tend to *something* as $\underline{x}$ tends to zero? Along x = 0 and y = 0, the limit is zero. But along $y = kx^2$ with non-zero k, work it out! Why does the method *work*? We want to see if there is an **L**, with $f(\underline{x})$ tending to L as $\underline{x}$ tends to $\underline{x}_0$ (*). A curve is a *continuous* x(t), $t \in [0,r]$, with $\lim_{t\to 0} f(x(t)) = x_0$. We **test** to see if $\lim_{t\to 0} f(\underline{x}(t))$ exists. If it doesn't, we stop, because (*) cannot hold for *all* L.

If $\lim_{x\to x_0} f(x) = L$, then $\lim_{t\to 0} f(x(t)) = L$ as well. If we **get** $\lim_{t\to 0} f(x(t)) = L$ (i.e. the limit exists and is L), then if $\lim_{x\to x_0} f(x)$ exists, it *must* be L. But if x_1 and x_2 are two curves through x_0, and $\lim_{t\to 0} f(x_1(t)) = L_1$ and $\lim_{t\to 0} f(x_2(t)) = L_2$; if L_1 is **not** equal to L_2, then $\lim_{x\to x_0} f(x)$ *cannot* exist!

➢ 5th October 1999

Continuity

Suppose that S is an **open** subset of $\mathbf{R}^m$. Let $\underline{x}_0 \in S$, f: $S \to \mathbf{R}^n$. Then we say that f is *continuous* at $\underline{x}_0$ if $\lim_{x\to x_0} f(x) = f(\underline{x}_0)$. We say that f is continuous on S if it is continuous at *all points* of S. If f: $S \to \mathbf{R}^n$, then $f(\underline{x}) = (f_1(\underline{x}), f_2(\underline{x}), \ldots, f_n(\underline{x}))$, and we can *specify* f by specifying the n **functions** $f_1,\ldots,f_n$ $(S \to \mathbf{R})$.

<u>FACT</u>: f is continuous (cts) on S iff all the f_i (i = 1...n) are cts and *real-valued functions*. In particular, the k^{th} co-ordinate function, p_k: $\mathbf{R}^n \to \mathbf{R}$, $p_k(\underline{x}) = x_k$, is cts (left as an exercise). Thinks: for n = 2, assume f_1 and f_2 are cts at $\underline{x}_0$. Given $\varepsilon > 0$, we need to find a $\delta > 0$ s.t. as *soon* as $\|x-x_0\| < \delta$, then we know that $\|f(\underline{x})-f(\underline{x}_0)\| < \varepsilon$. Taking the end result to *pieces*, we want to make $\|f(\underline{x})-f(\underline{x}_0)\| = ((f_1(\underline{x})-f_1(\underline{x}_0))^2+(f_2(\underline{x})-f_2(\underline{x}_0))^2)$ nice and small.

Pick $\delta_1 > 0$ so that if $|\underline{x}-\underline{x}_0| < \delta_1$, then $\|f_1(x)-f_1(x_0)\| < k\varepsilon$ (k is chosen later, and we can do this because f_1 is cts). Pick $\delta_2 > 0$, so that if $|\underline{x}-\underline{x}_0| < \delta_2$, then $|f_2(\underline{x})-f_2(\underline{x}_0)| < k\varepsilon$. If $\|\underline{x}-\underline{x}_0\| < \min(\delta_1,\delta_2)$, then $\|f(\underline{x})-f(\underline{x}_0)\| < ((k^2\varepsilon^2)+(k^2\varepsilon^2))^{1/2}$. Pick $k = {}^1/_{\sqrt{2}}$ and $\delta = \min(\delta_1,\delta_2)$. Now carry on.

FACTS: (1) If f: $S\to\mathbf{R}^n$, g: $T\to\mathbf{R}$, and $T\subseteq\mathbf{R}^n$ are continuous, then so is gf: $S_n f^{-1}(T)\to\mathbf{R}^p$. dom(gf) = the set of points for *which* gf(S) makes sense = $\{s \mid s\in S \text{ and } f(S)\in T\}$ ($s\in S$ so f(s) makes sense and $f(S)\in T$ so gf(S) is defined). Recall the *notation* $f^{-1}(T)=\{s \mid f(s)\in T\}$. Note: $f^{-1}(T)$ will be open if f is cts and T is open; and the *intersection* of 2 open sets is open. (2) If ($\underline{x}_n$) is a *sequence* of points in S, then $\lim_{n\to\infty}\underline{x}_n$ exists, and is equal to a. If f is cts, then $\lim_{n\to\infty} f(\underline{x}_n) = f(\underline{n})$. (If this is the 1 *variable* case, replace |...| by ||...||).

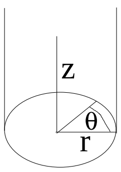
Back to differentiation. Single variable case: f: $S\to\mathbf{R}$, $S\subset\mathbf{R}^1$. S is *differentiable* at $c\in S$ if $\lim_{x\to c}\frac{f(x)-f(c)}{x-c}$ exists. ($x\in S$). The **value** of the *limit* is f'(c). For *small* (x-c), we have small $\frac{f(x)-f(c)}{x-c}$ - f'(c). So f(x)-f(c) = f'(c)(x-c) + a small error, which *tends* to 0 as (x-c) *tends* to 0. So $f(x)-f(c) = f'(c)(x-c) + e_c(x)(x-c)$, and $e_c(x)\to 0$ as $x\to 0$. *Rewrite*, with x = c+h, so that $f(c+h) = f(c)+f'(c)h+E_c(h)|h|$.

➢ 11th October 1999

Suppose that f: $S\to\mathbf{R}^n$, with $S\subseteq\mathbf{R}^m$ (an open subset). Let $\underline{c}\in S$. Pick an r > 0 *such that* $B(\underline{c};r)\subseteq S$ (for *safety*). Then $V\in\mathbf{R}^m$ *such* that ||v|| < r. **Definition**: We say that f is *differentiable* at $\underline{c}$ if there is a linear transformation T: $\mathbf{R}^m\to\mathbf{R}^n$, such that $f(\underline{c}+\underline{v}) = f(\underline{c})+T_c(\underline{v})+\|\underline{v}\|E_c(\underline{v})$, *where* $E_c(\underline{v})\to 0$ as $\underline{v}\to\underline{0}$. (*Given* $\varepsilon > 0$, there exists a $\delta > 0$ such *that* if $\|\underline{v}\| < \delta$, then $E_c(\underline{v}) < \varepsilon$). T_c is called the *total derivative* of t at c. Look at $\underline{v} = t\underline{u}$ for a *fixed* direction $\underline{u}$: $f(\underline{c}+t\underline{u}) = f(c)+T_c(t\underline{u})+|t|(\|\underline{u}\|)E_c(t\underline{u}) = f(c)+tT_c(\underline{u})+|t|(\|\underline{u}\|)E_c(t\underline{u})$. So *assuming* n = 1, $\lim_{t\to 0}(f(\underline{c}+t\underline{u})-f(c)/t) = \lim_{t\to 0}(T_c(\underline{u})\pm\|\underline{u}\|E_c(t\underline{u})) = T_c(\underline{u})+0$. **But** the LHS is $f'(\underline{c};\underline{u})$, the *directional derivative* of f at $\underline{c}$ in the direction of $\underline{u}$.

Recall: If L: $\mathbf{R}^m\to\mathbf{R}^n$, pick the *usual* bases $(e_k) = (0,\ldots,1,\ldots,0)$ (1 in the k^{th} place) in $\mathbf{R}^m$ and $\mathbf{R}^n$, so that we can *represent* L as a matrix $\underline{v}\in\mathbf{R}^m$, with $\underline{v} = \Sigma_{j=1}^m v_j\underline{e}_j^m$. Therefore, $\underline{v} = (v_1,\ldots,v_m)$. Now $L(\underline{v}) = L(\Sigma_{j=1}^m v_j e_j^m) = \Sigma_{j=1}^m v_j L(\underline{e}_j^m)$. But $L(\underline{e}_j^m) = \Sigma_{k=1}^n l_k\underline{e}_k^n$. (The **matrix** $n(l_{jk})$, where *this* holds). **Going** back to f: $S\to\mathbf{R}^n$, $T_c(\underline{e}_j^m) = (f_1(\underline{c};\underline{e}_j^m)), (f_2(\underline{c};\underline{e}_j^m)), \ldots, (f_n(\underline{c};\underline{e}_j^m)) = ((\delta f_1/\delta x_1)c),\ldots,(\delta f_n/\delta x_j)(c))$. If f is *differentiable* at c, then w.r.t. the *usual* bases, T_c has matrix $(\delta f_i/\delta x_j)$, i.e. the *Jacobian* matrix of $\underline{f}$ (at c). If f: $S\to\mathbf{R}$ (S in $\mathbf{R}^m$), the T_c is given by $T_c(\underline{u}) = f'(c;u) = \nabla f(c){:}\underline{u}$ $(\delta f/\delta x_1,\ldots,\delta f/\delta x_m)$.

➢ 12th October 1999



Example: Find the derivative of f at c (in the direction of $\underline{u}$), where $f(x,y,z) = x^3-xy^2-z$, $\underline{u} = (2,-3,6)$ & $\underline{c} = (1,1,0)$. $\nabla f = (3x^2-y^2, -2xy, -1)$. The *direction* of $\underline{u}$ is $\underline{u}/\|\underline{u}\|$. $\|\underline{u}\| = \sqrt{49} = 7$, so $\underline{u} = (^2/_7, -^3/_7, ^6/_7)$. $\nabla f(c) = (2,-2,1)$; $f'(c;u_{\text{normalised}}) = (2,-1,1).(^2/_7,-^3/_7,^6/_7) = ^4/_7+^6/_7+^6/_7$. **Example**: Let f: $\mathbf{R}^3\to\mathbf{R}^3$; $f(r,\theta,z) = (r\cos\theta, r\sin\theta, z)$. The **jacobian** matrix is as shown in *yellow*.

$$\begin{bmatrix}\cos\theta & -r\sin\theta & 0\\ \sin\theta & r\cos\theta & 0\\ 0 & 0 & 1\end{bmatrix}$$

The Chain Rule

Let f and g be *functions*. Assume h = f¤g is defined in a neighbourhood of a point $\underline{a}$. Assume g is differentiable at $\underline{a}$, with *total derivative* $g'(\underline{a})$ (a linear transformation). Set $\underline{b} = g(\underline{a})$. Assume f is differentiable at $\underline{b}$ with *total derivative* $f'(\underline{b})$ (a linear tranformation). Then h is **differentiable** at $\underline{a}$; it's total derivative is $f'(\underline{a}) = f'(\underline{b})$¤$g'(\underline{a})$, the **composite** linear transformation. If we fix co-ordinates, and write $Df(\underline{b})$ for the **Jacobian** matrix of f at $\underline{b}$, etc., we get $Df(\underline{a}) = Df(\underline{b})$¤$Dg(\underline{a})$.

Example: g: $\mathbf{R}\to\mathbf{R}^2$, f: $\mathbf{R}^2\to\mathbf{R}$. g(t) = (x(t), y(t)) in $\mathbf{R}^2$. f(x,y) in $\mathbf{R}$. $D_g(\underline{a}) = ({}^{dx/dt}{}_{dy/dt})_a$ (*a column* vector). $D_f(\underline{b}) = ({}^{\delta f}/_{\delta x}, {}^{\delta f}/_{\delta y})$. So h(t) = f(x(t), y(t)), ${}^{dh}/_{dt} = {}^{\delta f}/_{\delta x}{}^{\delta x}/_{\delta t} + {}^{\delta f}/_{\delta y}{}^{\delta y}/_{\delta t}$.

Tutorial Exercises

Q: For the following, give the **domain** (open or closed?); whether the function has a limit at (0,0); and decide whether the function could be made *continuous* at the origin. (a) ${}^{y^2}/_{(x^2+y^2)}$. (b) ${}^{(x-y)}/_{(x+y)}$. (c) ${}^{(x^3+y^3)}/_{(x^2+y^2)}$. (d) ${}^{(x+y)}/_{(x^2+y^2)}$. (e) ${}^{[\sin(x+y)]}/_{(x+y)}$. A: (a) Makes **sense** provided $x^2+y^2 \neq 0$. So the domain is $D_f = \mathbf{R}^2\backslash\{(0,0)\}$, which is an *open* set. The function cannot have a limit as $(x,y)\to(0,0)$, since going to (0,0) along the line x = y gives a limit of a ½, and going along x = 2y gives a limit of ${}^1/_5$. If it **did** have a limit *l*, that limit would be the answer along <u>any</u> path through the origin. In addition, the function *cannot* be made cts at the origin.

(b) Causes problems where x = -y. So $D_f = \{(x,y) \mid x \neq -y\}$. This is an **open** set (visualise the drawing). Along x = 0, the function has constant value -1. Along y = 0, it has value 1. So **there is no** limit; and it cannot be made continuous at the origin.

(c) *Goes* wrong at $x^2+y^2 = 0$ i.e. at (0,0) only. So $D_f = \mathbf{R}^2\backslash\{(0,0)\}$, open. Trying different paths and getting 0 suggests that the limit *should* be zero. Recall the definition of $\lim_{(x,y)\to(a,b)}f(x,y) = L$: Given $\varepsilon > 0$, there *exists a* $\delta > 0$ s.t. $\|(x,y)-(a,b)\| < \delta \Rightarrow |f(x,y)-L| < \varepsilon$. Convert the expression for f to polar **co-ordinates**. Using $x = r\cos\theta$ and $y = r\sin\theta$, f(x,y) becomes ${}^{r^3(\cos^3\theta+\sin^3\theta)}/_{r^2}$. (As $\|(x,y)\| = r$, we want to make this small when r is small). Given $\varepsilon > 0$, take $\delta = {}^{\varepsilon}/_2$, then $|f(x,y) - 0| = r|\cos^3\theta+\sin^3\theta| < 2r$. If $\|(x,y)\| < \delta$, then $|f(x,y)| < 2\|(x,y)\| \leq \varepsilon$, i.e. $\lim_{(x,y)\to\underline{0}}f(x,y) = 0$. Set f(0,0) = 0 to get a *continuous* function anywhere on $\mathbf{R}^2$.

(d) *Answer*: $D_f = \mathbf{R}^2\backslash\{(0,0)\}$; the limit does not exist as different paths give different "limits" — most tend to **infinity**. (e) As ${}^{\sin(u)}/_u\to0$ as $u\to0$, this has a *removable* discontinuity — set f(0,0) = 0 and you get a continuous function on $\mathbf{R}^2$.

