KORT SAMMANFATTNING AV FLERVARIABELKURSEN

LARS SVENSSON OCH ERIC NORDENSTAM

1. Normer på vektorrum

En avbildning $V \ni x \mapsto |x| \in \mathbb{R}$ från ett vektorrum över $\mathbb{R}$ (eller $\mathbb{C}$) kallas en norm om $\forall \lambda \in \mathbb{R}$ (eller $\mathbb{C}$) $\quad \forall x, y \in V$

(1) $|\lambda x| = |\lambda||x|$
(2) $|x + y| \leq |x| + |y|$ (triangelolikheten)
(3) $|x| \geq 0$
(4) $|x| = 0 \Rightarrow x = 0$

Ett vektorrum försett med en norm kallas ett normerat rum. Låt oss som vanligt införa notationen $B(p, r) = \{x \in V : |x - p| < r\}$ för bollen med radie r och centrum i p. Till ett normerat rum hör de vanliga topologiska begreppen som: om $M \subset V$ så är
p inre punkt i M om $B(p, r) \subset M$ för något $r > 0$,
p yttre punkt till M om p är inre punkt i $V \setminus M$,
p randpunkt till M om p är varken inre punkt eller yttre punkt till M.

Mängden av inre punkter till M betecknas $\overset{\circ}{M}$ och mängden av randpunkter ∂M. Unionen av M och dess rand kallas slutna höljet av M och betecknas $\bar{M}$. M kallas

(1) *öppen* om $\overset{\circ}{M} = M$.
(2) *sluten* om $\bar{M} = M$.
(3) *begränsad* om $M \subset B(0, r)$ för något $r > 0$.

2. Konvergens

En följd $\mathbb{N} \ni n \mapsto x_n \in V$ säges konvergera mot x då $n \to \infty$ om $\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ n \geq N \Rightarrow |x - x_n| < \varepsilon$.

2.1. **Cauchyföljder.** En följd $\mathbb{N} \ni n \mapsto x_n \in V$ kallas Cauchy om $\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ n, m \geq N \Rightarrow |x_n - x_m| < \varepsilon$.

Som vi vet så konvergerar varje Cauchyföljd i $\mathbb{R}^n$ eller $\mathbb{C}^n$, men detta är inte i allmänhet fallet om V är ett oändligtdimensionellt rum. Vi har därför följande viktiga definition.

Definition 2.2. Ett normerat rum i vilket varje Cauchyföljd konvergerar kallas ett *Banachrum*.

Med denna definition så visade vi tidigare i kursen att mängden av kontinuerliga funktioner från en kompakt i $\mathbb{R}^n$ till $\mathbb{R}$ med max-normen är ett Banachrum.

3. Operatornorm och ekvivalenta normer

Låt V och W vara två ändligtdimensionella vektorrum över $\mathbb{R}$ (eller $\mathbb{C}$). Mängden $L(V,W)$ av linjära avbildningar från V till W är själv ett vektorrum genom additionen och multiplikationen med skalär, definierade enligt

$$(S+\lambda T)(x) = S(T) + \lambda T(X)$$

för S, $T \in L(V,W)$, $x \in V$ och $\lambda \in \mathbb{R}$ (eller $\mathbb{C}$).

Vi ska införa en norm på $L(V,W)$, men först ska vi titta lite på V. Låt $v_1, \ldots, v_n$ vara en bas för V och låt $K : V \to \mathbb{R}^n$ (eller $\mathbb{C}^n$) vara koordinatavbildningen hörande till vår bas, dvs. om $K(v) = x = (x_1, \ldots, x_n)$ så $v = x_1v_1 + \cdots + x_nv_n$. Vi har då $|v| \leq |x_1||v_1| + \cdots + |x_n||v_n|$ som med Cauchys olikhet i $\mathbb{R}^n$ ger

$$|v| \leq (|v_1|^2 + \cdots + |v_n|^2)^{1/2}(|x_1|^2 + \cdots + |x_n|^2)^{1/2}.$$

Om $B = (|v_1|^2 + \cdots + |v_n|^2)^{1/2}$ och $|x|$ betecknar den Euklidiska normen i $\mathbb{R}^n$ (eller $\mathbb{C}^n$) så har vi alltså

$$|v| \leq B|x| \quad \forall v \in V,\ x = K(v).$$

Detta visar också att avbildningen

$$\mathbb{R}^n \ni x \mapsto |x_1v_1 + \cdots + x_nv_n| \in \mathbb{R}$$

är kontinuerlig och således antar ett minimum, säg A, på den kompakta mängden $\{x \in \mathbb{R}^n : |x| = 1\} = S$. Således har vi att

$$A \leq |x_1v_1 + \cdots + x_nv_n| \quad \forall x \in S.$$

Genom att skala x till godtycklig norm och notera att alla leden nedan är homogena under denna skalning ser vi att

$$A|x| \leq |x_1v_1 + \cdots + x_nv_n| = |v| \leq B|x| \quad \forall x \in \mathbb{R}^n.$$

Övning 1. Visa med hjälp av ovan att alla ändligtdimensionella normerade rum över $\mathbb{R}$ (eller $\mathbb{C}$) är Banachrum.

Övning 2. Visa också att om $|\cdot|_1$ och $|\cdot|_2$ är normer i ett vektorrum V av ändlig dimension, så existerar strikt positiva reella tal α och β sådana att $\alpha|x|_1 \leq |x|_2 \leq \beta|x|_1\ \forall x \in V$.

Vi säger att alla normer på ett ändligtdimensionellt vektorrum är ekvivalenta då de ger upphov till samma konvergensbegrepp och samma topologiska begrepp.

Låt $T \in L(V,W)$. Eftersom enhetsbollen i V är kompakt (=sluten och begränsad för ändligtdimensionella normerade vektorrum) och $V \ni x \mapsto$

$|T(x)| \in \mathbb{R}$ är kontinuerlig så antar $|T(x)|$ ett maximum där. Vi definierar operatornormen enligt

$$\|T\| = \max_{|x|\leq 1} |T(x)|.$$

Övning 3. Verifiera att $\|\cdot\|$ är en norm.

Om V och W har dimension n och m respektive så har $L(V, W)$ dimensionen nm varför i detta fall även $L(V, W)$ blir ett Banachrum.

Övning 4. Låt $T \in L(U, V)$, $S \in L(V, W)$ där U, V och W är ändligtdimensionella normerade rum. Visa att

$$\|S \circ T\| \leq \|S\| \cdot \|T\|.$$

4. Euklidiska normen på matriser

Vi låter $\mathbb{R}^{n\times m}$ ($\mathbb{C}^{n\times m}$) beteckna mängden av alla reella (komplexa) $n \times m$-matriser. På dessa har vi inre produkten

$$\begin{aligned} \mathbb{R}^{n\times m} \times \mathbb{R}^{n\times m} &\stackrel{\cdot}{\longrightarrow} \mathbb{R} \\ (A, B) &\longmapsto \operatorname{sp} B^T A = B \cdot A. \end{aligned}$$

Övning 5. Hur ser motsvarigheten ut i $\mathbb{C}^{n\times m}$?

Inre produktens symmetri följer av

$$\begin{aligned} A \cdot B = \operatorname{sp} A^T B = \sum_i (A^T B)_{ii} &= \sum_i \sum_j (A^T)_{ij} (B)_{ji} \\ &= \sum_i \sum_j A_{ij} B_{ij} = \operatorname{sp} B^T A = B \cdot A. \end{aligned}$$

Vi har följande variant på Cauchys olikhet.

Sats 4.1. *Om $A \in \mathbb{R}^{n\times m}$ och $B \in \mathbb{R}^{n\times m}$ så*

$$|AB| \leq |A||B|$$

där $|C| = (C \cdot C)^{1/2}$ är den euklidiska matrisnormen.

Bevis.

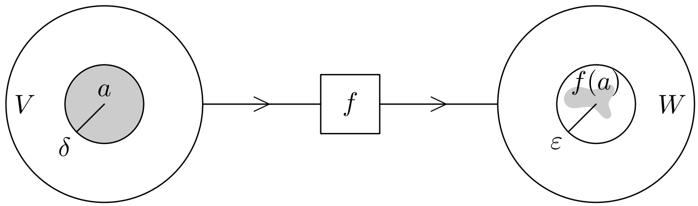
$$\begin{aligned} |AB|^2 &= \sum_{ij} ((AB)_{ij})^2 = \sum_{ij} \Big(\sum_k A_{ik} b_{kj} \Big)^2 \\ &\leq \text{Cauchy på } \mathbb{R}^n \leq \sum_{ij} \Big(\sum_k (A_{ik})^2 \Big) \Big(\sum_l (B_{lj})^2 \Big) \\ &= \sum_{ijkl} (A_{ik})^2 (B_{lj})^2 = |A|^2 |B|^2 \end{aligned}$$

$\square$

Anmärkning. Förväxla inte den vanliga Cauchys olikhet, $|A \cdot B| \leq |A||B|$ med olikheten i satsen då ju $A \cdot B$ betecknar inre produkten medan AB betecknar matrisprodukten mellan A och B. Den förra är definierad bara då A och B är matriser av samma format.

