De matematiska grunderna för den allmänna relativitetsteorin

Kapitel 1

Relativitetsteori

1.1 Inledning

1.2 Den speciella relativitetsteorin

1.3 Den allmänna relativitetsteorin

Kapitel 2

Introduktion till topologiska rum

2.1 Inledning

2.2 Om struktur

2.3 Definitionen av ett topologiskt rum

2.4 Grundläggande egenskaper

2.5 Parakompakthet

Kapitel 3

Mångfalder

3.1 Inledning

Mångfalder är en speciell typ av topologiska rum som lokalt påminner om delar av $\mathbf{R}^n$. De kan närmast ses som en generalisering av 2-dimensionella ytor i $\mathbf{R}^3$ (kom ihåg att en yta lokalt kan beskrivas med 2 parametrar så lokalt ser den ut som en (lite tillknycklad) del av $\mathbf{R}^2$). Denna ytterligare struktur gör mångfalder väl lämpade för att beskriva rumtider i allmän relativitetsteori.

3.2 Ett tankeexperiment

Idag vet vi nästan allt som finns att veta om jordytans form. Med hjälp av satellitmätningar kompletterade med markmätningar kan jordytans form beskrivas med stor noggrannhet ända ner på detaljnivå. Om vi går tillbaka i tiden har det sedan Maupertuis expedition till Lappland 1736 varit känt att jordytans form mycket väl approximeras av en oblat rotationsellipsoid (d v s ellipsoidaxeln som är parallell med rotationsaxeln är kortare än axlarna som är vinkelräta mot rotationsaxeln, 'jorden är tillplattad vid polerna'). Innan denna expedition ansågs jordytan vara i stort sett sfärisk och de äldsta beräkningarna av jordens radie gjordes år 240 fvt av Eratosthenes och hans värde ligger inom 5–15% från det korrekta värdet (osäkerheten i felmarginalen beror på att man idag inte vet exakt hur lång den gängse längdenheten, *stadion*, var).

Eratosthenes och andras mätningar var förresten välkända i Europa under medeltiden, vilket effektivt tar död på den seglivade myten att medeltidens folk trodde att jorden var platt.

Det är förresten inte heller sant att Columbus hade svårt att hitta finansiärer för att han hävdade att jorden var rund. Han hade svårt att hitta finansiärer för att han hävdade att jordradien var mycket mindre än vad mätningar gav vid handen, och att seglatsen till Indien därmed var mycket kortare än den är i verkligheten.

Låt oss nu göra tankeexperimentet att allt det ovanstående är okänt för oss och att vi vill undersöka jordytans form. Vi ställs då omedelbart inför följande problem: Om vi vill beskriva den lilla del av jordytan som för tillfället befinner sig inom vårt synfält måste beskrivningen innehålla mycket smådetaljer i form av berg, dalar m m. En sådan beskrivning blir helt ohanterlig om vi vill beskriva stora delar av, eller t o m hela, jordytan.

Ett sätt att lösa detta problem är att hitta ett matematiskt ramverk som är tillräckligt varierat för att kunna ge både detaljerade beskrivningar av småbitar av jordytan och mer 'utsmetade' beskrivningar av jordytan som helhet. 2-dimensionella ytor i $\mathbf{R}^3$ är precis ett sådant ramverk.

Minns att en 2-dimensionell yta Y i $\mathbf{R}^3$ är en delmängd till $\mathbf{R}^3$ som lokalt kan parametriseras med hjälp av 2 parametrar, d v s lokalt kan punkter (x, y, z) på Y skrivas

$$\left\{\begin{array}{lcl} x & = & x(s,t) \\ y & = & y(s,t) \\ z & = & z(s,t) \end{array}\right. , \ (s,t) \in D \subset \mathbf{R}^2,$$

Att parametriseringen är lokal innebär att det kanske inte finns någon enskild parametrisering som täcker hela Y, men att Y i alla fall kan skrivas som en union av överlappande bitar som alla kan parametriseras på detta sätt.

När vi samlat in tillräckligt mycket data om den del av jordytan vi är intresserade av och hittat en lämplig 2-dimensionell yta i $\mathbf{R}^3$ som modellerar denna del kan vi använda vår kunskap om 2-dimensionella ytor i $\mathbf{R}^3$ till att besvara t ex följande frågor:

- Vad är kortaste vägen mellan två givna punkter?
- Vilka tangentvektorer har ytan i en given punkt?
- Hur stor är ytans krökning i en given punkt?
- Finns det någon yta som beskriver *hela* jordytan på en gång?
- Är denna yta i så fall begränsad? Sluten? Ändras dess form med tiden?

Vårt mål med detta tankeexperiment är nu inte att komma på några revolutionerande resultat om jordytans form utan att hitta en beskrivning av universum, och detta problem liknar det vi ställdes inför i tankeexperimentet ovan. Vi vill dels kunna beskriva universum som helhet och dels mindre detaljer i universum (solsystemet, svarta hål m m). Kan detta göras på något liknande sätt?

Vi noterar att hela den av oss observerbara delen av universum är 3-dimensionell. Dessutom vet vi från den speciella relativitetsteorin att tid och rum är intimt sammanflätade, så det verkar naturligt att även baka in tiden i vår beskrivning av universum. Ett möjligt ramverk för våra modeller av (delar av) universum skulle alltså kunna vara 4-dimensionella 'ytor' i $\mathbf{R}^5$, eller kanske rentav i $\mathbf{R}^N$ för något $N \geq 5$.

Analogt med ovanstående kan en 4-dimensionell yta i $\mathbf{R}^N$ lokalt parametriseras med 4 parametrar, d v s en punkt $(x^0, x^1, \ldots, x^{N-1})$ (det ser kanske underligt ut att indexera koordinaterna med ett index uppe istället för nere men ignorera detta för tillfället, anledningen kommer att avslöjas senare) på ytan ges av

$$\left\{\begin{array}{rcl} x^0 & = & x^0(s,t,u,v) \\ x^1 & = & x^1(s,t,u,v) \\ \vdots & \vdots & \vdots \\ x^{N-1} & = & x^{N-1}(s,t,u,v) \end{array}\right. , \ (s,t,u,v) \in D \subset \mathbf{R}^4.$$

Vi ser att detta lite naiva resonemang leder till ett problem. Ingen har någonsin observerat att vårt universum skulle vara en delmängd av något omkringliggande $\mathbf{R}^N$ och då verkar det onaturligt att förutsätta existensen av ett sådant $\mathbf{R}^N$. Ovanstående ramverk duger alltså inte för våra syften.

Å andra sidan vill vi inte överge idén med 4-dimensionella ytor alltför lättvindigt. Trots allt verkar detta ramverk väl lämpat för besvara frågor, liknande de ovanstående om jordytan, om universum.

Det är alltså önskvärt att de matematiska objekt vi använder för att studera (delar av) universum har följande egenskaper:

- 4-dimensionella ytor i $\mathbf{R}^N$ där $N > 4$ är specialfall av sådana objekt.
- definitionen av ett sådant objekt refererar bara till objektet självt, och förutsätter *inte* existens av något omkringliggande $\mathbf{R}^N$.

Ett exempel på ett matematiskt objekt som uppfyller båda dessa krav är en s k *mångfald*.

3.3 Enhetssfären, del 1

Enhetssfären S är som bekant den yta i $\mathbf{R}^3$ som ges av ekvationen $x^2+y^2+z^2 = 1$. Som förberedelse inför definitionen av en mångfald ska vi först se hur vi kan generalisera definitionen av enhetssfären så att den inte refererar till något omkringliggande $\mathbf{R}^3$, utan bara till sfären själv.

Låt (θ, φ) vara vanliga sfäriska koordinater, d v s

$$\left\{ \begin{array}{lcl} x & = & \sin\theta\cos\varphi \\ y & = & \sin\theta\sin\varphi \\ z & = & \cos\theta \end{array} \right. , \ (\theta, \varphi) \in D.$$

Hur ska vi välja D? Låt oss vara lite förutseende. Vi kommer sannolikt att behöva derivera diverse funktioner med avseende på θ och φ. Detta innebär att vi måste kunna bilda differenskvoter i både θ och φ, vilket inte säkert är möjligt i randpunkter till D. Av detta skäl låter vi D vara en öppen mängd, t ex $D = \{(\theta, \varphi)\,;\ 0 < \theta < \pi\,,\ 0 < \varphi < 2\pi\}$.

Valet av D som en öppen mängd för dock med sig en kostnad eftersom de sfäriska koordinaterna då inte täcker hela S. Låt U vara den delmängd av S som beskrivs av de sfäriska koordinaterna med detta val av D. Då är U hela S förutom en halv storcirkel i den del av xz-planet där $x \geq 0$ med ändpunkter $(0, 0, 1)$ och $(0, 0, -1)$(jfr en meridian på jordytan).

Vi kompletterar därför beskrivningen av S med ytterligare ett sfäriskt koordinatsystem $(\bar\theta, \bar\varphi)$ där $\bar\theta$ utgår från positiva y-axeln istället för positiva z-axeln. Detta ger

$$\left\{ \begin{array}{lcl} x & = & \sin\bar\theta\sin\bar\varphi \\ y & = & \cos\bar\theta \\ z & = & \sin\bar\theta\cos\bar\varphi \end{array} \right. , \ (\bar\theta, \bar\varphi) \in \bar D.$$

där $\bar D = \{(\bar\theta, \bar\varphi)\,;\ 0 < \bar\theta < \pi\,,\ -\frac{\pi}{2} < \bar\varphi < \frac{3\pi}{2}\}$. Som ovan, låt $\bar U$ vara den del av S som täcks av ovanstående sfäriska koordinater med detta val av $\bar D$, d v s $\bar U$ är hela S förutom en halv storcirkel i den del av xy-planet där $x \leq 0$ med ändpunkter $(0,1,0)$ och $(0,0,-1)$. Då är $U \cup \bar{U} = S$ d v s de sfäriska koordinaterna $(\theta,\varphi) \in D$ och $(\bar{\theta},\bar{\varphi}) \in \bar{D}$ täcker tillsammans hela S. Lite informellt kan vi därför säga att företeelser på S kan översättas entydigt till företeelser på D och $\bar{D}$.

Om $p \in U \cap \bar{U}$ så att punkten p har både (θ,φ)- och $(\bar{\theta},\bar{\varphi})$-koordinater kan vi lösa ut de streckade koordinaterna som funktion av de ostreckade. Efter lite arbete får vi (övning)

$$\begin{aligned} \bar{\theta} &= \arccos(\sin\theta\sin\varphi) \\ \bar{\varphi} &= \begin{cases} \arctan(\tan\theta\cos\varphi) & \text{om} \quad 0 < \theta < \frac{\pi}{2} \\ \frac{\pi}{2} & \text{om } \theta = \frac{\pi}{2},\ \varphi \in \left]0,\frac{\pi}{2}\right[\cup \left]\frac{3\pi}{2},2\pi\right[\\ \arctan(\tan\theta\cos\varphi)+\pi & \text{om} \quad \frac{\pi}{2} < \theta < \pi \end{cases} \end{aligned} \tag{3.1}$$

Det viktiga här är inte exakt hur sambandet mellan streckade och ostreckade koordinater ser ut, utan att observera att de streckade koordinaterna blir $\mathcal{C}^\infty$ funktioner[1] av de ostreckade och omvänt (trivialt vad gäller $\bar{\theta}$, enkelt att verifiera vad gäller $\bar{\varphi}$).

Vi är nu klara att beskriva S helt utan referenser till något omkringliggande $\mathbf{R}^3$. Definiera en avbildning $\psi : U \to D$ genom att sätta $\psi(p) = (\theta,\varphi)$ där (θ,φ) är p:s sfäriska koordinater i det ostreckade systemet och analogt definierar vi $\bar{\psi} : \bar{U} \to \bar{D}$ genom att låta $\bar{\psi}(p) = (\bar{\theta},\bar{\varphi})$ där $(\bar{\theta},\bar{\varphi})$ är p:s sfäriska koordinater i det streckade systemet.

Funktionerna ψ och $\bar{\psi}$ kallas *koordinatsystem* eller *kartor* på S och mängden $\{\psi,\bar{\psi}\}$ som innehåller alla koordinatsystem vi behöver för att beskriva S kallas en *atlas*. Denna atlas ger nu, tillsammans med funktionen $\bar{\psi} \circ \psi^{-1} : D \to \bar{D}$ d v s sambandet mellan streckade och ostreckade koordinater i ekvation (3.1), en fullständig beskrivning av S och denna beskrivning refererar inte till något omkringliggande $\mathbf{R}^3$ utan bara till S själv.

Notera också att givet en godtycklig punkt $p \in S$ så finns en omgivning $U_0 \subset S$ till p sådan att $U_0 \subset U$ eller $U_0 \subset \bar{U}$, d v s så länge vi bara är intresserade av vad som händer lokalt, nära en viss punkt, behöver vi aldrig använda mer än ett koordinatsystem.

($\star$ En berättigad fråga är om det verkligen är nödvändigt att använda två koordinatsystem för att beskriva S. Går det att hitta ett listigt valt koordinatsystem som beskriver *hela* S?

Svaret på denna fråga är 'nej'. Det är en konsekvens av en sats som i folkmun kallas 'satsen om det håriga klotet', 'hairy ball-theorem' eller 'shaggy dog-theorem' (googla för mer detaljer), att man behöver minst två koordinatsystem för att beskriva hela S. Det koordinatsystem som kommer närmast att beskriva hela S är den stereografiska projektionen ('Riemannsfären', se kursen i komplex analys). Den ger ett koordinatsystem som täcker hela S utom en enda punkt, nämligen $(0,0,1)$. $\star$)

3.4 Definition av begreppet mångfald

Vi är visserligen bara intresserade av 4-dimensionella mångfalder (för att modellera universum eller delar därav) och 2-dimensionella (för att kunna använda vårt standardexempel, enhetssfären, för

[1] I denna skrift betecknar $\mathcal{C}^k$ klassen av funktioner som är k gånger kontinuerligt deriverbara och följaktligen betyder $\mathcal{C}^\infty$ klassen av funktioner som är kontinuerligt deriverbara hur många gånger som helst

att illustrera begrepp) men eftersom det inte innebär någon särskild extra ansträngning ger vi definitionen av en mångfald av godtycklig dimension:

Definition 3.1. Ett topologiskt rum M är en *n-dimensionell mångfald* av klass $\mathcal{C}^\infty$ om följande gäller:

1. Till varje $p \in M$ finns en omgivning U till p, en öppen mängd $D \subset \mathbf{R}^n$ och en bijektiv (inverterbar) funktion $\psi : U \to D$.

   ($\star$ Vi behöver också kräva att ψ är en homeomorfism, d v s att ψ och ψ^{-1} är kontinuerliga, sedda som avbildningar mellan de topologiska rummen M och $\mathbf{R}^n$ (se kap 2). $\star$)

2. Om $\psi : U \to D$ och $\bar{\psi} : \bar{U} \to \bar{D}$ är två av funktionerna från punkt 1 och $U \cap \bar{U} \neq \varnothing$ så är funktionen $\bar{\psi} \circ \psi^{-1} : D \to \bar{D}$ (vilken är en 'vanlig' vektorvärd funktion av n variabler) av klass $\mathcal{C}^\infty$.

3. Ovanstående krav på M visar sig inte vara tillräckligt restriktiva. Vi måste därför lägga till några ytterligare, ganska milda, regularitetskrav på M för att inte teorin ska urarta.

   ($\star$ Mer precist kräver vi att M är parakompakt och har hausdorffegenskapen (se kap 2). $\star$)

ψ kallas ett *koordinatsystem* eller en *karta* på M och mängden av alla koordinatsystem vi behöver för att beskriva M kallas en *atlas*. Observera att ψ är en vektorvärd funktion. Dess komponentfunktioner betecknas ofta $(x^a)_{a=0}^{n-1} = \left(x^0, x^1, \ldots, x^{n-1}\right)$ och kallas *koordinater* på U.

Observera också att koordinaterna indexeras med index uppe istället för nere. Detta beror på att vi senare kommer att behöva skriva upp formler innehållande många olika typer av indexerade kvantiteter. Att placera vissa index uppe och andra index nere är ett sätt att bringa lite ordning i det kaos som annars skulle uppstå. Läsaren uppmanas att redan nu vänja sig vid att koordinater indexeras med index uppe eftersom det underlättar i längden.

I konkreta exempel är det också vanligt att beteckna koordinater med varsin bokstav istället för att indexera dem, t ex genom att beskriva S med de sfäriska koordinaterna θ och φ. Dock behöver vi ofta använda formler där den indexerade formen av koordinaterna ingår och det är därför nödvändigt att ha en numrering av koordinaterna klar för sig, t ex $x^0 = \theta$, $x^1 = \varphi$ i fallet S.

Ibland behöver vi studera mångfalder av annan regularitet än $\mathcal{C}^\infty$, t ex mångfalder av klass $\mathcal{C}^k$ för något $k = 0, 1, 2, \ldots$ eller analytiska mångfalder. Vi modifierar då ovanstående definition genom att kräva att funktionerna på formen $\bar{\psi} \circ \psi^{-1}$ istället är av klass $\mathcal{C}^k$ respektive analytiska.

Det finns många exempel på mångfalder. $\mathbf{R}^n$ självt är förstås en n-dimensionell mångfald. Diskussionen i förra avsnittet visar också att enhetssfären S i $\mathbf{R}^3$ är en 2-dimensionell mångfald. T ex är ju funktionen $\bar{\psi} \circ \psi^{-1}$ precis den som ges av ekvation (3.1) och den är därför av klass $\mathcal{C}^\infty$. Funktionen $\psi \circ \bar{\psi}^{-1}$ har ett snarlikt utseende (vilket?) och är därför också av klass $\mathcal{C}^\infty$. S är alltså en mångfald som dessutom är en delmängd av en större mångfald, nämligen $\mathbf{R}^3$. Vi säger att S är en *delmångfald* till $\mathbf{R}^3$. Mer precist:

Definition 3.2. N sägs vara en *delmångfald* till M om M och N är mångfalder och $N \subset M$.

En k-dimensionell yta Y i $\mathbf{R}^n$ av klass $\mathcal{C}^\infty$ (d v s en yta där alla i parametriseringen ingående funktioner är av klass $\mathcal{C}^\infty$) är t ex en k-dimensionell delmångfald av $\mathbf{R}^n$ (inses på samma sätt som för S).

3.5 Regularitet hos funktioner

Vi vill också definiera regularitet för funktioner $f : M \to \mathbf{R}$ (eller $\mathbf{R}^m$ för något $m = 2, 3, \ldots$). Vi noterar att om $U \subset M$ är öppen och $\psi : U \to D \subset \mathbf{R}^n$ är är ett koordinatsystem på U så är $\tilde{f} = f \circ \psi^{-1} : D \to \mathbf{R}$ en vanlig flervariabelfunktion av koordinaterna $\left(x^0, x^1, \ldots, x^{n-1}\right)$. Observera också att om $\bar{\psi}$ är ett annat koordinatsystem på U så är $f \circ \psi^{-1} = \left(f \circ \bar{\psi}^{-1}\right) \circ \left(\bar{\psi} \circ \psi^{-1}\right)$, d v s $f \circ \psi^{-1}$ är $\mathcal{C}^\infty$ om och endast om $f \circ \bar{\psi}^{-1}$ är det. Det är därför naturligt att säga att f:s restriktion till U är av klass $\mathcal{C}^\infty$ (t ex) om $\tilde{f}$ är det. Eftersom hela M kan täckas av sådana U är följande definition naturlig:

Definition 3.3. $f : M \to \mathbf{R}$ sägs vara av klass $\mathbf{C}^\infty$ om $\tilde{f} = f \circ \psi^{-1}$ är av klass $\mathbf{C}^\infty$ för alla koordinatsystem ψ.

Definitionen av att f är av klass $\mathcal{C}^k$ eller att f är analytisk fås genom att istället kräva att $\tilde{f} \in \mathcal{C}^k$ resp. att $\tilde{f}$ är analytisk. Om istället $f : M \to \mathbf{R}^m$, $m = 2, 3, \ldots$ säger vi att $f \in \mathcal{C}^\infty$ om var och en av dess komponentfunktioner är $\mathcal{C}^\infty$.

Om $f : M \to \mathbf{R}$ är f alltså en regel som associerar ett reellt tal till varje punkt $p \in M$, medan $\tilde{f} = f \circ \psi^{-1}$ istället associerar samma tal till punkten p:s koordinater i koordinatsystemet ψ. Dessa funktioner innehåller alltså exakt samma information så länge vi bara betraktar koordinatsystemet ψ:s definitionsmängd U. Det är därför vanligt att identifiera $\tilde{f}$ med f och använda beteckningen f för båda funktionerna. En fördel med detta är att man får en enklare beteckningsapparat men en nackdel är att man själv måste hålla reda på att f ges av olika uttryck i olika koordinatsystem.

Exempel 3.4. Som illustration till ovanstående ska vi beräkna $\tilde{f}$ ur f i ett konkret fall. Låt $f : S \to \mathbf{R}$ vara den funktion som till varje punkt p på enhetssfären S i $\mathbf{R}^3$ ordnar p:s z-koordinat. Låt vidare ψ och $\bar{\psi}$ vara koordinatsystemen från avsnitt 3.3. Då är

$$\tilde{f}(\theta, \varphi) = f\left(\psi^{-1}(\theta, \varphi)\right) = z = \cos\theta\,,\ 0 < \theta < \pi\,,\ 0 < \varphi < 2\pi.$$

Som nämnts ovan identifierar vi ofta f med $\tilde{f}$ och skriver

$$f(\theta, \varphi) = \cos\theta\,,\ 0 < \theta < \pi\,,\ 0 < \varphi < 2\pi.$$

Om vi sätter $\tilde{\tilde{f}} = f \circ \bar{\psi}^{-1}$ får vi, i det streckade systemet

$$\tilde{\tilde{f}}\left(\bar{\theta}, \bar{\varphi}\right) = f\left(\bar{\psi}^{-1}\left(\bar{\theta}, \bar{\varphi}\right)\right) = z = \sin\bar{\theta}\cos\bar{\varphi}\,,\ 0 < \bar{\theta} < \pi, -\frac{\pi}{2} < \bar{\varphi} < \frac{3\pi}{2}.$$

Vi skriver

$$f\left(\bar{\theta}, \bar{\varphi}\right) = \sin\bar{\theta}\cos\bar{\varphi}\,,\ 0 < \bar{\theta} < \pi, -\frac{\pi}{2} < \bar{\varphi} < \frac{3\pi}{2}$$

eller möjligen

$$\bar{f}\left(\bar{\theta}, \bar{\varphi}\right) = \sin\bar{\theta}\cos\bar{\varphi}\,,\ 0 < \bar{\theta} < \pi, -\frac{\pi}{2} < \bar{\varphi} < \frac{3\pi}{2}$$

om vi vill markera koordinatbytet lite tydligare. □

Kapitel 4

Tensorer

4.1 Inledning

I detta kapitel ska vi generalisera en del begrepp från den linjära algebran. Vi ska t ex tala om vektorer definierade på en mångfald. Vi ska även generalisera begrepp som linjära avbildningar och skalärprodukter till mångfalder.

Den linjära algebran handlar som bekant om vektorrum och om ett vektorrum har ändlig dimension n är det, efter att vi infört en bas, väsentligen samma sak som $\mathbf{R}^n$, d v s vektorer adderas och multipliceras med tal genom att vektorernas komponenter (som är element i $\mathbf{R}^n$) adderas och multipliceras med tal. Lite informellt kan vi alltså säga att grundkursen i linjär algebra handlar om vektorer och linjära avbildningar på $\mathbf{R}^n$.

När vi ersätter $\mathbf{R}^n$ med en godtycklig mångfald ställs vi inför ett problem. En mångfald M har ingen naturlig vektorrumsstruktur. Det finns t ex inget naturligt sätt att addera två punkter på enhetssfären S och få en ny punkt på S. Vi är därför tvungna att på något sätt associera ett vektorrum till en punkt $p \in M$ innan vi kan tala om vektorer och linjära avbildningar (som i detta sammanhang kallas *tensorer*). Detta vektorrum kallas *tangentrummet* till M i p.

4.2 Enhetssfären, del 2

Även om enhetssfären S i $\mathbf{R}^3$ själv inte är ett vektorrum kan vi på ett naturligt sätt associera ett vektorrum till varje punkt $p \in S$, nämligen genom att betrakta enhetssfärens tangentplan T_pS i p. Vektorer i T_pS kallas *tangentvektorer* till S i p.

En kommentar om beteckningar: Tangentvektorer till S betecknas med latinska bokstäver skrivna i fetstil, t ex $\mathbf{v}$. Vektorer i $\mathbf{R}^3$ som inte nödvändigtvis tangerar S skrivs med vektorstreck. $\{\bar{e}_x, \bar{e}_y, \bar{e}_z\}$ betecknar t ex standardbasen i $\mathbf{R}^3$. Punkter betecknas med vanliga latinska bokstäver, t ex p. p:s ortsvektor betecknas, allt efter sammanhanget, med p eller med $\bar{r}(\theta, \varphi)$ om (θ, φ) är p:s koordinater i koordinatsystemet ψ. Observera dock att begreppet 'ortsvektor' bara är meningsfullt på delmängder av $\mathbf{R}^n$ och inte på mångfalder i allmänhet.

Observera att om $p \neq q$ så är $T_pS \neq T_qS$ i allmänhet. Tangentvektorer i p och tangentvektorer i q

är alltså olika typer av objekt och de kan därför inte relateras till varandra. Det är inte meningsfullt att bilda t ex $\mathbf{u} + \mathbf{v}$ om $\mathbf{u} \in T_pS$ och $\mathbf{v} \in T_qS$ om $p \neq q$. Vi måste tänka på tangentvektorer till S som fastsatta i den punkt där de tangerar S.

Ett problem med den 'vanliga' definitionen av tangentvektorer och tangentplan är att den inte går att generalisera till godtyckliga mångfalder eftersom definitionen inte bara refererar till S utan också till ett omkringliggande $\mathbf{R}^3$. Vi behöver därför hitta en beskrivning av tangentvektorer och tangentplan som bara refererar till S själv.

Låt $p \in S$ vara godtycklig. Med beteckningar från avsnitt 3.3 måste $p \in U$ eller $p \in \bar{U}$. Vi betraktar fallet $p \in U$, andra fallet behandlas snarlikt. Antag att p:s koordinater i koordinatsystemet ψ är (θ_0, φ_0) och låt Γ_θ vara en kurva genom p sådan att Γ_θ ligger helt på S och φ dessutom är konstant på S. En sådan kurva (en (del av en) meridian) får parameterframställningen

$$\begin{cases} x &= x(\theta, \varphi_0) &= \sin\theta\cos\varphi_0 \\ y &= y(\theta, \varphi_0) &= \sin\theta\sin\varphi_0 \\ z &= z(\theta, \varphi_0) &= \cos\theta \end{cases}.$$

Ortsvektorn $\bar{r}(\theta, \varphi_0)$ för en punkt på Γ_θ ges alltså av

$$\bar{r}(\theta, \varphi_0) = x(\theta, \varphi_0)\,\bar{e}_x + y(\theta, \varphi_0)\,\bar{e}_y + z(\theta, \varphi_0)\,\bar{e}_z = \sin\theta\cos\varphi_0\,\bar{e}_x + \sin\theta\sin\varphi_0\,\bar{e}_y + \cos\theta\,\bar{e}_z.$$

Dessutom är förstås $\bar{r}(\theta_0, \varphi_0) = p$.

Tangentvektorn $\mathbf{e}_\theta$ till Γ_θ i p ges därför av

$$\mathbf{e}_\theta = \bar{r}'_\theta(\theta_0, \varphi_0) = \left(\frac{\partial x}{\partial\theta}\,\bar{e}_x + \frac{\partial y}{\partial\theta}\,\bar{e}_y + \frac{\partial z}{\partial\theta}\,\bar{e}_z\right)\bigg|_{(\theta,\varphi)=(\theta_0,\varphi_0)} = \cos\theta_0\cos\varphi_0\,\bar{e}_x + \cos\theta_0\sin\varphi_0\,\bar{e}_y - \sin\theta_0\,\bar{e}_z.$$

På samma sätt låter vi $\mathbf{e}_\varphi$ vara tangentvektorn i p till en kurva Γ_φ på S där θ är konstant och Γ_φ är parametriserad med koordinaten φ. Då blir

$$\mathbf{e}_\varphi = \left(\frac{\partial x}{\partial\varphi}\,\bar{e}_x + \frac{\partial y}{\partial\varphi}\,\bar{e}_y + \frac{\partial z}{\partial\varphi}\,\bar{e}_z\right)\bigg|_{(\theta,\varphi)=(\theta_0,\varphi_0)} = -\sin\theta_0\sin\varphi_0\,\bar{e}_x + \sin\theta_0\cos\varphi_0\,\bar{e}_y.$$

Nu är $\mathbf{e}_\theta$ och $\mathbf{e}_\varphi$ tangentvektorer till S i p (ty de tangerar kurvor genom p som ligger helt i S) som dessutom är icke-parallella, så enligt satsen om rätt antal element är $\{\mathbf{e}_\theta, \mathbf{e}_\varphi\}$ en bas för tangentplanet till S i p. T_pS är alltså det linjära höljet till $\{\mathbf{e}_\theta, \mathbf{e}_\varphi\}$, d v s $T_pS = [\mathbf{e}_\theta, \mathbf{e}_\varphi] = \{$linjärkombinationer av $\mathbf{e}_\theta$ och $\mathbf{e}_\varphi\}$.

Observera att relationen $T_pS = [\mathbf{e}_\theta, \mathbf{e}_\varphi]$ i sig inte refererar till något omkringliggande $\mathbf{R}^3$. Om vi kan definiera vektorerna $\mathbf{e}_\theta$ och $\mathbf{e}_\varphi$ utan att referera till $\mathbf{R}^3$ är vi alltså klara. För att göra detta utnyttjar vi det faktum att definitionen av ett vektorrum är helt abstrakt, d v s den uttalar sig inte om vad för slags objekt vektorer ska vara, utan säger bara vad vi ska kunna göra med dem. Vi är alltså inte tvungna att definiera $\mathbf{e}_\theta$ och $\mathbf{e}_\varphi$ som riktade sträckor eller taltripplar utan vilka slags objekt som helst som kan adderas och multipliceras med tal duger. T ex kan vi definiera $\mathbf{e}_\theta$ och $\mathbf{e}_\varphi$ som lämpliga deriveringsoperatorer, nämligen riktningsderivator.

I flervariabelkursen definieras riktningsderivatan av en funktion $g : \mathbf{R}^n \to \mathbf{R}$ i punkten p med avseende på riktningen $\mathbf{v}$ som gränsvärdet

$$g'_{\mathbf{v}}(p) = \lim_{t \to 0} \frac{g(p + t\mathbf{v}) - g(p)}{t} \tag{4.1}$$

och det gäller att

$$g'_{\mathbf{v}}(p) = \operatorname{grad} g(p) \cdot \mathbf{v}$$

om $g \in \mathcal{C}^1$. I flervariabelkursen är vi främst intresserade av fallet $|\mathbf{v}| = 1$ eftersom riktningsderivatan av g då kan tolkas som g:s tillväxthastighet i $\mathbf{v}$:s riktning, men det finns inget som hindrar att vi låter $\mathbf{v}$ vara godtycklig och vi får på så sätt en mer generell riktningsderivata.

En stor fördel med denna generalisering är att om vi ser på en riktningsderivata som en operator $g \mapsto g'_{\mathbf{v}}(p)$, där p är fix så blir, på grund av ovanstående, denna operator linjär i $\mathbf{v}$. Mängden av alla riktningsderivator i en viss punkt i $\mathbf{R}^n$ bildar alltså ett n-dimensionellt vektorrum.

Sätt nu $r = \sqrt{x^2 + y^2 + z^2}$. Då är (r, θ, φ) 'vanliga' sfäriska koordinater på $\mathbf{R}^3$ i en omgivning av p. Låt $g : \mathbf{R}^3 \to \mathbf{R}$ vara en godtycklig funktion och låt f vara g:s restriktion till S (observera att även $f : S \to \mathbf{R}$ är godtycklig, eftersom g är det). Betrakta riktningsderivatan

$$\begin{aligned} g'_{\mathbf{e}_\theta}(p) &= \operatorname{grad} g(p) \cdot \mathbf{e}_\theta = \left[\left(\frac{\partial g}{\partial x}\, \bar{e}_x + \frac{\partial g}{\partial y}\, \bar{e}_y + \frac{\partial g}{\partial z}\, \bar{e}_z \right) \cdot \left(\frac{\partial x}{\partial \theta}\, \bar{e}_x + \frac{\partial y}{\partial \theta}\, \bar{e}_y + \frac{\partial z}{\partial \theta}\, \bar{e}_z \right) \right] \Bigg|_{(r,\theta,\varphi)=(1,\theta_0,\varphi_0)} \\ &= \left(\frac{\partial g}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial \theta} + \frac{\partial g}{\partial z}\frac{\partial z}{\partial \theta} \right) \Bigg|_{(r,\theta,\varphi)=(1,\theta_0,\varphi_0)} = /\text{ kedjeregeln }/ = \frac{\partial g}{\partial \theta}(1, \theta_0, \varphi_0) = \frac{\partial f}{\partial \theta}(p), \end{aligned}$$

där vi, i enlighet med tidigare konvention, identifierat f med $\tilde{f} = f \circ \psi^{-1}$. $g'_{\mathbf{e}_\theta}(p)$ beror alltså bara på g:s värden på S. På samma sätt fås

$$g'_{\mathbf{e}_\varphi}(p) = \frac{\partial f}{\partial \varphi}(p)$$

och om $\mathbf{v} = v^0 \mathbf{e}_\theta + v^1 \mathbf{e}_\varphi$ för reella tal v^0 och v^1 är

$$g'_{\mathbf{v}}(p) = g'_{v^0\mathbf{e}_\theta + v^1\mathbf{e}_\varphi}(p) = v^0 g'_{\mathbf{e}_\theta}(p) + v^1 g'_{\mathbf{e}_\varphi}(p) = v^0 \frac{\partial f}{\partial \theta}(p) + v^1 \frac{\partial f}{\partial \varphi}(p)$$

eftersom $g'_{\mathbf{v}}$ är linjär i $\mathbf{v}$ enligt ovan.

Alltså: Varje tangentvektor $\mathbf{v} = v^0 \mathbf{e}_\theta + v^1 \mathbf{e}_\varphi \in T_p S$ svarar mot en riktningsderivata $g \mapsto g'_{\mathbf{v}}(p)$ där $g : \mathbf{R}^3 \to \mathbf{R}$ och varje sådan riktningsderivata svarar enligt ovan mot en operator $f \mapsto v^0 \frac{\partial f}{\partial \theta}(p) + v^1 \frac{\partial f}{\partial \varphi}(p)$ där $f : S \to \mathbf{R}$ (och omvänt, i båda fallen) och dessa sistnämnda operatorer refererar inte till något omkringliggande $\mathbf{R}^3$.

Vi inför nu beteckningen

$$\mathbf{v}(f) = v^0 \frac{\partial f}{\partial \theta}(p) + v^1 \frac{\partial f}{\partial \varphi}(p). \tag{4.2}$$

Om vi *definierar* tangentplanet T_pS som mängden av alla operatorer på formen

$$f \mapsto \mathbf{v}(f) = v^0 \frac{\partial f}{\partial \theta}(p) + v^1 \frac{\partial f}{\partial \varphi}(p)$$

där $f : S \to R$ är denna definition alltså ekvivalent med den gamla, men har fördelen att den inte refererar till något omkringliggande $\mathbf{R}^3$. Denna definition är därför lämplig att generalisera till mångfalder.

4.3 Tangentvektorer

Vi är nu redo att definiera tangentvektorer till en godtycklig mångfald. Låt M vara en n-dimensionell mångfald och låt $p \in M$. Låt vidare $(x^a)_{a=0}^{n-1}$ vara ett koordinatsystem på M i en omgivning till p.

Definition 4.1. Med $\frac{\partial}{\partial x^a}\big|_p$ (eller, om det inte råder någon tvekan om vilken punkt som avses, $\frac{\partial}{\partial x^a}$) menas operatorn $f \mapsto \frac{\partial f}{\partial x^a}(p)$ där f är en $\mathcal{C}^\infty$ funktion definierad i en omgivning av p.

Anmärkning: Om man ska vara noga ges operatorn av $f \mapsto \frac{\partial}{\partial x^a}\left(f \circ \psi^{-1}\right)(p)$ där ψ är koordinatsystemet ovan, men enligt tidigare konvention identifierar vi f och $\tilde{f} = f \circ \psi^{-1}$. □

Definition 4.2. *Tangentrummet* T_pM till M i p definieras som det linjära höljet till mängden $\left\{\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^{n-1}}\right\}$ d v s

$$T_pM = \left[\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^{n-1}}\right] = \left\{\text{linjärkombinationer av } \left(\frac{\partial}{\partial x^a}\right)_{a=0}^{n-1}\right\}.$$

Elementen i T_pM kallas *tangentvektorer* till M i p. Basen $\left\{\left(\frac{\partial}{\partial x^a}\right)_{a=0}^{n-1}\right\}$ för T_pM kallas *koordinatbasen* hörande till koordinatsystemet $(x^a)_{a=0}^{n-1}$.

Att $\left\{\left(\frac{\partial}{\partial x^a}\right)_{a=0}^{n-1}\right\}$ utgör en bas för T_pM är förstås inte självklart även om mängden innehåller rätt antal element. För att rättfärdiga definitionen ovan behöver vi därför bevisa följande:

SATS 4.3. $\left\{\left(\frac{\partial}{\partial x^a}\right)_{a=0}^{n-1}\right\}$ *är linjärt oberoende.*

Bevis. Ansätt en linjär relation $\sum_{a=0}^{n-1} \lambda^a \frac{\partial}{\partial x^a} = 0$ där λ^a är reella tal. Vi ska visa att detta implicerar att $\lambda^a = 0$ för $a = 0, 1, \ldots n-1$.

Detta gör vi genom att, i tur och ordning, låta tangentvektorn $\sum_{a=0}^{n-1} \lambda^a \frac{\partial}{\partial x^a}$ verka på koordinatfunktionerna x^b, $b = 0, 1, \ldots, n-1$. Vi börjar med x^0. Då det gäller att $\frac{\partial x^0}{\partial x^0} = 1$ och att $\frac{\partial x^0}{\partial x^a} = 0$ om $a \neq 0$ ger ovanstående relation att

$$0 = \sum_{a=0}^{n-1} \lambda^a \frac{\partial x^0}{\partial x^a} = \lambda^0 \cdot 1 + \lambda^1 \cdot 0 + \ldots + \lambda^{n-1} \cdot 0 = \lambda^0,$$

och att övriga $\lambda^a = 0$ visas på samma sätt. □

Vi vet redan att $\left\{\left(\frac{\partial}{\partial x^a}\right)_{a=0}^{n-1}\right\}$ också spänner upp T_pM så $\left\{\left(\frac{\partial}{\partial x^a}\right)_{a=0}^{n-1}\right\}$ utgör en bas för T_pM och Definition 4.2 är alltså meningsfull.

Anmärkning: Även om vi definierat tangentvektorer som riktningsderivator finns det inget som hindrar att vi fortsätter *tänka på* tangentvektorer som riktade sträckor i $\mathbf{R}^n$.

I kursen i linjär algebra ges ju en helt abstrakt definition av ett vektorrum (d v s definitionen talar bara om vad man ska kunna göra med elementen i ett vektorrum, den säger ingenting om hur elementen ska vara definierade) och till syvende och sist är en vektor ingenting annat än ett element i ett vektorrum.

Så länge vi enbart är intresserade av vektorrumsegenskaper är det alltså likgiltigt om tangentvektorer är definierade som riktningsderivator eller riktade sträckor. □

Om $\mathbf{v}$ är en tangentvektor i p är $\mathbf{v}$ alltså en linjärkombination av $\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^{n-1}}$, d v s det finns reella tal v^a, $a = 0, 1, \ldots, n-1$ sådana att

$$\mathbf{v} = \sum_{a=0}^{n-1} v^a \frac{\partial}{\partial x^a}.$$

Verkan av en tangentvektor $\mathbf{v}$ i p på en godtycklig $\mathcal{C}^\infty$ funktion f ges alltså av

$$\mathbf{v}(f) = \sum_{a=0}^{n-1} v^a \left.\frac{\partial f}{\partial x^a}\right|_p . \tag{4.3}$$

Observera att detta är en direkt generalisering av ekvation (4.2).

Summor av detta slag dyker upp väldigt ofta i relativitetsteorin så vi tjänar i längden på att införa ett enklare skrivsätt för dem. Ett sådant skrivsätt infördes av Einstein själv och kallas därför *Einsteins summationskonvention.* Den säger att om ett index förekommer exakt två gånger i en term, en gång uppe och en gång nere, så ska detta index summeras över.[1] I detta sammanhang ska index *under* ett bråkstreck tolkas som om de var placerade motsatt sin riktiga plats. I fortsättningen används alltid denna konvention om inget annat sägs.

I ovanstående summa sitter indexet i v^a förstås uppe och indexet i faktorn $\frac{\partial}{\partial x^a}$ tolkas som att det sitter nere p g a bråkstrecket. När vi använder Einsteins summationskonvention skriver vi alltså

$$\mathbf{v} = v^a \frac{\partial}{\partial x^a} \tag{4.4}$$

och

$$\mathbf{v}(f) = v^a \left.\frac{\partial f}{\partial x^a}\right|_p . \tag{4.5}$$

istället för $\mathbf{v} = \sum_{a=0}^{n-1} v^a \frac{\partial}{\partial x^a}$ respektive $\mathbf{v}(f) = \sum_{a=0}^{n-1} v^a \left.\frac{\partial f}{\partial x^a}\right|_p$.

Övning: Skriv ned beviset av Sats 4.3 med användande av Einsteins summationskonvention.

[1]Som Einstein själv skämtsamt uttryckte saken: "I have made a great discovery in mathematics; I have suppressed the summation sign every time that the summation must be made over an index which occurs twice..."

Talen v^a kallas vektorns *komponenter* i koordinatsystemet x^a och det är vanligt att, något oegentligt, identifiera vektorn $\mathbf{v}$ med sina komponenter v^a och tala om 'vektorn v^a'. I så fall är det förstås viktigt att hålla reda på att komponenterna ändras ifall vi byter till nya koordinater $\bar{x}^a$. Detta är ämnet för följande sats:

SATS 4.4. *Låt $p \in M$ och låt x^a och $\bar{x}^a$ vara koordinatsystem i en omgivning av p. Om $\mathbf{v}$ är en tangentvektor till M i p och $\mathbf{v} = v^a \frac{\partial}{\partial x^a} = \bar{v}^a \frac{\partial}{\partial \bar{x}^a}$ så är*

$$\bar{v}^a = \frac{\partial \bar{x}^a}{\partial x^b} v^b. \tag{4.6}$$

Anmärkning: Glöm inte bort Einsteins summationskonvention! Med summatecknen utskrivna säger satsen att om

$$\mathbf{v} = \sum_{a=0}^{n-1} v^a \frac{\partial}{\partial x^a} = \sum_{a=0}^{n-1} \bar{v}^a \frac{\partial}{\partial \bar{x}^a}$$

så är

$$\bar{v}^a = \sum_{b=0}^{n-1} \frac{\partial \bar{x}^a}{\partial x^b} v^b,$$

för $a = 0, 1, \ldots, n-1$. □

Anmärkning: Observera också en viktig skillnad mellan indexen a och b i ekvationen

$$\bar{v}^a = \frac{\partial \bar{x}^a}{\partial x^b} v^b.$$

a är ett så kallat *fritt index* och ekvationen betyder att båda leden är lika för alla $a = 0, 1, \ldots, n-1$. En förutsättning för att en ekvation av detta slag över huvud taget ska vara meningsfull är att båda leden innehåller exakt samma uppsättning fria index och att dessa fria index sitter på samma plats (uppe eller nere) i båda leden. b är å andra sidan ett *summationsindex*, eller ett *dummy index*. Detta betyder att det är tillåtet att, precis som i alla summor, byta ut detta index mot en annan bokstav så länge denna bokstav inte används för att beteckna ett fritt index eller ett annat summationsindex. T ex är

$$\frac{\partial \bar{x}^a}{\partial x^b} v^b = \frac{\partial \bar{x}^a}{\partial x^c} v^c \neq \frac{\partial \bar{x}^c}{\partial x^a} v^a.$$

medan uttrycket $\frac{\partial \bar{x}^a}{\partial x^a} v^a$ helt saknar mening. □

Anmärkning: En tredje sak att observera är att när vi skriver $\frac{\partial \bar{x}^a}{\partial x^b}$ så menar vi egentligen $\frac{\partial \bar{x}^a}{\partial x^b}\Big|_p$. Denna förenkling av notationen är vanlig så länge det inte råder någon tvekan om vilken punkt som avses. □

Bevis av SATS 4.4. Enligt kedjeregeln är

$$\frac{\partial f}{\partial x^b} = \sum_{a=0}^{n-1} \frac{\partial \bar{x}^a}{\partial x^b} \frac{\partial f}{\partial \bar{x}^a} = \frac{\partial \bar{x}^a}{\partial x^b} \frac{\partial f}{\partial \bar{x}^a}$$

enligt summationskonventionen. Genom att låta deriveringsoperatorn $\mathbf{v}$ verka på en godtycklig $\mathcal{C}^\infty$ funktion f fås

$$\bar{v}^a \frac{\partial f}{\partial \bar{x}^a} = \mathbf{v}(f) = v^b \frac{\partial f}{\partial x^b} = v^b \frac{\partial \bar{x}^a}{\partial x^b} \frac{\partial f}{\partial \bar{x}^a}$$

Då f var godtycklig följer att

$$\bar{v}^a \frac{\partial}{\partial \bar{x}^a} = \mathbf{v} = v^b \frac{\partial \bar{x}^a}{\partial x^b} \frac{\partial}{\partial \bar{x}^a},$$

d v s vi har skrivit $\mathbf{v}$ som en linjärkombination av basvektorerna i basen $\left\{\frac{\partial}{\partial \bar{x}^0}, \frac{\partial}{\partial \bar{x}^1}, \ldots, \frac{\partial}{\partial \bar{x}^{n-1}}\right\}$ på två sätt. Eftersom dessa vektorer är linjärt oberoende är koefficienterna i linjärkombinationen entydigt bestämda, varför

$$\bar{v}^a = v^b \frac{\partial \bar{x}^a}{\partial x^b}.$$

$\square$

En uppsättning tal v^a som uppfyller denna transformationslag sägs transformeras *kontravariant*.