Q: Which of the **following** functions is continuous at the origin?: (a) $f(x,y) = {}^{xy}/_{(x^2+y^2)}$ if $(x,y) \neq (0,0)$; and f(x,y) = 0 if (x,y) = (0,0); (b) $f(x,y) = {}^{(x^3-y^3)}/_{(x^2+y^2)}$ if $(x,y) \neq (0,0)$, and f(x,y) = 0 if (x,y) = (0,0). A: (a) Along x = y, this gives a *limit* of a ½ as $(x,y)\to(0,0)$ (along **other** curves, you get different limits). So as ½ ≠ 0, f is not continuous at (0,0). (b) $\lim_{(x,y)\to(0,0)} f(x,y) = 0$ (similar to question 1c), so it is **continuous** at (0,0).

Q: Determine whether or not the *following function* is continuous everywhere in $\mathbf{R}^2$: $f(x,y) = {}^{(x^3-y^3)}/_{(x-y)}$ for $x \neq y$, and $f(x,y) = 3x^2$ for x = y. A: If $x \neq y$, $f(x,y) = {}^{x^3-y^3}/_{x-y} = x^2+xy+y^2$. This is cts *everywhere* in $\mathbf{R}^2\backslash\{(\alpha,\alpha) \mid \alpha \in \mathbf{R}\}$, so we need only check *continuity* on this "diagonal". $\lim_{(x,y)\to(a,a)} f(x,y) = 3a^2$. This is a claim. Sketch Proof: From the definition, we need to check that if given an $\varepsilon > 0$, there exists a $\delta > 0$, etc. Consider the first few "*obvious*" lines, and then split to consider the **two** cases x = y and $x \neq y$ separately.

Q: In the **discussion** of continuity of a function f, (f: $S\to\mathbf{R}^n$), it says "f is cts if it is cts at all points of S" and "f is cts at a *element* S if $\lim_{x\to a} f(x) = f(a)$. This *definition* of continuity uses the notion of a limit, so a reminder: $\lim_{x\to a} f(x) = f(a)$ means: Given $\varepsilon > 0$, there's a $\delta > 0$ s.t. if $|x-a| < \delta$, then $|f(x)-f(a)| < \varepsilon$.

For a **different** a, we may need a different δ. Example: $f(x) = {}^1/_x$ for $x \in \mathbf{R}\backslash\{0\}$. Fix $\varepsilon = 10$ and test for *the continuity* of f at each of the points $a_n = {}^1/_n$. ($n \in \mathbf{N}\backslash\{0\}$). In each case, give a suitable $\delta > 0$. Can you use a single δ for all a_n? In general, functions that are just *continuous* can have awkward properties. The stronger notion of uniform continuity is therefore introduced. In this, the SAME δ works for all points. Check that $f(x) = x^2$ for $x \in [-1,1]$ is *uniformly* continuous.

Q: **Generalising** the above to two variables, (i) Prove first that the function $f(x,y) = 5x+xy$ is cts at any (a,b) in $\mathbf{R}^2$; (ii) Prove that the function above is *uniformly* continuous in the square region $[0,1]\times[0,1]$, that is where $0\le x\le 1$ and $0\le y\le 1$.

A: Suppose we could **find** a δ which worked for all the a_n. Any smaller δ would therefore also work, so we can find **some** n with ${}^1/_n < \delta$, and we can assume that $n > 100$, say. Now **take** $x = {}^1/_n$, $|x-a_{2n}| = |{}^1/_n-{}^1/_{2n}| = {}^1/_{2n} < \delta$. But $|f(x)-f(a_{2n})| = |n-2n| = |n| > 100$. So **NO WAY** can we get the same δ everywhere. For $f(x) = x^2$, look at the *usual* direct proof of continuity. Given $\varepsilon > 0$, check the continuity at $a \in [-1,1]$. $|x^2-a^2| = |(x-a)(x+a)| \le 2|x-a|$ (since $|x+a| \le 2$ if $x,a \in [-1,1]$). So to get $|x^2-a^2| < \varepsilon$, we need only take $|x-a| < {}^\varepsilon/_2$ (which does *not* depend on a) so we have: f is **uniformly** continuous *everywhere* in [-1,1].

(b) Try the **same** approach. We are given $\varepsilon > 0$. Check for *continuity* at $(a,b) \in \mathbf{R}^2$. Examine: $|f(x,y)-f(a,b)| = |5(x-a)+(xy-ab)| \le 5|x-a|+|xy-ab|$. (*Triangle* inequality). (Note: if $\|(x,y)-(a,b)\|$ is small, we can *make* $5|x-a|$ small (it will be anyway!) so we have to **concentrate** on (xy-ab)). Now $|xy-ab| = |x(y-b)+b(x-a)| \le |x||y-b|+|b||x-a|$. (**Same** old trick — take it off and add it in again). If $\|(x,y)-(a,b)\|$ is small, say $< \delta_1$, then $|x-a| < \delta_1$.

(*) So $|x| < |a|+\delta_1$. Also, $|y-b| < \delta_1$ so $|xy-ab| \le \delta_1(|a|+\delta_1)+|b|\delta_1$. This can now be *made* smaller then ${}^\varepsilon/_6$, say, by picking δ_1 to be small enough (the *details* don't matter!) If we pick δ to be smaller than this δ_1, and also smaller than ${}^\varepsilon/_6$, we wil get $|f(x,y)-f(a,b)| < {}^5/_6\varepsilon+{}^\varepsilon/_6 = \varepsilon$. (But δ_1 will depend on (a,b) in *general* here!).

(ii) Now if we *restrict our* attention to $(x,y) \in [0,1]\times[0,1]$, then in *line* (*) above, $|x| < 1$ and $|b| < 1$. So we **have** $|xy-ab| \le 2\delta_1$, so we can choose δ_1 so that $2\delta_1 < {}^\varepsilon/_6$ *independently* of (a,b). (This is not the "**full**" elegant solution). Fact: **any** continuous function on a closed bounded subset of $\mathbf{R}^m$ is *uniformly continuous.*

➢ 18th October 1999

(2) Let $g(t) = (x(t), y(t), z(t))$; $g: S \to \mathbf{R}^3$; (*where* S is a subset); $f: \mathbf{R}^3\to\mathbf{R}$; h is a *composite:* $f(x(t), y(t), z(t)) = h(t)$. Then ${}^{dh}/_{dt} = {}^{\delta f}/_{\delta x}{}^{dx}/_{dt} + {}^{\delta f}/_{\delta y}{}^{dy}/_{dt} + {}^{\delta f}/_{\delta z}{}^{dz}/_{dt}$.

(3) Let f: f(x,y,z); $f: \mathbf{R}^3\to\mathbf{R}$; $x = x(r,s)$, $y = y(r,s)$, $z = z(r,s)$. **Write** $g(r,s) = (x(r,s), y(r,s), z(r,s))$. Let $g: \mathbf{R}^2\to\mathbf{R}^3$. Now look at the *definitions* for $D_g(\underline{a})$ on the right and $D_f(\underline{b}) = ({}^{\delta f}/_{\delta x}, {}^{\delta f}/_{\delta y}, {}^{\delta f}/_{\delta z})$. Let $h = f\cdot g: \mathbf{R}^2\to\mathbf{R}$. Then $D_h(\underline{a}) = ({}^{\delta h}/_{\delta r}, {}^{\delta h}/_{\delta s})$; $D_h(\underline{a})$ = see the *left*. So ${}^{\delta h}/_{\delta r} = {}^{\delta f}/_{\delta x}{}^{\delta x}/_{\delta r} + {}^{\delta f}/_{\delta y}{}^{\delta y}/_{\delta r} + {}^{\delta f}/_{\delta z}{}^{\delta z}/_{\delta r}$. **Similarly** for ${}^{\delta h}/_{\delta s}$.

$$D_g(a) = \begin{bmatrix} \frac{\delta x}{\delta r} & \frac{\delta x}{\delta s} \\ \frac{\delta y}{\delta r} & \frac{\delta y}{\delta s} \\ \frac{\delta z}{\delta r} & \frac{\delta z}{\delta s} \end{bmatrix}$$

$$\left[\frac{\delta f}{\delta r}\ \frac{\delta f}{\delta y}\ \frac{\delta f}{\delta z}\right] \begin{bmatrix} \frac{\delta x}{\delta r} & \frac{\delta x}{\delta s} \\ \frac{\delta y}{\delta r} & \frac{\delta y}{\delta s} \\ \frac{\delta z}{\delta r} & \frac{\delta z}{\delta s} \end{bmatrix}$$

Inverse Functions (1 Variable Case)

Let f be a function, f: S→**R**, S ⊆ **R** (open). Consider a c ∈ S such that f'(x) *exists* and is continuous in a neighbourhood of c. Assume that f'(c) ≠ 0. Then if f'(c) > 0, f is **monotonically** increasing near c. If f'(c) < 0, f is monotonically *decreasing* near c. If x > c and f'(c) > 0, (we need f(x) ≥ f(c)), **look** at f(x)-f(c) = f'(c)(x-c) + small error(|x-c|). (This small error *tends* to 0 as x→c). So as f'(c) > 0 and x-c > 0, f(x) will be greater than f(c). (for x *near enough* to c). Similarly for c > x. And **similarly** for f'(c) < 0: implies *monotonically decreasing*.

Thus near c, f will have a (**locally** defined) inverse function. If y = f(x), we can find a function g such that g(y) = x. *Now* f(x)-f(c) = f'(c)(x-c) + small error. So (x-c) = $^{f(x)-f(c)}/_{f'(c)}$ + small error. Write y = f(x) and b = f(c), so that g(y) = x; g(b) = c; g(y)-g(b) = $^{1}/_{f'(c)}$(y-b) + small error. So g'(b) exists and *equals* $^{1}/_{f'(c)}$ (**Detailed** argument omitted).

Note: As g is a (*local*) inverse for f near c, f·g = Id *near* f(c) = b, and g·f = Id *near* c. The **chain** rule gives g'(b) = $^{1}/_{f(c)}$ anyway. (I = D+D_g).

➢ 19th October 1999

Can we **generalise the** single variable case to *the many* variable case? Suppose that S ⊆ $\mathbf{R}^m$ (open); f = (f_1,...,f_n); S→$\mathbf{R}^n$; a ∈ S; and all $\delta f_i/\delta x_j$ are continuous on some *open set* containing a. Then f(x)-f(a) = f'(a)(x-a) + small error. (Note: f'(a) = D_f(a)). We write y = f(x); b = f(a) as *before*. Then y-b = D_f(a)(x-a) + small error (*****).

Think of it as (y-b) = T(x-a), with T *linear*. We can solve it **provided** det(T) ≠ 0, i.e. as long as T is invertible. Oops! — we will need n = m!. We therefore can hope to *solve* (*****) for x in terms of y provided that det D_f(a) ≠ 0, and then x will be **determined** by y locally near y = b.

Notation: *The Jacobian* determinant is J_f(a) = det(D_f(a)). Could also **use** J_f(x) = det(D_f(x)). J_f(x) will be a sum of *products* of the ($\delta f_i/\delta x_j$)(x)'s. Hence J_f(x) is a **continuous** function of x on some *open* neighbourhood of a. If J_f(a) ≠ 0, then J_f(x) ≠ 0 in a **neighbourhood** around a.

Consequence: f must be 1-to-1 near a. Justification: because f(x) = f(x'); 0 = f(x)-f(x') = D_f(c)(x-x'), [Using a **many** variable version of the *Mean Value Theorem* — details not given], where c lies **between** x and x'. So D_f(c) will be *invertible* provided all of x, x' and c are near **enough** to a. But then 0 = D_f(c)(x-x'), and so x-x' = 0, i.e. x = x'. Then we **expect** to have a (*locally* defined) inverse g of f near a: D_g(b) = $D_f(a)^{-1}$, where b = f(a). **Note**: the *notation* $\delta(f_1,...,f_m)/\delta(x_1,...,x_m)$ is sometimes used for J_f(x).

Curvilinear Co-ordinates

We will look at *Polar co-ordinates in* $\mathbf{R}^2$, *cylindrical co-ordinates in* $\mathbf{R}^3$, *and spherical co-ordinates in* $\mathbf{R}^3$. **Polar Co-ordinates.** (x,y)→(r,θ) [cartesian→polar]. Here, x = rcosθ and y = rsinθ. Or, r = $\sqrt{(x^2+y^2)}$ with r > 0, and θ = $\tan^{-1}(^{y}/_{x})$ with 0<θ<π.

Workshop 3

Recall that a function $f(x_1,...,x_n)$ is a convex function on a *convex set S* (in $\mathbf{R}^n$) if for any $x' \in S$ and $x'' \in S$, $f(cx'+(1-c)x'') \le cf(x') + (1-c)f(x'')$ **holds** for $0\le c\le 1$. Also, a **subset** $X \subseteq \mathbf{R}^n$ is convex if for any *pair of points* $\underline{x},\underline{y} \in X$ and any t with $0\le t\le 1$, then $t\underline{x}+(1-t)\underline{y} \in X$ (Thus X is **convex** if the line segment joining any *two points of X* is completely within X).

Recall that the **Hessian**, H, of $f(x_1,...,x_n)$, is the n×n matrix whose ij_{th} entry is $\delta^2 f/\delta x_i \delta x_j$. $H(\underline{c})$ is the Hessian **evaluated** at a point $\underline{c}$. With this, **Taylor's** theorem implies that $f(\underline{x}) = f(\underline{x}_0) + D_f(\underline{x}_0).(\underline{x}-\underline{x}_0) + {}^1/_{2!}(\underline{x}-\underline{x}_0)^T H(\underline{c})(\underline{x}-\underline{x}_0)$ for some $\underline{c}$ *between* $\underline{x}$ and $\underline{x}_0$. (Note: **additional** details such as differentiability, etc. on X, X being convex are needed. These are *omitted*).

Winston has the following theorem: (assuming X is *convex*): Suppose $f(x_1,...,x_n)$ has continuous second-order partial derivatives for each **point** $x = (x_1,...,x_n) \in X$. Then $f(x_1,...,x_n)$ is a convex function on X iff for each $x \in X$, all **principal minors** of H are non-negative.

This condition on the *principal minors* is more simply stated using the notion of a +ve (**semi**) **definite** matrix: A *real symmetric matrix A* is positive semi definite if $\underline{x}^T A\underline{x} \ge 0$; it is positive **definite** if $\underline{x}^T A\underline{x} > 0$. With this, the above **result** states: Suppose $f(x_1,...,x_n)$ has continuous second-order partial derivatives for each *point* $x = (x_1,...,x_n) \in X$. Then $f(x_1,...,x_n)$ is a convex function on X iff for each $\underline{c} \in X$, $H(\underline{c})$ is +ve semi definite.