5. Kontinuitet

En funktion $f : V \to W$ mellan två normerade vektorrum kallas kontinuerlig i punkten $a \in V$ om $\forall \varepsilon > 0\ \exists \delta > 0$ så att $f(B(a,\delta)) \subset B(f(a),\varepsilon)$.



Övning 6. Visa med hjälp av vårt tidigare resonemang om ekvivalenta normer att våra tidigare kända satser, om att kontinuerliga funktioner tar kompakter på kompakter och bågvisa sammanhängande mängder på dylika, gäller för ändligtdimensionella normerade vektorrum.

Övning 7. Låt $f : \mathbb{R}^n \to \mathbb{R}$ vara kontinuerlig överallt och att $|f(x)| \to \infty$ när $|x| \to \infty$. Visa att f antar ett minsta värde.

Övning 8. Låt $f : \mathbb{R}^n \to \mathbb{R}$ vara kontinuerlig överallt, $f(x) \to 0$ då $|x| \to \infty$ och att $\exists a \in \mathbb{R}^n$ med $f(a) > 0$. Visa att f antar ett största värde.

6. Derivation

En avbildning $f : V \to W$ mellan normerade rum säges vara deriverbar i a om det existerar en kontinuerlig linjär avbildning betecknad $f'(a) : V \to W$ sådan att

$$f(a+h) = f(a) + f'(a)(h) + o(h)$$

där $o(h) = |h|\varepsilon(h)$ för något $\varepsilon : V \to W$ sådan att $\varepsilon(h) \to 0 = \varepsilon(0)$ då $n \to 0$. Vi säger att f är deriverbar i $\Omega \subset V$ om $f'(a)$ existerar för varje $a \in \Omega$. Således kan f' uppfattas som en avbildning från Ω till $L(V,W)$.

7. Kedjeregeln

Sats 7.1. *Om $f : U \to V$ är deriverbar i a och $g : V \to W$ är deriverbar i $b = f(a)$ så är $g \circ f$ deriverbar i a med derivata $(g \circ f)'(a) = g'(b) \circ f'(a)$.*

Bevis. Låt $h \in V$ och sätt $k = f(a+h) - f(a)$. Då går k mot noll då h går mot noll och $k = f'(a)(h) + o(h)$. Vi har nu

$$\begin{aligned}(g \circ f)(a+h) &= g(b+k) = g(b) + g'(b)(k) + o(k) = \\ &= (g \circ f)(a) + (g'(b) \circ f'(a))(h) + g'(b)(o(h)) + o(k).\end{aligned}$$

Det räcker således att visa att

$$g'(b)(o(h)) + o(k) = o(h).$$

Om $k \neq 0$ och $h \neq 0$ så har vi

$$\frac{|g'(b)(o(h)) + o(k)|}{|h|} \leq \|g'(b)\| \frac{|o(h)|}{|h|} + \frac{|o(k)|}{|k|} \cdot \frac{\|f'(a)\||h| + |o(h)|}{|h|}$$

som $\to 0$ då $h \to 0$. Om $k = 0$ är saken trivial eftersom vi då har $(g \circ f)(a+h) = (g \circ f)(a)$. □

8. Medelvärdesolikheten

Sats 8.1. *Om $f : \mathbb{R}^n \to \mathbb{R}^m$ är deriverbar på segmentet $[x, y] = \{tx + (1-t)y : 0 \leq t \leq 1\}$ och $\|f'(z)\| \leq K$ då $z \in [x, y]$ så gäller*

$$|f(x) - f(y)| \leq K|x - y|$$

och om $m = 1$ så existerar $z \in [x, y]$ sådan att

$$f(x) - f(y) = f'(z)(x - y).$$

Bevis. Tag enhetsvektor e i $\mathbb{R}^m$ och definiera $g : [0, 1] \to \mathbb{R}$ genom $g(t) = f(tx + (1-t)y) \cdot e$. Av kedjeregeln följer att $g'(t) = (f'(tx + (1-t)y)(x - y)) \cdot e$. Av medelvärdessatsen på g följer då att $\exists \theta \in (0, 1)$ sådan att $g(1) - g(0) = (f(x) - f(y)) \cdot e = (f'(\theta x + (1-\theta)y)(x-y)) \cdot e$. Således har vi $|(f(x) - f(y)) \cdot e| \leq K|x - y|$ för varje enhetsvektor e som ger att $|f(x) - f(y)| \leq K|x - y|$. Om $m = 1$ väljer vi bara $e = 1$ och får för $z = \theta x + (1-\theta)y$ att $f(x) - f(y) = f'(z)(x - y)$. □

Anmärkning. Resultatet ovan gäller för godtyckliga normerade rum men beviset kräver något som vi väntar med.

Vi ska nu trimma vår olikhet en smula. Låt f vara en deriverbar funktion från $\mathbb{R}^n$ till $\mathbb{R}^m$ och definiera för fixt $y \in \mathbb{R}^n$ en funktion F genom

$$F(x) = f(x) - f(y) - f'(y)(x - y)$$

Då är $F(y) = 0$ och av kedjeregeln har vi $F'(x) = f'(x) - f'(y)$, varför medelvärdesolikheten ger $|F(x) - F(y)| \leq \max_{z \in [x,y]} |F'(z)||x - y|$, dvs. $|f(x) - f(y) - f'(y)(x - y)| \leq \max_{z \in [x,y]} |f'(x) - f'(y)||x - y|$

Anmärkning. Denna olikhet gäller också för alla normerade rum om max ersätts med supremum.

9. Matrisrepresentation av derivator

Låt $f : \mathbb{R}^n \to \mathbb{R}^m$ vara deriverbar i $a \in \mathbb{R}^n$ och välj standardbasvektorer $e_i \in \mathbb{R}^n$ och $e_j \in \mathbb{R}^m$. Definiera $g(t) = e_j \cdot f(a + te_i)$ och få enligt kedjeregeln $g'(0) = e_j \cdot (f'(a)(e_i))$ som är element på rad j och kolonn i i matrisrepressentationen för $f'(a)$ i standardbaserna för $\mathbb{R}^n$ och $\mathbb{R}^m$ respektive.

Vidare om $f_j(x) = e_j \cdot f(x)$ så får vi att

$$\frac{\partial f_j(x)}{\partial x_i} = e_j \cdot (f'(a)(e_i)).$$

10. Variabelnotation

Om x och y är variabler i $\mathbb{R}^n$ och $\mathbb{R}^m$ respektive som är relaterade enligt $y = f(x)$ för en deriverbar funktion $f : \mathbb{R}^n \to \mathbb{R}^m$ så skriver vi ofta

$$f'(x) = Df(x) = \frac{dy}{dx} = \frac{d\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}{d\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}.$$

Kedjeregeln blir då med matrismultiplikation

$$\mathbb{R}^n \to \mathbb{R}^m \to \mathbb{R}^k$$

$$x \ \mapsto y \ \mapsto z$$

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}.$$

11. Inversa funktionssatsen

Sats 11.1 (Inversa funktionssatsen i normerad form)**.** *Låt $f : \mathbb{R}^n \to \mathbb{R}^n$ vara kontinuerligt deriverbar i en omgivning av origo och antag att $f(0) = 0$ och $f'(0) = \mathrm{id} = I$. Då existerar två omgivningar U och V till origo samt en kontinuerligt deriverbar $g : V \to U$ sådan att $f \circ g = \mathrm{id}$ på V och $g \circ f = \mathrm{id}$ på U.*

Bevis. Då f' är kontinuerlig i origo existerar $r > 0$ sådan att $|f'(x) - I| \le \frac{1}{2}$ om $|x| < r$. Fixera ett $y \in \mathbb{R}^n$ med $|y| < r/2$ och definiera $\varphi(x) = y + x - f(x)$. Vi observerar att om x är en fixpunkt till φ, dvs. $\varphi(x) = x$, så har vi att $y = f(x)$, dvs. x är "inversa bilden av y" under f.

Vi visar först att φ avbildar $\bar{B}(0, r')$ in i sig själv om $r' = |y| \cdot 2 < r$. Så tag $|x| \le r'$. Då har vi $|\varphi(x)| = |\varphi(x) - \varphi(0) + \varphi(0)| \le |\varphi(x) - \varphi(0)| + |y| \le \max_{z \in [0,x]} |\varphi'(z)||x - 0| + |y| \le \frac{1}{2}r' + \frac{1}{2}r' = r'$. Enligt Banachs fixpunktssats existerar ett unikt $x = g(y)$ med $|x| \le 2|y|$ sådant att $f(x) = y$. Vi ska nu visa att g är kontinuerligt deriverbar i $B(0, r/2) = V$. Vi visar först att g är deriverbar i origo med $g'(0) = \mathrm{id} = I$. Om $f(x) = f(0) + f'(0)(x) + o(x) = x + o(x) = y$ så har vi $x = g(y) = y - o(x)$ men $g(0) = 0$ och

$$\frac{|o(x)|}{|y|} = \frac{|o(x)|}{|x|}\frac{|x|}{|y|} \leq 2\frac{|o(x)|}{|x|}$$

och om $y \to 0$ så $x \to 0$, så g är deriverbar i origo med $g'(0) = I$. Fixera nu $x_0 = g(y_0)$ och sätt $x = x_0 + s$ och definiera $y = f(x)$ och $F(s) = t = f'(x_0)^{-1}(f(x_0 + s) - y_0)$. Då har vi $F(0) = 0$ och $F'(0) = f'(x_0)^{-1}f'(x_0) = I$ varför enligt ovan F har en lokal invers G som är deriverbar i origo med $G'(0) = I$. Vi har då $G \circ F(s) = s = x - x_0 = G \circ f'(x_0)^{-1}(f(x) - y_0) \Rightarrow g(y) = x_0 + G \circ f'(x_0)^{-1}(y - y_0)$. Av kedjeregeln är högra ledet deriverbart i $y = y_0$ med derivatan $G'(0)f'(x_0)^{-1} = f'(x_0)^{-1}$. Detta visar att $g'(y_0)$ existerar och att $g'(y_0) = f'(x_0)^{-1} = f'(g(y_0))$. Detta visar också att g' är kontinuerlig eftersom g och f' är kontinuerliga. (Vi vet ju, sedan tidigare, att inverser till kontinuerliga avbildningar är kontinuerliga.) □

Vi ska nu generalisera satsen en smula.