Transformationslagen (4.6) (och många som liknar den) dyker upp i många sammanhang och vi gör därför följande definition:

Definition 4.5. En uppsättning tal v^a som uppfyller transformationslagen $\bar{v}^a = \frac{\partial \bar{x}^a}{\partial x^b} v^b$ vid ett koordinatbyte kallas en *kontravariant tensor*[2]

Kontravarianta tensorer (d v s en tangentvektors komponenter) är specialfall av *tensorer*. Dessa ska vi studera närmare i återstoden av detta kapitel.

Först ska vi dock införa en del begrepp som vi kommer att behöva i senare kapitel. Minns att en $\mathcal{C}^\infty$ *kurva* i $\mathbf{R}^n$ är en vektorvärd $\mathcal{C}^\infty$ funktion $x^a = x^a(\tau)$ där τ tillhör något intervall I. Denna definition lämpar sig väl för generalisering till en mångfald M. Vi måste dock iaktta viss försiktighet eftersom det inte säkert finns något koordinatsystem som täcker hela M.

Definition 4.6. Låt $I \subset \mathbf{R}$ vara ett intervall. En funktion $\gamma : I \to M$ kallas en $\mathcal{C}^\infty$ *kurva* på M om det för varje koordinatsystem ψ på M gäller att funktionen $\psi \circ \gamma : I \to \mathbf{R}^n$ är av klass $\mathcal{C}^\infty$. Om I också förses med en genomloppsriktning sägs γ vara *orienterad*.

Kurvor av annan regularitet än $\mathcal{C}^\infty$ definieras analogt.

Observera att för givet τ är $\gamma(\tau)$ en punkt $p \in M$ och $(\psi \circ \gamma)(\tau) = \psi(\gamma(\tau))$ är då p:s koordinater x^a i koordinatsystemet ψ. Som tidigare identifierar vi oftast p med sina koordinater och beskriver kurvan genom *parameterframställningen* $x^a = x^a(\tau)$, där vänsterledet ska tolkas som koordinaterna x^a för en punkt p på γ och högerledet ska tolkas som komponenterna till den vektorvärda funktionen $\psi(\gamma(\tau))$.

[2]Här skiljer sig terminologin åt mellan olika författare. En del föredrar att (som här) definiera tensorer utifrån deras transformationsegenskaper eftersom detta ger en förhållandevis enkel definition. Andra föredrar att göra teorin så koordinatfri som möjligt eftersom målet är att beskriva verkligheten, och verkligheten är rimligen densamma oavsett vilka koordinater vi använder för att beskriva den. Dessa författare använder därför en annan (ekvivalent i sak, men konceptuellt annorlunda) definition av begreppet kontravariant tensor. Om denna kan man läsa i t ex *Wald*: General Relativity.

Om kurvan γ_1 har parameterframställningen $x^a = x^a(\sigma)$, $\sigma \in J$ och genomlöper samma punkter lika många gånger och i samma ordning som γ säger man ofta att γ och γ_1 är olika parameterframställningar av *samma* kurva.

Låt γ vara en kurva på M med parameterframställning $x^a = x^a(\tau)$ och låt p vara en punkt på γ svarande mot parametervärdet τ_0 och sätt

$$v^a = \left(\frac{dx^a}{d\tau}\right)\bigg|_{\tau=\tau_0}.$$

Vi observerar först att v^a är kontravariant tensor i p, d v s v^a är komponenter i koordinatbasen till en vektor $\mathbf{v}$ i p, ty om $\bar{x}^a$ är nya koordinater i en omgivning av p så är

$$\bar{v}^a = \left(\frac{d\bar{x}^a}{d\tau}\right)\bigg|_{\tau=\tau_0} = \left(\frac{\partial\bar{x}^a}{\partial x^b}\frac{dx^b}{d\tau}\right)\bigg|_{\tau=\tau_0} = \frac{\partial\bar{x}^a}{\partial x^b}v^b,$$

enligt kedjeregeln. v^a uppfyller alltså transformationslagen (4.6) och är därför en kontravariant tensor, d v s komponenter till en tangentvektor $\mathbf{v}$ till M i p. $\mathbf{v}$ kallas en *tangentvektor* till γ.

Antag nu att f är en (snäll) funktion på M och bilda f:s restriktion g till γ, d v s sätt $g(\tau) = f(x^a(\tau))$ (där vi som vanligt identifierar punkter $\gamma(\tau)$ på γ med sina koordinater $x^a(\tau)$).

Vi beräknar $g'(\tau_0)$ med hjälp av kedjeregeln:

$$g'(\tau_0) = \left(\frac{d}{d\tau}f(x^a(\tau))\right)\bigg|_{\tau=\tau_0} = \left(\frac{\partial f}{\partial x^a}\frac{dx^a}{d\tau}\right)\bigg|_{\tau=\tau_0} = v^a\frac{\partial f}{\partial x^a} = \mathbf{v}(f).$$

Denna räkning motiverar varför $\mathbf{v}$ kallas en *tangentvektor* till γ. Under ganska milda villkor blir kurvan γ nämligen själv en 1-dimensionell mångfald och parametern τ blir ett koordinatsystem på γ. Tangentrummet till γ i p blir då endimensionellt och avbildningen $\frac{d}{d\tau} : g \mapsto g'(\tau_0)$ bildar koordinatbasen för detta tangentrum. Ovanstående räkning visar då att $\frac{d}{d\tau}$ direkt kan identifieras med vektorn $\mathbf{v}$ på M.

4.4 Kovarianta tensorer

Som tidigare låter vi M vara en n-dimensionell mångfald med lokala koordinater x^a i en omgivning av en godtycklig punkt $p \in M$. Alla partiella derivator som dyker upp i detta avsnitt underförstås (som tidigare) vara evaluerade i p. Låt nu dx^a betyda ett litet tillskott[3] i koordinaten x^a så att punkten q med koordinater $x^a + dx^a$ ligger nära p.

Definition 4.7. *Differentialen* till en funktion $f : M \to \mathbf{R}$ är $df = \frac{\partial f}{\partial x^a}dx^a$.

Observera placeringen av indexet. Då x^a har indexet uppe men står under bråkstrecket anses $\frac{\partial f}{\partial x^a}$ ha indexet nere och vi skriver därför ofta

$$\nabla_a f = \frac{\partial f}{\partial x^a}. \tag{4.7}$$

[3]En mer precis definition av dx^a ges i kapitel 6.

Med detta skrivsätt är alltså $\nabla_a f$ den a:te komponenten till df i basen dx^a. Vi kommer därför (något oegentligt) att säga att differentialen till f ges av $df_a = \nabla_a f = \frac{\partial f}{\partial x^a}$.

Låt oss nu byta koordinater till $\bar{x}^a$ och betrakta $\frac{\partial f}{\partial \bar{x}^a}$. Enligt kedjeregeln är

$$\frac{\partial f}{\partial \bar{x}^a} = \frac{\partial x^b}{\partial \bar{x}^a} \frac{\partial f}{\partial x^b}$$

eller, med $v_a = \frac{\partial f}{\partial x^a}$ och $\bar{v}_a = \frac{\partial f}{\partial \bar{x}^a}$,

$$\bar{v}_a = \frac{\partial x^b}{\partial \bar{x}^a} v_b. \tag{4.8}$$

Notera att detta *inte* är samma transformationslag som i Sats 4.4 även om den har vissa likheter. $v_a = \frac{\partial f}{\partial x^a}$ transformeras alltså *inte* kontravariant. Detta innebär att om v_a och $\bar{v}_a$ är komponenter till något matematiskt objekt V i koordinatsystemen x^a respektive $\bar{x}^a$ så kan V *inte* vara en tangentvektor till M för i så fall skulle ju v_a och $\bar{v}_a$ uppfylla transformationslagen (4.6) och inte (4.8).

En uppsättning tal v_a som uppfyller transformationslagen (4.8) sägs transformeras *kovariant* och vi gör följande definition:

Definition 4.8. En uppsättning tal v_a som uppfyller transformationslagen $\bar{v}_a = \frac{\partial x^b}{\partial \bar{x}^a} v_b$ sägs bilda en *kovariant tensor*.

Enligt ovan är tydligen differentialen till en funktion $f : M \to \mathbf{R}$ ett exempel på en kovariant tensor.

Lägg också märke till hur indexen är placerade. Kontravarianta tensorer har ett index uppe medan kovarianta tensorer har ett index nere.

($\star$ Ovanstående kortfattade beskrivning av kovarianta tensorer räcker gott och väl för våra syften i denna kurs, men många (inklusive jag själv) finner den otillfredsställande. Skälet till detta är följande.

Betrakta en tangentvektor (d v s en riktningsderivata) $\mathbf{v}$. Dess komponenter v^a och $\bar{v}^a$ i två olika koordinatsystem x^a respektive $\bar{x}^a$ uppfyller transformationslagen (4.6) och de utgör därför en kontravariant tensor. Omvänt: Givet en kontravariant tensor v^a så kan vi återskapa $\mathbf{v}$ genom

$$\mathbf{v} = v^a \frac{\partial}{\partial x^a}$$

och om vi istället återskapar $\mathbf{v}$ från $\bar{v}^a$ genom

$$\mathbf{v} = \bar{v}^a \frac{\partial}{\partial \bar{x}^a}$$

får vi precis samma tangentvektor $\mathbf{v}$.

Om vi försöker beskriva en kovariant tensor α_a på liknande sätt stöter vi på patrull eftersom vi inte har tillgång till något begrepp motsvarande tangentvektorn $\mathbf{v}$ att utgå från. För att ge en liknande beskrivning skulle vi behöva konstruera någon typ av basvektorer från koordinaterna x^a som vi t ex skulle kunna beteckna dx^a (jämför definitionen av differential ovan), och sedan definiera något slags matematiskt objekt α (som vi senare kommer att kalla en *kovektor*) som linjärkombinationer av dx^a så att

$$\alpha = \alpha_a dx^a.$$

Hur detta kan göras ska vi studera i detalj i kapitel 6. $\star$)

4.5 Allmänna tensorer

Ekvationerna (4.6) och (4.8) kan nu enkelt generaliseras till uppsättningar av tal indexerade med fler än ett index. Om t ex T^{ab} transformeras enligt

$$\bar{T}^{ab} = \frac{\partial \bar{x}^a}{\partial x^c}\frac{\partial \bar{x}^b}{\partial x^d}T^{cd} \tag{4.9}$$

(observera att detta är en dubbelsumma över c och d) vid ett koordinatbyte från x^a till $\bar{x}^a$ sägs T^{ab} vara en *kontravariant tensor av rang 2* och om istället T_{ab} transformeras enligt

$$\bar{T}_{ab} = \frac{\partial x^c}{\partial \bar{x}^a}\frac{\partial x^d}{\partial \bar{x}^b}T_{cd} \tag{4.10}$$

sägs T_{ab} vara en *kovariant tensor av rang 2*. Vi kan också blanda kovarianta och kontravarianta index. Om $T^a{}_b$ transformeras enligt

$$\bar{T}^a{}_b = \frac{\partial \bar{x}^a}{\partial x^c}\frac{\partial x^d}{\partial \bar{x}^b}T^c{}_d \tag{4.11}$$

sägs $T^a{}_b$ vara en *(blandad) tensor av typ (1,1)* och den sägs ha *kontravariant rang 1* och *kovariant rang 1*.

Anmärkning: Även i dessa formler underförstås att de partiella derivatorna ska evalueras i någon given men godtycklig punkt $p \in M$. En tensor är alltså, precis som en tangentvektor, kopplad till en specifik punkt på M. För att förtydliga säger man därför ofta att 'T_{ab} är en tensor i p' istället för 'T_{ab} är en tensor'. $\square$

I en blandad tensor är det viktigt att hålla reda på indexens placering eftersom vi så småningom kommer att definiera operationer som höjer och sänker index. Vi skriver därför *aldrig* T^a_b eftersom det i så fall inte är tydligt om vi menar $T^a{}_b$ eller $T_b{}^a$ och dessa tensorer är olika i allmänhet.

Ovanstående definitioner är alla specialfall definitionen av en allmän tensor:

Definition 4.9. Om talen $T^{a_1a_2\ldots a_s}{}_{b_1b_2\ldots b_t}$ transformeras enligt

$$\bar{T}^{a_1a_2\ldots a_s}{}_{b_1b_2\ldots b_t} = \frac{\partial \bar{x}^{a_1}}{\partial x^{c_1}}\frac{\partial \bar{x}^{a_2}}{\partial x^{c_2}}\cdots\frac{\partial \bar{x}^{a_s}}{\partial x^{c_s}}\frac{\partial x^{d_1}}{\partial \bar{x}^{b_1}}\frac{\partial x^{d_2}}{\partial \bar{x}^{b_2}}\cdots\frac{\partial x^{d_t}}{\partial \bar{x}^{b_t}}T^{c_1c_2\ldots c_s}{}_{d_1d_2\ldots d_t} \tag{4.12}$$

vid ett koordinatbyte från x^a till $\bar{x}^a$ sägs $T^{a_1a_2\ldots a_s}{}_{b_1b_2\ldots b_t}$ vara en *(blandad) tensor av typ (s,t)* och vi säger att $T^{a_1a_2\ldots a_s}{}_{b_1b_2\ldots b_t}$ har *kontravariant rang s* och *kovariant rang t*.

Låt nu S vara en allmän uppsättning tal, indexerade med ett antal kovarianta och kontravarianta index sådana att de inte nödvändigtvis är ordnade som i ovanstående definition med alla kontravarianta index till vänster och alla kovarianta index till höger. Som exempel kan vi betrakta $S^a{}_c{}^b$. Vi kan då bilda en ny uppsättning tal T, med indexen ordnade som i ovanstående definition, genom att helt enkelt flytta om indexen. I vårt exempel sätter vi alltså $T^{ab}{}_c = S^a{}_c{}^b$ och vi säger att även S är en tensor om T uppfyller ovanstående definition.

Blandade tensorer av typ (1,1) uppträder, förklädda och under annat namn i många tidigare kurser, vilket vi ska se i nästa exempel. Först måste vi dock göra några förberedelser.

Definition 4.10. $\delta_a{}^b = \begin{cases} 1 & \text{om } a = b \\ 0 & \text{om } a \neq b \end{cases}$ kallas *Kroneckers delta.*

SATS 4.11. *Kroneckers delta är en blandad tensor av typ (1,1).*

Bevis. Enligt definitionen har Kroneckers delta samma komponenter i alla koordinatsystem. Vi behöver alltså visa att

$$\frac{\partial x^c}{\partial \bar{x}^a}\frac{\partial \bar{x}^b}{\partial x^d}\delta_c{}^d = \bar{\delta}_a{}^b = \begin{cases} 1 & \text{om } a = b \\ 0 & \text{om } a \neq b \end{cases}.$$

Betrakta vänsterledet och notera att de enda termerna i dubbelsumman som ger något bidrag är de där $c = d$. Detta ger

$$\frac{\partial x^c}{\partial \bar{x}^a}\frac{\partial \bar{x}^b}{\partial x^d}\delta_c{}^d = \frac{\partial x^c}{\partial \bar{x}^a}\frac{\partial \bar{x}^b}{\partial x^c}.$$

Denna enkelsumma kan skrivas om till en enda term med hjälp av kedjeregeln eftersom

$$\frac{\partial x^c}{\partial \bar{x}^a}\frac{\partial \bar{x}^b}{\partial x^c} = \frac{\partial \bar{x}^b}{\partial \bar{x}^a}.$$

Nu är $\frac{\partial \bar{x}^b}{\partial \bar{x}^a} = 1$ om $a = b$ och $= 0$ annars, så det följer att $\frac{\partial x^c}{\partial \bar{x}^a}\frac{\partial \bar{x}^b}{\partial x^d}\delta_c{}^d = \bar{\delta}_a{}^b$. □

Eftersom Kroneckers delta har samma komponenter i alla koordinatsystem skriver vi normalt inte $\bar{\delta}_a{}^b$ utan vi låter $\delta_a{}^b$ beteckna tensorn oavsett koordinatsystem.

Exempel 4.12. Detta är ett långt exempel. Målet är att visa att definitionerna ovan gör tensorer till en naturlig generalisering av matrisbegreppet och att transformationslagarna är omskrivningar av basbytesfomlerna från kursen i linjär algebra.

Vi kommer konsekvent att låta kontravarianta index, d v s index uppe, beteckna radnummer och kovarianta index, d v s index nere beteckna kolonnummer.

En kontravariant tensor v^a kommer därför att representeras av kolonnmatrisen $v = \begin{bmatrix} v^0 \\ v^1 \\ \vdots \\ v^{n-1} \end{bmatrix}$ och en kovariant tensor u_a kommer att representeras av radmatrisen $u = [u_0\ u_1\ \cdots\ u_{n-1}]$. En blandad tensor $T^a{}_b$ av typ (1,1) kommer att representeras av en $n \times n$-matris T vars element på plats (a, b) är just $T^a{}_b$ d v s

$$T = \begin{bmatrix} T^0{}_0 & T^0{}_1 & \dots & T^0{}_{n-1} \\ T^1{}_0 & T^1{}_1 & \dots & T^1{}_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ T^{n-1}{}_0 & T^{n-1}{}_1 & \dots & T^{n-1}{}_{n-1} \end{bmatrix}$$

Låt nu $p \in M$, med lokala koordinater x^a i en omgivning till p och betrakta en linjär avbildning $\mathbf{F} : T_pM \to T_pM$ given av $\mathbf{w} = \mathbf{F}(\mathbf{v})$ där $\mathbf{v}$, $\mathbf{w} \in T_pM$.

Det visas i kursen i linjär algebra att om $\mathbf{v}$ och $\mathbf{w}$ har komponenterna $v = \begin{bmatrix} v^0 \\ v^1 \\ \vdots \\ v^{n-1} \end{bmatrix}$ respektive $w = \begin{bmatrix} w^0 \\ w^1 \\ \vdots \\ w^{n-1} \end{bmatrix}$ i basen $\left\{ \left(\frac{\partial}{\partial x^a}\right)_{a=0}^{n-1} \right\}$ så finns en $n \times n$-matris F sådan att $w = Fv$.

Vi låter $F^a{}_b$ vara elementet på plats (a, b) i matrisen F. Eftersom $\mathbf{w} = w^a \frac{\partial}{\partial x^a}$ så är

$$w^a \frac{\partial}{\partial x^a} = \mathbf{w} = \mathbf{F}(\mathbf{v}) = F^a{}_b v^b \frac{\partial}{\partial x^a}.$$

d v s ett ekvivalent sätt att skriva sambanden $\mathbf{w} = \mathbf{F}(\mathbf{v})$ och $w = Fv$ är $w^a = F^a{}_b v^b$, ty vektorerna $\frac{\partial}{\partial x^a}$ utgöt en bas för T_pM.

Det vi har gjort hittills i exemplet är förstås inget djupsinnigt. Det är självklart att elementen i en matris kan indexeras med två index, elementets radnummer och kolonnummer. Det som inte är självklart är hur indexen ska placeras. Varför inte lika gärna F_{ab} eller F^{ab}? Anledningen är att $F^a{}_b$ uppfyller transformationslagen för en blandad tensor av typ (1,1), vilket vi nu ska visa.

Inför nya koordinater $\bar{x}^a$ i en omgivning av p. Detta ger oss en ny koordinatbas $\frac{\partial}{\partial \bar{x}^a}$ för T_pM och vi söker bassambandet, d v s sambandet mellan de båda koordinatbaserna. Genom att låta $\frac{\partial}{\partial \bar{x}^a}$ verka på en godtycklig funktion g och använda kedjeregeln får vi

$$\frac{\partial g}{\partial \bar{x}^a} = \frac{\partial x^b}{\partial \bar{x}^a} \frac{\partial g}{\partial x^b} \, , \ \forall g$$

d v s bassambandet blir

$$\frac{\partial}{\partial \bar{x}^a} = \frac{\partial x^b}{\partial \bar{x}^a} \frac{\partial}{\partial x^b}.$$

Detta är ett gammalt välkänt samband i ny skepnad. För att inse detta byter vi beteckningar. Låt

$$\underline{\mathbf{e}} = \left[\frac{\partial}{\partial x^0} \ \frac{\partial}{\partial x^1} \cdots \frac{\partial}{\partial x^{n-1}} \right]$$

d v s $\underline{\mathbf{e}}$ är den vanliga basmatrisen för den ursprungliga basen. Analogt låter vi $\underline{\mathbf{f}}$ vara basmatrisen för den nya basen, d v s

$$\underline{\mathbf{f}} = \left[\frac{\partial}{\partial \bar{x}^0}\ \frac{\partial}{\partial \bar{x}^1} \cdots \frac{\partial}{\partial \bar{x}^{n-1}}\right]$$

Låt nu T vara den matris vars element på plats (a,b) är $\frac{\partial x^a}{\partial \bar{x}^b}$. Då kan bassambandet ovan skrivas på två ekvivalenta sätt

$$\frac{\partial}{\partial \bar{x}^a} = \frac{\partial x^b}{\partial \bar{x}^a}\frac{\partial}{\partial x^b} \iff \underline{\mathbf{f}} = \underline{\mathbf{e}}T,$$

där den andra formen bör vara bekant från kursen i linjär algebra.

Koordinatsambandet, d v s sambandet mellan vektorn $\mathbf{v}$:s komponenter i gamla och nya basen ges enligt ekvation (4.6) av

$$\bar{v}^a = \frac{\partial \bar{x}^a}{\partial x^b} v^b \iff \bar{v} = Uv,$$

där U:s element på plats (a,b) är $\frac{\partial \bar{x}^a}{\partial x^b}$.

Betrakta nu matrisprodukten UT. Dess element på plats (a,c) ges av

$$\frac{\partial \bar{x}^a}{\partial x^b}\frac{\partial x^b}{\partial \bar{x}^c} = \frac{\partial \bar{x}^a}{\partial \bar{x}^c} = \delta_c{}^a.$$

d v s $UT = I$ (enhetsmatrisen). Detta ger $U = T^{-1}$ och därför är koordinatsambandet

$$\bar{v} = T^{-1}v \iff v = T\bar{v}$$

vilket också bör vara bekant från kursen i linjär algebra.

Det är nu dags att återgå till den linjära avbildningen $\mathbf{F}$ och de ekvivalenta sambanden $\mathbf{w} = \mathbf{F}(\mathbf{v}) \Leftrightarrow w = Fv \Leftrightarrow w^a = F^a{}_b v^b$. Eftersom det inte är något speciellt med de ursprungliga koordinaterna x^a så kan vi förstås också skriva $\mathbf{w} = \mathbf{F}(\mathbf{v}) \Leftrightarrow \bar{w} = \bar{F}\bar{v}$ där $\bar{w} = T^{-1}w$ är $\mathbf{w}$:s komponenter i basen $\frac{\partial}{\partial \bar{x}^a}$ och $\bar{F}$ är $\mathbf{F}$:s avbildingsmatris i denna bas. Sätter vi in bas- och koordinatsambanden fås

$$\underline{\mathbf{f}}\,\bar{F}\bar{v} = \mathbf{F}(\mathbf{v}) = \underline{\mathbf{e}}\,Fv = \underline{\mathbf{f}}\,T^{-1}FT\bar{v}.$$

Eftersom $\underline{\mathbf{f}}$ är en bas och $\bar{v}$ är godtycklig följer sambandet

$$\bar{F} = T^{-1}FT \tag{4.13}$$

vilket också är en känd formel från linjär algebra. Låt nu $\bar{F}^a{}_b$ vara matrisen $\bar{F}$:s element på plats (a,b). Ekvation (4.13) skriven på summaform blir då

$$\bar{F}^a{}_b = \frac{\partial \bar{x}^a}{\partial x^c} F^c{}_d \frac{\partial x^d}{\partial \bar{x}^b}$$

vilket visar att $F^a{}_b$ transformeras som en blandad tensor av typ $(1,1)$. $\square$

Vi noterar att även om t ex en tensor T_{ab} av typ $(0,2)$ inte transformeras som en blandad tensor av typ (1,1) så kan det ändå vara ändamålsenligt att ordna dess komponenter i en matris. Vi kommer att se exempel på detta redan i nästa kapitel. Vi måste dock tolka en sådan matris med viss försiktighet. Låt v^a vara en kontravariant tensor och bilda $w_a = T_{ab}v^b$. Eftersom detta uttryck bara har ett fritt index som sitter nere förväntar vi oss att w_a är en kovariant tensor. Man kan visa (se nästa kapitel) att så också är fallet.

Låt nu T vara den $n \times n$-matris vars element på plats (a,b) är T_{ab} och låt som ovan v vara den kolonnmatris vars element på plats a är v^a. Matrisprodukten $w = Tv$ blir då en kolonnmatris vars element på plats a är $w_a = T_{ab}v^b$. Här måste vi vara försiktiga med tolkningen av elementen i kolonnmatrisen w.

I exemplet ovan representerade vi ju *kontravarianta* tensorer med kolonnmatriser eftersom index uppe associerades med radnummer. Kovarianta tensorer representerades istället av radmatriser.

Om vi väljer att representera en kovariant tensor av rang 2 med en matris måste vi alltså *själva* komma ihåg att kolonnvektorn w ovan ska associeras med en *kovariant* tensor w_a. Notationen ger oss ingen hjälp med att komma ihåg detta.

När vi ändå talar om matriser ska vi undanröja ett vanligt missförstånd gällande faktorernas ordning i uttryck av det slag vi studerat ovan:

Anmärkning: Låt A och B vara matriser vars element på plats (a,b) ges av $A^a{}_b$ respektive $B^a{}_b$ och bilda deras matrisprodukt $C = AB$. C:s element på plats (a,b) blir då

$$C^a{}_b = A^a{}_c B^c{}_b.$$

Observera att multiplikation av tal är kommutativ så vi kan lika gärna skriva

$$C^a{}_b = B^c{}_b A^a{}_c.$$

Vid första anblicken verkar kanske denna operation lite suspekt eftersom matrismultiplikation, till skillnad från multiplikation av tal, inte är kommutativ. Trots detta ser det kanske ut som om vi kastat om ordningen på matriserna A och B i matrisprodukten. Detta har vi emellertid inte alls gjort. För att inse detta, bilda matrisprodukten $D = BA$ och betrakta D:s element $D^a{}_b$ på plats (a,b):

$$D^a{}_b = B^a{}_c A^c{}_b \neq B^c{}_b A^a{}_c = C^a{}_b.$$

Faktorernas inbördes ordning i en matrisprodukt som är skriven på tensorform bestäms alltså helt av vilka index det summeras över. Om vi skriver A- eller B-faktorn först har ingen betydelse. □

4.6 Tensorfält

Från kursen i vektoranalys erinrar vi oss att en tillräckligt snäll funktion som till varje punkt i (någon delmängd av) $\mathbf{R}^3$ ordnar en vektor i $\mathbf{R}^3$, kallas ett vektorfält. Vi ska nu generalisera denna definition till en mångfald M.

Låt $\psi : U \to D \subset \mathbf{R}^n$ (där $U \subset M$ är öppen) vara ett koordinatsystem på U och betrakta en funktion som till varje punkt p i U ordnar en tangentvektor

$$p \mapsto \mathbf{v}(p) = v^a \left. \frac{\partial}{\partial x^a} \right|_p \in T_pM.$$

I fortsättningen kommer vi oftast att skriva $\mathbf{v}$ istället för $\mathbf{v}(p)$.

v^a blir då funktioner av p och det är klart att $\mathbf{v}$ helt bestäms av dessa funktioner. $\tilde{v}^a = v^a \circ \psi^{-1}$ blir då n st 'vanliga' reellvärda funktioner av koordinaterna x^a och vi säger att $\mathbf{v} \in \mathcal{C}^\infty(U)$ om $\tilde{v}^a \in \mathcal{C}^\infty(D)$ (jfr avsnitt 3.5).

Om $\mathbf{v} \in \mathcal{C}^\infty(U)$ säger vi att $\mathbf{v}$ är ett *tangentvektorfält* (eller ett *vektorfält*) på U. Vi säger också (något oegentligt) att v^a är ett (tangent-)vektorfält på U.

Minns att om $\mathbf{v} = v^a \frac{\partial}{\partial x^a}$ är en vektor i $p \in M$ så ges en (snäll) funktion f:s riktningsderivata m a p $\mathbf{v}$ av (se ekvation (4.3))

$$\mathbf{v}(f) = v^a \left. \frac{\partial f}{\partial x^a} \right|_p .$$

Om $\mathbf{v} = v^a \frac{\partial}{\partial x^a}$ istället är ett vektorfält kan vi fortfarande bilda riktningsderivatan

$$\mathbf{v}(f) = v^a \frac{\partial f}{\partial x^a},$$

som nu inte blir ett tal utan en ny $\mathcal{C}^\infty$ funktion av p.

Ovanstående definition av vektorfält är dock inte tillräckligt allmän för våra syften. Vi har nämligen förutsatt att $\mathbf{v}$:s definitionsmängd gått att täcka med ett enda koordinatsystem. Vi ska därför generalisera ovanstående definition till vektorfält definierade på hela M. Vektorfält definierade på mer allmänna delmängder till M kan sedan definieras på liknande sätt.

Betrakta därför en funktion som till varje punkt $p \in M$ ordnar en tangentvektor $\mathbf{v}(p) = v^a(p) \left. \frac{\partial}{\partial x^a} \right|_p \in T_pM$. För varje koordinatsystem $\psi : U \to D$ i M:s atlas betraktar vi sedan $\mathbf{v}$:s restriktion till U och vi säger att $\mathbf{v} \in \mathcal{C}^\infty(M)$ om $\mathbf{v} \in \mathcal{C}^\infty(U)$ för varje koordinatsystem i M:s atlas och vi säger att $\mathbf{v}$ (eller, något oegentligt, v^a) är ett *(tangent-)vektorfält* på M om $\mathbf{v} \in \mathcal{C}^\infty(M)$.

Om $p \mapsto T^{a_1 \ldots a_s}{}_{b_1 \ldots b_t}$ där $T^{a_1 \ldots a_s}{}_{b_1 \ldots b_t}$ är en tensor i p säger vi (analogt med ovanstående) att $T^{a_1 \ldots a_s}{}_{b_1 \ldots b_t} \in \mathcal{C}^\infty(U)$ om $\tilde{T}^{a_1 \ldots a_s}{}_{b_1 \ldots b_t} = T^{a_1 \ldots a_s}{}_{b_1 \ldots b_t} \circ \psi^{-1}$ (som är en uppsättning av n^{s+t} st 'vanliga' reellvärda funktioner av n variabler) är av klass $\mathcal{C}^\infty(U)$ och vi säger att $T^{a_1 \ldots a_s}{}_{b_1 \ldots b_t}$ är ett *tensorfält* på U.

I litteraturen förekommer tensorfält ofta medan tensorer definierade i en enstaka punkt förekommer ganska sällan. Det har därför blivit vanligt att kort och gott skriva 'tensor' eller 'vektor' när det egentligen är ett tensor- resp. vektorfält som avses. Detta leder ganska sällan till missförstånd, men viss uppmärksamhet och försiktighet från läsarens sida rekommenderas ändå.

Exempel 4.13. Vi såg ovan att partiell derivering av en funktion f gav en kovariant tensor $\frac{\partial f}{\partial x^a}$. En naturlig fråga är därför: Är en partiell derivata av ett tensorfält också ett tensorfält?

Svaret på denna fråga är (tyvärr) i allmänhet nej. För att inse detta studerar vi den partiella derivatan $\frac{\partial v^a}{\partial x^b}$ av ett vektorfält v^a. Om vi byter till nya koordinater $\bar{x}^a$ transformeras alltså v^a enligt ekvation (4.6). Vi undersöker nu hur $\frac{\partial v^a}{\partial x^b}$ transformeras vid detta koordinatbyte:

$$\frac{\partial \bar{v}^a}{\partial \bar{x}^b} = \frac{\partial x^d}{\partial \bar{x}^b}\frac{\partial}{\partial x^d}\left(v^c\frac{\partial \bar{x}^a}{\partial x^c}\right) = \frac{\partial x^d}{\partial \bar{x}^b}\frac{\partial \bar{x}^a}{\partial x^c}\frac{\partial v^c}{\partial x^d} + \frac{\partial x^d}{\partial \bar{x}^b}\frac{\partial^2 \bar{x}^a}{\partial x^d \partial x^c}v^c,$$

där vi använt kedjeregeln i första likheten.

Vi ser att första termen överensstämmer med transformationslagen (4.12) om $s = t = 1$, d v s med transformationslagen för en blandad tensor av typ $(1, 1)$ vilket, p g a indexens placering, är den typ vi förväntat oss att $\frac{\partial v^a}{\partial x^b}$ skulle ha haft om den hade varit en tensor.

Det är dock lätt att se att andra termen är nollskild i allmänhet (det är en god övning att själv konstruera ett explicit motexempel!) vilket innebär att $\frac{\partial v^a}{\partial x^b}$ *ej* är en tensor.

Partiell derivering är alltså ingen tensoroperation. Vi behöver alltså definiera derivator av tensorfält på något annat sätt. Hur detta kan göras ska vi diskutera i kapitel 7. □

Om $\mathbf{Y} = Y^a\frac{\partial}{\partial x^a}$ är ett vektorfält som verkar på en (snäll) funktion f fås

$$\mathbf{Y}(f) = Y^a\frac{\partial f}{\partial x^a},$$

som är en ny (snäll) funktion av p. Här kan vi förstås låta $\mathbf{Y}$, eller något annat vektorfält, verka på $\mathbf{Y}(f)$ och därmed bilda riktningsderivator av högre ordning. Av speciellt intresse är följande uttryck, till synes en riktningsderivata av andra ordningen.

Definition 4.14. *Kommutatorn* $[\mathbf{Y}, \mathbf{Z}]$ av vektorfälten $\mathbf{Y}$ och $\mathbf{Z}$ definieras av

$$[\mathbf{Y}, \mathbf{Z}](f) = \mathbf{Y}(\mathbf{Z}(f)) - \mathbf{Z}(\mathbf{Y}(f)).$$

En viktig anledning till att just detta uttryck är av intresse är att det inte alls ger en riktningsderivata av andra ordningen vilket man kanske kunde tro. Uttrycket definierar istället en rektningsderivata av *första* ordningen, d v s ett vektorfält.

SATS 4.15. *Om* $\mathbf{Y}$*,* $\mathbf{Z}$ *är vektorfält är* $[\mathbf{Y}, \mathbf{Z}]$ *ett vektorfält, d v s* $[\mathbf{Y}, \mathbf{Z}] = [\mathbf{Y}, \mathbf{Z}]^a\frac{\partial}{\partial x^a}$ *för några funktioner* $[\mathbf{Y}, \mathbf{Z}]^a$*. Vidare ges dessa av formeln*

$$[\mathbf{Y}, \mathbf{Z}]^a = Y^b\frac{\partial Z^a}{\partial x^b} - Z^b\frac{\partial Y^a}{\partial x^b}.$$

Bevis. Låt f vara en godtycklig (snäll) funktion och betrakta

$$\begin{aligned}[\mathbf{Y}, \mathbf{Z}](f) &= Y^b\frac{\partial}{\partial x^b}\left(Z^a\frac{\partial f}{\partial x^a}\right) - Z^b\frac{\partial}{\partial x^b}\left(Y^a\frac{\partial f}{\partial x^a}\right)\\ &= Y^b\frac{\partial Z^a}{\partial x^b}\frac{\partial f}{\partial x^a} + Y^bZ^a\frac{\partial^2 f}{\partial x^b\partial x^a} - Z^b\frac{\partial Y^a}{\partial x^b}\frac{\partial f}{\partial x^a} - Z^bY^a\frac{\partial^2 f}{\partial x^b\partial x^a}\\ &= \left(Y^b\frac{\partial Z^a}{\partial x^b} - Z^b\frac{\partial Y^a}{\partial x^b}\right)\frac{\partial f}{\partial x^a} + Y^aZ^b\left(\frac{\partial^2 f}{\partial x^a\partial x^b} - \frac{\partial^2 f}{\partial x^b\partial x^a}\right) = \left(Y^b\frac{\partial Z^a}{\partial x^b} - Z^b\frac{\partial Y^a}{\partial x^b}\right)\frac{\partial f}{\partial x^a}\end{aligned}$$

där vi döpt om några summationsindex i tredje likheten. Då f var godtycklig följer att

$$[\mathbf{Y},\mathbf{Z}] = \left(Y^b\frac{\partial Z^a}{\partial x^b} - Z^b\frac{\partial Y^a}{\partial x^b}\right)\frac{\partial}{\partial x^a}.$$

Vi ser att $[\mathbf{Y},\mathbf{Z}]$ är en linjärkombination av koordinatbasvektorerna $\frac{\partial}{\partial x^a}$ vilket visar att $[\mathbf{Y},\mathbf{Z}]$ är ett vektorfält. Direkt avläsning ger sedan att

$$[\mathbf{Y},\mathbf{Z}]^a = Y^b\frac{\partial Z^a}{\partial x^b} - Z^b\frac{\partial Y^a}{\partial x^b}.$$

$\square$

Kapitel 5

Tensoralgebra

5.1 Inledning

I detta kapitel ska vi se på en del algebraiska operationer som kan utföras på tensorer. Vi ska först lägga fast lite notation. I hela detta kapitel låter vi M vara en n-dimensionell mångfald och p får vara en godtycklig punkt på M. Vi antar också att x^a och $\bar{x}^a$ är ett ursprungligt respektive ett nytt koordinatsystem i en omgivning av p.

5.2 Tensoroperationer

Låt $T_a{}^b$, $U_a{}^b$, V_{ab} och $W_a{}^{bc}$ vara tensorer i p (d v s uppsättningar av tal som transformeras enligt formlerna i föregående avsnitt vid koordinatbyten). Låt oss fundera över vilka räkneoperationer vi skulle kunna definiera med hjälp av dessa.

Vi kan t ex bilda $T_a{}^b + U_a{}^b$ genom vanlig addition av tal för varje uppsättning värden på a och b. Eftersom $T_a{}^b$ och $U_a{}^b$ uppfyller samma transformationslag verkar det troligt att även summan gör det. Summan är i så fall en ny tensor av samma typ som de båda termerna (vi ska snart bevisa att denna intuition är korrekt).

Vi kan förstås också formellt bilda summan $T_a{}^b + V_{ab}$, men eftersom $T_a{}^b$ och V_{ab} uppfyller helt olika transformationslagar verkar det inte troligt att summan är en tensor (så är heller inte fallet i allmänhet, vilket är lätt att visa). Samma resonemang kan göras för summan $T_a{}^b + W_a{}^{bc}$.

Det verkar alltså som om tensorer kan adderas genom vanlig komponentvis addition av tal förutsatt att tensorerna är av samma typ.

Vad händer då om vi försöker multiplicera två tensorer? Det står i alla fall klart att vi måste iaktta viss försiktighet. Att skriva $T_a{}^b U_a{}^b$ ger ett odefinierat uttryck eftersom summationskonventionen förutsätter att upprepade index förekommer en gång uppe och en gång nere. Om vi byter ut indexen kan vi dock bilda produkten $T_a{}^b U_c{}^d$ vilket i alla fall ger ett väldefinierat uttryck. Vi kan också bilda produkter av tensorer av olika typ t ex $T_a{}^b V_{cd}$, $V_{ab} W_c{}^{de}$ eller varför inte $T_d{}^g V_{ca} W_e{}^{fb}$.

Man kan visa att alla dessa uttryck är tensorer. Låt oss därför formulera följande sats:

SATS 5.1. *Låt* $S^{a_1\dots a_s}{}_{b_1\dots b_t}$*,* $T^{a_1\dots a_s}{}_{b_1\dots b_t}$ *vara tensorer av typ* (s,t) *och låt* $U^{c_1\dots c_q}{}_{d_1\dots d_r}$ *vara en tensor av typ* (q,r)*. Låt dessutom* $a \in \mathbf{R}$*. Då gäller*

(i) $S^{a_1\dots a_s}{}_{b_1\dots b_t} + T^{a_1\dots a_s}{}_{b_1\dots b_t}$ *är en tensor av typ* (s,t)*.*

(ii) $aT^{a_1\dots a_s}{}_{b_1\dots b_t}$ *är en tensor av typ* (s,t)*.*

(iii) $S^{a_1\dots a_s}{}_{b_1\dots b_t}U^{c_1\dots c_q}{}_{d_1\dots d_r}$ *är en tensor av typ* $(s+q, t+r)$*.*

(iv) Om $S^{a_1\dots a_s}{}_{b_1\dots b_t} = T^{a_1\dots a_s}{}_{b_1\dots b_t}$ *i något koordinatsystem* x^e *så är* $\bar{S}^{a_1\dots a_s}{}_{b_1\dots b_t} = \bar{T}^{a_1\dots a_s}{}_{b_1\dots b_t}$ *i ett godtyckligt koordinatsystem* $\bar{x}^e$*.*

Anmärkning: Den sista punkten är viktig då den säger att om två tensorer är lika i ett koordinatsystem så är de lika i *alla* koordinatsystem. Tensorekvationer är därför ett lämpligt verktyg för att beskriva naturlagarna, eftersom vi inte förväntar oss att naturlagarna beror på vilket koordinatsystem vi använder.

□

Bevis. Sätt $V^{a_1\dots a_s}{}_{b_1\dots b_t} = S^{a_1\dots a_s}{}_{b_1\dots b_t} + T^{a_1\dots a_s}{}_{b_1\dots b_t}$. Vi vill visa att $V^{a_1\dots a_s}{}_{b_1\dots b_t}$ är en tensor, d v s att om $\bar{V}^{a_1\dots a_s}{}_{b_1\dots b_t} := \bar{S}^{a_1\dots a_s}{}_{b_1\dots b_t} + \bar{T}^{a_1\dots a_s}{}_{b_1\dots b_t}$, där $\bar{S}^{a_1\dots a_s}{}_{b_1\dots b_t}$ och $\bar{T}^{a_1\dots a_s}{}_{b_1\dots b_t}$ ges av transformationslagen (4.12) så är $\bar{V}^{a_1\dots a_s}{}_{b_1\dots b_t} = \frac{\partial \bar{x}^{a_1}}{\partial x^{c_1}} \cdots \frac{\partial \bar{x}^{a_s}}{\partial x^{c_s}} \frac{\partial x^{d_1}}{\partial \bar{x}^{b_1}} \cdots \frac{\partial x^{d_t}}{\partial \bar{x}^{b_t}} V^{c_1\dots c_s}{}_{d_1\dots d_t}$.

Men detta följer direkt:

$$
\begin{aligned}
\bar{V}^{a_1\dots a_s}{}_{b_1\dots b_t} &= \bar{S}^{a_1\dots a_s}{}_{b_1\dots b_t} + \bar{T}^{a_1\dots a_s}{}_{b_1\dots b_t} \\
&= \frac{\partial \bar{x}^{a_1}}{\partial x^{c_1}} \cdots \frac{\partial \bar{x}^{a_s}}{\partial x^{c_s}} \frac{\partial x^{d_1}}{\partial \bar{x}^{b_1}} \cdots \frac{\partial x^{d_t}}{\partial \bar{x}^{b_t}} S^{c_1\dots c_s}{}_{d_1\dots d_t} + \frac{\partial \bar{x}^{a_1}}{\partial x^{c_1}} \cdots \frac{\partial \bar{x}^{a_s}}{\partial x^{c_s}} \frac{\partial x^{d_1}}{\partial \bar{x}^{b_1}} \cdots \frac{\partial x^{d_t}}{\partial \bar{x}^{b_t}} T^{c_1\dots c_s}{}_{d_1\dots d_t} \\
&= \frac{\partial \bar{x}^{a_1}}{\partial x^{c_1}} \cdots \frac{\partial \bar{x}^{a_s}}{\partial x^{c_s}} \frac{\partial x^{d_1}}{\partial \bar{x}^{b_1}} \cdots \frac{\partial x^{d_t}}{\partial \bar{x}^{b_t}} (S^{a_1\dots a_s}{}_{b_1\dots b_t} + T^{a_1\dots a_s}{}_{b_1\dots b_t}) \\
&= \frac{\partial \bar{x}^{a_1}}{\partial x^{c_1}} \cdots \frac{\partial \bar{x}^{a_s}}{\partial x^{c_s}} \frac{\partial x^{d_1}}{\partial \bar{x}^{b_1}} \cdots \frac{\partial x^{d_t}}{\partial \bar{x}^{b_t}} V^{c_1\dots c_s}{}_{d_1\dots d_t},
\end{aligned}
$$

vilket visar första delen. *(ii)*, *(iii)* och *(iv)* visas på liknande sätt (övning). □

I ovanstående sats är det inget problem att byta ut alla tensorer i en punkt mot tensorfält (så t ex är summan av två tensorfält av samma typ ett nytt tensorfält). Alla bevis går igenom oförändrade. Vi kan t o m ersätta talet $a \in \mathbf{R}$ med en godtycklig $\mathcal{C}^\infty$ funktion f.