Q: Suppose that X is a **convex** subset of $\mathbf{R}^1$ (is an interval) and that we have the function $f: X\to\mathbf{R}^1$. Suppose that $f''(x)$ exists for all $x \in X$. Prove that f is a *convex function* iff $\mathbf{f''(x) \ge 0}$ for all $x \in X$. Hints: look at *both* sides; use the **one** variable form of the Taylor expansion. Fiddle using $f''(x) \ge 0$.

Q: Determine whether *each of the following functions* is convex, concave or neither: (a) $f(x) = x^3$, $x \ge 0$; (b) $f(x) = x^3$, $x \in \mathbf{R}$; (c) $f(x) = {}^1/_x$, $x > 0$; (d) $f(x) = x^a$, $(0\le a\le 1)$, $x > 0$.

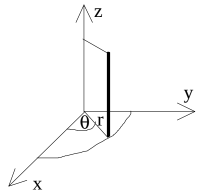
Q: Calculate the **Hessian**, $H(\underline{c})$, of $f(x,y) = x^4+2xy^2+y^2$. Find a **region** of the plane on which $H(\underline{c})$ is +ve *semidefinite* everywhere. (ii) Look at the above proof. Does it generalise to higher dimensions? Try to **prove** the previous stated theorem.

➢ 25th October 1999

Curvilinear Co-ordinates

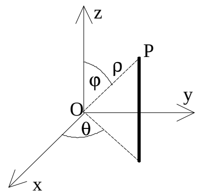
Set $g: S\to\mathbf{R}^2$; $S = \{(r,\theta) \mid 0<r, 0<\theta<2\pi\}$; $(r,\theta)\to(r\cos\theta, r\sin\theta)$. [So $f: \mathbf{R}^2\to\mathbf{R}$, say in x,y terms, becomes $f\bullet g$ in *terms* of (r,θ); g is the "*substitution*" for (x,y) in terms of (r,θ)]. Note: $D_g(r,\theta) = \begin{pmatrix}\cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta\end{pmatrix}$. All ${}^{\delta g}/_{\delta r}$ and ${}^{\delta g}/_{\delta\theta}$ are *continuous*. The Jacobian is $J_g(r,\theta) = \det(D_g(r,\theta)) = r > 0$ (where the r *comes* from $r\cos^2\theta+r\sin^2\theta$). So g is **locally** invertible.

Cylindrical Co-ordinates



We represent a *point* by (r, θ, z), where (1) (r,θ) is the expression for the *vertical* projection in polar co-ordinates; (2) z is the (old) **vertical** co-ordinate. Here, $x = r\cos\theta$, $y = r\sin\theta$, and z = z. Further, $r = \sqrt{(x^2+y^2)}$; $\theta = \tan^{-1}(y/x)$; and z = z.

Spherical Polar Co-ordinates



These are *represented* by (ρ, φ, θ), where ρ is the *distance* to P from O; φ is the *distance* between the line OP and the **positive** z-axis; and θ is as in cylindrical co-ordinates. Here, $x = r\cos\theta$, $y = r\sin\theta$, and z = z. (**These** come from r, θ, z). Also, $r = \rho\sin\varphi$; θ = θ; and $z = \rho\cos\varphi$. (**These** come from ρ, φ, θ.) And $\rho = \sqrt{(r^2+z^2)} = \sqrt{(x^2+y^2+z^2)}$. Therefore, $x = \rho\sin\varphi$; $y = \rho\sin\varphi\sin\theta$; and $z = \rho\cos\varphi$.

An Example on the Use of Change in Co-ordinates

Zaplace's equation in two dimensions is $\nabla^2 f = 0$, i.e. $\frac{\delta^2 f}{\delta x^2}+\frac{\delta^2 f}{\delta y^2} = 0$. What does this look like in *polars*? Now $\frac{\delta f}{\delta x} = \frac{\delta f}{\delta r}\frac{\delta r}{\delta x} + \frac{\delta f}{\delta \theta}\frac{\delta \theta}{\delta x}$, and *similarly* for $\frac{\delta f}{\delta y}$. Differentiate $x^2+y^2 = r^2$ w.r.t. x: $2x = 2r\frac{\delta r}{\delta x}$, so $\frac{\delta r}{\delta x} = \frac{x}{r} = \cos\theta$. Also, $\frac{\delta r}{\delta y} = \sin\theta$. **Now** $\frac{y}{x} = \tan\theta$ implies $\sec^2\theta\frac{\delta\theta}{\delta x} = -\frac{y}{x^2}$. Or $\frac{\delta\theta}{\delta x} = -\frac{y}{x^2}\cos^2\theta = -\frac{\sin\theta}{r}$. *Similarly,* $\frac{\delta\theta}{\delta y} = \frac{\cos\theta}{r}$, so that $\frac{\delta f}{\delta x} = \cos\theta\frac{\delta f}{\delta r} - \frac{\sin\theta}{r}\frac{\delta f}{\delta\theta}$.

Since f is *arbitrary,* $\frac{\delta}{\delta x} = \cos\theta\frac{\delta}{\delta r} - \frac{\sin\theta}{r}\frac{\delta}{\delta\theta}$. **So** $\frac{\delta^2 f}{dx^2} = ((\cos\theta\frac{\delta}{\delta r} - \frac{\sin\theta}{r}\frac{\delta}{\delta\theta}) \mid \cos\theta\frac{\delta f}{\delta r} - \frac{\sin\theta}{r}\frac{\delta f}{\delta\theta}) = \cos\theta\frac{\delta}{\delta r}(\cos\theta\frac{\delta f}{\delta r}) + \cos\theta\frac{\delta}{\delta r}(-\frac{\sin\theta}{r}\frac{\delta f}{\delta\theta}) - \frac{\sin\theta}{r}\frac{\delta}{\delta\theta}(\cos\theta\frac{\delta f}{\delta r}) - \frac{\sin\theta}{r}\frac{\delta}{\delta\theta}(-\frac{\sin\theta}{r}\frac{\delta f}{\delta\theta}) = \cos^2\theta\frac{\delta^2 f}{\delta r^2} - \sin\theta\cos\theta[-\frac{1}{r^2}\frac{\delta f}{\delta\theta}+\frac{1}{r}\frac{\delta^2 f}{\delta r\delta\theta}] + \ldots = \cos^2\theta\frac{\delta^2 f}{\delta r^2} + \frac{\sin^2\theta}{r^2}\frac{\delta^2 f}{\delta\theta^2} - \frac{2}{r}\sin\theta\cos\theta\frac{\delta^2 f}{\delta r\delta\theta} + \frac{2\sin\theta}{r^2}\cos\theta\frac{\delta f}{\delta\theta} + \frac{\sin^2\theta}{r}\frac{\delta f}{\delta r}$.

Similarly for $\frac{\delta^2 f}{\delta y^2}$, and then we *use* $\sin^2\theta+\cos^2\theta = 1$. This *gives* $\boldsymbol{\nabla^2 f = \frac{\delta^2 f}{\delta r^2} + \frac{1}{r^2}\frac{\delta^2 f}{\delta\theta^2} + \frac{1}{r}\frac{\delta f}{\delta r}}$. As an application, if you are modelling a *phenomenon* which has circular symmetry, then f can be assumed to be <u>independent</u> of θ. So $0 = \nabla^2 f = \frac{d^2 f}{dr^2}+\frac{1}{r}\frac{dr}{df}$; $r\frac{d^2 f}{dr^2}+\frac{df}{dr} = 0$, or $\frac{d}{dr}(r\frac{df}{dr}) = 0$. So $r\frac{df}{dr} = A$, a *constant;* $\frac{df}{dr} = \frac{A}{r}$, so $f(r) = A\ln(r)+B$ (where **B** is *another* constant).

➢ 26th October 1999

Two *useful things* left out of the syllabus: (1) **Taylor** Series: $f(x) = f(x_0) + \frac{df}{dx}(x_0)(x-x_0) + \frac{1}{2!}\frac{d^2 f}{dx^2}(c)(x-x_0)^2$, where the last term with c in it is the *error* term. **Generalising** to n-dimensions, $f(\underline{x}) = f(\underline{x}_0)+D_f(\underline{x}_0)(\underline{x}-\underline{x}_0)+\frac{1}{2}(\underline{x}-\underline{x}_0)^T H(c)(\underline{x}-\underline{x}_0)$, for some c lying *between* $\underline{x}$ and $\underline{x}_0$. (2) **Implicit Function Theorem:** Involves e.g. $x^2+y^2 = 1$; $y = \pm\sqrt{(1-x^2)}$.

Assignment 1 (Not Assessed)

Q: Find the *equation* of the tangent plane and the normal line to the quadric surface $3x^2+2y^2+z^2+4x+5y+6z = 3$ at the **point** P(-1,-1,1). A: $f_x = \frac{\delta f}{\delta x} = 6x+4$; $f_y = 4y+5$; $f_z = 2z+6$. At P, $f_x(P) = -2$; $f_y(P) = 1$; and $f_z(P) = 8$. So the **tangent** plane is $-2(x+1)+(y+1)+8(z-1) = 0$ or $-2x+y+8z-9 = 0$. (Note the use of $\nabla f(P)$ to get $\nabla f(P).(\underline{x}-\underline{x}_0) = 0$. Geometrically, $\nabla f(P)$ is a *normal direction* to the surface. This gives: ...). The **normal** line to the surface at P is $\underline{x} = \underline{x}_0+\nabla f(P)t$, i.e. $x = -1-2t$, $y = -1+t$; $z = 1+8t$.

Q: Verify that the *function* $f(x,y,z) = 3x^2+2y^2+z^2+4x+5y+6z$ is *convex* on $\mathbf{R}^3$. A: Same f as above. $\delta^2f/\delta x^2 = 6$; $\delta^2f/\delta x\delta y = 0$; $\delta^2f/\delta x\delta z = 0$; $\delta^2f/\delta y^2 = 4$; $\delta^2f/\delta z^2 = 2$. Other terms are *zero*. So $H = ({}^6 0_0\ {}^0 4_0\ {}^0 0_2)$. *Positive* values, diagonal matrix. So f is positive definite — from the description $x^T Hx > 0$ for all $\underline{x} \neq 0$. Positive **definite** and Semi **definite** matrices are very useful!

Q: For what *values* of λ is the following *function* convex on $\mathbf{R}^3$: $f(x,y,z) = -x^2-y^2-2z^2+\lambda xy$? A: The Hessian is $({}^{-2}\lambda_0\ {}^{\lambda}{-2}_0\ {}^0 0_{-4})$ and so will not be *positive definite* for any value of λ. Is the function *concave* somewhere?

➢ 1st November 1999

Section 2: Integration

Preliminaries

We are *given* [a,b]. A partition P of [a,b] is a **finite** set of points $P = \{x_0,...,x_n\}$, with $a = x_0 < x_1 <...< x_n = b$. The **norm** of P is the *largest* of the numbers $\Delta x_k = x_k-x_{k-1}$, and is written as |P|. A partition P^1 is *finer* than P (or is a refinement of P) **if** $P \subset P^1$ (so that P^1 has some *extra* points). Note: $(P \subset P^1) \Rightarrow (|P| \geq |P^1|)$.

We **have** (we are given) $\alpha: [a,b] \rightarrow \mathbf{R}$ which is *bounded* on [a,b]. (There exists an M such that $|\alpha(x)| \leq M$ for all $x \in [a,b]$). We write $\Delta\alpha^k = \alpha(x_k)-\alpha(x_{k-1})$ and $\Sigma_{k=1}^n \Delta\alpha_k = \alpha(b)-\alpha(a)$.

Definition of the Riemann-Stietjes Sum (R-S Sum) of f w.r.t. α

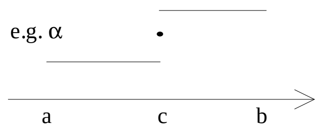
Suppose that P is a *partition* of [a,b] and that $t_k \in [x_{k-1}, x_k]$. A sum of the *form* $S(P, f, \alpha) = \Sigma_{k=1}^n f(t_k)\, \Delta\alpha_k$ is called an R-S sum *of f with respect to* α. We say that "f is integrable w.r.t. α on [a,b]" if there is a number **A** such that: for every $\varepsilon > 0$, there is a partition $P_\varepsilon \in \mathcal{P}$ [a,b], (where $\mathcal{P}$ is a curly P), the set of *partitions* of [a,b], such that if P is finer than P_ε, and for every *choice* of the t_k, then $|S(P,f,\alpha)-A| < \varepsilon$.

If such an *A exists,* it is unique, and we **write** $\int_a^b f d\alpha$. We also write $f \in \mathcal{R}(\alpha)$ if $\int_a^b f d\alpha$ exists. In the case that $\alpha(x) = x$, this is just the *Riemann integral*, $\int_a^b f dx$, and we *write* $f \in \mathcal{R}$.

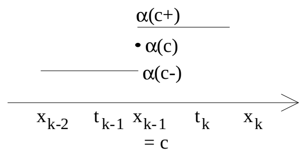
Properties: (1) If $f,g \in \mathcal{R}(\alpha)$ on [a,b], and if $c_1,c_2 \in \mathbf{R}$, then $c_1f+c_2g \in \mathcal{R}(\alpha)$ on [a,b] and $\int_a^b (c_1f+c_2g)d\alpha = c_1\int_a^b f d\alpha + c_2\int_a^b g d\alpha$. (2) If $f \in \mathcal{R}(\alpha)$ and $f \in \mathcal{R}(\beta)$, and we have $c_1,c_2 \in \mathbf{R}$, then $f \in \mathcal{R}(c_1\alpha+c_2\beta)$ and (of *course*) $\int_a^b f d(c_1\alpha+c_2\beta) = c_1\int_a^b f d\alpha + c_2\int_a^b f d\beta$. (3) If $a<c<b$, and if **two** of the three integrals below exist, then so *does the third*, and we have $\int_a^c f d\alpha + \int_c^b f d\alpha = \int_a^b f d\alpha$. (4) Integration by **Parts**. If $f \in \mathcal{R}(\alpha)$ on [a,b], and if $\alpha \in \mathcal{R}(f)$ on [a,b], then $\int_a^b f d\alpha + \int_a^b \alpha df = f(b)\alpha(b)-f(a)\alpha(a)$.

Reduction to Riemann Integrals: Suppose that $f \in \mathcal{R}(\alpha)$ on [a,b], and that α has a *continuous derivative* α' on [a,b], then $\int_a^b f(x)\alpha'(x)dx$ exists and *equals* $\int_a^b f d\alpha$. **Change of Variable:** Suppose that $f \in \mathcal{R}(\alpha)$ on [a,b], and that g is *strictly monotone* (strictly ⇑), then if $x > x'$, we have $g(x) > g(x')$. (*Similarly* for ⇓ and continuous defined *on* [c,d]). Now assume that $g(c) = a$ and that $g(d) = b$. Let $h(x) = f(g(x))$ and $\beta(x) = \alpha(g(x))$ for $x \in [c,d]$. Then $h \in \mathcal{R}(\beta)$ and $\int_c^d h d\beta = \int_a^b f d\alpha$.