Sats 11.2 (Inversa funktionssatsen)**.** *Låt f vara en kontinuerligt deriverbar avbildning från en omgivning av a i $\mathbb{R}^n$ till $\mathbb{R}^m$.*

- *(i) Om $f'(a)$ är injektiv så existerar en kontinuerligt deriverbar avbildning g från en omgivning till $f(a)$ sådan att $g \circ f$ är identiteten i närheten av a. Vi säger att f har en lokal vänsterinvers.*
- *(ii) Om $f'(a)$ är surjektiv så existerar en lokal högerinvers g till f i en omgivning av $f(a)$ sådan att $f \circ g$ är identiteten i närheten av $f(a)$.*
- *(iii) Om $f'(a)$ är bijektiv så har f en lokal invers.*

Bevis. *(i)* Låt A vara linjär : $\mathbb{R}^m \to \mathbb{R}^n$ sådan att $A \circ f'(a) = \mathrm{id} = I$ på $\mathbb{R}^m$. Bilda $F(x) = A(f(a + x) - f(a))$ då har vi $F(0) = 0$ och $F'(0) = I$. Alltså existerar en kontinuerligt deriverbar G i närheten av origo med $G(0) = 0$ och $G'(0) = I$ sådan att $x = G \circ A(f(a + x) - f(a))$ dvs. $x - a = G \circ A(f(x) - f(a))$ om x ligger nära a, eller $x = a + G \circ A(f(x) - f(a))$ Sätt nu $g(y) = a + G \circ A(y - f(a))$. Så får vi $g \circ f(x) = x$, som visar *(i)*.

(ii) Låt $B : \mathbb{R}^m \to \mathbb{R}^n$ vara linjär sådan att $f'(a)B = \mathrm{id} = I$ på $\mathbb{R}^m$. Bilda $F(x) = f(a + Bx) - f(a)$ då är F kontinuerligt deriverbar nära origo och $F(0) = 0, F'(0) = f'(a)B = I$. Resten av beviset överlämnas som övning. □

12. Implicita funktionssatsen

Låt $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ vara kontinuerligt deriverbar i närheten av $(a, b) \in \mathbb{R}^m \times \mathbb{R}^n$ och anta att derivatan av

$$\mathbb{R}^n \ni y \mapsto f(a, y) \in \mathbb{R}^n$$

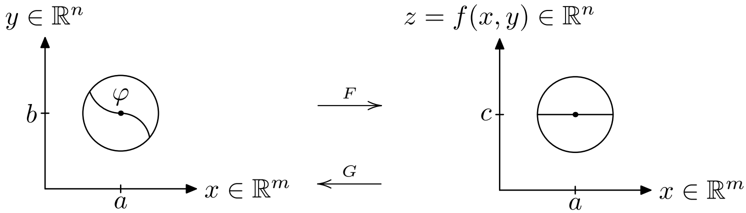
är inverterbar i b. Då existerar en kontinuerligt deriverbar funktion φ definierad i en omgivning till a sådan att $b = \varphi(a)$ och där $f(x, \varphi(x)) = f(a, b)$ för x nära a.

Bevis. Bilda $\mathbb{R}^m \times \mathbb{R}^n \xrightarrow{F} \mathbb{R}^m \times \mathbb{R}^n$, $(x, y) \mapsto (x, f(x, y))$. Om vi inför variabelnotation $z = f(x, y))$ och arbetar med kolonnvektorer fås

$$F' = \frac{d\begin{bmatrix} x \\ z \end{bmatrix}}{d\begin{bmatrix} x \\ y \end{bmatrix}} = \begin{bmatrix} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial y} \\ \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{bmatrix}$$

där I_m är identitetsmatrisen av formatet $m \times m$. $\det F' = \det \frac{dz}{dy} \neq 0$ i punkten $(x, y) = (a, b)$. Således har F en lokal invers G.

Betrakta figuren.



Låt P beteckna projektionen

$$\mathbb{R}^m \times \mathbb{R}^n \ni (s, t) \mapsto t \in \mathbb{R}^n$$

och definiera

$$\varphi(x) = P \circ G(x, c) \quad \text{för } x \text{ nära } a.$$

Då har vi $F \circ G(x, c) = (x, c)$ dvs. $F(x, \varphi(x)) = (x, f(x, \varphi(x))) = (x, c)$ som visar att $f(x, \varphi(x)) = c$ och att $\varphi(a) = P \circ G(a, f(a, b)) = P \circ G \circ F(a, b) = b$. $\square$

13. Parameteriserade mångfalder

Låt $f : \Omega \to \mathbb{R}^n$ vara en deriverbar funktion där Ω är en öppen del av $\mathbb{R}^m$ och där $1 \leq m \leq n$. Antag att $f'(x)$ är injektiv $\forall x \in \Omega$. Då kallas $f(\Omega)$ en m-dimensionell mångfald (yta) i $\mathbb{R}^n$. Om $\Omega \ni a \mapsto f(a) = p \in f(\Omega) = S$ så kallas $p + \operatorname{Im} f'(a) = T_pS$ för tangentrummet till S i p.

14. Mångfalder som nivåytor

Låt $f : \Omega \to \mathbb{R}^m$ vara deriverbar $(m \leq n)$ i $\Omega \subset \mathbb{R}^n$ där $f'(p)$ är surjektiv för varje p i den öppna mängden Ω. Definiera

$$S = \{x \in \Omega : f(x) = 0\}.$$

Om $p \in S$ så definieras

$$T_pS = \{x \in \mathbb{R}^n : f'(p)(x-p) = 0\}.$$

Vi ser att $T_pS = p + \ker f'(p)$ så T_pS, som kallas tangentrummet till S i punkten p, är en affin mängd i $\mathbb{R}^n$ av dimensionen $n-m$. Om $m=1$ är S en hyperyta och om $m>1$ är S skärningen av m stycken hyperytor. Eftersom $f'(p)$ antogs vara surjektiv så har matrisen

$$f'(p) = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}_{x=p,y=f(x)}$$

linjärt oberoende rader. Men då kan vi välja ut m stycken kolonner, säg de m första, så att

$$\det \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_m} \end{bmatrix}_{x=a} \neq 0$$

Men då följer av implicita funktionssatsen att variablerna $x_{m+1}, \ldots, x_n$ kan lösas ut som funktion av $x_1, \ldots, x_m$ ur ekvationen $f(x) = 0$, nära a. Säg att $(x_{m+1}, \ldots, x_n) = \varphi(x_1, \ldots, x_m)$. Det betyder att S kan parameteriseras enligt

$$(x_1, \ldots, x_m) \to (x_1, \ldots, x_m, \varphi(x_1, \ldots, x_m))$$

lokalt nära $a = (a_1, \ldots, a_n)$.

15. Taylors formel

Antag att $f : \mathbb{R}^m \to \mathbb{R}$ är N gånger deriverbar i $a \in \mathbb{R}^n$ och definiera för fixt $h \in \mathbb{R}^n$ en avbildning $g : \mathbb{R} \to \mathbb{R}$ genom $g(t) = f(a+th)$. Kedjereglen ger då

$$\begin{aligned} g'(t) = f'(a+th)(h) &= \begin{bmatrix} \frac{\partial f}{\partial x_1}(a+th) & \ldots & \frac{\partial f}{\partial x_n}(a+th) \end{bmatrix} h \\ &= h_1D_1f(a+th) + \cdots + h_nD_nf(a+th) \\ &= (h_1D_1 + \cdots + h_nD_n)f(a+th) \end{aligned}$$

där $D_j = \frac{\partial}{\partial x_j}$.