Ett annat sätt att bilda en ny tensor ur en given är att starta med en blandad tensor, t ex $T_{ab}{}^{cd}$ och sedan byta ut ett av indexen uppe mot samma bokstav som ett av indexen nere och få (t ex) $U_a{}^d = T_{ab}{}^{bd}$. Vi har alltså plockat ut de komponenter av tensorn för vilka $b = c$ och summerat dessa. $U_a{}^d$ sägs då vara en *kontraktion* av tensorn $T_{ab}{}^{cd}$ och man visar på liknande sätt som ovan att $U_a{}^d$ är en tensor om $T_{ab}{}^{cd}$ är det. En allmän kontraktion definieras på samma sätt, *mutatis mutandis*. För fullständighets skull ger vi den allmänna definitionen:

Definition 5.2. Låt $T^{a_1\dots a_s}{}_{b_1\dots b_t}$ vara en blandad tensor av typ (s,t), $s,t \neq 0$ och låt i,j vara heltal sådana att $1 \leq i \leq s$, $1 \leq j \leq t$. *Kontraktionen* av T med avseende på det i:te kontravarianta och det j:te kovarianta indexet är då

$$U^{a_1\dots a_{i-1}a_{i+1}\dots a_s}{}_{b_1\dots b_{j-1}b_{j+1}\dots b_t} = T^{a_1\dots a_{i-1}ca_{i+1}\dots a_s}{}_{b_1\dots b_{j-1}cb_{j+1}\dots b_t}.$$

SATS 5.3. *En kontraktion av en tensor är en tensor.*

Bevis. Övning. □

Anmärkning: Observera att uttrycket $F^a{}_b v^b$ som förekommer i Exempel 4.12 kan ses som en kontraktion av tensorn $F^a{}_b v^c$. Observera också att kontraktionen av tensorn $F^a{}_b$ (som enligt Exempel 4.12 kan tolkas som elementen i avbildningsmatrisen F till en linjär avbildning $\mathbf{F}$ i den aktuella basen) blir

$$F^a{}_a = \text{Summan av huvuddiagonalelementen i } F = \mathrm{Tr}\,(F) = \text{Spåret av } F.$$

Spåret av en matris dyker upp i många tillämpningar och det är alltså inget annat än en kontraktion av motsvarande tensor. □

Definition 5.4. Låt T_{ab} vara en tensor. Om $T_{ab} = T_{ba}$ sägs T_{ab} vara *symmetrisk* och om $T_{ab} = -T_{ba}$ sägs T_{ab} vara *antisymmetrisk* (eller *skevsymmetrisk*).

En tensor är förstås varken symmetrisk eller antisymmetrisk i allmänhet. Däremot kan en godtycklig tensor av typen $(0,2)$ (eller $(2,0)$, resonemanget är detsamma) skrivas som en summa av en symmetrisk och en antisymmetrisk tensor. För att se det gör vi följande definition:

Definition 5.5. Den symmetriska tensorn

$$T_{(ab)} = \frac{1}{2}\,(T_{ab} + T_{ba})$$

kallas den *symmetriska delen* av T_{ab} och den antisymmetriska tensorn

$$T_{[ab]} = \frac{1}{2}\,(T_{ab} - T_{ba})$$

kallas den *antisymmetriska delen* av T_{ab}.

Man ser nu lätt att

$$T_{ab} = T_{(ab)} + T_{[ab]}.$$

Dessutom är denna uppdelning entydig:

SATS 5.6. *Om $T_{ab} = S_{ab} + A_{ab}$ där S_{ab} är symmetrisk och A_{ab} är antisymmetrisk så är $S_{ab} = T_{(ab)}$ och $A_{ab} = T_{[ab]}$.*

Bevis. Enligt ovan är $S_{ab} + A_{ab} = T_{(ab)} + T_{[ab]} \Leftrightarrow S_{ab} - T_{(ab)} = T_{[ab]} - A_{ab} =: W_{ab}$. Detta visar att W_{ab} är både symmetrisk och antisymmetrisk, varför $W_{ab} = W_{ba} = -W_{ab}$ (där vi först använt symmetrin och sedan antisymmetrin), d v s $W_{ab} = 0$. Det följer att $S_{ab} = T_{(ab)}$ och $A_{ab} = T_{[ab]}$. $\square$

Symmetriska och antisymmetriska delar kan också generaliseras till tensorer med godtyckligt många index. Antag att $T_{a_1 \ldots a_t}$ är en tensor av typ $(0, t)$. Låt vidare π vara en permutation av talen $1, \ldots, t$, d v s $\pi(1), \pi(2), \ldots, \pi(q)$ är en uppräkning av heltalen från 1 till t i någon ordningsföljd. Vi definierar då symmetriska delen av $T_{a_1 \ldots a_t}$ som

$$T_{(a_1 \ldots a_t)} = \frac{1}{t!} \sum_{\text{alla } \pi} T_{a_{\pi(1)} \ldots a_{\pi(t)}}.$$

T ex blir då

$$T_{(abc)} = \frac{1}{3!} \left(T_{abc} + T_{acb} + T_{bca} + T_{bac} + T_{cab} + T_{cba} \right).$$

och vi ser att $T_{(abc)}$ är symmetrisk vid omkastning av vilket indexpar som helst.

Givet en permutation π, låt $q(\pi)$ vara antalet transpositioner (d v s platsbyten) som går åt för att bilda π från den ursprungliga ordningsföljden $1, 2, \ldots t$. Vi säger att π är *udda* om $q(\pi)$ är det och vi säger att π är *jämn* om $q(\pi)$ är det.

T ex är permutationen $\pi(1) = 2$, $\pi(2) = 3$, $\pi(3) = 1$ en jämn permutation av talen $1, 2, 3$ eftersom den kan bildas genom transpositionerna $1, 2, 3 \mapsto 2, 1, 3 \mapsto 2, 3, 1$.

Vi definierar nu antisymmetriska delen av $T_{a_1 \ldots a_t}$ genom (jfr definitionen av en determinant)

$$T_{[a_1 \ldots a_t]} = \frac{1}{t!} \sum_{\text{alla } \pi} (-1)^{q(\pi)} T_{a_{\pi(1)} \ldots a_{\pi(t)}},$$

så t ex är

$$T_{[abc]} = \frac{1}{3!} \left(T_{abc} - T_{acb} + T_{bca} - T_{bac} + T_{cab} - T_{cba} \right)$$

och vi ser att $T_{[abc]}$ är antisymmetrisk vid omkastning av vilket indexpar som helst.

Att en tensor är summan av sin symmetriska och antisymmetriska del gäller dock bara för tensorer med 2 index. Vi ser ju t ex direkt ur ovanstående formler för tensorer med 3 index att

$$T_{(abc)} + T_{[abc]} = \frac{1}{3} \left(T_{abc} + T_{bca} + T_{cab} \right) \neq T_{abc}.$$

Vi kan också symmetrisera (eller antisymmetrisera) över en delmängd av en tensors index. T ex är

$$T_{a(bc)d} = \frac{1}{2} \left(T_{abcd} + T_{acbd} \right).$$

Om vi behöver symmetrisera (eller antisymmetrisera) över index som inte står intill varandra skrivs de index som ska uteslutas ur (anti)symmetriseringen mellan lodräta streck, sålunda:

$$T_{(a|b|c)} = \frac{1}{2} \left(T_{abc} + T_{cba} \right),$$

där alltså symmetriseringen sker över indexparet ac men inte över b.

5.3 Enhetssfären, del 3

Nästa punkt på programmet är att definiera en skalärprodukt, också kallad en *metrik*, på T_pM, och som vanligt får inte definitionen referera till något omkringliggande $\mathbf{R}^n$. Vi ska först titta på hur den vanliga euklidiska skalärprodukten i $\mathbf{R}^3$ ger oss en skalärprodukt på T_pS där S är enhetssfären. Därefter ska vi ge en alternativ beskrivning av denna, som enbart refererar till S själv. I hela detta avsnitt kommer vi att använda samma beteckningar som i avsnitten 3.3 och 4.2.

Låt $p \in S$. Då är $p \in U$ eller $p \in \bar{U}$ och utan att förlora i allmängiltighet kan vi (som tidigare) anta att $p \in U$ och att p har koordinater (θ_0, φ_0) i koordinatsystemet ψ. Om $\mathbf{u}, \mathbf{v} \in T_pS$ så är enligt tidigare $\mathbf{u} = u^0\mathbf{e}_\theta + u^1\mathbf{e}_\varphi$ och $\mathbf{v} = v^0\mathbf{e}_\theta + v^1\mathbf{e}_\varphi$ för reella tal u^0, u^1, v^0, v^1. Skalärprodukten mellan $\mathbf{u}$ och $\mathbf{v}$ ges då av

$$\mathbf{u} \cdot \mathbf{v} = \left(u^0\mathbf{e}_\theta + u^1\mathbf{e}_\varphi\right) \cdot \left(v^0\mathbf{e}_\theta + v^1\mathbf{e}_\varphi\right) = u^0v^0\mathbf{e}_\theta \cdot \mathbf{e}_\theta + u^0v^1\mathbf{e}_\theta \cdot \mathbf{e}_\varphi + u^1v^0\mathbf{e}_\varphi \cdot \mathbf{e}_\theta + u^1v^1\mathbf{e}_\varphi \cdot \mathbf{e}_\varphi.$$

Vi behöver alltså räkna ut skalärprodukterna $g_{00} := \mathbf{e}_\theta \cdot \mathbf{e}_\theta$, $g_{01} := \mathbf{e}_\theta \cdot \mathbf{e}_\varphi$, $g_{10} := \mathbf{e}_\varphi \cdot \mathbf{e}_\theta$ och $g_{11} := \mathbf{e}_\varphi \cdot \mathbf{e}_\varphi$. Vi får t ex

$$g_{11} = \mathbf{e}_\varphi \cdot \mathbf{e}_\varphi = (-\sin\theta_0 \sin\varphi_0\, \bar{e}_x + \sin\theta_0 \cos\varphi_0\, \bar{e}_y) \cdot (-\sin\theta_0 \sin\varphi_0\, \bar{e}_x + \sin\theta_0 \cos\varphi_0\, \bar{e}_y) = \sin^2\theta_0,$$

och vi verifierar lika lätt att $g_{00} = 1$ och $g_{01} = g_{10} = 0$. Detta ger att

$$\mathbf{u} \cdot \mathbf{v} = g_{00}u^0v^0 + g_{01}u^0v^1 + g_{10}u^1v^0 + g_{11}u^1v^1 = g_{ab}u^av^b,$$

där $g_{00} = 1$, $g_{01} = g_{10} = 0$ och $g_{11} = \sin^2\theta_0$ och vi observerar att denna formel *inte* refererar till något omkringliggande $\mathbf{R}^3$ (även om vi använde oss av ett omkringliggande $\mathbf{R}^3$ för att härleda den).

Om inte bara $p \in U$ utan också $p \in \bar{U}$ och har koordinater $(\bar{\theta}_0, \bar{\varphi}_0)$ i koordinatsystemet $\bar{\psi}$ och vi upprepar samma konstruktion fås på samma sätt

$$\mathbf{u} \cdot \mathbf{v} = \bar{g}_{ab}\bar{u}^a\bar{v}^b,$$

där $\bar{g}_{00} = 1$, $\bar{g}_{01} = \bar{g}_{10} = 0$ och $\bar{g}_{11} = \sin^2\bar{\theta}_0$ och $\bar{u}^a$ och $\bar{v}^a$ är $\mathbf{u}$:s resp. $\mathbf{v}$:s komponenter i kordinatbasen till $\bar{\psi}$. Enligt transformationslagarna gäller då

$$g_{ab}u^av^b = \mathbf{u} \cdot \mathbf{v} = \bar{g}_{cd}\bar{u}^c\bar{v}^d = \bar{g}_{cd}u^a\frac{\partial\bar{x}^c}{\partial x^a}v^b\frac{\partial\bar{x}^d}{\partial x^b},$$

för alla u^a och v^b. Låt t ex $u^0 = v^0 = 1$ och $u^1 = v^1 = 0$ så fås $g_{00} = \bar{g}_{cd}\frac{\partial\bar{x}^c}{\partial x^0}\frac{\partial\bar{x}^d}{\partial x^0}$. På samma sätt visas att

$$g_{ab} = \bar{g}_{cd}\frac{\partial\bar{x}^c}{\partial x^a}\frac{\partial\bar{x}^d}{\partial x^b}$$

för alla värden på a och b. Kontrahera båda leden med $\frac{\partial x^a}{\partial\bar{x}^e}\frac{\partial x^b}{\partial\bar{x}^f}$ så fås

$$g_{ab}\frac{\partial x^a}{\partial\bar{x}^e}\frac{\partial x^b}{\partial\bar{x}^f} = \bar{g}_{cd}\frac{\partial\bar{x}^c}{\partial x^a}\frac{\partial\bar{x}^d}{\partial x^b}\frac{\partial x^a}{\partial\bar{x}^e}\frac{\partial x^b}{\partial\bar{x}^f} = /\text{kedjeregeln}/ = \bar{g}_{cd}\frac{\partial\bar{x}^c}{\partial\bar{x}^e}\frac{\partial\bar{x}^d}{\partial\bar{x}^f} = \bar{g}_{cd}\delta_e{}^c\delta_f{}^d = \bar{g}_{ef}.$$

Detta är transformationslagen för en tensor av typ $(0,2)$ så g_{ab} definierad på detta sätt är alltså en tensor av typ $(0,2)$.

5.4 Metriska tensorn

Inspirerade av förra avsnittet gör vi följande definition:

Definition 5.7. Ett symmetriskt tensorfält $g_{ab} = g_{ba}$ är en *metrik* om g_{ab} är icke-singulär i varje punkt, d v s om $g_{ab}u^a v^b = 0$ för alla v^b om och endast om $u^a = 0$. *Skalärprodukten* mellan u^a och v^a definieras $g_{ab}u^a v^b$ och vi säger att u^a och v^a är *ortogonala* om $g_{ab}u^a v^b = 0$.

Denna definition skiljer sig lite från definitionen av skalärprodukt som används i kursen i linjär algebra. Här kräver vi *inte* att skalärprodukten är positivt definit, d v s att $g_{ab}v^a v^b > 0$ för alla $v^a \neq 0$. Det kan alltså hända att $g_{ab}v^a v^b < 0$ eller att $g_{ab}v^a v^b = 0$ trots att $v^a \neq 0$.

Om metriken g_{ab} skulle vara positivt definit sägs g_{ab} vara en *Riemannsk* metrik.

Exempel 5.8. Då den vanliga euklidiska skalärprodukten på $\mathbf{R}^n$ ges av

$$g_{ab}u^a v^b = u^0 v^0 + u^1 v^1 + \ldots + u^{n-1} v^{n-1}$$

ser vi att $g_{ab} = 1$ om $a = b$ och $g_{ab} = 0$ om $a \neq b$. Ordnar vi g_{ab}:s komponenter i en matris får vi alltså en diagonalmatris med ettor i huvuddiagonalen och nollor för övrigt, d v s enhetsmatrisen. Vi skriver $g_{ab} = \mathrm{diag}(1, 1, \ldots, 1)$.

Då speciellt

$$g_{ab}u^a u^b = (u^0)^2 + (u^1)^2 + \ldots + (u^{n-1})^2 > 0$$

om något $u^a \neq 0$ följer att g_{ab} är positivt definit och därmed en Riemannsk metrik på $\mathbf{R}^n$. □

Exempel 5.9. Metriken g_{ab} given av $g_{00} = 1$, $g_{01} = g_{10} = 0$, $g_{11} = \sin^2\theta$ på delmängden U av enhetssfären (se avsnitt 5.3) är en Riemannsk metrik då denna metrik ger samma skalärprodukt mellan två vektorer $\mathbf{u}$ och $\mathbf{v} \in T_pM$, $p \in U$ som den vanliga euklidiska skalärprodukten på $\mathbf{R}^3$, och denna är Riemannsk enligt föregående exempel. □

Exempel 5.10. Metriken $\eta_{ab} = \mathrm{diag}(1, -1, -1, -1)$ på $\mathbf{R}^4$ kallas *Minkowskimetriken*. Skalärprodukten mellan två godtyckliga vektorer u^a och v^a ges då av

$$\eta_{ab}u^a v^b = u^0 v^0 - u^1 v^1 - u^2 v^2 - u^3 v^3.$$

Om $u^a = (1, 0, 0, 0)$ och $v^a = (0, 1, 0, 0)$ är alltså $\eta_{ab}u^a u^b = 1$ men $\eta_{ab}v^a v^b = -1$. η_{ab} är alltså indefinit och därmed *ej* Riemannsk. Dessutom, om $w^a = (1, 1, 0, 0)$ är $g_{ab}w^a w^b = 0$ trots att $w^a \neq 0$. Minkowskimetriken spelar en central roll i både speciell och allmän relativitetsteori. □

Vi kommer att använda följande terminologi:

Definition 5.11. En vektor v^a är *tidsartad* om $g_{ab}v^a v^b > 0$, *rumsartad* om $g_{ab}v^a v^b < 0$ och *ljusartad* eller en *nollvektor*[1] om $g_{ab}v^a v^b = 0$.

[1] Observera att ordet 'nollvektor' olyckligtvis får dubbla betydelser. Dels kan det (som här) betyda en vektor v^a sådan att $g_{ab}v^a v^b = 0$ och dels kan det betyda den vektor vars samtliga komponenter $v^a = 0$ vilket det inte är frågan om här. På engelska undviks dubbeltydigheten genom att de förstnämnda vektorerna kallas 'null vectors' och den sistnämnda kallas 'the zero vector', men på svenska finns tyvärr inte motsvarande möjlighet.

Observera att denna karakterisering gäller punktvis. Ett vektorfält kan mycket väl vara rumsartat i vissa punkter och tidsartat i vissa.

Om g_{ab} är en metrik är alltså g_{ab} icke-singulär, d v s $g_{ab}u^a v^b = 0$ för alla v^b om och endast om $u^a = 0$. Genom att sätta in v^b bestående av en etta och resten nollor ser vi att detta är ekvivalent med att $g_{ab}u^a = 0$ om och endast om $u^a = 0$. Det linjära ekvationssystemet $g_{ab}u^a = 0$ har alltså entydig lösning. Om vi låter G betyda den matris som har elementet g_{ab} på plats (a, b) är alltså $\det(G) \neq 0$ varför G är inverterbar.

Låt g^{ab} vara elementet på plats (a, b) i G^{-1}. Då är $g_{ab}g^{bc}$ elementet på plats (a, c) i $GG^{-1} = I$, d v s $g_{ab}g^{bc} = \delta_a{}^c$. Av denna anledning kallas g^{ab} den *inversa metriken*.

Övning: Visa att g^{ab} är ett tensorfält om g_{ab} är det.

Observera att metriken g_{ab} kan ses som en linjär avbildning som avbildar en tangentvektor (kontravariant tensor) på en kovariant tensor genom $v^a \mapsto g_{ab}v^b$. Denna avbildning är dessutom inverterbar då $g^{ac}(g_{cb}v^b) = \delta_b{}^a v^b = v^a$.

Metriken och dess invers ger oss alltså en 1-1-motsvarighet mellan tangentvektorer och kovarianta tensorer. För att göra notationen mer ekonomisk kommer vi i fortsättningen att beteckna den kovarianta tensorn $g_{ab}v^b$ med samma stambokstav som tangentvektorn v^a, fast med indexet sänkt.

Definition 5.12. Givet en tangentvektor v^a sätter vi

$$v_a = g_{ab}v^b \qquad \text{(sänkning av index).}$$

På samma sätt kan den inversa metriken ses som en inverterbar linjär avbildning från kovarianta tensorer till tangentvektorer genom $u_a \mapsto g^{ab}u_b$. Vi definierar därför höjning av index på liknande sätt:

Definition 5.13. Givet en godtycklig kovariant tensor u_a sätter vi

$$u^a = g^{ab}u_b \qquad \text{(höjning av index).}$$

Dessa definitioner kan, på uppenbart sätt, generaliseras till allmänna tensorer. För en blandad tensor $T_a{}^b$ (t ex) gäller det att

$$T_{ab} = g_{bc}T_a{}^c \quad , \quad T^{ab} = g^{ac}T_c{}^b.$$

Här ser vi varför det är viktigt att inte skriva indexen i blandade tensorer rakt under varandra. Om vi skulle skriva T_a^b är det inte klart om vi menar $T_a{}^b$ eller $T^b{}_a$ och eftersom det gäller att

$$T^b{}_a = g^{bd}g_{ac}T_d{}^c \neq T_a{}^b,$$

skulle detta skrivsätt kunna leda till missförstånd.

Observera att denna notation blir konsistent i meningen att om vi först höjer ett index och sedan sänker det igen (eller tvärtom) så får vi tillbaka den tensor vi startade med (varför?).

I Definition 4.7 definierade vi differentialen df av en funktion f genom att kräva att df:s komponenter ges av $(df)_a = \nabla_a f = \frac{\partial f}{\partial x^a}$ och vi såg att dessa komponenter transformerades kovariant. Vi gör nu följande definition

Definition 5.14. *Gradienten*, $\nabla^a f$, till en funktion f definieras

$$\nabla^a f := g^{ab}\nabla_b f = g^{ab}\frac{\partial f}{\partial x^b}.$$

Anmärkning: Enligt kursen i flervariabelanalys har gradienten och differentialen till en funktion f samma komponenter, nämligen $\frac{\partial f}{\partial x^a}$. Anledningen till detta är förstås att hela flervariabelkursen utspelar sig i $\mathbf{R}^n$ försedd med den euklidiska metriken, så att g^{ab} = enhetsmatrisen. På en godtycklig mångfald har gradienten och differentialen till f olika komponenter i allmänhet. □

Definition 5.7 generaliserar begreppet ortogonalitet till allmänna mångfalder och i kursen i linjär algebra har vi sett att räkningar och resonemang ofta underlättas om vi arbetar i en bas i vilken basvektorerna är normerade och inbördes ortogonala (en s k ON-bas). Vi vill därför generalisera begreppet ON-bas till allmänna mångfalder, men först behöver vi ytterligare terminologi.

Låt p vara en punkt på M och låt $\{\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_{n-1}\}$ vara en bas för T_pM. Detta innebär att var och en av dessa basvektorer kan skrivas som en linjärkombination av vektorerna i koordinatbasen. Speciellt finns tal $(e_0)^a$ sådana att

$$\mathbf{e}_0 = (e_0)^a\frac{\partial}{\partial x^a},$$

och analogt för övriga basvektorer, d v s vi kan skriva

$$\mathbf{e}_i = (e_i)^a\frac{\partial}{\partial x^a}\ ,\ i = 0, 1, \ldots, n-1. \tag{5.1}$$

Som tidigare kommer vi att identifiera basvektorerna med sina komponenter och tala om vektorn $(e_i)^a$. Det är då viktigt att komma ihåg att indexet i talar om vilken av vektorerna i ovanstående bas vi menar och att indexet a pekar ut en speciell komponent i koordinatbasen, hos denna vektor.

Låt nu $M = \mathbf{R}^n$ och låt g_{ab} vara den vanliga euklidiska metriken. Enligt Exempel 5.8 är g_{ab} positivt definit. Vi kan då normera en vektor $v^a \neq 0$ genom att sätta

$$\hat{v}^a = \frac{1}{\sqrt{g_{cd}v^cv^d}}v^a$$

så att

$$g_{ab}\hat{v}^a\hat{v}^b = \frac{g_{ab}v^av^b}{\left(\sqrt{g_{cd}v^cv^d}\right)^2} = 1.$$

På en godtycklig mångfald M är g_{ab} inte nödvändigtvis positivt definit vilket innebär att endast vektorer v^a med $g_{ab}v^av^b > 0$ d v s tidsartade vektorer kan normeras på detta enkla sätt. Om v^a istället är rumsartad så att $g_{ab}v^av^b < 0$ kan vi istället sätta

$$\hat{v}^a = \frac{1}{\sqrt{-g_{cd}v^cv^d}}v^a$$

och få

$$g_{ab}\hat{v}^a\hat{v}^b = \frac{g_{ab}v^av^b}{\left(\sqrt{-g_{cd}v^cv^d}\right)^2} = -1.$$

En nollvektor v^a med $g_{ab}v^a v^b = 0$ går däremot inte att normera på något meningsfullt sätt.

Vi är nu klara att generalisera definitionen av en ON-bas till en godtycklig mångfald.

Definition 5.15. $\{(e_i)^a\,,\ i = 0, 1, \ldots, n-1\}$ är en ON-bas för T_pM det gäller att

$$g_{ab}(e_i)^a(e_j)^b = \begin{cases} \pm 1 & \text{om } i = j \\ 0 & \text{om } i \neq j \end{cases}\ ,$$

d v s basvektorerna är normerade och inbördes ortogonala.

Anmärkning: Villkoret att basvektorerna är normerade och inbördes ortogonala säkerställer att $\{(e_i)^a\,,\ i = 0, 1, \ldots, n-1\}$ verkligen är en bas för T_pM. Detta bevisas på samma sätt som motsvarande resultat i linjära algebran (hur?). □

En direkt konsekvens av denna definition är att en ON-bas inte innehåller några nollvektorer.

Det är kanske inte självklart att det alltid existerar en ON-bas för T_pM om M är en godtycklig mångfald med godtycklig (eventuellt indefinit) metrik, men faktum är att givet en godtycklig tidsartad eller rumsartad vektor v^a kan vi, även i fallet indefinit metrik, konstruera en ON-bas sådan att $(e_0)^a = \hat{v}^a$ med hjälp av Gram-Schmidts ortogonaliseringsprocess (se t ex Janfalk: Linjär algebra).

SATS 5.16. *Om $(e_i)^a$ och $(\bar{e}_i)^a$ är två ON-baser så innehåller de lika många tidsartade respektive rumsartade vektorer.*

Anmärkning: Detta resultat är ekvivalent med den s k *tröghetslagen*. Se Janfalk: Linjär algebra. □

Bevis. Antag motsatsen, d v s att p st av vektorerna $(e_i)^a$ och q st av vektorerna $(\bar{e}_i)^a$ är tidsartade, där $p \neq q$. Utan att förlora i allmängiltighet kan vi anta att $p < q$ och att vektorerna i båda baserna är ordnade så att de första p resp. q vektorerna i vardera basen är tidsartade.

Då finns en vektor

$$v^a = v^i(e_i)^a = v^0(e_0)^a + \ldots + v^{n-1}(e_{n-1})^a = \bar{v}^i(\bar{e}_i)^a = \bar{v}^0(\bar{e}_0)^a + \ldots + \bar{v}^{n-1}(\bar{e}_{n-1})^a$$

sådan att $v^a \neq 0$ men $v^0 = \ldots = v^{p-1} = 0$ och $\bar{v}^q = \ldots = \bar{v}^{n-1} = 0$ ty detta är $p + (n-q) = n - (q-p) < n$ st linjära och homogena villkor på en vektor i ett n-dimensionellt vektorrum; detta system är alltså homogent och underbestämt varför lösningarna bildar ett vektorrum med dimension minst 1. Då $(e_i)^a$ är en ON-bas fås

$$\begin{aligned} g_{ab}v^a v^b &= g_{ab}\left(v^0(e_0)^a + \ldots + v^{n-1}(e_{n-1})^a\right)\left(v^0(e_0)^b + \ldots + v^{n-1}(e_{n-1})^b\right) \\ &= (v^0)^2 + \ldots + (v^{p-1})^2 - (v^p)^2 - \ldots - (v^{n-1})^2 = -(v^p)^2 - \ldots - (v^n)^2 < 0 \end{aligned}$$

men å andra sidan är också $(\bar{e}_i)^a$ en ON-bas, vilket ger

$$\begin{aligned} g_{ab}v^a v^b &= g_{ab}\left(\bar{v}^0(\bar{e}_0)^a + \ldots + \bar{v}^{n-1}(\bar{e}_{n-1})^a\right)\left(\bar{v}^0(\bar{e}_0)^b + \ldots + \bar{v}^{n-1}(\bar{e}_{n-1})^b\right) \\ &= (\bar{v}^0)^2 + \ldots + (\bar{v}^{q-1})^2 - (\bar{v}^q)^2 - \ldots - (\bar{v}^{n-1})^2 = (\bar{v}^0)^2 + \ldots + (\bar{v}^{q-1})^2 > 0 \end{aligned}$$

vilket är en motsägelse. □

Antalet tidsartade och rumsartade vektorer som ingår i en ON-bas är alltså karakteristiskt för metriken och vi gör därför följande definition:

Definition 5.17. Antalet tidsartade och rumsartade vektorer som ingår i en ON-bas kallas metrikens *signatur*.

Signaturen brukar anges på lite olika sätt i olika skrifter. Om t ex $\dim(M) = n = 5$ och en ON-bas innehåller 3 st tidsartade och 2 st rumsartade vektorer kan signaturen anges som $(3, 2)$, som $3+2$ eller som $(+++--)$ (vi kommer att använda detta skrivsätt i denna skrift). Om det inte råder någon tvekan om att $n = 5$ anges signaturen ibland som skillnaden mellan antalet tidsartade och antalet rumsartade vektorer, d v s som 1 i detta fall.

I både allmän och speciell relativitetsteori är vi främst intresserade av metriker med signatur $(+---)$[2] vilket motiverar

Definition 5.18. En 4-dimensionell mångfald M försedd med en metrik av signatur $(+---)$ kallas en *rumtid*.

Enligt avsnitt 5.3 är koordinatbasen inte nödvändigtvis en ON-bas. I detta avsnitt var visserligen koordinatbasvektorerna $\frac{\partial}{\partial\theta}$ och $\frac{\partial}{\partial\varphi}$ inbördes ortogonala men $\frac{\partial}{\partial\varphi}$ var inte normerad (dess skalärprodukt med sig själv blev ju $\sin^2\theta$ istället för 1). Det är också enkelt att konstruera koordinatsystem i vilka koordinatbasvektorerna inte ens är inbördes ortogonala (de flesta linjära koordinatbyten i $\mathbf{R}^2$ ger ett koordinatsystem med denna egenskap).

Följande sats gäller dock

SATS 5.19. *Givet en godtycklig mångfald M med metrik g_{ab} och en godtycklig punkt $p \in M$ så existerar ett koordinatsystem $\bar{x}^a$ i en omgivning till p sådant att koordinatbasen $\frac{\partial}{\partial\bar{x}^a}$ är en ON-bas för T_pM så i dessa koordinater ges metriken av en diagonalmatris med diagonalelement ± 1 i punkten p.*

Bevis. Låt x^a vara ett koordinatsystem i en omgivning av p. Enligt ovan existerar tal $(e_i)^a$ sådana att $\mathbf{e}_i = (e_i)^a \frac{\partial}{\partial x^a}$ utgör en ON-bas för T_pM. Definiera nya koordinater $\bar{x}^i$ i en omgivning av p genom sambandet $x^a = (e_i)^a \bar{x}^i$ (varför är detta ett koordinatbyte?). Enligt kedjeregeln är då

$$\frac{\partial}{\partial\bar{x}^i} = \frac{\partial x^a}{\partial\bar{x}^i}\frac{\partial}{\partial x^a} = (e_i)^a \frac{\partial}{\partial x^a} = \mathbf{e}_i$$

så koordinatbasen $\frac{\partial}{\partial\bar{x}^i}$ utgör en ON-bas för T_pM. Enligt transformationslagen för tensorer av typ $(0, 2)$ ges metrikens komponenter i punkten p i de nya koordinaterna av

$$\bar{g}_{ij} = g_{ab}\frac{\partial x^a}{\partial\bar{x}^i}\frac{\partial x^b}{\partial\bar{x}^j} = g_{ab}(e_i)^a(e_j)^b = \begin{cases} \pm 1 & \text{om } i = j \\ 0 & \text{om } i \neq j \end{cases}$$

eftersom $(e_i)^a$ utgjorde en ON-bas i p. □

[2] Signaturerna $(-+++)$ och $(+++-)$ förekommer också i litteraturen. De ger en helt ekvivalent beskrivning förutsatt att begreppen tidsartad och rumsartad också byter betydelse.

Anmärkning: Det är omöjligt att överdriva vikten av att inse att den senaste satsen *bara gäller punktvis*. Utanför p är det inte säkert (det är inte ens särskilt troligt) att $\frac{\partial}{\partial \bar{x}^i}$ utgör en ON-bas för tangentrummet och då blir inte heller metriken en diagonalmatris utanför p. Existensen av en koordinatbas som också är en ON-bas på en hel öppen mängd visar sig vara ett mycket starkt krav att ställa på en mångfald. Mer om detta i senare kapitel. □

En konsekvens av denna sats är att om M är en rumtid och $p \in M$ så finns det koordinater i en omgivning av p sådana att $g_{ab}|_p = \text{diag}(1, -1, -1, -1)$, d v s $g_{ab}|_p = \eta_{ab}$ där η_{ab} är Minkowskimetriken (se Exempel 5.10). Det finns därför anledning att undersöka Minkowskimetriken lite närmare.

Exempel 5.20. Låt M vara en rumtid och låt $p \in M$ vara godtycklig. Som i Sats 5.19 finns koordinater $x^a = (t, x, y, z)$ sådana att $g_{ab} = \text{diag}(1, -1, -1, -1) = \eta_{ab}$ i punkten p. Låt $x^a = x^a(s)$ vara en parametrisering av en kurva Γ genom p med tangentvektor $v^a = \frac{dx^a}{ds}\big|_p$ i p.

Vi studerar först fallet att v^a är en nollvektor. I punkten p är alltså

$$(v^t)^2 - (v^x)^2 - (v^y)^2 - (v^z)^2 = g_{ab}v^a v^b = \eta_{ab}v^a v^b = 0,$$

där v^t, v^x, v^y, v^z är v^a:s komponenter i koordinatbasen. Lösningarna till denna ekvation bildar en dubbelkon i T_pM med spetsen i origo, kallad *ljuskonen* i p. De två halvorna av ljuskonen i p kännetecknas av att $v^t > 0$ i ena halvan och $v^t < 0$ i den andra.

Betrakta nu en partikel som följer Γ och använd relativistiska enheter d v s $c = 1$ där c är ljusets hastighet (eller, mer terminologiskt korrekt, ljusets fart) i vakuum. Partikelns fart i det ögonblick den passerar p ges då av

$$v = \frac{\sqrt{(v^x)^2 + (v^y)^2 + (v^z)^2}}{|v^t|} = 1.$$

Om en partikel följer en kurva Γ vars tangentvektor är en nollvektor rör sig alltså partikeln med ljusets hastighet, vilket motiverar varför en nollvektor också kallas en ljusartad vektor och varför ovanstående kon kallas ljuskonen.

Om v^a istället är tidsartad, d v s

$$(v^t)^2 - (v^x)^2 - (v^y)^2 - (v^z)^2 = g_{ab}v^a v^b > 0$$

är alltså v^a en vektor som pekar in i någon av ljuskonens båda halvor. Vilken av dem bestäms av tecknet på v^t precis som för nollvektorer.

Vi inför följande terminologi.

Definition 5.21. En nollvektor eller tidsartad vektor v^a är *framtidsriktad* om $v^t > 0$ och *dåtidsriktad*[3] om $v^t < 0$.

Observera att begreppen framtidsriktad och dåtidsriktad är koordinatberoende. Om vi byter till tidskoordinaten $\bar{t} = -t$ så byter vi framtidsriktade vektorer mot dåtidsriktade och vice versa.

[3] Äras den som äras bör. Denna översättning av ordet 'past-pointing' är inte min egen utan den härrör från Lars Alexandersson.

Farten hos en partikel som följer en bana Γ vars tangent v^a i p är tidsartad ges av

$$v = \frac{\sqrt{(v^x)^2 + (v^y)^2 + (v^z)^2}}{|v^t|} < 1,$$

så om tangenten till partikelbanan är tidsartad rör sig partikeln med en fart lägre än ljusets. Alla partiklar som observerats rör sig alltså längs banor vars tangentvektorer är tidsartade eller nollvektorer.

Till sist undersöker vi fallet att v^a är rumsartad så att

$$(v^t)^2 - (v^x)^2 - (v^y)^2 - (v^z)^2 = g_{ab}v^a v^b < 0$$

d v s v^a pekar i en riktning som ligger utanför ljuskonen.

Om v^a är tangent till en partikels bana ges partikelns fart av

$$v = \frac{\sqrt{(v^x)^2 + (v^y)^2 + (v^z)^2}}{|v^t|} > 1,$$

så partikeln rör sig fortare än vad ljuset gör i vakuum. En sådan partikel kallas en *tachyon* och har aldrig observerats. $\square$

Kapitel 6

(⋆ Abstrakta tensorer ⋆)

6.1 Inledning

Syftet med kapitel 4 och 5 har varit att att introducera tensorbegreppet på ett sätt som inte kräver så mycket matematiska förkunskaper och samtidigt inte heller tummar alltför mycket på den matematiska stringensen. Därför har fokus legat på vad som kan *göras* med tensorer istället för på vad tensorer *är*, i någon djupare mening. Som tidigare sagts räcker detta gott och väl för att få en stabil matematisk grund för den allmänna relativitetsteorin.

För många känns dock detta angreppssätt inte helt tillfredsställande. T ex förväntar vi oss inte att naturlagarna ska bero på vilket koordinatsystem vi valt för att ange punkters läge i en rumtid. Som vi har sett är detta visserligen inget problem så länge naturlagarna formuleras som tensorekvationer, eftersom sådana är giltiga i *alla* koordinatsystem så fort de är giltiga i något.

Frågan är dock varför det ska vara nödvändigt att införa ett koordinatsystem över huvud taget. Naturlagarna borde väl gälla även om vi, med en dåres envishet, vägrar att införa ett koordinatsystem.

Därför kan det kännas onaturligt att definiera begreppet 'tensor' utifrån koordinattransformationslagen (4.12). Mer naturligt vore i så fall att göra denna definition helt koordinatfri[1].

I detta kapitel ska vi därför gå lite mer på djupet och förklara vad tensorer *är*, och inte bara vad vi kan göra med dem. T ex ska vi se på en koordinatfri definition av en tensor. Vi ska också rättfärdiga vårt lite slappa och oegentliga språkbruk när vi talar om 'vektorn v^a' istället för 'vektorn $\mathbf{v} = v^a \frac{\partial}{\partial x^a}$ med komponenter v^a i koordinatbasen'. På vägen kommer vi också att lösa mysteriet som beskrivs i slutet av avsnitt 4.4, d v s om nu en kontravariant tensor v^a är komponenter till en tangentvektor $\mathbf{v}$, vad är då en kovariant tensor α_a komponenter till för slags objekt?

På några ställen i kapitlet blir vi tvungna att modifiera terminologin och notationen från de föregående kapitlen lite grand. Jag hoppas att detta inte leder till förvirring.

Till sist några ord på vägen: Hela detta kapitel består av överkursmaterial och är också mer

[1] Vissa anser t o m att definitionen av begreppet 'mångfald' bör vara helt koordinatfri. Hur en sådan definition ska se ut ligger bortom horisonten för denna skrift, men en introduktion ges i *Ludvigsen*: General Relativity – A Geometric Approach.

matematiskt krävande än kapitel 4 och 5. Var alltså förberedd på att behöva kämpa lite extra med det. Den extra ansträngningen är dock, i min mening, väl värd sitt pris.

6.2 Kovektorer

Låt oss nu förflytta oss lite bakåt i tiden (rummet?, rumtiden?, skriften?), till slutet på avsnitt 4.3. Vi antar alltså att vi redan definierat mångfalder och tangentvektorer, infört koordinatbasen och härlett transformationslagen (4.6). Vi ska ägna resten av detta kapitel till att åter bygga upp teorin från senare delen av kapitel 4 och från kapitel 5, men vi ska göra det på ett annat sätt än i dessa kapitel.

Låt M vara en godtycklig n-dimensionell mångfald och låt $p \in M$. Låt x^i vara lokala koordinater i en omgivning U av p (av skäl som kommer att framgå i nästa avsnitt använder vi i detta kapitel endast bokstäverna i, j, k, l, m och i nödfall även n (att n i så fall *inte* står för mångfaldens dimension bör vara uppenbart) för att beteckna numeriska index som antar värdena $0, 1, \ldots, n-1$).

Vi gör följande definition:

Definition 6.1. En linjär avbildning $\alpha : T_pM \to \mathbf{R}$ kallas en *kovektor* (eller en *dualvektor*) i p. Mängden av alla kovektorer i p kallas *dualrummet* till T_pM (eller *kotangentrummet* i p) och betecknas T_pM^*.

I detta kapitel betecknas kovektorer oftast med grekiska bokstäver $(\alpha, \beta, \ldots)$, men även andra typer av beteckningar kommer att förekomma.

Vi noterar också att T_pM^* blir ett vektorrum eftersom mängden av linjära avbildningar är sluten under både addition och multiplikation med tal.

Anmärkning: I andra grenar av matematiken, t ex inom funktionalanalysen, kallas en kovektor en *linjär funktional* och den är ofta komplexvärd istället för reellvärd. Dessutom behöver den inte vara definierad just på tangentrummet till en mångfald utan den kan vara definierad på ett godtyckligt vektorrum, vars dimension inte ens behöver vara ändlig. □

Exempel 6.2. Låt $M = \mathbf{R}^2$ och låt L vara en rät linje i T_pM som går genom origo och har en riktningsvektor $\mathbf{a} \neq \mathbf{0}$.

Låt $\mathbf{b}$ vara en annan vektor i T_pM som inte är parallell med $\mathbf{a}$. Givet en godtycklig vektor $\mathbf{v} \in T_pM$ gör vi följande konstruktion: Låt $\mathbf{v}$ utgå från origo. Drag en rät linje parallell med $\mathbf{b}$ genom $\mathbf{v}$:s spets. Denna linje skär L i exakt en punkt, d v s det finns ett reellt tal $\alpha(\mathbf{v})$ (som förstås beror på $\mathbf{v}$) sådant att denna punkts ortsvektor ges av $\alpha(\mathbf{v})\mathbf{a}$. Denna vektor kallas $\mathbf{v}$:s (sneda) projektion på $\mathbf{a}$ parallellt med $\mathbf{b}$.

Betrakta avbildningen $\alpha : \mathbf{v} \mapsto \alpha(\mathbf{v})$. α är då reellvärd och det är lätt att verifiera (Gör det genom att rita en figur!) att α är linjär i $\mathbf{v}$. Det följer att α är en kovektor i p. □

Låt nu $\left(\frac{\partial}{\partial x^j}\right)_{j=0}^{n-1}$ vara koordinatbasen för T_pM. Som tidigare kan en godtycklig vektor $\mathbf{v} \in T_pM$ skrivas $\mathbf{v} = v^j \frac{\partial}{\partial x^j}$ för några tal v^j. Då är β, given av $\beta(\mathbf{v}) = v^0$ linjär och reellvärd, och därmed en kovektor i p och vi ser att

$$\beta\left(\frac{\partial}{\partial x^j}\right) = \begin{cases} 1 & \text{om } j = 0 \\ 0 & \text{om } j \neq 0 \end{cases} = \delta_j{}^0.$$

Denna kovektor är, som vi kommer att se nedan, väldigt användbar och det är därför praktiskt att ge den en egen beteckning. Vi gör därför följande definition

Definition 6.3. Kovektorerna dx^i, $i = 0, 1, \ldots, n-1$ definieras av sambandet

$$dx^i\left(\frac{\partial}{\partial x^j}\right) = \delta_j{}^i. \tag{6.1}$$

Observera att vi här frångår principen att beteckna kovektorer med grekiska bokstäver. Eftersom dessa kovektorer är konstruerade direkt från koordinaterna x^i är det naturligt att låta detta avspegla sig i beteckningen. Vi ser också att om $\mathbf{v} = v^j \frac{\partial}{\partial x^j}$ så är, p g a lineariteten hos dx^i,

$$dx^i(\mathbf{v}) = dx^i\left(v^j \frac{\partial}{\partial x^j}\right) = v^j dx^i\left(\frac{\partial}{\partial x^j}\right) = v^j \delta_j{}^i = v^i,$$

analogt med kovektorn β ovan.