What happens with *discontinuities*? A diagram is shown of an example of a step function. Here, for $a \le x < c$, we have $\alpha(a)$; for $x = c$ we have $\alpha(c)$, and for $c < x \le b$, we have $\alpha(b)$. **Assume we have** $f: [a,b] \to \mathbf{R}$, and at c, *one of f and* α is continuous from the left at c; and at **least** one is continuous from the right. Then $f \in \mathcal{R}(\alpha)$ on [a,b], and $\int_a^b f d\alpha = f(c)[\alpha(c_+)-\alpha(c_-)]$ (This is the case f = *continuous*).



Idea: Let $P \in \mathcal{P}[a,b]$, and assume that $c \in P$ w.l.o.g. (w.l.o.g. = *without loss of generality*). Then, $S(P,f,\alpha) = \Sigma f(t_k)\Delta\alpha_k = f(t_{k-1})[\alpha(c)-\alpha(c_-)] + f(t_k)[\alpha(c_+)-\alpha(c)]$. (See *diagram*). Now let $|P| \to 0$, so that $t_k \to c$ from the *right*, and $t_{k-1} \to c$ from the *left*. For simplicity, **assume that f is continuous**. As $|P| \to 0$, $f(t_k) \to f(c)$, and $f(t_{k-1}) \to f(c)$. So $S(P,f,\alpha) \to f(c)(\alpha(c_+)-\alpha(c_-))$.



➢ 8th November 1999

Define a **step** function. Let a be *defined* on [a,b] so that it is discontinuous at a finite number of points c_k. ($a \le c_1 < c_2 < \ldots, c_n \le b$, and 'a' is a *constant* on each subinterval (c_{k-1}, c_k)). In this case, α is called a step function. The **number** $\alpha(c_{k+})-\alpha(c_{k-})$ is called the *jump* at c_k. If $c_1 = a$, the **jump** at c_1 is $\alpha(c_{1+})-\alpha(c_1)$. If $c_n = b$, the jump at c_n is $\alpha(c_{n-})-\alpha(c_n)$.

Theorem: (Reduction of an R-S integral to a *finite* sum): If a is a step function on [a,b], with jump α_k at c_k, so that not *both* f and α are discontinuous from the **left** or from the **right** at c_k, then $\int_a^b f d\alpha$ exists, and equals $\Sigma_{k=1}^n f(c_k)\alpha_k$. Example: *Greatest* integer function: $\lfloor x \rfloor = n$ if $n \le x$, and n is the **largest** such $n \in \mathbf{Z}$. Suppose that $\Sigma_n^{k=1} a_k$ is a *finite* sum.

Define f on [0, n] by $f(x) = a_k$ if $k-1 < x \le k$, and $f(x) = 0$ if $x = 0$. Then $\Sigma_{k=1}^n a_k = \Sigma_{k=1}^n f(k) = \int_0^n f(x) d\lfloor x \rfloor$. **Application**: *Euler's summation formula*. If f has a *continuous* derivative f' on [a,b], then $\Sigma_{a<n\le b} f(n) = \int_a^b f(x)dx + (\int_a^b f'(x)((x))dx) + f(a)((a))-f(b)((b))$, where $((x)) = x-\lfloor x \rfloor$ is the *decimal* part function. When a and b are **integers**, this gives $\Sigma_{n=1}^b f(n) = \int_a^b f(x)dx + \int_a^b f'(x)(x-\lfloor x \rfloor-{}^1/_2)dx + {}^{f(a)-f(b)}/_2$.

Proof: Use *integration* by parts for f, and $\alpha(x) = x-\lfloor x \rfloor$. For example, $\Sigma_{k=1}^n 1/k^s = (1/n^{s-1}) + s\int_1^n (\lfloor x \rfloor/x^{s+1})dx$. ($s \neq 1$). So $\Sigma_{k=1}^n {}^1/_k = \ln(n) + C + O({}^1/_n)$ (where C is a *constant* and $O({}^1/_n)$ is "of order ${}^1/_n$"). This is very useful in *analytic* number theory.

Functions of Bounded Variation

Let $P \in \mathcal{P}[a,b]$, $f: [a,b] \to \mathbf{R}$, and write $\Delta f_k = f(x_k)-f(x_{k-1})$. If there is a *positive number M* such that $\Sigma |\Delta f_k| < M$ for all partitions of [a,b], then f is of **bounded variation** on [a,b]. For example, any *monotone* function is of Bounded Variation (BV) on [a,b]. Example: **define** $V_f(a,b) = {}^{\text{l.u.b.}}{}_{P \in \mathcal{P}[a,b]} (\Sigma |\Delta f_k|)$. The total *variation* of f on [a,b] is $V(x) = V_f(a,x)$ for $a<x<b$; $V(a) = 0$. Therefore, V(x) is **monotonically increaing** on [a,b] and V-f is *monotonically* increasing on [a,b]. So f is of BV on [a,b] iff f = g-h for two *monotonically increasing* functions on [a,b]. (Since f = V-(V-f)).

Workshop (Weeks 6-7)

Finding **maxima** of functions: non-linear programming. We will look at ways of finding maxima (or minima) of *functions of many variables*. The method is known as the method of steepest ascent. It is used a lot within **neural** networks to search for stable/turning points of an “energy” function which will give a *reasonable* solution to an optimisation problem. Unless the function is concave, then this will only be a local optimum, but the method DOES usually improve guesses or estimates.

The method uses the **Gradient:** given f: $\mathbf{R}^n \to \mathbf{R}$, we set $\nabla f(\underline{x}) = (\delta f/\delta x_1(\underline{x}), \ldots, \delta f/\delta x_n(\underline{x}))$ to be the *gradient* vector of f. We will usually **normalise** this to get its direction, $\nabla f(\underline{x})/\|\nabla f(\underline{x})\|$. Q: Find the **gradient** when $f(x_1,x_2) = x_1^2+x_2^2$, and hence the *direction* of $\nabla f(3,4)$. What does the *graph* of the curve $f(x_1,x_2) = f(3,4)$ look like? Represent the **direction** of $\nabla f(3,4)$ on it as a *vector* based at (3,4).

A: $\delta f/\delta x_1 = 2x_1$, $\delta f/\delta x_2 = 2x_2$. So $\nabla f(x) = (2x_1, 2x_2)$, with $\nabla f(3,4) = (6,8)$. $\|\nabla f(3,4)\| = \sqrt{(6^2+8^2)} = \sqrt{100} = 10$. So the **direction** is $({}^6/_{10}, {}^8/_{10})$. Now $f(x_1,x_2) = f(3,4)$ implies that $x_1^2+x_2^2 = 25$. This is a circle of *radius* 5, centred at the origin. The direction of $\nabla f(3,4)$ at (3,4) will be *perpendicular* to the **tangent** to the circle at that point.

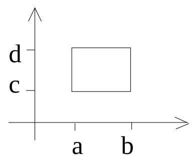
Investigate in general: If we move x a small length δ in the direction defined by a unit (column) vector d (so that x changes to $\underline{x}+\delta\underline{d}$), estimate the *change* in f, i.e. estimate $f(\underline{x}+\delta\underline{d})-f(\underline{x})$ in terms of ∇f, d and δ. **Show** that if $\nabla f(\underline{x}).\underline{d}/\|\nabla f(\underline{x})\| > 0$, then moving in the *direction of* δ will **increase** the value of $f(\underline{x})$, but if it is < 0, we can expect $f(\underline{x})$ to *decrease*.

➢ 9th November 1999

Theorem: (a) If f is *continuous* on [a,b], and $\alpha \in$ BV on [a,b], then $f \in R(\alpha)$ on [a,b]. (b) If f is *continuous* on [a,b], then $f \in R$ on [a,b]. (*Riemann* integrable). (c) If f is of BV on [a,b], then $f \in R$ on [a,b].

Interchanging the Order of Integration

Let $R = \{(x,y) \mid a \le x \le b, c \le y \le d\} = [a,b]\times[c,d]$. Assume α is of BV on [a,b]; β is of BV on [c,d]; and f is *continuous* on R. Set $F(y) = \int_a^b f(x,y)d\alpha(x)$, and $G(x) = \int_c^d f(x,y)d\beta(y)$. Then $F \in R(\beta)$ on [c,d] and $G \in R(\alpha)$ on [a.b] imply that $\int_a^b (\int_c^d f(x,y)d\beta(y))d\alpha(x) = \int_c^d(\int_a^b f(x,y)d\alpha(x))d\beta(y)$.

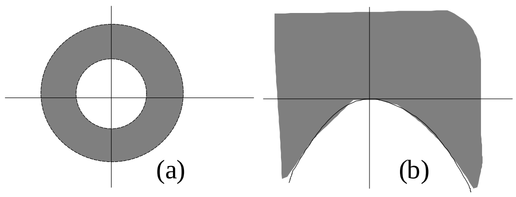


Measure. Let $S \in \mathbf{R}$. By a *Lebesgue* covering of S, we mean a collection $T = \{I_1,I_2,\ldots\}$ of open *intervals* so that $\cup^{\infty}_{k=1} I_R \supseteq S$. Let $L(I_k)$ be the *length* of I_k, and set $L(T) = \Sigma^{\infty}_{k=1} L(I_k)$, provided *this* is < ∞. Set $\bar{m}(S)$ = g.l.b.{L(T) | T is a *Lebesgue* covering of S}. This is called the **outer** Lebesgue measure of S. If $\bar{m} = 0$, S is said to be of *measure* zero.

Theorem (Lebesgue). If f is *defined* and *bounded* on [a,b]; and let D denote the set of **discontinuities** of f in [a,b], then $f \in \mathbf{R}$ on [a,b] iff $\bar{m}(D) = 0$. **Applications**: (1) If $f \in R$ on [a,b], then $f \in [c,d]$ for any *subinterval* $[c,d] \subset [a,b]$, |f| is R on [a,b], and $f^2 \in R$ on [a,b],... (2) If $f \in \mathbf{R}$ and $g \in \mathbf{R}$ on [a,b], then *provided* g is bounded away from O on [a,b], then $f/g \in R$ on [a,b].

Assignment 1

Q: For (a) $f(x,y) = \ln\{(25-x^2-y^2)(x^2+y^2-4)\}$ and (b) $f(x,y) = \sqrt{(x^2+2y)}$, give (i) the *domain*, (ii) sketch D_f in the usual x-y plane, (iii) give the *image* or *range* of f: A: (a) (i) The things in the brackets must be *positive*, so $25 > x^2+y^2 > 4$ **or** $25 < x^2+y^2$ *and* $x^2+y^2 < 4$ (oops: impossible!). So we need the first one. Such points satisfying this form an *annulus* with outer radius 5; inner radius 2.



(iii) For any $(x,y) \in D_f$, $(25-(x^2+y^2))(x^2+y^2-4) > 0$. But *setting* $r^2 = x^2+y^2$, what is the maximum value of the expression $(25-r^2)(r^2-4) = 29r^2-r^4-100 = F(r)$, say? $dF/dr = 58r-4r^3 = 0$ for maximum or minimum. This has *two* solutions: $r = 0$ or $r^2 = 58/4 = 14.5$, i.e. $r \approx 3.8$. Thus $0 < F(r) \leq (25-14.5)(14.5-4) = (10.5)^2 = 110.25$. Thus $-\infty < f(x,y) \leq \ln(110.25)$. (Since $F(r) > 0$ but we can get *arbitrarily* near to 0).

(b) We need $x^2+2y \geq 0$, i.e. $y \geq -\frac{1}{2}x^2$. So $D_f = \{(x,y) \mid y \geq -\frac{1}{2}x^2\}$. $A = \sqrt{(x^2+2y)}$ can take any *non-negative* value, (even keeping y = 0 we get $\sqrt{x^2}$!), so the **image** of f is $\mathbf{R}^+$. Note that when talking about what values the *domain* has, use "**accepts**" rather than "**takes**" and "consists of" rather than "contains". Also use [a,b] etc., for *specifying* ranges, not $-5<x<6$, etc.

Q: Does the **function** $f(x,y) = (x^2+y)/(3x^4+2y^2)$ have a limit at the *origin*? Justify your answer. (ii) Show that $\lim_{x\to1, y\to2} (x^2+y^2) = 5$. A: Suppose one looks *along* the y-axis, (so x = 0 here), then $f(0,y) = \frac{1}{2}y$. This tends to $\pm\infty$ as y tends to 0, so f **cannot** have a limit at the origin.

(ii) (Plan ahead — we want to end up with a statement: "*given* $\varepsilon > 0$, there is a δ such that if $\|(x,y)-(1,2)\| < \delta$, then $|f(x,y)-5| < \varepsilon$". If $\|(x,y)-(1,2)\| < \delta$, then $|x-1| < \delta$ and $|y-2| < \delta$. See what this implies for a *general* δ. If $|x-1| < \delta$, then $1-\delta < x < 1+\delta$, and we expect $\delta < 1$, so in this *interval,* x is positive. Thus $1-2\delta < 1-2\delta+\delta^2 < x^2 < 1+2\delta+\delta^2 < 1+3\delta$.

Similarly, if $|y-2| < \delta$, then $4-4\delta < y^2 < 4+5\delta$. Putting these together gives $|x^2+y^2-5| < 8\delta$, and the only *restriction* was $\delta < 1$. To get this $< \varepsilon$, choose $\delta \leq \varepsilon/10.$ Now put the building blocks together. Here is a *possible* way: Given $\varepsilon > 0$, set $\delta = \min(\varepsilon/10, 1)$. If $\|x-(1,2)\| < \delta$, then (a) $1-\delta < x < 1+\delta$, and (b) $2-\delta < y < 2+\delta$. (a) implies that $-2\delta < x^2-1 < 3\delta$ whilst (b) implies that $4\delta < y^2-4 < 5\delta$. Hence $|x^2+y^2-5| < 8\delta < \varepsilon$ as required, i.e. $\lim_{x\to1 y\to2} (x^2+y^2) = 5$.

Q: Suppose that X is a **convex** set in **R** (so is an interval). *Suppose* that f: $X \to \mathbf{R}^1$ is a convex function, and that x_0 is a point at which f has a local *minimum*. (No assumption about the differentiability of f is needed: "x_0 is a local minimum" means "there is some $\varepsilon > 0$ such that if $x \in X$ and $|x-x_0| < \varepsilon$, then $f(x_0) \leq f(x)$"). Prove that $f(x_0)$ is a *global minimum* (so no value of f is **less** than this one).