Låt oss införa notationen

$$h \cdot D = h_1D_1 + \cdots + h_nD_n$$

som en linjär avbildning som tar kontinuerligt deriverbara funktioner (på $\mathbb{R}^n$) till kontinuerliga funktioner. Vi skriver då $g'(t) = (h \cdot D)f(a+th)$. Genom att sätta $f_1 = (h \cdot D)f$ och $g_1(t) = f_1(a+th)$ så ser vi att $g''(t) = g_1'(t) = (h \cdot D)f_1(a+th) = (h \cdot D)^2 f(a+th)$. Förfarandet itereras och ger

$$g^{(k)}(t) = (h \cdot D)^k f(a+th).$$

Men

$$g(t) = g(0) + g'(0)t + \cdots + \frac{1}{(N-1)!} g^{(N-1)}(0)t^{N-1} + \frac{1}{N!} g^{(N)}(\theta t).$$

Således har vi visat att

$$f(a+h) = \sum_{k=0}^{N-1} \frac{1}{k!} (h \cdot D)^k f(a) + \frac{1}{N!} (h \cdot D)^N f(p).$$

16. Några olika skrivsätt

Translatet $T_h f : \mathbb{R}^m \to \mathbb{R}$ definieras av

$$(T_h f)(x) = f(x+h).$$

Om f är oändligt deriverbar och om Taylorserien konvergerar överallt kan vi alltså skriva

$$(T_h f)(a) = \sum_0^\infty \frac{1}{k!} (h \cdot D)^k f(a)$$

eller om vi tar bort a

$$T_h f = \sum_0^\infty \frac{1}{k!} (h \cdot D)^k f$$

eller om vi tar bort f

$$T_h = \sum_0^\infty \frac{1}{k!} (h \cdot D)^k.$$

Det är också naturligt att definiera

$$e^{h \cdot D} = \sum_0^\infty \frac{1}{k!} (h \cdot D)^k.$$

Taylor formel får då utseendet

$$T_h = e^{h \cdot D},$$

Vackert va!

Det finns också ett annat sätt att skriva Taylors formel, nämnligen med s.k. multiindexnotation: En lista $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ kallas ett multiindex. Om $\alpha, \beta \in \mathbb{N}^n$ så definieras summan

$$\alpha + \beta = (\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n)$$

och graden

$$|\alpha| = \alpha_1 + \cdots + \alpha_n$$

och fakulteten

$$\alpha! = (\alpha_1!)(\alpha_2!) \cdots (\alpha_n!).$$

Om $x = (x_1, \ldots, x_n)$ är en lista av variabler definieras x upphöjt till α av

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

dvs. ett monom av grad $|\alpha|$. Om $D = (D_1, \ldots, D_n)$ är listan av de partiella deriveringarna sätter vi $D^\alpha = (D_1^{\alpha_1}, \ldots, D_n^{\alpha_n})$. Med ovanstående notation kan Taylors formel skrivas precis som i envariabelkursen

$$f(a+th) = \sum_{|\alpha|<N} \frac{1}{\alpha!} D^\alpha f(a) h^\alpha + \text{rest}$$

där

$$\text{rest} = \sum_{|\alpha|=N} \frac{1}{\alpha!} D^\alpha f(p) h^\alpha$$

för något $p \in [a, a+h]$.

Övning 9. Visa detta!

$$(x_1, x_2, x_3, x_4)^{(1,2,0,3)} = x_1 x_2^2 x_4^3, \tag{1}$$

$$(1,2,0,3)! = (1!)(2!)(0!)(3!) = 1 \cdot 2 \cdot 1 \cdot 6 = 12. \tag{2}$$

17. Max och min

Definition 17.1. En punkt a kallas kritisk till $f : \mathbb{R}^n \to \mathbb{R}$ om $f'(a) = 0$.

Vi har omedelbart följande enkla sats.

Sats 17.2. *Om f har ett lokalt min eller max i a så är a kritisk.*

Bevis. Sätt för en enhetsvektor e i $\mathbb{R}^n$ $g(t) = f(a+te)$. Då har g lokalt max eller min i 0 så $g'(0) = f'(a)e = 0$. Men e var godtycklig så $f'(a) = 0$. $\square$

Låt nu $f : \mathbb{R}^n \to \mathbb{R}$ vara två gånger deriverbar. Enligt Taylors har vi då

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2}(h \cdot D)^2 f(p)$$

för något p mellan a och $a+h$. Men

$$(h \cdot D)^2 f(x) = (\sum_{i,j} h_i h_j D_i D_j) f(x)$$

$$= \sum_{i,j} h_i (D_i D_j f(x)) h_j = h^T f''(x) h$$

där $f''(x)$ betecknar den s.k. Hessianen

$$f''(x) = \begin{bmatrix} D_1 D_1 f(x) & \ldots & D_1 D_n f(x) \\ \vdots & & \vdots \\ D_n D_1 f(x) & \ldots & D_n D_n f(x) \end{bmatrix}.$$

Alltså har vi

$$f(a+h) = f(a) + f'(a)(h) + \frac{1}{2} h^T f''(p) h.$$

Anmärkning. Om $D = (D_1 \dots D_n)$ så kan vi symboliskt skriva

$$f'' = D^T Df = \begin{bmatrix} D_1 \\ \vdots \\ D_n \end{bmatrix} \begin{bmatrix} D_1 & \dots & D_n \end{bmatrix} f.$$

18. Karakterisering av kritiska punkter

Låt a vara kritisk till $f : \mathbb{R}^n \to \mathbb{R}$. Då har vi att

$$f(a+h) = f(a) + \frac{1}{2} h^T f''(p)h, \quad p \in (a, a+h).$$

Vi har omedelbart följande.

Sats 18.1. *Om f'' är positivt definit i en omgivning av a så har f ett lokalt minimum i a. Om $f''(p)$ är negativt definit i en omgivning av a så har f ett lokalt maximum i a.*

Anmärkning. Om f'' är kontinuerlig i a och om $f''(a)$ är pos. def. resp neg. def. i a så är även $f''(p)$ pos. def. resp neg. def. i en omgivning av a på grund av kontinuiteten.

Vi har också

Sats 18.2. *Om f'' är kontinuerlig i a och om* $\det f''(a) \neq 0$ *samt att $f''(a)$ har både strikt positiva och negativa egenvärden, dvs. är indefinit, så har f en sadelpunkt i a (dvs. varken lokalt max eller min).*

Bevis. Övning. Använd att egenvärdena till en matris varierar kontinuerligt med matrisens element. □

Anmärkning. I det fall att $\det A = 0$ kräver en helt annan analys för bestämning av vilken karaktär en kritisk punkt har ($A = f''$). Detta problem är mycket svårt och leder in i singularitetsteori och s.k. katastrofteori.

Om man ska avgöra om en symmetrisk $n \times n$-matris A är pos. def, neg. def eller indefinit så har man stor nytta av följande resultat, som vi visade i linjär algebra kursen.

19. Positivt och negativt definita matriser

Sats 19.1. *Om A är en reell symmetrisk $n \times n$-matris och om D_k definieras som determinanten av matrisen $[a_{ij}]_{1 \leq i,j \leq k}$ så gäller att A är positivt definit precis om $D_k > 0$ $\forall k = 1, 2, \dots, n$.*

Bevis. Vi påminner om att en matris är positivt definit precis om $h^T Ah > 0$ för alla nollskilda vektorer h i $\mathbb{R}^n$. Detta är också ekvivalent med att alla egenvärden till A är strikt positiva. Antag först att A är pos. def. Betrakta underrummet $E_k = \text{linspan}\{e_1, \dots, e_k\}$. Om $0 \neq h = h_1 e_1 + \dots + h_k e_k \in E_k$ så har vi $h^T Ah = \sum_{i,j=1}^k h_i a_{ij} h_j > 0$. Detta betyder att den matris vi får om vi stryker alla rader och kolonner i A utom de k första är pos. def. Men då är alla egenvärden hos denna positiv och därmed dess determinant.

Antag nu omvänt $D_k > 0$ för alla $k \leq n$. Vi använder nu induktion över n och antar att satsen är visad för matriser av mindre format än $n \times n$ (trivialt för 1×1). Då är alltså A positivt definit på E_{n-1}. Om A inte är positivt definit så måste minst två egenvärden till A vara negativa (eller ett med algebraisk multiplicitet minst 2). Låt L vara ett tvådimensionellt underrum i $\mathbb{R}^n$ som spänns av motsvarande egenvektorer. Eftersom $\dim L + \dim E_{n-1} > n$ så existerar en nollskild $h \in L \cap E_{n-1}$ Men då har vi en motsägelse eftersom $h \in L \to h^T A h < 0$ och $h \in E_{n-1} \to h^T A h > 0$. $\square$

Om $D_1 < 0$, $D_2 > 0$, $D_3 < 0$, ... så är A negativt definit (ty då är ju $-A$ positivt definit). I övriga fall, då $D_n = \det A \neq 0$ är A indefinit.

20. Extremvärdesproblem med bivilkor

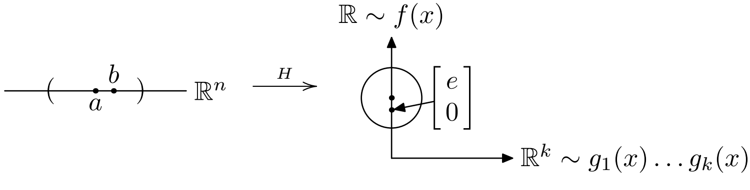
Låt f, g_1, ..., g_k vara kontinuerligt deriverbara funktioner från $\mathbb{R}^n$ till $\mathbb{R}$ och anta att $k < n$. Problem: maximera eller minimera $f(x)$ då $g_1(x) = \cdots = g_k(x) = 0$. Vi har nu följande användbara sats.