Med denna definition gäller

SATS 6.4. *Mängden $B = \left\{\left(dx^i\right)_{i=0}^{n-1}\right\}$ är en bas för T_pM^*.*

Bevis. Vi visar först att B är linjärt oberoende. Ansätt därför en linjär relation $\lambda_i dx^i = 0$, där 0 betyder nollavbildningen, d v s den linjära avbildning som avbildar varje tangentvektor på talet 0. För alla $j = 0, 1, \ldots, n-1$ gäller då

$$0 = \lambda_i dx^i\left(\frac{\partial}{\partial x^j}\right) = \lambda_i \delta_j{}^i = \lambda_j,$$

vilket visar att B är linjärt oberoende.

För att visa att B spänner upp T_pM^* låter vi $\alpha \in T_pM^*$ vara godtycklig. Vi sätter sedan $\alpha_i = \alpha\left(\frac{\partial}{\partial x^i}\right)$. Om nu $\mathbf{v} = v^j \frac{\partial}{\partial x^j}$ är en godtycklig vektor i T_pM så fås

$$\begin{aligned} \alpha(\mathbf{v}) &= \alpha\left(v^j \frac{\partial}{\partial x^j}\right) = v^j \alpha\left(\frac{\partial}{\partial x^j}\right) = v^j \alpha_j = \alpha_i v^j \delta_j{}^i = \alpha_i v^j dx^i\left(\frac{\partial}{\partial x^j}\right) \\ &= \alpha_i dx^i\left(v^j \frac{\partial}{\partial x^j}\right) = \alpha_i dx^i(\mathbf{v}). \end{aligned} \tag{6.2}$$

Eftersom $\mathbf{v}$ var godtycklig följer att $\alpha = \alpha_i dx^i$, d v s B spänner upp T_pM^*. □

Mängden B ovan kallas *dualbasen* till koordinatbasen $\left\{\left(\frac{\partial}{\partial x^j}\right)_{j=0}^{n-1}\right\}$ och speciellt ser vi att tangentrummet T_pM och dess dualrum T_pM^* har samma dimension.

Låt nu $f \in \mathcal{C}^\infty$ och $\mathbf{v} \in T_pM$ vara godtyckliga. Låter vi $\mathbf{v}$ verka på f får vi enligt ekvation 4.5 funktionen

$$\mathbf{v}(f) = v^j \frac{\partial f}{\partial x^j}, \tag{6.3}$$

där v^j är $\mathbf{v}$:s komponenter i koordinatbasen. Betraktad på detta sätt är alltså $\mathbf{v}$ en avbildning som avbildar en $\mathcal{C}^\infty$ funktion f på en annan $\mathcal{C}^\infty$ funktion $\mathbf{v}(f)$.

Ett annat sätt att tolka ekvation 6.3 är att se f som fix och $\mathbf{v}$ som variabel. Vi får då en avbildning $\mathbf{v} \mapsto \mathbf{v}(f)$ och vi observerar att denna variabel bestäms helt och hållet av funktionen f och den är också linjär i $\mathbf{v}$. Denna avbildning är alltså en kovektor och vi definierar:

Definition 6.5. *Differentialen* till f är den kovektor df som definieras av sambandet

$$df(\mathbf{v}) = \mathbf{v}(f).$$

Då df är en kovektor kan den förstås skrivas som en linjärkombination av dualbasvektorerna dx^j:

SATS 6.6.

$$df = \frac{\partial f}{\partial x^j} dx^j$$

Bevis. Om vi låter df verka på en godtycklig tangentvektor $\mathbf{v}$ med komponenter v^i i koordinatbasen fås

$$df(\mathbf{v}) = \mathbf{v}(f) = v^i \frac{\partial f}{\partial x^i} = \frac{\partial f}{\partial x^j} \delta_i{}^j v^i = \frac{\partial f}{\partial x^j} dx^j \left(\frac{\partial}{\partial x^i} \right) v^i = \frac{\partial f}{\partial x^j} dx^j \left(v^i \frac{\partial}{\partial x^i} \right) = \frac{\partial f}{\partial x^j} dx^j(\mathbf{v}),$$

där likheterna följer av, i tur och ordning, Definition 6.5, ekvation 6.3, Definition 4.10, ekvation 6.1 samt lineariteten hos dx^j. Då $\mathbf{v}$ var godtycklig följer det att $df = \frac{\partial f}{\partial x^j} dx^j$. □

En konsekvens av denna sats är att vår nya definition av differential tydligen är ekvivalent med den gamla (Definition 4.7).

Ovan har vi alltså definierat en kovektor som en linjär avbildning som avbildar tangentvektorer på reella tal. En naturlig fråga att ställa sig är då: Vad händer om vi upprepar proceduren och definierar en 'ko-kovektor' (eller vad vi nu ska kalla den) som en linjär avbildning som avbildar kovektorer på reella tal. Mängden av alla 'ko-kovektorer' bildar då ett vektorrum T_pM^{**}, kallat *bidualen* till T_pM.

Vi observerar att för varje tangentvektor $\mathbf{v} \in T_pM$ kan vi sätta

$$\mathbf{V}(\alpha) := \alpha(\mathbf{v}), \tag{6.4}$$

vilket definierar en 'ko-kovektor' $\mathbf{V} \in T_pM^{**}$ eftersom denna avbildning är reellvärd och linjär i α. Enligt ovan har dessutom T_pM^* och T_pM^{**} samma dimension, d v s T_pM^{**} har samma dimension som T_pM.

Det följer att T_pM^{**} inte kan innehålla några andra element än de som ges av ekvation (6.4). En 'ko-kovektor' är alltså 'samma sak' som en tangentvektor d v s T_pM och T_pM^{**} är samma[2] vektorrum.

Vektorrum med denna egenskap sägs vara *reflexiva* och ett resonemang nästan identiskt med ovanstående visar att alla vektorrum av ändlig dimension är reflexiva[3].

Låt oss nu byta till nya koordinater $\bar{x}^j$ i en omgivning av p. Vi får då nya koordinatbasvektorer $\frac{\partial}{\partial \bar{x}^j}$ och därmed också nya dualbasvektorer $d\bar{x}^i$. Därför finns tal $\bar{\alpha}_i$ sådana att kovektorn $\alpha = \alpha_i dx^i$ kan skrivas $\alpha = \bar{\alpha}_i d\bar{x}^i$ och vi söker ett samband mellan α_i och $\bar{\alpha}_i$.

SATS 6.7. *Om $\alpha = \alpha_i dx^i = \bar{\alpha}_i d\bar{x}^i$ så är*

$$\bar{\alpha}_i = \frac{\partial x^j}{\partial \bar{x}^i} \alpha_j. \tag{6.5}$$

Bevis. Enligt räkningen i ekvation (6.2) gäller att $\alpha(\mathbf{v}) = v^k \alpha_k$ för alla tangentvektorer $\mathbf{v} = v^j \frac{\partial}{\partial x^j}$. Följdaktligen är också $\alpha(\mathbf{v}) = \bar{v}^m \bar{\alpha}_m$. Enligt transformationslagen (4.6) är då

$$v^k \alpha_k = \bar{v}^m \bar{\alpha}_m = \frac{\partial \bar{x}^m}{\partial x^k} v^k \bar{\alpha}_m,$$

för alla v^k. Låt nu $v^0 = 1$, övriga $v^k = 0$. Detta ger $\alpha_0 = \frac{\partial \bar{x}^m}{\partial x^0} \bar{\alpha}_m$ och på samma sätt visas att $\alpha_j = \frac{\partial \bar{x}^m}{\partial x^j} \bar{\alpha}_m$ för alla övriga värden på j. Kontrahera detta samband med $\frac{\partial x^j}{\partial \bar{x}^i}$ så fås

$$\frac{\partial x^j}{\partial \bar{x}^i} \alpha_j = \frac{\partial x^j}{\partial \bar{x}^i} \frac{\partial \bar{x}^m}{\partial x^j} \bar{\alpha}_m = \frac{\partial \bar{x}^m}{\partial \bar{x}^i} \bar{\alpha}_m = \delta_i{}^m \bar{\alpha}_m = \bar{\alpha}_i$$

□

Observera att i detta kapitels framställning är ekvation (6.5) en sats, medan den identiska ekvationen (4.8) i avsnitt 4.4 är en definition.

6.3 Abstrakta indexnotationen

Låt oss nu stanna upp och tänka efter ett ögonblick. Hittills har vi definierat tangentvektorer $\mathbf{v} \in T_pM$ och kovektorer $\alpha \in T_pM^*$ och om vi tänker tillbaka på de föregående kapitlen inser vi att vi behöver definiera tensorer av godtycklig typ (s, t). Hur ska vi skaffa oss en effektiv betecknings-apparat för objekt av så många olika slag?

Helt klart är att det inte går att fortsätta som vi hittills gjort, d v s växla mellan olika alfabet, olika typsnitt och liknande. En sådan beteckningsapparat skulle snabbt bli alldeles för komplicerad och vi måste därför komma på någonting bättre.

[2]Mer precist sägs avbildningen som avbildar $\mathbf{v}$ på $\mathbf{V}$ ovan vara en *isomorfism* om alla element i T_pM^{**} kan fås på detta sätt. Vektorrummen T_pM och T_pM^{**} sägs då vara *isomorfa*.

[3]För vektorrum av oändlig dimension är saken mer komplicerad. Många viktiga rum är reflexiva men många andra är det inte. Reflexivitet (och mycket annat) studeras närmare i ämnet funktionalanalys.

I de föregående kapitlen definierades tangentvektorer och kovektorer utifrån vilka transformationslagar deras komponenter uppfyller vid koordinatbyten (och även allmänna tensorer faktiskt, det kommer vi att se i nästa avsnitt). Som vi tidigare diskuterat är denna definition visserligen enkel och förutsätter inte så mycket förkunskaper, men den är också otillfredsställande på många sätt.

Trots detta kan ingen förneka att denna definition ger oss en mycket effektiv beteckningsapparat. Den använder inga exotiska typsnitt utan bara 'vanliga' bokstäver. Dessutom räcker det att kasta en snabb blick på t ex $T^{ij}{}_{klm}$ för att se att detta är en tensor av typen $(2, 3)$. Ingen ytterligare precisering behövs.

En idé är därför att kopiera denna beteckningsapparat, men istället låta de indexerade kvantiteterna stå för tensorerna själva och inte för deras komponenter. Detta kan göras genom att introducera den s k *abstrakta indexnotationen*.

Antag att $p \in M$ och att $\mathbf{v} \in T_pM$. Vi inför nu beteckningarna T_pM^a, T_pM^b, $\ldots$ för olika kopior av T_pM. Observera att här är a, b, $\ldots$ inga numeriska index (minns att i detta kapitel betecknas numeriska index med bokstäverna i, j, k, l, m och i nödfall n (som i så fall ej ska förväxlas med $n = \dim M$)).

Det är alltså inte meningsfullt att t ex sätta $a = 0$. Indexet a (t ex) ska istället ses som en abstrakt 'etikett' som talar om att T_pM^a är just det vanliga tangentrummet i p och inte något annat vektorrum.

Eftersom T_pM^a är en kopia av T_pM finns förstås en tangentvektor motsvarande $\mathbf{v}$ i T_pM^a. Denna tangentvektor betecknas v^a. Vi använder alltså samma latinska stambokstav som i ursprungliga vektorn $\mathbf{v}$, dock utan fetstil.

Vi förser också vektorn med samma index som på kopian T_pM^a, ett s k *abstrakt index*. Detta index ska också placeras likadant som på T_pM^a, d v s uppe.

Med denna notation är det alltså helt ekvivalent att skriva $\mathbf{v} \in T_pM$ och $v^a \in T_pM^a$. Vi är förstås inte heller tvungna att använda just bokstaven a för att beteckna det abstrakta indexet utan vi kan också t ex skriva $v^b \in T_pM^b$, $v^e \in T_pM^e$ o s v.

Vi kan t o m nöja oss med att skriva v^a, och alltså utelämna '$\in T_pM^a$', eftersom det abstrakta indexet ^{a} talar om för oss att v^a är ett element i T_pM^a.

Observera att nollvektorn skrivs utan något abstrakt index. Motsvarigheten till påståendet $\mathbf{u} = \mathbf{0}$ är alltså $u^a = 0$ och inte $u^a = 0^a$ i den abstrakta indexnotationen.

Antag nu att $u^a, v^a \in T_pM^a$. Eftersom T_pM^a är ett vektorrum kan vi förstås bilda $u^a + v^a$ och detta blir också en vektor i T_pM^a. Analogt, om $u^b, v^b \in T_pM^b$ så är $u^b + v^b \in T_pM^b$ väldefinierad.

Om däremot $u^a \in T_pM^a$ och $v^b \in T_pM^b$ är visserligen u^a och v^b båda tangentvektorer till M i p, men då u^a och v^b tillhör *olika* kopior av T_pM, och därmed också olika vektorrum är deras summa $u^a + v^b$ *ej* definierad. Vi kan alltså bara addera vektorer om de är försedda med *samma* abstrakta index.

Så här långt ser abstrakta indexnotationen identisk ut med notationen i de föregående kapitlen, men vi kommer nu till en viktig skillnad. Betrakta någon av koordinatbasvektorerna $\frac{\partial}{\partial x^i} \in T_pM$. Här är i inget abstrakt index utan ett numeriskt. $\frac{\partial}{\partial x^i}$ betecknar ju den tangentvektor som verkar på en godtycklig funktion f genom att derivera f med avseende på koordinat nummer i, d v s x^i.

Hur ska vi då beteckna den motsvarande tangentvektorn i kopian T_pM^a? Vi måste förse vektorn med ett abstrakt index a. Vi betecknar därför denna vektor $\left(\frac{\partial}{\partial x^i}\right)^a$.

Observera skillnaden mellan de båda indexen i denna beteckning. a:et är ett abstrakt index. Detta index antar inga värden. Det talar istället om för oss att vi har med ett element i T_pM^a att göra, d v s en tangentvektor. i:et är å andra sidan ett numeriskt index. i betecknar något av talen $0, 1, \ldots, n-1$ där $n = \dim M$, och talar om för oss vilken av de n koordinatbasvektorerna vi menar.

Låt $v^a \in T_pM^a$ vara godtycklig. Då vektorerna $\left(\frac{\partial}{\partial x^i}\right)^a$ utgör en bas för T_pM^a enligt definition 4.2 och sats 4.3 finns (som tidigare) tal v^i (där i förstås är ett numeriskt index) så att

$$v^a = v^i \left(\frac{\partial}{\partial x^i}\right)^a, \tag{6.6}$$

där vi (förstås) summerar över i p g a Einsteins summationskonvention.

Ingenting hindrar nu att vi låter v^a betyda ett vektorfält (minns att ett vektorfält associerar varje punkt p i någon delmängd till M till en tangentvektor v^a i p så att v^a:s komponenter i godtycklig koordinatbas blir funktioner av klass $\mathcal{C}^\infty$) istället för en tangentvektor i en punkt $p \in M$, men vi ser att i så fall blir ekvation (6.6) ej allmängiltig. Den gäller bara i mängden U d v s i den omgivning till p där koordinatsystemet x^i är definierat. I andra delar av M ges v^a visserligen av liknande formler, men med koordinatsystemet utbytt.

Den abstrakta indexnotationen är alltså mer generell än indexnotationen i de föregående kapitlen i följande viktiga avseende: Om ett tangentvektorfält v^a identifieras med sina komponenter v^i (som i de tidigare kapitlen) gäller alla samband och formler vi skriver upp endast *lokalt*, närmare bestämt i en öppen mängd i M som kan beskrivas med ett enda koordinatsystem. I den abstrakta indexnotationen går det däremot utmärkt att skriva upp samband som gäller i hur stora mängder som helst, t ex i *hela* M.

Nästa steg blir att utvidga den abstrakta indexnotationen till kovektorer och för att göra detta sneglar vi på indexnotationen i de föregående kapitlen. Där betecknade vi en kovariant tensor av typ $(0, 1)$ (d v s komponenterna till en kovektor) med ett index nere.

Vi låter därför T_pM_a, T_pM_b, ... vara olika kopior av T_pM^*. Om $\alpha \in T_pM^*$ betecknar vi motsvarande kovektor i T_pM_a med α_a. Motsvarande kovektor i T_pM_b betecknas α_b o s v.

Vi behöver också förse kovektorerna $dx^i \in T_pM^*$ i dualbasen till koordinatbasen med abstrakta index. Motsvarande kovektor i T_pM_a betecknas $(dx^i)_a$. Observera att samma skillnad som ovan, mellan de båda indexen, råder här.

Minns att T_pM^* är dualrummet till T_pM, d v s T_pM^* är mängden av alla linjära avbildningar från T_pM till $\mathbf{R}$. T_pM_a, T_pM_b, ... blir därmed också dualrum till T_pM och/eller någon av dess kopior T_pM^a, T_pM^b, ...

Den konvention vi kommer att använda är att en kopia av T_pM^* försedd med ett abstrakt index är dualrum till den kopia av T_pM som är försedd med samma abstrakta index. Exempelvis är alltså T_pM_a dualrum till T_pM^a, d v s T_pM_a är mängden av alla linjära avbildningar från T_pM^a till $\mathbf{R}$.

Om $\alpha \in T_pM^*$ och $\mathbf{v} \in T_pM$ är godtyckliga kan vi som bekant låta kovektorn α verka på tangentvektorn $\mathbf{v}$ och få det reella talet $\alpha(\mathbf{v})$. Uttrycker vi detta med abstrakta indexnotationen är alltså $\alpha_a \in T_pM_a$ och $v^a \in T_pM^a$ godtyckliga och låter vi α_a verka på v^a fås det reella talet $\alpha_a(v^a)$. Här kan förstås det abstrakta indexet a bytas ut mot vilken annan bokstav som helst.

Parenteserna i detta uttryck utelämnas i allmänhet så vi skriver $\alpha_a v^a$ istället för $\alpha_a(v^a)$. Observera att denna upprepning av det abstrakta indexet a *inte* innebär någon summering. Indexet a antar ju inte ens några värden att summera över. Upprepningen betyder endast att v^a ska ges som argument till avbildningen α_a så att denna verkar på v^a och returnerar det reella talet $\alpha_a v^a$.

Observera att T_pM_a *inte* är dualrum till T_pM^b eftersom de abstrakta indexen inte matchar. Uttrycket $\alpha_a(v^b)$ är alltså odefinierat. Motsvarande uttryck utan parenteser, $\alpha_a v^b$ kommer vi att definiera i nästa avsnitt. Det kommer då att beteckna den s k *tensorprodukten* av α_a och v^b och är därför *inte* samma sak som $\alpha(\mathbf{v})$.

I förra avsnittet såg vi att dualrummet till T_pM^* var 'samma' vektorrum som T_pM i följande mening: Givet ett godtyckligt element $\mathbf{V}$ i dualrummet T_pM^{**} till T_pM^* så finns en tangentvektor $\mathbf{v} \in T_pM$ sådan att $\mathbf{V}(\alpha) = \alpha(\mathbf{v})$ för alla kovektorer α. Något oegentligt kan vi alltså identifiera $\mathbf{V}$ och $\mathbf{v}$ och skriva $\mathbf{v}(\alpha) = \alpha(\mathbf{v})$. Med abstrakta index skrivs detta samband $v^a(\alpha_a) = \alpha_a(v^a)$ eller, om parenteserna utelämnas $v^a\alpha_a = \alpha_a v^a$. För att abstrakta indexnotationen ska bli konsistent *måste* vi alltså identifiera uttrycken $v^a\alpha_a$ och $\alpha_a v^a$.

Dualbasvektorn dx^i definierades av sambandet $dx^i\left(\frac{\partial}{\partial x^j}\right) = \delta_j{}^i$. Med abstrakta index skrivs detta samband

$$(dx^i)_a \left(\frac{\partial}{\partial x^j}\right)^a = \left(\frac{\partial}{\partial x^j}\right)^a (dx^i)_a = \delta_j{}^i, \tag{6.7}$$

om vi, som ovan, utelämnar parenteserna när en kovektor verkar på en vektor.

Om $\alpha_a = \alpha_i(dx^i)_a$ och $v^a = v^j\left(\frac{\partial}{\partial x^j}\right)^a$ är godtyckliga ger denna ekvation samt linearitet att

$$\alpha_a v^a = \alpha_a\left(v^j\left(\frac{\partial}{\partial x^j}\right)^a\right) = v^j\alpha_a\left(\left(\frac{\partial}{\partial x^j}\right)^a\right) = v^j\alpha_i(dx^i)_a\left(\frac{\partial}{\partial x^j}\right)^a = v^j\alpha_i\delta_j{}^i = v^i\alpha_i = \alpha_i v^i,$$

där sista likheten följer av att α_i och v^i är vanliga reella tal. Observera att detta väsentligen är samma räkning som i ekvation (6.2). Notera också hur lika de båda ytterleden ser ut trots att de betecknar helt olika saker. I vänstra ledet står ju det tal vi erhåller när kovektorn α_a verkar på vektorn v^a och i högra ledet står det tal vi får när vi multiplicerar talen α_i och v^i med varandra och sedan summerar över i.

Den abstrakta indexnotationen imiterar alltså även i detta avseende den indexnotation vi använt i tidigare kapitel.

Vi påminner om att differentialen till en funktion f är en kovektor df som definieras av sambandet $df(\mathbf{v}) = \mathbf{v}(f) = v^j\frac{\partial f}{\partial x^j}$ för alla tangentvektorer $\mathbf{v} \in T_pM$. I abstrakta indexnotationen ska diferentialen df förses med ett index nere och skrivs därför df_a. Definitionen av differentialen skrivs då

$$df_a v^a = \mathbf{v}(f) = v^j\frac{\partial f}{\partial x^j} \text{ för alla } v^a \in T_pM^a$$

och Sats 6.6 lyder

$$df_a = \frac{\partial f}{\partial x^j}(dx^j)_a.$$

Anmärkning: Om vi låter en tangentvektor $\mathbf{v}$ verka på en funktion f fås enligt ovan funktionen $\mathbf{v}(f) = v^j \frac{\partial f}{\partial x^j}$. Med abstrakta index verkar det naturligt att skriva denna funktion som $v^a(f)$ och så görs också i många skrifter. Observera dock att $v^a(f)$ fortfarande är försedd med ett index uppe och därför, strikt tolkat, borde vara en tangentvektor och inte en funktion.

På grund av denna inkonsistens kommer vi att undvika beteckningen $v^a(f)$ i denna skrift och istället skriva $\mathbf{v}(f)$ eller $v^a df_a = df_a v^a$. $\square$

Vi påminner också om den alternativa beteckningen $\nabla_a f$ för differentialen till f. Det gäller alltså att

$$\nabla_a f = df_a = \frac{\partial f}{\partial x^j}(dx^j)_a$$

d v s komponenterna $\nabla_j f$ till $\nabla_a f$ är alltså $\nabla_j f = \frac{\partial f}{\partial x^j}$ (jämför ekvation 4.7).

6.4 Multilinjära avbildningar

I detta avsnitt ska vi definiera allmänna tensorer och dessutom utvidga den abstrakta indexnotationen till dessa. Först några praktiska definitioner.

Definition 6.8. Om A och B är två mängder definieras den *kartesiska produkten* $A \times B$ som

$$A \times B = \{(a, b)\,;\ a \in A \text{ och } b \in B\}\,.$$

T ex gäller att $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$ eftersom $\mathbf{R}^2$ är mängden av alla ordnade par av reella tal.

Om vi vill ta kartesiska produkten av fler faktorer än 2 behöver vi iaktta en smula försiktighet. Mängden $(A \times B) \times C$ består av alla element på formen $((a, b), c)$ där $a \in A$, $b \in B$ och $c \in C$.

Mängden $A \times (B \times C)$ består i stället av alla element på formen $(a, (b, c))$ där återigen $a \in A$, $b \in B$ och $c \in C$. Då elementen i $(A \times B) \times C$ och $A \times (B \times C)$ är objekt av helt olika typ är förstås $(A \times B) \times C \neq A \times (B \times C)$ i allmänhet.

Å andra sidan kan både $((a, b), c)$ och $(a, (b, c))$ identifieras med den ordnade trippeln (a, b, c) på ett naturligt sätt. Man säger att $(A \times B) \times C$ och $A \times (B \times C)$ är *(naturligt) isomorfa* och avbildningen som identifierar $((a, b), c)$ med $(a, (b, c))$ kallas en *(naturlig) isomorfism.*

Av detta skäl kommer vi att tillåta oss att skriva $A \times B \times C$ och elementen i denna mängd kommer vi att beteckna (a, b, c). Generaliseringen till godtyckligt (men fortfarande ändligt) många faktorer görs helt analogt.

Definition 6.9. Låt U och V vara godtyckliga vektorrum. En avbildning $T : U \times V \to \mathbf{R}$ sägs vara en *multilinjär avbildning* om det för alla $\mathbf{u}_1, \mathbf{u}_2 \in U$, för alla $\mathbf{v}_1, \mathbf{v}_2 \in V$ och för alla rella tal, λ_1, λ_2, μ_1 och μ_2 gäller att

$$\begin{aligned} T(\lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2, \mathbf{v}_1) &= \lambda_1 T(\mathbf{u}_1, \mathbf{v}_1) + \lambda_2 T(\mathbf{u}_2, \mathbf{v}_1) \\ T(\mathbf{u}_1, \mu_1 \mathbf{v}_1 + \mu_2 \mathbf{v}_2) &= \mu_1 T(\mathbf{u}_1, \mathbf{v}_1) + \mu_2 T(\mathbf{u}_1, \mathbf{v}_2), \end{aligned}$$

d v s om T är linjär i båda sina argument.

Multilinjära avbildningar med fler än två argument definieras analogt. En multilinjär avbildning med exakt två argument sägs då ofta vara *bilinjär*.

En direkt konsekvens av denna definition är att

$$\begin{aligned}T(\lambda_1\mathbf{u}_1+\lambda_2\mathbf{u}_2,\mu_1\mathbf{v}_1+\mu_2\mathbf{v}_2) &= \lambda_1 T(\mathbf{u}_1,\mu_1\mathbf{v}_1+\mu_2\mathbf{v}_2)+\lambda_2 T(\mathbf{u}_2,\mu_1\mathbf{v}_1+\mu_2\mathbf{v}_2)\\ &= \lambda_1\mu_1 T(\mathbf{u}_1,\mathbf{v}_1)+\lambda_1\mu_2 T(\mathbf{u}_1,\mathbf{v}_2)+\lambda_2\mu_1 T(\mathbf{u}_2,\mathbf{v}_1)+\lambda_2\mu_2 T(\mathbf{u}_2,\mathbf{v}_2).\end{aligned}$$

På samma sätt fås generaliseringen

$$T\left(\sum_{i=1}^{l}\lambda_i\mathbf{u}_i,\sum_{j=1}^{m}\mu_j\mathbf{v}_j\right)=\sum_{i=1}^{l}\sum_{j=1}^{m}\lambda_i\mu_j T(\mathbf{u}_i,\mathbf{v}_j).$$

I avsnitt 6.2 definierade vi kovektorer som linjära avbildningar från tangentrummet till de reella talen, och vi såg att en kovektors komponenter i dualbasen $(dx^i)_{i=0}^{n-1}$ uppfyllde transformationslagen (4.8), d v s i terminologin från avsnitt 4.4 utgör en kovektors komponenter en kovariant tensor.

Vi vill nu ge en liknande definition av en allmän tensor av typen (s,t), d v s vi vill definiera en tensor på ett sådant sätt att dess komponenter i en lämplig bas uppfyller transformationslagen (4.12). För att inte drunkna i otympliga indexformler söker vi först efter en lämplig definition av en tensor av typ $(1,1)$.

Antag därför att $T^i{}_j$ är en uppsättning tal som uppfyller transformationslagen (4.11) i någon punkt $p\in M$ och låt $\alpha=\alpha_i dx^i$ och $\mathbf{v}=v^j\frac{\partial}{\partial x^j}$ vara en godtycklig kovektor respektive vektor i p.

Bilda uttrycket

$$T^i{}_j\alpha_i v^j.$$

Vad beror detta uttryck av? Det är klart att $T^i{}_j\alpha_i v^j$ beror av α_i och v^j och därmed på α och $\mathbf{v}$. Dessutom beror både $T^i{}_j$, α_i och v^j på vårt val av koordinatsystem. En ytterligare möjlighet är därför att $T^i{}_j\alpha_i v^j$ också beror på valet av koordinatsystem.

Låt nu $\bar{x}^i$ vara ett nytt, godtyckligt koordinatsystem och betrakta

$$\bar{T}^k{}_l\bar{\alpha}_k\bar{v}^l=T^i{}_j\frac{\partial\bar{x}^k}{\partial x^i}\frac{\partial x^j}{\partial\bar{x}^l}\alpha_m\frac{\partial x^m}{\partial\bar{x}^k}v^n\frac{\partial\bar{x}^l}{\partial x^n}=T^i{}_j\alpha_m v^n\frac{\partial x^m}{\partial x^i}\frac{\partial x^j}{\partial x^n}=T^i{}_j\alpha_m v^n\delta_i{}^m\delta_n{}^j=T^i{}_j\alpha_i v^j,$$

där vi använt kedjeregeln i andra likheten.

$T^i{}_j\alpha_i v^j$ beror alltså *ej* på vårt val av koordinatsystem och definierar därför en funktion T av enbart α och $\mathbf{v}$. Det är också lätt att se att denna funktion är linjär i både α och $\mathbf{v}$.

Detta resonemang visar att varje uppsättning tal $T^i{}_j$ som uppfyller transformationslagen (4.11) svarar mot en multilinjär avbildning $T:T_pM^*\times T_pM\to\mathbf{R}$ given av sambandet

$$T(\alpha,\mathbf{v})=T^i{}_j\alpha_i v^j.$$

Omvänt: Låt $T:T_pM^*\times T_pM\to\mathbf{R}$ vara godtycklig och sätt $T^i{}_j=T\left(dx^i,\frac{\partial}{\partial x^j}\right)$. Då är

$$T(\alpha,\mathbf{v})=T\left(\alpha_i dx^i,v^j\frac{\partial}{\partial x^j}\right)=\alpha_i v^j T\left(dx^i,\frac{\partial}{\partial x^j}\right)=T^i{}_j\alpha_i v^j,$$

och här måste $T^i{}_j$ uppfylla transformationslagen (4.11). Detta kan visas med ett resonemang snarlikt det i beviset av Sats 6.7 (Gör detta!).

Det råder alltså en $1-1$-motsvarighet mellan uppsättningar av tal $T^i{}_j$ som uppfyller transformationslagen (4.11) (d v s det vi tidigare kallade för en blandad tensor av typ $(1,1)$) och multilinjära avbildningar $T : T_pM^* \times T_pM \to \mathbf{R}$. Vi gör därför följande definition:

Definition 6.10. T är en *(blandad) tensor av typ* $(1,1)$ *i* $p \in M$ om $T : T_pM^* \times T_pM \to \mathbf{R}$ är multilinjär. Talen $T^i{}_j = T\left(dx^i, \frac{\partial}{\partial x^j}\right)$ kallas tensorns *komponenter* i koordinatsystemet x^i.

Då gäller också

SATS 6.11. *T är en tensor av typ $(1,1)$ i p om och endast om komponenterna $T^i{}_j$ definierade ovan uppfyller transformationslagen (4.11) och är sådana att*

$$T(\alpha, \mathbf{v}) = T^i{}_j \alpha_i v^j,$$

för alla $\alpha = \alpha_i dx^i \in T_pM^*$ och $\mathbf{v} = v^j \frac{\partial}{\partial x^j} \in T_pM$.

Bevis. Se ovanstående resonemang. □

Vi vill nu utvidga den abstrakta indexnotationen till tensorer av typ $(1,1)$. En sådan tensor tar enligt ovan två argument, en kovektor α och en vektor $\mathbf{v}$ och returnerar talet $T(\alpha, \mathbf{v})$. Ett sätt att tydliggöra detta synsätt är att istället för T skriva $T(\cdot,\cdot)$ där punkterna betyder att en kovektor kan stoppas in på den första punktens plats och att en vektor kan stoppas in på den andra.

Om vi börjar med att betrakta den andra punktens plats, där vi ska stoppa in en vektor för att få ut ett reellt tal, så är det en situation vi känner igen. En kovektor $\alpha = \alpha(\cdot)$ verkar ju på precis detta sätt, och med abstrakta index skrivs ju en sådan α_b, d v s med ett index nere. Det är därför naturligt att ersätta den andra punkten i $T(\cdot,\cdot)$ med ett abstrakt index nere.

Vi har också sett att en vektor (försedd med ett abstrakt index uppe) kan ses som en avbildning som avbildar kovektorer på reella tal. Det är därför lika naturligt att ersätta den första punkten i $T(\cdot,\cdot)$ med ett abstrakt index uppe.

Av skäl som kommer att framgå senare kräver vi också att olika bokstäver används för de båda abstrakta indexen. En tensor T av typ $(1,1)$ skrivs alltså $T^a{}_b$ (t ex) i den abstrakta indexnotationen.

Analogt med skrivsättet $\alpha(\mathbf{v}) = \alpha_a v^a$ skriver vi

$$T(\alpha, \mathbf{v}) = T^a{}_b \alpha_a v^b.$$

Vi upprepar alltså samma abstrakta index, fast på motsatt plats, på den kovektor resp. vektor som tensorn verkar på. Vi observerar också att enligt Sats 6.11 så gäller att

$$T(\alpha, \mathbf{v}) = T^a{}_b \alpha_a v^b = T^i{}_j \alpha_i v^j,$$

där $T^i{}_j = T\left(dx^i, \frac{\partial}{\partial x^j}\right)$. Även här ser alltså den abstrakta indexnotationen likadan ut som indexnotationen i tidigare kapitel.

Innan vi ger oss i kast med den allmänna definitionen av en tensor av typ (s,t) ska vi studera några ytterligare aspekter av en tensor av typ $(1,1)$. Låt alltså T vara en sådan tensor och låt α vara en godtycklig men fix kovektor och betrakta avbildningen $\tilde{T} : \mathbf{v} \mapsto T(\alpha, \mathbf{v})$.

$\tilde{T}$ är förstås linjär (eftersom T är bilinjär) och avbildar tangentvektorer $\mathbf{v} \in T_pM$ på reella tal. $\tilde{T}$ är alltså en kovektor för varje val av α. Man skriver ofta denna kovektor som $\tilde{T} = T(\alpha, \cdot)$.

Ett annat sätt att se på tensorn T är därför att se den som en linjär avbildning som avbildar kovektorn α på kovektorn $\tilde{T}$.

Uttryckt med abstrakta index är alltså $\tilde{T}_b$ en kovektor som avbildar vektorn v^b på talet

$$\tilde{T}(\mathbf{v}) = \tilde{T}_b v^b.$$

Men å andra sidan är ju

$$\tilde{T}(\mathbf{v}) = T(\alpha, \mathbf{v}) = T^a{}_b \alpha_a v^b,$$

så $\tilde{T}_b v^b = T^a{}_b \alpha_a v^b$ för alla v^b. Det är därför naturligt att införa beteckningen $\tilde{T}_b = T^a{}_b \alpha_a$ för $\tilde{T}_b$.

Vi har nu två olika beteckningar för kovektorn $\tilde{T}$, $T(\alpha, \cdot)$ resp. $T^a{}_b \alpha_a$ och det är instruktivt att jämföra dem. Att vi har fixerat första argumentet i T till α syns i den första beteckningen genom att vi stoppat in α på första platsen i parentesen och i den andra genom att första abstrakta indexet i T är samma som det abstrakta indexet på α. Att andra argumentet till T fortfarande inte är fixerat syns i den första beteckningen genom att denna plats är markerad med en punkt och i den andra genom att T:s andra abstrakta index b är fritt, d v s det finns inget annat index b med motsatt placering i uttrycket. Då indexet dessutom sitter nere ser vi att det är en vektor som kan stoppas in på denna plats, och inte en kovektor. Detta är inte lika tydligt i den första beteckningen.

Ett tredje sätt att se på en tensor T av typ $(1,1)$ är genom att istället fixera $\mathbf{v}$ och låta α vara fri. Då fås avbildningen $\hat{T} = T(\cdot, \mathbf{v})$ som avbildar en kovektor α på det reella talet $T(\alpha, \mathbf{v})$.

Minns att en tangentvektor $\mathbf{v}$ avbildar just kovektorer på reella tal genom (det något oegentligt skrivna) sambandet $\mathbf{v}(\alpha) = \alpha(\mathbf{v})$. Det följer att $\hat{T}$ är en tangentvektor och vi kan därför se tensorn T som en linjär avbildning som avbildar vektorn $\mathbf{v}$ på vektorn $\hat{T}$. Det är detta synsätt som ligger till grund för Exempel 4.12.

Med abstrakta index skriver vi $\hat{T}(\alpha) = \hat{T}^a \alpha_a = T^a{}_b \alpha_a v^b$ och $\hat{T}^a = T^a{}_b v^b$, helt analogt med $\tilde{T}$ ovan.

Sammanfattningsvis finns det alltså (minst?) tre ekvivalenta sätt att se på en tensor T av typ $(1,1)$:

(i) T är en linjär avbildning $T : T_pM^* \times T_pM \to \mathbf{R}$ sådan att $(\alpha, \mathbf{v}) \mapsto T(\alpha, \mathbf{v})$ eller, med abstrakta index, $(\alpha_a, v^b) \mapsto T^a{}_b \alpha_a v^b$.

(ii) T är en linjär avbildning $T : T_pM^* \to T_pM^*$ sådan att $\alpha \mapsto T(\alpha, \cdot) = \tilde{T}$ eller, med abstrakta index, $\alpha_a \mapsto T^a{}_b \alpha_a = \tilde{T}_b$.

(iii) T är en linjär avbildning $T : T_pM \to T_pM$ sådan att $\mathbf{v} \mapsto T(\cdot, \mathbf{v}) = \hat{T}$ eller, med abstrakta index, $v^b \mapsto T^a{}_b v^b = \hat{T}^a$.

Som ovan, låt $\alpha_a = \alpha_i (dx^i)_a$, $v^b = v^j \left(\frac{\partial}{\partial x^j}\right)^b$ och $T^i{}_j = T\left(dx^i, \frac{\partial}{\partial x^j}\right) = T^a{}_b (dx^i)_a \left(\frac{\partial}{\partial x^j}\right)^b$. Vi söker nu komponenterna till vektorn $\hat{T}^a$, d v s vi söker tal $\hat{T}^i$ så att $\hat{T}^a = \hat{T}^i \left(\frac{\partial}{\partial x^i}\right)^a$. Om vi låter vektorn $\hat{T}^a$ verka på dualbasvektorn $(dx^i)_a$ så fås

$$\hat{T}^k \left(\frac{\partial}{\partial x^k}\right)^a (dx^i)_a = \hat{T}^a (dx^i)_a = T^a{}_b (dx^i)_a v^b = v^j T^a{}_b (dx^i)_a \left(\frac{\partial}{\partial x^j}\right)^b = T^i{}_j v^j,$$

men å andra sidan är $\left(\frac{\partial}{\partial x^k}\right)^a (dx^i)_a = \delta_k{}^i$ enligt definitionen av dualbas, vilket ger

$$\hat{T}^k \left(\frac{\partial}{\partial x^k}\right)^a (dx^i)_a = \hat{T}^k \delta_k{}^i = \hat{T}^i.$$

Komponenterna $\hat{T}^i$ till $\hat{T}^a = T^a{}_b v^b$ ges alltså av

$$\hat{T}^i = T^i{}_j v^j,$$

så även här fås komponentformeln genom att helt enkelt byta ut de abstrakta indexen mot numeriska index. På samma sätt visas att komponenterna $\tilde{T}_j$ till $\tilde{T}_b$ ges av

$$\tilde{T}_j = T^i{}_j \alpha_i.$$

I ovanstående räkning använde vi Kroneckers delta $\delta_i{}^j$ som definierades i Definition 4.10. Då denna definition bara refererar till komponenterna hos Kroneckers delta ger vi nu en alternativ definition som vi sedan ska bevisa är ekvivalent med Definition 4.10. I denna definition använder vi oss av synsätt *(iii)* ovan.

Definition 6.12. Identitetsavbildningen $\delta : T_pM \to T_pM$ given av $\delta(\mathbf{v}) = \mathbf{v}$ (d v s $\delta_a{}^b v^a = v^b$) för alla $\mathbf{v} \in T_pM$ kallas *Kroneckers delta*.

Då gäller

SATS 6.13. *Kroneckers delta är en tensor av typ* $(1,1)$ *och dess komponenter* $\delta_i{}^j := \delta_a{}^b \left(\frac{\partial}{\partial x^i}\right)^a (dx^j)_b$ *ges av*

$$\delta_i{}^j = \begin{cases} 1 & \textit{om } i = j \\ 0 & \textit{om } i \neq j \end{cases}.$$

Bevis. Första delen följer direkt av att Kroneckers delta är en linjär avbildning på T_pM samt resonemanget ovan, kallat synsätt *(iii)*.

Vi noterar att vi i detta kapitel redan använt beteckningen

$$\delta_i{}^j = \begin{cases} 1 & \text{om } i = j \\ 0 & \text{om } i \neq j \end{cases}. \tag{6.8}$$

Låt oss därför temporärt kalla komponenterna till tensorn $\delta_a{}^b$ för $d_i{}^j$. Vi ska då visa att $d_i{}^j = \delta_i{}^j$ där $\delta_i{}^j$ ges av ekvation (6.8).

Enligt definitionen av en tensors komponenter är

$$d_i{}^j = \delta_a{}^b \left(\frac{\partial}{\partial x^i}\right)^a (dx^j)_b = \left(\frac{\partial}{\partial x^i}\right)^b (dx^j)_b = \delta_i{}^j,$$

där andra likheten följer av att $\delta_a{}^b v^a = v^b$ för alla vektorer v^a och den tredje likheten följer av ekvation (6.7) eller (ekvivalent) ekvation (6.1). □

Vår nya definition av Kroneckers delta är alltså helt ekvivalent med den gamla.

Det är nu dags att generalisera ovanstående definition av en tensor av typ $(1,1)$ till en definition av en allmän tensor av typ (s,t). Med ett resonemang analogt med det ovanstående kan vi motivera följande definition:

Definition 6.14. T är en *tensor av typ* (s,t) *i* $p \in M$ om

$$T : \underbrace{T_pM^* \times \cdots \times T_pM^*}_{s \text{ st}} \times \underbrace{T_pM \times \cdots \times T_pM}_{t \text{ st}} \to \mathbf{R}$$

är en multilinjär avbildning. Talen

$$T^{i_1 i_2 \ldots i_s}{}_{j_1 j_2 \ldots j_t} = T\left(dx^{i_1}, dx^{i_2}, \ldots, dx^{i_s}, \frac{\partial}{\partial x^{j_1}}, \frac{\partial}{\partial x^{j_2}}, \ldots, \frac{\partial}{\partial x^{j_t}}\right) \tag{6.9}$$

kallas tensorns *komponenter* i koordinatsystemet x^i.

Liksom tidigare är det här inte nödvändigt att T:s argument kommer i precis ovanstående ordning med alla kovektorer först och alla vektorer sist. Som i avsnitt 4.5 tillåter vi omordningar av argumenten, men liksom i avsnitt 4.5 nöjer vi oss med att skriva upp ovanstående formulering.

Observera att en kovektor är definierad som en linjär avbildning från T_pM till $\mathbf{R}$ och att en tangentvektor kan ses som en linjär avbildning från T_pM^* till $\mathbf{R}$. Kovektorer och vektorer är alltså specialfall av tensorer (av typ $(0,1)$ resp. $(1,0)$) enligt denna definition. Detta rättfärdigar att vi kallade dem tensorer i förra kapitlet.

Det gäller också att

SATS 6.15. *T är en tensor av typ (s,t) i p om och endast om komponenterna $T^{i_1 i_2 \ldots i_s}{}_{j_1 j_2 \ldots j_t}$ definierade ovan uppfyller transformationslagen (4.12) och är sådana att*

$$T(\alpha^1, \alpha^2, \ldots, \alpha^s, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t) = T^{i_1 i_2 \ldots i_s}{}_{j_1 j_2 \ldots j_t} (\alpha^1)_{i_1} (\alpha^2)_{i_2} \ldots (\alpha^s)_{i_s} (v_1)^{j_1} (v_2)^{j_2} \ldots (v_t)^{j_t}$$

för alla $\alpha^1 = (\alpha^1)_{i_1} dx^{i_1}, \ldots, \alpha^s = (\alpha^s)_{i_s} dx^{i_s} \in T_pM^*$, $\mathbf{v_1} = (v_1)^{j_1} \frac{\partial}{\partial x^{j_1}}, \ldots, \mathbf{v_t} = (v_t)^{j_t} \frac{\partial}{\partial x^{j_t}} \in T_pM$.

Bevis. Analogt med beviset för Sats 6.11. □

Den abstrakta indexnotationen fungerar likadant som för tensorer av typ $(1,1)$ och precis som dessa kan en tensor av typ (s,t) ses som flera avbildningar i en.

