FFM234, Vektorfält och klassisk fysik - Föreläsningsanteckningar

Tünde Fülöp, Institutionen för fysik, Chalmers, Göteborg, Sverige

17 september 2021

Mål Målet med föreläsningen är att ge en introduktion till tensorer. (Kap 12 i kompendiet, kap 4 i Arfken.)

Vi har diskuterat skalärer och vektorer som varierar i tid och rum och hoppades att vi skulle klara oss att beskriva allt i naturen. Men så är tyvärr inte fallet. Många gånger behöver vi en generalisering av begreppen skalär och vektor. Användbarheten är störst vid matematisk behandling av flerdimensionella objekt, då utskrivning i komponentform annars skulle göra beräkningarna besvärliga.

När Einstein la fram sin speciella relativitetsteori klarade han sig med en matematisk beskrivning baserad på skalärer och vektorer. Men senare visade Minkowski att Newtons gravitationsteori inte är förenlig med den speciella relativitetsteorin. För system som påverkas av krafter behöver man en mer allmän relativitetsteori.

Einstein insåg i diskussioner med kollegor att han behövde överge det traditionella euklidiska rummet för en beskrivning som tillåter rummet att kröka sig i närheten av materia. För att kunna hantera gravitationen matematiskt med hjälp av kroklinjiga koordinater i detta krökta rum behövde han ett nyutveklat verktyg - differentialgeometrin. När han publicerade sin allmänna relativitetsteori blev han en av pionjärerna att använda dess formalism som baserar sig på tensorer.

Tensorer är alltså viktiga inom teoretisk fysik, men har också tillämpningar inom andra områden, till exempel plasmafysik, hållfasthetslära, och strömningsmekanik.

Som vi kommer att se tensorer är objekt vars komponenter transformeras enligt speciella regler. Tensoranalys möjliggör kompakt skrivna samband som används vid byte av koordinatsystem.

Vi kommer att betrakta transformationer i rummet mellan kartesiska system. Formalismen kan dock generaliseras till större symmetrier, t.ex. Lorentzinvarians i speciell relativitetsteori vilken involverar rum-tiden (fyra dimensioner).

Repetition

- Indexnotation: skriv vektorn $\mathbf{A}$ som A_i, vilket vi kan tolka som en lista över komponenterna, $\{A_i\}_{i=1}^3$.

- Indexregeln: en bokstavsindex som förekommer endast en gång i en term är giltigt för vart och ett av värdena 1 till n på detta index, där n är antal dimensioner. Ett sådant index kallas fritt index. Det kan bytas ut mot vilken som helst icke upptagen bokstav, dock måste detta byte ske samtidigt i alla termer i ett visst uttryck.

- Einsteins summationskonvention: summation är underförstådd så snart ett index förekommer två gånger. Ett sådant index kallas stumt index. Det kan inom en viss term bytas mot annan bokstav som inte är upptagen i termen.

- Skalärprodukt kan alltså skrivas

$$\mathbf{A} \cdot \mathbf{B} = A_i B_i$$

- En matris representeras av sina matriselement. Det första indexet är rad- och det andra kolumnindexet.

- Matris-vektor-multiplikationen

$$[\mathbf{Mv}]_i = M_{ij} v_j. \tag{1}$$

  Notera att det har ett fritt index (eftersom det är en vektor).

- $\delta_{ij} A_j = \sum_{j=1}^3 \delta_{ij} A_j = A_i$, dvs Kroneckers delta leder till att ett summationsindex “elimineras”; den plockar ut en term ur summan.

- Permutationssymbolen eller Levi-Civita tensorn:

$$\varepsilon_{ijk} = \begin{cases} 1, & ijk \text{ jämn permutation av 123,} \\ -1, & ijk \text{ udda permutation av 123,} \\ 0, & \text{annars, dvs. om minst två index tar samma värde.} \end{cases} \tag{2}$$

  Byter tecken om man byter plats på två index.

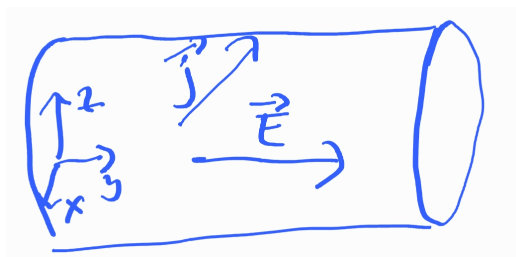
- Kryssprodukten skrivs alltså

$$[\mathbf{A} \times \mathbf{B}]_i = \varepsilon_{ijk} A_j B_k. \tag{3}$$

- Permutationssymbolen har egenskapen: $\varepsilon_{ilm}\varepsilon_{ijk} = \delta_{lj}\delta_{mk} - \delta_{lk}\delta_{mj}$ där alla δ_{xy} är Kroneckerdeltan.