Sats 20.1. *Om $a \in \mathbb{R}^n$ är ett max eller min till f under bivillkoren $g_1 = \cdots = g_k = 0$ så måste $f'(a), g_1'(a), \ldots, g_k'(a)$ vara linjärt beroende.*

Bevis. Antag motsatsen, dvs. att de är linjärt oberoende och definiera en hjälpfunktion $H : \mathbb{R}^n \to \mathbb{R} \times \mathbb{R}^k$ genom

$$\mathbb{R}^n \ni x \mapsto H(x) = \begin{bmatrix} f(x) \\ g_1(x) \\ \vdots \\ g_k(x) \end{bmatrix} \in \mathbb{R} \times \mathbb{R}^k.$$

Betrakta figuren.



Om $f'(a), g_1'(a), \ldots, g_k'(a)$ är linjärt oberoende så är $H'(a)$ surjektiv. Enligt inversa funktionssatsen har då H en lokal högerinvers, G säg, i närheten av $H(a)$. Om $\begin{bmatrix} e \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ tillhör denna "närhet" så uppfyller $b = G(\begin{bmatrix} e \\ 0 \\ \vdots \\ 0 \end{bmatrix})$ bivillkoren $g_1(b) = \cdots = g_k(b) = 0$ och $f(b) = e$. Genom att välja $e > f(a)$ respektive $e < f(a)$ ser vi att a varken är max- eller minpunkt. $\square$

21. Frenets formler

Låt $\mathbb{R} \ni t \overset{x}{\mapsto} x(t) \in \mathbb{R}^n$ vara en n gånger deriverbar kurva och låt $e_1(t), \ldots, e_n(t)$ vara den ortonormala bas för $\mathbb{R}^n$ som ges av Gram-Schmidts ortogonaliseringsförfarande tillämpad på $\dot{x}(t), \ddot{x}(t), \ldots$. Vi antar härvid att $\dot{x}(t), \ddot{x}(t), \ldots, x^{(n)}(t)$ är linjärt oberoende (det s.k. generiska fallet). Per konstruktion har vi att för $1 \leq k \leq n$,

$$\operatorname{linspan}\{x'(t), \ldots, x^{(k)}(t)\} = \operatorname{linspan}\{e_1(t), \ldots, e_k(t)\}.$$

Vidare förjer av konstruktionen att

$$\begin{aligned}
e_1'(t) &= \lambda_{11}(t)e_1(t) + \lambda_{12}(t)e_2(t) \\
e_2'(t) &= \lambda_{21}(t)e_1(t) + \lambda_{22}(t)e_2(t) + \lambda_{23}(t)e_3(t) \\
&\vdots \\
e_{k-1}'(t) &= \lambda_{k-1,1}(t)e_1(t) + \cdots + \lambda_{k-1,k}(t)e_k(t) \\
&\vdots \\
e_n'(t) &= \lambda_{n1}(t)e_1(t) + \cdots + \lambda_{nn}(t)e_n(t).
\end{aligned} \tag{3}$$

Vidare följer av $e_i(t) \cdot e_j(t) = \delta_{ij}$ att

$$e_i(t) \cdot e_j'(t) + e_i'(t) \cdot e_j(t) = 0 \quad \forall i, j.$$

Men enligt (3) har vi även att $e_i(t) \cdot e_j'(t) = \lambda_{ji}(t)$, där vi definierar $\lambda_{ji}(t) = 0$ om $1 + j < i$. Alltså har vi $\lambda_{ij}(t) = -\lambda_{ji}(t)$ $\forall i, j$ och följdaktligen

$$\begin{aligned}
e_1'(t) &= \lambda_{12}(t)e_2(t) \\
e_2'(t) &= \lambda_{21}(t)e_1(t) + \lambda_{23}(t)e_3(t) \\
&\vdots \\
e_{k-1}'(t) &= \lambda_{k-1,k-2}(t)e_{k-2}(t) + \lambda_{k-1,k}(t)e_k(t) \\
&\vdots \\
e_n'(t) &= \lambda_{n,n-1}(t)e_{n-1}(t).
\end{aligned}$$

Dessa formler brukar kallas Frenet's. Inför nu den ortogonala matrisen

$$E(t) = [e_1(t), \ldots, e_n(t)]$$

med $e_j(t)$ som den j:te kolumnen, och den antisymmetriska bandmatrisen

$$\Lambda = \begin{bmatrix} 0 & \lambda_{12} & 0 & 0 & \dots & 0 & 0 \\ -\lambda_{12} & 0 & \lambda_{23} & 0 & \dots & 0 & 0 \\ 0 & -\lambda_{23} & 0 & \lambda_{34} & \dots & 0 & 0 \\ 0 & 0 & -\lambda_{34} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda_{n-1,n} \\ 0 & 0 & 0 & 0 & \dots & -\lambda_{n-1,n} & 0 \end{bmatrix}.$$

Matrisen är bara nollskilld på sub och superdiagonalen.

Övning 10. Uttryck Frenets formel som en relation mellan matriserna $E(t)$, $E(t)'$ och $\Lambda(t)$.

Anmärkning. För fallet $n = 3$ brukar $\lambda_{12} = \kappa$ kallas krökning och $\lambda_{23} = \tau$ torsion. Λ får då formen

$$\begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix}$$

.

För praktiskt bruk kan det vara enklare att använda formlerna för κ och τ som står i kursboken.

22. Högre ordningens derivator

Låt $f : V \to W_0$ vara en oändligt deriverbara funktion mellan ändligtdimensionella normerade rum. Låt vidare $L(V, W)$ beteckna mängden av linjära avbildningar från V till W. Vi inför också beteckningen $L_n(V, W)$ för mängden av multilinjära avbildninngar från $V^n = V \times V \times \cdots \times V$ till W.

Vi har nu

$$\begin{aligned} f^{(0)} &= f : V \to W_0 \\ f^{(1)} &= f' : V \to L(V, W_0) = W_1 \\ f^{(1)} &= f'' : V \to L(V, W_1) = W_2 \\ &\vdots \\ f^{(n)} &= f^{(n-1)\prime} : V \to L(V, W_{n-1}) = W_n \\ &\vdots \end{aligned}$$

Vi ska nu konstruera en kanonisk avbildning från W_n till $L_n(V, W_0)$ och sedan ska vi visa att bilden av denna endast innehåller symmetriska multilinjära avbildningar. Uppenbarligen är $W_1 = L_1(V, W_0) = L(V, W_0)$. Antag vi har konstruerat en avbildning $W_{n-1} \xrightarrow{\sim} L_{n-1}(V, W_0)$ betecknad $T \mapsto \tilde{T}$. Vi ska nu definiera $W_n \xrightarrow{\sim} L(V, W_0)$ så tag $T \in W_n = L(V, W_{n-1})$

och $\widetilde{(h_1, h_2, \ldots, h_n)} \in V^n$. Definiera nu $\tilde{T}$ genom $\tilde{T}(h_1, h_2, \ldots, h_n) = T(h_n)(h_1, h_2, \ldots, h_{n-1})$.

Övning 11. Verifiera att $\tilde{T}$ är multilinjär. Vi definierar nu $D^n f$ genom

$$(D^n f)(x) = (\widetilde{f^{(n)}(x)})$$

och får alltså att $D^n f : V \to L_n(V, W_0)$.

Övning 12. Visa att Taylorpolynomet av f i punkten a kan skrivas

$$f(a+h) = f(a) + Df(a)(h) + \frac{1}{2!} D^2 f(a)(h,h) + \ldots$$

Vi ska nu visa att $D^n f(a)$ är en symmetrisk multilinjär avbildning. En multilinjär $T : V^n \to W$ kallas symmetrisk om för varje permutation π:

$$T(h_1, h_2, \ldots, h_n) = T(h_{\pi(1)}, h_{\pi(2)}, \ldots, h_{\pi(n)})$$

gäller för alla $(h_1, \ldots, h_n)$ i V^n.

Vi börjar med ett lemma.

Lemma 22.1. *Om $f : \mathbb{R}^2 \to \mathbb{R}$ är sådan att alla partiella derivatorna $\partial_1 f$, $\partial_2 f$, $\partial_1 \partial_2 f$ och $\partial_2 \partial_1 f$ existerar i en omgivning av (a, b) och där $\partial_1 \partial_2 f$ eller $\partial_2 \partial_1 f$ är kontinuerlig i (a, b) så $\partial_1 \partial_2 f(a, b) = \partial_2 \partial_1 f(a, b)$.*

Bevis. Fixera h och k och definiera

$$\begin{aligned} g(y) &= f(a+h, y) - f(a, y), \\ F(h,k) &= g(b+k) - g(b) \\ &= f(a+h, b+k) - f(a, b+k) - f(a+h, b) + f(a, b). \end{aligned}$$

Av medelvärdessatsen följer då att

$$F(h,k) = g'(b+tk)k = (\partial_2 f(a+h, b+tk) - \partial_2 f(a, b+tk))k$$

för något $0 < t < 1$. Medelvärdessatsen ger vidare

$$F(h,k) = \partial_1 \partial_2 f(a+sh, b+tk)kh$$

för något $0 < s < 1$. För $hk \neq 0$ har vi alltså

$$\frac{f(a+h, b+k) - f(a, b+k) - f(a+h, b) + f(a, b)}{hk} = \partial_1 \partial_2 f(a+sh, b+tk).$$