Som en illustration betraktar vi en tensor T av typ $(2,1)$. Vi låter alltså T vara en multilinjär avbildning $T : T_pM^* \times T_pM^* \times T_pM \to \mathbf{R}$ och med abstrakta index skrivs T som $T^{ab}{}_c$. Om vi fixerar kovektorerna α_a och β_b fås avbildningen $T^{ab}{}_c\alpha_a\beta_b$ d v s avbildningen som skrivs $T(\alpha,\beta,\cdot)$ utan abstrakta index. Denna avbildning är linjär och avbildar vektorer på reella tal, d v s den är en kovektor vilket också syns genom att $T^{ab}{}_c\alpha_a\beta_b$ har ett fritt index c och detta index sitter nere. Vi kan alltså se T som en linjär avbildning $T : T_pM^* \times T_pM^* \to T_pM^*$ genom att $(\alpha_a,\beta_b) \mapsto T^{ab}{}_c\alpha_a\beta_b$.

Vi kan också fixera v^c och få avbildningen $T^{ab}{}_cv^c$ som är en tensor av typ $(2,0)$. Ett annat sätt att se på T är alltså som en linjär avbildning som avbildar tangentvektorer på tensorer av typ $(2,0)$. Många andra ekvivalenta möjligheter finns också, förstås.

Låt T vara en tensor av typ $(0,2)$ (t ex). Med abstrakta index skriver vi

$$T(\mathbf{u},\mathbf{v}) = T_{ab}u^av^b.$$

Observera att det abstrakta indexet ^{a} på u^a talar om att u^a ska stoppas in på det första argumentets plats i tensorn T. På samma sätt ser vi på indexet hos v^b att denna vektor ska stoppas in på andra argumentets plats. Det kan alltså inte leda till några missförstånd om vi kastar om ordningen mellan u^a och v^b och skriver

$$T(\mathbf{u},\mathbf{v}) = T_{ab}u^av^b = T_{ab}v^bu^a.$$

Som jämförelse skriver vi också upp $T(\mathbf{v},\mathbf{u})$ med abstrakta index:

$$T(\mathbf{v},\mathbf{u}) = T_{ab}v^au^b \neq T_{ab}v^bu^a = T(\mathbf{u},\mathbf{v}).$$

I Definition 5.4 definierade vi symmetriska respektive antisymmetriska tensorer av typ $(0,2)$ utifrån deras komponenters egenskaper. En alternativ definition som inte refererar till tensorns komponenter är följande:

Definition 6.16. Låt T vara en tensor av typ $(0,2)$. Om $T(\mathbf{u},\mathbf{v}) = T(\mathbf{v},\mathbf{u})$ för alla $\mathbf{u},\mathbf{v} \in T_pM$ sägs T vara *symmetrisk* och om istället $T(\mathbf{u},\mathbf{v}) = -T(\mathbf{v},\mathbf{u})$ sägs T vara *antisymmetrisk* (eller *skevsymmetrisk*).

Med abstrakta index skrivs en tensor T av typ $(0,2)$ som T_{ab} och att T_{ab} är symmetrisk innebär då att

$$T_{ab}u^av^b = T(\mathbf{u},\mathbf{v}) = T(\mathbf{v},\mathbf{u}) = T_{ab}v^au^b = T_{ba}v^bu^a = T_{ba}u^av^b$$

för alla vektorer u^a och v^b. Här följer näst sista likheten av att index som inte är fria får döpas om.

T är alltså symmetrisk om och endast om $T_{ab} = T_{ba}$ och på samma sätt visas att T är antisymmetrisk om och endast om $T_{ab} = -T_{ba}$.

Vi definierar nu symmetriska och antisymmetriska delen av T på samma sätt som i Definition 5.5

Definition 6.17. Den symmetriska tensorn

$$T_{(ab)} = \frac{1}{2}\left(T_{ab} + T_{ba}\right)$$

kallas den *symmetriska delen* av T_{ab} och den antisymmetriska tensorn

$$T_{[ab]} = \frac{1}{2}(T_{ab} - T_{ba})$$

kallas den *antisymmetriska delen* av T_{ab}.

och, precis som tidigare, är $T_{ab} = T_{(ab)} + T_{[ab]}$. Denna uppdelning är dessutom entydig:

SATS 6.18. *Om $T_{ab} = S_{ab}+A_{ab}$ där S_{ab} är symmetrisk och A_{ab} är antisymmetrisk så är $S_{ab} = T_{(ab)}$ och $A_{ab} = T_{[ab]}$.*

Bevis. Snarlikt komponentbeviset av Sats 5.6. Genomför själv detaljerna som övning. □

Låt oss betrakta den allmänna definitionen av en tensors komponenter, ekvation (6.9). Med abstrakta index skrivs denna

$$T^{i_1 i_2 \ldots i_s}{}_{j_1 j_2 \ldots j_t} = T^{a_1 a_2 \ldots a_s}{}_{b_1 b_2 \ldots b_t}(dx^{i_1})_{a_1}(dx^{i_2})_{a_2} \ldots (dx^{i_s})_{a_s} \left(\frac{\partial}{\partial x^{j_1}}\right)^{b_1} \left(\frac{\partial}{\partial x^{j_2}}\right)^{b_2} \cdots \left(\frac{\partial}{\partial x^{j_t}}\right)^{b_t} \tag{6.10}$$

Vid första anblicken verkar det kanske som om dessa komponenter är något väsensskilt från komponenterna v^j av en kovektor v^a definierade av sambandet

$$v^a = v^j \left(\frac{\partial}{\partial x^j}\right)^a. \tag{6.11}$$

Så är dock inte fallet. För att se detta, låt v^a verka på en godtycklig dualbasvektor $(dx^i)_a$:

$$v^a (dx^i)_a = v^j \left(\frac{\partial}{\partial x^j}\right)^a (dx^i)_a = v^j \delta_j{}^i = v^i.$$

Genom att låta en godtycklig kovektor α_a verka på en godtycklig koordinatbasvektor $\left(\frac{\partial}{\partial x^i}\right)^a$ kan man också visa (se räkningarna i beviset av Sats 6.4) att $\alpha_i = \alpha_a \left(\frac{\partial}{\partial x^i}\right)^a$. Ekvation (6.9) för en tensors komponenter gäller alltså även för tangentvektorer och kovektorer.

Efter denna insikt är det naturligt att ställa sig frågan: Om nu alla tensorers komponenter (inklusive vektorers och kovektorers komponenter) ges av ekvation (6.9), är det också så att en motsvarighet till ekvation (6.11) gäller för en allmän tensor? M a o, kan vi konstruera några baselement från koordinaterna x^i så att en allmän tensor ges av en linjärkombination av dessa baselement, med tensorns komponenter som koefficienter?

Svaret på dessa frågor är ja, vilket vi kommer att visa i avsnitt 6.6, men innan dess ska vi se lite grand på tensorfält.

6.5 Tensorfält

Tidigare har vi definierat ett vektorfält som en funktion som till en punkt p ordnar en tangentvektor $\mathbf{v}(p)$ i p så att $\mathbf{v}$:s komponenter v^i blir $\mathcal{C}^\infty$ funktioner. Allmänna tensorfält har definierats på liknande sätt.

I detta avsnitt ska vi se på en mer koordinatfri definition av vektorfält och andra tensorfält och visa att denna definition är ekvivalent med den gamla.

Vi noterar först att om $f \in \mathcal{C}^\infty$ och $v^i \in \mathcal{C}^\infty$ så är också $v^i \frac{\partial f}{\partial x^i} \in \mathcal{C}^\infty$. Omvänt: Om $v^i \frac{\partial f}{\partial x^i} \in \mathcal{C}^\infty$ för alla $f \in \mathcal{C}^\infty$ kan vi speciellt sätta $f = x^j$ och få $v^j = v^i {\delta_i}^j = v^i \frac{\partial x^j}{\partial x^i} \in \mathcal{C}^\infty$. Vi definierar därför

Definition 6.19. Antag att $\mathbf{v} = \mathbf{v}(p) \in T_pM$ för alla p. $\mathbf{v}$ sägs då vara ett *vektorfält* på M om $\mathbf{v}(f) \in \mathcal{C}^\infty$ för alla $f \in \mathcal{C}^\infty$.

Det följer direkt av ovanstående resonemang att denna definition är ekvivalent med den gamla. Observera också att denna definition funkar oavsett hur många koordinatsystem som behövs för att täcka $\mathbf{v}$:s definitionsmängd (jämför diskussionen i avsnitt 4.6).

Genom att bygga vidare på denna definition kan vi nu ge en koordinatfri definition av ett kovektorfält.

Definition 6.20. Antag att $\alpha = \alpha(p) \in T_pM^*$ för alla p. α sägs då vara ett *kovektorfält* på M om $\alpha(\mathbf{v}) \in \mathcal{C}^\infty$ för alla vektorfält $\mathbf{v}$.

Abstrakta indexnotationen ger följande ekvivalenta och mer koncisa formulering:

Definition 6.21. $\alpha_a = \alpha_a(p)$ är ett *kovektorfält* på M om $\alpha_a v^a \in \mathcal{C}^\infty$ för alla vektorfält v^a.

Tensorfält av typ (s, t) definieras helt analogt

Definition 6.22. ${T^{a_1\ldots a_s}}_{b_1\ldots b_t} = {T^{a_1\ldots a_s}}_{b_1\ldots b_t}(p)$ är ett *tensorfält av typ* (s, t) om

$${T^{a_1\ldots a_s}}_{b_1\ldots b_t} \alpha^1_{a_1} \ldots \alpha^s_{a_s} v^{b_1}_1 \ldots v^{b_t}_t \in \mathcal{C}^\infty$$

för alla kovektorfält $\alpha^1_{a_1}, \ldots, \alpha^s_{a_s}$ och alla vektorfält $v^{b_1}_1, \ldots, v^{b_t}_t$.

Vi ska nu ge en användbar karakterisering av tensorfält och för att inte tynga ner framställningen för mycket ska vi hålla oss till tensorfält av förhållandevis enkel typ.

Låt α, $\mathbf{v}$ och f vara ett godtyckligt kovektorfält, vektorfält respektive en godtycklig snäll funktion. Låt $p \in M$ vara godtycklig. Då är

$$\alpha(f\mathbf{v})|_p = \alpha(p)\,(f(p)\mathbf{v}(p)) = f(p)\alpha(p)\,(\mathbf{v}(p)) = f\ \alpha(\mathbf{v})|_p\,,$$

så $\alpha(f\mathbf{v}) = f\alpha(\mathbf{v})$ eftersom p var godtycklig. Vidare, om $\mathbf{u}$ är ett godtyckligt vektorfält är $\alpha(\mathbf{u}+\mathbf{v}) = \alpha(\mathbf{u}) + \alpha(\mathbf{v})$. En avbildning α som har dessa båda egenskaper sägs vara $\mathcal{C}^\infty$*-linjär* så det följer att om α är ett kovektorfält så är α också $\mathcal{C}^\infty$-linjär.

Hur är det då med omvändningen? Om α är $\mathcal{C}^\infty$-linjär följer det då att α är ett kovektorfält? Svaret på denna fråga är ja, och detta är ämnet för nästa sats.

SATS 6.23. *Låt α vara en avbildning som avbildar vektorfält på $\mathcal{C}^\infty$ funktioner. Då är α ett kovektorfält om och endast om α är $\mathcal{C}^\infty$-linjär.*

Bevis. Ena riktningen är redan klar så antag att α är $\mathcal{C}^\infty$-linjär. α är då självklart linjär i $\mathbf{v}$ så det som återstår att visa är att $\alpha(\mathbf{v})|_p$ endast beror på $\mathbf{v}(p)$ och inte på några andra av $\mathbf{v}$:s värden.

Låt därför p vara en godtycklig punkt och låt $\mathbf{u}$ och $\mathbf{v}$ vara godtyckliga vektorfält sådana att $\mathbf{u}(p) = \mathbf{v}(p)$. Vi måste då bevisa att $\alpha(\mathbf{u})|_p = \alpha(\mathbf{v})|_p$ eller, med $\mathbf{w} = \mathbf{u} - \mathbf{v}$, att $\alpha(\mathbf{w})|_p = 0$. Inför koordinater x^i i en omgivning U av p.

Låt w^i vara $\mathbf{w}$:s komponenter i detta koordinatsystem. Observera att nu är w^i inte längre tal utan $\mathcal{C}^\infty$ funktioner av koordinaterna x^j och då $\alpha(\mathbf{w})|_p = 0$ följer att alla $w^i = 0$ i p.

Då α är $\mathcal{C}^\infty$-linjär följer att

$$\alpha(\mathbf{w})|_p = \alpha\left(w^i \frac{\partial}{\partial x^i}\right)\bigg|_p = \left(w^i \alpha\left(\frac{\partial}{\partial x^i}\right)\right)\bigg|_p = 0.$$

Det följer att $\alpha(\mathbf{v})|_p$ endast beror på $\mathbf{v}(p)$ och eftersom p var godtycklig är α därmed ett kovektorfält. □

Anmärkning: Beviset ovan är inte riktigt vattentätt. Det visar att $\alpha(\mathbf{v})|_p$ inte beror på $\mathbf{v}(q)$ om $q \neq p$ och $q \in U$. Detta innebär att beviset fungerar i fallet att hela M kan täckas med ett enda koordinatsystem, d v s om vi kan ta $U = M$. Om detta inte går måste resonemanget ovan modifieras och för att inte tynga ner framställningen för mycket ger vi här bara huvuddragen av beviset i det allmänna fallet.

Med *stödet* till en funktion χ (betecknas $\operatorname{supp}(\chi)$) menas slutna höljet av den mängd där $\chi \neq 0$. Med beteckningar som ovan kan man visa att det finns en $\mathcal{C}^\infty$ funktion χ med $\operatorname{supp}(\chi) \subset U$ och $\chi = 1$ i en omgivning av p.

Observera att om f är en $\mathcal{C}^\infty$ funktion på M så är även χf det och dessutom är $\chi f = f$ i en omgivning av p och $\chi f = 0$ utanför U. Vi kommer dessutom att tolka χf som 0 utanför U även om f bara är definierad på U. Samma sak om χ multipliceras med ett vektorfält som bara är definierat på U.

Detta ger att

$$\alpha(\mathbf{w})|_p = \left(\chi^2 \alpha(\mathbf{w})\right)\big|_p = \alpha\left(\chi^2 \mathbf{w}\right)\big|_p = \alpha\left(\chi w^i \chi \frac{\partial}{\partial x^i}\right)\bigg|_p = \left(\chi w^i \alpha\left(\chi \frac{\partial}{\partial x^i}\right)\right)\bigg|_p = 0,$$

återigen med beteckningar som ovan. Detta bevisar det allmänna fallet. □

Detta resultat går utan problem att generalisera till andra typer av tensorfält:

SATS 6.24. *Låt T avbilda ett godtyckligt antal vektorfält och kovektorfält på reellvärda, $\mathcal{C}^\infty$ funktioner. Då är T ett tensorfält om och endast om T är $\mathcal{C}^\infty$-linjär i vart och ett av sina argument.*

För allas trevnad avstår vi från att ge ett fullständigt bevis av detta.

6.6 Tensorprodukten

Innan vi kan besvara frågorna i slutet på avsnitt 6.4 behöver vi införa ett nytt begrepp, *tensorprodukten* av två tensorer och precis som i förra avsnittet kommer vi att studera ett någorlunda enkelt fall i detalj och sedan skissera hur generaliseringen till allmänna tensorer ser ut. Närmare bestämt studerar vi först tensorprodukten av två kovektorer.

Definition 6.25. Låt $\alpha, \beta \in T_pM^*$. *Tensorprodukten* $\alpha \otimes \beta$ av α och β är den tensor av typ $(0,2)$ som ges av

$$\alpha \otimes \beta(\mathbf{u}, \mathbf{v}) = \alpha(\mathbf{u})\beta(\mathbf{v})$$

för alla $\mathbf{u}, \mathbf{v} \in T_pM$.

Att tensorprodukten verkligen är en tensor följer direkt av att högerledet $\alpha(\mathbf{u})\beta(\mathbf{v})$ är linjärt i både $\mathbf{u}$ och $\mathbf{v}$. Observera att tensorprodukten *inte* är kommutativ eftersom

$$\alpha \otimes \beta(\mathbf{u}, \mathbf{v}) = \alpha(\mathbf{u})\beta(\mathbf{v}) \neq \beta(\mathbf{u})\alpha(\mathbf{v}) = \beta \otimes \alpha(\mathbf{u}, \mathbf{v})$$

i allmänhet.

Vi ska nu hitta ett bra sätt att beteckna tensorprodukten i den abstrakta indexnotationen och det har ju tidigare varit ett framgångsrecept att kopiera den indexnotation vi använder för att beteckna tensorns komponenter. Låt därför α_i och β_j vara komponenterna till kovektorerna α resp. β så att $\alpha = \alpha_i dx^i$ och $\beta = \beta_j dx^j$. Enligt ekvation (6.9) gäller det att $\alpha_i = \alpha\left(\frac{\partial}{\partial x^i}\right)$ och $\beta_j = \beta\left(\frac{\partial}{\partial x^j}\right)$. Enligt samma ekvation får $\alpha \otimes \beta$ komponenterna

$$(\alpha \otimes \beta)_{ij} = \alpha \otimes \beta\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \alpha\left(\frac{\partial}{\partial x^i}\right)\beta\left(\frac{\partial}{\partial x^j}\right) = \alpha_i\beta_j.$$

I abstrakta indexnotationen skrivs kovektorerna α och β med varsitt index nere, t ex α_a och β_b medan $\alpha \otimes \beta$ ska förses med två abstrakta index nere (varför?). En första, provisorisk, beteckning blir därför $(\alpha \otimes \beta)_{ab}$. Vi upprepar alltså de abstrakta indexen på α_a och β_b med indexens inbördes ordning bibehållen. P g a ovanstående komponentformel är det dock naturligt att förenkla denna beteckning till

$$(\alpha \otimes \beta)_{ab} = \alpha_a\beta_b.$$

I abstrakta indexnotationen skriver vi alltså tensorprodukten på samma sätt som om det rörde sig om vanlig multiplikation av tal. Vi noterar också att definitionen av tensorprodukt blir

$$(\alpha \otimes \beta)_{ab}u^a v^b = \alpha_a u^a \beta_b v^b.$$

Observera att ovanstående komponentformel också kan skrivas

$$(\alpha \otimes \beta)_{ij} = \alpha_i\beta_j = \beta_j\alpha_i,$$

ty i detta fall är det ju verkligen en vanlig multiplikation av tal som står i högra ledet. Om vi ska vara konsekventa med att låta abstrakta indexnotationen imitera våra komponentformler måste vi alltså låta

$$\alpha_a\beta_b = \beta_b\alpha_a.$$

Av samma skäl måste vi sätta

$$(\alpha \otimes \beta)_{ab} u^a v^b = \alpha_a u^a \beta_b v^b = \alpha_a \beta_b u^a v^b = u^a \beta_b \alpha_a v^b = \dots$$

d v s i den abstrakta indexnotationen spelar faktorernas ordning i en tensorprodukt ingen roll.

Anmärkning: Efter detta uttalande kan det tyckas vara dags att dra i nödbromsen! Har vi just härlett en motsägelse? Å ena sidan spelar faktorernas ordning ingen roll när vi skriver upp en tensorprodukt i den abstrakta indexnotationen, men å andra sidan visade vi ju ovan att tensorprodukten *inte* var kommutativ. Hur går detta ihop sig?

Lugn i stormen, allt är i sin ordning! Faktum är nämligen att i den abstrakta indexnotationen har inte faktorernas inbördes ordning något med kommutativitet att göra. Denna information är istället lagrad i de abstrakta indexen.

Låt oss än en gång skriva upp definitionen av $\alpha \otimes \beta$ med abstrakta index:

$$(\alpha \otimes \beta)_{ab} u^a v^b = \alpha_a u^a \beta_b v^b = \alpha_a \beta_b u^a v^b.$$

Definitionen av $\beta \otimes \alpha$ lyder däremot $\beta \otimes \alpha(\mathbf{u}, \mathbf{v}) = \beta(\mathbf{u})\alpha(\mathbf{v})$, eller med abstrakta index

$$(\beta \otimes \alpha)_{ab} u^a v^b = \beta_a u^a \alpha_b v^b = \beta_a \alpha_b u^a v^b$$

för alla vektorer u^a och v^b. Det följer att

$$(\beta \otimes \alpha)_{ab} = \beta_a \alpha_b \neq \beta_b \alpha_a = \alpha_a \beta_b = (\alpha \otimes \beta)_{ab}.$$

Här finns alltså ingen motsägelse eftersom omkastning av faktorernas ordning i en tensorprodukt svarar mot att endast motsvarande *index* byter plats i den abstrakta indexnotationen.

Tensorprodukten är alltså inte kommutativ oavsett vilken notation vi använder.

□

Sedan tidigare har vi infört T_pM^a som beteckning för en kopia av T_pM i vilken alla element förses med ett abstrakt index a där uppe. På samma sätt var ju T_pM_b en kopia av T_pM^* i vilken alla element förses med ett abstrakt index b nere.

Låt oss generalisera denna beteckning genom att låta T_pM försedd med någon indexkombination (t ex T_pM_{ab}) stå för vektorrummet bestående av alla tensorer som har samma kombination av abstrakta index. T_pM_{ab} betyder då mängden av alla tensorer T_{ab} av typen $(0, 2)$. Speciellt är då $(\alpha \otimes \beta)_{ab} = \alpha_a \beta_b \in T_pM_{ab}$.

Med hjälp av tensorprodukten och vårt koordinatsystem x^i kan vi konstruera en bas för T_pM_{ab} och vi kommer på köpet att se att komponenterna T_{ij} av en tensor $T_{ab} \in T_pM_{ab}$ spelar samma roll som komponenterna till en vektor eller kovektor; de dyker nämligen upp som koefficienter när T_{ab} skrivs som en linjärkombination av baselementen. Följande sats gäller nämligen:

SATS 6.26. *Mängden* $B = \left\{(dx^i \otimes dx^j)_{ab} = (dx^i)_a (dx^j)_b, i, j = 0, \dots, n-1\right\}$ *utgör en bas för* T_pM_{ab}*. Vidare kan en godtycklig tensor* $T_{ab} \in T_pM_{ab}$ *skrivas*

$$T_{ab} = T_{ij} (dx^i)_a (dx^j)_b$$

där T_{ij} är T_{ab}:s komponenter givna av ekvation (6.10), d v s

$$T_{ij} = T_{ab}\left(\frac{\partial}{\partial x^i}\right)^a\left(\frac{\partial}{\partial x^j}\right)^b.$$

Bevis. Vi visar först att B är linjärt oberoende. Ansätt därför en linjär relation $\lambda_{kl}(dx^k)_a(dx^l)_b = 0$ där 0 i högerledet betyder den tensor av typen $(0,2)$ som avbildar varje par av tangentvektorer på talet 0. Låter vi denna tensor verka på den i:te resp. j:te koordinatbasvektorn fås

$$0 = \lambda_{kl}(dx^k)_a(dx^l)_b\left(\frac{\partial}{\partial x^i}\right)^a\left(\frac{\partial}{\partial x^j}\right)^b = \lambda_{kl}(dx^k)_a\left(\frac{\partial}{\partial x^i}\right)^a(dx^l)_b\left(\frac{\partial}{\partial x^j}\right)^b = \lambda_{kl}\delta_i{}^k\delta_j{}^l = \lambda_{ij}$$

för alla i och j vilket visar att B är linjärt oberoende.

Låt nu $u^a, v^a \in T_pM^a$ vara godtyckliga vektorer. Enligt Sats 6.15 är då $T_{ab}u^av^b = T_{ij}u^iv^j$ där u^i och v^i är u:s resp. v:s komponenter i koordinatbasen. Vi är alltså klara om vi kan visa att $T_{ij}(dx^i)_a(dx^j)_bu^av^b = T_{ij}u^iv^j$ (varför?).

Betrakta därför

$$T_{ij}(dx^i)_a(dx^j)_bu^av^b = T_{ij}(dx^i)_a(dx^j)_bu^k\left(\frac{\partial}{\partial x^k}\right)^a v^l\left(\frac{\partial}{\partial x^l}\right)^b = T_{ij}\delta_k{}^i\delta_l{}^ju^kv^l = T_{ij}u^iv^j,$$

och resultatet följer. □

Innan vi tittar på en helt allmän tensorprodukt ska vi se på tensorprodukten av en vektor och en kovektor. Minns att en tangentvektor $\mathbf{v}$ kan ses som en linjär avbildning från T_pM^* till $\mathbf{R}$ genom sambandet $\mathbf{v}(\alpha) = \alpha(\mathbf{v})$ eller, med abstrakta index, $v^a\alpha_a = \alpha_av^a$. Vi definierar

Definition 6.27. Låt $\mathbf{v} \in T_pM$, $\alpha \in T_pM^*$. *Tensorprodukten* $\mathbf{v} \otimes \alpha$ av $\mathbf{v}$ och α är den tensor av typ $(1,1)$ som ges av

$$\mathbf{v} \otimes \alpha(\beta, \mathbf{u}) = \mathbf{v}(\beta)\alpha(\mathbf{u})$$

för alla $\beta \in T_pM^*$ och $\mathbf{u} \in T_pM$.

Med abstrakta index skriver vi

$$(\mathbf{v} \otimes \alpha)^a{}_b\beta_au^b = v^a\beta_a\alpha_bu^b$$

och

$$(\mathbf{v} \otimes \alpha)^a{}_b = v^a\alpha_b.$$

Observera att även här är faktorernas ordning oväsentlig, men indexens placering är mycket viktig.

Liksom i förra fallet kan vi konstruera en bas för $T_pM^a{}_b$ med hjälp av tensorprodukter av element från koordinatbasen och dess dualbas:

SATS 6.28. *Mängden* $B = \left\{ \left(\frac{\partial}{\partial x^i} \otimes dx^j\right)^a{}_b = \left(\frac{\partial}{\partial x^i}\right)^a (dx^j)_b, i, j = 0, \dots, n-1 \right\}$ *utgör en bas för* $T_pM^a{}_b$*. Vidare kan en godtycklig tensor* $T^a{}_b \in T_pM^a{}_b$ *skrivas*

$$T^a{}_b = T^i{}_j \left(\frac{\partial}{\partial x^i}\right)^a (dx^j)_b$$

där $T^i{}_j$ *är* $T^a{}_b$*:s komponenter givna av ekvation (6.10), d v s*

$$T^i{}_j = T^a{}_b (dx^i)_a \left(\frac{\partial}{\partial x^j}\right)^b .$$

Bevis. Snarlikt beviset för Sats 6.26. Genomför själv detaljerna som övning. □

Givet en vektor $\mathbf{v}$ och en kovektor α kan vi alltså bilda talet $\alpha(\mathbf{v}) = \mathbf{v}(\alpha)$ och dessutom tensorn $\mathbf{v} \otimes \alpha$. Det är därför naturligt att betrakta en avbildning som avbildar $\mathbf{v} \otimes \alpha$ på $\alpha(\mathbf{v})$. Denna avbildning kallas *kontraktionen* eller *spåret* av tensorn $\mathbf{v} \otimes \alpha$ och betecknas Tr (efter engelskans 'trace'). Med abstrakta index är alltså Tr : $v^a\alpha_b \mapsto v^a\alpha_a = \alpha_a v^a$ och vi ser att Tr är linjär i både v^a och α_b.

Enligt Sats 6.28 kan alla tensorer i $T_pM^a{}_b$ skrivas som en summa av tensorer på formen $v^a\alpha_b$ och det är därför frestande att utvidga definitionen av Tr till hela $T_pM^a{}_b$ genom linearitet. Vi måste dock iaktta viss försiktighet eftersom en sådan framställning inte är entydig. Först måste vi alltså visa att olika framställningar ger samma värde på spåret.

För att göra detta frångår vi tillfälligt den abstrakta indexnotationen eftersom vi då får ett 'renare' resonemang.

Antag därför att en tensor T av typ $(1, 1)$ kan skrivas

$$T = \sum_{k=1}^{r} \mathbf{v}_k \otimes \alpha^k \quad \text{resp.} \quad T = \sum_{l=1}^{q} \mathbf{u}_l \otimes \beta^l,$$

för vektorer $\mathbf{v}_k$, $\mathbf{u}_l$ och kovektorer α^k, β^l. Observera att vi inte använder summationskonventionen här eftersom summorna inte nödvändigtvis innehåller $n = \dim M$ termer. Vi föredrar därför att explicit skriva ut summatecknen. Den första likheten ger oss uttrycket $\sum_{k=1}^{r} \alpha^k(\mathbf{v}_k)$ för spåret av T och den andra ger oss uttrycket $\sum_{l=1}^{q} \beta^l(\mathbf{u}_l)$ för spåret av T, och vi behöver visa att dessa två uttryck är lika.

Vi studerar den första av ovanstående likheter. Låt $(v_k)^j$ och $(\alpha^k)_i$ vara komponenterna till $\mathbf{v}_k$ resp. α^k. Enligt ekvation (6.10) är då $(v_k)^j = \mathbf{v}_k(dx^j)$ och $(\alpha^k)_i = \alpha^k\left(\frac{\partial}{\partial x^i}\right)$. Om vi låter T verka på dx^i och $\frac{\partial}{\partial x^j}$ fås

$$T\left(dx^i, \frac{\partial}{\partial x^j}\right) = \sum_{k=1}^{r} \mathbf{v}_k \otimes \alpha^k \left(dx^i, \frac{\partial}{\partial x^j}\right) = \sum_{k=1}^{r} \mathbf{v}_k\left(dx^i\right) \alpha^k \left(\frac{\partial}{\partial x^j}\right) = \sum_{k=1}^{r} (v_k)^i (\alpha^k)_j,$$

och den andra likheten ger på samma sätt och med analoga beteckningar

$$T\left(dx^i, \frac{\partial}{\partial x^j}\right) = \sum_{l=1}^{q} (u_l)^i (\beta^l)_j,$$

varför

$$\sum_{k=1}^{r}(v_k)^i(\alpha^k)_j = \sum_{l=1}^{q}(u_l)^i(\beta^l)_j$$

för alla i och j. Sätt nu $j = i$ och summera över i så fås

$$\sum_{k=1}^{r}(v_k)^i(\alpha^k)_i = \sum_{l=1}^{q}(u_l)^i(\beta^l)_i \Leftrightarrow \sum_{k=1}^{p}\alpha^k(\mathbf{v}_k) = \sum_{l=1}^{q}\beta^l(\mathbf{u}_l).$$

Olika sätt att skriva en tensor T som en summa av tensorprodukter ger alltså samma värde på $\mathrm{Tr}(T)$. Följande definition är alltså meningsfull:

Definition 6.29. *Spåret* (eller *kontraktionen*), $\mathrm{Tr}(T)$, av en tensor T av typ $(1,1)$ definieras av att Tr är linjär och $\mathrm{Tr}(\mathbf{v}\otimes\alpha) = \alpha(\mathbf{v})$.

Med abstrakta index gäller alltså $\mathrm{Tr}(v^a\alpha_b) = v^a\alpha_a = \alpha_a v^a$. Det är därför naturligt att införa beteckningen $T^a{}_a$ för $\mathrm{Tr}(T^a{}_b)$.

Vi söker nu en komponentformel för $\mathrm{Tr}(T^a{}_b)$ så låt $T^i{}_j$ vara komponenterna till $T^a{}_b$. Enligt Sats 6.28 gäller då att

$$T^a{}_b = T^i{}_j\left(\frac{\partial}{\partial x^i}\right)^a(dx^j)_b.$$

Definitionen av $\mathrm{Tr}(T^a{}_b)$ samt linearitet ger då att

$$T^a{}_a = \mathrm{Tr}(T^a{}_b) = \mathrm{Tr}\left(T^i{}_j\left(\frac{\partial}{\partial x^i}\right)^a(dx^j)_b\right) = T^i{}_j\left(\frac{\partial}{\partial x^i}\right)^a(dx^j)_a = T^i{}_j\delta_i^j = T^i{}_i,$$

så även här imiterar den abstrakta indexnotationen utseendet på motsvarande komponentformel.

Nu är det dags att vi tar oss an det allmänna fallet. Först definierar vi tensorprodukten av två godtyckliga tensorer.

Definition 6.30. Låt S och T vara tensorer av typ (q,r) resp. (s,t). Då är *tensorprodukten*, $U = S\otimes T$, av S och T en tensor av typ $q+s, r+t$ definierad av

$$U(\alpha_1,\dots,\alpha_q,\mathbf{u}^1,\dots,\mathbf{u}^r,\beta_1,\dots,\beta_s,\mathbf{v}^1,\dots,\mathbf{v}^t) = S(\alpha_1,\dots,\alpha_q,\mathbf{u}^1,\dots,\mathbf{u}^r)T(\beta_1,\dots,\beta_s,\mathbf{v}^1,\dots,\mathbf{v}^t).$$

Generaliseringen till fler faktorer än två är uppenbar.

Motsvarigheten till Sats 6.26 och 6.28 gäller också:

SATS 6.31. *Mängden*

$$B = \left\{\left(\frac{\partial}{\partial x^{i_1}}\right)^{a_1}\cdots\left(\frac{\partial}{\partial x^{i_s}}\right)^{a_s}(dx^{j_1})_{b_1}\dots(dx^{j_t})_{b_t}, i_1,\dots,i_s, j_1\dots j_t = 0,\dots,n-1\right\}$$

utgör en bas för $T_pM^{a_1\dots a_s}{}_{b_1\dots b_t}$. Vidare kan en godtycklig tensor $T^{a_1\dots a_s}{}_{b_1\dots b_t} \in T_pM^{a_1\dots a_s}{}_{b_1\dots b_t}$ skrivas

$$T^{a_1\dots a_s}{}_{b_1\dots b_t} = T^{i_1\dots i_s}{}_{j_1\dots j_t}\left(\frac{\partial}{\partial x^{i_1}}\right)^{a_1}\cdots\left(\frac{\partial}{\partial x^{i_s}}\right)^{a_s}(dx^{j_1})_{b_1}\dots(dx^{j_t})_{b_t}$$

där $T^{i_1\ldots i_s}{}_{j_1\ldots j_t}$ *är* $T^{a_1\ldots a_s}{}_{b_1\ldots b_t}$*:s komponenter givna av ekvation (6.9), d v s*

$$T^{i_1\ldots i_s}{}_{j_1\ldots j_t} = T^{a_1\ldots a_s}{}_{b_1\ldots b_t}\left(\frac{\partial}{\partial x^{i_1}}\right)^{a_1}\ldots\left(\frac{\partial}{\partial x^{i_s}}\right)^{a_s}(dx^{j_1})_{b_1}\ldots(dx^{j_t})_{b_t}.$$

Kontraktioner av en tensor av godtycklig typ (s,t) fungerar på samma sätt som kontraktionen av en tensor av typ $(1,1)$, men saken kompliceras av att det går att kontrahera över vilka indexpar som helst, förutsatt att ena indexet sitter uppe och det andra nere. Av barmhärtighetsskäl avstår vi från att formulera några allmänna resultat och exemplifierar istället med en tensor $T^{ab}{}_c$ av typ $(2,1)$.

Betrakta först en tensor på formen $u^a v^b \alpha_c$. En sådan kan kontraheras på två olika sätt: Dels genom att kontrahera över indexen a och c och därmed bilda tensorn $u^a v^b \alpha_a$ och dels genom att kontrahera över indexen b och c och därmed bilda tensorn $u^a v^b \alpha_b$. Det finns alltså två olika sätt att kontrahera en tensor på denna form.

Enligt Sats 6.31 kan en godtycklig tensor $T^{ab}{}_c$ av typ $(2,1)$ skrivas som en summa av tensorer på formen $u^a v^b \alpha_c$ och vi utvidgar därför definitionen av kontraktion till allmänna tensorer genom linearitet.

Precis som för tensorer av typ $(1,1)$ är dock en sådan framställning inte entydig och vi måste därför visa att olika framställningar ger samma värde på kontraktionen av $T^{ab}{}_c$. Detta kan göras på liknande sätt som i $(1,1)$-fallet ovan och detaljerna lämnas därför som övning.

En godtycklig tensor $T^{ab}{}_c$ av typ $(2,1)$ kan alltså kontraheras på två olika sätt, dels över indexen a och c och dels över indexen b och c. De kontraherade tensorerna betecknas $T^{ab}{}_a$ respektive $T^{ab}{}_b$ och är i allmänhet olika.

Vi behöver också komponentformler för dessa kontraktioner. Låt därför $T^{ij}{}_k$ vara komponenterna till $T^{ab}{}_c$ i något koordinatsystem. Det bör inte komma som någon överraskning att följande gäller:

$$T^{ab}{}_a = T^{ij}{}_i \quad , \quad T^{ab}{}_b = T^{ij}{}_j.$$

Beviset för detta resultat är snarlikt beviset för motsvarande resultat för tensorer av typ $(1,1)$. Genomför själv detaljerna som övning.

6.7 Metriska tensorn

Vi är nu klara att ge en koordinatfri definition av en metrik.

Definition 6.32. Ett symmetriskt tensorfält g av typ $(0,2)$ (d v s $g(\mathbf{u},\mathbf{v}) = g(\mathbf{v},\mathbf{u})$ för alla vektorfält $\mathbf{u}$ och $\mathbf{v}$) är en *metrik* om g är icke-singulär i varje punkt, d v s om $g(\mathbf{u},\mathbf{v}) = 0$ för alla $\mathbf{v}$ om och endast om $\mathbf{u} = \mathbf{0}$. *Skalärprodukten* mellan $\mathbf{u}$ och $\mathbf{v}$ definieras $g(\mathbf{u},\mathbf{v})$ och vi säger att $\mathbf{u}$ och $\mathbf{v}$ är *ortogonala* om $g(\mathbf{u},\mathbf{v}) = 0$.

Denna definition kan förstås också översättas till den abstrakta indexnotationen:

Definition 6.33. Ett symmetriskt tensorfält $g_{ab} = g_{ba}$ är en *metrik* om g_{ab} är icke-singulär i varje punkt, d v s om $g_{ab}u^a v^b = 0$ för alla v^b om och endast om $u^a = 0$. *Skalärprodukten* mellan u^a och v^a definieras $g_{ab}u^a v^b$ och vi säger att u^a och v^a är *ortogonala* om $g_{ab}u^a v^b = 0$.

Observera att Definition 5.7 ser helt identisk ut med denna. Detta är förstås inte förvånande eftersom vi konstruerat den abstrakta indexnotationen så att den ska bli identisk med den numeriska indexnotation vi använde i kapitel 5.

Vi ska nu titta lite närmare på villkoret att metriken g är icke-singulär. Detta betyder enligt ovan att $g(\mathbf{u}, \mathbf{v}) = 0$ för alla $\mathbf{v}$ om och endast om $\mathbf{u} = \mathbf{0}$, d v s avbildningen $g(\mathbf{u}, \cdot) : T_pM \to \mathbf{R}$ (som är en kovektor (varför?)) är nollavbildningen om och endast om $\mathbf{u} = \mathbf{0}$. I abstrakta index är alltså $\tilde{u}_b := g_{ab}u^a = 0$ om och endast om $u^a = 0$.

Avbildningen $g_{ab} : T_pM^a \to T_pM_b$ som avbildar u^a på $\tilde{u}_b$ har alltså trivialt nollrum så enligt en sats från linjära algebran är denna avbildning injektiv och har en invers. Vidare är T_pM^a och T_pM_b ändligt-dimensionella och har samma dimension vilket ger att inversen blir definierad på hela T_pM_b. Denna invers blir en tensor av typ (2,0) (varför?) som kallas *den inversa metriken* och betecknas g^{ab}, så att $g^{ab}\tilde{u}_b = u^a$.

Metriken g_{ab} och den inversa metriken g^{ab} associerar alltså en vektor u^a med en entydigt bestämd kovektor $\tilde{u}_b$ och omvänt. Om M:s metrik är en gång för alla given är det därför vanligt att se på u^a och $\tilde{u}_b$ som olika aspekter av samma objekt och därför utelämna $\tilde{}$ och skriva

$$u_b = g_{ab}u^a \tag{6.12}$$

och

$$u^a = g^{ab}u_b. \tag{6.13}$$

Dessa båda formler brukar kallas sänkning respektive höjning av index (jämför Definition 5.12 och 5.13) och de generaliseras till allmänna tensorer som i avsnitt 5.4.

Vi noterar också att

$$\delta_a{}^b u^a = u^b = g^{bc}u_c = g^{bc}g_{ac}u^a$$

för alla vektorer u^a vilket ger att $g^{bc}g_{ac} = \delta_a{}^b$. I komponenter är alltså $g^{jk}g_{ik} = \delta_i{}^j$ d v s om vi ordnar komponenterna g_{ij} och g^{ij} i två symmetriska $n \times n$-matriser så blir dessa matriser varandras inverser.

Vi är nu klara att definiera begreppet ON-bas (jämför Definition 5.15):

Definition 6.34. $\{(e_i)^a\,,\ i = 0, 1, \ldots, n-1\}$ (där i alltså är ett numeriskt index och a är ett abstrakt index) är en ON-bas för T_pM det gäller att

$$g_{ab}(e_i)^a(e_j)^b = \begin{cases} \pm 1 & \text{om } i = j \\ 0 & \text{om } i \neq j \end{cases},$$

d v s basvektorerna är normerade och inbördes ortogonala.

Vi är nu nästan klara med det vi föresatte oss i avsnitt 6.1. Det enda som återstår är sista delen av avsnitt 5.4 efter Definition 5.15) och denna del går igenom nästan oförändrad om man bara tolkar om index från början av alfabetet som abstrakta index istället för numeriska. Läsaren uppmanas att gå tillbaka och läsa om denna del och tänka igenom de resonemang som förs samt tolka om alla formler och beräkningar i termer av abstrakta index.

Några ord på vägen innan du gör detta: Kom ihåg att abstrakta indexnotationen är konstruerad så att alla formler ska se helt identiska ut med de 'vanliga', numeriska indexfornler vi använt i avsnitt 6.1. Enda undantaget är formler som innehåller koordinatbasvektorer. Betrakta t ex ekvation 5.1 med följande utseende (här är alltså även a ett numeriskt index):

$$\mathbf{e}_i = (e_i)^a \frac{\partial}{\partial x^a}\,,\ i = 0, 1, \ldots, n-1.$$

I abstrakta indexnotationen skrivs basvektorn $\mathbf{e}_i$ som $(e_i)^a$ så istället måste vi skriva

$$(e_i)^a = (e_i)^j \left(\frac{\partial}{\partial x^j}\right)^a\,,\ i = 0, 1, \ldots, n-1,$$

där $(e_i)^j$ alltså betecknar komponenterna till $(e_i)^a$ i koordinatbasen.

Existens av global Riemannsk metrik ...

Existens av global Lorentzmetrik ...

6.8 Slutord

Vi är nu klara med det vi föresatte oss i inledningen till detta kapitel. Där gjorde vi ett antal påståenden och vi ska nu verifiera att alla dessa påståenden nu är rättfärdigade:

- 'T ex ska vi se på en koordinatfri definition av en tensor.'

  I Definition 6.14 definierar vi tensorer som linjära avbildningar. Denna definition refererar inte i sig till något koordinatsystem. Ett koordinatsystem behövs först då vi vill beräkna tensorns komponenter.

- 'Vi ska också rättfärdiga vårt lite slappa och oegentliga språkbruk när vi talar om 'vektorn v^a' istället för 'vektorn $\mathbf{v} = v^a \frac{\partial}{\partial x^a}$ med komponenter v^a i koordinatbasen'. '

  I och med införandet av den abstrakta indexnotationen har vi sett att v^a är en fullt respektabel och konsistent beteckning för vektorn själv. Vi är alltså inte tvungna att tolka ^{a} som ett numeriskt index och v^a som vektorns komponenter.

- 'På vägen kommer vi också att lösa mysteriet som beskrivs i slutet av avsnitt 4.4, d v s om nu en kontravariant tensor v^a är komponenter till en tangentvektor $\mathbf{v}$, vad är då en kovariant tensor α_a komponenter till för slags objekt?'

  Begreppet 'kovektor' var vad vi sökte när vi ställde denna fråga. Med abstrakta indexnotationen skulle vi säga att om talen α_i transformeras kovariant utgör de komponenter till kovektorn $\alpha_a = \alpha_i(dx^i)_a$ i koordinatbasen.

Vi avslutar kapitlet genom att än en gång konstatera att i indexformlerna i både tidigare och senare kapitel går det bra att tolka indexen som abstrakta *eller* numeriska allt efter behag. Den enda inkonsistensen gäller koordinatbasvektorerna som måste förses med ett extra abstrakt index d v s vi skriver $\left(\frac{\partial}{\partial x^i}\right)^a$ istället för $\frac{\partial}{\partial x^i}$.

Kapitel 7

Tensoranalys

7.1 Inledning

I detta kapitel ska vi generalisera derivatabegreppet till tensorfält. Det är inte uppenbart hur en sådan generalisering ska se ut. Vi kommer att hämta inspiration från geometri i $\mathbf{R}^3$ och på enhetssfären S till att definiera begreppet 'kovariant derivata'. Det går att tänka på kovarianta derivatan av ett tensorfält som ett tensorfält som håller reda på det ursprungliga tensorfältets alla riktningsderivator.

Ett alternativt (men närbesläktat) sätt att generalisera derivatabegreppet ges av Lie-derivatan. Den är viktig för att beskriva symmetrier i rumtiden.