A: Suppose that X is **convex**; $X \subseteq \mathbf{R}$; f: $X \to \mathbf{R}^1$ is a *convex* function; and x_0 is a point at which f has a local **minimum**. "X is convex" means "if $x,x' \in X$ are any *two* points, and $t \in [0,1]$, then $x_t = tx+(1-t)x' \in X$". "f is a *convex* function" means "if $x,x' \in X$, and $t \in [0,1]$, then $f(x_t) \leq tf(x)+(1-t)f(x')$". "f has a *local minimum* at x_0" means "if $x,x' \in X$ and $t \in [0,1]$ such that $|x-x_0| < \varepsilon$, then $f(x_0) \leq f(x)$".

The idea of the proof is to go *from an arbitrary* x along the line to x_0, until we get to near x_0 — where we know about the **relative** sizes of $f(x_t)$ and $f(x_0)$ by the above. Then we use *algebra*. <u>**Proof**</u>: Suppose that $x \in X$ ($x \neq x_0$). We want to show that $f(x) \geq f(x_0)$. As $x \in X$, we know that for any $t \in [0,1]$, $x_t = tx+(1-t)x_0$ is *also* in X.

If we pick $0 < t < \varepsilon/|x-x_0|$, where ε is given by the *local* maximality condition (above), then $|x_t-x_0| = t(x-x_0) < \varepsilon$. So by the above, $f(x_t) \geq f(x)$ (---(a)). By the *convexity* of f, $tf(x)+(1-t)f(x_0) \geq f(x_0)$. *Rearranging* this gives $tf(x) \geq tf(x_0)$. But we (cunningly) picked $t > 0$, so this *implies* that $f(x) \geq f(x_0)$ as required.

Q: Determine whether each of the *following* functions is convex, concave or neither. Then find and classify any **stationary** points. (i) $f(x_1,x_2) = 2x_1^3+3x_1x_2+x_2^2$, $x \in \mathbf{R}^2$; (ii) $f(x_1,x_2,x_3) = x_1^2+x_2^2-2x_3^2+0.5x_1x_2$, $x \in \mathbf{R}^3$. A: f is convex iff for all $x \in S$, all *principle minors* are ≥ 0. f is concave iff all non-zero principal k×k minors have the same sign as $(-1)^k$.

(i) $H = \begin{pmatrix} 12x_1 & 3 \\ 3 & 2 \end{pmatrix}$. So as x_1 could be +ve or -ve, this is not *going* to be "definite" at all. More exactly, if $12x_1 \geq 0$, we might hope for +ve definite, but $\det(H) = 24x_1-9$, so f is only convex on the set $x_1 \geq {}^9/_{24}$. Stationary Points: f has a local min at $({}^3/_4, -{}^9/_8)$, and a saddle point at (0,0). (ii) $H = \begin{pmatrix} 2 & 0.5 & 0 \\ 0.5 & 2 & 0 \\ 0 & 0 & -4 \end{pmatrix}$. The 1st principle minors are 2, 2 and -4, so the function in neither convex **nor** concave.

Q: For what *values* of a, b and c is $ax_1^2+bx_1x_2+cx_2^2$ a *convex* function on $\mathbf{R}^2$? A: $H = \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}$. We need $a \geq 0$; $c \geq 0$ and $b^2-4ac \leq 0$.

Q: Find all *local minima*, *local maxima* and *saddle* points for $f(x_1,x_2) = x_1^3-x_2^2+x_1x_2-3x_2$ on $\mathbf{R}^2$. A: $\delta f/\delta x_1 = 3x_1^2+x_2$. If $\delta f/\delta x_1 = 0$, then $x_2 = -3x_1^2$ (---(*)). $\delta f/\delta x_2 = -2x_2+x_1-3$. If $\delta f/\delta x_2 = 0$, then $x_2 = {}^1/_2(x_1-3)$ (---(**)). *Solving* (*) & (**), we get $6x_1^2+x_1-3 = 0$; $x_1 = {}^{-1\pm\sqrt{(1+72)}}/_{12} = {}^{-1\pm\sqrt{73}}/_{12}$. We will need *approximate values* of this: $x_1 \approx -0.795$ or 0.629).

Let us write $\underline{x}^{(1)}$ and $\underline{x}^{(2)}$ for $\underline{x}^{(1)} = ({}^{-1-\sqrt{73}}/_{12}, -{}^{37}/_{24}-{}^{\sqrt{73}}/_{24})$ and $\underline{x}^{(2)} = ({}^{-1+\sqrt{73}}/_{12}, -{}^{37}/_{24}+{}^{\sqrt{73}}/_{24})$. The *Hessian*, $H(\underline{x})$, is $H(\underline{x}) = \begin{pmatrix} 6x_1 & 1 \\ 1 & -2 \end{pmatrix}$. The *principle minors* are as follows: (1st): $6x_1$ and -1; (2nd): $-12x_1-1$ (i.e. $\det(H(\underline{x}))$).

At $\underline{x}^{(1)}$, both the first principal *minors* are negative (since $x_1^{(1)} \approx -0.8$), whilst the *second* principal minor is approximately 6.2. The **signs** are also $(-1)^k$: the 1st ones (k = 1) are *negative*; the 2nd one (k = 2) is *positive*. Thus $H(\underline{x})$ is positive definite **near** $\underline{x}^{(1)}$, and f has a *local* maximum at $\underline{x}^{(1)}$ (f is *concave* at $\underline{x}^{(1)}$). At $\underline{x}^{(2)}$, $6x_1^{(2)}$ is positive (≈ 3.6), but -2 is *negative*. The 2nd principal minor is negative. Thus it is a **saddle** point. f is neither concave *nor* convex near $x^{(2)}$. Alternatively, apply the earlier test. This is really the same here as we have 2 variables.

➢ 15th November 1999

Workshop (Weeks 6 & 7) Continued

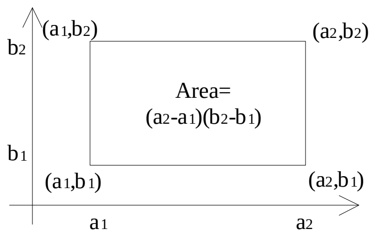
Method of steepest ascent. Begin at some *point* $\underline{v}_0$. For non-negative t, let $\underline{v}_t = \underline{v}_0 + t\nabla f(\underline{v}_0)$. (i) Find $t_0 \geq 0$ such that t_0 *maximises* $f(\mathbf{v}_t)$. Set $\mathbf{v}_1 = \mathbf{v}_{t0}$. (ii) Ask "Is $\|\nabla f(\mathbf{v}_1)\|$ *smaller* than some preset bound?", e.g. 0.01. If "yes", **terminate**, as $\nabla f(\mathbf{v}_1)$ being that small means that $\mathbf{v}_1$ is *approximately* a stationary point. If "no", then move from $\mathbf{v}_1$ in the direction of $\nabla f(\mathbf{v}_1)$, and repeat. When the process *stops*, with $\mathbf{v}_n$, say, then $\mathbf{v}_n$ is an *approximate* stationary point of f.

Predicting change with differentials. *Suppose* we know that f: S→**R** is differentiable, and also *know* $f(\underline{x}_0)$ and $Df(\underline{x}_0)$. Can we estimate $f(\underline{x}_0+\Delta\underline{x})$ for *small* values of $\Delta\underline{x}$? If $\Delta\underline{x}_0$ is small, (i.e. $\|\Delta\underline{x}_0\|$ is small), then f and its *linear approximation* will not differ by **very** much, and we can estimate the *change* $\Delta f = f(\underline{x}_0+\Delta\underline{x})-f(\underline{x}_0)$ to be $L(\underline{x}_0+\Delta\underline{x})-L(\underline{x}_0)$, where $L(\underline{x}) = f(\underline{x}_0)+\Delta f(\underline{x}_0)(\underline{x}-\underline{x}_0)$.

Of course, this gives an *estimate* for Δf given by $\nabla f(\underline{x}_0).\Delta\underline{x}$. As this is a good *estimate* "in the limit as $\|\Delta\underline{x}_0\|\to 0$", we use *differential* notation: $df = \nabla f(\underline{x}_0).dx$, where $\underline{dx} = (dx_1,\ldots,dx_n)$, if a *co-ordinate* system is given. In other words, $df = \Sigma_{i=1}^{n} (\delta f/\delta x_i)dx_i$. Q: Let $z = (x/y)^{1/3}$. Find the approximate **maximum** percentage error in measuring z if there is a *possible* 1% error in measuring x, and a *possible* 0.5% error in measuring y.

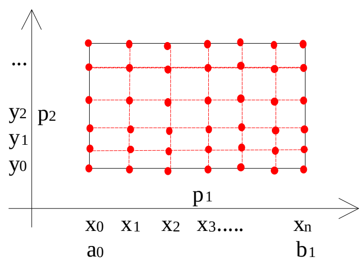
➢ 15th November 1999

Multiple Integrals



To **generalise** integration to many variables, we will need a *replacement* for the length of an interval [a,b] (i.e. (b-a)) for many dimensions. In dimension 3, we use the area of a **rectangle** as shown. In 3 dimensions, the volume is $(a_2-a_1)(b_2-b_1)(c_2-c_1)$. What is the measure "A" of a *bounded* n-dimensional interval in $\mathbf{R}^n$, if $A_1,\ldots,A_n$ *are bounded* intervals in **R**?

A: $A = A_1\times\ldots\times A_n = \{(x_1,\ldots,x_n) \mid x_k \in A_k, k = 1,\ldots,n\}$. The A_i may be *open, closed or half-open* (but will usually be **bounded**). We say that A is an *n-dimensional* interval. Each A_k has *finite* length, e.g. $A_k = [a_k,b_k]$, where $b_k-a_k < \infty$. We *write* $\mu(A_k) = b_k-a_k$. Let us *define* $\mu(A) = \mu(A_1)\ldots\mu(A_n)$, the *measure* of A.



Partitions of A. Let $A = A_1\ldots A_n$ be closed and bounded. If P_k is a partition of A_k, let $P = P_1\times\ldots\times P_n$. (P is a *partition* of A). It is **fairly** clear **when** P' < P ('<' = finer than) occurs, because we are using *partitions* of A. ($P \subseteq P'$).

Riemann Sum: Suppose that f is *defined* and *bounded* on an $I \subset \mathbf{R}$. (I is a *generalised* closed bounded subinterval of $\mathbf{R}^n$). If $P \in \mathcal{P}(I)$, ($\mathcal{P}(I)$ = partitions of $I,\dots,I_m$), where P divides I into *m* subintervals I_k; and $\underline{t}_k \in I_k$, then $S(P,f) = \Sigma_{k=1}^{m} f(t_k)\mu(I_k)$ is called a *Riemann* sum.

Riemann Integrable Functions

f is said to be *Riemann integrable* on I ($f \in \mathcal{R}$ on I) whenever there is a *real number M* such that for every $\varepsilon > 0$, then **there** is a $P_\varepsilon \in \mathcal{P}(I)$ such that if $P < P_\varepsilon$, and $|S(P,f)-M| < \varepsilon$ for all *Riemann sums* S(P,f), then M is *denoted by* $\int_I f d\underline{x}$, or $\int_I f(\underline{x})d\underline{x}$, or $\int_I f(x_1,\dots,x_n)dx_1dx_2\dots dx_n$. Also, for **small** n, it can be written as a *"multiple integral", e.g.* $\int\int_I f(x_1,x_2)dx_1dx_2$.

Multiple Integrals: Theory

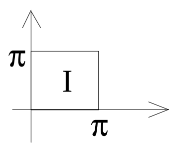
The **basic** idea is to first integrate w.r.t. the first variable dx_1, pretending all other variables are constants, so that you get an *expression* without x_1's in it. Then you repeat with x_2, etc. This is relatively easy when the region you are integrating is a **rectangle**. But often, the limits of integration of the *inner* integral depend on another variable. Luckily, this does not make much difference — when you come to substitute into the **expression** for the integral, you just substitute those "variable values".

Often, the integral is not in a form which is **easy** to evaluate. Write the region in another form if the problem is the order of integration. For example, if the *inner integral* is w.r.t. y, with the region expressed with a constraint on y in terms of x, draw the region, and rework the constraint to give a constraint on x in terms of y. Now try the integral the *"other way around"*, integrating first w.r.t. x and then w.r.t. y. This **generalises** to higher dimensions, but is harder — you cannot draw *higher* dimensional regions!

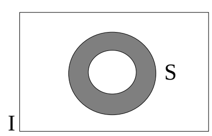
➢ 16th November 1999

A **subset** $T \subset \mathbf{R}^n$ is said to be of *measure zero* if for every $\varepsilon > 0$, T can be *covered* by a collection $\{I_k\}_{k=1}^{\infty}$ of *n-dimensional intervals*, with $\Sigma\mu(T_k) < \varepsilon$. **Theorem**: If f is defined and bounded on a *closed bounded* $I \subset \mathbf{R}^n$, then $f \in \mathcal{R}$ on I iff the set of *discontinuities* of f has measure zero.

Example of Multiple Integrals by Iterated Integrations



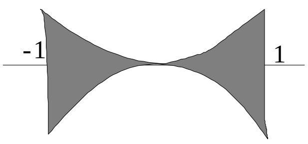
Example: $I = [0,\pi]\times[0,\pi]$, $f(x,y) = \sin^2x\sin^2y$. So $\int\int_I \sin^2x\sin^2y\,dydx = \int_0^\pi(\int^\pi \sin^2x\sin^2y\,dy)dx = \int_0^\pi \sin^2x(\int_0^\pi \sin^2y\,dy)dx = \int_0^\pi \sin^2x[\int_0^\pi(^{1-\cos 2y}/_2)dy)dx = \int^\pi \sin^2x[^y/_2 - ^{\sin 2y}/_4]_0^\pi dx = ^\pi/_2\int_0^\pi \sin^2x\,dx = (^\pi/_2)^2$.



Could we have **handled** an integral over some *less regular region* — other than a rectangle? For example, if $S = \{(x,y) \mid 1 \le x^2+y^2 \le 2\}$ and $f(x,y) = \sqrt{(x^2+y^2-1)}$, what should $\int_S f\,dxdy$ mean? Idea: Pick some $I \supset S$ (in *general*) and let χ_S: $I \to \mathbf{R}$ be defined by $\chi(s) = \{1$ if $s\in S$; 0 if $s\notin S\}$, the *characteristic function* of S. If χ_S is *Riemann integrable* on I, then set g: $I\to\mathbf{R}$ to be $g(\underline{x}) = \{f(\underline{x})$ if $x\in S$; 0 if $x\notin S\}$, and define $\int_S fd\underline{x} = \int_I gd\underline{x}$. In our circumstances, the subset S will *always* be such that $\int_I \chi_S dx$ exists, but more *general cases* can arise.