Exempel Låt oss börja med ett exempel om elektrisk ledningsförmåga:

Ett elektriskt fält $\mathbf{E}$ är pålagt ett material.



Strömtätheten $\mathbf{j}$ kan fås av Ohms lag, som är på formen $\mathbf{j} = \sigma \mathbf{E}$. Om vi har ett isotropt material, så att konduktiviteten σ är skalär och oberoende av riktning kommer $\mathbf{j}$ ha samma riktning som elektriska fältet.

Men det behöver inte alls vara så i det allmänna fallet. Om materialet är anisotropt, som det ofta är i många fall, behöver strömmen i materialet inte var riktad parallellt med elektriska fältet. En matematisk modell som beskriver detta fall måste generaliseras till att σ blir en matris.

Om E-fältet som i figuren är riktat i y-riktningen har vi alltså $\mathbf{E} = (0, E_y, 0)$ och strömtäthetsvektorn blir $\mathbf{j} = (\sigma_{xy}E_y, \sigma_{yy}E_y, \sigma_{zy}E_y)$. (Obs! ej tensornotation för indexen.) Med denna representation tillåter vi möjligheten att ström kan gå även i x och z riktningarna.

Om vi har

$$\sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4}$$

avser vi ett material där konduktiviteten är hög i x-riktningen, men där ingen ström alls kan flyta i z-riktningen (detta kan hända t.ex. i ett kristallgitter där de vertikala lagren är elektriskt isolerade).

Vi kan skriva ekvationen på indexräkningsformen

$$j_i = \sigma_{ij} E_j \tag{5}$$

Detta är nu en tensorekvation.

Ohms lag på den här formen gäller för det specifika koordinatsystemet vi har valt. Hur förändras sambandet då vi övergår till ett annat koordinatsystem? Det är just koordinattransformationer som det blir viktigt att tala om matrisen som konduktivitetstensorn.

Antag att ekvationen gäller i det kartesiska koordinatsystemet K för strömtätheten i materialet i figuren. I ett annat kartesiskt koordinatsystem K' vars origo sammanfaller med det i K (men är roterat) skriver vi på motsvarande sätt

$$j'_i = \sigma'_{ij} E'_j. \tag{6}$$

Hur kommer konduktivitetsmatrisen se ut i det systemet? Vi behöver relatera de primmade komponenterna till de ursprungliga. Med hjälp av indexnotation kan en godtycklig vektor i K med basen $(\hat{e}_1, \hat{e}_2, \hat{e}_3)$ skrivas

$$\mathbf{v} = v_1 \hat{e}_1 + v_2 \hat{e}_2 + v_3 \hat{e}_3 = v_j \hat{e}_j. \tag{7}$$

I det primmade koordinatsystemet K' med basen $(\hat{e}'_1, \hat{e}'_2, \hat{e}'_3)$ kan samma vektor skrivas

$$\mathbf{v} = v'_1 \hat{e}'_1 + v'_2 \hat{e}'_2 + v'_3 \hat{e}'_3 = v'_j \hat{e}'_j. \tag{8}$$

Eftersom för ortonormala basvektorer gäller $\hat{e}'_i \cdot \hat{e}'_j = \delta_{ij}$ har vi från ovanstående ekvation (8)

$$\mathbf{v} \cdot \hat{e}'_i = v'_j \hat{e}'_j \cdot \hat{e}'_i = v'_j \delta_{ij} = v'_i. \tag{9}$$

Från ekvation (7) däremot får vi

$$\mathbf{v} \cdot \hat{e}'_i = v_j \hat{e}_j \cdot \hat{e}'_i \equiv v_j L_{ij} \tag{10}$$

där vi definierat **transformationsmatrisen** L_{ij} med komponenterna

$$L_{ij} = \hat{e}_j \cdot \hat{e}'_i = \cos\left(\hat{e}_j, \hat{e}_i{}'\right). \tag{11}$$

En godtycklig vektors komponenter i koordinatsystemet K' relaterade till koordinaterna i K blir då

$$v'_i = L_{ij} v_j. \tag{12}$$

Notera att vi genom att derivera ekv. (12) kan skriva matrisen $\mathbf{L}$ som $L_{ij} = \frac{\partial v_i'}{\partial v_j}$.

Från linjär algebra vet vi att matrisen L_{ij} är ortogonal, dvs den uppfyller att $\mathbf{L}\mathbf{L}^t = \mathbf{I}$.