Låt nu för fixt $k \neq 0$ $h \to 0$. Vi får då

$$\frac{\partial_1 f(a, b+k) - \partial_1 f(a, b)}{k} = \lim_{h \to 0} \frac{F(h,k)}{hk}.$$

Låt nu $k \to 0$ och vi får

$$\partial_2 \partial_1 f(a, b) = \lim_{k \to 0} \lim_{h \to 0} \frac{F(h,k)}{hk}$$

på samma sätt fås

$$\partial_1 \partial_2 f(a, b) = \lim_{h \to 0} \lim_{k \to 0} \frac{F(h,k)}{hk}.$$

Men

$$\frac{F(h,k)}{hk} = \partial_1\partial_2 f(a+sh, b+tk) \to \partial_1\partial_2 f(a,b) \text{ om } (h,k) \to (0,0)$$

eftersom $\partial_1\partial_2 f$ är kontinuerlig i (a,b). Således har vi visat att de bägge upprepade gränsvärdena sammanfaller, dvs. att

$$\partial_1\partial_2 f(a,b) = \partial_2\partial_1 f(a,b)$$

som visar lemmat. $\square$

Sats 22.2. *Om $f : V \to W$ är två gånger kontinuerligt deriverbar i en omgivning av $a \in V$ så är $D^2 f(a)$ en symmetrisk bilinjär avbildning från $V \times V$ till W.*

Bevis. Det räcker att visa att

$$(f''(a)(h))(k) = (f''(a)(k))(h) \quad \forall h,k \in V.$$

Fixera $h,k \in V$ och en linjär avbildning $\alpha : W \to \mathbb{R}$. Definiera $g : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ genom $g(x,y) = \alpha \circ f(a + xh + yk)$. Då är g två gånger kontinuerligt deriverbar i närheten av $(0,0)$ och

$$\begin{aligned}
& \partial_1 g(x,y) = \alpha(f'(a+xh+yk)h) \\
\Rightarrow \quad & \partial_1 g(0,y) = \alpha(f'(a+yk)h) = \\
& = \alpha(f'(a)h + f''(a)(yk)(h) + o(yk)h) \\
\Rightarrow \quad & \partial_2\partial_1 g(0,0) = \alpha(f''(a)(k)(h))
\end{aligned}$$

På samma sätt fås

$$\partial_1\partial_1 g(0,0) = \alpha(f''(a)(k)(h))$$

Men $\partial_1\partial_2 g(0,0) = \partial_2\partial_1 g(0,0)$ ger då att

$$\begin{aligned}
& \alpha(f''(a)(h)(k) - f''(a)(k)(h)) = 0 \quad \forall\alpha \\
\Rightarrow \quad & f''(a)(h)(k) = f''(a)(k)(h) \quad \forall h,k \in V
\end{aligned}$$

$\square$

23. Variabelsubstitution i multipelintegraler

Låt D och E vara öppna delmängder i $\mathbb{R}^n$ sådana att $\bar{D}$ och $\bar{E}$ är kompakta.

Antag att $\varphi : \bar{D} \to \bar{E}$ är en kontinuerlig deriverbar bijektion sådan att $\det\varphi' > 0$ i $\bar{D}$. Antag vidare att $f : \bar{E} \to \mathbb{R}$ är sådan att Riemannintegralerna $\int_E f$ och $\int_D f \circ \varphi \det\varphi'$ existerar. (Det skulle duga om f vore kontinuerlig.)

Då gäller följande

Sats 23.1.

$$\int_{\varphi(D)} f = \int_D f \circ \varphi \det\varphi'$$

Idén med vårt bevis är att först visa satsen för avbildningar φ som bara ändrar en av variablerna: Vi inför

Definition 23.2. En avbildning $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ kallas enkel om

$$\varphi(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_n)$$

där $y_1 = \psi(x_1, x_2, \ldots, x_n)$ och $y_k = x_k$ om $1 < k \leq n$. (Eller om detta uppnås efter en omnumrering av variablerna.)

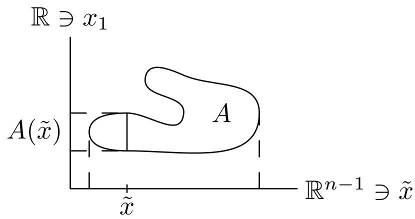
Vi noterar att

$$\det \varphi' = \partial_1 \psi.$$

Lemma 23.3. *Om φ är enkel så gäller sats 23.1.*

Bevis. Låt oss först för en delmängd $A \subset \mathbb{R}^n$ definiera

$$\begin{aligned} \tilde{A} &= \{\tilde{x} \in \mathbb{R}^{n-1} : (x_1, \tilde{x}) \in A \text{ för något } x_1 \in \mathbb{R}\} \\ A(\tilde{x}) &= \{x_1 \in \mathbb{R} : (x_1, \tilde{x}) \in A\} \end{aligned}$$



$$\begin{aligned} \int_E f &= \int_{y \in E} f(y_1, \tilde{y})\, dy_1 d\tilde{y} = \\ &= \int_{y \in E} f(y_1, y_2, \ldots, y_n)\, dy_1 dy_2 \ldots dy_n = \\ &= \{\text{Fubini}\} = \int_{\tilde{y} \in \tilde{E}} \Big(\int_{y_1 \in E(\tilde{y})} f(y_1, \tilde{y})\, dy_1 \Big)\, d\tilde{y} = \\ &= \int_{\tilde{x} \in \tilde{E}} \Big(\int_{y_1 \in E(\tilde{x})} f(y_1, \tilde{x})\, dy_1 \Big)\, d\tilde{x} \end{aligned}$$

I den inre integralen, för fixt $\tilde{x}$, gör vi substitutionen $y_1 = \psi(x_1, \tilde{x})$ varvid $dy_1 = \partial_1 \psi(x_1, \tilde{x}) dx_1$ och $y_1 \in E(\tilde{x}) \Leftrightarrow (y_1, \tilde{x}) \in E \Leftrightarrow (\psi(x_1, \tilde{x}), \tilde{x}) \in E \Leftrightarrow \varphi(x_1, \tilde{x}) \in E \Leftrightarrow (x_1, \tilde{x}) \in D \Leftrightarrow x_1 \in D(\tilde{x})$. Alltså får vi

$$\begin{aligned} \int_E f &= \int_{\tilde{x} \in \tilde{D}} \Big(\int_{x_1 \in D(\tilde{x})} f(\psi(x_1, \tilde{x}), \tilde{x}) \partial_1 \psi(x_1, \tilde{x}) dx_1 \Big) d\tilde{x} \\ &= \int_D (f \circ \varphi) \partial_1 \psi = \int_D f \circ \varphi \det \varphi'. \end{aligned}$$

$\square$

Vi visar nu att om sats 23.1 gäller för φ_1 och φ_2 så gäller den för $\varphi_1 \circ \varphi_2$ ty låt $D \xrightarrow{\varphi_2} D_1 \xrightarrow{\varphi_1} E$ vara bijektioner som i satsen.

Antag att $\int_{\varphi_1(D_1)} f_1 = \int_{D_1} f_1 \circ \varphi_1 \det \varphi_1'$ och att $\int_{\varphi_2(D)} f_2 = \int_D f_2 \circ \varphi_2 \det \varphi_2'$ för godtyckliga $f_1 : E \to \mathbb{R}$ och $f_2 : D_1 \to \mathbb{R}$. Då har vi om $\varphi = \varphi_1 \circ \varphi_2$ och $f : E \to \mathbb{R}$ att

$$\int_E f = \int_{\varphi_1(D_1)} f = \int_{D_1} f \circ \varphi_1 \det \varphi_1' = \int_{\varphi_2(D)} f \circ \varphi_1 \det \varphi_1' =$$
$$\int_D f \circ \varphi_1 \circ \varphi_2((\det \varphi_1') \circ \varphi_2) \det \varphi_2'.$$

Men av kedjeregeln har vi

$$\varphi'(x) = (\varphi_1 \circ \varphi_2)'(x) = (\varphi_1' \circ \varphi_2(x)) \circ \varphi_2'(x)$$

dvs.

$$\det \varphi'(x) = \det(\varphi_1' \circ \varphi_2(x)) \det \circ \varphi_2'(x).$$

Således har vi visat att

$$\int_{\varphi(D)} f = \int_D f \circ \varphi \det \varphi'.$$

Det som återstår att visa är att varje avbildning $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ kan skrivas som en sammansättning av högst n stycken enkla avbildningar. Detta visas med hjälp av implicita funktionssatsen. Det blir naturligtvis ett lokalt resultat och eventuellt måste vi omnumrera koordinaterna vilket naturligtvis inte spelar något roll eftersom satsen är trivial för permutationer.

Vi behöver

Lemma 23.4. *Låt $\varphi = (\varphi_1, \varphi_2, \ldots, \varphi_n)$ vara en kontinuerligt deriverbar funktion från en öppen omgivning av $a \in \mathbb{R}^n$ till $\mathbb{R}^n$ sådan att $\partial_1 \varphi_1(a) \neq 0$. Då kan φ faktoriseras enligt $\varphi = \alpha \circ \beta$ där vi för x nära a har*

$$(x_1, x_2, \ldots, x_n) \overset{\beta}{\mapsto} (y_1, x_2, \ldots, x_n) \overset{\alpha}{\mapsto} (y_1, y_2, \ldots, y_n)$$

där $y_i = \varphi_i(x)$.