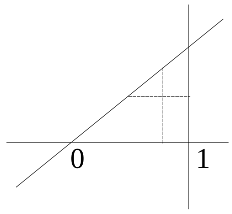
Order of Integration

Fubini's Theorem: (but with a *rectangular* region): If $f(x,y)$ is *continuous* on $I = [a,b]\times[c,d]$, then $\int_S f d\underline{x} = \int_c^d(\int_a^b f(x,y)dx)dy = \int_a^b(\int_c^d f(x,y)dy)dx$. Not **all** regions are rectangular. *Suppose* that g_1, g_2: $[a,b]\to\mathbf{R}$ are *continuous*, and let $S = \{(x,y) \mid x \in [a,b], g_1(x) \leq y \leq g_2(x)\}$ (so here we are *assuming that* $g_1(x) \leq g_2(x)$ for all x).

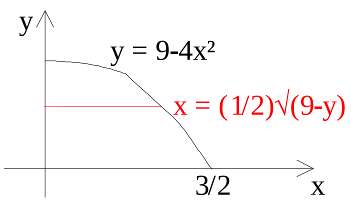


Example: Let $g_1(x) = -x^2$ and $g_2(x) = x^2$ for $x \in [-1,1]$. **Assuming** that f: $S\to\mathbf{R}$ is continuous, then $\int_S f d\underline{x} = \int_a^b(\int_{g_1(x)}^{g_2(x)} f(x,y)dy)dx$. If **S** is a above, and $f(x,y) = xe^{-y}$, then the integral is $\int_S xe^{-y}d\underline{x} = \int_{-1}^1(\int_{-x^2}^{x^2} xe^{-y}dy)dx$. **Evaluate** the inner bracket: $\int_{-x^2}^{x^2} xe^{-y}dy = x\int_{-x^2}^{x^2} e^{-y}dy = x[-e^{-y}]_{-x^2}^{x^2} = x(-e^{x^2}+e^{-x^2})$. So the integral is $\int_{-1}^1 x(e^{-x^2}-e^{x^2})dx$ = ...etc.

➢ 23rd November 1999



Here, the region worked well with the *order* of integration. This does not **always** happen. Example: Let S be the region in $\mathbf{R}^2$ bounded by the x-axis, the line y = x, and the line x = 1 (as shown). Then $S = \{(x,y) \mid x \in [0,1], 0 \leq y \leq x\}$. This is the same as $S = \{(x,y) \mid y \in [0,1]\ y \leq x \leq 1\}$. Take $f(x,y) = {}^{\sin(x)}/_{x}$. Look at $\int_S f d\mathbf{x} = \int_0^1(\int_y^1 {}^{\sin x}/_x\, dx)dx$. Reverse, giving $\int_0^1(\int_0^x {}^{\sin x}/_x\, dy)dx = \int_0^1[{}^{\sin x}/_x y]_0^x\, dx = \int_0^1 \sin(x)dx = -\cos(1)+1$.



Typical Question: Sketch the **region** of integration and write an equivalent double integral with the order of integration *reversed*. Let Int $= \int_0^{3/2}\int_0^{9-4x^2} 16xdydx = \int_0^{3/2}[16xy]_0^{9-4x^2}dx = \int_0^{3/2} 16x(9-4x^2)dx = \ldots$ **Reversing** the order, we get Int $= \int_0^9\int_0^{1/2\sqrt{(9-y)}} 16xdxdy = \int_0^9[8x^2]_0^{1/2\sqrt{(9-y)}} dy = \int_0^9 2(9-y)dy = \ldots$

Tutorial

Q: Evaluate the **following** integrals: (a) $\int_{-1}^1\int_0^2 (2x^2+2xy+3y^2)dydx$, (b) $\int_0^3\int_1^2 xye^{(2x+3y)}dxdy$. A: (a) Evaluating the *inner* integral: $\int_0^2 2x^2+2xy+3y^2dy = [2x^2y+xy^2+y^3]_0^2 = [4x^2+4x+8]$. Now we have $\int_{-1}^1 4x^2+4x+8dx = [{}^4/_3x^3+2x^2+8x]_{-1}^1 = [({}^4/_3+2+8)-(-{}^4/_3+2-8)] = {}^{34}/_3+{}^{22}/_3 = {}^{56}/_3$.

(b) $\int_0^3\int_1^2 xye^{(2x+3y)}dxdy = \int_0^3\int_1^2 xe^{2x}ye^{3y}dxdy = \int_0^3 ye^{3y}dy \int_1^2 xe^{2x}dx$. *Evaluate* the first integral by parts: Let $u = y$, ${}^{du}/_{dx} = 1$; ${}^{dv}/_{dx} = e^{3y}$, $v = {}^1/_3e^{3y}$. So $\int_0^3 ye^{3y}dy = [{}^y/_3e^{3y}]_0^3-\int_0^3{}^1/_3e^{3y} = [{}^y/_3e^{3y}]_0^3-[{}^1/_9e^{3y}]_0^3 = [{}^y/_3e^{3y}]_0^3-[{}^1/_9e^9-{}^1/_9] = [e^9-0]-[{}^1/_9e^9-{}^1/_9] = e^9-{}^1/_9e^9+{}^1/_9 = ({}^8/_9e^9+{}^1/_9)$. Similarly, $\int_1^2 xe^{2x} = ({}^3/_4e^4-{}^1/_4e^2)$. So the *answer* is $({}^8/_9e^9+{}^1/_9)({}^3/_4e^4-{}^1/_4e^2)$.

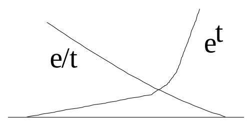
Assignment 2

Q: Given the **mapping** $x = u^2+v^2$; $y = 2uv$, and the function $w = w(x,y)$, express $\nabla^2 w$ in terms of *partial* derivatives with respect to u and v. A: The answer is too *detailed* to give in full here. Technique: find ${}^{\partial u}/_{\partial x}$ and ${}^{\partial v}/_{\partial x}$ by *implicit* differentiation; then find ${}^{\partial^2 w}/_{\partial x^2}$ and ${}^{\partial^2 w}/_{\partial y^2}$ **Tips**: Keep any *factors* outside, group *related* terms, *split* the calculation into bits, and *reassemble* at the end. Look out for any *shortcuts*, e.g. $v^2-u^2 = -(u^2-v^2)$.

Q: Compute the **Jacobian** matrix of each of the following transformations, and determine where *local* inverses exist: (a) $x = e^u\cos v$, $y = \sin v$; (b) $x = u^2+v^2$, $y = u+v$. A: (a) $J = \begin{bmatrix} e^u\cos v & -e^u\sin v \\ 0 & \cos v \end{bmatrix}$. $\text{Det}(J) = e^u\cos^2 v > 0$, except when $v = {}^{2k+1}/_{2}\pi$, $k = 0, \pm1, \ldots$. Thus *local inverses* exist except at **odd** multiples of ${}^{\pi}/_{2}$ in the v-direction. (b) $J = \begin{pmatrix} 2u & 2v \\ 1 & 1 \end{pmatrix}$. $\text{Det}(J) = 2(u-v)$. Local inverses exist **except** on the line $u = v$.

Q: Perform 2 iterations of the method of **steepest** ascent in an attempt to maximise $f(x_1,x_2) = (x_1+x_2)e^{-(x_1+x_2)}-x_1$, starting at the *point* (0,1). A: $\nabla f = ((1-(x_1+x_2))e^{-(x_1+x_2)}-1, (1-(x_1+x_2)e^{-(x_1+x_2)})$. $\underline{v}_0 = (0,1)$, $\nabla f(0,1) = (-1,0)$. *Note*: $\|\nabla f(\underline{v}_0)\| = 1$. And $v_t^{(0)} = \underline{v}_0+t\nabla f(\underline{v}_0) = (-t,1)$. $f(\underline{v}^{(0)}) = (1-t)e^{-(1-t)}+t$.

Next, find the **maximum** value of $f(v_0^{(t)})$ as a function of t. A: ${}^{df}/_{dt}(\underline{v}_t^{(0)}) = (1-t)e^{-(1-t)}-e^{-(1-t)}+1 = -te^{-(1-t)}+1$. For stationary values, this will be *zero*, so at $e^t = {}^{e}/_{t}$, we have the *diagram* shown. The two curves meet at 1 point **only**, and observation says that t = 1 is the answer. (The *double* derivative test shows that this is a max.).



$\underline{v}_1 = (-1,1) = \underline{v}_1^{(0)}$. $\underline{v}_t^{(1)} = (-1,1+t)$; ${}^{df}/_{dt}(\underline{v}_t^{(1)}) = (1-t)e^{-t}$; $\underline{v}_2 = (-1,2)$. $\nabla f(\underline{v}_1) = (0,1)$; $f(\underline{v}_t^{(1)}) = te^{-t}+1$, which equals zero *when* t = 1. (*Again* a maximum). Note that $\nabla f(v_2) = (-1,0)$. Although there is **no** maximum, our iteration *is* increasing the value of f (see the table).

v	(0,1)	(-1,1)	(-1,2)	(-2,2)
f(v)	e^{-1}	1	$1+e^{-1}$	2

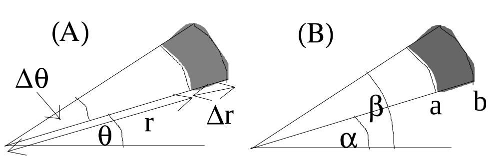
Q: **Euler's** theorem on homogenous functions is: Let f be defined on an *open* set S in $\mathbf{R}^n$. We say that f is *homogeneous* of degree p over S if $f(\lambda\underline{x}) = \lambda p f(\underline{x})$ for every real λ and *every* $x \in S$ such *that* $\lambda\underline{x} \in S$. Euler's *theorem* on homogeneous functions in part states that in this case, if f is <u>differentiable</u> at $\underline{x}$, then $\underline{x}.\nabla f(\underline{x}) = pf(\underline{x})$. **Prove** This.

A: Let f be **homogenous** of degree p on S. Fix $x \in S$. Let $g(\lambda) = f(\lambda x)$. By *first principles,* $g'(1) = \lim_{h\to0} {}^{g(1+h)-g(1)}/_{h} = \lim_{h\to0} {}^{f((1+h)x)-f(x)}/_{h}$ (---(*)) = (since f is *homogenous of degree* p) $= \lim_{h\to0} {}^{((1+h)^{\wedge}p-1)}/_{h} f(x)$ = (.. as in $(x+h)^p = x^p+px^{p-1}+\ldots$, *the binomial theorem* is used in the derivation of ${}^{d}/_{dx}(x^p) = px^{p-1}$) $= \lim_{h\to0}(p + \text{terms in } h)f(x) = pf(x)$.

Going **back** to (---(*)), this is $\lim_{h\to0} f(\underline{x}+h\underline{x})-f(\underline{x})/h$. But this is the **definition** of the directional derivative *of f at $\underline{x}$* in the direction of $\underline{x}$, and we have that this is $\underline{x}.\nabla f(\underline{x})$. Putting this **together**, $\underline{x}.\nabla f(\underline{x}) = pf(\underline{x})$, i.e. if $\underline{x} \in \mathbf{R}^n$, then $\Sigma_{i=1}^{n} x_i\,(\partial f/\partial x_i) = pf(\underline{x})$.

Double Integration in Polar Co-ordinates

Looking at the **diagram** on the right, (A), the area of a segment with angle $\Delta\theta$ and radius r is ${}^{1}/_{2}\Delta\theta r^2$. The area of a segment with angle $\Delta\theta$ and radius $(r+\Delta r)$ is ${}^{1}/_{2}\Delta\theta(r+\Delta r)^2$. The *shaded* area is $r\Delta r\theta$ + terms in $\Delta r^2\Delta\theta$. In the *Riemann sum* with a partition of some sector as shown, (B), $S(P,f) = \Sigma f(\underline{t}_k)\mu(S_k) = \Sigma f(\underline{t}_k)r_{k-1}\Delta r_k\Delta\theta_k$ + small terms. As the *norm* of the partition tends to zero, this approaches $\int f\,d\mu = \int_{\alpha}^{\beta}\int_{a}^{b} f(r,\theta)r\,dr\,d\theta$.



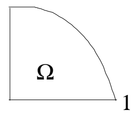
Example: Area of the sector: $\iint_S d\mu = \int^{\beta}_{\alpha}\int^{b}_{a} r\,dr\,d\theta = \int^{\beta}_{\alpha}[r^2/2]^b_a d\theta = [b^2/2 - a^2/2](\beta-\alpha)$. We can adopt this to the *general* case of a change in co-ordinates. The basic element of *volume/area/measure* is $dx_1,\ldots,dx_n$ = **determinant** of a matrix with dx_i on the diagonals and *zeroes* everywhere else. dx_i is an "*infinitesimal*" change in the x_i direction.

If we **change** from $x_1,\ldots,x_n$ co-ordinates to $y_1,\ldots,y_n$ co-ordinates, where $x_k = g_k(y_1,\ldots,y_n)$, then $dx_k = \Sigma^n_{i=1} (\partial g_k/\partial y_i)dy_i$, i.e. as shown on the *right*. So, $dx_1\ldots dx_n = J_g(y)dy_1\ldots dy_n$. Examples: (1) **Polars**: $x_1 = r\cos\theta$, $x_2 = r\sin\theta$, $J = \partial(x_1,x_2)/\partial(r,\theta) = r$. The integral becomes $\iint f(r,\theta)r\,dr\,d\theta$.

$$\begin{pmatrix} dx_1 & \cdots & 0 \\ 0 & \cdots & dx_n \end{pmatrix} = Dg.\begin{pmatrix} dy_1 & \cdots & 0 \\ 0 & \cdots & dy_n \end{pmatrix}$$

(2) **Cylindrical polars**: $(x_1,x_2,x_3) \leftrightarrow (r,\theta,z)$. $J = r$ again. $\iiint f dx_1 dx_2 dx_3 = \iiint f(r,\theta,z)r\,dr\,d\theta\,dz$. (3) Spherical polars: $(x_1,x_2,x_3) \leftrightarrow (\rho,\varphi,\theta)$. $J = |\rho^2\sin\varphi|$ (the volume is +ve so we do need the |...|). The *integral* becomes $\iiint f.|\rho^2\sin\varphi|d\rho\,d\varphi\,d\theta$. Look in *Thomas & Finney* for other examples.