7.2 Partiella derivator

Vi drar oss till minnes att ett tangentvektorfält $\mathbf{v}$ med komponenter v^a definierades som (se Definition 4.2) en riktningsderivata $\mathbf{v} = v^a \frac{\partial}{\partial x^a}$ d v s $\mathbf{v}$ avbildar en funktion f på funktionen $\mathbf{v}(f) = v^a \frac{\partial f}{\partial x^a}$. Enligt ekvation (4.7) kan vi alltså skriva

$$\mathbf{v}(f) = v^a \nabla_a f.$$

$\nabla_a f$, d v s differentialen till f innehåller all information som behövs för att bestämma vilken riktningsderivata till f som helst. Genom ovanstående samband definierar den också en avbildning som avbildar $\mathbf{v}$:s komponenter v^a på $\mathbf{v}(f)$.

Sammanfattningsvis: $\nabla_a f$ är en kovariant tensor som 'håller reda på f:s riktningsderivator' genom att riktningsderivatan av f med avseende på vektorn v^a ges av $v^a \nabla_a f$. Av detta skäl kallas $\nabla_a f$ ibland för *kovarianta derivatan* av f.

Målet med detta kapitel är, som tidigare sagts, att vi ska lära oss att derivera ett allmänt tensorfält och det verkar naturligt att börja med att försöka derivera enklast möjliga typ av tensorfält, d v s ett vektorfält u^a. Vägledda av ovanstående diskussion om differentialen verkar det naturligt att ställa frågan: Går det att hitta ett tensorfält $\nabla_b u^a$ av typ $(1, 1)$ kallat den *kovarianta derivatan* av u^a sådant att riktningsderivatan av u^a med avseende på vektorn v^b (vad nu detta betyder, begreppet riktningsderivata av ett vektorfält på en allmän mångfald är ju inte heller definierat ännu) ges av $v^b\nabla_b u^a$?[1]

Enligt ekvation 4.7 var ju $\nabla_a f = \frac{\partial f}{\partial x^a}$. Vi frågar oss därför om det är lämpligt att definiera $\nabla_b u^a$ som $\frac{\partial u^a}{\partial x^b}$. Detta första försök faller emellertid platt. $\frac{\partial u^a}{\partial x^b}$ är nämligen inte en tensor i allmänhet, vilket vi visade i Exempel 4.13.

Första försöket att definiera en kovariant derivata av ett vektorfält har alltså misslyckats och vi behöver nya uppslag för att komma vidare. Om vi bara kunde undersöka ett konkret fall där vi kan se geometriskt vad som vore en bra definition av kovariant derivata. I så fall skulle vi säkert kunna se hur detta ska generaliseras till en allmän mångfald ...

7.3 Enhetssfären, del 4 med mera

I detta kapitel ska vi försöka konstruera en kovariant derivata till ett vektorfält i tre konkreta fall, d v s en tensor $\nabla_b u^a$ av typ $(1,1)$ sådan att riktningsderivatan (vad vi nu ska mena med det) av ett vektorfält u^a med avseende på vektorn v^a ges av $v^b\nabla_b u^a$. Vi börjar med att studera $\mathbf{R}^3$, därefter ska vi se på enhetssfären S och till sist en godtycklig yta Y i $\mathbf{R}^3$.

Först några ord om tangentrummet till $\mathbf{R}^3$. Om $\{\bar{e}_x, \bar{e}_y, \bar{e}_z\}$ är standardbasen för $\mathbf{R}^3$ ges riktningsderivatan av en funktion f med avseende på $\bar{e}_x$ av $f'_{\bar{e}_x}(p) = \bar{e}_x \cdot \operatorname{grad} f(p) = \frac{\partial f}{\partial x}$ och analogt för övriga standardbasvektorer. Koordinatbasen för $T_p\mathbf{R}^3$ ges alltså av de vanliga partiella derivatorna $\left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right\}$ (formellt är ju elementen i tangentrummet riktningsderivator och inte geometriska vektorer). Genom att identifiera $\frac{\partial}{\partial x}$ med $\bar{e}_x$ o s v blir alltså $\mathbf{R}^3$ identisk med[2] sitt tangentrum $T_p\mathbf{R}^3$ i p.

Analogt med ekvation (4.1) definierar vi

Definition 7.1. Låt $\mathbf{u}$ vara ett vektorfält definierat i en omgivning av $p \in \mathbf{R}^3$ och låt $\mathbf{v} \in T_p\mathbf{R}^3$. Riktningsderivatan $\partial_{\mathbf{v}}\mathbf{u}(p)$ av $\mathbf{u}$ med avseende på $\mathbf{v}$ definieras då

$$\partial_{\mathbf{v}}\mathbf{u}(p) = \lim_{t\to 0} \frac{\mathbf{u}(p+t\mathbf{v}) - \mathbf{u}(p)}{t}.$$

Om $\mathbf{u} = u^0\bar{e}_x + u^1\bar{e}_y + u^2\bar{e}_z$ är alltså

$$\begin{aligned}\partial_{\mathbf{v}}\mathbf{u}(p) &= \lim_{t\to 0} \frac{\mathbf{u}(p+t\mathbf{v}) - \mathbf{u}(p)}{t} \\ &= \lim_{t\to 0}\left(\frac{u^0(p+t\mathbf{v}) - u^0(p)}{t}\bar{e}_x + \frac{u^1(p+t\mathbf{v}) - u^1(p)}{t}\bar{e}_y + \frac{u^2(p+t\mathbf{v}) - u^2(p)}{t}\bar{e}_z\right).\end{aligned}$$

Enligt tidigare är

$$\lim_{t\to 0} \frac{f(p+t\mathbf{v}) - f(p)}{t} = f'_{\mathbf{v}}(p) = \mathbf{v}(f)\,,\ \forall f$$

[1] Tyvärr är inte terminologin konsekvent här heller. I många framställningar kallas $v^b\nabla_b u^a$ för *kovarianta derivatan av u^a med avseende på v^a*. ∇_a kallas i dessa framställningar en *förbindelse*.

[2] Egentligen 'isomorf med'. Jämför diskussionen i slutet av avsnitt 6.2.

vilket ger att

$$\partial_{\mathbf{v}}\mathbf{u}(p) = \left(\mathbf{v}(u^0)\bar{e}_x + \mathbf{v}(u^1)\bar{e}_y + \mathbf{v}(u^2)\bar{e}_z\right)\Big|_p .$$

Eftersom $\bar{e}_x$, $\bar{e}_y$ och $\bar{e}_z$ är konstanta vektorer sker alltså derivering av vektorfält i $\mathbf{R}^3$ komponentvis. Vi observerar att om även $\mathbf{v}$ är ett vektorfält i en omgivning av p så blir

$$\partial_{\mathbf{v}}\mathbf{u} = \mathbf{v}(u^0)\bar{e}_x + \mathbf{v}(u^1)\bar{e}_y + \mathbf{v}(u^2)\bar{e}_z$$

ett nytt vektorfält som också är definierat i en omgivning av p och om $(\partial_{\mathbf{v}}\mathbf{u})^a$ är komponenterna till $\partial_{\mathbf{v}}\mathbf{u}$ i standardbasen är enligt ovan

$$(\partial_{\mathbf{v}}\mathbf{u})^a = \mathbf{v}(u^a).$$

Vi är nu klara att definiera kovarianta derivatan av ett vektorfält $\mathbf{u}$ i $\mathbf{R}^3$.

Definition 7.2. Antag att $\mathbf{u}$ och $\mathbf{v}$ har komponenter u^a respektive v^a. Om det finns en uppsättning funktioner $\partial_b u^a$ sådana att

$$v^b\partial_b u^a = (\partial_{\mathbf{v}}\mathbf{u})^a$$

för alla vektorfält $\mathbf{u}$ och $\mathbf{v}$ kallas $\partial_b u^a$ den *kovarianta derivatan* av u^a.

Anmärkning: På en godtycklig mångfald kommer kovarianta derivatan av u^a att betecknas $\nabla_b u^a$. Med tanke på de följande exemplen (enhetssfären samt en godtycklig yta Y i $\mathbf{R}^3$) är det dock praktiskt att ha en separat beteckning för kovariant derivata i $\mathbf{R}^3$. Vi väljer därför att skriva den kovarianta derivatan[3] i $\mathbf{R}^3$ som $\partial_b u^a$ istället för $\nabla_b u^a$. □

Vi ska nu bestämma ett uttryck för $\partial_b u^a$. Enligt ovan är

$$v^b\partial_b u^a = (\partial_{\mathbf{v}}\mathbf{u})^a = \mathbf{v}(u^a) = v^b\frac{\partial u^a}{\partial x^b}$$

för alla v^b vilket ger

$$\partial_b u^a = \frac{\partial u^a}{\partial x^b}.$$

Observera att ovanstående resonemang bara visar att detta samband gäller i det givna koordinatsystemet (x, y, z). Det är dock inte svårt att visa (Gör detta!) att sambandet också gäller i ett godtyckligt *kartesiskt* koordinatsystem, d v s ett koordinatsystem (x', y', z') sådant att koordinatbasen $\left\{\bar{e}_{x'}, \bar{e}_{y'}, \bar{e}_{z'}\right\}$ är en ON-bas bestående av *konstanta* vektorer.

Sammanfattningsvis är alltså kovariant derivata och partiell derivata samma sak i kartesiska koordinatsystem för $\mathbf{R}^3$. Det är dock viktigt att observera att i andra koordinatsystem, t ex i sfäriska eller cylindriska koordinater, är kovariant derivata och partiell derivata helt olika begrepp.

[3]Här gäller det att se upp med beteckningarna igen. I många skrifter (t ex *d'Inverno-Vickers:* Introducing Einstein's relativity — A deeper understanding) är $\partial_b u^a$ istället en beteckning för den partiella derivatan $\frac{\partial u^a}{\partial x^b}$ och denna beteckning används inte bara i $\mathbf{R}^3$ utan också på en godtycklig mångfald. Här reserverar vi beteckningen $\partial_b u^a$ för kovarianta derivatan i $\mathbf{R}^3$.

Exempel 7.3. *(Enhetssfären, del 4)* Vi vänder oss nu till enhetssfären S. Beteckningarna i detta avsnitt kommer att vara desamma som i de tidigare kapitlen om S. Vi kommer också (som i avsnitt 5.3) att identifiera koordinatbasvektorerna $\frac{\partial}{\partial\theta}$ och $\frac{\partial}{\partial\varphi}$ med de geometriska vektorerna $\mathbf{e}_\theta$ och $\mathbf{e}_\varphi$.

Låt $p \in S$ vara godtycklig och som tidigare kan vi anta att $p \in U$ och har koordinater (θ, φ) i koordinatsystemet ψ. Ortsvektorn $\bar{r} = \bar{r}(\theta, \varphi)$ för p är

$$\bar{r}(\theta, \varphi) = \sin\theta\cos\varphi\bar{e}_x + \sin\theta\sin\varphi\bar{e}_y + \cos\theta\bar{e}_z.$$

och koordinatbasvektorerna ges av

$$\mathbf{e}_\theta = \quad \cos\theta\cos\varphi\bar{e}_x + \cos\theta\sin\varphi\bar{e}_y - \sin\theta\bar{e}_z \tag{7.1}$$

$$\mathbf{e}_\varphi = -\sin\theta\sin\varphi\bar{e}_x + \sin\theta\cos\varphi\bar{e}_y \tag{7.2}$$

Godtyckliga tangentvektorfält $\mathbf{u}$ och $\mathbf{v}$ definierade i en omgivning av p ges av

$$\begin{aligned}\mathbf{u} &= u^0\mathbf{e}_\theta + u^1\mathbf{e}_\varphi = u^0(\cos\theta\cos\varphi\bar{e}_x + \cos\theta\sin\varphi\bar{e}_y - \sin\theta\bar{e}_z) + u^1(-\sin\theta\sin\varphi\bar{e}_x + \sin\theta\cos\varphi\bar{e}_y)\\ \mathbf{v} &= v^0\mathbf{e}_\theta + v^1\mathbf{e}_\varphi = v^0(\cos\theta\cos\varphi\bar{e}_x + \cos\theta\sin\varphi\bar{e}_y - \sin\theta\bar{e}_z) + v^1(-\sin\theta\sin\varphi\bar{e}_x + \sin\theta\cos\varphi\bar{e}_y)\end{aligned}$$

där u^0, u^1, v^0 och v^1 är godtyckliga funktioner av θ och φ. Vi låter nu $\mathbf{U}$ och $\mathbf{V}$ vara vektorfält definierade i en omgivning Ω till p i $\mathbf{R}^3$ sådana att $\mathbf{U}|_S = \mathbf{u}$ och $\mathbf{V}|_S = \mathbf{v}$, där $\mathbf{U}|_S$ betyder $\mathbf{U}$:s restriktion till den del av S som också ligger i Ω.

Sätt $r = \sqrt{x^2 + y^2 + z^2}$. Då blir (r, θ, φ) de vanliga sfäriska koordinaterna på Ω och koordinatbasvektorn $\bar{e}_r$ hörande till koordinaten r ges av

$$\bar{e}_r = \sin\theta\cos\varphi\bar{e}_x + \sin\theta\sin\varphi\bar{e}_y + \cos\theta\bar{e}_z.$$

Om vi utvidgar definitionsmängden av $\mathbf{e}_\theta$ och $\mathbf{e}_\varphi$ genom att kräva att de ges av formlerna (7.1) och (7.2) i hela Ω blir $\{\mathbf{e}_\theta, \mathbf{e}_\varphi, \bar{e}_r\}$ en bas för $\mathbf{R}^3$ (egentligen tangentrummet till $\mathbf{R}^3$, se disussionen i förra exemplet) som är definierad i hela Ω. Vektorfältet $\mathbf{U}$ kan då skrivas

$$\mathbf{U} = U^0\mathbf{e}_\theta + U^1\mathbf{e}_\varphi + U^2\bar{e}_r,$$

där funktionerna U^k uppfyller $U^0\big|_S = u^0$, $U^1\big|_S = u^1$ och $U^2\big|_S = 0$ och motsvarande för $\mathbf{V}$.

Vi söker nu riktningsderivatan av $\mathbf{U}$ med avseende på $\mathbf{V}$ i punkter på S. Med beteckningar från förra exemplet är denna $(\partial_\mathbf{V}\mathbf{U})|_S$. På grund av linearitet i $\mathbf{V}$ blir

$$(\partial_\mathbf{V}\mathbf{U})|_S = \left(V^0\partial_{\mathbf{e}_\theta}\mathbf{U} + V^1\partial_{\mathbf{e}_\varphi}\mathbf{U} + V^2\partial_{\bar{e}_r}\mathbf{U}\right)\Big|_S = v^0\left(\partial_{\mathbf{e}_\theta}\mathbf{U}\right)|_S + v^1\left(\partial_{\mathbf{e}_\varphi}\mathbf{U}\right)\Big|_S,$$

ty $V^0\big|_S = v^0$, $V^1\big|_S = v^1$ och $V^2\big|_S = 0$.

Enligt förra exemplet är kovariant derivata = partiell derivata i standardbasen för $\mathbf{R}^3$. Kedjeregeln ger därför att

$$(\partial_{\mathbf{e}_\theta}\mathbf{U})|_S = \left(\cos\theta\cos\varphi\frac{\partial\mathbf{U}}{\partial x} + \cos\theta\sin\varphi\frac{\partial\mathbf{U}}{\partial y} - \sin\theta\frac{\partial\mathbf{U}}{\partial z}\right)\bigg|_S = \left(\frac{\partial\mathbf{U}}{\partial\theta}\right)\bigg|_S = \frac{\partial\mathbf{u}}{\partial\theta},$$

där sista likheten följer av att deriveringen sker tangentiellt med S och derivatan beror därför bara av $\mathbf{U}$:s värden på S. En snarlik räkning ger också att $\left(\partial_{\mathbf{e}_\varphi}\mathbf{U}\right)\big|_S = \frac{\partial \mathbf{u}}{\partial \varphi}$. Detta ger

$$\begin{aligned}
(\partial_{\mathbf{e}_\theta}\mathbf{U})|_S &= \frac{\partial \mathbf{u}}{\partial \theta} = \frac{\partial u^0}{\partial \theta}\mathbf{e}_\theta + u^0\frac{\partial \mathbf{e}_\theta}{\partial \theta} + \frac{\partial u^1}{\partial \theta}\mathbf{e}_\varphi + u^1\frac{\partial \mathbf{e}_\varphi}{\partial \theta} = \frac{\partial u^0}{\partial \theta}\mathbf{e}_\theta + \frac{\partial u^1}{\partial \theta}\mathbf{e}_\varphi \\
&\quad + u^0\left(-\sin\theta\cos\varphi\bar{e}_x - \sin\theta\sin\varphi\bar{e}_y - \cos\theta\bar{e}_z\right) + u^1\left(-\cos\theta\sin\varphi\bar{e}_x + \cos\theta\cos\varphi\bar{e}_y\right) \\
&= \frac{\partial u^0}{\partial \theta}\mathbf{e}_\theta + \left(\frac{\partial u^1}{\partial \theta} + (\cot\theta)u^1\right)\mathbf{e}_\varphi - u^0\bar{r}.
\end{aligned}$$

Vi ser här att de första termerna visserligen bildar en tangentvektor till S men att den sista termen är ortogonal mot S. $\frac{\partial \mathbf{u}}{\partial \theta}$ är alltså *inte* en tangentvektor till S trots att både $\mathbf{u}$ och $\mathbf{e}_\theta$ är tangentvektorfält till S. Vi får också

$$\begin{aligned}
\left(\partial_{\mathbf{e}_\varphi}\mathbf{U}\right)\big|_S &= \frac{\partial \mathbf{u}}{\partial \varphi} = \frac{\partial u^0}{\partial \varphi}\mathbf{e}_\theta + u^0\frac{\partial \mathbf{e}_\theta}{\partial \varphi} + \frac{\partial u^1}{\partial \varphi}\mathbf{e}_\varphi + u^1\frac{\partial \mathbf{e}_\varphi}{\partial \varphi} = \frac{\partial u^0}{\partial \varphi}\mathbf{e}_\theta + \frac{\partial u^1}{\partial \varphi}\mathbf{e}_\varphi \\
&\quad + u^0\left(-\cos\theta\sin\varphi\bar{e}_x + \cos\theta\cos\varphi\bar{e}_y\right) + u^1\left(-\sin\theta\cos\varphi\bar{e}_x - \sin\theta\sin\varphi\bar{e}_y\right) \\
&= \frac{\partial u^0}{\partial \varphi}\mathbf{e}_\theta + \left(\frac{\partial u^1}{\partial \varphi} + (\cot\theta)u^0\right)\mathbf{e}_\varphi - u^1\left(\sin\theta\cos\varphi\bar{e}_x + \sin\theta\sin\varphi\bar{e}_y\right)
\end{aligned}$$

För att uttrycka den sista termen i kända vektorer observerar vi att

$$\bar{r} + \cot\theta\mathbf{e}_\theta = \left(\sin\theta + \frac{\cos^2\theta}{\sin\theta}\right)\cos\varphi\bar{e}_x + \left(\sin\theta + \frac{\cos^2\theta}{\sin\theta}\right)\sin\varphi\bar{e}_y = \frac{1}{\sin\theta}\left(\cos\varphi\bar{e}_x + \sin\varphi\bar{e}_y\right),$$

vilket ger

$$\left(\partial_{\mathbf{e}_\varphi}\mathbf{U}\right)\big|_S = \left(\frac{\partial u^0}{\partial \varphi} - (\sin\theta\cos\theta)u^1\right)\mathbf{e}_\theta + \left(\frac{\partial u^1}{\partial \varphi} + (\cot\theta)u^0\right)\mathbf{e}_\varphi - \sin^2\theta\cdot u^1\bar{r}.$$

Riktningsderivatan av $\mathbf{U}$ med avseende på $\mathbf{V}$ ges alltså av

$$\begin{aligned}
(\partial_{\mathbf{V}}\mathbf{U})|_S &= v^0\left(\partial_{\mathbf{e}_\theta}\mathbf{U}\right)|_S + v^1\left(\partial_{\mathbf{e}_\varphi}\mathbf{U}\right)\big|_S = \left(v^0\frac{\partial u^0}{\partial \theta} + v^1\frac{\partial u^0}{\partial \varphi} - (\sin\theta\cos\theta)u^1v^1\right)\mathbf{e}_\theta \\
&\quad + \left(v^0\frac{\partial u^1}{\partial \theta} + v^1\frac{\partial u^1}{\partial \varphi} + \cot\theta(u^1v^0 + u^0v^1)\right)\mathbf{e}_\varphi - (u^0v^0 + (\sin^2\theta)u^1v^1)\bar{r}.
\end{aligned}$$

Vi ser nu att $(\partial_{\mathbf{V}}\mathbf{U})|_S$ bara beror på $\mathbf{U}$:s och $\mathbf{V}$:s värden på S, d v s på våra ursprungliga vektorfält $\mathbf{u}$ och $\mathbf{v}$. Det spelar alltså ingen roll *hur* vi utvidgar $\mathbf{u}$ och $\mathbf{v}$ till den öppna mängden $\Omega \subset \mathbf{R}^3$. Det är därför befogat att införa beteckningen $\partial_{\mathbf{v}}\mathbf{u}$ istället för $(\partial_{\mathbf{V}}\mathbf{U})|_S$. Vi sätter alltså

$$\begin{aligned}
\partial_{\mathbf{v}}\mathbf{u} &= \left(v^0\frac{\partial u^0}{\partial \theta} + v^1\frac{\partial u^0}{\partial \varphi} - (\sin\theta\cos\theta)u^1v^1\right)\mathbf{e}_\theta + \left(v^0\frac{\partial u^1}{\partial \theta} + v^1\frac{\partial u^1}{\partial \varphi} + \cot\theta(u^1v^0 + u^0v^1)\right)\mathbf{e}_\varphi \\
&\quad - (u^0v^0 + (\sin^2\theta)u^1v^1)\bar{r}. \qquad (7.3)
\end{aligned}$$

Nu är $\partial_{\mathbf{v}}\mathbf{u}$ visserligen ett vektorfält som är definierat på (en del av) S, men det är *inte* ett tangentvektorfält till S ty den sista termen är nollskild i allmänhet. Om vi vill definiera riktningsderivatan av $\mathbf{u}$ med avseende på $\mathbf{v}$ som ett tangentvektorfält till S duger alltså inte ekvation (7.3). Det finns dock ett naturligt sätt att skapa en tangentvektor till S från uttrycket i (7.3), nämligen genom att ortogonalprojicera denna vektor på T_pS.

Om riktningsderivatan av $\mathbf{u}$ med avseende på $\mathbf{v}$ betecknas $\nabla_{\mathbf{v}}\mathbf{u}$ sätter vi alltså

$$\nabla_{\mathbf{v}}\mathbf{u} = \left(v^0\frac{\partial u^0}{\partial\theta} + v^1\frac{\partial u^0}{\partial\varphi} - (\sin\theta\cos\theta)u^1v^1\right)\mathbf{e}_\theta + \left(v^0\frac{\partial u^1}{\partial\theta} + v^1\frac{\partial u^1}{\partial\varphi} + \cot\theta(u^1v^0 + u^0v^1)\right)\mathbf{e}_\varphi.$$

Om komponenterna till $\nabla_{\mathbf{v}}\mathbf{u}$ betecknas $v^b\nabla_b u^a$ och koordinaterna $(\theta,\varphi) = (x^0, x^1)$ får vi

$$v^b\nabla_b u^0 = v^0\frac{\partial u^0}{\partial x^0} + v^1\frac{\partial u^0}{\partial x^1} - (\sin\theta\cos\theta)u^1v^1 = v^b\frac{\partial u^0}{\partial x^b} - (\sin\theta\cos\theta)u^1v^1$$

och

$$v^b\nabla_b u^1 = v^0\frac{\partial u^1}{\partial x^0} + v^1\frac{\partial u^1}{\partial x^1} + \cot\theta(u^1v^0 + u^0v^1) = v^b\frac{\partial u^1}{\partial x^b} + \cot\theta(u^1v^0 + u^0v^1).$$

Dessa ekvationer kan sammanfattas till en om vi inför betckningen ${\Gamma^a}_{bc}$, $a, b, c = 0, 1$ genom att sätta ${\Gamma^0}_{11} = -\sin\theta\cos\theta$, ${\Gamma^1}_{01} = {\Gamma^1}_{10} = \cot\theta$ och övriga ${\Gamma^a}_{bc} = 0$. Då är ${\Gamma^a}_{bc}$ symmetrisk i sina nedre index, d v s ${\Gamma^a}_{bc} = {\Gamma^a}_{cb}$ och

$$v^b\nabla_b u^a = v^b\left(\frac{\partial u^a}{\partial x^b} + {\Gamma^a}_{bc}u^c\right).$$

En lämplig definition av kovarianta derivatan $\nabla_b u^a$ är alltså

$$\nabla_b u^a = \frac{\partial u^a}{\partial x^b} + {\Gamma^a}_{bc}u^c, \tag{7.4}$$

där ${\Gamma^0}_{11} = -\sin\theta\cos\theta$, ${\Gamma^1}_{01} = {\Gamma^1}_{10} = \cot\theta$ och övriga ${\Gamma^a}_{bc} = 0$ *ifall detta uttryck definierar en tensor!!* Observera att detta inte är självklart eftersom de sfäriska koordinaterna dyker upp explicit i uttrycket för $\nabla_b u^a$.

Vi noterar också att om $\nabla_b u^a$ *är* en tensor så är med nödvändighet ${\Gamma^a}_{bc}$ *inte* en tensor (varför?). I själva verket är den transformationslag som ${\Gamma^a}_{bc}$ måste uppfylla för att $\nabla_b u^a$ ska bli en tensor ganska komplicerad och vi återkommer till den i nästa exempel. Där ska vi också se att uttrycket ovan för $\nabla_b u^a$ verkligen definierar en tensor på S.

Ekvation (7.4) skulle alltså duga som definition av kovariant derivata på S. Den är dock svår att direkt generalisera till en godtycklig mångfald eftersom det inte är klart hur ${\Gamma^a}_{bc}$ ska konstrueras på en godtycklig mångfald. Nästa exempel ger dock en viktig ledtråd om detta. □

Exempel 7.4. Låt Y vara en godtycklig $\mathcal{C}^\infty$, 2-dimensionell yta[4] i $\mathbf{R}^3$ och låt (x^0, x^1) vara koordinater på Y. Det är visserligen inte säkert att det finns något koordinatsystem som täcker hela Y

[4] Detta villkor kan försvagas till t ex $\mathcal{C}^2$ om man så vill.

(jämför enhetssfären S), men om vi fixerar en godtycklig punkt $p \in Y$ finns i alla fall en omgivning $U \subset Y$ till p som kan beskrivas med ett enda koordinatsystem.

Vi söker nu riktningsderivatan av $\mathbf{u}$ med avseende på $\mathbf{v}$ där $\mathbf{u}$ och $\mathbf{v}$ är godtyckliga vektorfält på U och vi följer receptet från förra exemplet. Här finns dock ett problem som inte är helt enkelt att övervinna. I förra exemplet kunde vi dra nytta av att koordinaterna θ och φ är väldefinierade även utanför S och att vi dessutom har tillgång till den sfäriska koordinaten r så att (r, θ, φ) blir ett koordinatsystem i en 3-dimensionell omgivning Ω av $U \subset S$.

I detta exempel är koordinaterna (x^0, x^1) bara definierade på Y. Vi måste därför börja med att utvidga definitionsmängden till (x^0, x^1) till någon 3-dimensionell omgivning Ω till U i $\mathbf{R}^3$ och sedan hitta x^2 så att (x^0, x^1, x^2) är ett koordinatsystem på Ω. Man kan visa att detta är möjligt, men för att inte tynga ner framställningen utelämnar vi detaljerna.

Givet detta koordinatsystem för Ω kan $\mathbf{u}$ och $\mathbf{v}$ utvidgas till vektorfält $\mathbf{U}$ och $\mathbf{V}$ d v s $\mathbf{U}|_S = \mathbf{u}$ och $\mathbf{V}|_S = \mathbf{v}$, definierade på en omgivning Ω till p i $\mathbf{R}^3$.

Vi utnyttjar att vektorfältet $\mathbf{U}$ på $\Omega \subset \mathbf{R}^3$ får deriveras komponentvis med avseende på $\mathbf{V}$ i ett kartesiskt koordinatsystem. Detta ger som i förra exemplet en riktningsderivata i $\mathbf{R}^3$ som betecknas $\partial_{\mathbf{V}}\mathbf{U}$ och på liknande sätt som där finner vi att i punkter på Y beror $\partial_{\mathbf{V}}\mathbf{U}$ inte på *hur* utvidgningen gjordes utan $\partial_{\mathbf{V}}\mathbf{U}$ beror bara på $\mathbf{U}$:s och $\mathbf{V}$:s värden på Y, d v s bara på de ursprungliga vektorfälten $\mathbf{u}$ och $\mathbf{v}$. Som i förra exemplet sätter vi därför

$$\partial_{\mathbf{v}}\mathbf{u} = (\partial_{\mathbf{V}}\mathbf{U})|_Y .$$

Precis som i förra exemplet skriver vi tangentvektorer till Y med fetstil och andra vektorer i $\mathbf{R}^3$ med vektorstreck. Ortsvektorn $\bar{r}$ för punkten p på Y med koordinater (x^0, x^1) ges då av

$$\bar{r} = \bar{r}(x^0, x^1) = x(x^0, x^1)\bar{e}_x + y(x^0, x^1)\bar{e}_y + z(x^0, x^1)\bar{e}_z.$$

Precis som på enhetssfären bildar vektorerna

$$\mathbf{e}_0 = \frac{\partial \bar{r}}{\partial x^0} \quad , \qquad \mathbf{e}_1 = \frac{\partial \bar{r}}{\partial x^1}$$

en bas för T_pY (eller mer precist: dessa geometriska vektorer i $\mathbf{R}^3$ identifieras med riktningsderivatorna $\frac{\partial}{\partial x^0}$ och $\frac{\partial}{\partial x^1}$ som bildar koordinatbasen för T_pY i detta koordinatsystem).

Vi kan alltså skriva $\mathbf{u} = u^a\mathbf{e}_a = u^a\frac{\partial \bar{r}}{\partial x^a}$ och $\mathbf{v} = v^a\frac{\partial \bar{r}}{\partial x^a}$ och då derivering i ett kartesiskt koordinatsystem sker komponentvis får vi

$$\partial_{\mathbf{v}}\mathbf{u} = v^b\frac{\partial}{\partial x^b}\left(u^a\frac{\partial \bar{r}}{\partial x^a}\right) = v^b\frac{\partial u^a}{\partial x^b}\frac{\partial \bar{r}}{\partial x^a} + \frac{\partial^2 \bar{r}}{\partial x^b \partial x^a}u^a v^b = v^b\frac{\partial u^a}{\partial x^b}\mathbf{e}_a + \frac{\partial^2 \bar{r}}{\partial x^b \partial x^c}u^c v^b.$$

Vi noterar att $\partial_{\mathbf{v}}\mathbf{u}$ inte nödvändigtvis är ett tangentvektorfält till Y. Som i förra exemplet definierar vi därför riktningsderivatan $\nabla_{\mathbf{v}}\mathbf{u}$ som ortogonalprojektionen av $\partial_{\mathbf{v}}\mathbf{u}$ på T_pY och vi ser att första termen redan är en tangentvektor till Y. Det räcker alltså att ortogonalprojicera den andra.

Metriken g_{ab} på Y definieras $g_{ab} = \mathbf{e}_a \cdot \mathbf{e}_b$ för $a, b = 0, 1$, där $\cdot$ står för den vanliga euklidiska skalärprodukten i $\mathbf{R}^3$ och om $\mathbf{u} = u^a\mathbf{e}_a$, $\mathbf{v} = v^a\mathbf{e}_a \in T_pY$ är $\mathbf{u} \cdot \mathbf{v} = g_{ab}u^a v^b$ (jämför konstruktionen i avsnitt 5.3).

Låt oss nu studera problemet att hitta ortogonalprojektionen $\mathbf{q}$ på T_pY, av en godtycklig vektor $\bar{w} \in \mathbf{R}^3$. Observera att basvektorerna $\mathbf{e}_d$ inte är ortogonala i allmänhet, så den vanliga projektionsformeln fungerar inte här. Istället utnyttjar vi att vektorn $\bar{w} - \mathbf{q}$ ska vara ortogonal mot T_pY och därmed mot båda basvektorerna $\mathbf{e}_d$, d v s $\bar{w} \cdot \mathbf{e}_d = \mathbf{q} \cdot \mathbf{e}_d$ för $d = 0, 1$. Då $\mathbf{q} \in T_pY$ är $\mathbf{q} = q^b\mathbf{e}_b$ vilket ger att $\mathbf{q} \cdot \mathbf{e}_d = q^b\mathbf{e}_b \cdot \mathbf{e}_d = g_{db}q^b$. Kontraktion med g^{ad} ger sedan att

$$g^{ad}\mathbf{q} \cdot \mathbf{e}_d = g^{ad}g_{db}q^b = {\delta_b}^a q^b = q^a.$$

Det gäller alltså att

$$q^a = g^{ad}\mathbf{q} \cdot \mathbf{e}_d = g^{ad}\bar{w} \cdot \mathbf{e}_d = g^{ad}\bar{w} \cdot \frac{\partial \bar{r}}{\partial x^d},$$

d v s

$$\mathbf{q} = q^a\mathbf{e}_a = g^{ad}\bar{w} \cdot \frac{\partial \bar{r}}{\partial x^d}\mathbf{e}_a.$$

I vårt fall är $\bar{w} = \frac{\partial^2 \bar{r}}{\partial x^b \partial x^c} u^c v^b$ vilket ger följande uttryck för $\nabla_\mathbf{v}\mathbf{u}$:

$$\nabla_\mathbf{v}\mathbf{u} = \left(v^b \frac{\partial u^a}{\partial x^b} + g^{ad} \frac{\partial^2 \bar{r}}{\partial x^b \partial x^c} \cdot \frac{\partial \bar{r}}{\partial x^d} u^c v^b \right) \mathbf{e}_a.$$

Vi vill nu eliminera ortsvektorn $\bar{r}$ från detta uttryck eftersom det är den enda delen av uttrycket som refererar till ett omkringliggande $\mathbf{R}^3$. Detta kan göras på följande sätt: Betrakta

$$\frac{\partial g_{cd}}{\partial x^b} = \frac{\partial}{\partial x^b}\left(\frac{\partial \bar{r}}{\partial x^c} \cdot \frac{\partial \bar{r}}{\partial x^d} \right) = \frac{\partial^2 \bar{r}}{\partial x^b \partial x^c} \cdot \frac{\partial \bar{r}}{\partial x^d} + \frac{\partial^2 \bar{r}}{\partial x^b \partial x^d} \cdot \frac{\partial \bar{r}}{\partial x^c}.$$

Permutation av indexen ger sedan att

$$\frac{\partial g_{db}}{\partial x^c} = \frac{\partial^2 \bar{r}}{\partial x^c \partial x^d} \cdot \frac{\partial \bar{r}}{\partial x^b} + \frac{\partial^2 \bar{r}}{\partial x^c \partial x^b} \cdot \frac{\partial \bar{r}}{\partial x^d}$$

och

$$\frac{\partial g_{bc}}{\partial x^d} = \frac{\partial^2 \bar{r}}{\partial x^d \partial x^b} \cdot \frac{\partial \bar{r}}{\partial x^c} + \frac{\partial^2 \bar{r}}{\partial x^d \partial x^c} \cdot \frac{\partial \bar{r}}{\partial x^b}.$$

Addera de två första av dessa och subtrahera den tredje så fås

$$\frac{\partial g_{cd}}{\partial x^b} + \frac{\partial g_{db}}{\partial x^c} - \frac{\partial g_{bc}}{\partial x^d} = 2\frac{\partial^2 \bar{r}}{\partial x^b \partial x^c} \cdot \frac{\partial \bar{r}}{\partial x^d},$$

där vi har utnyttjat att Y och därmed $\bar{r}$ är av klass $\mathcal{C}^2$ så att blandade partiella andraderivator är lika.

Om vi inför beteckningen

$$\Gamma^a{}_{bc} = \frac{1}{2} g^{ad} \left(\frac{\partial g_{cd}}{\partial x^b} + \frac{\partial g_{db}}{\partial x^c} - \frac{\partial g_{bc}}{\partial x^d} \right) \tag{7.5}$$

så är alltså

$$\nabla_{\mathbf{v}}\mathbf{u} = \left(v^b\frac{\partial u^a}{\partial x^b} + \Gamma^a{}_{bc}u^cv^b\right)\mathbf{e}_a.$$

$\Gamma^a{}_{bc}$ kallas *Christoffelsymbolerna* på Y och vi observerar att $\Gamma^a{}_{bc} = \Gamma^a{}_{cb}$ så Christoffelsymbolerna är symmetriska med avseende på de nedre indexen.

Om vi (som tidigare) låter $v^b\nabla_b u^a$ beteckna komponenterna till $\nabla_{\mathbf{v}}\mathbf{u}$ är alltså

$$v^b\nabla_b u^a = \left(\frac{\partial u^a}{\partial x^b} + \Gamma^a{}_{bc}u^c\right)v^b$$

för alla v^b, varför det är lämpligt att definiera *kovarianta derivatan* $\nabla_b u^a$ som

$$\nabla_b u^a = \frac{\partial u^a}{\partial x^b} + \Gamma^a{}_{bc}u^c. \tag{7.6}$$

Anmärkning: Christoffelsymbolerna

$$\Gamma^a{}_{bc} = \frac{1}{2}g^{ad}\left(\frac{\partial g_{cd}}{\partial x^b} + \frac{\partial g_{db}}{\partial x^c} - \frac{\partial g_{bc}}{\partial x^d}\right)$$

utgör *ej* en tensor. Faktum är att genom att utnyttja kedjeregeln samt att g^{ad} och g_{cd} är tensorer kan man visa (Gör detta själv som övning!) att Christoffelsymbolerna transformeras enligt

$$\bar{\Gamma}^a{}_{bc} = \frac{\partial \bar{x}^a}{\partial x^d}\frac{\partial x^e}{\partial \bar{x}^b}\frac{\partial x^f}{\partial \bar{x}^c}\Gamma^d{}_{ef} - \frac{\partial x^d}{\partial \bar{x}^b}\frac{\partial x^e}{\partial \bar{x}^c}\frac{\partial^2 \bar{x}^a}{\partial x^d \partial x^e} \tag{7.7}$$

vid byte till ett nytt koordinatsystem $\bar{x}^a$. Observera att första termen är transformationslagen för en tensor av typ $(1,2)$, men att den andra termen förstås är nollskild i allmänhet.

Ingen av termerna i högerledet till ekvation (7.6) är alltså tensorer, men med hjälp av (7.7), (4.6) och kedjeregeln kan man visa (Gör även detta som övning!) att *hela* högerledet till (7.6) *är* en tensor. $\nabla_b u^a$ definierad av (7.6) är alltså en tensor precis som vi önskade. □

Övning: Enligt förra exemplet ges Christoffelsymbolerna i sfäriska koordinater på enhetssfären S av $\Gamma^0{}_{11} = -\sin\theta\cos\theta$, $\Gamma^1{}_{01} = \Gamma^1{}_{10} = \cot\theta$ och övriga $\Gamma^a{}_{bc} = 0$. Visa att detta är precis vad en direkt beräkning av $\Gamma^a{}_{bc}$ från ekvation (7.5) ger.

Var noga med att inte missa någon av Christoffelsymbolerna. För en 2-dimensionell yta finns det 8 st varav några visserligen inte är oberoende av de övriga p g a symmetrin $\Gamma^a{}_{bc} = \Gamma^a{}_{cb}$. □

I nästa avsnitt ska vi definiera kovariant derivata på en godtycklig mångfald M försedd med en metrik g_{ab}. Eftersom ekvationerna (7.5) och (7.6) ger en definition av kovariant derivata som inte refererar till något omkringliggande $\mathbf{R}^3$ skulle dessa formler gå bra att direkt generalisera till M.

En sådan definition skulle dock inte vara särskilt upplysande. På en godtycklig mångfald som inte ligger i något omkringliggande $\mathbf{R}^n$ har vi ju inte ovanstående geometriska tolkning av ekvationerna (7.5) och (7.6) som ortogonalprojektionen på M av den 'vanliga' derivatan i $\mathbf{R}^n$.

Av pedagogiska skäl kommer vi därför, till en början, att karakterisera den kovarianta derivatan på M utifrån diverse andra egenskaper. Längre fram i nästa avsnitt kommer vi dock att se att denna karakterisering är ekvivalent med ekvationerna (7.5) och (7.6).

7.4 Kovarianta derivator

Låt nu M vara en godtycklig mångfald med lokala koordinater x^a och metrik g_{ab}. Låt $\mathbf{u} = u^a \frac{\partial}{\partial x^a}$ vara ett godtyckligt vektorfält på M. Vi vill nu definiera en kovariant derivata $\nabla_b u^a$ av u^a. Som vi tidigare konstaterat är $\frac{\partial u^a}{\partial x^b}$ ingen tensor i allmänhet men rimligen bör väl kovariant derivata ändå ha något med vanlig derivata att göra?! En naturlig ansats är därför

$$\nabla_b u^a = \frac{\partial u^a}{\partial x^b} + \text{korrektionsterm som gör H.L. till en tensor,}$$

och frågan är hur korrektionstermen ska se ut.

En viktig egenskap hos andra typer av derivator är linearitet och om vi vill att kovariant derivata ska vara linjär i u^a har vi inget annat val än att låta korrektionstermen vara linjär i u^a. Dessutom vill vi att $\nabla_b u^a$ ska vara en tensor av typ $(1,1)$. Vi definierar därför

Definition 7.5. Om

$$\nabla_b u^a = \frac{\partial u^a}{\partial x^b} + \Gamma^a{}_{bc} u^c \tag{7.8}$$

är en tensor av typ $(1,1)$ kallas $\nabla_b u^a$ en *kovariant derivata* av u^a. Koefficienterna $\Gamma^a{}_{bc}$ kallas en *förbindelse*. Vidare, om $\mathbf{v} = v^a \frac{\partial}{\partial x^a}$ är ett godtyckligt vektorfält kallas vektorfältet $\nabla_{\mathbf{v}} \mathbf{u}$ med komponenter

$$(\nabla_{\mathbf{v}} \mathbf{u})^a := v^b \nabla_b u^a$$

för *riktningsderivatan* av u^a med avseende på v^b.[5]

Antag att $\nabla_b u^a$ är en kovariant derivata. Vi vet sedan tidigare att $\frac{\partial u^a}{\partial x^b}$ inte är en tensor så det följer att $\Gamma^a{}_{bc}$ inte kan vara en tensor heller. Om vi byter till ett nytt koordinatsystem $\bar{x}^a$ och utnyttjar att både u^a och $\nabla_b u^a$ uppfyller transformationslagarna för tensorer så kan vi härleda en transformationslag för förbindelsen $\Gamma^a{}_{bc}$. Vi får

SATS 7.6. $\nabla_b u^a = \frac{\partial u^a}{\partial x^b} + \Gamma^a{}_{bc} u^c$ *är en tensor om och endast om* $\Gamma^a{}_{bc}$ *uppfyller transformationslagen*

$$\bar{\Gamma}^a{}_{bc} = \frac{\partial \bar{x}^a}{\partial x^d} \frac{\partial x^e}{\partial \bar{x}^b} \frac{\partial x^f}{\partial \bar{x}^c} \Gamma^d{}_{ef} - \frac{\partial x^d}{\partial \bar{x}^b} \frac{\partial x^e}{\partial \bar{x}^c} \frac{\partial^2 \bar{x}^a}{\partial x^d \partial x^e} \tag{7.9}$$

Bevis. Övning. □

Anmärkning: Observera att transformationslagen (7.9) är identisk med (7.7). □

($\star$ Om du har läst kapitel 6 behöver vi undanröja ett möjligt missförstånd. Betrakta avbildningen $(\alpha_a, v^b, u^c) \mapsto \Gamma^i{}_{jk} \alpha_i v^j u^k$ där α_i är komponenterna till kovektorn α_a o s v. Denna avbildning är förstås linjär i alla sina argument och därför, enligt Definition 6.14, en tensor. Samma sak gäller

[5] Här varierar terminologin mellan olika skrifter igen. Som tidigare nämnts kallas $v^b \nabla_b u^a$ i vissa skrifter för *kovarianta derivatan* av u^a med avseende på v^b och ∇_b kallas då en *förbindelse*. Koefficienterna $\Gamma^a{}_{bc}$ kallas i detta fall *förbindelsekoefficienter*.

förstås för avbildningen $(\alpha_a, v^b) \mapsto \frac{\partial u^i}{\partial x^j}\alpha_i v^j$. Båda termerna i (7.5) är alltså tensorer vilket strider mot det som sas ovan.

Denna motsägelse är dock bara skenbar. Den är en konsekvens av skillnaden i synsätt mellan kapitel 6 och övriga kapitel. Med synsättet i kapitel 6 är det helt korrekt att $\frac{\partial u^a}{\partial x^b}$ med komponenter $\frac{\partial u^i}{\partial x^j}$ i koordinatsystemet x^i är en tensor. Denna tensors komponenter i ett annat koordinatsystem $\bar{x}^i$ ges dock av

$$\frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^l}{\partial \bar{x}^j}\frac{\partial u^k}{\partial x^l} \neq \frac{\partial \bar{u}^i}{\partial \bar{x}^j}.$$

enligt transformationslagen (4.11) för en tensor av typ $(1, 1)$.