Vi kan nu skriva upp en vektors komponenter i koordinatsystemet K uttryckt i koordinaterna i K'. Man behöver bara ta inversen av L_{ij}. Eftersom transformationsmatrisen är ortogonal, är inversen lika med transponatet $L_{ij}^{-1} = L_{ij}^t = L_{ji}$ (notera att j och i har bytt plats) så vi får

$$v_i = L_{ji} v_j'. \tag{13}$$

En användbar relation till senare är att skriva $\mathbf{L}\mathbf{L}^t = \mathbf{I}$ med hjälp av indexnotation:

$$\delta_{ij} = \left(L^t L\right)_{ij} = \left(L^t\right)_{ik} L_{kj} = L_{ki} L_{kj}. \tag{14}$$

Nu när vi vet hur man transformerar fram och tillbaka mellan koordinatsystemen, kan vi bestämma formen på konduktivitetsmatrisen σ i det nya systemet K'.

Eftersom $\mathbf{j}$ och $\mathbf{E}$ är vektorer kan vi nu skriva

$$j_i' = L_{ip} j_p \ \text{ från (12)} \tag{15}$$

$$E_q = L_{jq} E_j' \ \text{ från (13)} \tag{16}$$

och får om vi använder ekvationer (5) $j_p = \sigma_{pq} E_q$ och (6) $j_i' = \sigma_{ij}' E_j'$:

$$\sigma_{ij}' E_j' = j_i' = L_{ip} j_p = L_{ip} \sigma_{pq} E_q = L_{ip} \sigma_{pq} L_{jq} E_j' \tag{17}$$

vilket leder till

$$\left(\sigma_{ij}' - L_{ip} L_{jq} \sigma_{pq}\right) E_j' = 0 \tag{18}$$

Detta ska gälla för alla vektorer $\mathbf{E}$. Uttrycket i parentesen måste vara noll

$$\sigma_{ij}' = L_{ip} L_{jq} \sigma_{pq} \tag{19}$$

vilket visar hur konduktivitetsmatrisen förändras vid byte av koordinatsystem.

σ är exempel på en tensor, som är en storhet som transformeras enligt ovanstående vid byte av kartesiskt koordinatsystem. Men ekvationen gäller mycket generellt vid byte av koordinatsystem, man kan visa att tensorekvationer är invarianta vid byte av koordinatsystem, det vill säga samma ekvation gäller i alla koordinatsystem.

Definition: tensorer En tensor av ordningen p (eller rank p) $A_{i_1 i_2 \dots i_p}$ uppfyller transformationslagen

$$A'_{i_1 i_2 \dots i_p} = L_{i_1 j_1} L_{i_2 j_2} \dots L_{i_p j_p} A_{j_1 \dots j_p} \tag{20}$$

där L är transformationsmatrisen. Transformationsregeln är att när man byter system skall varje index multipliceras med det andra indexet på matrisen L.

En skalär är en tensor av nollte ordningen, då har vi en enda komponent $A' = A$ som har samma värde i alla koordinatsystem.

Kommentar

Det är viktigt att förstå att det är transformationsregeln som är det viktiga. Det räcker inte med att "s är ett tal".

En vektor är tensor av första ordningen (rank 1). En vektor $\mathbf{v}$ är en uppsättning objekt (funktioner) som beter sig likadant som ortvektorns komponenter när vi byter system, dvs.

$$v'_i = L_{ij} v_j \tag{21}$$

För att beskriva en tensor räcker det alltså att ange komponenterna i ett bestämt koordinatsystem. Då är komponenterna bestämda i alla andra koordinatsystem genom ekvation (20).

Konduktivitetsmatrisen som vi hade i exemplet är en andra ordningens tensor.

Sammanfattning: tensorbegreppet

Fysikaliska lagar skall inte bero på i vilket koordinatsystem de beskrivs. Detta kan vi åstadkomma genom att skriva dessa lagar som en likhet mellan två objekt vilka vi vet transformerar likadant under en koordinattransformation. Vi kallar sådana objekt för *tensorer*. En skalär är en tensor av rank-0, en vektor är en tensor av rank-1.

Exempel: Skalärprodukt Är $\mathbf{u} \cdot \mathbf{v}$ en skalär (dvs tensor rank-0)? Hur beter den sig vid ett koordinatbyte?

$$u'_i v'_i = L_{ij} u_j L_{ik} v_k = \delta_{jk} u_j v_k = u_k v_k \tag{22}$$

Resultatet är alltså oberoende av koordinatsystem, dvs det är en skalär.

Exempel: Produkt av tensorer Produkten av två tensorer är också en tensor.