Example: $I = \iint_\Omega xy\,dx\,dy$. Convert to *polar* co-ordinates: $\Omega = \{(r,\theta) \mid 0\le r\le 1, 0\le\theta\le \pi/2\}$. So $I = \int^{\theta=\pi/2}_{\theta=0}\int^{r=1}_{r=0} r^3\cos\theta\sin\theta\,dr\,d\theta = 1/8$. Example: to evaluate $\int^{\infty}_{-\infty} e^{-x^2}dx = I$, evaluate $I^2 = (\int^{\infty}_{-\infty} e^{-x^2}dx)(\int^{\infty}_{-\infty} e^{-y^2}dy) = \int^{\infty}_{-\infty}\int^{\infty}_{-\infty} e^{-(x^2+y^2)}dx\,dy$. Let $D_b = \{(x,y) \mid 0 \le x^2+y^2 \le b^2\}$. So $\iint_{D_b} e^{-(x^2+y^2)}dx\,dy = \int^{\theta=2\pi}_{\theta=0}\int^{r=b}_{r=0} e^{-r^2}r\,dr\,d\theta = (\int^{2\pi}_0 d\theta)[-1/2\,e^{-r^2}]^b_0 = 2\pi.(1-e^{-b^2}/2) = \pi-\pi e^{-b^2}$. Now let $b\to\infty$, so $I^2 = \lim\iint_{D_b} xy\,dx\,dy = \pi$. Therefore, $I = \sqrt{\pi}$. Summary: $\int^{\infty}_{-\infty}e^{-x^2}dx = \sqrt{\pi}$.



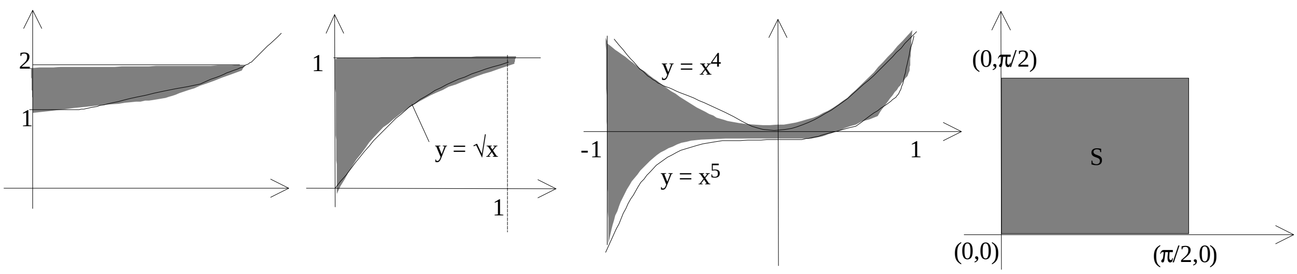
➢ 30th November 1999

Section 3: Constrained Optimisation

This is where we *maximise* or *minimise* a function f subject to some constraints. If we have a general function f with no constraints, we could use the **Hessian** or the **steepest ascent** methods. If we have a linear function f with linear constraints, we typically use linear programming (G2M85). If we want to optimise (x,y,z) subject to a constraint of the form g(x,y,z) = 0, we handle by the method of *Lagrange multipliers*. Example: find the **optima** of $x^3+15xy^{24}-16zxy$ on the surface of $x^2+y^2+z^2-1 = 0$. Method: Find *values* of x, y, z and λ that **simultaneously** satisfy the equations $\nabla f+\lambda\nabla g = 0$ and $g(x,y,z) = 0$. (Variants with *n-variables* and *m constraints* exist). The **first** condition is interpreted as $\partial f/\partial x = -\lambda\,\partial g/\partial x$; $\partial f/\partial y = -\lambda\,\partial g/\partial y$; $\partial f/\partial z = -\lambda\,\partial g/\partial z$. Now we solve, *remembering* that g(x,y,z) = 0.

Assignment 3

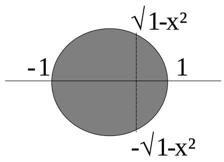
Q: For the following integrals, sketch the region and *evaluate* the integral. (If necessary, change the **order** of integration). (i) $\int^2_1\int^{\ln y}_0 e^{-x}dx\,dy$; (ii) $\int^1_0\int^1_{\sqrt{x}} \sin(y^3+1/2)dy\,dx$; (iii) $\iint_S (x+y)dx\,dy$, where $S = \{(x,y): -1\le x\le 1$, y lies between x^4 and $x^5\}$; (iv) $\int^{\pi/2}_0\int^{\pi/2}_0 \cos(x+y)dx\,dy$.



A: (i) The **lower** x-limit is $x = \ln y$, so $y = e^x$. There is no need to change the order of integration. So $\int_1^2\int_0^{\ln y} e^{-x}dxdy = \int_1^2[-e^{-x}]_0^{\ln y}dy = \int_1^2\{1-{}^1/_y\}dy = [y-\ln y]_1^2 = 2-\ln2-1+\ln1 = 1-\ln2$. (ii) We need to *change* the order. Note that if $y = \sqrt{x}$, then $x = y^2$. So $\int_0^1\int_{\sqrt{x}}^1 \sin({}^{y^3+1}/_2)dydx = \int_0^1\int_0^{y^2} \sin({}^{y^3+1}/_2)dxdy = \int_0^1[x\sin({}^{y^3+1}/_2)]_0^{y^2}dy = \int_0^1 y^2\sin({}^{y^3+1}/_2)dy = [-{}^2/_3\cos({}^{y^3+1}/_2)]_0^1 = {}^2/_3(\cos{}^1/_2-\cos1)$.

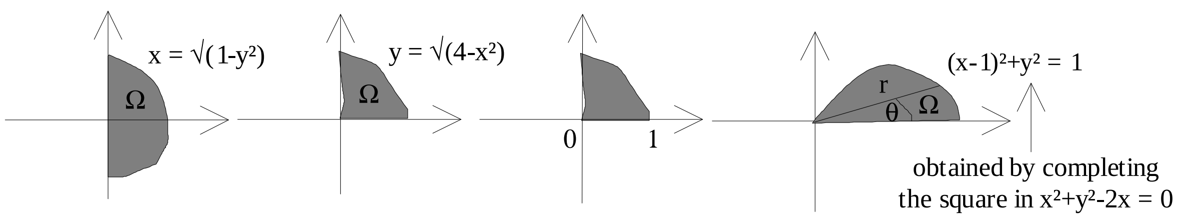
(iii) Within the *x-range* $(-1\le x\le1)$, we have $x^4 \geq x^5$, so no problems. $\iint_S (x+y)dxdy = \int_{-1}^1\int_{x^5}^{x^4} (x+y)dydx = \int_{-1}^1 [xy+{}^{y^2}/_2]_{x^5}^{x^4}dx = \int_{-1}^1 (x^5+(x^8/2)-x^6-(x^{10}/2))dx = [(x^6/6)+(x^9/18)-(x^7/7)-(x^{11}/22)]_{-1}^1 = {}^1/_9-{}^2/_7-{}^1/_{11} = -{}^{184}/_{693}$. (iv) $\int_0^{\pi/2}\int_0^{\pi/2} \cos(x+y)dxdy = \int_0^{\pi/2}[\sin(x+y)]_0^{\pi/2}dy = \int_0^{\pi/2}(\sin(y+{}^\pi/_2)-\sin(y))dy = [-\cos(y+{}^\pi/_2)+\cos(y)]_0^{\pi/2} = -\cos({}^\pi/_2+{}^\pi/_2) + \cos({}^\pi/_2) + \cos({}^\pi/_2) - \cos(0) = 0$.

Q: Find the **volume** of the paraboloid $z = x^2+y^2$ under the cylinder $x^2+y^2 \le 1$. A: The required integral is $\iint_S zdxdy = \int_{-1}^1\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (x^2+y^2)dydx = \int_{-1}^1(x^2y+{}^{y^3}/_3)_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}dx = 2\int_{-1}^1(x^2(1-x^2)^{1/2}+{}^1/_3(1-x^2)^{3/2})dx = \int_{-1}^1(1-x^2)^{1/2}(2x^2+{}^2/_3(1-x^2))dx = \int_{-1}^1{}^2/_3(1-x^2)^{1/2}dx+{}^4/_3\int_{-1}^1x^2(1-x^2)dx$. To evaluate this either plough on, or *use substitutions*, polars, *volumes* of revolution, etc.



Tutorial

For each of the following, change the "*Cartesian* integral" to an equivalent "*polar* integral" and then evaluate. (i) $\int_{-1}^1\int_0^{\sqrt{1-y^2}} \sqrt{(x^2+y^2)}dxdy$; (ii) $\int_0^2\int_0^{\sqrt{4-x^2}} \sqrt{(x^2+y^2)}dydx$; (iii) $\int_0^1\int_0^{\sqrt{1-x^2}} \sin(x^2+y^2)dydx$; (iv) $\int_0^2\int_0^{\sqrt{2x-x^2}} xdydx$.

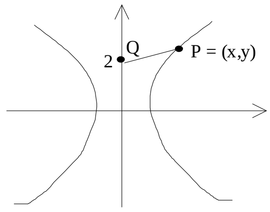


A: (i) Let $I = \int_{-1}^1\int_0^{\sqrt{1-y^2}} \sqrt{(x^2+y^2)}dxdy$. If $x = r\cos\theta$ and $y = r\sin\theta$ (*polars*), then $I = \int_0^1\int_{-\pi/2}^{\pi/2} \sqrt{(r^2\cos^2\theta+r^2\sin^2\theta)}.rdrd\theta = \int_0^1\int_{-\pi/2}^{\pi/2} r.rdrd\theta$ (with $\{(r,\theta) \mid 0\le r\le1, -{}^\pi/_2\le\theta\le{}^\pi/_2\}$) $= \int_0^1(\int_{-\pi/2}^{\pi/2}r^2dr)d\theta = {}^\pi/_3$. (ii) $\int_0^2\int_0^{\sqrt{4-x^2}} \sqrt{(x^2+y^2)}dydx = \int_0^{\pi/2}\int_0^2 r^2drd\theta = \ldots = {}^4/_3\pi^3$. (iii) $\int_0^1\int_0^{\sqrt{1-x^2}} \sin(x^2+y^2)dydx = \int_0^{\pi/2}\int_0^1 \sin(r^2).rdrd\theta = {}^\pi/_4(1-\cos1)$. (iv) Let $I = \int_0^2\int_0^{\sqrt{2x-x^2}} xdydx$. There are several ways to answer this question; one is to let $x = 1+r\cos\theta$ and $y = r\sin\theta$. So $I = \int_0^\pi\int_0^1 (1+r\cos\theta)rdrd\theta$ (Do check the *Jacobean*) $= \ldots = \int_0^\pi[{}^1/_2-{}^1/_3\cos\theta]d\theta = \ldots$

➢ 6th December 1999

Example: Maximise the volume of a **box** with fixed surface area. *So we maximise* $f(x,y,z) = xyz$ subject to $2(xy+yz+xz) = A$. Set $g(x,y,z) = xy+yz+xz-{}^A/_2$. (So $g = 0$ is the *constraint*). There are **four** equations: $yz = -\lambda(y+z)$; $xz = -\lambda(x+z)$; $xy = -\lambda(x+y)$; and $xy+yz+xz = {}^A/_2$. So we have $-\lambda = {}^{yz}/_{y+z} = {}^{xz}/_{x+z} = {}^{xy}/_{x+y}$.

Therefore, $yz(x+z) = xz(y+z)$, i.e. $yz^2 = xz^2$, or $z^2(x-y) = 0$. The *stationary points* will imply that $z = 0$ (which will not give us a **maximum**) or $x = y$. Symmetrically, we also have $x = z$ and $u = z$. The constraint then gives $3x^2 = {}^A/_2$; $x = \sqrt{({}^A/_6)}$. (The -ve square root has no *relevance* here). So $V_{max} = ({}^A/_6)^{3/2}$.



Example: $x^2-y^2 = 1$, a *hyperbola*. What is the shortest distance to the hyperbola from the point (0,2)? A: Here, $f(x,y) = x^2+(y-2)^2$, and $g(x,y) = x^2-y^2-1$. Note *that* $x \neq 0$ because $g(x,y) \neq 0$. The **Lagrange** conditions give $2x = -2\lambda x$ and $2(y-2) = 2\lambda y$. From $2x = -2\lambda x$, $\lambda = -1$. So $2(y-2) = -2y$; $4y = 4$, *implying that* $y = 1$. Substituting into g **gives** $x^2-1-1 = 0$; $x = \pm\sqrt{2}$, and $f(\pm\sqrt{2}, 1) = 3$. Exercise: Find the **critical** points of $f(x,y,z) = xyz$ subject to $x^2+y^2+z^2$, and *classify* them (there are lots of them).

➢ 7th December 1999

Hessian Methods for Constrained Optimisation

Let $l = f(x,y) + \lambda g(x,y)$. Find the *partials* of l w.r.t. x, y and λ; and equate to **zero**. So we have ${}^{\partial f}/_{\partial x}+\lambda{}^{\partial g}/_{\partial x} = 0$ and ${}^{\partial f}/_{\partial y} + \lambda{}^{\partial g}/_{\partial y} = 0$, with $g(x,y) = 0$. Solve for x and y. The **maximisation/minimisation** of f, subject to $g(x,y) = 0$, is the same as *unconstrained optimisation* of $l(x,y)$. The second order conditions come from the Bordered Hessian of f and g, which is shown on the right. Evaluate this at the critical points.

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial g}{\partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial g}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & 0 \end{pmatrix}$$

If $\det(H) < 0$, then f has a constrained **minimum** at the critical point. If $\det(H) > 0$, f has a constrained **maximum**. Note that there are *"side conditions"* on f and g to get them to work. So we will only use it on simple examples. Example: optimise $f(x,y,z) = 4x^2+3xy+6y^2$ such that $x+y-56 = 0$.

Here, $l = 4x^2+3xy+6y^2 + \lambda(x+y-56)$, etc., giving $8x+3y+\lambda = 0$, $3x+12y+\lambda = 0$, and $x+y = 56$ — which *imply* that $x^* = 36$ and $y^* = 20$. (The Solutions). The Bordered Hessian is $H = ({}^{8}3_{1}\ {}^{3}12_{1}\ {}^{1}1_{0})$. The determinant is $\det(H) = -14 < 0$, so we have a *minimum*. Another example: optimise $f(x,y) = 4x^2-2xy+6y^2$ subject to $x+y = 72$. Here, $l = f+\lambda g$, and $x^* = 42$; $y^* = 30$.

In the above, **note** that $\lambda = -276$ and the value of l_{opt} is 9936. If we increase the *constant term* of g by ε, (e.g. $x+y = 72+\varepsilon$, ε small), then l increases by 273ε. We have the **General Property of Lagrange Multipliers**: λ is the change in the *value* of the maximum when a small unit change in the **constraint** is made.

Note: this is *interpreted* as saying the following: "λ^* measures a property related to the **sensitivity** of the optimum value of f to *small* changes in g". More investigation of this will be required to get any more detail.