Med detta synsätt är alltså både $\frac{\partial u^i}{\partial x^j}$ och $\frac{\partial \bar{u}^i}{\partial \bar{x}^j}$ komponenter till tensorer, men de är komponenter till *olika* tensorer. Med samma resonemang inser man att ${\Gamma^i}_{jk}$ och ${\bar{\Gamma}^i}_{jk}$ också blir komponenter till olika tensorer.

Sammanfattningsvis: Med synsättet i kapitel 6 är alltså båda termerna i ekvationen $\nabla_b u^a = \frac{\partial u^a}{\partial x^b} + {\Gamma^a}_{bc}u^c$ visserligen tensorer men om vi byter koordinatsystem, så att

$$\nabla_b u^a = \bar{\nabla}_b \bar{u}^a = \frac{\partial \bar{u}^a}{\partial \bar{x}^b} + {\bar{\Gamma}^a}_{bc}\bar{u}^c$$

så är termerna i detta uttryck helt andra tensorer än termerna i vårt ursprungliga uttryck.

Förvirrande? Kanske det, men här finns också en viktig lärdom att dra. En matematisk teori är något som vi själva bygger upp genom de definitioner vi gör. Dessa definitioner får konsekvenser, och dessa kallar vi satser. Hade vi definierat våra begrepp på något annat sätt hade vi också fått andra satser. Detta betyder att det inte finns något 'korrekt' sätt att definiera ett matematiskt begrepp, t ex en tensor. Det betyder bara att vi får leva med den definition vi väljer och hålla till godo med de satser och resultat vi faktiskt kan bevisa.

Den till synes naturliga frågan 'Men hur är det egentligen? Är ${\Gamma^a}_{bc}$ en tensor eller ej?' är alltså felställd. Med synsättet i kapitel 6 är svaret 'ja' och med synsättet i övriga kapitel är svaret 'nej' och detta är ingen motsägelse utan en konsekvens av att vi byggt upp teorin på lite olika sätt i de båda framställningarna. $\star$)

Av transformationslagen (7.9) följer att om vi vet ${\Gamma^a}_{bc}$ i ett koordinatsystem x^a, så vet vi också hur ${\bar{\Gamma}^a}_{bc}$ ser ut i ett godtyckligt koordinatsystem $\bar{x}^a$. Däremot är ${\Gamma^a}_{bc}$:s värden i det ursprungliga koordinatsystemet x^a fortfarande helt godtyckliga. Kovarianta derivatan av u^a är alltså långt ifrån att vara entydigt bestämd av Definition 7.5 eftersom denna definition innehåller n^3 (där som vanligt $n = \dim M$) helt godtyckliga funktioner. Vårt nästa projekt blir därför att lägga på ytterligare villkor på ∇_b så att denna (och därmed ${\Gamma^a}_{bc}$) blir entydigt bestämd.

Vi börjar med en definition.

Definition 7.7. ${T^a}_{bc} = {\Gamma^a}_{bc} - {\Gamma^a}_{cb}$ kallas *torsionstensorn* till förbindelsen ${\Gamma^a}_{bc}$.

Vi ser att torsionstensorn är antisymmetrisk i sina nedre index, d v s ${T^a}_{bc} = -{T^a}_{cb}$. Om nu ${T^a}_{bc}$ *inte* är en tensor så är detta en mycket dålig definition. Det är därför säkrast att verifiera att ${T^a}_{bc}$ verkligen *är* en tensor innan vi går vidare.

Proposition 7.8. *Torsionstensorn* $T^a{}_{bc} = \Gamma^a{}_{bc} - \Gamma^a{}_{cb}$ *är en tensor av typ* $(1,2)$.

Bevis. Vid byte till nya koordinater $\bar{x}^a$ ger transformationslagen (7.9) att

$$\begin{aligned}\bar{T}^a{}_{bc} = \bar{\Gamma}^a{}_{bc} - \bar{\Gamma}^a{}_{cb} &= \frac{\partial \bar{x}^a}{\partial x^d}\frac{\partial x^e}{\partial \bar{x}^b}\frac{\partial x^f}{\partial \bar{x}^c}\Gamma^d{}_{ef} - \frac{\partial x^d}{\partial \bar{x}^b}\frac{\partial x^e}{\partial \bar{x}^c}\frac{\partial^2 \bar{x}^a}{\partial x^d \partial x^e} - \frac{\partial \bar{x}^a}{\partial x^d}\frac{\partial x^f}{\partial \bar{x}^c}\frac{\partial x^e}{\partial \bar{x}^b}\Gamma^d{}_{fe} + \frac{\partial x^e}{\partial \bar{x}^c}\frac{\partial x^d}{\partial \bar{x}^b}\frac{\partial^2 \bar{x}^a}{\partial x^e \partial x^d} \\ &= \frac{\partial \bar{x}^a}{\partial x^d}\frac{\partial x^e}{\partial \bar{x}^b}\frac{\partial x^f}{\partial \bar{x}^c}\left(\Gamma^d{}_{ef} - \Gamma^d{}_{fe}\right) = \frac{\partial \bar{x}^a}{\partial x^d}\frac{\partial x^e}{\partial \bar{x}^b}\frac{\partial x^f}{\partial \bar{x}^c}T^d{}_{ef},\end{aligned}$$

d v s $T^a{}_{bc}$ uppfyller transformationslagen för en tensor av typ $(1,2)$. □

I Exempel 7.3 och 7.4 beräknade vi explicita uttryck för $\Gamma^a{}_{bc}$ under antagandet att kovariant derivering ges av 'vanlig' derivering i $\mathbf{R}^3$ följt av ortogonalprojektion på ytorna S respektive Y. I båda fallen blev $\Gamma^a{}_{bc}$ symmetrisk, d v s $\Gamma^a{}_{bc} = \Gamma^a{}_{cb}$ vilket innebär att torsionstensorn $T^a{}_{bc} = 0$ i båda dessa exempel. Det är därför naturligt att kräva att $T^a{}_{bc} = 0$ även på en allmän mångfald.

Så småningom kommer vi också att kräva att $T^a{}_{bc} = 0$ och alltså inskränka oss till att studera symmetriska förbindelser, men låt oss vänta en liten stund med att lägga på detta krav. Torsionstensorn dyker nämligen upp i en del andra sammanhang och det är instruktivt att se hur.

Låt $\mathbf{Y}$ och $\mathbf{Z}$ vara två vektorfält med komponenter Y^a respektive Z^a. Kommutatorn (se Definition 4.14) $[\mathbf{Y},\mathbf{Z}]$ av $\mathbf{Y}$ och $\mathbf{Z}$ definierades av att $[\mathbf{Y},\mathbf{Z}](f) = \mathbf{Y}(\mathbf{Z}(f)) - \mathbf{Z}(\mathbf{Y}(f))$ för alla funktioner f. Enligt Sats 4.15 är $[\mathbf{Y},\mathbf{Z}]$ ett vektorfält med komponenter

$$[\mathbf{Y},\mathbf{Z}]^a = Y^b\frac{\partial Z^a}{\partial x^b} - Z^b\frac{\partial Y^a}{\partial x^b}.$$

och det gäller att $[\mathbf{Y},\mathbf{Z}] = -[\mathbf{Z},\mathbf{Y}]$.

Kommutatorn till $\mathbf{Y}$ och $\mathbf{Z}$ är alltså ett nytt vektorfält som är antisymmetriskt i $\mathbf{Y}$ och $\mathbf{Z}$ dess komponenter bildas från förstaderivatorna av Y^a och Z^a. Observera att precis samma sak kan sägas om vektorfältet $\nabla_{\mathbf{Y}}\mathbf{Z} - \nabla_{\mathbf{Z}}\mathbf{Y}$, med komponenter $Y^b\nabla_b Z^a - Z^b\nabla_b Y^a$, och det är naturligt att fråga sig om det finns något samband mellan dessa vektorfält. Låt oss därför betrakta

$$\begin{aligned}Y^b\nabla_b Z^a - Z^b\nabla_b Y^a &= Y^b\frac{\partial Z^a}{\partial x^b} + \Gamma^a{}_{bc}Y^bZ^c - Z^b\frac{\partial Y^a}{\partial x^b} - \Gamma^a{}_{bc}Z^bY^c \\ &= Y^b\frac{\partial Z^a}{\partial x^b} - Z^b\frac{\partial Y^a}{\partial x^b} + \Gamma^a{}_{bc}Y^bZ^c - \Gamma^a{}_{cb}Z^cY^b = [\mathbf{Y},\mathbf{Z}]^a + T^a{}_{bc}Y^bZ^c.\end{aligned}$$

Det gäller alltså att

$$T^a{}_{bc}Y^bZ^c = Y^b\nabla_b Z^a - Z^b\nabla_b Y^a - [\mathbf{Y},\mathbf{Z}]^a,$$

för alla Y^b och Z^c. Detta visar

Proposition 7.9. *Om torsionstensorn* $T^a{}_{bc} = 0$ *d v s om* $\Gamma^a{}_{bc}$ *är symmetrisk så gäller att*

$$\nabla_{\mathbf{Y}}\mathbf{Z} - \nabla_{\mathbf{Z}}\mathbf{Y} = [\mathbf{Y},\mathbf{Z}]$$

för alla vektorfält $\mathbf{Y}$ *och* $\mathbf{Z}$. *I komponenter:*

$$Y^b\nabla_b Z^a - Z^b\nabla_b Y^a = [\mathbf{Y},\mathbf{Z}]^a.$$

Hittills har vår kovarianta derivata ∇_b bara verkat på vektorfält men vi vill förstås kunna derivera tensorfält av godtycklig typ. För att en sådan utvidgning ska bli entydig måste vi ställa lite krav på ∇_a och följande krav verkar naturliga:

- Additivitet: $\nabla_a(S+T) = \nabla_a S + \nabla_a T$ för godtyckliga tensorfält S och T.
- Leibniz regel: $\nabla_a(ST) = S\nabla_a T + T\nabla_a S$ för godtyckliga tensorfält S och T.
- ∇_a kommuterar med kontraktion. T ex ska det alltså gälla att

$$\nabla_a({T_b}^b) = \text{kontraktionen av } \nabla_a {T_b}^c \text{ över indexen } b \text{ och } c.$$

- $\nabla_a f = \frac{\partial f}{\partial x^a}$ = differentialen av f för alla skalära funktioner f.

Observera att det sista kravet är Definition 4.7 och alltså inget nytt. Vi tar med det i denna lista för fullständighets skull. Observera också att homogenitet (och därmed linearitet) är en konsekvens av dessa krav.

Givet ${\Gamma^a}_{bc}$ bestämmer dessa krav samt ekvation (7.8) entydigt hur ∇_a verkar på ett godtyckligt tensorfält. Vi härleder ett uttryck för derivatan av ett godtyckligt kovariant tensorfält α_a. Låt u^a vara ett godtyckligt vektorfält. P g a kraven ovan gäller då

$$\nabla_b(\alpha_a u^a) = u^a \nabla_b \alpha_a + \alpha_a \nabla_b u^a = u^a \nabla_b \alpha_a + \alpha_a \frac{\partial u^a}{\partial x^b} + \alpha_a {\Gamma^a}_{bc} u^c = u^a \nabla_b \alpha_a + \alpha_a \frac{\partial u^a}{\partial x^b} + \alpha_c {\Gamma^c}_{ba} u^a.$$

Å andra sidan är $\alpha_a u^a$ en vanlig skalär funktion vilket ger

$$\nabla_b(\alpha_a u^a) = \frac{\partial}{\partial x^b}(\alpha_a u^a) = \alpha_a \frac{\partial u^a}{\partial x^b} + u^a \frac{\partial \alpha_a}{\partial x^b}.$$

Identifikation av dessa uttryck ger nu att

$$u^a \nabla_b \alpha_a = u^a \frac{\partial \alpha_a}{\partial x^b} - {\Gamma^c}_{ba} \alpha_c u^a,$$

för alla vektorer u^a. Det följer att

$$\nabla_b \alpha_a = \frac{\partial \alpha_a}{\partial x^b} - {\Gamma^c}_{ba} \alpha_c. \tag{7.10}$$

Vi kan nu härleda ännu en viktig egenskap hos torsionstensorn. Låt f vara en godtycklig funktion och betrakta den kovarianta andraderivatan

$$\nabla_a(\nabla_b f) = \nabla_a \left(\frac{\partial f}{\partial x^b} \right) = \frac{\partial^2 f}{\partial x^a \partial x^b} - {\Gamma^c}_{ab} \frac{\partial f}{\partial x^c}$$

Den blandade kovarianta andraderivatan ges alltså av

$$(\nabla_a \nabla_b - \nabla_b \nabla_a) f = \frac{\partial^2 f}{\partial x^a \partial x^b} - {\Gamma^c}_{ab} \frac{\partial f}{\partial x^c} - \frac{\partial^2 f}{\partial x^b \partial x^a} + {\Gamma^c}_{ba} \frac{\partial f}{\partial x^c} = {T^c}_{ba} \nabla_c f,$$

vilket bevisar

Proposition 7.10. *Om torsionstensorn $T^a{}_{bc} = 0$ d v s om $\Gamma^a{}_{bc}$ är symmetrisk så gäller att*

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f \tag{7.11}$$

för alla funktioner f av klass $\mathcal{C}^2$.

Vi sammanfattar det vi hittills vet om torsionstensorn $T^a{}_{bc} = \Gamma^a{}_{bc} - \Gamma^a{}_{cb}$:

- Om Y är en yta i $\mathbf{R}^3$ och riktningsderivatan av ett vektorfält definieras som ortogonalprojektionen på T_pY av den 'vanliga' riktningsderivatan i $\mathbf{R}^3$ så blir förbindelsen $\Gamma^a{}_{bc}$ symmetrisk, d v s torsionstensorn $T^a{}_{bc} = 0$.

- Om $\Gamma^a{}_{bc}$ är symmetrisk så att torsionstensorn $T^a{}_{bc} = 0$ så gäller

$$\nabla_{\mathbf{Y}} \mathbf{Z} - \nabla_{\mathbf{Z}} \mathbf{Y} = [\mathbf{Y}, \mathbf{Z}]$$

  för alla vektorfält $\mathbf{Y}$ och $\mathbf{Z}$, d v s

$$Y^b \nabla_b Z^a - Z^b \nabla_b Y^a = [\mathbf{Y}, \mathbf{Z}]^a$$

  enligt Proposition 7.9.

- Om $\Gamma^a{}_{bc}$ är symmetrisk så att torsionstensorn $T^a{}_{bc} = 0$ så gäller

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f$$

  för alla funktioner f av klass $\mathcal{C}^2$, enligt Proposition 7.10.

Vi ser att vi har starka skäl att kräva att förbindelsen $\Gamma^a{}_{bc}$ är symmetrisk så att torsionstensorn $T^a{}_{bc} = 0$. Låt oss därför från och med nu anta att alla förbindelser $\Gamma^a{}_{bc}$ är symmetriska om inget annat sägs[6].

Antalet oberoende okända komponenter i $\Gamma^a{}_{bc}$ minskar då från n^3 till $\frac{n^2(n+1)}{2}$.

Vi har hittills undersökt hur kovarianta derivatan ∇_a verkar på kontra- respektive kovarianta tensorer av rang 1. Låt oss gå vidare till tensorer av högre rang, t ex en tensor T^{ab} av typ $(2, 0)$. Låt oss först titta på ett specialfall, nämligen en tensor på den enkla formen $T^{ab} = u^a v^b$ där u^a och v^b är godtyckliga vektorfält.

Leibniz regel ger då

$$\begin{aligned}
\nabla_c T^{ab} &= \nabla_c (u^a v^b) = v^b \nabla_c u^a + u^a \nabla_c v^b = v^b \left(\frac{\partial u^a}{\partial x^c} + \Gamma^a{}_{cd} u^d \right) + u^a \left(\frac{\partial v^b}{\partial x^c} + \Gamma^b{}_{cd} v^d \right) \\
&= \frac{\partial}{\partial x^c} (u^a v^b) + \Gamma^a{}_{cd} u^d v^b + \Gamma^b{}_{cd} u^a v^d = \frac{\partial T^{ab}}{\partial x^c} + \Gamma^a{}_{cd} T^{db} + \Gamma^b{}_{cd} T^{ad}.
\end{aligned}$$

[6] I allmän relativitetsteori är förbindelsen alltid symmetrisk, men det har gjorts försök att generalisera allmän relativitetsteori genom att tillåta nollskild torsion, t ex i så kallad *Einstein-Cartanteori*. Se t ex *Penrose-Rindler:* Spinors and spacetime, vol 1.

Vi observerar att högerledet är additivt (t o m linjärt) i T och då en godtycklig tensor T^{ab} kan skrivas som en summa av ändligt många tensorer av ovanstående enkla form (se till att du inser varför!) följer det att

$$\nabla_c T^{ab} = \frac{\partial T^{ab}}{\partial x^c} + \Gamma^a{}_{cd} T^{db} + \Gamma^b{}_{cd} T^{ad}, \tag{7.12}$$

för alla tensorer T^{ab} av typ $(2,0)$.

Anmärkning: Det är visserligen så att en given tensor T^{ab} kan skrivas som en ändlig summa av termer på formen $u^a v^b$ på många olika sätt, men det är inte svårt att visa att (gör gärna detta själv som övning!) högerledet i ekvation (7.12) ger samma resultat i alla dessa fall. □

Observera strukturen i denna formel. Om du håller för indexet b och bara läser de två första termerna är formeln identisk med formeln (7.8) för kovarianta derivatan av ett vektorfält T^a. Om du istället håller för indexet a och läser första och tredje termen fås istället kovarianta derivatan av ett vektorfält T^b.

Förutom den partiella derivatan av T^{ab} får vi alltså två termer i vilka förbindelsen $\Gamma^a{}_{bc}$ kontraheras med ett av T:s index i taget.

Samma mönster upprepas om vi istället deriverar en tensor av typ $(0,2)$.

Övning: Visa att

$$\nabla_c T_{ab} = \frac{\partial T_{ab}}{\partial x^c} - \Gamma^d{}_{ca} T_{db} - \Gamma^d{}_{cb} T_{ad},$$

för alla tensorer T_{ab} av typ $(0,2)$. Kovarianta derivatan av en tensor av godtycklig typ följer också samma mönster:

SATS 7.11. *Om $T^{a\ldots b}{}_{c\ldots d}$ är ett godtyckligt tensorfält så gäller*

$$\nabla_e T^{a\ldots b}{}_{c\ldots d} = \frac{\partial T^{a\ldots b}{}_{c\ldots d}}{\partial x^e} + \Gamma^a{}_{ef} T^{f\ldots b}{}_{c\ldots d} + \ldots + \Gamma^b{}_{ef} T^{a\ldots f}{}_{c\ldots d} - \Gamma^f{}_{ec} T^{a\ldots b}{}_{f\ldots d} - \ldots - \Gamma^f{}_{ed} T^{a\ldots b}{}_{c\ldots f}. \tag{7.13}$$

7.5 Parallelltransport samt enhetssfären, del 5

I vår strävan att göra förbindelsen $\Gamma^a{}_{bc}$ entydigt bestämd har vi så här långt bestämt oss för att $\Gamma^a{}_{bc}$ ska uppfylla transformationslagen (7.9) samt vara symmetrisk $\Gamma^a{}_{bc} = \Gamma^a{}_{cb}$ vilket lämnar $\Gamma^a{}_{bc}$ med $\frac{n^2(n+1)}{2}$ frihetsgrader (där som vanligt $n = \dim M$).

I detta avsnitt ska vi lägga på ytterligare villkor på $\Gamma^a{}_{bc}$ och därmed uppnå målet entydighet. Vi ska därför se på ett geometriskt exempel för att hitta lämpliga villkor.

Exempel 7.12. Betrakta $\mathbf{R}^n$. Låt x^a vara kartesiska koordinater på $\mathbf{R}^n$ och låt $\gamma : x^a = x^a(\tau)$ vara en kurva i $\mathbf{R}^n$. Vektorfältet $X^a = \frac{dx^a}{d\tau}$ är då en tangentvektor till γ.

Låt p vara en punkt på γ. Om $\mathbf{u}_0$ är en tangentvektor i p kan denna vektor, på ett naturligt sätt, parallelltransporteras (d v s parallellförflyttas) längs γ. Vi erhåller på detta sätt ett vektorfält $\mathbf{u}$ på γ sådant att $\mathbf{u}|_p = \mathbf{u}_0$ och vektorfältets komponenter u^a i det kartesiska koordinatsystemet x^a, är konstanta längs γ.

Observera att om vi byter till godtyckliga koordinater $\bar{x}^a$ så blir $\mathbf{u}$:s nya komponenter $\bar{u}^a$ *inte* nödvändigtvis konstanta. Konstanta komponenter är därför inte ett koordinatoberoende sätt att karakterisera ett parallelltransporterat vektorfält.

Vi vill nu hitta en karakterisering av ett parallelltransporterat vektorfält som inte förutsätter kartesiskt koordinatsystem. Låt därför $\mathbf{u}$ vara ett vektorfält på γ med komponenter u^a i koordinatbasen till x^a. Om u^a är parallelltransporterat längs γ är alltså u^a konstanta på γ, d v s $\frac{du^a}{d\tau} = 0$. Kedjeregeln ger då att

$$0 = \frac{du^a}{d\tau} = \frac{dx^b}{d\tau}\frac{\partial u^a}{\partial x^b} = X^b \partial_b u^a \Leftrightarrow \partial_{\mathbf{X}} \mathbf{u} = 0.$$

I $\mathbf{R}^n$ gäller alltså att ett vektorfält $\mathbf{u}$ med komponenter u^a i ett kartesiskt koordinatsystem är parallelltransporterat längs en kurva γ med tangentvektor $\mathbf{X}$ om och endast om riktningsderivatan av $\mathbf{u}$ med avseende på $\mathbf{X}$ är noll, d v s $\partial_{\mathbf{X}} \mathbf{u} = 0$. □

Detta sista resultat lämpar sig väl för generalisering till en godtycklig mångfald:

Definition 7.13. Om ∇_c är en kovariant derivata på en godtycklig mångfald M sägs vektorfältet u^a vara *parallelltransporterat* längs en kurva $\gamma \subset M$ med tangentvektor X^c om $X^c \nabla_c u^a = 0$.

Med denna definition kan också en vektor i en punkt $p \in \gamma$ på ett entydigt sätt parallelltransporteras längs γ:

SATS 7.14. *Givet en tangentvektor u_0^a i en punkt p på en kurva γ med parametrisering $x^a = x^a(\tau)$ så finns ett entydigt bestämt vektorfält u^a på γ sådant att $u^a|_p = u_0^a$ och u^a är parallelltransporterat längs γ.*

Bevis. Enligt definitionen av parallelltransport ska det gälla att

$$0 = \frac{dx^c}{d\tau}\nabla_c u^a = \frac{dx^c}{d\tau}\left(\frac{\partial u^a}{\partial x^c} + \Gamma^a{}_{bc} u^b\right) = \frac{du^a}{d\tau} + \Gamma^a{}_{bc}\frac{dx^c}{d\tau} u^b.$$

Funktionerna u^a löser alltså ett linjärt differentialekvationssystem och då även ett begynnelsevillkor (d v s $u^a|_p = u_0^a$) är givet följer det av teori för differentialekvationer att lösningen är entydig. u^a bestäms alltså entydigt av de givna villkoren.

Detta visar satsen i fallet att *hela* γ ligger i någon öppen mängd som kan beskrivas med ett enda koordinatsystem. Det återstår alltså att undersöka fallet att γ är för 'lång' så att det inte finns någon sådan öppen mängd. Vi nöjer oss med att skissa den delen av resonemanget.

Låt $q \in \gamma$ vara godtycklig och betrakta den del av γ som förbinder p och q. Dela upp denna del i överlappande bitar som var och en ligger i sådana öppna mängder. Man kan visa att det räcker med ändligt många sådana bitar.

Ovanstående resonemang visar nu att u_0^a kan parallellförflyttas längs den första av dessa bitar på ett entydigt sätt. Då första och andra biten överlappar kan därmed parallelltransporten fortsättas entydigt längs andra biten o s v.

Detta visar att parallelltransport är väldefinierad och entydigt bestämd mellan p och den godtyckliga punkten q och satsen följer.

□

Anmärkning: Ovanstående resultat måste tolkas med viss försiktighet om γ skär sig själv, d v s om $x^a(\tau_1) = x^a(\tau_2)$ för några $\tau_1 \neq \tau_2$. Noga räknat blir då u^a väldefinierat och entydigt bestämt, sett som funktion av τ men det är fullt möjligt (t o m högst troligt) att u^a inte blir en väldefinierad funktion av punkterna på γ, nämligen genom att $u^a(\tau_1) \neq u^a(\tau_2)$ trots att parametervärdena τ_1 och τ_2 svarar mot samma punkt på γ.

□

En konsekvens av detta är att parallelltransport ofta är *beroende av vägen.*

Exempel 7.15. *(Enhetssfären, del 5)* Betrakta enhetssfären S och låt p, q och r vara punkterna med kartesiska koordinater $(x, y, z) = (1, 0, 0)$, $(0, 1, 0)$ resp. $(0, 0, 1)$. Låt γ_1 vara kvartscirkelbågen från p till r.

Då är vektorn $\bar{e}_z$ tangent till γ_1 i p. Om denna vektor parallelltransporteras längs γ_1 till r resulterar detta i vektorn $-\bar{e}_x$ i r. Detta kan räknas ut med formeln ovan eftersom vi härlett uttryck för Christoffelsymbolerna i Exempel 7.3, men det bör också vara intuitivt klart. Om inte, läs igenom tankeexperimentet som börjar direkt efter detta exempel innan du fortsätter.

Låt nu γ_2 och γ_3 vara kvartscirkelbågarna från p till q resp. från q till r. Om vi parallelltransporterar $\bar{e}_z$ i p längs γ_2 fås på samma sätt vektorn $\bar{e}_z$ i q och om denna parallelltransporteras längs γ_3 fås vektorn $-\bar{e}_y$ i r, d v s *inte* samma vektor som när vi parallelltransporterade $\bar{e}_z$ i p till punkten r längs kurvan γ_1. Parallelltransport är alltså beroende av vägen mellan p och r.

Betrakta nu parallelltransport längs den slutna kurvan $-\gamma_1 + \gamma_2 + \gamma_3$. Starta med vektorn $-\bar{e}_x$ i punkten r. Parallelltransport av denna längs $-\gamma_1$ ger enligt ovan vektorn $\bar{e}_z$ i p och parallelltransport längs $\gamma_2 + \gamma_3$ ger enligt ovan vektorn $-\bar{e}_y \neq -\bar{e}_x$ i r.

Det resulterande parallelltransporterade vektorfältet är alltså *inte* är entydigt bestämt av *punkterna* på $\gamma := -\gamma_1 + \gamma_2 + \gamma_3$, men väl av *parametervärdet* i en tänkt parametrisering av γ. □

Låt oss nu göra ett tankeexperiment. Vi ser på specialfallet att Y är en 2-dimensionell yta i $\mathbf{R}^3$ och vi tänker oss att vi väljer ut en tangentvektor $u_0 \in T_pY$ där $p \in Y$ och parallelltransporterar den längs någon kurva γ med tangentvektor X^a genom p. Enligt satsen ovan får vi då ett entydigt bestämt vektorfält u^a, definierat på γ sådant att $u^a\big|_p = u_0$ och $X^b \nabla_b u^a = 0$ på hela γ.

Enligt den geometriska tolkningen av kovariant derivata på en yta i Exempel 7.4 betyder denna ekvation att den 'vanliga' derivatan i $\mathbf{R}^3$ av u^a överallt är ortogonal mot Y, d v s sett som ett vektorfält i $\mathbf{R}^3$ kommer u^a *endast* att variera så att u^a 'följer med' alla ojämnheter längs γ och hela tiden förblir tangentiell till Y. u^a får alltså *inte* variera på något annat sätt.

T ex kan vi tänka på Y som ett landskap fullt av berg och dalar och på γ som en väg genom landskapet. Om vi dessutom tänker på en bil som kör med helt jämn fart längs vägen γ kan vi tänka på vektorfältet u^a som något avlångt föremål (t ex en bräda) fastsatt på bilens tak.

Jag slår vad om att om du skulle titta på brädan före och efter parallelltransporten skulle du bli mäkta förvånad om brädans längd hade ändrats under parallelltransporten. Lika förvånad tror jag du skulle bli om du istället spikade ihop två brädor så att de bildar en vinkel α mellan sig, och sedan fann att den vinkeln ändrats under parallelltransporten.

Ett annat sätt att säga ovanstående är att vi förväntar oss att parallelltransport ska bevara längder och vinklar, d v s parallelltransport ska bevara skalärprodukter. Vi vill nu undersöka om denna förväntan är sann och följande lemma är då till hjälp:

Lemma 7.16. *Om Y är en yta i $\mathbf{R}^3$ är följande ekvivalenta:*

- *Parallelltransport bevarar skalärprodukter på Y, d v s om γ är en godtycklig kurva på Y med parameterframställning $x^a = x^a(\tau)$ och tangentvektor $X^b = \frac{dx^b}{d\tau}$ så är $g_{ab}u^a v^b$ oberoende av τ om $X^b\nabla_b u^a = X^b\nabla_b v^a = 0$ på hela γ, d v s om u^a och v^b är parallelltransporterade längs γ.*
- *$\nabla_c g_{ab} = 0$ på hela Y*

Bevis. Låt $p \in Y$ och låt γ vara en godtycklig kurva genom p. Då ger kedjeregeln att

$$\frac{d}{d\tau}(g_{ab}u^a v^b) = \frac{dx^c}{d\tau}\frac{\partial}{\partial x^c}(g_{ab}u^a v^b) = X^c u^a v^b \nabla_c g_{ab} + g_{ab}v^b X^c \nabla_c u^a + g_{ab}u^a X^c \nabla_c v^b,$$

och vi ser att de två sista termerna är noll eftersom både u^a och v^b är parallelltransporterade längs γ. $g_{ab}u^a v^b$ är alltså oberoende av τ om och endast om

$$X^c u^a v^b \nabla_c g_{ab} = 0,$$

för alla u^a och v^b som är parallelltransporterade längs γ. Detta är självklart uppfyllt om $\nabla_c g_{ab} = 0$ vilket visar ena riktningen.

För att visa andra riktningen använder vi Sats 7.14. Enligt denna följer att $X^c u^a v^b \nabla_c g_{ab} = 0$ för alla tangentvektorer u^a och v^b i p d v s $X^c \nabla_c g_{ab} = 0$ i p och då γ var godtycklig följer att $\nabla_c g_{ab} = 0$ i p som också var godtycklig och lemmat följer. □

Ett nödvändigt och tillräckligt villkor för att parallelltransport ska bevara skalärprodukter är alltså att $\nabla_c g_{ab} = 0$ på hela M. Men med hjälp av Sats 7.11 och ekvation (7.5) kan vi beräkna $\nabla_c g_{ab}$ (Gör detta själv som övning!).

Man finner då att villkoret $\nabla_c g_{ab} = 0$ faktiskt är uppfyllt överallt på Y och parallelltransport bevarar alltså skalärprodukter, som förväntat.

På en allmän mångfald M gör vi därför följande definition:

Definition 7.17. En förbindelse $\Gamma^a{}_{bc}$ sägs vara *metrisk* om den tillhörande kovarianta derivatan ∇_c uppfyller $\nabla_c g_{ab} = 0$ överallt på M.

Av resonemanget ovan följer

SATS 7.18. *På en yta Y i $\mathbf{R}^3$ är förbindelsen given av Christoffelsymbolerna (d v s ekvation (7.5)) både metrisk och symmetrisk.*

Målet med hela denna diskussion var att hitta lämpliga villkor på förbindelsen $\Gamma^a{}_{bc}$ som gör den entydigt bestämd. I och med följande sats är detta mål nått.

SATS 7.19. *På en godtycklig mångfald M finns en entydigt bestämd metrisk och symmetrisk förbindelse. Denna ges av Christoffelsymbolerna*

$$\Gamma^a{}_{bc} = \frac{1}{2}g^{ad}\left(\frac{\partial g_{cd}}{\partial x^b} + \frac{\partial g_{db}}{\partial x^c} - \frac{\partial g_{bc}}{\partial x^d}\right).$$

Bevis. Detaljerna i beviset lämnas som övning, men följande stolpar kan vara till hjälp

- Börja med entydigheten. Antag alltså att $\Gamma^a{}_{bc}$ är symmetrisk och metrisk så att $\nabla_c g_{ab} = 0$.
- Lös ut $\frac{\partial g_{ab}}{\partial x^c}$ ur detta samband.
- Teckna $\frac{\partial g_{cd}}{\partial x^b}$, $\frac{\partial g_{db}}{\partial x^c}$ och $-\frac{\partial g_{bc}}{\partial x^d}$ med hjälp av detta samband och addera.
- Lös ut $\Gamma^a{}_{bc}$ ur detta samband. Var noga med att visa att lösningen är entydig. Detta visar entydigheten.
- För existensen, låt uttrycket från förra punkten definiera $\Gamma^a{}_{bc}$. Visa sedan att $\Gamma^a{}_{bc} = \Gamma^a{}_{cb}$ och $\nabla_c g_{ab} = 0$ följer av denna definition.

□

Definition 7.20. Den entydigt bestämda metriska och symmetriska förbindelsen given av Christoffelsymbolerna

$$\Gamma^a{}_{bc} = \frac{1}{2} g^{ad} \left(\frac{\partial g_{cd}}{\partial x^b} + \frac{\partial g_{db}}{\partial x^c} - \frac{\partial g_{bc}}{\partial x^d} \right).$$

kallas *Levi-Civita-förbindelsen.*

Anmärkning: Om $\Gamma^a{}_{bc}$ är Levi-Civita-förbindelsen så följer det, med samma resonemang som i beviset av Lemma 7.16 att motsvarande parallelltransport (given av Definition 7.5) bevarar skalärprodukter. □

($\star$ Noga räknat har vi bara visat att Levi-Civita förbindelsen existerar och är entydigt bestämd om hela M kan beskrivas av ett och samma koordinatsystem. Om detta inte går visar ovanstående resonemang bara att Levi-Civita förbindelsen existerar *lokalt.*

Resonemanget kan dock modifieras lite grand, och visar då att så länge metriken själv är globalt definierad (d v s definierad på hela M) så blir också Levi-Civita förbindelsen det. Se t ex *do Carmo*: Riemannian Geometry. $\star$)

7.6 Liederivator, kongruenser och symmetrier

Kapitel 8

Geodeter

8.1 Inledning

Bland alla kurvor i $\mathbf{R}^n$ har de räta linjerna en särställning. De är de enda kurvorna med egenskapen att om p och q är två punkter på kurvan och $\mathbf{u}$ och $\mathbf{v}$ är tangentvektorer till kurvan i p respektive q så är $\mathbf{u}$ och $\mathbf{v}$ med nödvändighet parallella. I detta avsnitt ska vi generalisera denna egenskap hos räta linjer till godtyckliga mångfalder. De kurvor vi då får kommer att kallas *geodeter*.

8.2 Geodeter

Betrakta $\mathbf{R}^n$ försett med ett kartesiskt koordinatsystem x^a och en metrik g_{ab} som får vara av godtycklig signatur. En rät linje L i $\mathbf{R}^n$ är som bekant en kurva med parametrisering $x^a = k^a\tau + m^a$, där k^a, m^a är konstanter sådana att minst ett $k^a \neq 0$.

Observera att med denna parametrisering blir kurvtangenten given av $\frac{dx^a}{d\tau} = k^a$ ett konstant, nollskilt vektorfält längs L. Omvänt: Om kurvtangenten $\frac{dx^a}{d\tau} = k^a$ är konstant och nollskild följer det att $x^a = k^a\tau + m^a$ för några konstanter m^a. Detta visar

Proposition 8.1. *En kurva L i $\mathbf{R}^n$ är en (del av en) rät linje om och endast om kurvtangenten $\frac{dx^a}{d\tau} = k^a$ är ett konstant och nollskilt vektorfält längs L.*

Anmärkning: I fortsättningen kommer vi ofta att betrakta olika sätt att parametrisera en given kurva Γ. Om $x^a = x^a(\sigma)$ är en parametrisering av en kurva Γ och vi inför en ny parameter τ genom $\sigma = \sigma(\tau)$ kommer vi att inskränka oss till att betrakta fallet att både funktionen $\sigma(\tau)$ och dess invers är av klass $\mathcal{C}^1$. Detta innebär speciellt att $\sigma'(\tau) \neq 0$ överallt på Γ (Varför?). □

Låt $x^a = x^a(\sigma)$ och $x^a = x^a(\sigma(\tau))$ vara parametriseringar av samma räta linje L och betrakta tangentvektorn $\frac{dx^a}{d\sigma}$. Kedjeregeln ger då att $\frac{dx^a}{d\sigma} = \frac{dx^a}{d\sigma}\sigma'(\tau)$ d v s $\frac{dx^a}{d\sigma}$ är en konstant vektor om och endast om $\sigma'(\tau)$ är en nollskild konstant.

Två godtyckliga parametrar τ och σ sådana att båda tangentvektorerna $\frac{dx^a}{d\sigma}$ och $\frac{dx^a}{d\tau}$ är nollskilda, konstanta vektorfält, uppfyller alltså sambandet $\sigma = a\tau + b$ där $a \neq 0$ och b är konstanter. Ett sådant samband kallas *affint*. Vi gör därför följande definition:

Definition 8.2. Låt $x^a = x^a(\tau)$ vara en parametrisering av en rät linje L i $\mathbf{R}^n$. τ sägs vara en *affin parameter* till L om tangentvektorn $\frac{dx^a}{d\tau}$ är en konstant vektor.

Resonemanget ovan visar då att

Proposition 8.3. *Om $x^a = x^a(\tau)$ är en parametrisering av en rät linje L i $\mathbf{R}^n$ och τ är en affin parameter så ges alla affina parametrar σ till L av $\sigma = a\tau + b$ där $a \neq 0$ och b är konstanter.*

Observera att om $x^a = x^a(\tau)$ är en affin parametrisering av en rät linje L så att $X^a = \frac{dx^a}{d\tau}$ är konstant på L så är $\frac{d^2x^a}{d\tau^2} = 0$ d v s

$$0 = \frac{d^2x^a}{d\tau^2} = \frac{d}{d\tau}\left(\frac{dx^a}{d\tau}\right) = \frac{dx^b}{d\tau}\frac{\partial}{\partial x^b}\left(\frac{dx^a}{d\tau}\right) = X^b\partial_b X^a.$$

Följande gäller alltså:

Proposition 8.4. *Om den räta linjen L har parametriseringen $x^a = x^a(\tau)$ och $X^a = \frac{dx^a}{d\tau}$ så är τ en affin parameter om och endast om $X^b\partial_b X^a = 0$ d v s om linjens tangentvektor $X^a = \frac{dx^a}{d\tau}$ är parallelltransportertad längs L.*

Detta resultat följer också direkt av utredningen i Exempel 7.12.

Nu är det förstås ingen naturlag att linjer måste parametriseras med affina parametrar, så låt oss undersöka vad som gäller för en godtycklig parameter. Låt därför $x^a = x^a(\sigma)$ vara en affin parametrisering av en rät linje L så att $\frac{d^2x^a}{d\sigma^2} = 0$. Inför en ny, godtycklig, parameter τ genom parameterbytet $\sigma = \sigma(\tau)$.

$x^a = x^a(\sigma(\tau))$ är då en ny, inte nödvändigtvis affin, parametrisering av L. Kedjeregeln ger

$$\frac{d^2x^a}{d\tau^2} = \frac{d}{d\tau}\left(\sigma'(\tau)\frac{dx^a}{d\sigma}\right) = \sigma''(\tau)\frac{dx^a}{d\sigma} + \sigma'(\tau)^2\frac{d^2x^a}{d\sigma^2} = \frac{\sigma''(\tau)}{\sigma'(\tau)}\frac{dx^a}{d\tau}$$

Om $X^a = \frac{dx^a}{d\tau}$ gäller alltså att

$$X^b\partial_b X^a = \frac{d^2x^a}{d\tau^2} = \frac{\sigma''(\tau)}{\sigma'(\tau)}\frac{dx^a}{d\tau} = \frac{\sigma''(\tau)}{\sigma'(\tau)}X^a.$$

Detta visar ena riktningen av följande resultat:

Proposition 8.5. *$x^a = x^a(\tau)$ är en parametrisering av en rät linje om och endast om det finns en funktion $\lambda(\tau)$ sådan att vektorfältet $X^a = \frac{dx^a}{d\tau}$ uppfyller $X^b\partial_b X^a = \lambda(\tau)X^a$.*

Anmärkning: Ett ekvivalent sätt att skriva detta villkor är $\frac{d^2x^a}{d\tau^2} = \lambda(\tau)\frac{dx^a}{d\tau}$. □

Bevis. ⇒) Följer direkt av ovanstående genom att sätta $\lambda(\tau) = \frac{\sigma''(\tau)}{\sigma'(\tau)}$.

$\Leftarrow$) Observera att vi är klara om vi hittar ett parameterbyte $\sigma = \sigma(\tau)$ sådant att $\frac{d^2x^a}{d\sigma^2} = 0$. Enligt räkningarna ovan gäller att

$$\lambda(\tau)X^a = X^b\partial_b X^a = \frac{d^2x^a}{d\tau^2} = \sigma''(\tau)\frac{dx^a}{d\sigma} + \sigma'(\tau)^2\frac{d^2x^a}{d\sigma^2} \Leftrightarrow \lambda(\tau)\sigma'(\tau)\frac{dx^a}{d\sigma} = \sigma''(\tau)\frac{dx^a}{d\sigma} + \sigma'(\tau)^2\frac{d^2x^a}{d\sigma^2}.$$

Låt $\Lambda(\tau)$ vara någon primitiv funktion till $\lambda(\tau)$ och sätt $\sigma'(\tau) = e^{\Lambda(\tau)}$. Då är $\sigma''(\tau) = \lambda(\tau)\sigma'(\tau)$ vilket ger $\frac{d^2x^a}{d\sigma^2} = 0$ och resultatet följer. $\square$

Vi skulle alltså kunna ha tagit ekvationen $X^b\partial_b X^a = \lambda(\tau)X^a$, där $X^a = \frac{dx^a}{d\tau}$, som definition av en rät linje. Villkoret för att linjen hade varit affint parametriserad hade då varit att $\lambda(\tau) = 0$.

Detta alternativa sätt att definiera en rät linje går naturligt att generalisera genom att ∂_a, d v s den kovarianta derivatan i $\mathbf{R}^n$ ersätts med den kovarianta derivatan ∇_a på en allmän mångfald. Namnet 'rät linje' är dock (som vi ska se) olämpligt i detta sammanhang. De räta linjernas motsvarigheter på en allmän mångfald kommer att kallas *geodeter*:

Definition 8.6. En kurva Γ med parametrisering $x^a = x^a(\tau)$ och tangentvektor $X^a = \frac{dx^a}{d\tau}$ kallas en *geodet* om det finns en funktion $\lambda(\tau)$ sådan att $X^b\nabla_b X^a = \lambda(\tau)X^a$. Vidare sägs τ vara en *affin parameter* om $\lambda(\tau) = 0$.

Anmärkning: Om $x^a = x^a(\tau)$ är en affint parametriserad geodet Γ är alltså $X^b\nabla_b X^a = 0$, d v s kurvtangenten X^a är parallelltransporterad längs Γ. Räkningar liknande de som redan gjorts ovan visar också (genomför detaljerna själv som övning) att σ är en annan affin parameter längs Γ om och endast om $\sigma = a\tau + b$ för konstanter $a \neq 0$ och b. $\square$

Om Γ är en affint parametriserad geodet är alltså dess tangentvektor X^a parallelltransporterad längs Γ. Enligt anmärkningen direkt efter Definition 7.20 är alltså skalärprodukten $g_{ab}X^aX^b$ konstant längs Γ. Detta innebär speciellt att om $g_{ab}X^aX^b > 0$ i någon punkt $p \in \Gamma$ så är $g_{ab}X^aX^b > 0$ längs hela Γ. Med andra ord: Om vektorfältet X^a är tidsartat i någon punkt på Γ så är X^a tidsartat längs hela Γ.

Vi kontrollerar vad som händer om Γ ej är affint parametriserad, d v s om $X^b\nabla_b X^a = \lambda(\tau)X^a$. Produktregeln samt kedjeregeln ger då

$$\frac{d}{d\tau}\left(g_{ab}X^aX^b\right) = \frac{dx^c}{d\tau}\nabla_c\left(g_{ab}X^aX^b\right) = X^c g_{ab}\left(X^b\nabla_c X^a + X^a\nabla_c X^b\right) = 2\lambda(\tau)g_{ab}X^aX^b,$$

där vi också har använt att förbindelsen är metrisk. Denna första ordningens linjära differentialekvation kan lösas med integrerande faktor, vilket ger $g_{ab}X^aX^b = Ce^{2\Lambda(\tau)}$ där $\Lambda(\tau)$ är någon primitiv till $\lambda(\tau)$ och C är en godtycklig konstant.