Bevis. Det räcker att visa att $(y_2, \ldots, y_n)$ kan lösas ut som en kontinuerligt deriverbar funktion av $(y_1, x_2, \ldots, x_n)$ i närheten av vår punkt. Betrakta därför avbildningen

$$((y_1, x_2, \ldots, x_n), (x_1, y_2, \ldots, y_n)) \mapsto (\varphi_1(x) - y_1, \ldots, \varphi_n(x) - y_n)$$

från $\mathbb{R}^n \times \mathbb{R}^n$ med värden i $\mathbb{R}^n$. Då har vi att

$$\frac{\partial(\varphi_1(x) - y_1, \ldots, \varphi_n(x) - y_n)}{\partial(x_1, y_2, \ldots, y_n)} = \begin{bmatrix} \partial_1\varphi_1(x) & 0 & 0 & \ldots & 0 \\ \partial_1\varphi_2(x) & -1 & 0 & \ldots & 0 \\ \partial_1\varphi_3(x) & 0 & -1 & \ldots & 0 \\ \vdots & & & & \\ \partial_1\varphi_n(x) & -1 & 0 & \ldots & -1 \end{bmatrix}$$

är invertibel i vår aktuella punkt eftersom dess determinant, som är $(-1)^{n-1}\partial_1\varphi_1(a)$, är nollskilld. Implicita funktionssatsen fullbordar beviset. $\square$

Genom ett kompakthetsresonemang, som lämnas som övning, kan man visa att vårat D kan delas upp i ett ändligt antal mindre områden $D_1, D_2, \dots, D_N$ sådana att $\cup D_j = D$ och $D_i \cap D_j$ saknar inre punkter om $i \neq j$ och sådana att φ på varje D_j kan skrivas som en sammansättning av enkla avbildningar således har vi

$$\int_{\varphi(D_j)} f = \int_{D_j} f \circ \varphi \det \varphi' \quad \forall j$$

som ger

$$\int_{\varphi(D)} f = \sum \int_{\varphi(D_j)} f = \sum \int_{D_j} f \circ \varphi \det \varphi' = \int_D f \circ \varphi \det \varphi' \quad \forall j.$$

Detta visar satsen.

24. Volymer av parallellepipeder

Om A är en $n \times m$-matris med kolonner $a_1, a_2, \dots, a_m$ i $\mathbb{R}^n$ och $m \leq n$, så spänner dessa upp en parallellepiped som vi betecknar

$$\operatorname{Ep}(a_1, \dots, a_m) = \{\sum t_i a_i : 0 \leq t_i \leq 1\}.$$

Vi vill nu definiera dess m-dimensionella volym.

Till detta väljer vi enhetsvektorer $b_{m+1}, \dots, b_n$ som är parvis ortogonala och som dessutom är ortogonala mot alla a_i:na. Bilda också $n \times (n-m)$-matrisen B med kolonnerna $b_{m+1}, \dots, b_n$. Vi har då att, om $C = [A, B]$ är den $n \times n$-matris vi får en kolonnerna a_1, a_2, $\dots$, a_m, b_{m+1}, $\dots$, b_n,

$$(\det C)^2 = \det C^T \det C = \det(C^T C).$$

Men

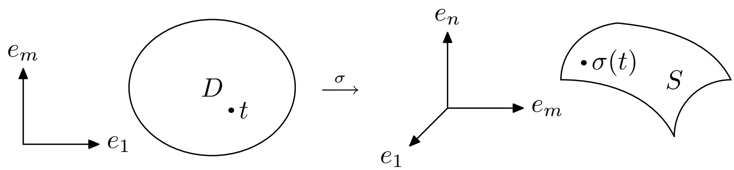
$$C^T C = \begin{bmatrix} A^T \\ B^T \end{bmatrix} \begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} A^T A & 0 \\ 0 & I \end{bmatrix}$$

så $(\det C)^2 = det(A^T A)$ är således oberoende av valet av b:na. Vi definierar den m-dimensionella volymen av $\operatorname{Ep}(a_1, \dots, a_m)$ genom

$$\operatorname{Vol}_m \operatorname{Ep}(a_1, \dots, a_m) = (\det A^T A)^{1/2}.$$

25. Ytor och deras areor

Låt $\mathbb{R}^m \supset D \xrightarrow{\sigma} S \subset \mathbb{R}^n$ vara en kontinuerligt deriverbar och bijektiv avbildning sådan att σ' är injektiv i hela D, där $m \leq n$:



Vi definierar den m-dimensionella volymen av S eller arean eller måttet kort och gott genom

$$\int_S dS = \int_D (\det \sigma'^T \sigma')^{1/2}. \tag{4}$$

Övning 13. Visa att integralen i (4) bara beror av S, dvs. antag att om $\tilde\sigma = \sigma \circ \alpha$ där $\alpha : \tilde D \to D$ är kont deriverbar så är

$$\int_S dS = \int_{\tilde D} (\det \tilde\sigma'^T \tilde\sigma')^{1/2}.$$

Anmärkning. Vi skriver lite formellt att det oriktade areamåttet ges av

$$dS = (\det \sigma'(t)^T \sigma'(t))^{1/2} dt_1 \dots dt_m$$

i parametriseringen

$$D \ni (t_1, \dots, t_m) \mapsto \sigma(t_1, \dots, t_m) \in S.$$

26. Flödesintegralen

Låt $\mathbb{R}^{n-1} \supset D \xrightarrow{\sigma} S \subset \mathbb{R}^n$ vara en parametrisering av hyperytan S och antag $\hat N : S \to S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$ är normal till S. Detta betyder att $\hat N(\sigma(u))$ är ortogonal mot $\operatorname{Im} \sigma'(u)$ för alla u i D.

Definition 26.1 (Orientering)**.** Vi säger att σ är en *positiv parametrisering* till ytan S med normal $\hat N$ om $\det[\sigma', \hat N \circ \sigma] > 0$ i hela D.

Låt oss föreställa oss en inkompressibel vätska som strömmar i $\mathbb{R}^3$ och som vi tiden t och punkten x har hastigheten $F(x,t)$ och att densiteten är konstant lika med ett överallt. Hur mycket vätska strömmar då genom en tänkt yta S per tidsenhet om vi räknar vätskan som strömmar med positiv inre produkt med normalen som positivt flöde.

Vi resonerar nu först en smula heuristiskt. Betrakta en infinitesimal parallellepiped

$$u + \operatorname{Ep}[e_1 du_1, \dots, e_{n-1} du_{n-1}]$$

i D. Denna avbildas, via σ, på epipeden

$$\sigma(u) + \operatorname{Ep}[\sigma'(u) e_1 du_1, \dots, \sigma'(u) e_{n-1} du_{n-1}].$$

Denna senare, som är en liten epiped på ytan S, flyttas av flödet under den infinitesimala tiden dt, vektorn $F(\sigma(u), t)dt$. Men då blir flödet genom denna epiped under tiden dt lika med determinanten

$$\det[\sigma'(u)e_1 du_1, \dots, \sigma'(u)e_{n-1}du_{n-1}, F(\sigma(u), t)dt].$$

Således verkar det inte helt orimligt att definiera flödet ut genom S med normal $\hat{N}$ per tidsenhet genom

$$\int_D \det[\sigma', F(\sigma(u), t)]du_1 \dots du_{n-1}$$

och flödet som passerar ytan S mellan tidpunkterna a och b blir

$$\int_a^b \left(\int_D \det[\sigma', F(\sigma(u), t)]du_1 \dots du_n \right) dt.$$

Vi ska hör inte syssla med flöden som beror av tiden, utan bara med stationära flöden, då F är oberoende av t. Vi definierar då flödet av vektorfältet $F : \mathbb{R}^n \to \mathbb{R}^n$ ut genom hyperytan S med normal $\hat{N}$ genom

$$\int_D \det[\sigma', F \circ \sigma]$$

om σ är positiv parametrising.

Vi ska nu visa att denna integral är oberoende av val av positiv parametrisering. Låt nämnligen $\mathbb{R}^{n-1} \supset \tilde{D} \xrightarrow{\alpha} D \xrightarrow{\sigma} S$ där vi antar $\det \alpha' > 0$ och sätt $\tilde{\sigma} = \sigma \circ \alpha$. Vi har då

$$\begin{aligned}
\int_{\alpha(\tilde{D})} \det[\sigma', F \circ \sigma] &= \int_{\tilde{D}} \det[\sigma' \circ \alpha, F \circ \sigma \circ \alpha] \det \alpha' = \\
&= \int_{\tilde{D}} \det[\sigma' \circ \alpha, F \circ \tilde{\sigma}] \det \begin{bmatrix} \alpha' & 0 \\ 0 & 1 \end{bmatrix} = \\
&= \int_{\tilde{D}} \det[(\sigma' \circ \alpha)\alpha', F \circ \tilde{\sigma}] = \\
&= \int_{\tilde{D}} \det[\tilde{\sigma}', F \circ \tilde{\sigma}]
\end{aligned}$$

där $\left[\begin{smallmatrix} \alpha' & 0 \\ 0 & 1 \end{smallmatrix}\right]$ betecknar en $n \times n$-matrisvärd funktion på D.