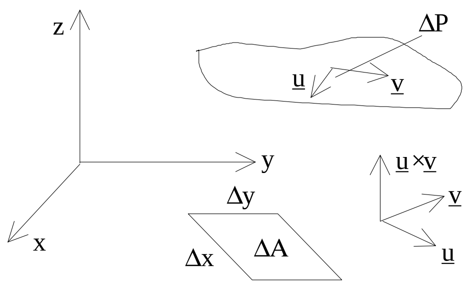
Section 4: Surfaces & Surface Integrals

Let S be a surface in space, *perhaps* represented in the form $z = f(x,y)$, (---(1)) $[(x,y) \in \mathbf{R} \subseteq \mathbf{R}^2]$, or by an *implicit* function $F(x,y,z) = 0$ (---(2)), e.g. ${}^{x^2}/_{a^2}+{}^{y^2}/_{b^2}+{}^{z^2}/_{c^2}-1 = 0$ (an *ellipsoid*), or in parametric form: $x = f(u,v)$, $y = g(u,v)$, $z = h(u,v)$. (---(3)) $[(u,v) \in \mathbf{R} \subseteq \mathbf{R}^2]$. The area in *form* (1) is $A = \int\int_{R_{x,y}} \sqrt{1+(\frac{\partial z}{\partial x})+(\frac{\partial z}{\partial y})^2}\, dydx$, and we have similar expressions in the other cases.

➢ 13th December 1999

Note: the lecture material from now on is not examinable.

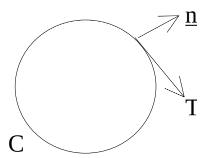
Let $F(x,y,z) = 0$. The area of a **surface** in a small plane is $\approx |\underline{u}\times\underline{v}|$, where $\underline{u}$ and $\underline{v}$ are tangent vectors: $\underline{u} \approx \frac{\partial F}{\partial x}\Delta x e_1$, $\underline{v} \approx \frac{\partial F}{\partial y}\Delta y e_2$. The area ΔP above ΔA is $\frac{\Delta A}{|\cos\gamma|}$, where γ is the angle between ∇F and the z-axis. **Therefore**, $|\cos\gamma| = |\underline{\nabla F}.\underline{p}|/||\nabla F||$ where $p = e_3$, the unit vector in the z direction. So the *surface* area is $\iint_{R_{xy}} \frac{||\nabla F||}{|\nabla F.p|}dxdy$, provided we have $\nabla F.p \neq 0$.



Example: Let $F(x,y,z) = x^2+y^2-z$. (A paraboloid). Then $\nabla F = (2x,2y,-1)$; $|\nabla F| = \sqrt{(4x^2+4y^2+1)}$; $|\nabla F.p| = |-1| = 1$; and $S = \iint_R \sqrt{(4x^2+4y^2+1)}dxdy$. (Compare with the *other* formula). **Surface Integrals.** Let g be defined by g: $S\rightarrow\mathbf{R}$, where S is given by $F(x,y,z) = 0$. The Surface Integral is $\iint_S g dA = \iint_{R_{xy}} g(x,y,z) \frac{||\nabla F||}{|\nabla F.p|}dxdy$. **Vector Fields.** Let F be defined by F: $S\rightarrow\mathbf{R}^3$, where $S \subset \mathbf{R}^3$. So $F = M\underline{i}+N\underline{j}+P\underline{k}$. Example: If F: $S\rightarrow\mathbf{R}$ is differentiable, then $\nabla F = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z})$ is the *gradient vector field* of F on S.

Flux

Consider Flux across a *plane curve*. Let C be a smooth closed curve, defined by $x = g(t)$ and $y = h(t)$, with $t \in [a,b]$. The tangent vector T is defined by $(\frac{dx}{ds}\underline{i}+\frac{dy}{ds}\underline{j})$. Take $\underline{k}$, a *unit vector* in the z direction. The outward unit normal $\underline{n}$, of C, is defined by $\underline{n} = \underline{T}\times\underline{k} = \frac{dy}{ds}\underline{i}-\frac{dx}{ds}\underline{j}$. If f: $S\rightarrow\mathbf{R}^3$ is a vector field with $C \subset S$, the *flux of f across C* is given by the formula $\int_C \underline{F}.\underline{n} = \int_C(M\frac{dy}{ds}-N\frac{dx}{ds})ds$.



Example: Let $\underline{F} = (x-y)\underline{i}+x\underline{j}$, and let $C = \{x^2+y^2-1\}$. Parametrise: $r = \cos t\underline{i}+\sin t\underline{j}$, for $0\leq t\leq 2\pi$; $M = \cos t-\sin t$ $(= x-y)$; $N = \cos t$. So $\frac{dy}{dt} = \cos t$ and $\frac{dx}{dt} = -\sin t$. **Flux**: $\int_C(M\frac{dy}{ds}-N\frac{dx}{ds})ds = \int_0^{2\pi} \cos^2 t-\sin t\cos t+\cos t\sin t)dt = \ldots = \pi$. Note that $\pi > 0$, so the net flux is *outward*.

Green's Theorem

Green's Theorem (in the **plane**): $\iint_C \underline{F}.\underline{n}\ dS = \iint_R (\frac{\partial M}{\partial x}+\frac{\partial N}{\partial y})dxdy$ (the region *enclosed* by C). There is a 3 dimensional version: given a surface S in $\mathbf{R}^3$, assume that we can choose a *unit normal vector* field $\underline{n}$ (S can be **oriented**). The *flux* across S is given by $\iint_S \underline{F}.\underline{n}\ d\sigma$, where $d\sigma = \frac{||\nabla f||}{|\nabla f.p|}dA$ as *before*, with $S = \{(x,y,z) \mid f(x,y,z) = 0\}$.

➢ 14th December 1999

Example: $F(x,y,z) = x\underline{i}+y\underline{j}$. Calculate the flux of F **across** the surface $x^2+y^2+z^2 = a^2$ ($\underline{n}$ is outward). The *outward normal vector* field $\underline{n}$ is given by the gradient ∇f, where $f(x,y,z) = x^2+y^2+z^2-a^2$ after *normalisation*. So $\nabla f = (2x,2y,2z)$, and points outwards (good!). It follows that $\underline{n} = \frac{\nabla f}{||\nabla f||} = \frac{(2x,2y,2z)}{\sqrt{4x^2+4y^2+4z^2}} = \frac{(2x,2y,2z)}{\sqrt{4a^2}} = (\frac{x}{a}, \frac{y}{a}, \frac{z}{a})$.

Now $\underline{F}.\underline{n} = \frac{1}{a}(x,y,0).(x,y,z) = \frac{1}{a}(x^2+y^2)$. The flux *out* of S is $\frac{1}{a}\iint_S (x^2+y^2)d\sigma$. The usual *parametrisation* of S is $r(u,v) = (a\cos u\cos v, a\sin u \sin v)$. $(0\leq u\leq 2\pi)$, $(-\frac{\pi}{2}\leq v\leq\frac{\pi}{2})$, etc.

Stoke's Theorem (in R^3)

Given C, a *boundary* curve of an orientable surface S, then $\int_C \underline{F}.d\underline{r} = \iint_S(\underline{\nabla}\times\underline{F}).\underline{n}\, d\sigma$.

Gauss' Divergence Theorem

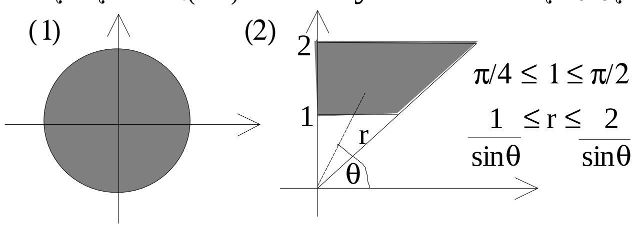
$\iint_S \underline{F}.\underline{n}\, d\sigma = \iiint_D \underline{\nabla}.\underline{F}\, dV$ (*The Boundary* of D is S). Here, $\underline{F}$ is a vector field across the *boundary* of the region D. Notes: $\underline{\nabla}.\underline{F} = ({}^{\partial}/_{\partial x},{}^{\partial}/_{\partial y},{}^{\partial}/_{\partial z}).(M,N,P) = {}^{\partial}/_{\partial x}M+{}^{\partial}/_{\partial y}N+{}^{\partial}/_{\partial z}P$. And so

$$\underline{\nabla}\times\underline{F} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ M & N & P \end{vmatrix}.$$

A matrix $\left({}^{A}_{B}\ {}^{B}_{C}\right)$ is positive semi-definite if $A \geq 0$, $C \geq 0$, and $AC-B^2 \geq 0$. **Similar** conditions apply for negative definite, etc.).

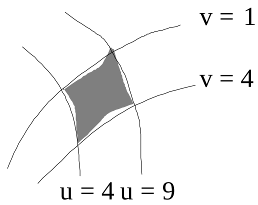
Assignment 4

Q: By changing to a *polar integral*, evaluate $I = \int_{-1}^{1}\int_{-\sqrt{(1-x^2)}}^{\sqrt{(1-x^2)}} e^{-(x^2+y^2)}dydx$. A: $I = \int_{\theta=0}^{2\pi}\int_0^1 re^{-r^2}drd\theta = 2\pi[e^{-r^2}/2]_0^1 = \pi(e^{-1}-1)$. Q: Evaluate $I = \int_1^2\int_0^y {}^{x\sqrt{(x^2+y^2)}}/_y dxdy$. A: $I = \int_{\theta=\pi/4}^{\pi/2}\int_{1/\sin\theta}^{2/\sin\theta} {}^{r\cos\theta.r}/_{\sin\theta}\, rdrd\theta = \int_{\pi/4}^{\pi/2} {}^{\cos\theta}/_{\sin\theta}[{}^{r^3}/_3]_{1/\sin\theta}^{2/\sin\theta}d\theta = {}^1/_3\int_{\pi/4}^{\pi/2} {}^{\cos\theta}/_{\sin\theta}({}^8/_{\sin^3\theta}-{}^1/_{\sin^3 q})d\theta = \int_{\pi/4}^{\pi/2} 7/3\ \cos\theta/\sin^4\theta\ d\theta = [-{}^7/_9\sin^3\theta]_{\pi/4}^{\pi/2} = -{}^7/_9(1-2\sqrt{2}) = {}^7/_9(2\sqrt{2}-1)$.



Q: Evaluate the following using *cyllindrical co-ordinates:* $I = \int_0^1\int_0^{\sqrt{(1-x^2)}}\int_0^2 \cos(x^2+y^2)dzdydx$. A: $I = 2\int_0^1\int_0^{\sqrt{(1-x^2)}} \cos(x^2+y^2)dydx = 2\int_0^{\pi/2}\int_0^1 r\cos r^2 drd\theta = \pi\int_0^1 r\cos(r^2)dr = \pi[\sin(r^2)]_0^1 = \pi\sin(1)$.

Q: Evaluate $\iint_\Omega xydxdy$, where Ω is the region in the *1st quadrant* bounded by the curves $x^2+y^2 = 4$, $x^2+y^2 = 9$, $x^2-y^2 = 1$ and $x^2-y^2 = 4$. (Use the **transition** $u = x^2+y^2$ and $v = x^2-y^2$). A: We know that $u+v = 2x^2$ and $u-v = 2y^2$ so $x = \sqrt{({}^{u+v}/_2)}$ and $y = \sqrt{({}^{u-v}/_2)}$. So $J(u,v) = \left|{}^{\partial x/\partial u}_{\partial y/\partial u}\ {}^{\partial x/\partial v}_{\partial y/\partial v}\right|$. Calculate this *directly* (${}^{\partial x}/_{\partial u} = {}^1/_{2\sqrt{2}}(u+v)^{-1/2}$, ${}^{\partial x}/_{\partial v} = {}^1/_{2\sqrt{2}}(u+v)^{-1/2}$) to get $J(u,v) = {}^1/_{4\sqrt{(u^2-v^2)}}$. Therefore, the *integral is* $\int_1^4\int_4^9 J(u,v)\,xy = \int_1^4\int_4^9 {}^{\sqrt{(u^2-v^2)}}/_2 \times {}^1/_{4\sqrt{(u^2-v^2)}}\,dudv = {}^1/_8\int_1^4\int_4^9 dudv = {}^{15}/_8$.



Exam Paper: January 2000

SECTION 1 (Compulsory)

(1) (a) Calculate the Hessian of the function:
$f(x, y) = -xye^{-(x^2+y^2)/2}$.
Find and classify all the critical points of f. **[12 marks]**

(b) The volume V of a sphere of radius R is given in cartesian coordinates by
$V = 2\iint_\Omega \sqrt{(R^2 - (x^2 + y^2))}\,dxdy$,
where Ω is the disc of radius R centred at the origin. Use polar coordinates to show that $V = {}^4/_3\pi R^3$. **[8 marks]**

SECTION 2 (Answer 2 out of 4 questions)

(2) By using the method of Lagrange multipliers, minimise xyz subject to the constraint $8x^3-8y^3+27z^3 = 1$, with $x \geq 0$, and $z \geq 0$, whilst $y \leq 0$. **[15 marks]**

(3) (a) Evaluate the triple integral $\int_0^2 \int_0^x \int_0^{4-x^2} xyz \, dz\, dy\, dx$. **[8 marks]**

(b) Use the transformation u = x+y, v = 2x-y to simplify the integral $\iint_\Omega (x+y)^2 dxdy$ where Ω is the parallelogram bounded by the lines x+y=0, x+y=1, 2x-y=0, 2x-y=3. **[7 marks]**

(4) (a) Determine whether or not the following function is continuous at the origin. If the function is continuous then prove it, otherwise provide reasons for the lack of continuity: $f(x, y) = \frac{2xy}{x^2+y^2}$, if $(x, y) \neq (0, 0)$; $f(0, 0) = 0$. **[6 marks]**

(b) A function $f: S \to \mathbb{R}$ is defined on a convex subset S of $\mathbb{R}^n$. You are told that at some point $\mathbf{x_0} \in S$, f has a local minimum (so there is some ε > 0 such that if $\|\mathbf{x}-\mathbf{x_0}\| < \varepsilon$ then $f(\mathbf{x}) \geq f(\mathbf{x_0})$.) Supposing that f is a convex function on S, prove that f has a global minimum at $\mathbf{x}_0$. **[9 marks]**

(5) Consider the mapping given by $u = e^x\cosh(y)$, $v = e^x\sinh(y)$.

(i) Find the Jacobian $\partial(u, v)/\partial(x, y)$ of the transformation.
Determine where local inverses exist. **[6 marks]**

(ii) Solve for x, y in terms of u, v to find the inverse mapping and calculate its Jacobian, giving your answer in terms of u, v and also in terms of x, y: $\partial(x, y)/\partial(u, v)$. **[6 marks]**

Finally verify $\partial(u, v)/\partial(x, y)\ \partial(x, y)/\partial(u, v) = I$, the two-by-two identity matrix. **[3 marks]**

(Questions done: 1, 3, 5)