- $c_{ij} = a_i b_j$ är också en tensor.
- $u_i = M_{ij} v_j$ är också en tensor.

För det andra fallet har vi

$$M'_{ij} v'_j = L_{ik} L_{jl} M_{kl} L_{jm} v_m = L_{ik} \delta_{lm} M_{kl} v_m = L_{ik} M_{kl} v_l = L_{ik} u_k$$

(där vi i första steget har använt att $\mathbf{L}$ är ortogonal). Om M_{ij} och v_i är tensorer är alltså även u_i det.

Det allmänna beviset av att multiplikation mellan tensorer av godtycklig rang är också tensor går likadant.

Exempel: Kryssprodukt Är $(\mathbf{a} \times \mathbf{b})_i = \varepsilon_{ijk} a_j b_k$ en tensor? Dvs är resultatet en vektor?

Räknereglerna ovan ger att vi bara måste visa att ε_{ikl} är en tensor. Standardmetoden för att visa om en storhet är en tensor eller inte är att undersöka om den transformeras som en tensor enligt definitionen. Transformationsreglerna säger:

$$\varepsilon'_{ijk} = L_{il} L_{jm} L_{kn} \varepsilon_{lmn} \tag{23}$$

Kom ihåg att ε_{ijk}

$$\varepsilon_{ijk} = \begin{cases} 1, & ijk \text{ jämn permutation av 123,} \\ -1, & ijk \text{ udda permutation av 123,} \\ 0, & \text{annars, dvs. om minst två index tar samma värde.} \end{cases} \tag{24}$$

Vi kan då skriva

$$\varepsilon'_{ijk} = L_{i1} L_{j2} L_{k3} + L_{i2} L_{j3} L_{k1} + L_{i3} L_{j1} L_{k2} - L_{i2} L_{j1} L_{k3} - L_{i1} L_{j3} L_{k2} - L_{i3} L_{j2} L_{k1} \tag{25}$$

eller på annat sätt

$$\varepsilon'_{ijl} = \begin{vmatrix} L_{i1} & L_{i2} & L_{i3} \\ L_{j1} & L_{j2} & L_{j3} \\ L_{k1} & L_{k2} & L_{k3} \end{vmatrix} \tag{26}$$

Om vi nu använder transformationsregeln $v'_i = L_{ij} v_j$ på enhetsvektorns komponenter, dvs uttrycker varje $\hat{e}'_i$ i den gamla basen $\hat{e}_i$ så får vi $\hat{e}'_i =$

$(\hat{e}_i{}' \cdot \hat{e}_m)\hat{e}_m = L_{im}\hat{e}_m$ och vi kommer ihåg att för allmänna vektorer v_i, w_i, u_i gäller att

$$(v_1, v_2, v_3) \cdot ((w_1, w_2, w_3) \times (u_1, u_2, u_3)) = \begin{vmatrix} v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \end{vmatrix}$$

så ser vi att vi kan skriva

$$\varepsilon'_{ijl} = \begin{vmatrix} L_{i1} & L_{i2} & L_{i3} \\ L_{j1} & L_{j2} & L_{j3} \\ L_{k1} & L_{k2} & L_{k3} \end{vmatrix} = \hat{e}'_i \cdot (\hat{e}'_j \times \hat{e}'_k) \tag{27}$$

eftersom $\hat{e}'_i = (L_{i1}, L_{i2}, L_{i3})$ och på samma sätt för de övriga basvektorerna.

Vidare har vi $\hat{e}'_i \cdot (\hat{e}'_j \times \hat{e}'_k) = 0$ om två index är lika, $= 1$ om i, j, l är jämna permutationer av $1, 2, 3$ och -1 om de är udda permutationer. Så vi har

$$\varepsilon'_{ijk} = \varepsilon_{ijk}$$

vilket betyder att Levi–Civita-tensorn är en *invariant tensor*. En annan invariant tensor är Kroneckers delta. Visa själv att $\delta'_{ij} = \delta_{ij}$.

Exempel: Tröghetstensorn Ett klassiskt exempel är tröghetstensorn,

$$I_{ij} = \int_V dV\, \rho(r^2\delta_{ij} - x_i x_j), \tag{28}$$

som relaterar rörelsemängdsmomentet till rotationsvektorn enligt $L_i = I_{ij}\omega_j$. I och med att den innehåller upprepade kryssprodukter är det enklare att härleda den i tensorformalism.