27. Gauss sats

Låt $F : \mathbb{R}^n \to \mathbb{R}^n$ vara ett vektorfält och Ω en kropp i $\mathbb{R}^n$ med en randyta S med utåtgående normal $\hat{N}$. Definiera $\varphi_t(x) = x + tF(x)$. Enligt variabelsubstitutionsformeln har vi då att

$$\int_{\varphi_t(\Omega)} 1 = \int_\Omega \det(\mathrm{id} + tF') = \operatorname{Vol} \varphi_t(\Omega).$$

Detta gäller åtminstonde för t litet eftersom $\det(\mathrm{id} + tF') \to 1$ om $t \to 0$. Låt $\lambda_1, \ldots, \lambda_n$ vara egenvärdena till F'. Då har $\mathrm{id} + tF'$ egenvärdena $1 + t\lambda_1, \ldots, 1 + t\lambda_n$, varför

$$\begin{aligned}\det(\mathrm{id} + tF') &= (1 + t\lambda_1) \ldots (1 + t\lambda_n) \\ &= 1 + t(\lambda_1 + \cdots + \lambda_n) + O(t^2) = 1 + t \operatorname{sp} F' + O(t^2).\end{aligned}$$

Alltså har vi för $t = 0$ att $\frac{d}{dt} \operatorname{Vol} \varphi_t(\Omega) = \int_\Omega \operatorname{sp} F'$.

Definition 27.1. Divergensen av vektorfältet $\operatorname{div} F$ är spåret av F'.

$$\operatorname{div} F = \operatorname{sp} F'$$

Övning 14. Visa att $\frac{d}{dt}|_{t=0} \operatorname{Vol} \varphi_t(D) = \int_D \det[\sigma', F \circ \sigma]$ dvs. flödet av vektorfältet F ut ur S.

Vi har således

Sats 27.2 (Gauß). *Flödet av ett vektorfält $\bar{F}$ ut ur en yta S som omsluter en kropp Ω i $\mathbb{R}^n$ är lika med integralen av divergensen av F över Ω.*

$$\int_D \det[\sigma', F \circ \sigma] = \int_\Omega \operatorname{div} F$$

om $\mathbb{R}^{n-1} \supset D \xrightarrow{\sigma} S$ och om $\det[\sigma', \hat{N} \circ \sigma] > 0$, där $\hat{N}$ är den utåtgående normalen till ytan S.

Övning 15. Visa genom att tillämpa Gauß sats för $n = 2$ på vektorfältet

$$\bar{F} = \begin{bmatrix} Q \\ -P \end{bmatrix}$$

den s.k. Greens formel

$$\int_\Gamma P dx + Q dy = \iint_D (Q_x + P_y) dx dy$$

där Γ är en kurva i planet som ör positivt orienterad och omsluter D.

27.3. **Specialfallet $\mathbb{R}^3$.** (Fysiknotation med vektorstreck) Låt S vara en tvådimensionell yta i $\mathbb{R}^3$ som parametriseras enligt

$$\mathbb{R}^2 \supset D \ni (u, v) \xrightarrow{\bar{r}} \bar{r}(u, v) \in \mathbb{R}^3$$

och låt $\bar{F} : \mathbb{R}^3 \to \mathbb{R}^3$ vara vårt vektorfält givet at $\bar{F} = (P, Q, R)$. Då har vi $\operatorname{sp} \bar{F}' = P_x + Q_y + R_z$, där koordinaterna i $\mathbb{R}^3$ kallas (x, y, z) som vanligt. Vi har också då att

$$\begin{aligned}\det[\sigma', F \circ \sigma] &= \det[\frac{\partial \bar{r}}{\partial u}, \frac{\partial \bar{r}}{\partial v}, \bar{F} \circ \bar{r}] = \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & P \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & Q \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & R \end{bmatrix} \\ &= \left(\frac{\partial \bar{r}}{\partial u} \times \frac{\partial \bar{r}}{\partial v} \right) \cdot \bar{F}(\bar{r}(u, v)).\end{aligned}$$

Vi inför då också $d\bar{s} = \frac{\partial \bar{r}}{\partial u} \times \frac{\partial \bar{r}}{\partial u} dudv$. Gauß sats blir då

$$\int_D \det[\sigma', F \circ \sigma] = \oiint_S \bar{F} \cdot d\bar{s} = \iiint_\Omega \operatorname{div} \bar{F}\, dV = \int_\Omega \operatorname{div} F$$

där

$$\oiint_S \bar{F} \cdot d\bar{s} = \iint_D \bar{F}(\bar{r}(u,v)) \cdot \left(\frac{\partial \bar{r}(u,v)}{\partial u} \times \frac{\partial \bar{r}(u,v)}{\partial v} \right) dudv.$$

28. Stokes sats

Låt $\bar{F} = (P, Q, R)$ vara ett vektorfält $\mathbb{R}^3 \to \mathbb{R}^3$. Då är $\bar{F}' - \bar{F}'^T$ en antisymmetrisk avbildning med matrisrepressentationen

$$\begin{bmatrix} P_x & P_y & P_z \\ Q_x & Q_y & Q_z \\ R_x & R_y & R_z \end{bmatrix} - \begin{bmatrix} P_x & Q_x & R_x \\ P_y & Q_y & R_y \\ P_z & Q_z & R_z \end{bmatrix} = \begin{bmatrix} 0 & P_y - Q_x & P_z - Rx \\ Q_x - P_y & 0 & Q_z - R_y \\ R_x - P_z & R_y - Q_z & 0 \end{bmatrix}.$$

Vi vet att kryssprodukter i $\mathbb{R}^3$ kan representeras med antisymmetriska 3×3-matriser och viceversa.

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} bz - cy \\ cx - az \\ zy - bx \end{bmatrix} = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Således har vi att

$$(\bar{F}' - \bar{F}'^T)\bar{v} = \begin{bmatrix} R_y - Q_z \\ P_z - R_x \\ Q_x - P_y \end{bmatrix} \times \bar{v}.$$

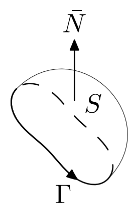
Vi definierar rotationen av vektorfälltet $\bar{F}$ som

$$\operatorname{rot} \bar{F} = \operatorname{rot} \begin{bmatrix} P \\ Q \\ R \end{bmatrix} = \begin{bmatrix} R_y - Q_z \\ P_z - R_x \\ Q_x - P_y \end{bmatrix}.$$

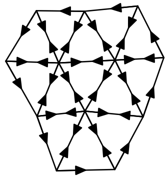
Vi får att $\operatorname{rot} \bar{F}$ inte beror av vårt speciella val av ON-system, eftersom $\bar{F}' - \bar{F}'^T$ bara beror av vårt val av inre produkt, som behövs för definition av transponatet.

Sats 28.1 (Stokes)**.** *Låt $\bar{F}$ vara ett kontinuerligt deriverbart vektorfält på $\mathbb{R}^3$ och låt S vara en orienterade slät yta med randkurva Γ och normaltfält $\bar{N}$ riktade på vanligt sätt (figur). Då gäller*

$$\iint_S \operatorname{rot} \bar{F} \cdot d\bar{s} = \oint_\Gamma \bar{F} \cdot d\bar{r}.$$



Bevis. Enligt definitionen på integralen räcker det att visa satsen då S består av ett ändligt antal plana trianglar.



Kalla trianglarna S_1, S_2, …, S_m och deras respektive randkurvor Γ_1, Γ_2, …, Γ_m. Eftersom

$$\iint_S \operatorname{rot} \bar{F} \cdot d\bar{s} = \sum_i \iint_{S_i} \operatorname{rot} \bar{F} \cdot d\bar{s}$$

och

$$\oint_\Gamma \bar{F} \cdot d\bar{r} = \sum_i \int_{\Gamma_i} \bar{F} \cdot \bar{r}.$$

Så det räcker att visa satsen för en triangel som vi kallar S. Då vidare $\operatorname{rot} \bar{F}$ är oberoende av koordinatsystem, så länge det är ON, kan vi utan inskränkning anta att S ligger i xy-planet.

Låt nu $D = \{\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right] \in \mathbb{R}^2 : \left[\begin{smallmatrix} x \\ y \\ 0 \end{smallmatrix}\right] \in S\}$ och kalla den positivt orienterade randen till D för $\mathcal{C}$. Då parametriseras S enligt

$$D \ni \begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \in S$$

och vi har

$$d\bar{s} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} dxdy$$

varför

$$\operatorname{rot} \bar{F} \cdot d\bar{s} = (Q_x(x, y, 0) - P_y(x, y, 0))dxdy.$$

Greens formel i planet ger då

$$\begin{aligned} \iint_S \operatorname{rot} \bar{F} \cdot d\bar{s} &= \iint_D (Q_x(x, y, 0) - P_y(x, y, 0))dxdy = \\ &= \int_{\mathcal{C}} P(x, y, 0)dx + Q(x, y, 0)dy = \\ &= \int_\Gamma Pdx + Qdy + Rdz = \oint_\Gamma \bar{F} \cdot d\bar{r} \end{aligned}$$

som visar satsen. □

E-mail address: eno@math.kth.se