Om nu X^a är tidsartad, d v s $g_{ab}X^aX^b > 0$, i någon punkt på Γ följer att $C > 0$ så $g_{ab}X^aX^b > 0$ på hela Γ. Även om Γ inte är affint parametriserad gäller alltså att X^a är tidsartad på hela Γ om X^a är tidsartad i någon punkt på Γ.

Fallen att X^a är rumsartad eller en nollvektor i någon punkt på Γ kan förstås behandlas på liknande sätt, vilket visar

SATS 8.7. *Om $x^a = x^a(\tau)$ är en parametrisering av en geodet Γ på en godtycklig mångfald M är $X^a = \frac{dx^a}{d\tau}$ antingen tidsartad på hela Γ, rumsartad på hela Γ eller en nollvektor på hela Γ.*

Beroende på vilket av dessa fall som gäller sägs Γ vara en *tidsartad geodet*, *rumsartad geodet* eller en *nollgeodet*.
Anmärkning: Om Γ är en rumsartad eller tidsartad geodet är alltså $X^a = \frac{dx^a}{d\tau} \neq 0$ i alla punkter på Γ. Vi kommer i fortsättningen att kräva att detta villkor är uppfyllt även om Γ är en nollgeodet. På detta sätt undviker vi en del onödiga och ointressanta komplikationer. $\Box$

Om vi vill bestämma geodeterna till en given mångfald är det praktiskt att skriva om definitionen lite grand. Betrakta därför

$$X^b \nabla_b X^a = \frac{dx^b}{d\tau}\left(\frac{\partial}{\partial x^b}\left(\frac{dx^a}{d\tau}\right) + \Gamma^a{}_{bc}\frac{dx^c}{d\tau}\right) = \frac{d}{d\tau}\left(\frac{dx^a}{d\tau}\right) + \Gamma^a{}_{bc}\frac{dx^b}{d\tau}\frac{dx^c}{d\tau} = \frac{d^2x^a}{d\tau^2} + \Gamma^a{}_{bc}\frac{dx^b}{d\tau}\frac{dx^c}{d\tau}.$$

Detta visar

SATS 8.8. *$x^a = x^a(\tau)$ är en parametrisering av en geodet om $x^a(\tau)$ uppfyller*

$$\frac{d^2x^a}{d\tau^2} + \Gamma^a{}_{bc}\frac{dx^b}{d\tau}\frac{dx^c}{d\tau} = \lambda(\tau)\frac{dx^a}{d\tau}, \tag{8.1}$$

för någon funktion $\lambda(\tau)$. Vidare är τ en affin parameter om och endast om

$$\frac{d^2x^a}{d\tau^2} + \Gamma^a{}_{bc}\frac{dx^b}{d\tau}\frac{dx^c}{d\tau} = 0. \tag{8.2}$$

Ekvationerna (8.1) och (8.2) brukar kallas den *geodetiska ekvationen*[1]. Observera också att i $\mathbf{R}^n$ med ett kartesiskt kordinatsystem x^a är $\Gamma^a{}_{bc} = 0$ så de geodetiska ekvationerna övergår till $\frac{d^2x^a}{d\tau^2} = \lambda(\tau)\frac{dx^a}{d\tau}$ respektive $\frac{d^2x^a}{d\tau^2} = 0$, d v s samma ekvationer som vi arbetade med i den inledande diskussionen ovan.

En viktig egenskap hos räta linjer i $\mathbf{R}^n$ är att givet en punkt $p \in \mathbf{R}^n$ och en vektor v_p^a i p så finns det en entydigt bestämd rät linje genom p som är parallell med v_p^a. Vi ska nu se att ett liknande resultat gäller för geodeter på an allmän mångfald.

Låt $p \in M$ vara godtycklig och låt x_p^a vara p:s koordinater. Låt vidare v_p^a vara en godtycklig vektor i p och betrakta begynnelsevärdesproblemet

$$\left\{\begin{array}{rcl} \frac{d^2x^a}{d\sigma^2} + \Gamma^a{}_{bc}\frac{dx^b}{d\sigma}\frac{dx^c}{d\sigma} & = & 0 \\ x^a(0) & = & x_p^a \\ \frac{dx^a}{d\sigma}(0) & = & v_p^a \end{array}\right. .$$

Enligt teori för differentialekvationer har detta problem en entydig lösning $x^a(\sigma)$ definierad i (åtminstone ett litet) öppet intervall kring 0.

[1] Ja, båda två. Inte heller här är terminologin konsekvent.

Låt oss se efter vad som händer om vi byter affin parameter genom att sätta $\sigma = a\tau + b$ där $a \neq 0$ och b är konstanter. För tydlighets skull inför vi en ny beteckning $\tilde{x}^a(\tau) = x^a(\sigma(\tau))$. Då blir $\tilde{x}^a(-b/a) = x^a\left(\sigma(-b/a)\right) = x^a(0) = x^a_p$ och

$$\frac{d\tilde{x}^a}{d\tau}(-b/a) = \frac{dx^a}{d\sigma}\left(\sigma(-b/a)\right)\frac{d\sigma}{d\tau} = \frac{dx^a}{d\sigma}(0) \cdot a = av^a_p.$$

Ett byte av affin parameter innebär alltså en translation av det parametervärde som svarar mot punkten p, en omskalning av geodetens tangentvektor i p samt eventuellt en omkastning av orienteringen hos geodeten. Detta visar

SATS 8.9. *Givet en punkt $p \in M$ och vektor v^a_p i p så finns en entydigt bestämd affint parametriserad geodet genom p vars tangentvektor i p är parallell med v^a_p.*

Även denna egenskap hos räta linjer i $\mathbf{R}^n$ delas alltså av geodeter, i alla fall lokalt. Geodeten i denna sats blir ju inte säkert definierad för *alla* värden på den affina parametern.

8.3 Enhetssfären, del 6

Efter att ha läst föregående avsnitt har du kanske fått intrycket att geodeter är mer eller mindre samma sak som räta linjer i $\mathbf{R}^n$. Så är dock inte fallet, vilket vi ska se genom att återigen besöka vår favoritmångfald, enhetssfären.

Exempel 8.10. *(Enhetssfären, del 6)* Här kommer nästa episod i följetongen om enhetssfären S. Vi söker geodeterna på S och vi kommer att använda samma notation som i de tidigare episoderna.

I Exempel 7.3 beräknade vi Christoffelsymbolerna på S och fick ${\Gamma^0}_{11} = -\sin\theta\cos\theta$, ${\Gamma^1}_{01} = {\Gamma^1}_{10} = \cot\theta$ och övriga ${\Gamma^a}_{bc} = 0$. Ekvationerna $\frac{d^2x^a}{d\tau^2} + {\Gamma^a}_{bc}\frac{dx^b}{d\tau}\frac{dx^c}{d\tau} = 0$ för en affint parametrierad geodet på S blir då

$$\begin{cases} \frac{d^2\theta}{d\tau^2} - \sin\theta\cos\theta\left(\frac{d\varphi}{d\tau}\right)^2 &= 0 \\ \frac{d^2\varphi}{d\tau^2} + 2\cot\theta\frac{d\theta}{d\tau}\frac{d\varphi}{d\tau} &= 0 \end{cases}. \tag{8.3}$$

För att hitta alla geodeter behöver vi hitta allmänna lösningen till detta system. Detta visar sig dock vara svårt att göra genom direkt manipulation av systemets ekvationer. Vi försöker därför gå en annan väg.

Betrakta ekvatorn E på S. På E är $\theta = \frac{\pi}{2}$ d v s första ekvationen är identiskt uppfylld och den andra övergår till $\frac{d^2\varphi}{d\tau^2} = 0 \Leftrightarrow \varphi = a\tau + b$ där villkoret $\frac{d\varphi}{d\tau} \neq 0$ ger att $a \neq 0$. E är alltså en geodet på S och de olika möjliga värdena på a och b svarar mot friheten att välja olika affina parametriseringar av E.

Vi har nu hittat *en* geodet på S. Kan vi hitta flera? Observera att ekvatorn är ett exempel på en *storcirkel* på S d v s en cirkel på S med samma medelpunkt som S. Av symmetriskäl förväntar vi oss därför att *alla* storcirklar på S ska vara geodeter. Låt oss försöka bevisa detta.

Låt $\bar{E}$ vara en godtycklig storcirkel på S. Det visar sig bli besvärligt att försöka parametrisera $\bar{E}$ i det givna koordinatsystemet och sätta in parametriseringen i systemet ovan. Vi konstruerar därför ett nytt koordinatsystem. Notera att geodetiska ekvationen är en tensorekvation, så om vi hittar en lösning i ett koordinatsystem så blir denna en lösning i alla koordinatsystem.

Konstruera först ett kartesiskt koordinatsystem $(\bar{x}, \bar{y}, \bar{z})$ i $\mathbf{R}^3$ genom att låta origo i det nya koordinatsystemet sammanfalla med origo i det gamla, genom att låta $\bar{x}$- och $\bar{y}$-axeln ligga i samma plan som $\bar{E}$ och genom att låta $\bar{z}$-axeln vara vinkelrät mot $\bar{x}\bar{y}$-planet, d v s det plan som innehåller $\bar{E}$.

Definiera sedan sfäriska koordinater $(\bar{x}^0, \bar{x}^1) = (\bar{\theta}, \bar{\varphi})$ relativt detta nya koordinatsystem d v s definiera $\bar{\theta}$ och $\bar{\varphi}$ genom

$$\begin{cases} \bar{x} &= \sin\bar{\theta}\cos\bar{\varphi} \\ \bar{y} &= \sin\bar{\theta}\sin\bar{\varphi} \\ \bar{z} &= \cos\bar{\theta} \end{cases} .$$

Eftersom detta är exakt samma uttryck som i de 'vanliga' sfäriska koordinaterna med den lilla skillnaden att koordinaterna är streckade istället för ostreckade ges förstås metrikens komponenter $\bar{g}_{ab}$ samt Christoffelsymbolerna $\bar{\Gamma}^a{}_{bc}$ av samma uttryck som vi hittade i Exempel 5.3 och 7.3, fast med streckade koordinater istället för ostreckade.

Detta innebär att geodetiska ekvationen i de nya koordinaterna blir identisk med systemet (8.3) sånär som på att ostreckade koordinater ska bytas mot streckade. På samma sätt som ovan följer därför att $\bar{E}$, d v s ekvatorn på S i det streckade koordinatsystemet kännetecknad av villkoret $\bar{\theta} = \frac{\pi}{2}$, är en geodet på S och då $\bar{E}$ var en godtycklig storcirkel på S följer att alla storcirklar på S är geodeter på S.

Därmed har vi hittat många geodeter på S. Kan det finnas ännu fler? Svaret på denna fråga är nej, vilket kan inses på följande sätt: Låt $p \in S$ och $v_p^a \in T_pS$ vara godtyckliga. Då finns en storcirkel $\bar{E}$ genom p sådan att v_p^a tangerar $\bar{E}$. Men enligt Sats 8.9 är detta den enda geodeten med denna egenskap. Det kan alltså inte finnas några andra geodeter på S än storcirklarna vi redan hittat. □

8.4 Variationskalkyl, geodeter och enhetssfären, del 7.

Ett viktigt verktyg inom både fysik och matematik är den s k *variationskalkylen* och i detta kapitel ska vi ge en introduktion till denna. Eftersom vi i denna skrift endast kommer att använda variationskalkyl som ett verktyg för att komma åt den geodetiska ekvationen, och därmed också Christoffelsymbolerna, blir denna introduktion ganska kortfattad. För en mer utförlig behandling, se t ex Yakimenko: Lecture Notes in Analytical Mechanics, Craggs: Calculus of Variations eller Oden-Reddy: Variational Methods in Theoretical Mechanics.

Vi börjar med att illustrera variationskalkylens idé med ett exempel:

Exempel 8.11. Vad är den kortaste kurvan mellan två punkter i planet?

Alla vet säkert svaret på denna fråga, men låt oss låtsas att vi inte vet och se om vi kan hitta den kortaste kurvan på något sätt.

Då inget koordinatsystem är givet är vi alltså fria att själva konstruera ett. Placera origo i den ena av de givna punkterna och välj koordinataxlar så att den andra punkten hamnar i första kvadranten, d v s den andra punkten får koordinaterna (a, b) för $a, b > 0$.

Låt oss anta att det finns en kortaste kurva (ej självklart!) och att denna kurva är en funktionskurva av klass $\mathcal{C}^2$ (att bevisa att så är fallet ligger bortom horisonten för denna skrift). I så fall kan vi formulera problemet så här:

Minimera funktionalen[2] $I(y) = \int_0^a \sqrt{1 + y'(x)^2}\, dx$ där $y = y(x) \in \mathcal{C}^2$ och $y(0) = 0$, $y(a) = b$.

Detta är ett optimeringsproblem av ett lite annat slag än de som brukar behandlas i kurser i analys. I dessa kurser beror den funktion som ska minimeras eller maximeras oftast av ändligt många variabler. Här är variabeln som får variera en hel kurva. Vi kan faktiskt se det som att vart och ett av funktionen y:s värden i intervallet $]a, b[$ är en variabel och med det synsättet ska vi alltså minimera en funktion av oändligt många (t o m överuppräkneligt många) variabler.

Beskrivet på detta sätt låter problemet nästan oöverstigligt svårt, men räddning finns inom räckhåll. Antag att y är en funktion som löser ovanstående problem och låt η vara en godtycklig $\mathcal{C}^2$-funktion med $\eta(0) = \eta(a) = 0$. Låt vidare ε vara ett reellt tal. Då är $\bar{y} = y + \varepsilon\eta$ också en $\mathcal{C}^2$-funktion vars graf förbinder $(0, 0)$ och (a, b) för alla ε så enligt vårt antagande gäller $I(y) \leq I(\bar{y}) = I(y + \varepsilon\eta)$ för alla värden på ε. $I(y + \varepsilon\eta)$ har alltså ett lokalt minimum för $\varepsilon = 0$, varför derivatan av $I(y + \varepsilon\eta)$ med avseende på ε blir 0 då $\varepsilon = 0$.

Då derivering innanför integraltecknet är tillåten i detta fall (se Persson-Böiers: Analys i flera variabler) ges denna derivata av

$$0 = \frac{d}{d\varepsilon} \int_0^a \sqrt{1 + (y' + \varepsilon\eta')^2}\, dx \Bigg|_{\varepsilon=0} = \int_0^a \frac{(y' + \varepsilon\eta')\eta'}{\sqrt{1 + (y' + \varepsilon\eta')^2}}\, dx \Bigg|_{\varepsilon=0} = \int_0^a \frac{y'\eta'}{\sqrt{1 + (y')^2}}\, dx$$
$$= \left[\frac{y'\eta}{\sqrt{1 + (y')^2}} \right]_0^a - \int_0^a \frac{d}{dx} \left(\frac{y'}{\sqrt{1 + (y')^2}} \right) \eta\, dx,$$

där vi partialintegrerat i sista steget. Nu är den utintegrerade biten 0 eftersom $\eta(0) = \eta(a) = 0$ enligt förutsättning, så det följer att andra termen är noll för alla η. Då $y \in \mathcal{C}^2$ följer det att (Hur? Nyttig övning att själv genomföra resonemanget.)

$$\frac{d}{dx} \left(\frac{y'}{\sqrt{1 + (y')^2}} \right) = 0 \Leftrightarrow \frac{y'}{\sqrt{1 + (y')^2}} = C_0 \Leftrightarrow y' = C,$$

där C_0 och C är godtyckliga konstanter. Detta, samt villkoren $y(0) = 0$, $y(a) = b$ ger nu att $y = \frac{b}{a}x$. Den kortaste kurvan som förbinder punkterna $(0, 0)$ och (a, b) är alltså sträckan mellan dessa punkter, som förväntat.

□

[2] Med en *funktional* menas en avbildning från något vektorrum (oftast ett funktionsrum av något slag) till $\mathbf{R}$ (eller $\mathbf{C}$). Funktionaler studeras närmare i ämnet Funktionalanalys.

Vi ska nu generalisera idéerna från detta exempel. Låt därför y vara en funktion av en parameter som vi kallar u och antag att y:s värden i två punkter u_1 och u_2 är fixerade. Vi kommer att använda beteckningen $\dot{y} = \frac{dy}{du}$.

Betrakta funktionalen $I(y) = \int_{u_1}^{u_2} L(y, \dot{y}, u)\, du$, där $L(y, \dot{y}, u)$ kallas en *lagrangefunktion*, och antag att vi söker y som minimerar (eller maximerar) denna funktional.

Observera att det absolut **inte** är säkert att ett sådant y existerar, detta är något som måste undersökas separat. **Om** ett minimum (eller maximum) existerar kan vi dock använda samma idé som i förra exemplet för att hitta det, och det är denna idé som kallas *variationskalkyl*.

Låt (som ovan) η vara en godtycklig $\mathcal{C}^2$-funktion med $\eta(u_1) = \eta(u_2) = 0$ och låt ε vara ett reellt tal. Om y minimerar (eller maximerar) $I(y)$ är $I(y) \leq I(y + \varepsilon\eta)$ för alla ε. Om lagrangefunktionen L är så snäll att derivering innanför integraltecknet är tillåten ger detta att

$$0 = \frac{d}{d\varepsilon} \int_{u_1}^{u_2} L(y + \varepsilon\eta, \dot{y} + \varepsilon\dot{\eta}, u)\, du \bigg|_{\varepsilon=0} = \int_{u_1}^{u_2} \left(\frac{\partial L}{\partial y}(y + \varepsilon\eta, \dot{y} + \varepsilon\dot{\eta}, u)\eta + \frac{\partial L}{\partial \dot{y}}(y + \varepsilon\eta, \dot{y} + \varepsilon\dot{\eta}, u)\dot{\eta} \right) du \bigg|_{\varepsilon=0}$$
$$= \int_{u_1}^{u_2} \left(\frac{\partial L}{\partial y}(y, \dot{y}, u)\eta + \frac{\partial L}{\partial \dot{y}}(y, \dot{y}, u)\dot{\eta} \right) du = \left[\frac{\partial L}{\partial \dot{y}}\eta \right]_{u_1}^{u_2} + \int_{u_1}^{u_2} \left(\frac{\partial L}{\partial y} - \frac{d}{du}\left(\frac{\partial L}{\partial \dot{y}} \right) \right) \eta\, du,$$

där vi partialintegrerat i sista steget. Då $\eta(u_1) = \eta(u_2) = 0$ försvinner den utintegrerade termen och då $\eta \in \mathcal{C}^2$ var godtycklig följer, på samma sätt som ovan att

$$\frac{d}{du}\left(\frac{\partial L}{\partial \dot{y}} \right) - \frac{\partial L}{\partial y} = 0. \tag{8.4}$$

Denna ekvation kallas *Euler-Lagranges ekvation*.

Om istället lagrangefunktionen L beror av flera variabler y^k, $k = 0, \ldots, n-1$ och deras derivator $\dot{y}^k$ fås på samma sätt *Euler-Lagranges ekvationer*

$$\frac{d}{du}\left(\frac{\partial L}{\partial \dot{y}^k} \right) - \frac{\partial L}{\partial y^k} = 0\,,\ k = 0, \ldots, n-1. \tag{8.5}$$

Euler-Lagranges ekvationer utgör ett effektivt sätt att plocka fram geodetiska ekvationen för en given metrik. För att se detta, låt x^a vara ett godtyckligt koordinatsystem.

Observera att om g_{ab} vore en Riemannsk (d v s positivt definit) metrik så kan vi definiera en lagrangefunktion genom $L = \left(g_{ab}\dot{x}^a\dot{x}^b\right)^{1/2}$ och motsvarande funktional $I = \int_{u_1}^{u_2} L(x^a, \dot{x}^a, u)\, du$ blir då längden av kurvan $x^a = x^a(u)$, $u_1 \leq u \leq u_2$. Att minimera denna funktional är alltså ekvivalent med att hitta den kortaste kurvan mellan två punkter på en mångfald med metrik g_{ab}.

För en Lorentzmetrik är situationen inte lika enkel. Till att börja med måste vi skilja på om vi betraktar tidsartade kurvor, rumsartade kurvor eller nollkurvor eftersom $g_{ab}\dot{x}^a\dot{x}^b$ då är > 0, < 0 resp. $= 0$.

Fallet nollkurvor är speciellt intressant. För en nollkurva är ju som sagt $g_{ab}\dot{x}^a\dot{x}^b = 0$ vilket ger $I = 0$. I:s värde är alltså oberoende av om Euler-Lagranges ekvationer är uppfyllda eller inte. Det är alltså inte meningsfullt att tala om att vi minimerar (eller maximerar) I i detta fall.

Det är därför bättre att tänka på kurvor som uppfyller Euler-Lagranges ekvationer som kurvor som gör I *stationär*. Att I är stationär för en viss kurva betyder att om denna kurva störs lite grand och ε är ett mått på storleken hos störningen, så blir förändringen av I:s värde inte $O(\varepsilon)$ utan $O(\varepsilon^2)$.

Exempel 8.12. Vi ska i detta exempel studera tidsartade kurvor, d v s kurvor som överallt uppfyller villkoret $g_{ab}\dot{x}^a\dot{x}^b > 0$. Övriga fall kan behandlas på liknande sätt. Definiera som ovan lagrangefunktionen $L = \left(g_{ab}\dot{x}^a\dot{x}^b\right)^{1/2}$ och motsvarande funktional $I = \int_{u_1}^{u_2} L(x^a, \dot{x}^a, u)\, du$. Vi söker nu de kurvor som gör denna funktional stationär, d v s de kurvor som uppfyller Euler-Lagranges ekvationer.

Vi börjar med att fixera, den hittills godtyckliga, parametern u. Antag alltså att $x^a = x^a(u)$ är en parametrisering av en kurva som löser Euler-Lagranges ekvationer. Med denna parametrisering insatt är $L = \left(g_{ab}\dot{x}^a\dot{x}^b\right)^{1/2} =: h(u)$ en funktion av enbart u. Definiera en ny parameter τ som en lösning till differentialekvationen $\frac{d\tau}{du} = \sqrt{h(u)}$. Då är

$$g_{ab}\frac{dx^a}{d\tau}\frac{dx^b}{d\tau} = g_{ab}\dot{x}^a\dot{x}^b\left(\frac{du}{d\tau}\right)^2 = h(u)\cdot\left(\frac{1}{\sqrt{h(u)}}\right)^2 = 1,$$

d v s om vi byter beteckning genom att låta $\dot{x}^a$ stå för $\frac{dx^a}{d\tau}$ istället för $\frac{dx^a}{du}$ så är $g_{ab}\dot{x}^a\dot{x}^b = 1$. Parametern τ brukar kallas *egentiden* längs den tidsartade kurvan $x^a = x^a(\tau)$.

Efter detta parameterbyte är tydligen $g_{ab}\dot{x}^a\dot{x}^b$, och därmed även L, oberoende av τ om $x^a = x^a(\tau)$ löser Euler-Lagranges ekvationer, d v s om $\frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{x}^a}\right) - \frac{\partial L}{\partial x^a} = 0$. Detta ger

$$0 = 2L\frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{x}^a}\right) - 2L\frac{\partial L}{\partial x^a} = \frac{d}{d\tau}\left(2L\frac{\partial L}{\partial \dot{x}^a}\right) - 2L\frac{\partial L}{\partial x^a} = \frac{d}{d\tau}\left(\frac{\partial L^2}{\partial \dot{x}^a}\right) - \frac{\partial L^2}{\partial x^a}.$$

Kurvan $x^a = x^a(\tau)$ löser alltså Euler-Lagranges ekvationer även för lagrangefunktionen $L^2 = g_{bc}\dot{x}^b\dot{x}^c$, vilket underlättar de kommande beräkningarna en del.

Observera att $L^2 = g_{bc}\dot{x}^b\dot{x}^c$ innehåller termer av två olika slag, nämligen termer där $b = c$ och termer där $b \neq c$ och dessa sistnämnda termer förekommer i dubbel uppsättning eftersom g_{bc} är symmetrisk. Detta ger att $\frac{\partial L^2}{\partial \dot{x}^a} = 2g_{ab}\dot{x}^b$ och vi får

$$\frac{d}{d\tau}\left(\frac{\partial L^2}{\partial \dot{x}^a}\right) = \frac{d}{d\tau}\left(2g_{ab}\dot{x}^b\right) = 2g_{ab}\ddot{x}^b + 2\frac{\partial g_{ab}}{\partial x^c}\dot{x}^b\dot{x}^c = 2g_{ab}\ddot{x}^b + \left(\frac{\partial g_{ab}}{\partial x^c} + \frac{\partial g_{ac}}{\partial x^b}\right)\dot{x}^b\dot{x}^c,$$

där sista likheten följer av att faktorn $\dot{x}^b\dot{x}^c$ är symmetrisk i b och c.

Euler-Lagranges ekvationer kan alltså skrivas

$$0 = \frac{d}{d\tau}\left(\frac{\partial L^2}{\partial \dot{x}^a}\right) - \frac{\partial L^2}{\partial x^a} = 2g_{ab}\ddot{x}^b + \left(\frac{\partial g_{ab}}{\partial x^c} + \frac{\partial g_{ac}}{\partial x^b} - \frac{\partial g_{bc}}{\partial x^a}\right)\dot{x}^b\dot{x}^c,$$

och efter kontraktion med $\frac{1}{2}g^{ad}$ känner vi igen formeln för Christoffelsymbolerna och får

$$0 = \frac{1}{2}g^{ad}\cdot 2g_{ab}\ddot{x}^b + \frac{1}{2}g^{ad}\left(\frac{\partial g_{ab}}{\partial x^c} + \frac{\partial g_{ac}}{\partial x^b} - \frac{\partial g_{bc}}{\partial x^a}\right)\dot{x}^b\dot{x}^c = \ddot{x}^d + \Gamma^d{}_{bc}\dot{x}^b\dot{x}^c.$$

Sammanfattningsvis: De tidsartade kurvor $x^a = x^a(\tau)$ som löser Euler-Lagranges ekvationer för lagrangefunktionen $L = \left(g_{ab}\dot{x}^a\dot{x}^b\right)^{1/2}$ är de kurvor som uppfyller de båda villkoren

$$\ddot{x}^a + \Gamma^a{}_{bc}\dot{x}^b\dot{x}^c = 0 \,, \quad g_{ab}\dot{x}^a\dot{x}^b = 1,$$

d v s enligt Sats 8.8 är dessa kurvor affint parametriserade, tidsartade geodeter.

□

Fallen med rumsartade kurvor resp. nollkurvor kan behandlas på liknande sätt (genomför själv detaljerna!) och enligt Sats 8.7 är dessa tre fall de enda som kan förekomma. Vi sammanfattar resultatet i en sats.

SATS 8.13. *Låt $L = \left|g_{ab}\dot{x}^a\dot{x}^b\right|^{1/2}$ och låt γ vara en kurva som gör motsvarande funktional I stationär. Då finns en parametrisering av γ sådan att*

$$\ddot{x}^a + \Gamma^a{}_{bc}\dot{x}^b\dot{x}^c = 0,$$

och $g_{ab}\dot{x}^a\dot{x}^b = 1$ om γ är tidsartad, $g_{ab}\dot{x}^a\dot{x}^b = -1$ om γ är rumsartad och $g_{ab}\dot{x}^a\dot{x}^b = 0$ om γ är en nollkurva. I samtliga fall är γ en affint parametriserad geodet.

Euler-Lagranges ekvationer är ofta det snabbaste och enklaste sättet att beräkna Christoffelsymbolerna eftersom dessa kan läsas av direkt från geodetiska ekvationen som enligt ovanstående sats är ekvivalent med Euler-Lagranges ekvationer.

Exempel 8.14. *(Enhetssfären, del 7)* Med beteckningar från tidigare episoder i denna följetong är

$$L^2 = g_{ab}\dot{x}^a\dot{x}^b = \dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2.$$

Då enhetssfären är tvådimensionell finns det två Euler-Lagrange ekvationer. Den första ges av

$$0 = \frac{d}{d\tau}\left(\frac{\partial L}{\partial\dot{\theta}}\right) - \frac{\partial L}{\partial\theta} = 2\frac{d\dot{\theta}}{d\tau} - 2\sin\theta\cos\theta\dot{\varphi}^2 = 2\ddot{\theta} - 2\sin\theta\cos\theta\dot{\varphi}^2,$$

och den andra ges av

$$0 = \frac{d}{d\tau}\left(\frac{\partial L}{\partial\dot{\varphi}}\right) - \frac{\partial L}{\partial\varphi} = \frac{d}{d\tau}\left(2\sin^2\theta\dot{\varphi}\right) - 0 = 2\sin^2\theta\frac{d\dot{\varphi}}{d\tau} + 4\sin\theta\cos\theta\cdot\dot{\theta}\dot{\varphi},$$

vilket ger de geodetiska ekvationerna

$$\left\{\begin{array}{rcl} \ddot{\theta} - \sin\theta\cos\theta\,\ddot{\varphi}^2 & = & 0 \\ \ddot{\varphi} + 2\cot\theta\,\dot{\theta}\dot{\varphi} & = & 0 \end{array}\right. , \tag{8.6}$$

och vi ser att detta är samma ekvationer som (8.3), precis som förväntat.

Vi får nu Christoffelsymbolerna genom direkt avläsning från (8.2). Vi ser att $\Gamma^0{}_{11} = -\sin\theta\cos\theta$, $\Gamma^1{}_{01} = \Gamma^1{}_{10} = \cot\theta$ och att alla övriga $\Gamma^a{}_{bc} = 0$.

□

Kapitel 9

Krökning

9.1 Inledning

I Exempel 7.15 såg vi att parallelltransport på enhetssfären S var beroende av vägen. Detta är en av många effekter av att S inte är platt utan krökt. I detta kapitel ska vi se på en del andra konsekvenser av krökning och även ge detta begrepp en precis betydelse.

Vi kommer att se att krökningen hos en mångfald beskrivs av ett tensorfält $R^a{}_{bcd}$ kallat Riemanns krökningstensor. Från denna kan vi bilda Weyltensorn $C^a{}_{bcd}$, Riccitensorn R_{ab} och Einsteintensorn G_{ab} som alla spelar ledande roller i den allmänna relativitetsteorin.

9.2 Riemanns krökningstensor

Från flervariabelanalysen är vi vana vid att partiella andraderivator kommuterar, d v s $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ för alla $f \in \mathcal{C}^2$. Enligt Proposition 7.10 gäller dessutom att $\nabla_a \nabla_b f = \nabla_b \nabla_a f$ så länge förbindelsen är torsionsfri (d v s om $\Gamma^a{}_{bc} = \Gamma^a{}_{cb}$). Torsionsfria kovarianta andraderivator kommuterar tydligen också, i alla fall så länge det är $\mathcal{C}^\infty$ funktioner som deriveras två gånger.

Hur förhåller det sig då om vi istället deriverar ett vektorfält två gånger och jämför de blandade andraderivatorna? Detta ska vi nu ta reda på. Låt därför v^a vara ett godtyckligt vektorfält och betrakta

$$\begin{aligned}
\nabla_c \nabla_d v^a &= \frac{\partial}{\partial x^c} (\nabla_d v^a) + \Gamma^a{}_{bc} \nabla_d v^b - \Gamma^b{}_{dc} \nabla_b v^a \\
&= \frac{\partial}{\partial x^c} \left(\frac{\partial v^a}{\partial x^d} + \Gamma^a{}_{bd} v^b \right) + \Gamma^a{}_{bc} \left(\frac{\partial v^b}{\partial x^d} + \Gamma^b{}_{ed} v^e \right) - \Gamma^b{}_{dc} \nabla_b v^a \\
&= \frac{\partial^2 v^a}{\partial x^c \partial x^d} + \left(\frac{\partial}{\partial x^c} \Gamma^a{}_{bd} \right) v^b + \Gamma^a{}_{bd} \frac{\partial v^b}{\partial x^c} + \Gamma^a{}_{bc} \frac{\partial v^b}{\partial x^d} + \Gamma^a{}_{ec} \Gamma^e{}_{bd} v^b - \Gamma^b{}_{dc} \nabla_b v^a,
\end{aligned}$$

där vi döpt om summationsindexen i näst sista termen. Liknande räkningar ger (förstås!)

$$\nabla_d \nabla_c v^a = \frac{\partial^2 v^a}{\partial x^d \partial x^c} + \left(\frac{\partial}{\partial x^d} \Gamma^a{}_{bc} \right) v^b + \Gamma^a{}_{bc} \frac{\partial v^b}{\partial x^d} + \Gamma^a{}_{bd} \frac{\partial v^b}{\partial x^c} + \Gamma^a{}_{ed} \Gamma^e{}_{bc} v^b - \Gamma^b{}_{cd} \nabla_b v^a.$$

Vid subtraktion av dessa ekvationer försvinner många termer (p g a att partiella andraderivator kommuterar, att förbindelsen är symmetrisk och som en konsekvens av att summationsindex får döpas om) och vi får

$$(\nabla_c \nabla_d - \nabla_d \nabla_c) v^a = \left(\frac{\partial}{\partial x^c} \Gamma^a{}_{bd} - \frac{\partial}{\partial x^d} \Gamma^a{}_{bc} + \Gamma^a{}_{ec} \Gamma^e{}_{bd} - \Gamma^a{}_{ed} \Gamma^e{}_{bc} \right) v^b.$$

Med hjälp av transformationslagen 7.7 för Christoffelsymbolerna kan man visa (Gör detta själv som övning!) att uttrycket inom parentes transformeras som en tensor av typ $(1,3)$ och vi gör därför följande definition.

Definition 9.1. *Riemanns krökningstensor* $R^a{}_{bcd}$[1] definieras

$$R^a{}_{bcd} = \frac{\partial}{\partial x^c} \Gamma^a{}_{bd} - \frac{\partial}{\partial x^d} \Gamma^a{}_{bc} + \Gamma^a{}_{ec} \Gamma^e{}_{bd} - \Gamma^a{}_{ed} \Gamma^e{}_{bc}. \tag{9.1}$$

($\star$ Att Riemanntensorn verkligen *är* en tensor kan bevisas på ett enklare sätt än att härleda en transformationslag för den. Enligt Sats 6.24 räcker det nämligen att visa att avbildningen

$$v^a \mapsto (\nabla_c \nabla_d - \nabla_d \nabla_c) v^a,$$

är $\mathcal{C}^\infty$-linjär. Låt därför $f \in \mathcal{C}^\infty$ vara godtycklig och betrakta

$$\begin{aligned}(\nabla_c \nabla_d - \nabla_d \nabla_c)(f v^a) &= \nabla_c \left(v^a \nabla_d f + f \nabla_d v^a \right) - \nabla_d \left(v^a \nabla_c f + f \nabla_c v^a \right) \\ &= v^a (\nabla_c \nabla_d - \nabla_d \nabla_c) f + f (\nabla_c \nabla_d - \nabla_d \nabla_c) v^a = f (\nabla_c \nabla_d - \nabla_d \nabla_c) v^b,\end{aligned}$$

ty första termen i näst sista ledet är 0 då förbindelsen är symmetrisk. Det följer att avbildningen ovan är $\mathcal{C}^\infty$-linjär vilket innebär att $R^a{}_{bcd}$ är ett tensorfält. $\star$)

För ett godtyckligt vektorfält v^a gäller alltså att

$$(\nabla_c \nabla_d - \nabla_d \nabla_c) v^a = R^a{}_{bcd} v^b.$$

Om istället v_a är ett godtyckligt kovariant tensorfält ger liknande räkningar som ovan att

$$(\nabla_c \nabla_d - \nabla_d \nabla_c) v_a = -R^b{}_{acd} v_b.$$

För tensorer av ko- eller kontravariant rang > 1 samt för blandade tensorer är formlerna mer komplicerade. Vi ger inte någon allmän formel här utan vi nöjer oss med att titta på några specialfall. Om T^{ab} är en tensor (egentligen ett tensorfält, men detta språkbruk är vanligt i litteraturen så jag vill vänja läsaren vid det) av typ $(2,0)$ så är

$$(\nabla_c \nabla_d - \nabla_d \nabla_c) T^{ab} = R^a{}_{ecd} T^{eb} + R^b{}_{ecd} T^{ae},$$

[1] Observera att indexens placering varierar i litteraturen. Även här har det alltså visat sig svårt att komma överens om en global konvention.

och om istället $T^a{}_b$ är en tensor av typ $(1, 1)$ så är

$$(\nabla_c \nabla_d - \nabla_d \nabla_c) T^a{}_b = R^a{}_{ecd} T^e{}_b - R^e{}_{bcd} T^a{}_e.$$

Dessa formler visas på liknande sätt som de tidigare och generalisering till tensorer med annan indexkonfiguration sker analogt.

Genom att titta på högerledet i Definition 9.1 ser vi att $R^a{}_{bcd} = -R^a{}_{bdc}$ d v s Riemanntensorn är antisymmetrisk över sitt sista indexpar. Riemanntensorn har också en del andra algebraiska symmetrier och vi sammanfattar dessa i följande sats:

SATS 9.2. *För Riemanns krökningstensor $R^a{}_{bcd}$ gäller att*

(i) $R^a{}_{bcd} = -R^a{}_{bdc}$

(ii) $R^a{}_{bcd} + R^a{}_{cdb} + R^a{}_{dbc} = 0$

(iii) $R^a{}_{[bcd]} = 0$

(iv) $R_{abcd} = R_{cdab}$*, där (som vanligt)* $R_{abcd} = g_{ae} R^e{}_{bcd}$

(v) $R_{abcd} = -R_{bacd}$.

Beviset av denna sats lämnas som övning för läsaren. *(ii)* och *(iv)* går t ex att visa direkt från Definition 9.1 eller med hjälp av s k *geodetiska koordinater* (se *d'Inverno-Vickers*: Introducing Einstein's Relativity — A deeper understanding) medan *(iii)* och *(v)* är konsekvenser av övriga symmetrier.

Betrakta nu Riemanntensorn R_{abcd}, där vi sänkt första indexet med metriken. Då denna har 4 index som alla löper från 0 till $n - 1$, där n är mångfalden M:s dimension, har Riemanntensorn potentiellt n^4 oberoende komponenter. Om vi betraktar specialfallet att M är en fyrdimensionell rumtid har R_{abcd} alltså potentiellt $4^4 = 256$ oberoende komponenter. Detta antal reduceras dock avsevärt av ovanstående symmetrier.

Betrakta först antisymmetrin $R_{abcd} = -R_{bacd}$ över första indexparet. Den ger omedelbart att alla komponenter i R_{abcd} där $a = b$ är lika med 0 och att hälften av komponenterna där $a \neq b$ är lika med den andra hälften, fast med ombytt tecken. Det finns därmed bara 6 st oberoende möjligheter för indexparet ab, nämligen 01, 02, 03, 12, 13 och 23. På samma sätt ger antisymmetrin $R_{abcd} = -R_{abdc}$ över andra indexparet att indexparet cd bara har 6 oberoende möjligheter, nämligen samma 6 möjligheter som ab, vilket ger 36 hittills oberoende möjligheter för hela indexuppsättningen $abcd$.

Symmetrin $R_{abcd} = R_{cdab}$ över indexparen reducerar sedan antalet oberoende möjligheter till 21 st. Till sist visar det sig att den återstående symmetrin $R_{abcd} + R_{acdb} + R_{adbc} = 0$ bara ger ett ytterligare villkor på komponenterna R_{abcd}. Detta inses lättast genom att helt enkelt lista de olika möjligheterna. Detta reducerar antalet oberoende komponenter i R_{abcd} till 20 st.

Förutom de algebraiska symmetrierna i Sats 9.2 uppfyller Riemanntensorn också en differentiell symmetri kallad *Bianchi-identiteten*:

SATS 9.3. *Det gäller att* $\nabla_a R_{debc} + \nabla_b R_{deca} + \nabla_c R_{deab} = 0$.

Denna kan också bevisas genom byte till geodetiska koordinater (Övning!).

I kommande avsnitt kommer vi att se att om $R_{abcd} = 0$ så har M mycket gemensamt med $\mathbf{R}^n$, så mycket att man ofta säger att M lokalt 'är' $\mathbf{R}^n$. Vi ska så småningom undersöka detta i detalj. Här nöjer vi oss med att göra följande definition:

Definition 9.4. En mångfald M sägs vara *platt* om dess Riemanntensor $R_{abcd} = 0$.

9.3 Ricci, Einstein och Weyl

I detta avsnitt ska vi se på en användbar uppdelning av Riemanntensorn i enklare tensorer. Vi börjar med följande definition.

Definition 9.5. $R_{ab} = R^c{}_{acb}$ kallas *Riccitensorn*[2] och den skalära funktionen $R = R_a{}^a = g^{ab} R_{ab}$ kallas *Ricciskalären.*

Det är en omedelbar konsekvens av Riemanntensorns symmetrier att $R_{ab} = R_{ba}$ så Riccitensorn är symmetrisk.

Vi ska nu studera symmetriska tensorer lite närmare. Låt därför $T_{ab} = T_{ba}$ vara en godtycklig symmetrisk tensor på en mångfald med dimension n. Låt T beteckna spåret av T_{ab}, d v s sätt $T = T_a{}^a = g^{ab} T_{ab}$. Bilda en ny tensor $\tilde{T}_{ab} = T_{ab} + f \cdot g_{ab} T$ där f är en (tills vidare) godtycklig funktion. Låt oss beräkna spåret av $\tilde{T}_{ab}$:

$$\tilde{T}_a{}^a = T_a{}^a + f \cdot g_a{}^a T = T + f \cdot \delta_a{}^a T = (1 + fn)T,$$

och vi ser att $\tilde{T}_{ab}$ blir spårfri precis då $f = -\frac{1}{n}$. Vi definierar därför

Definition 9.6. Om $T_{ab} = T_{ba}$ är en godtycklig symmetrisk tensor kallas $\tilde{T}_{ab} = T_{ab} - \frac{1}{n} g_{ab} T$, där $T = T_a{}^a$, den *spårfria delen* av T_{ab}.

En godtycklig symmetrisk tensor $T_{ab} = \tilde{T}_{ab} + \frac{1}{n} g_{ab} T$ kan alltså alltid delas upp i en summa av en spårfri, symmetrisk tensor och en tensor som överallt är proportionell mot metriken g_{ab}. Dessutom blir uppdelningen entydig:

Proposition 9.7. *Om* T_{ab} *är symmetrisk och* $T_{ab} = U_{ab} + V g_{ab}$ *där* $U_{ab} = U_{ba}$, $U_a{}^a = 0$ *och* V *är en funktion så är* $U_{ab} = \tilde{T}_{ab} = T_{ab} - \frac{1}{n} g_{ab} T$ *och* $V = \frac{1}{n} T$ *där* $T = T_a{}^a$.

Beviset lämnas som övning till läsaren.

Kombinerar vi ovanstående med Definition 5.5 ser vi att en godtycklig, inte nödvändigtvis symmetrisk tensor T_{ab} kan delas upp enligt

$$T_{ab} = T_{[ab]} + T_{(ab)} = T_{[ab]} + \tilde{T}_{ab} + \frac{1}{n} g_{ab} T,$$

[2] Här följer den vanliga brasklappen. Indexens placering är inte standardiserad. Var alltså noga med att kolla upp vilken konvention som gäller i just den skrift du för tillfället läser.

där $T_{[ab]}$ är antisymmetrisk, $\tilde{T}_{ab} = T_{(ab)} - \frac{1}{n} g_{ab} T$ är symmetrisk och spårfri och $T = g^{ab} T_{(ab)} = g^{ab} T_{ab} = T_a{}^a$ är en funktion. Denna uppdelning är dessutom entydig enligt ovan samt Sats 5.6.

Från T_{ab} kan vi också bilda en tensor med samma spårfria del som T_{ab} men där spåret bytt tecken:

Definition 9.8. Om $T_{ab} = \tilde{T}_{ab} + \frac{1}{n} g_{ab} T$, där $T = T_a{}^a$, är en godtycklig symmetrisk tensor kallas $\check{T}_{ab} = \tilde{T}_{ab} - \frac{1}{n} g_{ab} T = T_{ab} - \frac{2}{n} g_{ab} T$ den till T_{ab} *spårreflekterade tensorn* eller *spårreflektionen* av T_{ab}.

Vi återgår nu till Riccitensorn:

Definition 9.9. $\tilde{R}_{ab} = R_{ab} - \frac{1}{n} g_{ab} R$ kallas den *spårfria Riccitensorn* och spårreflektionen av Riccitensorn

$$G_{ab} := \check{R}_{ab} = R_{ab} - \frac{2}{n} g_{ab} R,$$

kallas *Einsteintensorn.*

Observera att Einsteintensorn är symmetrisk, d v s $G_{ab} = G_{ba}$.

Vi kommer mest att intressera oss för fallet att $n = 4$, där alltså $\tilde{R}_{ab} = R_{ab} - \frac{1}{4} g_{ab} R$ och $G_{ab} = R_{ab} - \frac{1}{2} g_{ab} R$.

Man kan visa (övning) att Einsteintensorn uppfyller den *kontraherade Bianchi-identiteten*:

SATS 9.10. *Det gäller att* $\nabla^a G_{ab} = g^{ac} \nabla_c G_{ab} = 0$.