Ett litet volymelement dV av en stel kropp har massan $dm = \rho dV$, och om kroppen roterar med en rotationsvektor ω_i har det hastigheten

$$v_i = (\vec{\omega} \times \mathbf{r})_i = \varepsilon_{ijk}\omega_j x_k.$$

Dess rörelsemängd är

$$\mathrm{d}p_i = \mathrm{d}m v_i = \mathrm{d}m \varepsilon_{ijk}\omega_j x_k.$$

Bidraget till rörelsemängdsmomentet från volymelementet är

$$\mathrm{d}L_i = \varepsilon_{ijk} x_j \mathrm{d}p_k = \mathrm{d}m \varepsilon_{ijk} x_j \varepsilon_{klm}\omega_l x_m \tag{29}$$

Eftersom

$$\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$$

får vi

$$\begin{aligned} \mathrm{d}L_i = \mathrm{d}m(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})x_j x_m \omega_l &= \mathrm{d}m(r^2\omega_i - x_i x_j \omega_j) \\ &= \mathrm{d}V \rho(r^2\delta_{ij} - x_i x_j)\omega_j \end{aligned} \tag{30}$$

och totalt blir detta

$$L_i = \int_V \mathrm{d}V\, \rho(r^2\delta_{ij} - x_i x_j)\omega_j \equiv I_{ij}\omega_j, \tag{31}$$

där vi alltså definierar tröghetstensorn i det sista steget.

Derivator i indexnotation (kap 5.4) Vektoroperatorn ∇ har komponenterna $\frac{\partial}{\partial x_i}$, som vi ofta skriver ∂_i. Gradienten av ett skalärt fält skrivs $[\nabla\phi]_i = \partial_i\phi$. Divergensen av ett vektorfält är $\nabla \cdot \mathbf{F} = \partial_i F_i$ och rotationen $[\nabla \times \mathbf{F}]_i = \varepsilon_{ijk}\partial_j F_k$. Laplaceoperatorn verkande på ett skalärt fält är $\Delta\phi = \nabla \cdot (\nabla\phi) = \partial_i\partial_i\phi$.

Exempel: Magnetiskt tryck och spänning En ledare som leder en strömtäthet $\mathbf{j}(\mathbf{r})$ och då har ett motsvarande magnetfält $\mathbf{B}(\mathbf{r})$ påverkas av Lorentzkraften med krafttäthet $\mathbf{j}\times\mathbf{B}$. Vi kommer att skriva om den i termer av enbart magnetfältet $\mathbf{B}$ och hittar en fysikalisk tolkning för de olika termerna.

Vi antar att tidsberoendet är svagt, så vi kan använda Ampères lag utan förskjutningsström, $\nabla \times \mathbf{B} = \mu_0\mathbf{j}$, alltså är krafttätheten

$$\frac{1}{\mu_0}(\nabla \times \mathbf{B}) \times \mathbf{B}.$$

Rotationen $[\nabla \times \mathbf{B}]_i = \varepsilon_{ijk}\partial_j B_k$. Bortsett från den ointressanta för-faktorn μ_0^{-1}, kan vi skriva komponenterna av krafttätheten

$$(\nabla \times \mathbf{B}) \times \mathbf{B}$$

med indexnotation (med m som fritt index),

$$\epsilon_{mil}\epsilon_{ijk}\partial_j(B_k)B_l = \epsilon_{ilm}\epsilon_{ijk}B_l\partial_j B_k.$$

Detta kan förenklas till

$$(\delta_{lj}\delta_{mk} - \delta_{lk}\delta_{mj})B_l\partial_j B_k = B_l\partial_l(B_m) - B_l\partial_m B_l.$$

Första termen i vektornotation är $(\mathbf{B} \cdot \nabla)\mathbf{B}$, medan den andra termen kan förenklas om man inser att $B_l \partial_m B_l = \frac{1}{2}\partial_m(B_l B_l)$, som är $\frac{1}{2}\nabla(B^2)$. Alltså

$$\mathbf{j} \times \mathbf{B} = -\frac{1}{2\mu_0}\nabla B^2 + \frac{1}{\mu_0}(\mathbf{B} \cdot \nabla)\mathbf{B}.$$

Den första termen kan identifieras som den negativa gradienten av energitätheten i det magnetiska fältet $(B^2/(2\mu_0))$. På samma sätt som tryckgradienten $-\nabla p$ i en gas utför en kraft (gasen vill expandera till regioner med lägre tryck), och den termiska energitätheten är p, så har magnetfältet också ett tryck som är proportionell mot B^2, och vill expandera till regioner med lägre magnetfältsstyrka. Den andra termen kan identifieras som en sorts spänning eller ”styvhet”, som (mindre uppenbart) gör att fältlinjerna visar ett viss motstånd mot böjning. De föredrar att vara raka linjer.